A COMPREHENSIVE FRAMEWORK FOR MUNICIPAL SOLID WASTE FORECASTING & ROUTE OPTIMIZATION: A COMPARATIVE ANALYSIS OF DIFFERENT MACHINE LEARNING MODELS FOR THE CITY OF SURAT

A Thesis submitted to Gujarat Technological University

for the Award of

Doctor of Philosophy

in

Civil Engineering

by

Rathod Abhijit Ramjibhai

209999912005

Under supervision of

Dr. Vinodkumar Manilal Patel



GUJARAT TECHNOLOGICAL UNIVERSITY, AHMEDABAD

OCTOBER – 2026

© Rathod Abhijit Ramjibhai

DECLARATION

I declare that the thesis entitled **A Comprehensive Framework for Municipal Solid Waste Forecasting & Route Optimization: A Comparative Analysis of different Machine Learning Models for the City of Surat** submitted by me for the degree of Doctor of Philosophy is the record of research work carried out by me during the period from **February 2021** to **July 2026** under the supervision of **Dr. Vinodkumar Manilal Patel** and this has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other University or other institution of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis. I shall be solely responsible for any plagiarism or other irregularities, if noticed in the thesis.

Signature of the Research Scholar: ..... Date: **13/10/2026**

Name of Research Scholar: **Rathod Abhijit Ramjibhai**

Place: **Ahmedabad**

CERTIFICATE

I certify that the work incorporated in the thesis **A Comprehensive Framework for Municipal Solid Waste Forecasting & Route Optimization: A Comparative Analysis of different Machine Learning Models for the City of Surat** submitted by **Mr. Rathod Abhijit Ramjibhai** was carried out by the candidate under my supervision/guidance. To the best of my knowledge: (i) the candidate has not submitted the same research work to any other institution for any degree/diploma, Associateship, Fellowship or other similar titles (ii) the thesis submitted is a record of original research work done by the Research Scholar during the period of study under my supervision, and (iii) the thesis represents independent research work on the part of the Research Scholar.

Signature of the Research Supervisor:

Date: **13/10/2026**

Name of Research Supervisor: **Dr. Vinodkumar Manilal Patel**

Place: **Ahmedabad**

Course-work Completion Certificate

This is to certify that **Mr. Rathod Abhijit Ramjibhai** enrolment no. is **209999912005** enrolled for PhD program in the **Civil Engineering** branch of Gujarat Technological University, Ahmedabad.

☐ He has been exempted from the course-work (successfully completed during M. Phil Course)

☐ He has been exempted from Research Methodology Course only (successfully completed during M. Phil Course)

☑ He has successfully completed the PhD course work for the partial requirement for the award of PhD Degree. His performance in the course work is as follows-

Grade Obtained in Research Methodology [PHD2001]	Grade Obtained in Research and Publication Ethics [PHD2002]	Grade Obtained in Self-Study Course/Contact Program [PHD2003]
BB	**AA**	**AB**

Supervisor's Sign

(Dr. Vinodkumar Manilal Patel)

Originality Report Certificate

It is certified that PhD Thesis titled **A Comprehensive Framework for Municipal Solid Waste Forecasting & Route Optimization: A Comparative Analysis of different Machine Learning Models for the City of Surat** by **Mr. Rathod Abhijit Ramjibhai** has been examined by us. We undertake the following:

a) Thesis has significant new work / knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.

b) The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.

c) There is no fabrication of data or results which have been compiled/analyzed.

d) There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

e) The thesis has been checked using **Turnitin** (copy of originality report attached) and found within limits as per GTU Plagiarism Policy and instructions issued from time to time (i.e. permitted similarity index <=10%).

Signature of the Research Scholar: ..... Date: **13/10/2026**

Name of Research Scholar: **Rathod Abhijit Ramjibhai**

Signature of the Research Supervisor: ..... Date: **13/10/2026**

Name of Research Supervisor: **Dr. Vinodkumar Manilal Patel**

Place: **Ahmedabad**

Originality Report

Abhijit Rathod

A comprehensive framework for municipal solid waste forecasting & route optimization: A comparative analysis of di...

Quick Submit

Quick Submit

Gujarat Technological University

Document Details

Submission ID
trn:oid:::1:3647574412

Submission Date
Sep 13, 2026, 5:27 PM GMT+5:30

Download Date
Sep 13, 2026, 5:38 PM GMT+5:30

File Name
209999912005_Thesis_Final-Copy.docx

File Size
30.6 MB

361 Pages

95,974 Words

586,814 Characters

turnitin Page 1 of 366 - Cover Page Submission ID trn:oid:::1:3647574412

8% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text
- Small Matches (less than 14 words)

Match Groups

- 59 Not Cited or Quoted 8%
  Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%
  Matches that are still very similar to source material
- 0 Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

2% Internet sources
2% Publications
8% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.



Ph.D. Thesis Non- Exclusive License to

GUJARAT TECHNOLOGICAL UNIVERSITY

In consideration of being a Research Scholar at Gujarat Technological University, and in the interests of the facilitation of research at the University and elsewhere, I, **Mr. Rathod Abhijit Ramjibhai** having Enrollment No. **209999912005** hereby grant a non-exclusive, royalty free and perpetual license to the University on the following terms:

a. The University is permitted to archive, reproduce and distribute my thesis, in whole or in part, and/or my abstract, in whole or in part (referred to collectively as the "Work") anywhere in the world, for non-commercial purposes, in all forms of media;

b. The University is permitted to authorize, sub-lease, sub-contract or procure any of the acts mentioned in paragraph (a);

c. The University is authorized to submit the Work at any National / International Library, under the authority of their "Thesis Non-Exclusive License";

d. The Universal Copyright Notice (©) shall appear on all copies made under the authority of this license;

e. I undertake to submit my thesis, through my university, to any Library and Archives. Any abstract submitted with the thesis will be considered to form part of the thesis.

f. I represent that my thesis is my original Work, does not infringe any rights of others, including privacy rights, and that I have the right to make the grant conferred by this non-exclusive license.

g. If third party copyrighted material was included in my thesis for which, under the terms of the Copyright Act, written permission from the copyright owners is required, I have obtained such permission from the copyright owners to do the acts mentioned in paragraph (a) above for the full term of copyright protection.

h. I understand that the responsibility for the matter as mentioned in the paragraph (g) rests with the authors / me solely. In no case shall GTU have any liability for any acts / omissions / errors / copyright infringement from the publication of the said thesis or otherwise.

i. I retain copyright ownership and moral rights in my thesis, and may deal with the copyright in my thesis, in any way consistent with rights granted by me to my university in this non-exclusive license.

j. GTU logo shall not be used /printed in the book (in any manner whatsoever) being published or any promotional or marketing materials or any such similar documents.

k. The following statement shall be included appropriately and displayed prominently in the book or any material being published anywhere: "The content of the published work

is part of the thesis submitted in partial fulfilment for the award of the degree of Ph.D. in Civil Engineering of the Gujarat Technological University".

l. I further promise to inform any person to whom I may hereafter assign or license my copyright in my thesis of the rights granted by me to my university in this nonexclusive license. I shall keep GTU indemnified from any and all claims from the Publisher(s) or any third parties at all times resulting or arising from the publishing or use or intended use of the book / such similar document or its contents.

m. I am aware of and agree to accept the conditions and regulations of Ph.D. including all policy matters related to authorship and plagiarism.

Date: **13/10/2026**

Place: **Ahmedabad**

Signature of the Research Scholar

Recommendation of the Supervisor: **YES, RECOMMENDED**

Signature of the Research Supervisor

Thesis Approval Form

The viva-voce of the PhD Thesis submitted by **Mr. Rathod Abhijit Ramjibhai (209999912005)** entitled **A Comprehensive Framework for Municipal Solid Waste Forecasting & Route Optimization: A Comparative Analysis of different Machine Learning Models for the City of Surat** was conducted on: **13/10/2026 (Tuesday)** at Gujarat Technological University.

(Please tick any one of the following options)

☑ The performance of the candidate was satisfactory. We recommend that he be awarded the PhD degree.

☐ Any further modifications in research work recommended by the panel after 3 months from the date of first viva-voce upon request of the Supervisor or request of Independent Research Scholar after which viva-voce can be re-conducted by the same panel again.

(Briefly specify the modifications suggested by the panel)
-----Not Applicable-----

☐ The performance of the candidate was unsatisfactory. We recommend that he/she should not be awarded the PhD degree.

(The panel must give justifications for rejecting the research work)
-----Not Applicable-----

Dr. Vinodkumar Manilal Patel
Name and Signature of Supervisor

External Examiner 1
1) (External Examiner 1) **Dr. Deepak Singh**

External Examiner 2
2) (External Examiner 2) **Dr. Bhoomi R. Andharia**

ABSTRACT

Indian cities are urbanising faster than their municipal solid waste (MSW) systems can adapt, with collection infrastructure, forecasting practice, and routing formulations often assuming levels of data availability and GPS accuracy that most Urban Local Bodies (ULBs) do not possess. This thesis develops an integrated three-part framework—forecasting, route optimisation, and decision support—built and tested on data from the Surat Municipal Corporation (SMC), Gujarat, to address these deficiencies within a single empirically grounded programme of research.

The forecasting component constructs a 419-feature engineering pipeline from 186 months of SMC waste generation records (June 2009–December 2024) and comparatively evaluates seventeen models spanning linear and regularised regression, kernel methods, ensemble learning, deep learning architectures, and classical time series approaches under an eight metric evaluation protocol. Across this uniform tabular setting, Ordinary Least Squares regression achieves the strongest overall performance ($R = 0.9972$, MAPE = 1.43), outperforming all tested deep learning architectures, including LSTM, and demonstrating that systematic feature engineering can supersede architectural complexity in data-scarce municipal forecasting contexts. A separate interrupted time series analysis on an independent 196-month raw dataset isolates the structural impact of the COVID-19 disruption, with linear models incorporating a pandemic indicator achieving up to a 16% reduction in RMSE relative to non-indicator baselines.

The route optimisation component unfolds in three stages. The first stage, published in the International Journal of Applied Mathematics (2025), couples an LSTM forecaster with a capacitated vehicle routing (CVRP) solver on a synthetic network scaled to Surat and demonstrates substantial reductions in collection distance and fleet size as a proof of concept. The second stage routes forecasted demand through eight GPS verified transfer stations and the Khajod disposal site, quantifying an 82% theoretical efficiency potential relative to current operations and highlighting transfer stations as operational bottlenecks rather than passive depots. The third stage analyses 662,803 real operational collection records from October 2025, revealing that only 8.6% of unique collection point names can be geocoded to coordinates and motivating IMPACT, a five component GPS-free framework validated on

the full operational dataset. Using an XGBoost classifier with SHAP attribution, IMPACT demonstrates that contractor noncompliance, rather than fleet capacity, is the dominant cause of missed collection stops across all nine zones.

The decision support component operationalises the forecasting and optimisation outputs within a four-layer architecture—Data Ingestion, Model Execution, Analytics, and User Interface—implemented as a Streamlit proof of concept with verified stakeholder dashboards and data sources. The system is explicitly designed to be GPS-free and IoT-free, integrating monthly forecasting and near-real-time optimisation outputs via a demand scalar and role-specific interfaces for SMC personnel, and is structured for replication in Indian ULBs that lack GPS tracked fleets or dense GIS infrastructure. The framework is aligned with the Smart Cities Mission, Swachh Bharat Mission 2.0, and Sustainable Development Goals 11, 12, and 13, and delivers eighteen original methodological, empirical, and practical contributions mapped to three formally approved research objectives and five major literature gaps across forecasting, feature engineering, routing, GIS, and integrated decision support.

Collectively, the research provides a compliance first governance recommendation for a tier-II Indian municipality, validated on large-scale operational data, and demonstrates that integrated, GPS-free, machine learning driven forecasting and routing, supported by a deployable decision support system, can substantially improve the reliability, efficiency, and policy alignment of MSW collection systems in Surat and comparable Indian and developing economy urban contexts.

Acknowledgement

Pursuing this PhD has been a transformative journey, and it would not have been possible without the invaluable support, encouragement, and guidance of many individuals. I take this opportunity to sincerely express my gratitude to all those who, directly or indirectly, supported me throughout the completion of my doctoral journey.

First and foremost, I express my deep sense of gratitude to my research supervisor **Dr. Vinodkumar M. Patel**, Professor, Civil Engineering Department, Shantilal Shah Engineering College, Bhavnagar for his invaluable guidance, constant encouragement, moral support, and wholehearted co-operation during this research work. His painstaking efforts in reviewing the manuscript and his interest in the research work are beyond words of appreciation. I am deeply grateful to him for always being there for me and I feel truly privileged to have the opportunity to work under his guidance and share this journey with him.

I extend my sincere gratitude to Doctoral Progress Committee members, **Dr. Umeshkumar Khare** and **Dr. Yogeshkumar S. Patel** for their valuable guidance, constructive suggestions, and constant encouragement during the course of my research. Their insightful feedback and critical observations have been instrumental in refining the quality of this work and in strengthening the overall direction of my doctoral study.

I sincerely thank the Officials of the Surat Municipal Corporation, Surat for providing the essential datasets used in this research. I also gratefully acknowledge the Google for providing the Google Collaboratory platform, which provided the necessary computational resources and GPU acceleration used to conduct the experiments in this study.

I gratefully acknowledge the blessings of my elder brother, **Hitendra R. Rathod,** whose values and inspiration have always guided me in this journey. I am profoundly grateful to my wife **Mrs. Kusum Rathod** for her unconditional love, patience and unwavering support, which gave me strength throughout this journey. I am thankful to my parents' **Mr. Ramjibhai** and **Mrs. Dhirajben Rathod** for their blessings, endless motivation and constant encouragement during doctoral journey. The affection, strength, and encouragement of my entire family have been invaluable in helping me persevere and complete this research work.

A special mention goes to my beloved son, **Swayam Rathod**, whose innocent smile and unwavering love brought me immense strength and joy. His innocent curiosity and radiant smiles were my strongest inspiration throughout this academic journey.

My heartfelt thanks to my dear friends and colleagues, **Prof. (Dr.) Utkarsh Nigam, Prof. (Dr.) Nevil Trambadia, Prof. (Dr.) Girish Jagad, Prof. (Dr.) Bhagyesh Contractor, Prof. Prashant Ramani** and **Prof. Alkeshkumar A. Rabari** whose encouragement, companionship and support have been a source of strength and joy throughout this research journey.

I am sincerely thankful to the **PhD Section** of Gujarat Technological University for their timely assistance, administrative support, and cooperation during my research journey.

In the end, I remain grateful to all who directly or indirectly supported, encouraged, and inspired me throughout this doctoral journey.

Thank you all for your invaluable contributions to my PhD research on "A Comprehensive Framework for Municipal Waste Forecasting & Route Optimization: A Comparative Analysis of different Machine Learning Models for the City of Surat"

"Many have not been mentioned, but none forgotten"

Rathod Abhijit Ramjibhai

Dedicated to my dear brother

Shri Hitendra R. Rathod

Dedicated to my wife

Mrs. Kusum Rathod

Dedicated to my son

Swayam

Table of Contents

DECLARATION i

CERTIFICATE ii

Course-work Completion Certificate iii

Originality Report Certificate iv

Ph.D. Thesis Non- Exclusive License to vii

Thesis Approval Form ix

ABSTRACT x

Acknowledgement xii

Table of Contents xv

List of Abbreviation and Symbols xxiii

List of Figures xxvi

List of Tables xxix

List of Appendices xxxii

1.1 Background of the Study 1

1.2 The Surat Case Study Context 2

1.2.1 City Profile and Waste Generation Trends 2

1.2.2 Operational Challenges in MSW Management 4

1.3 Research Gap and Problem Statement 6

1.3.1 Forecasting Deficiency in Indian Municipal Context 6

1.3.2 Route Optimization Integration Gap 7

1.3.3 Decision Support System Integration Deficit 8

1.3.4 Integrated Problem Statement 9

1.4 Research Objectives 10
1.5 Research Questions 10
1.6 Significance and Contributions of the Study 12
1.6.1 Theoretical Contributions 12
1.7 Practical Contributions 14
1.7.1 Environmental and Sustainability Contributions 16
1.7.2 Societal and Health Impact 17
1.8 Scope and Delimitations 18
1.8.1 Research Scope 18
1.8.2 Delimitations 19
1.9 Thesis Organization 20
2.1 Introduction 24
2.2 Municipal Solid Waste Generation Forecasting 27
2.2.1 Traditional Statistical Approaches 28
2.2.2 Machine Learning-Based Forecasting 30
2.2.3 Deep Learning-Based Forecasting 31
2.2.4 Evolution of Forecasting Research 32
2.2.5 Critical Analysis 34
2.3 Feature Engineering and Explainable Artificial Intelligence 37
2.3.1 Importance of Feature Engineering in MSW Forecasting 38
2.3.2 Socioeconomic and Demographic Variables 38
2.3.3 Environmental and Climatic Variables 39
2.3.4 Temporal Feature Engineering 39
2.3.5 Feature Selection Methods 40
2.3.6 Explainable Artificial Intelligence and SHAP 41
2.3.7 Critical Analysis 42
2.4 Machine Learning and Deep Learning Models for MSW Forecasting 44
2.4.1 Supervised Machine Learning Models 45
2.4.2 Deep Learning Models 46

2.4.3 Comparative Performance Trends and Model Ranking Evidence 47
2.4.4 Critical Analysis 48
2.5 Evaluation Metrics and Validation Protocols in MSW Forecasting 52
2.5.1 Error and Accuracy Metrics 53
2.5.2 Goodness-of-Fit and Skill Metrics 54
2.5.3 Validation Schemes and Protocol Design 55
2.5.4 Critical Analysis 55
2.6 COVID-19 Impact on Municipal Solid Waste Systems 57
2.6.1 Global COVID-19 Impact Studies on MSW Quantity and Composition 57
2.6.2 COVID-19 Effects on Collection and Operational Systems 58
2.6.3 COVID-19 Effects on Forecasting Accuracy 59
2.6.4 Interrupted Time Series and Difference-in-Differences Approaches 59
2.6.5 Critical Analysis 60
2.7 Route Optimization in Waste Collection 62
2.7.1 Classical VRP and CVRP Formulations 63
2.7.2 Metaheuristic Solution Methods 63
2.7.3 Inventory Routing and Dynamic Collection 64
2.7.4 Multi-Objective and Green Routing 66
2.8 GIS and Spatial Applications in MSW Management 69
2.8.1 GIS-Based Route Optimization Studies 69
2.8.2 GIS-MCDM for Facility and Collection-Point Siting 70
2.8.3 Spatial Data Challenges in Developing Cities 70
2.9 Decision Support Systems and Smart-City Waste Management 72
2.9.1 Classical Decision Support Systems for SWM 72
2.9.2 IoT-Enabled Smart Waste Systems 73
2.9.3 Digital Twin Decision Support Systems 74
2.9.4 Stakeholder Interfaces, KPIs, and SDG Alignment 74
2.10 Integrated Forecasting–Optimization–DSS Frameworks 76
2.10.1 Level 2 Integration: Forecasting and Optimization 76
2.10.2 Level 3 Integration: Adding GIS 77
2.10.3 Level 4 Integration: Adding DSS 77

2.10.4 Level 5 Integration: Full Smart-City and Policy Alignment 77
2.11 Critical Analysis of Existing Research 82
2.11.1 MSW Forecasting Methods 82
2.11.2 Feature Engineering and XAI 83
2.11.3 ML/DL Model Comparison and Benchmarking 83
2.11.4 COVID-19 Impact and Policy-Sensitive Forecasting 84
2.11.5 Route Optimization and Transfer Station Integration 85
2.11.6 GIS and Spatial Realism 86
2.11.7 DSS, Smart-City Systems, and Integrated Frameworks 86
2.12 Research Gap Identification 87
2.13 Chapter Summary 92
3.1 Research Philosophy and Overall Design 96
3.2 Four-Phase Integrated Research Methodology Framework 99
3.2.1 Phase 1: Data Collection, Preprocessing, and Feature Engineering 100
3.2.2 Phase 2: Forecasting Model Development and Comparative Evaluation 100
3.2.3 Phase 3: Vehicle Routing Optimization 100
3.2.4 Phase 4: Decision Support System Design 101
3.3 Data Collection Strategy 101
3.4 Data Preprocessing and Cleaning 105
3.5 Feature Engineering Framework 106
3.6 Model Development and Validation Protocol 109
3.7 Daily Demand Disaggregation Methodology 113
3.8 CVRP Formulation and Optimisation Procedure 115
3.9 DSS Architecture and Implementation Plan 120
4.1 Introduction and Chapter Scope 123
4.2 Descriptive Analysis of Waste Generation Time Series 124

4.3 Feature Space Overview 127

4.4 Comparative Performance of Forecasting Models 129

4.4.1 Group A: Linear and Regularised Models 131
4.4.2 Group B: Kernel Methods 133
4.4.3 Group C: Gradient Boosting Ensembles 135
4.4.4 Group D: Bagging Ensemble 136
4.4.5 Group E: Deep Learning 137
4.4.6 Group F: Classical Time Series Baselines 139

4.5 Model Ranking and Robustness Analysis 140

4.5.1 Tier Classification System 140
4.5.2 Statistical Significance Testing 142
4.5.3 Computational Efficiency Analysis 143

4.6 Feature Importance and Sensitivity Analysis 146

4.6.1 Cross-Model Feature Importance Consensus 146
4.6.2 Ablation Study Results 148
4.6.3 COVID-19 Pandemic Impact Analysis 150

4.7 Discussion 154

4.7.1 Why Feature Engineering Matters More Than Architecture 154
4.7.2 Statistical versus Machine Learning versus Deep Learning Performance 155
4.7.3 Practical Implications for Municipal Forecasting 157

5.1 Introduction: The Three-Stage Research Evolution 159

5.2 Stage 1 - Published CVRP Proof of Concept 161

5.2.1 LSTM Forecasting on the Raw SMC Monthly Dataset 161
5.2.2 Synthetic CVRP Infrastructure Validated Against National Standards 162
5.2.3 LSTM Forecasting Results 163
5.2.4 CVRP Optimization Results 164
5.2.5 Limitations and Motivation for Stage 2 166

5.3 Stage 2 - Transfer Station Transportation Efficiency Assessment 166

5.3.1 Dataset: Eight Real Transfer Station GPS Coordinates 166
5.3.2 Methodology: Haversine Distance Computation and Demand Disaggregation 167

5.3.3 Results: Baseline Distance and Theoretical Optimization Potential 168
5.3.4 Limitations and Motivation for Stage 3 169
5.4 Stage 3 - Full Operational Data Investigation 170
5.4.1 Dataset: October 2025 SMC Operational Route Data 170
5.4.2 Geocoding Investigation: Nominatim Pilot Study 170
5.4.3 Geocoding Failure Analysis 171
5.4.4 Research Pivot: The GPS-Free Operational Framework 173
5.4.5 Framework Architecture and Design Rationale 174
5.4.6 KPI Baseline Establishment 177
5.4.7 TWR-A: Time Window Recalibration - Global Shift Analysis 180
5.4.8 TWR-B: Operational Time Window Framework 183
5.4.9 VAO: Vehicle Allocation Optimization 184
5.4.10 MSRS: Missed Stop Route Cause System 187
5.4.11 WDB+RB: Ward Demand Balancing and Route Balancing 190
5.4.12 FIO: Forecasting Integration Optimization 193
5.5 Discussion 196
5.5.1 The Research Evolution as Methodological Contribution 196
5.5.2 Why Compliance Dominates Over Capacity: Central Finding and Policy Implications 197
5.5.3 Forecasting-to-Operations Integration: How Forecast Accuracy Drives Routing Efficiency 199
5.5.4 Limitations and Directions for Future Research 200
6.1 Introduction and Chapter Scope 202
6.2 Unified DSS Architecture 204
6.2.1 Data Ingestion Layer 208
6.2.2 Model Execution Layer 209
6.2.3 Analytics Layer 210
6.2.4 User Interface Layer 211
6.2.5 Conditional Formatting of Tables 211
6.2.6 GPS-Free Architecture 211
6.3 Stakeholder-Specific Dashboards and Interfaces 212
6.4 Real-Time Replanning and Scenario Analysis 224

6.5 Four-Phase Implementation Roadmap for SMC Deployment 229
6.6 Role-Based Access Control and System Governance 234
6.6.1 Operational Access Tier: 235
6.7 Alignment with Municipal Deployment Requirements 236
6.8 Validation and Performance Assessment 242
6.9 Discussion 244
7.1 Introduction and Chapter Scope 248
7.2 Integrated Framework Performance 249
7.2.1 System-Level Performance Summary 249
7.2.2 Component Interdependencies and Amplification Effects 251
7.2.3 Integrated Contribution Assessment 252
7.3 Forecasting and Operational Optimization Synergies 258
7.3.1 Demand Propagation: From Monthly Forecast to Daily Operations 258
7.3.2 Forecast Accuracy Sensitivity on Operational Efficiency 259
7.3.3 Cross-Temporal Integration Architecture 261
7.4 Comparison with Existing Integrated Approaches 261
7.4.1 Framework-Level Comparison Matrix 261
7.4.2 What This Framework Provides That Prior Work Does Not 264
7.4.3 Position in the Integration Maturity Spectrum 266
7.5 Practical Implications for Surat and Comparable Cities 267
7.5.1 Surat-Specific Action Recommendations 267
7.5.2 Generalizability to Tier-II Indian Cities 268
7.5.3 Transferability to Developing-Economy Urban Contexts 270
7.6 Alignment with Smart Cities Mission, Swachh Bharat Mission, and SDGs 271
7.6.1 SDG 11 - Sustainable Cities and Communities 271
7.6.2 SDG 12 - Responsible Consumption and Production 272
7.6.3 SDG 13 - Climate Action 273
7.6.4 Smart Cities Mission and Swachh Bharat Mission 2.0 Alignment 274

7.7 Robustness and Sensitivity Discussion 274
7.7.1 Internal Consistency Across Components 274
7.7.2 Scope Boundaries and Acknowledged Limitations 276
7.7.3 Sensitivity of Integrated Conclusions to Individual Assumptions 279
7.8 Implementation and Policy Implications 280
7.8.1 The Compliance-First Governance Narrative 280
7.8.2 Capital Reallocation Case: From Fleet Procurement to Contractor Governance 281
7.8.3 Replication Framework for Indian Urban Local Bodies 283
7.9 Chapter Summary 284
8.1 Summary of Research 286
8.1.1 Summary of Major Contributions 287
8.1.2 Methodological Contributions 288
8.1.3 Empirical Contributions 290
8.1.4 Practical Contributions 292
8.2 Conclusions by Research Objective 296
8.2.1 Research Objective 1: Machine Learning-Based MSW Forecasting Framework 296
8.2.2 Research Objective 2: Multi-Stage Route Optimization and Operational Efficiency Framework 298
8.2.3 Research Objective 3: Integrated Decision Support System 299
8.3 Recommendations for Municipal Authorities 301
8.4 Limitations of the Current Research 303
8.5 Future Research Directions 305
8.5.1 Multi-City Validation of the Integrated Framework 305
8.5.2 IoT Integration and Real-Time Data Streams 306
8.5.3 Multi-Objective Optimization and Reinforcement Learning 306
8.5.4 Circular Economy and Extended Producer Responsibility 307
8.6 Chapter Summary 308

List of Abbreviation and Symbols

%	Percent
°	Degree
°C	Degree Celsius
α	Alpha
β	Beta
κ	Kappa (Contractor Compliance Coefficient)
μ	Mean
σ	Standard Deviation
Σ	Summation (Sigma)
CO_2	Carbon Dioxide
kg	Kilogram
km	Kilometre
km^2	Square Kilometre
L	Litre
m	Metre
m^2	Square Metre
mt/d	Metric Tonnes per Day
t	Metric Tonne
t/day	Tonnes per Day
AI	Artificial Intelligence
AMC	Ahmedabad Municipal Corporation
ANN	Artificial Neural Network
ANFIS	Adaptive Neuro-Fuzzy Inference System
ARIMA	AutoRegressive Integrated Moving Average
ARIMAX	AutoRegressive Integrated Moving Average with Exogenous Variables
ARMS	Adaptive Route Management System
AUC	Area Under the Curve
Avg.	Average
CARTOSAT	Cartography Satellite
CatBoost	Categorical Boosting Algorithm
CB	Community Bin
CNN	Convolutional Neural Network
COVID-19	Coronavirus Disease 2019
CPCB	Central Pollution Control Board
CP-SAT	Constraint Programming Satisfiability Solver
CPI	Consumer Price Index
CSV	Comma-Separated Values
CV	Coefficient of Variation
CVRP	Capacitated Vehicle Routing Problem
DL	Deep Learning

DM	Diebold–Mariano Statistical Test
DSS	Decision Support System
EDA	Exploratory Data Analysis
ETA	Estimated Time of Arrival
FIO	Forecasting Integration Optimization
GBM	Gradient Boosting Machine
GDP	Gross Domestic Product
GIS	Geographic Information System
GoI	Government of India
GPS	Global Positioning System
GRU	Gated Recurrent Unit
GRU-Attention	Gated Recurrent Unit with Attention Mechanism
GUI	Graphical User Interface
HPO	Hyperparameter Optimization
ILP	Integer Linear Programming
IMPACT	Integrated Multi-level Planning for Adaptive Collection Timing
IoT	Internet of Things
IJAM	International Journal of Applied Management
ITS	Interrupted Time Series
JETS	Journal of Engineering and Technological Sciences
JSWTM	Journal of Solid Waste Technology and Management
KPI	Key Performance Indicator
Lasso	Least Absolute Shrinkage and Selection Operator
LightGBM	Light Gradient Boosting Machine
LR	Linear Regression
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLP	Multi-Layer Perceptron
MoHUA	Ministry of Housing and Urban Affairs
MSRS	Missed Stop Root Cause System
MSW	Municipal Solid Waste
NARX	Nonlinear AutoRegressive with Exogenous Inputs
NSE	Nash–Sutcliffe Efficiency
OLS	Ordinary Least Squares
Optuna	Hyperparameter Optimization Framework
OR-Tools	Google Operations Research Tools
OSRM	Open Source Routing Machine
Pareto	Pareto Optimality
PoC	Proof of Concept
POI	Point of Interest

Prophet	Facebook Prophet Forecasting Model
R^2	Coefficient of Determination
RBAC	Role-Based Access Control
RBI	Reserve Bank of India
RF	Random Forest
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
RO	Research Objective
RQ	Research Question
SARIMA	Seasonal AutoRegressive Integrated Moving Average
SBM	Swachh Bharat Mission
SBM 2.0	Swachh Bharat Mission 2.0
SCM	Smart Cities Mission
SDG	Sustainable Development Goal
SHAP	SHapley Additive exPlanations
SMC	Surat Municipal Corporation
SSP	Shared Socioeconomic Pathways
Streamlit	Python Framework for Interactive Web Applications
SVR	Support Vector Regression
TWR	Time Window Reliability
UAT	User Acceptance Testing
ULB	Urban Local Body
VAO	Vehicle Allocation Optimization
VMC	Vadodara Municipal Corporation
VRP	Vehicle Routing Problem
VRPTW	Vehicle Routing Problem with Time Windows
WDB	Workload Distribution Balancing
WGMod	Weather-Augmented Gradient Boosting Model
XGBoost	Extreme Gradient Boosting

List of Figures

FIGURE 1.1: Study Area Map of Surat Municipal Corporation 5

FIGURE 2.1 Research landscape of the Present Thesis 27

FIGURE 2.2 Feature engineering frameworks in MSW forecasting literature: 44

FIGURE 2.3 Five-phase evolution of municipal solid waste generation forecasting research, ... 51

FIGURE 2.4 Taxonomy of Machine Learning and Deep Learning Model Families for MSW Forecasting 52

FIGURE 2.5 - Route Optimization Research Landscape 93

FIGURE 2.6 - GIS and DSS Integration Spectrum in MSW Literature 94

FIGURE 2.7 - Integrated Framework Component Coverage Map 95

FIGURE 2.8 - Research Gap Hierarchy: 26 Gaps Mapped to Research Objectives 95

FIGURE 3.1 Integrated four-phase research methodology framework for the Surat MSW management study. 99

FIGURE 3.2 Comprehensive feature engineering pipeline 107

FIGURE 4.1 Monthly MSW generation time series for Surat (June 2009 – December 2024), with 12-month rolling mean trend overlay and COVID-19 disruption period annotation (March 2020 – March 2022). Data source: Surat Municipal Corporation (SMC) operational records. N = 186 monthly observations. 125

FIGURE 4.2 Model performance landscape - R^2 values for all 17 models arranged by tier, with model family colour coding and tier boundary annotations. 141

FIGURE 4.3 Actual versus predicted scatter plots for OLS, GPR, and Ridge regression on the test set (n = 37 chronological months), with ±5% and ±10% prediction band overlays. 142

FIGURE 4.4 Residual diagnostic analysis - time series of residuals, distribution histograms, and normal Q-Q plots for Tier 1 models (OLS, GPR, Ridge). Shapiro-Wilk test p-values annotated. 145

FIGURE 4.5 Aggregated cross-model feature importance - normalised consensus scores from LightGBM, XGBoost, and Random Forest, with cumulative importance curve across 14 annotated feature domain groups. 147

FIGURE 4.6 Ablation study - change in R^2 relative to the full-feature OLS baseline when each feature category is systematically excluded. 149

FIGURE 4.7 COVID-19 impact visualisation - monthly MSW time series with structural break annotations (March 2020 – March 2022) and comparative actual versus predicted trajectories for LR with and without COVID dummy variable. 152

FIGURE 5.1 Three-Stage Route Optimization Research Evolution Diagram 175

FIGURE 5.3 KPI Baseline Dashboard - Nine-Zone Compliance Overview 177

FIGURE 5.4 TWR-A Scheduled versus Actual Hourly Distribution and Offset Histogram 181

FIGURE 5.5 TWR-A Sensitivity Analysis - ONTIME versus Shift Amount 182

FIGURE 5.6 VAO Fleet Allocation Comparison - Current versus Model C 187

FIGURE 5.7 MSRS SHAP Factor Group Attribution - Global and Zone-Level Decomposition 188

FIGURE 5.8 WDB+RB Demand Distribution Before and After FFD Balancing 191

FIGURE 5.9 FIO Integration Workflow - Forecast to Operations 194

FIGURE 5.10 IMPACT Components Summary Dashboard 195

FIGURE 6.1: DSS Four-Layer Architecture Diagram 205

FIGURE 6.2: IMPACT DSS Wireframe 01 - Executive Dashboard (left panel) and Forecast Dashboard (right panel). 221

FIGURE 6.3: IMPACT DSS Wireframe 02 - KPI Baseline Dashboard (left panel) and Optimization Dashboard (right panel). 222

FIGURE 6.4: IMPACT DSS Wireframe 03 - Compliance Dashboard (left panel) and IMPACT Framework Dashboard (right panel).....223

FIGURE 7.1: Integrated Framework Performance Summary Infographic.....258

List of Tables

TABLE 1.1: Research Objectives, Methods, Data Sources, and Key Outputs Across Thesis Chapters 23

TABLE 2.1 Evolution of Municipal Solid Waste Forecasting Research 32

TABLE 2.2 Summary of Key Municipal Solid Waste Generation Forecasting Studies 34

TABLE 2.3 Feature Engineering Approaches in Municipal Solid Waste Forecasting Studies 42

TABLE 2.4 Machine Learning and Deep Learning Approaches for Municipal Solid Waste Generation Forecasting 48

TABLE 2.5 Evaluation Metrics and Validation Protocols for MSW Forecasting 56

TABLE 2.6 COVID-19 Impact Studies on Municipal Solid Waste Generation 61

TABLE 2.7 Route Optimization Studies in MSW (VRP, CVRP, IRP) 67

TABLE 2.8 GIS-Based Waste Management Studies 71

TABLE 2.9 Decision Support Systems and Smart-City MSW Frameworks 74

TABLE 2.10 Integrated Forecasting–Optimization–DSS Framework Classification 78

TABLE 2.11 Research Gap–Research Objective Mapping Matrix 90

TABLE 3.1 Annual average MSW composition by waste category, Surat Municipal Corporation, 2009–2024 (adapted from [63]) 102

TABLE 3.2 CVRP infrastructure specifications for Stage 1 proof-of-concept analysis (adapted from [3], IJAM Table 1). 117

TABLE 3.3 LSTM feature engineering summary for Stage 1 route optimization analysis (adapted from [62], IJAM Table 2). 118

TABLE 4.1 Eight-metric performance evaluation of 10 successful forecasting models on the held-out test set. 130

TABLE 4.2 Eight-metric performance evaluation of 7 less-optimal forecasting models. Note: Holt-Winters MAPE is undefined (∞) due to near-zero observed values causing division instability..... 131

TABLE 4.3 Computational efficiency metrics for nine models with available timing data. 143

TABLE 4.4 Cross-model feature importance comparison - top feature groups, normalised importance scores from Random Forest and XGBoost. Adapted from JSWTM paper Table 5 [63]. 148

TABLE 4.5 COVID-19 impact analysis - model performance with and without pandemic binary indicator on the JSWTM raw dataset (196 months). Adapted from JSWTM paper Table 2 [63]. 151

TABLE 4.6 Comparative performance of JSWTM framework models versus literature benchmarks for MSW forecasting. Adapted from JSWTM paper Table 7 [63]..... 153

TABLE 5.1 Stage 1 CVRP Proof-of-Concept: LSTM Forecasting and Route Optimization Performance 164

TABLE 5.2 Stage 2 Transfer Station Transportation Efficiency: Demand Profile and Distance to Khajod Disposal Site 168

TABLE 5.3 Geocoding Failure Mode Taxonomy — West Zone Pilot Study (99 Routes, 162 Unique Endpoint Names)..... 171

FIGURE 5.2 IMPACT System Architecture Diagram 176

TABLE 5.4 October 2025 KPI Baseline — Service Status Distribution (455,827 Stop Visits, 949 Vehicles) 177

TABLE 5.5 KPI Baseline by Administrative Zone — All Nine Zones (October 2025, 455,827 Stop Visits) 179

TABLE 5.6 TWR-A Zone-Level Optimal Time Window Shifts and ONTIME Improvements (Strict-Window Baseline) 182

TABLE 5.7 TWR-B Zone-Level Recommended Time Window Shifts, Standardised Window Widths, and Adjustment Priority (All Nine Zones) 183

TABLE 5.8 VAO Zone-Level κ Coefficients, Fleet Status, and Miss Rate Decomposition (October 2025) 185

TABLE 5.9 MSRS Model Performance Comparison — Logistic Regression, Random Forest, and XGBoost 187

TABLE 5.10 MSRS Zone-Level SHAP Attribution — Five Factor Groups (% Contribution to Missed-Stop Prediction) 189

TABLE 5.11 WDB+RB Intra-Zone Demand Coefficient of Variation Before and After First Fit Decreasing (FFD) Rebalancing — All Nine Zones 191

TABLE 6.1. Forecast application guidelines by planning scenario for SMC solid waste management operations. 206

TABLE 6.2. DSS stakeholder requirements matrix for SMC solid waste management decision support system. 213

TABLE 6.3. Six dashboard modules of the IMPACT DSS: functions, key elements, and target audience. 215

TABLE 6.4. FIO demand scalar scenario rules and operational responses for SMC solid waste management. 226

TABLE 6.5. Four-phase SMC implementation roadmap for the IMPACT Decision Support System. 231

TABLE 6.6. Alignment of thesis contributions with United Nations Sustainable Development Goals. 239

TABLE 7.1: Integrated Research Contribution Summary for the Surat MSW Framework (Contributions C1–C18 mapped to Research Objectives and Verified Evidence) 253

TABLE 7.2: Comparative Assessment of the Surat Integrated Framework Against Selected Prior Integrated MSW Management Frameworks.....262

TABLE 7.3: Transferability Assessment of the Surat MSW Framework to Comparable Tier-II Indian Cities and Developing-Economy Urban Contexts.....268

TABLE 7.4: Scope Boundaries, Acknowledged Limitations, and Mitigation Approaches for the Integrated Surat MSW Framework.....276

TABLE 8.1: Summary of Research Contributions (C1–C18) Organized by Contribution Type, Research Objective, and Evidential Basis. 294

List of Appendices

Appendix A Forecasting Pipeline – Python Code…….324

Appendix B List of Publications……...375

CHAPTER 1

Introduction

1.1 Background of the Study

Management of municipal solid waste (MSW) has become one of the main infrastructure problems facing cities around the world in the twenty first century. As of 2024 global MSW production is at an approximate level of 2.01 billion tons annually, and will increase to an estimated 3.40 billion tons annually by 2050 unless the rates of urbanization and consumption continue to grow unabated [1]. Much of this growth will be concentrated in rapidly growing developing economies which have not kept pace with their populations in terms of infrastructure expansion, and whose financial capabilities (i.e., those of urban local bodies (ULBs)) are often insufficient to meet the requirements of managing such quantities of waste. Consequently, there exists a pressing need for effective methods that allow municipalities to collect, process and manage large volumes of waste.

India provides a major example of this situation at a continental scale. At present, India produces approximately 62 million tons of municipal solid waste each year from over 4041 urban local bodies [2]. The Municipal Solid Waste Management Rules (2016) [3] specify the requirement that ULBs shall implement door-to-door segregation of collected waste, and that waste shall be processed and disposed of in a timely manner while reducing dependency on landfills; however, actual implementation of these rules by many Indian ULBs has been limited and difficult to assess [3]. Given that both the Swachh Bharat mission 2.0 (sbm 2.0) [4] and the Smart Cities Mission (SCM) [5] are national programs which require ULBs to demonstrate improved efficiencies in waste collection based upon verifiable data related to collections efficiency, vehicle utilization and compliance with regulatory standards, the existing policy environment is creating increasing pressure on ULBs to develop systematic analytical frameworks for assessing waste management activities.

At the current stage of research on MSW management in India, the research question is not what algorithm performs best when used alone. Published studies have investigated statistical time-series models, machine learning approaches, and heuristics for vehicle routing optimization separately, and although each line of investigation has produced useful insights individually, none have provided an integrated end-to-end analytical framework -- one that starts with accurate predictions of waste generation demands, converts forecasts into operational routing parameters, and implements both within an organizational decision-making structure that can be implemented by Indian municipalities lacking gps tracking devices or structured geographic information systems (gis).

This integration gap -- i.e., the lack of a single analytical framework integrating demand prediction, route optimization, and decision support at the scale of an Indian tier ii municipality -- represents the central focus of the research presented in this dissertation.

To bridge this integration gap, the dissertation presents a three-component analytical framework for surat municipal corporation (SMC), state of Gujarat: a 17-model machine learning forecasting framework utilizing systematic feature engineering across 186 months of historical MSW generation data; a three-stage vehicle route optimization and operational efficiency framework concluding in the gps-free IMPACT (integrated multi-level Planning for Adaptive Collection Timing) framework tested against real October 2025 operational data; and an integrated Decision Support System designed for incremental implementation at SMC's operations and institutionally-based context. three peer reviewed articles published in refereed journals represent the primary evidence base for the research reported in this dissertation: JETS [6] ; jswtm [7] ; ijam [8].

1.2 The Surat Case Study Context

1.2.1 City Profile and Waste Generation Trends

With an estimated population of approximately 6.9 million as of 2024, Surat is the second largest city in the state of Gujarat and one of the fastest growing metropolitan centers in India. It is administered by Surat municipal corporation (SMC); it is characterized by its dense textile and diamond manufacturing sectors, rapid increases in residential development; and commensurate to these developments, Surat also exhibits a high and variable rate of

municipal solid waste generation. Between June 2009 and December 2024, monthly MSW generated in surat grew from approximately 3800 metric tons/day (mt/d) to approximately 6200 mt/d - a 63.2% increase during the same 186 month time frame. These continued upward trends in municipal solid waste generation are caused by population growth; economic growth; and corresponding increases in per capita consumption. Therefore, an ongoing planning challenge faced by SMC is that its static collection infrastructure does not keep pace with its changing levels of municipal solid waste generation without continuous analytical support.

The dataset forming the basis for the research reported herein - referred to hereafter as the "JETS canonical dataset" - is a 186-month time series database that was converted into a 419-dimensional analytical dataset containing temporal lag features (lags 1-12 months); rolling statistical metrics (windows: 3-, 6-, 12-month); socio-economic indices; climate-related variables; Google Mobility indexes; festival/public events calendar binary flags; covid-19 pandemic binary dummies; percent change year-over-year features; and cross-term feature interactions. This dataset served as the primary source of evidence supporting the forecasting component of this thesis.

Surat has a total of nine administrative zones; thirty election wards; and approximately 131 administrative sub-wards (these function as sub-ward office areas under the 9 larger zones to manage basic local civic infrastructure). There are Eight transfer stations with verifiable gps coordinates serving as intermediary consolidation facilities between door-to-door primary collection service areas and the Khajod disposal facility - a facility operated since 2014 located approximately sixteen kilometers from the center of Surat. Surat is identified as one of the beneficiary cities under the Smart Cities Mission program, thereby establishing a political context where analytical investments to improve waste management efficiency are mandated by policy and publicly accountable [4]. The combination of a rich longitudinal dataset; detailed operational records associated with its waste collection infrastructure; and an active policy reform environment make surat an appropriately tractable CASE for developing the integrated framework described in this thesis.

1.2.2 Operational Challenges in MSW Management

Although Surat is ranked very high nationally in cleanliness surveys; nevertheless, a systematic analysis of operational route data from October 2025 demonstrated several serious systemic shortcomings associated with collecting waste efficiently throughout surat. The dataset consisted of 662,803 operational records including 455,827 POI stop records; 3,341 named collection routes; 936 unique active-in-service vehicle identifiers; and all ninety-one wards located across nine administrative zones - thus providing the most detailed operational documentation available on surat's waste collection system to date. Baseline key performance indicator (KPI) analysis determined an ontime completion rate of 22.0% using SMC's Status column labeling; and an ontime completion rate of 6.1% using independent parsing window analysis applied to raw arrival times. Both measures were calculated according to different definitions regarding adherence to schedules; and therefore, both are legitimate measures.

The single most important finding from baseline KPI analysis - and one that altered significantly the research direction - was identification of contractor non-compliance as the greatest operational constraint affecting surat's waste collection efforts. An xgbboost classifier trained on real October 2025 operational data obtained an Area under receiver operating curve (roc-auc) value of .940 and a recall rate value of .881 on classification results using the Missed Stop Root Cause System (MSRS) module component of IMPACT to classify causes of missed collection stops. SHAP (Shapley Additive Explanations) feature attribution analyses showed contractor non-compliance accounted for seventy-five point seven two percent (%), vehicle capacity-related constraints accounted for one point six five (%)%, and scheduling constraints accounted for five point four two (%)% of overall feature attributions as root causes of missed stops respectively. Contractor non-compliance explanations were similarly consistent across all nine waste collection zones -- thus providing strong spatial evidence supporting this explanation.

The Central Zone provided particularly compelling evidence for compliance-dominance hypotheses. Although it had recorded the lowest demand/vehicle ratio across all SMC networks (twenty-one thousand nine hundred ten stop-demand units/vehicle) -- suggesting it should be expected to exhibit excellent collection performance -- it had the highest missed stop rate across all zones in surat at nine point three one (%). Thus, this demonstrates a direct

contradiction between resource availability and operational performance expectations -- contradicting expectations that failed collections in Indian municipalities are generally due to inadequate fleet capacities. Instead, evidence supports contractor governance failure as the primary constraint -- vehicles are dispatched; routes are assigned; however, contracted collection obligations are continually violated.

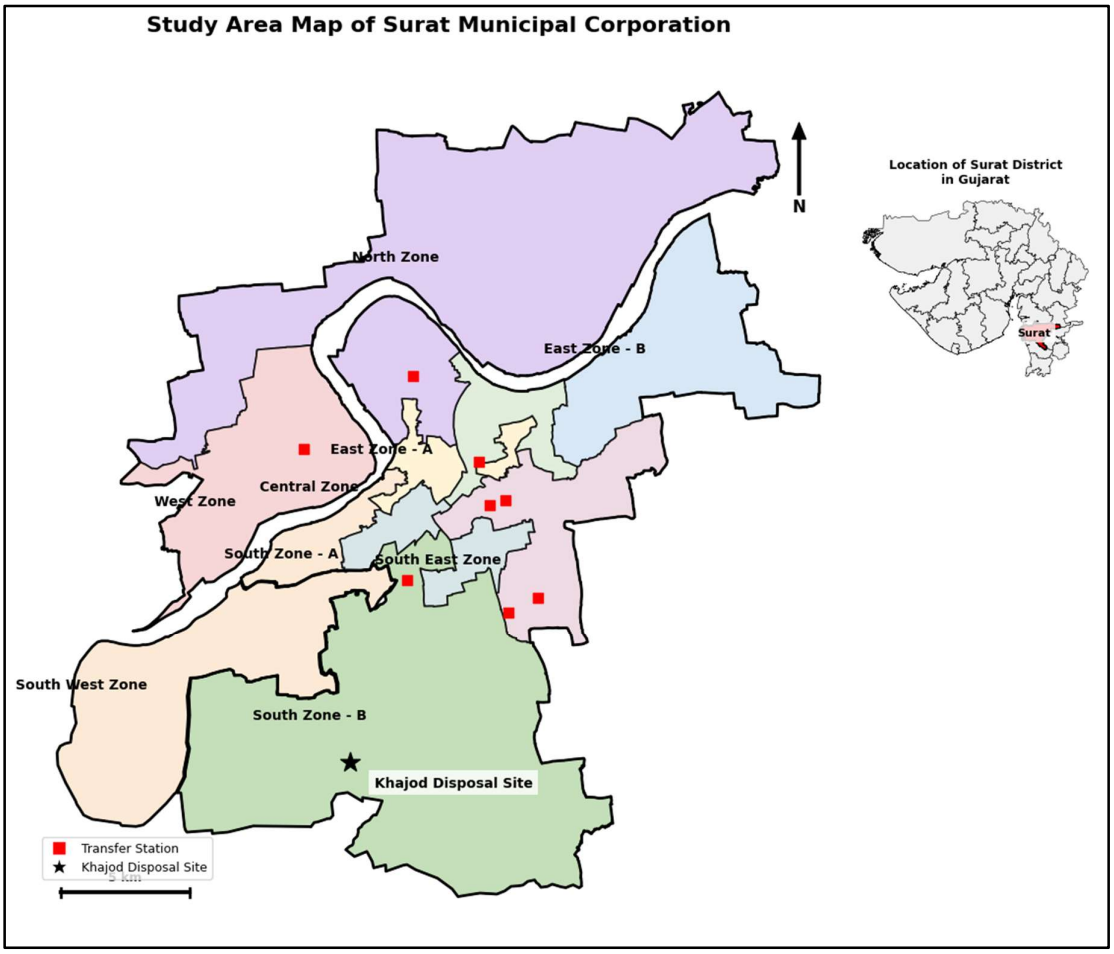


FIGURE 1.1: Study Area Map of Surat Municipal Corporation

Zones are colour-coded by administrative boundary. Transfer station locations are indicated by verified GPS coordinates. Source: SMC administrative data; map generated using Python matplotlib and Surat shapefiles.

Fig.1.1 shows our study area map of Surat Municipal Corporation Showing Nine Waste Collection Zones, Eight Transfer Stations, and the Khajod Disposal Site.

A complementary data constraint also influenced research design. A geocoding pilot performed on west zone (162 names representing unique point-of-interest locations from ninety-nine routes) using Nominatim open-source geocoder resulted in successful usable

coordinate acquisition rate of only Eight point six percent (%) -- far below typical rates ranging from seventy to eighty percent (%) commonly observed for well-defined address formats. Failure was attributed to informal naming conventions employed at surat's beat-level address data -- e.g., collection points defined by local landmarks; community colloquialisms; incomplete ward identifiers; and missing pin codes common in Indian urban addressing practices. Geocoding limitations precluded conventional vehicle routing problems with Time Windows (vrp-tws) formulations for use with October 2025 data; consequently motivated development of gps-free IMPACT framework as primary operational optimization contribution to this thesis.

1.3 Research Gap and Problem Statement

1.3.1 Forecasting Deficiency in Indian Municipal Context

There is a substantial body of research documenting shortcomings in published forecasting literature for India's municipalities. three major limitations were documented in the research supporting this thesis. First, the published comparative analysis is limited. Although there are many studies comparing two to three models and most use only very simple models (e.g., linear regression, sarima, machine learning) there are few if any comparisons of multiple families of models in a unified framework on India's cities' data. Therefore, practicing managers have no empirical basis to select among models based upon the characteristics of their own municipality. Second, there is little evidence to suggest that India's municipalities have invested sufficient resources in feature engineering to improve the accuracy of forecasts. All but a handful of published studies employed some form of lagged feature, and those few that attempted to include additional forms of data did so in an ad hoc manner. Consequently, the relative value of investing in higher quality feature sets compared to better models remains an unanswered question. Third, while numerous studies employing interrupted-time-series (its) methodology have demonstrated how to measure the impact of exogenous shocks (such as natural disasters) on long-run trends in waste generation; no prior study used its methodology to quantify the impacts of COVID-19 on waste generation trends in India. Taken together, these three shortcomings identify a forecasting gap that RO-1 was designed to bridge.

1.3.2 Route Optimization Integration Gap

Another limitation identified in the research supporting this thesis relates to the application of route optimization techniques to Indian municipalities. four related shortages were identified in chapter 2 (Sections 2.11 and 2.12). The First shortage refers to the reliance on unrealistically simplistic representations of physical infrastructure. Specifically, nearly all prior published studies employing route optimization techniques to optimize the operation of Indian municipalities' waste collection fleets have utilized unrealistic simplifications of their physical infrastructures. As an example, although published studies utilizing route optimization techniques to optimize the operation of Indian municipalities' waste collection fleets may utilize realistic depictions of the geography and layout of Indian cities, they also rely on unrealistic depictions of the actual locations of collection events. To date, no prior published study has validated its results using an actual operational dataset. The Second shortage is that there is a lack of understanding regarding the feasibility of open-source geocoding on the urban addresses for India's municipalities. While most prior published route optimization studies assumed that a GPS-coordinate could be obtained from an address record, this is not true for Surat's beat level data. Only 8.6% of 162 unique location names in the West Zone pilot produced usable coordinates via nominatim. This shortage identifies a gap between the assumptions about formats of addresses made by authors of prior published route optimization studies and the realities associated with collecting and processing data at the municipal level in India's urban areas. The Third shortage relates to the lack of modeling contractor compliance. Contractor compliance is empirically identified as the primary source of missed stops (as identified in the SHAP importance metric), and yet it is completely ignored by all prior published route optimization models for Indian MSW. As a result, simply reallocating vehicles does little to reduce missed stop rates. In fact, the fleet reallocations made by Surat's municipal government resulted in almost zero reduction in missed stop rates (-0.0008%). Finally, while there are several analytical studies focusing on optimizing the routing for vehicles collecting municipal solid waste in Indian cities, there is virtually no work focused on analyzing the transfer stations that connect door-to-door collections to disposal facilities. Using Haversine distance calculation in stage 2 of the optimization research, we estimated that Surat has 82% theoretical potential for improvement (i.e., baseline = 926.8 km/day vs. Analytical lower bound = 166.8 km/day).

1.3.3 Decision Support System Integration Deficit

Finally, while researchers have developed several Decision support systems (DSS) for municipal waste management, including logistics-scheduling tools, integrated smart-city platforms with IoT sensors and real time GPS tracking, digital twin architectures and others – they often ignore critical aspects of Indian municipal waste management. Specifically, there are three significant barriers preventing the successful implementation of DSS for Indian municipal waste management that are illustrated below.

Firstly, despite considerable progress in recent years towards developing DSS for municipal waste management, all existing solutions are based on the design assumption that GPS telemetry and structured GIS infrastructure will exist. However, given that nearly 90 percent of India's 4041 urban local bodies (ulb) do not collect GPS tracked fleets or maintain structured GIS address registries for waste collection purposes, this design assumption is a barrier to deploying DSS solutions to the vast majority of target users. There is no prior published solution that combines real operational data and verified stakeholder interfaces with a structured deployment plan suitable for GPS free DSS for Indian municipal waste management.

Secondly, while published DSS designs for municipal waste management provide detailed descriptions of operational stakeholders and their roles; they fail to incorporate role-based access control. Municipal waste management decisions involve multiple institutional levels (operational supervisors monitoring daily routes performance; management level staff responsible for monitoring zone KPIs and contractor compliance; commissioners responsible for making fleet procurement and policy decisions) each institutional level requires different data granularity levels, analytical depths and interface designs. Published academic prototypes typically provide a single analytic function without designing for multiple institutional diversity of Decision-making processes within municipalities.

Lastly, phased deployment planning is consistently absent from published academic prototypes of DSS solutions for municipal waste management. Published solutions are designed to illustrate research demonstrations rather than integrate into existing processes used by organizations such as SMC that involve procurement training governance change management etc. The lack of an implementable deployment roadmap aligned with Indian

institutional requirements represents another gap between academic design and municipal practice.

1.3.4 Integrated Problem Statement

In summary, while considerable progress has been achieved toward improving waste management practices through better forecasting and route optimization technologies; these improvements remain largely isolated from one another. Reliable forecasting is essential for effective planning and efficient resource utilization; however, forecasting alone cannot guarantee optimal routing unless accompanied by real-time traffic information, accurate representation of collection activities and proper consideration of spatial relationships between individual collection routes. Similarly, optimized routing cannot ensure efficient service delivery unless accompanied by adequate forecasting to determine likely demand patterns throughout service area. Finally, neither improved forecasting nor optimized routing provides assurance that municipal agencies will adopt new technologies unless accompanied by well-designed Decision support systems that facilitate adoption and effective use. Therefore, while important advancements have been made toward enhancing waste management capabilities; much additional development remains before these technologies become broadly useful in practice. The three research objectives outlined above represent a necessary step toward filling these remaining knowledge gaps by providing a comprehensive analytical approach that links reliable demand forecasting with operational route optimization and integrates both into a Decision support system suitable for Indian municipalities' needs. The research objectives that framed this study are presented below in Section 1.3. The following Sections present details regarding how these objectives were addressed in this thesis. These details are organized into five chapters that collectively outline how each objective was researched and what contributions this research makes to enhance our knowledge base regarding waste management practices. The remainder of this chapter summarizes key findings from previous studies regarding waste management practices in developing countries and describes why a more comprehensive framework is needed to guide future research.

1.4 Research Objectives

The research was organized around three formal objectives that each focus on a component of integrated MSW management.

RO-1: To develop and compare a forecasting framework using machine learning for monthly generation of MSW in Surat. This includes systematic pipeline for feature engineering, comparative analysis of 17 models from statistical families, ensemble families and deep learning families, an eight metric evaluation protocol and an analysis of impact due to interruption by COVID-19 using time series.

RO-2: Design and evaluate a framework for collection by SMC that progresses from a published proof of concept using LSTM CVRP to real assessment of transfer station efficiency and then a GPS free IMPACT framework validated against October 2025 operational data.

RO-3: Design and validate an integrated Decision Support System operationalizing forecasting and optimization outputs for deployment by SMC. This system has a four layer architecture and integrates forecasting into operations via scalar demand, stakeholder specific interfaces and phased deployment strategy that aligns with Smart Cities Mission, Swachh Bharat Mission 2.0 and SDGs 11, 12 and 13.

Together these three objectives form a unified analytical framework that starts with monthly demand forecasting and goes through operational route optimization to designing decision support systems for SMC deployment. RO-1 is discussed mainly in Chapter 4 with methodological foundations in Chapter 3; RO-2 is mainly addressed in Chapter 5; and RO3 is mainly dealt with in Chapter 6. Chapters 7 and 8 bring together and conclude all three objectives through integrated analysis across components and formal conclusions.

1.5 Research Questions

The three research objectives were converted into the following three research questions which corresponded with each of the above mentioned three research objectives.

RQ-1 (corresponding to RO-1): Which machine learning model family achieves the highest forecast accuracy for monthly municipal solid waste generation in Surat using

statistical, ensemble and deep learning families when using a 419 feature engineering pipeline and does systematic feature engineering quality outperform model architectural complexity in data constrained contexts for municipal forecasts?

It was shown by the research that Ordinary Least Squares (ols) regression applied to the 419 feature engineered JETS dataset constructed from 186 months of historical SMC data provided an R^2 value of .9972 and mean absolute percentage error (MAPE) of 1.43%, representing the best performance among all 17 evaluated models including all deep learning architectures. A tabular cross sectional evaluation framework employed in Chapter 4 produced an R^2 value of -3.04 for the long short-term memory (lstm) network established that model architectural complexity is not compensatory to constraints involving data and context misalignment in monthly datasets containing short panels. The finding that systematic feature engineering quality superseded architectural complexity in this forecasting application represents both the direct answer to rq-1, as well as a substantive theoretical contribution made by this thesis [6].

RQ-2 (corresponding to RO-2): Can a multi-stage operational and planning vehicle routing optimization framework for improving operational efficiencies be created and progressively verified – from an existing LSTM-CVRP proof-of-concept based on literature, through real-world evaluation of operational efficiency at transfer stations, to GPS-free operation — to identify and address the systemic factors causing missed waste collection stops in Surat?

The three stages of the development of this study, documented in Chapter 5 showed that the systemic cause of missed stops in Surat's solid waste collection system was due to contractors' failure to comply with requirements; not lack of capacity in fleets. The first stage (published in IJAM 2025 [8]), showed that there could have been a 50.31% reduction in the total distance traveled by a truck operating on a Surat scale representative network using a proposed LSTM-CVRP. The second stage used Haversine distance analysis on the actual eight GPS coordinates of transfer stations in Surat to find that theoretically, an optimal solution could have reduced the daily total kilometers traveled by trucks by 82%. The third stage developed the IMPACT framework as GPS-free operation, verified against the 662,803 row operational dataset collected in October 2025 and found through validation of the MSRS component that missed collection stop failures were attributed to contractor non-compliance

with the requirements for 75.72% of SHAP feature importance attribution among all nine zones.

RQ-3 (corresponding to RO-3): Design of a GPS-free decision support system which is operationally deployable that integrates on a monthly basis forecasted and in real time optimized outputs from SMC stakeholders across three tiers of personnel. What are the conditions under which this could align with SMART CITIES MISSION SWACHH BHARAT MISSION 2.0 AND SUSTAINABLE DEVELOPMENT GOALS 11, 12 AND 13 [9].

The study developed and tested a four-layer GPS free DSS architecture (Layers Data Ingestion, Layer Model Execution, Layers Analytics, Layer User Interface) by linking OLS forecasts at a monthly level with IMPACT operational data using a Forecasting Integration Optimization (FIO) demand scalar $s_d = \widehat{F_d}/F_{\text{base}}$. It was also established via implementation and demonstration of a Streamlit prototype (six modules dashboards; fourteen verified csv data source) across the three tier levels of stakeholder. The proposed framework has been mapped against SMART CITIES MISSION and SWACHH BHARAT MISSION 2.0's operational standards. Furthermore, it has been linked to SDG 11, 12 & 13 as per specific measurable outcomes documented in Chapter 7.

1.6 Significance and Contributions of the Study

Research resulted in the development of eighteen novel contributions (C1-C18) to forecasting method, operational optimization, and decision support systems. The contributions are organized as follows with respect to their type of contribution. All references cited represent completed research that has been peer reviewed, or verified through the author's thesis.

1.6.1 Theoretical Contributions

C1 - First 17 model comparative forecast approach for an Indian city: A new comparative MSW forecast approach was developed and implemented using 419 engineering features for an Indian city and represents the first comparative framework of its kind. This new framework consists of six different classes of algorithms including: linear and regularized models (Ordinary Least Squares, Ridge, Lasso, Elastic Net), kernel-based methods

(Gaussian Process Regression and Support Vector Regression), gradient boosted ensembles (XGBoost, LightGBM, CatBoost), bagged ensemble (Random Forest), neural networks (Tuned Multi Layer Perceptron, Long Short Term Memory Recurrent Neural Network, Gated Recurrent Unit Attention Recurrent Neural Network, Convolutional Long Short Term Memory Recurrent Neural Network), and classical time series methodologies (Seasonal Auto Regressive Integrated Moving Average Model, Holt Winters Exponential Smoothing Model, Facebook's Prophet Methodology). The breadth of this comparative framework is greater than that found in any previous study on MSW forecasting for an Indian city and provides a basis of evaluation for all subsequent algorithmic comparisons [6].

C3 - Evaluation of 8-Metric Multidimensional Standardized Protocol for Model Assessment in Evaluating Forecasting Models of Municipal Water Demand: An eight-metric standardized evaluation protocol was used as a uniform evaluation metric for each of the 17 forecasted models assessed: R^2, MAE, RMSE, MAPE, CV, Theil's U Statistic, Mean Daily Absolute Error (MDA), and Residual Standard Deviation. Use of multiple metrics assists with cross study comparability problems created by the use of a single metric in previous studies. The metrics may also be easily applied to future studies assessing forecasting models for municipalities [6].

C5 - A new empirical result showing that feature engineering is better than complex architectures: The least complex model in the test (a simple OLS model) performed best of all Deep Learning architectures tested with the exact same 419 features. The results were as follows: LSTM had an R^2 value of -3.04; GRU-Attention had an R^2 value of -3.65; ConvLSTM had an R^2 value of -3.03. Thus, this study resolved one of the major open questions in the literature on Applied Machine Learning for Municipal Environmental Data, and will inform Model Selection Guidelines across Indian and Developing-Economy Urban Local Bodies [6].

C9 - First documented evidence of a system-wide geocoding failure in an open source geocoder using municipal address data from india; the Nominatim geocoder had only 8.6% of Surat's beat level address data successfully geocoded with coordinates -a significant decrease compared to the typical 70-80% expected with structured addresses. Therefore, this study demonstrates that vrptw problems utilizing gps-coordinates are unlikely to be deployed

to over 90% of Indian urban local bodies (ulb) without either a geocoding remediation of the existing address data or an alternative gps-free solution.

C11- rejection of the fleet inadequacy hypothesis. Contractor compliance failure was found to have accounted for 75.72% of Contractor SHAP feature importance attribution, accounting for the reasons why collection stops were missed. Vehicle capacity factors had only 1.65% of Contractor SHAP feature importance attribution, across all nine zones, as verified through an XGBoost MSRS classifier having an AUC=0.940 and recall=88.1%. The central zone provided a counter-intuitive example that cannot be explained based upon a capacity driven model (the lowest number of vehicles per unit of demand at 21,909 but highest percentage of collection stops missed at 9.31%).

C18 - adaptive research design – methodological contribution. The use of evidence-based adaptive research design allowed for converting a 8.6% failure rate in geocoding to produce the gps free IMPACT framework. The resulting gps-free framework is deployable to approximately 90% of the Indian urban local bodies (ULBs), making the research failure a research success.

1.7 Practical Contributions

C2 - feature engineering pipeline

The systematic pipeline that includes temporal lag features (lagged by 1 through 12 months), rolling statistic features (over 3, 6, and 12 month periods), socio-economic variables, climate variables, Google Mobility variable, festival/event indicator, COVID-19 indicator, interaction terms, and year over year percent change variables provides a direct template for engineers developing similar pipelines for other Indian cities with monthly waste generation datasets.

C4 - it's for COVID-19

Applying it to the COVID-19 structural break (March 2020 - March 2022) resulted in a 16% reduction in root mean squared error (RMSE) compared to linear regression and a 10% improvement in RMSE with respect to XGBoost when the COVID-19 dummy variable was

added into the model [2]. The same method can be applied to other cities in India that have waste generation records during the pandemic.

C6 - characterization of computational resources

OLS requires .01 seconds to train completely whereas CatBoost requires 28.5 seconds - a 2,850 fold difference in time-to-train for models with at least equal predictive performance. For SMC and similarly sized ULBs lacking dedicated computing infrastructure, this result clearly supports OLS as the viable option for deploying an msw forecasting system.

C7 - proof of concept of lstm-cvrp (ijam, 2025)

This paper presents the first published integration framework of lstm-cvrp for optimizing msw routes in india. The study showed that integrating lstm-cvrp reduced distances traveled per day by 50.31%, and also reduced vehicle fleet size by 24.0% for a Surat scale network [8], while achieving r^2=0.93 using a time series chronological split validation framework. It is important to note that the results of this stage one work are considered to be qualitative "proof-of-concept" results relevant to the test network studied herein, rather than actual operational results for SMC's current live collection network.

C8 - gap analysis for efficiency of transfer stations

Haversine distance calculations were performed on the gps locations of the eight active transfer stations in Surat and found there exists an estimated 82% theoretical optimization opportunity for reducing distances between transfer stations (from 926.8 km/day to 166.8 km/day).

C10 - framework IMPACT -gps-free operational optimization

IMPACT represents the first gps-free msw operational optimization framework tested on a large-scale real dataset (662,803 records; nine zones; October 2025) and consists of five components: twr-a/b, VAO, MSRS, wdb+rb, FIO. Due to its gps-free design, IMPACT can be implemented in approximately 90% of Indian ULBs lacking gps telemetry systems.

C12 - quantification and capital reallocation case for vehicle surplus

ILP optimization performed via google or-tools cp-sat identified a total surplus of 260 vehicles within SMC's existing fleet. Additionally, reallocating these vehicles resulted in a -0.0008 % point change in missed-stop-rates. These results suggest that SMC's current investment strategy should shift from purchasing additional vehicles towards improving contract oversight systems.

C14 - architecture of Four-layer DSS

A complete Four-layer DSS architecture (Data Ingestion layer, model execution layer, analytics layer, user interface layer) was developed which described how OLS forecasting would integrate with each of the IMPACT components across three different user interface layers representing three distinct levels of stakeholders' interactions with the DSS.

C15 - DSS prototype implementation using streamlit

Streamlit version 1.28 has been utilized to develop a DSS prototype having six dashboard modules (executive dashboard, forecast dashboard, IMPACT dashboard, Compliance Dashboard, Route Dashboard, scenario Analysis Dashboard) supported by fourteen verified CSV files representing the input/output relationships between the various components of the IMPACT framework. The gps-free design enables users/stakeholders to engage directly with their respective stakeholders with regard to smart city planning/reporting purposes and smart city mission requirements.

C16 - Three-phase SMC deployment roadmap with role-based access control: Phase 1 (data infrastructure), Phase 2 (dashboard deployment and training), and Phase 3 (operational integration) with RBAC specified across three stakeholder tiers (route supervisors, zone management, municipal commissioners), aligned with Smart Cities Mission and SBM 2.0.

1.7.1 Environmental and Sustainability Contributions

C13 - Route Optimization Results in Environmental Benefits (Stage 1 CVRP): Optimizing routes using a representative Surat-scale synthetic network resulted in a yearly carbon dioxide savings of 46.8 tonnes and daily fuel reductions of 49.3 liters [8] — all environmental impacts are stated to be based upon a representation of the proof-of-concept network and therefore should not be taken as actual operational results for SMC’s live collection service.

C17 - Validation of an integrated framework with respect to Smart Cities Mission, SBM 2.0, and SDGs 11, 12, and 13: GPS-free deployment addresses SDG 11 (Sustainable Cities); 260 vehicle surplus quantification and demand-driven routing addresses SDG 12 (Responsible Consumption); and the 46.8 t CO_2 / year reduction addresses SDG 13 (Climate Action). Each claim refers to specific measured research-based outcome data as opposed to generic goal aligned language referenced in chapter seven.

1.7.2 Societal and Health Impact

Four types of impacts directly traceable to the empirical results of this study exist. Workload Distribution Balancing and Route Balancing (WDB + RB) resulted in removing all routes in Surat's waste collection network that were found to be over-loaded statistically (232 routes greater than 1.5 times the standard deviation away from the mean). The coefficient of variation of stop loads across the entire vehicle fleet decreased by 57.7%, from 1.42 to .60. This change in workload distribution has two major implications beyond simply increasing the operational efficiency of the collection route process. First, it improves the equity of how workloads are distributed among vehicles. Second, it also addresses the issue of environmental and health inequity. Overloaded routes typically lead to delayed or missed collections; these delays disproportionately affect specific geographic zones and communities, leading to increased levels of environmental and health inequities.

The first narrative provided by the compliance-first governance paradigm was implemented through the Decision Support System (DSS) allowing SMC to provide evidence based reasoning for shifting their focus from managing the fleet to holding contractors accountable. The Central Zone clearly represents the highest priority zone requiring a governance intervention due to its low contractor compliance coefficient ($\kappa = .31$) and high missed stop rates (.0931) even though the Central Zone had the lowest demand per vehicle ratio throughout the network. Therefore, a targeted enforcement strategy focused solely within the Central Zone would likely represent a lower-cost solution compared to expanding the size of the overall fleet throughout the network.

The non-GPS design architecture allows for an immediate application of the IMPACT framework and Streamlit DSS prototype into at least 90 percent of India's Urban Local Bodies (ULBs) who lack GPS telemetry. Prior to the implementation of the IMPACT

framework and DSS, Indian ULBs lacked access to the analytical optimized solutions available for improving waste management operations.

Lastly, repeated failures in performing routine waste collections create serious public health issues because they result in the creation of areas where pathogens can develop, increase opportunities for vectors such as rodents, cockroaches, etc., and degrade public space quality in high density urban areas. The research identifies contractor compliance as the primary upstream constraint preventing successful waste collection services and provides a governance-centered action framework that addresses the root cause of waste collection service delivery problems versus addressing only symptom(s).

1.8 Scope and Delimitations

1.8.1 Research Scope

Research is defined by the following geographical, temporal, methodological and publication boundaries. Geographical boundaries include Surat municipality, Gujarat State, India; specifically, this includes nine areas used for waste collection, thirty administrative wards, and approximately 130 administrative sub-ward units all under control of the Surat Municipal Corporation. Also, the eight transfer stations with confirmed GPS locations, and the Khajod waste disposal site were incorporated in the optimization infrastructure analysis.

Temporal boundaries include the 186 month time period between June 2009 and December 2024 as provided by the SMC and converted into the JETS canonical dataset of 419 engineered features. The entire 186 month time frame was utilized across the entire thesis. Temporal boundaries also include a COVID-19 disrupted-time series analysis in Section 4.5.3 using only the JSWTM 196 month raw dataset, but the 196 month is labeled as such in Section 4.5.3. Therefore it should not be referenced in any other sections or analyses in the thesis. The October 2025 SMC operational route data (662,803 records) established temporal boundaries for operational optimization analysis found in Chapter 5.

Methodological boundaries involved evaluating eighteen different machine learning and statistical models that represented six algorithms. Route optimization had a progression through three research phases; Phase One (LSTM-CVRP using a validated synthetic infrastructure of Surat, published in IJAM 2025 [8]), Phase Two (analysis of Haversine

distances using GPS coordinates at eight transfer stations), and Phase Three (IMPACT framework without GPS that has five sequential components utilizing October 2025 real data). The Decision Support System (DSS) component consisted of an architecture specification that included four layers, and a Streamlit prototype application with six dashboard modules and fourteen verified CSV data sources. Peer reviewed evidence based publications include JETS [6], JSWTM [7], and IJAM [8].

1.8.2 Delimitations

The delimitations that follow detail the specific limitations associated with the scope of this study. As such, the study is specifically confined to a single city, namely Surat. Although the methodology has been developed to be transferable, no studies have been undertaken to validate the applicability of the methodology in other cities, and this is clearly stated as a major area of further research in Chapter 8.

A 50.31% decrease in distances traveled and a corresponding reduction of approximately 46.8 tonnes CO_2 per year (based upon the 2019 emissions factor of 3.18 kg CO_2/tonne of waste collected) resulted from application of Stage 1 CVRP methodology (i.e., CVRP modelling) to synthetic, MoHUA/CPCB-validated Surat-scale collection infrastructure. This does not reflect travel over the operational routing network used by the Surat Municipal Corporation (SMC), nor does it represent "real-world" operational conditions. Therefore, all findings resulting from Stage 1 methodology are consistently described as "proof-of-concept" results throughout the study.

As part of addressing the limitation inherent in applying a simulated (synthetic) data set in Stage 1 (as opposed to actual operational data), the research deliberately employed a three-stage research design. Each successive stage represents increasing realism in terms of simulation: Stage 1 utilizes a simulated data set; Stage 2 employs real, operational infrastructure; and Stage 3 utilizes real, operational data.

In addition to computational validation on real operational data (see Section 5.7), the IMPACT framework and Streamlit DSS prototype were also not deployed within SMC during the course of this study. In order to clarify this point, the thesis explicitly describes the prototype as being a "proof-of-concept artefact"; therefore, the three-phase deployment plan presented in Chapter 6 represents the next step in deploying the prototype into

operational use. Real-time IoT sensor integration was not implemented. Finally, the forecasting component of the system utilizes monthly data granularity, which reflects SMC's long-standing historical reporting structure; however, variations in intramonthly demand are not currently included within the scope of this study and are identified as a potential future area of research.

While the research identifies multi-objective optimization (MOO) utilizing an ε-constraint Pareto frontier formulation across all of the IMPACT components (simultaneously optimizing across time window compliance, vehicle allocation, workload equity and demand integration) as a potentially tractable extension, it was determined during this research project to be beyond its scope. Instead, individual IMPACT components utilize single-objective formulations.

The complete set of limitations arising from these scope boundaries is formally documented in Section 8.4, and the corresponding future research agenda — prioritising multi-city validation of the compliance-dominance finding, IoT integration, multi-objective optimisation, and circular economy extensions — is set out in Section 8.5.

1.9 Thesis Organization

This thesis will be organised into eight chapters. Chapters 1–3 will provide the background and methodological foundation for the research. Chapters 4–6 will present the results of the three research components. Chapter 7 will synthesise the findings of the three research components. Chapter 8 will contain the conclusion of this thesis.

Table 1.1 below lists the research objectives, methodologies, data sources, and main products for each of the main chapters.

The next several paragraphs outline the purpose and contributions of each of the nine chapters.

Chapter 1 - Introduction Provides both international and Indian national perspectives on the MSW management issue. Identifies the Surat case study as the focus area for this research using specific operationally-collected data and identifies some of the problems that exist there. Defines the four dimensional research gap identified for this research. Outlines the

three frozen Research Objectives (RO-1, RO-2, RO-3) along with their associated Research Questions. Lists all eighteen original Contributions (C1-C18), organised by Contribution Type. Defines the boundaries of the research. Provides a summary of how this thesis is structured.

Chapter 2 - Literature Review Presents a systematic review of the MSW management literature. Organised around six themes: Evolution of MSW Generation Forecasting; Feature Engineering and Explainability Frameworks; Machine Learning and Deep Learning Model Comparison; Evaluation Metrics and Protocols; Analysis of the Impact of Covid-19 on MSW Management; Vehicle Routing Optimization Metaheuristics; Geographic Information System Applications in Waste Management; Decision Support Smart City Frameworks; Integrated Forecasting-Optimization-DSS Classification Frameworks. Concludes with an identification of five Major Research Gaps: Forecasting Benchmarking and Structural Break Methodology due to Covid-19; Feature Engineering and SHAP Attribution; ML-to-Routing Integration; Routing Framework Dependence on GPS Data; Design of Integrated DSS Systems. Maps each of these five Major Research Gaps to its corresponding Research Objective (via TABLE 2.11).

Chapter 3 - Research Methodology outlines an integrated research design that is composed of four phases. These include Phase 1 (Data Collection, Preprocessing and Feature Engineering), Phase 2 (Forecasting Framework – 17 Models), Phase 3 (Route Optimization - Three Stages and Design of IMPACT Framework), and Phase 4 (Design of Data Science Architecture and Prototype of the Streamlit DSS). Section 3.6 specifically resolves the differences observed between the two forecasts, i.e., IJAM time series forecast $R^2 = 0.93$ vs. Tabular Data $R^2 = -3.04$ as a function of model-context alignment.

Chapter 4 - Forecasting Model Development and Comparative Analysis represents the entire scope of RO-1. It includes comparative performance results of all 17 models tested over six algorithmic families, a three-tiered model classification system, Diebold-Mariano significance testing, computational efficiency testing, and SHAP feature importance results obtained by comparing each pair of models. Section 4.5.3 describes the COVID-19 ITS analysis using the JSWTM 196-month raw dataset. A key finding from this analysis is presented in Section 4.6 (Contributions C1-C6): OLS $R^2 = 0.9972$, MAPE = 1.43%, which demonstrated superior performance when compared to all other 17 models.

Chapter 5 - Operational Optimization Framework demonstrates the complete scope of RO-2 using a three-stage research development process. Stage 1 (IJAM 2025 [8]): Demonstrated LSTM R^2 = 0.93 and CVRP 50.31% distance reduction on a synthetic network. Stage 2: Identified an 82% theoretical optimization potential based upon analysis of transfer stations. Stage 3: Applied the IMPACT framework to real data containing 662,803 rows - identified compliance was dominant (75.72% SHAP), identified a 260-vehicle surplus (ILP OPTIMAL), reduced CV from 1.42 → 0.60. Contributions C7-C13, C18 are directly related to these findings.

Chapter 6 - Decision Support System Integration represents the complete scope of RO-3. This chapter presents the four-layer DSS architecture, six stakeholder-specific dashboard modules, and three tier persona design. Additionally, it describes the incorporation of FIO demand scalars into the DSS architecture as $S_d = \widehat{F_d}/F_{\text{base}}$. Also included in this chapter is a description of the Streamlit Proof-of-Concept prototype (six modules, fourteen verified CSV files), three-phase Software Maintenance Cycle (SMC) deployment plan with Role-Based Access Control (RBAC), and SDG Alignment Matrix with measured outcome linkages. Contributions C14-C17 are directly related to these findings.

Chapter 7 - Results, Synthesis, and Discussion presents a synthesis of the different components of the research study that differ from the individual chapters' analyses and discussions. The chapter identifies the quantitative relationship between forecasting and optimizing operations, compares the integrated framework proposed herein to existing published frameworks, identifies practical implications for Solid Waste Management Corporations (SMCs) and transferable applicability to similar cities, identifies SDG alignments with specific measured outcome linkages, and identifies institutional governance implications of compliance-dominant finding and capital reallocation example.

Chapter 8 - Conclusions and Future Research closes all three research objective with measurable outcomes, summarizes all eighteen contributions (C1-C18) related to the research objectives in a structured format, presents actionable recommendations for SMCs and other similar Indian municipalities, details the research limitations and identifies four future research areas including multi-city validations; integration of Internet of Things sensors and real-time data streaming architectures; multi-objective Pareto optimization

across components of IMPACT framework; and circular economy extensions connecting the framework to material recovery and waste hierarchy applications.

TABLE 1.1: *Research Objectives, Methods, Data Sources, and Key Outputs Across Thesis Chapters*

Research Objective	**Primary Methods**	**Data / Evidence Source**	**Key Outputs (with metrics)**	**Primary Chapter**
RO-1: Develop and comparatively evaluate a machine learning forecasting framework: 419-feature engineering, 17-model comparison, 8-metric evaluation, COVID-19 ITS analysis	17-model comparison (6 algorithm families); systematic 419-feature engineering; 8-metric evaluation protocol; Diebold-Mariano significance testing; interrupted-time-series COVID-19 analysis	JETS canonical dataset: 186 months (June 2009–December 2024), 419 features; JSWTM raw dataset: 196 months (§4.5.3 COVID analysis ONLY)	OLS best: R^2=0.9972, MAPE=1.43%; feature engineering > deep learning (LSTM R^2=−3.04); COVID ITS: 16% RMSE reduction (LR), 10% (XGBoost); Contributions C1–C6	Ch 4 (method: Ch 3)
RO-2: Design and evaluate a multi-stage route optimization framework: LSTM-CVRP proof-of-concept → transfer station assessment → GPS-free IMPACT framework	Stage 1: LSTM + CVRP (Google OR-Tools); Stage 2: Haversine distance on real GPS coordinates; Stage 3: IMPACT (TWR-A/B, VAO, MSRS, WDB+RB, FIO); ILP CP-SAT; XGBoost SHAP	Stage 1: raw SMC data + synthetic network; Stage 2: 8 real transfer station GPS coordinates; Stage 3: 662,803-row October 2025 real operational dataset	Stage 1: 50.31% distance reduction, 24% fleet reduction, 46.8t CO_2/year (synthetic); Stage 2: 82% theoretical potential; Stage 3: compliance 75.72% SHAP >> capacity 1.65%; 260-vehicle surplus; CV 1.42→0.60; Contributions C7–C13, C18	Ch 5 (method: Ch 3)
RO-3: Design and validate a GPS-free integrated DSS: four-layer architecture, FIO demand scalar, stakeholder interfaces, phased roadmap, policy alignment	Four-layer DSS architecture design; FIO demand scalar $s_d = \widehat{F_d}/F_{\text{base}}$; Streamlit PoC (v1.28); RBAC design; three-phase roadmap; SDG 11/12/13 outcome mapping	14 verified CSV data sources; IMPACT outputs; OLS forecast outputs; Smart Cities Mission and SBM 2.0 documentation; SDG 2030 Agenda	GPS-free 4-layer DSS; 6 dashboard modules; 3 stakeholder tiers; 3-phase deployment roadmap; SDG 11/12/13 validated with measured outcomes; Contributions C14–C17	Ch 6 (integration: Ch 5)

CHAPTER 2

Literature Review

2.1 Introduction

Waste management is currently among the greatest resource and governance issues facing municipalities globally today. As a result of increasing levels of global urbanisation, the quantity and variability of waste generated in municipalities is outgrowing what existing paradigms of waste management are able to manage effectively. Since more than half of all people live in urban environments today, and the resultant impact on municipal infrastructure such as sanitation services, waste collection, transfer stations, and landfills will be felt predominantly by municipalities in South Asia, sub-Saharan Africa and Latin America; given the confluence of rapid population growth, rapidly expanding informal settlement development, inadequate institutional capacities and severely constrained public funding in these regions, the operational challenges associated with managing municipal solid waste are both urgent and technologically difficult.

In terms of municipal solid waste generation forecasting it also takes on strategic importance. Forecasting allows municipal governments to forecast future demand on their collection infrastructure prior to deficiencies occurring operationally, allow them to calibrate the deployment of vehicles based upon actual waste generation volume rather than pre-determined schedules or routes, provide sufficient capacity for their transfer stations and meet the regulatory compliance required under national regulations governing municipal solid waste management. However, there continues to exist significant gaps between what reliable forecasting would enable municipal systems to accomplish and what they currently receive. In fact, nearly two-thirds of all urban solid waste management systems globally are operating on fixed schedule collection methods which bear no systematic relation to projected volumes of waste generation. The continued use of fixed scheduling methodologies despite dynamic, seasonally varied and pandemic disrupted waste generation

reflects a current unmet need at the nexus of machine learning, operations research and urban governance to integrate advanced demand forecasting with operational route optimization and stakeholder-focused decision-making into a deployable system.

There are many variables that contribute to the nature of municipal solid waste generation. Variables such as household income, population density, Gross Domestic Product (GDP), Urbanisation Rate, Seasonal Weather Patterns, Public Holidays and Consumer Behavior Changes all influence the amount of waste generated and do so in a manner that is often non-linear. The COVID-19 Pandemic clearly illustrated how large scale social disruption can significantly alter the amounts of waste generated. The engineering of informative features from these heterogeneous data sources, and the principled selection of those features that genuinely improve predictive accuracy rather than introducing noise, constitutes a research challenge that the existing literature has addressed only partially. The most comprehensive prior study identified in the evidence base for this chapter used more than fifty candidate variables [10]; the present thesis extends this to 419 engineered features within a systematic selection and attribution pipeline. That scale of feature engineering is, as the review in Section 2.3 establishes, unprecedented in the published MSW literature.

Route optimization in solid waste collection presents an equally substantive research domain. The vehicle routing problem and its extensions - the capacitated vehicle routing problem (CVRP), the inventory routing problem (IRP), and their multi-objective and stochastic variants - provide formal mathematical frameworks for improving collection efficiency. Metaheuristic solution methods including genetic algorithms, particle swarm optimization, and tabu search have been applied to MSW routing with documented gains in distance reduction, fleet utilization, and fuel consumption. The integration of machine-learning-driven demand forecasts into routing formulations, however, has occurred only in nascent form, with the most methodologically advanced example to date - the Geneva non-stationary stochastic IRP - employing statistical seasonal decomposition rather than multi-model ML forecasting as the demand input [11]. The specific challenges of routing in developing-world urban environments, where detailed Geographic Information System data are frequently unavailable, street-level geocoding fails on informal settlement addresses, and transfer station operations dominate the logistics topology, have received limited systematic attention.

Decision support systems occupy the third dimension of an integrated MSW management framework. The evolution from classical scenario-based planning models through IoT-enabled smart-waste platforms to real-time urban digital twin systems represents a substantial advance in capability, yet the published literature demonstrates that most deployed systems remain either operationally focused - routing optimization without strategic analytics - or strategically focused - macro-level planning without operational coupling. The most advanced integrated system identified in the evidence base for this chapter, the urban digital twin deployed for the City of Tshwane, South Africa [12], requires extensive IoT infrastructure unavailable in most Indian tier-two cities and does not incorporate multi-model machine-learning-based demand forecasting, formal COVID structural break analysis, or alignment with India's Swachh Bharat Mission and Smart Cities Mission policy frameworks. No published framework, as the analysis in Sections 2.9 and 2.10 of this chapter establishes, combines all required components within the operational realities of an Indian tier-two urban context.

Chapter 2 reviews the published literature across nine thematic domains that constitute the intellectual foundation of the present thesis. Sections 2.2 through 2.5 examine the evolution of MSW generation forecasting from classical statistical methods through ML and deep learning, covering feature engineering, model comparison, and evaluation protocols. Section 2.6 reviews COVID-19 impacts on MSW systems and the interrupted time series designs used to quantify structural disruption. Section 2.7 surveys vehicle routing and inventory routing formulations applied to solid waste collection. Section 2.8 examines GIS-based spatial applications with particular attention to the Indian urban context. Section 2.9 reviews decision support systems and smart-city platforms. Section 2.10 synthesises integrated forecasting–optimization–DSS frameworks through a five-level classification taxonomy. Section 2.11 provides critical analysis across seven thematic domains, and Section 2.12 presents the five major research gaps that motivate the three research objectives of this thesis.

As specified in FIGURE 2.1 below, these thematic domains intersect at three research objectives: RO-1 addresses forecasting and feature engineering (Sections 2.2, 2.3, 2.4, 2.5, 2.6); RO-2 addresses route optimization and spatial analysis (Sections 2.7, 2.8); and RO-3 addresses decision support systems and integrated frameworks (Sections 2.9, 2.10). The literature review is structured to serve each objective in sequence, establishing the evidential basis for the methodological choices made in Chapters 3 through 6.

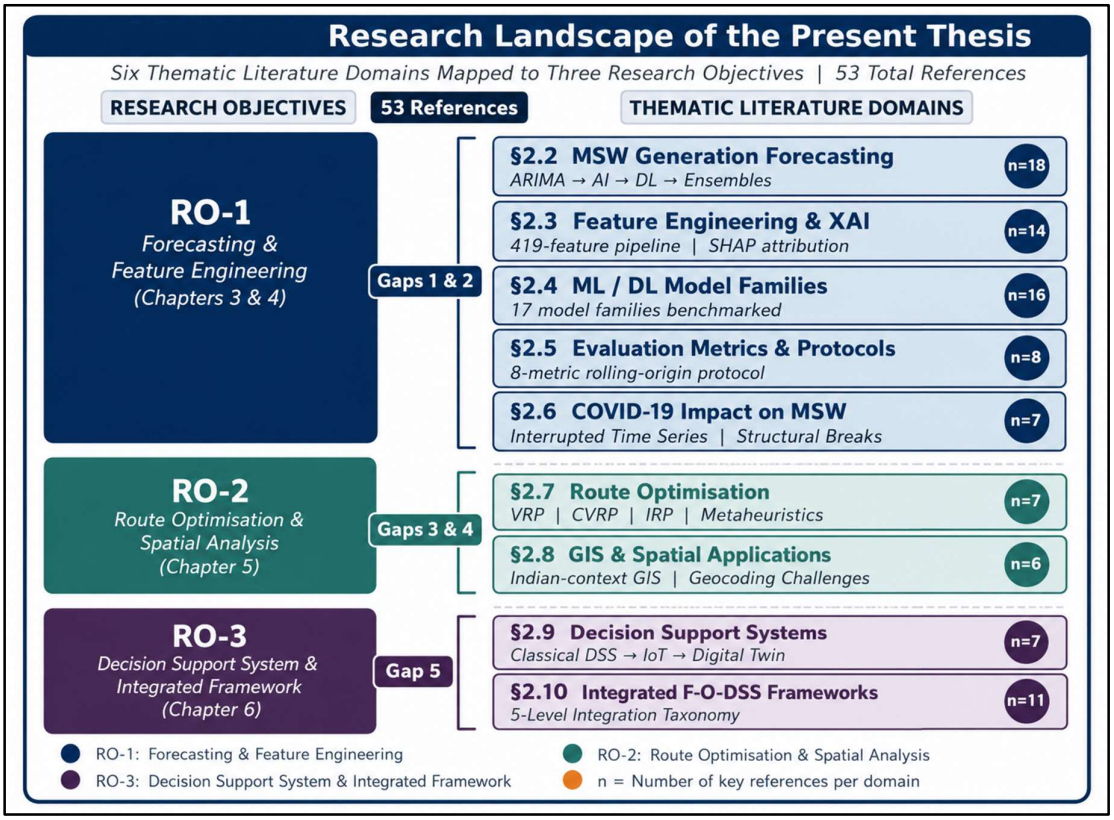


FIGURE 2.1 Research landscape of the Present Thesis

2.2 Municipal Solid Waste Generation Forecasting

The forecasting of municipal solid waste generation has evolved through three methodologically distinct phases over the past three decades: an initial period of classical statistical modelling, a transitional period in which machine learning approaches demonstrated systematic advantages over statistical baselines, and a contemporary period in which deep learning architectures and hybrid ensemble frameworks have pushed forecasting accuracy to levels previously unattainable. Developing an understanding of how far we have progressed toward the goal of developing a comprehensive model to describe the behavior of the soil gas is also important to establish what has been accomplished through the research reported in the literature. This section will trace the development of models based on the major families of methods used to develop these models. In addition to tracing the

development of models as it occurs within each family of methods, this section includes a final section that synthesizes the development of all model types and identifies those areas where additional work is required to develop a comprehensive model.

2.2.1 Traditional Statistical Approaches

By the end of the twentieth century and throughout the early years of the 21st century, traditional statistical techniques dominated in forecasting municipal solid waste generation. This was largely due to the lack of powerful computers available to municipal researchers and the lack of access by non-technical professionals to machine learning software. The use of AutoRegressive Integrated Moving Average (ARIMA) models based upon the Box-Jenkins methodology for univariate time series modeling was widespread for waste generation prediction. The major reasons for this popularity include the parsimonious nature of ARIMA; the well-established methods for determining whether residuals are autocorrelated; and the relative ease at which ARIMA parameters can be determined, estimated, and validated when working with small monthly datasets common in municipal archives. A SARIMA(p,d,q)(P,D,Q) model including an autoregressive component of order P, differencing of order D, and a moving average component of order Q as well as additional components to account for seasonality would represent a reasonable representation of short-term waste generation trends when demographic and economic conditions remain fairly stable.

In addition to accounting for short run changes, seasonal ARIMA (SARIMA) models also provide a way to handle seasonal patterns in waste generation found in many environments due to climate or culture. For example, festivals, school breaks and increased consumption during warmer months contribute to peak periods in waste generation but are often difficult to model using SARIMA without some type of seasonal differencing. In developing countries, there is often a presumption of stationarity of time series generated from historical records. Although theoretically it can be challenging to verify this assumption, in practice stationarity can be achieved using standard differencing procedures on short historical records. Multiple linear regression provided an alternative modelling approach that connected waste generation directly to observable socioeconomic predictors - population, household income, gross domestic product per capita, and urbanization rate - without

requiring the time-series machinery of Box-Jenkins identification. Polynomial trend analysis and log-linear regression permitted long-horizon projections from comparatively simple input requirements, making such models attractive for strategic municipal planning where computational capacity was constrained and long-run trajectories were of greater interest than short-term operational accuracy.

The addition of exogenous variables to ARIMA specifications, producing ARIMAX formulations, represented a partial synthesis of these approaches. Younes et al. [13] demonstrated the NARX (Nonlinear Autoregressive with Exogenous inputs) architecture as an early extension that permitted the modelling of nonlinear input-output relationships within a recurrent structure, providing forecasting accuracy superior to ARIMA for Malaysian municipal waste data where temporal self-correlation and socioeconomic predictors interacted in non-additive ways. This NARX application marked one of the earliest departures from strictly linear statistical modelling in the MSW forecasting literature, foreshadowing the subsequent ML transition.

The fundamental limitation of statistical methods, however, became apparent as urban waste generation systems grew in structural complexity. ARIMA and its variants presuppose stationarity - a requirement that is rarely satisfied in the time series of rapidly urbanizing cities, where population influx, income growth, policy changes, and shifts in consumption behavior continuously alter the underlying data-generating process. Multiple regression imposes linearity on relationships between predictors and waste generation that are demonstrably nonlinear: the effect of per-capita income on waste generation follows a curve that rises with early economic development and eventually plateaus as post-consumer recycling behaviors diffuse, a relationship that linear models represent only as a first-order approximation. Adamović [14] provided empirical evidence showing how structural changes (i.e., discrete changes to the mean or trend of the underlying data generating process) at specific calendar dates violated the assumptions of time-invariant stationary conditions found within classical forecasting methodologies and resulted in forecast deterioration at the exact point(s) of potential greatest consequence for operational decision-making. Adamovic's results regarding the superior performance of machine learning based models sensitive to structural breaks versus traditional ARIMA-based forecasting techniques during periods of significant economic transition established a paradigmatic basis for using machine learning as an alternative forecasting methodology.

2.2.2 Machine Learning-Based Forecasting

The transition from statistical to machine-learning based forecasting in MSW area has been supported by the foundational comparative study conducted by Abbasi and El Hanandeh [15]. They compared Artificial Neural Networks (ANN), Support Vector Machines (SVM), Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and ARIMA model as base line, using historical data from Logan City, Australia. Their main finding — that all three families of AI models are better than ARIMA for long term forecasting of MSW; and that ANFIS is slightly better than others thanks to its hybrid fuzzy logic architecture which allowed to process both linguist and numeric information contained in socio-economic driving factors — defined a pattern that has been widely used by many researchers since then. The impact of the study of Abbasi and El Hanandeh does not lie only on the results obtained; but also on the methodology they proposed: namely the use of the same dataset, and the same evaluation protocol, to compare several families of AI models at the same time, giving special emphasis to socio-economic and time-related factors.

Following Abbasi and El Hanandeh, ensemble tree methods have progressively replaced single algorithmic approaches as the primary choice of machine learning architectures for MSW forecasting. Among these, Random Forest (RF) has become the most commonly used one among ensemble methods due to its ability to provide competitive level of predictive accuracy while allowing to easily evaluate the importance of features of the input data by means of out-of-bag permutation scoring. Yang et al. [16] used a multi-model machine learning approach — combining Random Forest with XGBoost and other ensemble variants — for Chinese municipal waste data; and integrated SHapley Additive exPlanations (SHAP) to determine the most relevant predictors of waste generation within each family of models. Their study demonstrated how gradient-boosted tree methods produce more precise point forecasts than Random Forest alone for the non-linear, multi-predictors structure typical of Chinese urban waste data. Zhang et al. [17] further generalized this approach, introducing as external variables the Shared Socio-Economic Pathways (SSP) scenario variables. Thus, they generated probabilistic MSW trajectories under different future development scenarios. XGBoost, LightGBM and CatBoost were specifically highlighted for their efficiency in terms of computing resources needed when dealing with high dimensional feature sets, as well as their ability to natively handle missing values and categorical variables; and their

advantages in terms of generalization capability with respect to single decision-tree models in structured table-based prediction problems.

Nguyen et al. [18] compared six machine learning algorithms — including RF, XGBoost, SVM, and ANN variants — on Vietnamese municipal waste data; and found that XGBoost and Random Forest show the best consistency in terms of performance across cross validation folds. Lu et al. [19] introduced a weather-augmented gradient-boosted regression tree model called WGMod; and showed that adding meteorological variables improves significantly forecast accuracy in cities showing strong seasonal variations. Namoun et al. [20] applied Bayesian hyperparameter optimization via the Optuna framework to tuned gradient boosting ensembles, demonstrating that systematic hyperparameter search produced accuracy gains beyond those achievable through manual tuning.

2.2.3 Deep Learning-Based Forecasting

The application of deep learning architectures to municipal solid waste generation forecasting represents the most substantive methodological advance in the field during the 2016–2024 period. Long Short-Term Memory (LSTM) networks, which implement gated memory cells that can retain relevant information across extended temporal spans while selectively forgetting irrelevant context, were adapted for MSW time-series prediction on the basis of their structural advantage in capturing long-range temporal dependencies. Gated Recurrent Units (GRU), which implement a computationally lighter version of the gating mechanism with fewer parameters than LSTM, attracted parallel interest as an accuracy-efficiency trade-off for moderate-scale municipal datasets.

A comparison study by Mudannayake et al. [21] , was an analysis of all prior studies of machine learning (ML) and deep learning (DL) architectures on Municipal Solid Waste (MSW) forecasting. The authors evaluated nine ML/DL families (ANN, SVR, RF, KNN, XGB, ARIMA, LSTMs, GRUs, CNN-LSTMs), at multiple city locations with different forecasting horizons. Their results showed that the two DL families (LSTMs and GRUs) consistently had better performance than their traditional statistical counterparts (ARIMA and others) and shallower ML families (RF and SVR) in monthly forecasts of municipal solid waste. Therefore, this study provided the basis for future comparisons using DL. Cubillos et al. [22] extended the application of LSTMs from single-site waste forecasting to

multi-site household waste forecasting, across multiple Danish municipalities. This study demonstrated that DL can be used to analyze data generated from multiple locations with common characteristics by encoding each location's unique features into the input layer(s). Niu et al. [23] conducted a structured comparison of LSTM, ARIMA, and ANN for Chinese provincial MSW data, confirming deep learning superiority.

Hybrid architectural strategies extended deep learning performance along two complementary paths. For accuracy and interpretability, Lin et al. [19] augmented a CNN-LSTM pipeline with self-attention for daily Shanghai waste data, demonstrating that attention layers selectively re-weighted historical time steps to improve prediction accuracy; Ghanbari et al. [24] combined EEMD signal decomposition with MARS regression; and Jayaraman et al. [25] proposed MSWDeepStack - a stacking ensemble of LSTM, GRU, and CNN-LSTM under a gradient-boosted meta-learner - confirming that ensemble combination of architectures outperformed any individual constituent. Uncertainty quantification entered the domain through the Bayesian ANN ensemble of Xia et al. [26]. For pandemic robustness, Wan et al. [27] demonstrated that models trained with explicit COVID-19 disruption indicators substantially outperformed pre-pandemic-trained counterparts, establishing that a three-phase temporal structure - pre-COVID baseline, lockdown disruption, and post-lockdown recovery - was necessary to capture the pandemic's effect on waste generation, a finding that directly motivates the ITS structural break analysis of this thesis.

2.2.4 Evolution of Forecasting Research

The reviewed literature illustrates a progression from one stage to another by way of five stages of methodological development in forecasting municipal solid waste (MSW) generation; that is, each stage builds upon the previous as well as addresses the inherent limitations of the previous stage. Table 2.1 provides an overview of these progressive stages of development, along with representative model families, common input categories for MSW models, and key attributes of each stage.

TABLE 2.1 Evolution of Municipal Solid Waste Forecasting Research

Era	Dominant Model Families	Typical Input Categories	Key Advantages	Characteristic Limitations
Pre-2010 (Statistical)	ARIMA, SARIMA, OLS regression, polynomial trend, ARIMAX	Population, GDP, historical waste volumes	Interpretable; low computational demand; well-established diagnostic protocols	Linearity and stationarity assumptions; no feature interaction; structural break sensitivity
2010–2016 (Early AI / Soft Computing)	ANN, SVM, ANFIS, NARX	Socioeconomic indicators, basic temporal lags	Outperforms ARIMA; captures nonlinearity; accommodates exogenous inputs	Black-box output; small feature sets (typically 3–8 variables); no systematic feature selection; no SHAP
2016–2022 (ML Ensembles)	Random Forest, XGBoost, LightGBM, CatBoost, SVR, GBM	Population, GDP, income, weather, festival calendars, policy variables	High accuracy on tabular data; robust to outliers; native importance metrics	Temporal dependency treated implicitly via lags; importance metric bias
2018–2024 (Deep Learning)	LSTM, GRU, CNN-LSTM, attention hybrids, MLP, stacking ensembles	Temporal sequences, socioeconomic + demographic + weather	Explicit sequence modelling; multi-step capability; attention-based interpretability	Data-hungry; computationally intensive; interpretability deficit; overfitting risk on short archives
2022–present (Integrated Analytics)	SHAP-augmented ML, Prophet+ML hybrid, MI-SHAP pipelines, large-feature-set ensembles	Full feature pipelines: demographic + weather + COVID + temporal + policy + lag + rolling statistics	Integrates prediction and attribution; COVID-aware; systematic feature selection	Few published studies; no 17-model benchmark; limited developing-country context

2.2.5 Critical Analysis

In addition to three substantive limitations that exist in the body of literature for forecasting studies as described above, the aforementioned three limitations provide the basis for the research objectives identified in this paper and have therefore provided motivation for conducting Research Objective 1. There are no peer-reviewed studies that evaluate ten or more different types of model families using one data set of municipal solid waste under one multi-metric evaluation criteria -- specifically Ridge, Lasso, Elastic Net, and Gaussian Process Regression, and LSTM and GRU, with respect to each other, and all under conditions of an Indian Tier Two City with rich features (Gap F1). In addition there are no peer-reviewed studies that integrate formal Interrupted Time Series COVID-19 Structural Break analysis methodologies into a comprehensive Machine Learning Pipeline (Gap F2). Finally, it should be noted that the bulk of the existing literature regarding the use of machine learning in modeling municipal solid waste generation is based upon either developed economy or Chinese urban settings; therefore there is limited representation from India's Tier Two Cities (Gap F5). Full critical development of these limitations can be found in section 2.11.1.

TABLE 2.2 Summary of Key Municipal Solid Waste Generation Forecasting Studies

Author(s)	Year	Study Area	Method	Data / Frequency	Key Findings	Limitations
Younes et al. [13]	2015	Malaysia	NARX	Hospital waste; monthly	NARX outperforms ARIMA; exogenous inputs critical	Single institution; limited to healthcare waste
Abbasi & El Hanandeh [15]	2016	Logan City, Australia	ANN, SVM, ANFIS vs ARIMA	Socioeconomic; monthly	All AI models outperform ARIMA; ANFIS marginally best	3 model families; < 10 variables; single city

Adamović [14]	2017	EU-27 countries	ANN with structural break analysis	Annual macro-economic data	Structural breaks degrade classical ANN; break-sensitive models outperform	Annual resolution; no COVID era
Singh & Satija [28]	2018	Faridabad, India	ANN	Municipal records; monthly	ANN achieves competitive accuracy for Indian tier-2 city	Single model; no DL or ensemble; no SHAP
M. Cubillos et al. [22]	2020	Denmark (multi-site)	LSTM	Household waste; weekly	LSTM handles multi-site data; multi-scale prediction demonstrated	No routing integration; no COVID
Ayeleru et al. [29]	2021	Johannesburg, SA	ANN vs SVM	Municipal data; monthly	ANN marginally superior to SVM in developing-city context	Two models only; no SHAP
Yang et al. [16]	2021	Multiple Chinese cities	RF, XGBoost + SHAP	Socioeconomic; monthly	SHAP identifies GDP and population as top drivers	Limited model families; no COVID
Niu et al. [23]	2021	China	LSTM vs ARIMA vs ANN	Monthly generation data	LSTM consistently outperforms ARIMA and ANN	No DL hybrids; no COVID

Nguyen et al. [18]	2021	Vietnam	6 ML models	Municipal waste; monthly	XGBoost and RF most robust across folds	Single context; no SHAP
Mudannayake et al. [21]	2022	Multi-city	9 ML/DL models	Multiple city monthly records	LSTM/GRU best for monthly MSW; 9-model benchmark	No OLS/Ridge/GPR; no SHAP; no COVID; no routing
Lin et al. [30]	2022	Shanghai, China	Attention CNN-LSTM	Daily waste generation	Attention mechanism improves accuracy and provides approximate interpretability	No multi-city validation; no COVID
Zhang et al. [17]	2022	China (provincial)	ML ensemble + SSP scenarios	Socioeconomic + SSP pathways	SSP variables improve long-horizon projections	Scenario uncertainty not quantified
Lu et al. [19]	2022	Multiple Chinese cities	WGMod GBRT + weather features	Meteorological + demographic; monthly	Weather inputs measurably improve accuracy	No SHAP; no COVID variables
Namoun et al. [20]	2022	Multi-city	Optuna-tuned gradient boosting	MSW generation; monthly	Hyperparameter optimization improves accuracy	No SHAP; single evaluation metric
Xia et al. [26]	2022	China	Bayesian ANN ensemble	MSW generation; monthly	Bayesian uncertainty quantification provides calibrated	No SHAP; no DL comparison

					prediction intervals	
Ghanbari et al. [24]	2023	Multi-city	EEMD-MARS hybrid	MSW time series	Pre-processing with EEMD improves MARS accuracy	No SHAP
Wan et al. [27]	2024	City TBC	Prophet + ML hybrid	MSW + COVID-era data	COVID-aware ML outperforms pre-COVID models; three-phase temporal structure required	No formal ITS; no SHAP on COVID variables
Jayaraman et al. [25]	2025	Singapore	MSWDeepStack (LSTM+GRU+CNN-LSTM stacking)	Municipal waste; monthly	Stacking ensemble outperforms individual DL base models	Single city; no SHAP; no COVID

2.3 Feature Engineering and Explainable Artificial Intelligence

The process of feature engineering - i.e., the systematic manipulation (transformations), creation (synthesis) and selection of input variables for predictive models - is considered to be the dominant factor influencing the performance of machine learning models on structured datasets found in real world environments; therefore, it generally takes precedence over architectural design when working with tabular prediction problems. As such, the principles outlined above have significant applicability to municipal solid waste generation forecasting. The quantities and compositions of waste generated by an urban area's population are the end product of a complex system driven by multiple socio-economic, demographic, environmental and behavioral factors, where each factor produces signals that are dispersed among various data sources with different temporal resolutions, spatial

aggregations, and levels of reliability. The subsequent sections detail the types of variables used, temporal engineering techniques, feature selection techniques and explainability methodologies currently employed in MSW feature engineering.

2.3.1 Importance of Feature Engineering in MSW Forecasting

In addition to providing an extensive review of the published literature, it has been clearly shown how accurately the accuracy of the predictions is directly dependent upon the quality of the features used. This was illustrated by Abbasi and El Hanandeh [15] , where they showed that adding socioeconomic attributes to their models along with time-based features resulted in measurable improvements in the accuracy of the results from the AI model compared to those results produced by models that only included temporal based inputs. Additionally, Yang et al. [16] were able to show that, according to SHAP values ranking, socioeconomic inputs significantly exceeded temporal based input contributions toward improving XGBoost model predictive accuracy. Lastly, Zhao et al. [10] were able to demonstrate the superiority of using a larger, uncurated collection of over fifty potential inputs toward producing superior predictive accuracy versus utilizing smaller curated collections in a provincial China setting. On the other hand, Z. Ceylan [31] demonstrated the converse argument; i.e., that a small collection of only four well-selected socioeconomic attributes generated exceptional levels of predictive accuracy ($R^2 = 0.98 – 0.99$), for Turkish Provinces when appropriately regularized Bayesian Group Penalized Regression was utilized. Therefore, collectively, these studies established that feature creation may be equally important as modeling in regards to achieving high performance MSW forecasts.

2.3.2 Socioeconomic and Demographic Variables

Socioeconomic and demographic variables constitute the most extensively studied category of predictors in MSW forecasting, reflecting the longstanding theoretical understanding that waste generation is fundamentally a function of economic activity, population size, and consumption behavior. Smailbegovic et al. [32] conducted a European-scale multi-country ML and SHAP analysis covering more than thirty countries, identifying annual wages and urban population ratio as the two strongest SHAP contributors to MSW generation across the entire sample, providing cross-regional evidence for the primacy of income and

urbanization as predictors. Zhao et al. [10] substantially expanded the scope of socioeconomic variable exploration in a Chinese provincial MSW study, incorporating more than fifty candidate inputs across categories including land-use intensity, transportation infrastructure access, service sector employment, tourism activity, and policy implementation indicators. Singh and Satija [28] contributed the Indian-context evidence, confirming that the socioeconomic driver structure documented in developed-country and Chinese studies is transferable to Indian urban contexts.

2.3.3 Environmental and Climatic Variables

Environmental and climatic variables have received increasing attention in MSW forecasting as researchers have recognized that waste generation exhibits measurable seasonal variation driven by temperature, precipitation, humidity, and associated behavioral changes. Lu et al. [19] provided the most systematic evidence for the predictive value of meteorological inputs in their WGMod GBRT framework, demonstrating through an ablation study that weather features - specifically temperature, precipitation, and wind speed - improved forecast accuracy for Chinese cities with marked seasonal weather variation. Smailbegovic et al. [32] confirmed the relevance of climatic indicators at the European multi-country scale, finding that weather-related variables contributed to SHAP rankings in countries with pronounced seasonal contrasts but were subordinate to socioeconomic drivers at the aggregate cross-country level.

2.3.4 Temporal Feature Engineering

Temporal feature engineering -- constructing lag-variables; rolling-statistical measures; trend-indicators; and calendar-encoded variables systematically from the raw time series of waste-generation data -- is considered to be a major method whereby Machine Learning Models capture the autoregressive and seasonality inherent within municipal waste generation data.

Nourani et al.[33] have developed the most systematic procedure for selecting temporally-related variables in the literature for municipal solid waste (MSW) forecasting. They utilized a two-stage filter-wrapper pipeline involving both Mutual Information (MI) screening and

SHAP based refinement, and showed that their selected lag-structure provided greater out-of-sample generalization than did either the complete lag-set or some arbitrary lag-window. This demonstrates the theoretical basis for the use of systematic methods to select lags as opposed to simply choosing one using experience or intuition.

Namoun et al. [20] also demonstrated that rolling means and variances computed at different window-lengths contributed significantly to the overall performance of the ensemble models.

The cross-domain literature on Time-Series Feature Engineering includes additional evidence regarding the effectiveness of various feature-engineering techniques for improving the predictive accuracy of Water Demand Forecasting. For example, Guo et al. [34], used Discrete Wavelet Transform Decomposition as part of their modeling process (combined with Random Forest and Temporal Convolutional Network Modeling), and found that breaking down the demand series into separate components greatly improved forecasting accuracy. W. Hao et al., [35], who were analyzing Urban Water Demand in Milan utilizing Wavelet Decomposition, LightGBM, and Long Short-Term Memory Networks (LSTM), demonstrated that using rolling window-features computed on decomposed series resulted in superior results compared to computing such features on the original series.

2.3.5 Feature Selection Methods

A fundamental challenge in the forecasting literature of municipal solid waste (MSW) is to select from a large set of potential candidate features the most predictive subset. The literature has identified four main methodological families of techniques used to address this challenge: Filter Methods (Statistical Association Metrics); Wrapper Methods (Recursive Model Evaluation); Embedded Methods (Selection as Part of the Training Objective), and Projection Methods (Principal Components Analysis (PCA)). Prior to using k-means clustering and SHAP attribution to identify clusters in their South African MSW data, Adeleke and Jen [36], reduced the dimensionality of their twelve variable socio-economic input dataset with PCA. They demonstrated that PCA effectively concentrated the predictive variation of their original dataset into fewer orthogonal dimensions. Using Mutual Information as a filter step to find nonlinear statistical relationships between the candidate

features and the prediction targets, Nourani et al. [33]. In addition, Zhang et al. [17] identified correlation filters and Random Forest Permutation Importance as the two most frequently used feature selection techniques across the larger body of machine learning literature pertaining to MSW. Documenting a consistent pattern where published studies either use simple correlation-based feature selection or include all features available without formally selecting them, Liang et al. [37] documented their findings by surveying feature selection in the larger solid waste machine learning review literature.

2.3.6 Explainable Artificial Intelligence and SHAP

The adoption of explainable artificial intelligence (XAI) methods in MSW forecasting has accelerated substantially since 2021, driven by both technical advances in attribution methodology and by growing recognition that municipal governance contexts require model outputs to be interpretable by non-technical decision-makers. SHapley Additive exPlanations (SHAP), derived from cooperative game theory's Shapley values, has emerged as the dominant XAI framework in the MSW forecasting literature because of its theoretically grounded properties - efficiency, symmetry, dummy, and linearity - that collectively ensure that the attribution of the prediction to individual features is both internally consistent and directly comparable across model families.

Collectively, these SHAP studies established three methodological properties of direct relevance to this thesis. Global–local attribution consistency - demonstrated by Adeleke and Jen [36] for their PCA+ANFIS pipeline in South Africa and extended by Smailbegovic et al. [32] across more than thirty European countries, where annual wages and the urban-to-rural population ratio emerged as the two dominant cross-national drivers - provides assurance that SHAP captures genuine feature influence rather than model-level averaging artefacts. Non-linear interaction detection by Yang et al. [16] revealed through SHAP interaction plots that GDP's effect on waste generation is amplified in high-density urban zones, demonstrating a diagnostic capability inaccessible to standard importance rankings. Cross-architecture stability was confirmed by Nourani et al. [33], who found that importance rankings for their MI-selected feature set were consistent between FFNN and LSTM architectures - suggesting that importance structure reflects data properties rather than model-specific biases - while Que et al. [38] confirmed SHAP’s applicability to non-

stationary urban resource flow time series in the water demand domain, establishing methodological precedent directly relevant to this thesis.

2.3.7 Critical Analysis

The reviewed literature on feature engineering and explainable AI in MSW forecasting has established several important empirical foundations. Four substantive gaps, however, persist. Gap FE-1 identifies the systematically narrow exogenous feature sets (≤50 variables in all reviewed studies vs. 419 in this thesis). Gap FE-2 identifies the limited encoding of COVID-19 and policy variables as systematically engineered multi-phase features. Gap FE-3 identifies the absence of systematic multi-window rolling statistics exploration comparable to established practice in water demand forecasting [34], [35]. Gap FE4 identifies the limited embedding of SHAP as a core, cross-model pipeline component: every published MSW study applying SHAP does so within a single model family, preventing cross-architecture attribution comparison. The 419-feature pipeline with cross-model SHAP attribution developed in Chapter 4 constitutes the primary response to all four gaps.

TABLE 2.3 Feature Engineering Approaches in Municipal Solid Waste Forecasting Studies

Study	Variable Categories	Feature Engineering Method	Feature Selection Method	Interpretability Method	Key Limitation
Abbasi & El Hanandeh [15]	Socioeconomic (GDP, population, income); temporal lags	Basic lag construction	Domain knowledge selection	None	< 10 variables; no SHAP
Yang et al. [16]	Socioeconomic + demographic; temporal	Lag variables; seasonal dummies	RF importance ranking	SHAP (single model family)	Limited variable pool; SHAP on ML only
Z. Ceylan [31]	4 socioeconomic variables	Direct inclusion	Theory-driven + BGPR regularization	None	Only 4 variables; no

					temporal engineering
Zhang et al. [17]	Socioeconomic + SSP scenario variables	Scenario pathway encoding	Domain knowledge + correlation	None	Scenario-dependency uncertainty not quantified
Singh & Satija [28]	Population, income, expenditure (India)	Direct inclusion	Domain knowledge	None	Very limited feature engineering
Lu et al. [19]	Meteorological + demographic	Weather feature integration via GBRT	Ablation study	None	No SHAP; no policy variables
Smailbegovic et al. [32]	Socioeconomic + urbanization; multi-country	Panel data feature construction	Statistical pre-screening	SHAP (KRR model)	European context; no temporal lag engineering
Namoun et al. [20]	Temporal lags; rolling statistics	Multi-window rolling mean and variance	Optuna hyperparameter optimization	None	No socioeconomic or weather categories
Nourani et al. [33]	Temporal lags; socioeconomic inputs	Systematic lag construction; MI pre-filtering	Mutual Information filter + SHAP wrapper (MI-SHAP)	SHAP (FFNN and LSTM)	Two model families only; no rolling statistics
Zhao et al. [10]	50+ variables: socioeconomic + land-use + tourism + policy	Large candidate pool generation; BP-NN training	Implicit through BP-NN weight optimization	None	No SHAP; no COVID
Adeleke & Jen [36]	12 socioeconomic + demographic variables	PCA dimensionality reduction + k-means	PCA-based dimensionality reduction	SHAP (global + local; ANFIS only)	12 variables only; single model family
Wan et al. [27]	Socioeconomic + COVID	COVID dummy	Domain knowledge +	None	No systematic feature

	policy indicators + temporal	variables; Prophet changepoint features	automated changepoint		selection; no formal ITS
Liang et al. [37]	Full variable taxonomy (review scope)	Review of feature types across MSW studies	Review finding: correlation and RF dominate	Review finding: SHAP emerging	Review study - no original feature engineering

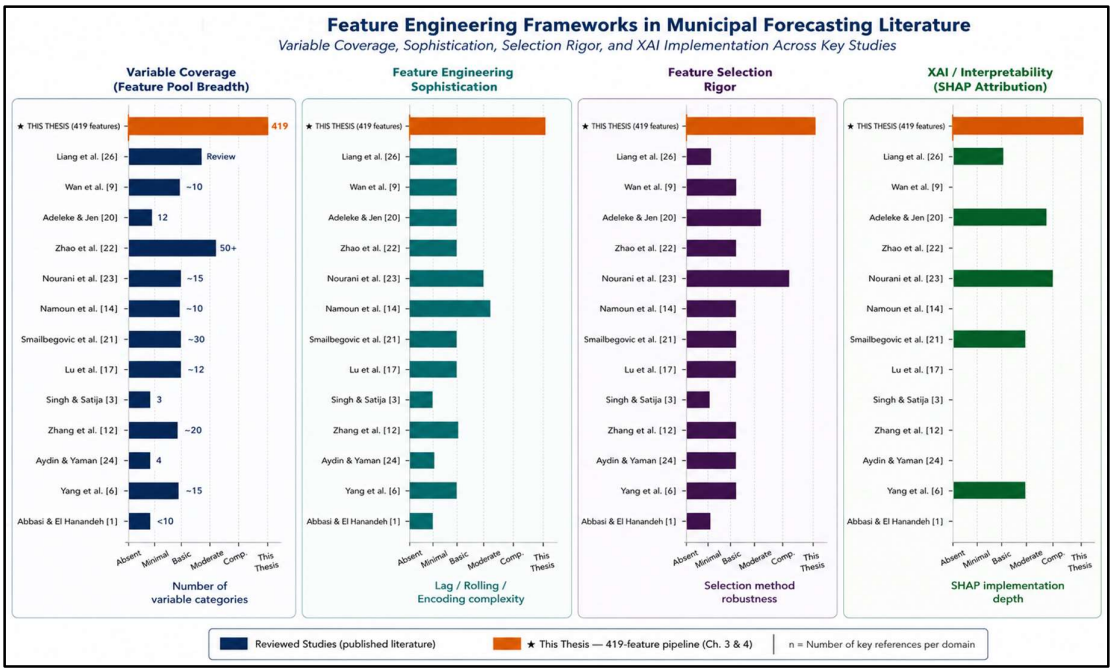


FIGURE 2.2 Feature engineering frameworks in MSW forecasting literature:

Fig. 2.2 shows variable coverage and pipeline completeness across key studies, compared with the 419-feature pipeline proposed in this thesis.

2.4 Machine Learning and Deep Learning Models for MSW Forecasting

Section 2.2 traced the chronological evolution of municipal solid waste generation forecasting from classical statistical methods through machine learning and deep learning architectures. Presently, this section is designed to provide an analytical technical comparison of the structural properties of the model families that are widely used in the literature regarding forecasting of municipal solid waste (MSW) as well as their comparative performance within several benchmark studies and the balance between interpretability and

computation that affects the choice of which families to use in operationally-based forecasting systems.

2.4.1 Supervised Machine Learning Models

Random Forest builds an entire forest of many different decision trees. Each of these decision trees is created using a random selection of samples from the original training set. When creating each internal node, Random Forest will choose only from the group of features that have been randomly chosen at that time. The practical benefits of Random Forests include being able to handle many different inputs (features) and outliers fairly well. Random Forests can also be used to get an idea of which features are important in a model by calculating how much the mean decrease in impurity is for each feature. For many years Random Forest has been one of the best methods for making predictions on a tabular input. Many researchers report good results with Random Forests and they don't require a lot of work in order to optimize their hyperparameters.

Nguyen et al. [18] ranked Random Forest as second-best in their evaluation of six models (the first-best was Decision Tree), in terms of average performance on five Vietnamese MSW datasets. Yang et al. [16] applied both Random Forest and XGBoost to Chinese MSW data, and reported that while their SHAP scores differed significantly between models (XGBoost generated larger absolute values than did RF), their Mean Decrease in Impurity scores agreed moderately closely.

Support Vector Regression applies the structural risk minimization principle to regression problems. It seeks to find a predictive function that lies entirely inside an epsilon insensitive band defined by all points in the training data.

Abbasi and El Hanandeh [15] showed Support Vector Regression (SVR) to be a strong competitor against other machine learning approaches for predicting MSW generation, including ARIMA and ANFIS. Specifically, Abbasi and El Hanandeh found SVR to generate more accurate predictions than ARIMA but less accurately than ANFIS based upon data collected in Logan City. Later, Ayeleru et al. [29] compared ANN and SVM for the case of South Africa's waste management practices. They concluded that ANN performed slightly better than SVM.

Between 2017-2022 the gradient boosting family became the most widely-used supervised machine learning methodology for predicting MSW generation. XGBoost extended basic gradient boosting by adding penalty terms to regularize the complexity of individual trees and to allow for parallel processing of candidates for new tree nodes. Both Yang et al [16] nd Nguyen et al. [18] independently determined XGBoost to be either the top or co-top performer among various competing machine learning methodologies in their studies comparing multiple machine learning methodologies on Chinese and Vietnamese MSW datasets. Additionally, Jayaraman et al. [39] demonstrated that employing a Gannet search algorithm for optimizing the hyperparameters for XGBoost led to improvements in accuracy over that obtained through manual or Optuna-based hyperparameter optimization techniques for Singapore municipal waste data. LightGBM enhanced the gradient boosting family by providing a histogram-based method for quickly selecting potential nodes for splitting trees during training and by allowing for leaf-wise tree construction. Lu et al. [19] further demonstrated in their WGMod GBRT framework that gradient boosting ensemble models produce their characteristic improvements in accuracy primarily due to the quality of the feature engineering provided to support them. Finally, Namoun et al. [20] showed that applying Bayesian hyperparameter optimization using Optuna resulted in significant improvements in accuracy relative to manually-tuned gradient boosting.

2.4.2 Deep Learning Models

The most common forms of deep learning are used to make MSW generation forecasts. Feed-forward neural network (ANN) and multi-layer perceptrons (MLP) are the simplest forms; recurrent networks (RNNs), such as long short-term memory (LSTM) or gated recurrent units (GRU); and hybrid models using convolutional feature extraction techniques combined with RNNs for time-series modeling.

Multilayer Perceptron (MLP) artificial neural networks were among the earliest alternative methods to ARIMA developed specifically for use in the MSW domain. Singh and Satija [28] confirmed that MLPS can provide non-linear function approximations of tabular socioeconomic predictors superior to traditional linear statistical approaches. This established the architectural basis for subsequent recurrent and deep learning architectures.

LSTMs are able to handle problems associated with vanishing gradients when working with long sequences. LSTMs do so by introducing three gates per cell (i.e., input, forget and output). For MSW generation at the monthly level - which will include both annual demographic changes and decade-long economic regime shifts - LSTMs are well-suited to handle these large-scale temporal structures. Mudannayake et al. [21], in comparing nine different models on forecasting tasks across multiple cities and multi-step ahead predictions, identified LSTMs as being the best performing model overall. Cubillos et al. [22] extend the LSTM architecture to forecast household waste from multiple sites in Denmark.

GRUs reduce the number of parameters needed for an LSTM by replacing the separate input and forget gates with a single "reset" gate. Mudannayake et al. [21] report that the performance of GRUs are similar to those achieved using LSTMs when forecasting monthly MSW generation.

There have been two main ways researchers have attempted to create hybrid deep learning architectures. First, many researchers have implemented convolutional neural networks (CNNs) in combination with RNNs (e.g., CNN-LSTM): Lin et al. [30] implement CNN-LSTMs to improve their forecast performance on daily MSW generation in Shanghai. They demonstrate that adding self-attention layers can also provide additional increases in performance beyond what was possible with a standalone LSTM. Second, some researchers have used ensembling-based methods to combine the outputs of multiple base models together using stacking-based architectures: Jayaraman et al. [25] provide an example of how they created a MSWDeepStack framework that uses stacking ensembles to combine the outputs of base models composed of LSTMs, GRUs and CNN-LSTMs via a gradient-boosted learner. Similarly, Xia et al. [26] use Bayesian ANNs to quantify the uncertainty associated with an ensemble of deep learning models.

2.4.3 Comparative Performance Trends and Model Ranking Evidence

Comparative empirical research provides a basis for making several generalizations about the relative effectiveness of various models in forecasting Municipal Solid Waste (MSW). First, all of the reviewed comparisons show that Deep Learning (DL) techniques-primarily Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU)-are superior to traditional statistical techniques for forecasting MSW regardless of whether they were tested as part of

an evaluation process using different scenarios. However, the superiority of DL techniques over other Shallow Machine Learning (SML) methods depends upon the specific context in which they are compared. For example, Mudannayake et al. [21], compare LSTM to Random Forest (RF) and Extreme Gradient Boosting (XGB) on monthly data from multiple cities; while Nguyen et al. [18] find RF and XGB to be competitive with or better than Artificial Neural Networks (ANN) on Vietnamese data for which there was less historical information. This contextual nature of the advantages/ disadvantages associated with DL suggests that this advantage will increase with longer datasets and/or stronger autocorrelative relationships among time series variables, but may decrease if either dataset length is short or tabular feature engineering is rich.

2.4.4 Critical Analysis

The Four key structural limitations in terms of evidentiary value of the model-comparison literature cited above can be summarised as follows.

Firstly, the benchmark scope of the literature is limited to only nine model families; mudannayake et al. [21], provide the largest single comparative exercise. Furthermore, mudannayake et al., [21] exclude other forms of machine learning including but not limited to regularized linear models and Gaussian Process Regression. Therefore, the question of whether increased machine learning complexity costs outweighs benefits for high feature density datasets such as those collected from india tier-2 cities remains unanswered (Gap F1). Secondly, there is a lack of geographical diversity within the literature concerning india tier two cities which are represented solely by an earlier ANN study of Singh & Satija [28] (gap F5). Lastly, there are inconsistencies in evaluation methodology/protocols used within each study cited herein. More specifically, while root mean squared error (RMSE) was found to dominate in most studies cited herein, no studies referenced herein employed or referenced either Theil's U or mean displacement accuracy (MDA) as part of their respective evaluation methodologies.

TABLE 2.4 Machine Learning and Deep Learning Approaches for Municipal Solid Waste Generation Forecasting

Author(s)	Year	Model(s)	FE Sophistication	Evaluation Metrics	Main Performance Finding	Key Limitation
Abbasi & El Hanandeh [15]	2016	ANN, SVM, ANFIS vs ARIMA	Low (< 10 variables)	RMSE, MAE, R^2	All AI models > ARIMA; ANFIS marginally best	3 ML families only; no DL; no SHAP
Ayeleru et al. [29]	2021	ANN vs SVM	Low (5 socioeconomic variables)	RMSE, R^2	ANN marginally superior to SVM	2 models only; no SHAP
Yang et al. [16]	2021	RF, XGBoost + SHAP	Medium (SE + temporal)	R^2, RMSE, MAE	XGBoost best; SHAP confirms GDP and population as top drivers	Limited model families; no DL; no COVID
Niu et al. [23]	2021	LSTM, ARIMA, ANN	Low (temporal; monthly)	RMSE, MAE	LSTM consistently outperforms ARIMA and ANN	No DL hybrids; no FE; no COVID
Nguyen et al. [18]	2021	6 ML models (RF, XGBoost, SVM, ANN, KNN, DT)	Low (SE; monthly Vietnam)	RMSE, MAE	XGBoost and RF most robust	No DL; no SHAP
Cubillos et al. [22]	2020	LSTM (multi-site)	Medium (multi-site encoding; weekly)	RMSE, MAE	LSTM accommodates multi-site prediction	No routing; no COVID
Mudannayake et al. [21]	2022	9 models: LSTM, GRU, CNN-LSTM, ANN, RF, SVR, KNN, XGBoost, ARIMA	Low (temporal + SE; monthly multi-city)	RMSE, MAE, R^2	LSTM/GRU best overall; 9-model benchmark established	No OLS/Ridge/GPR; no SHAP; no COVID; no routing

Lin et al. [30]	2022	Attention CNN-LSTM	Medium (daily SE + temporal)	RMSE, MAE	Attention CNN-LSTM outperforms LSTM	No multi-city validation; no COVID
Namoun et al. [20]	2022	Optuna-tuned gradient boosting	Medium (temporal + rolling statistics)	RMSE, MAE	Bayesian hyperparameter optimization exceeds manual tuning	No SHAP; no DL comparison
Xia et al. [26]	2022	Bayesian ANN ensemble	Low (SE + temporal)	RMSE + prediction intervals	Bayesian intervals better calibrated than point estimates	No SHAP; no DL comparison
Lu et al. [19]	2022	WGMod GBRT + weather features	Medium (meteorological + SE; monthly)	RMSE, MAE	Weather inputs measurably improve accuracy	No SHAP; no COVID
Ghanbari et al. [24]	2023	EEMD-MARS hybrid	Low-Medium (decomposed series)	RMSE, MAE	EEMD pre-processing improves MARS accuracy	Limited to MARS; no SHAP
Zhang et al. [17]	2022	ML ensemble + SSP scenarios	High (SSP pathways + SE + temporal)	RMSE, R^2	SSP-augmented ensemble best for long-horizon scenario forecasting	Scenario uncertainty not quantified
Jayaraman et al. [39]	2023	GSO-XGBoost	Medium (SE + temporal; Singapore)	RMSE, MAPE	Metaheuristic optimization outperforms standard tuning	Single city; no DL; no COVID

Wan et al. [27]	2024	Prophet + ML hybrid (COVID-aware)	Medium (SE + COVID indicators)	RMSE, MAPE	COVID-aware models significantly outperform pre-COVID trained models	No formal ITS; no SHAP on COVID
Jayaraman et al. [25]	2025	MSWDeepStack (LSTM + GRU + CNN-LSTM stacking)	High (multi-DL stacking)	RMSE, MAPE	Stacking ensemble outperforms all individual base models	Single city; no SHAP; no COVID

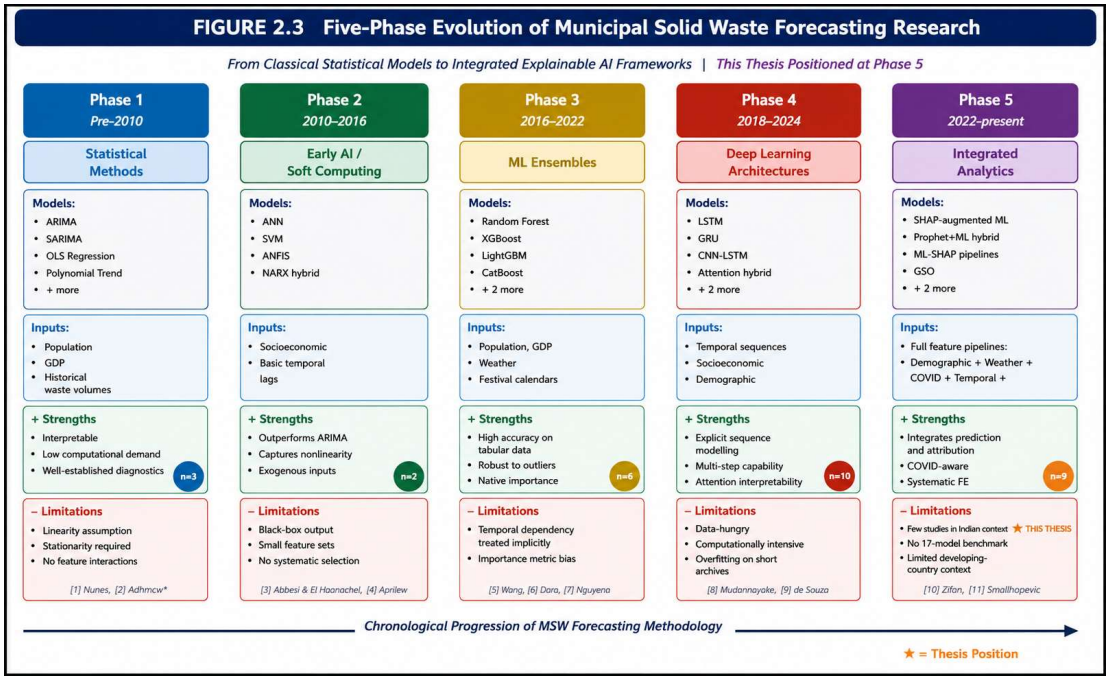


FIGURE 2.3 Five-phase evolution of municipal solid waste generation forecasting research,

Fig. 2.3 shows different models from traditional statistical models to integrated explainable AI frameworks, with representative studies and the methodological positioning of this thesis.

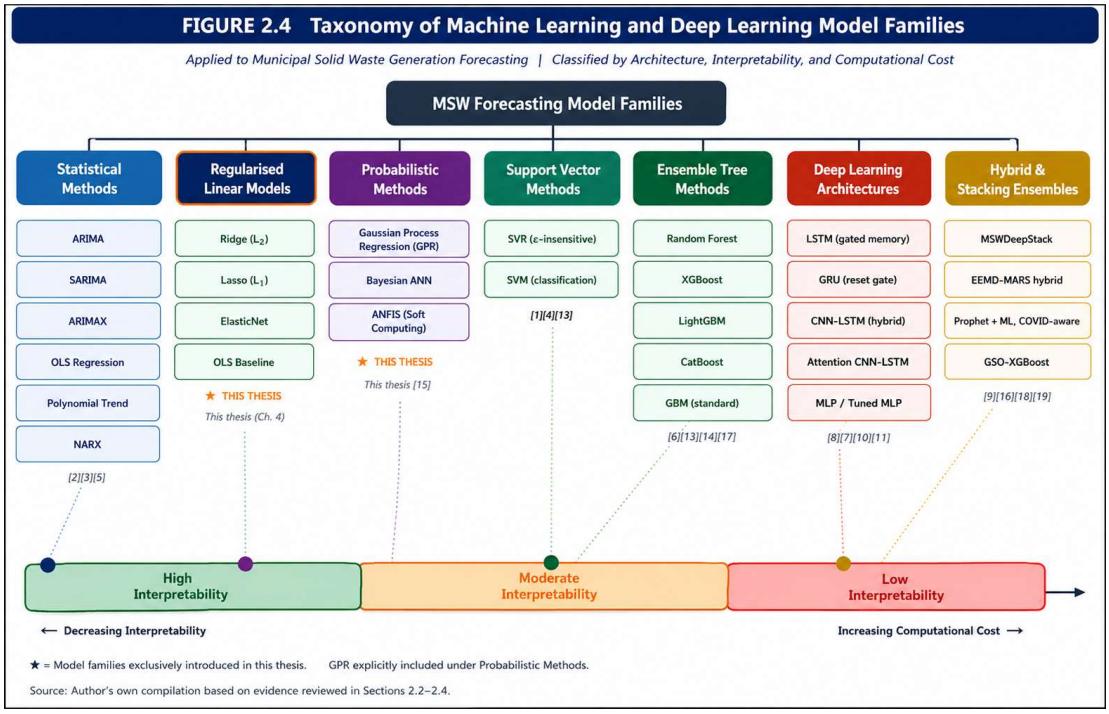


FIGURE 2.4 Taxonomy of Machine Learning and Deep Learning Model Families for MSW Forecasting

Fig. 2.4 shows hierarchical taxonomy of machine learning and deep learning model families applied to municipal solid waste generation forecasting, classified by architectural complexity, interpretability level, and computational cost, with representative studies from the evidence base. Source: Author's own compilation.

2.5 Evaluation Metrics and Validation Protocols in MSW Forecasting

The selection and application of evaluation metrics is not a peripheral concern in machine learning forecasting research but a central methodological decision that determines what performance dimension the evaluation system measures, what model behaviors it rewards or penalizes, and whether models trained and selected on one dataset will generalize to operational conditions. Misleading academic ranking through incorrect evaluation metrics may be one effect of evaluating an MSW generation forecasting system (which is used for determining when to deploy fleets, how much capacity will be needed at a given time, and what data to use to show that fleets are meeting their obligations as per regulatory requirements) but operational mis-calibration of the physical waste collection infrastructure would likely be another. This section provides definitions and critical examinations of the methods used by the largest number of researchers to evaluate the performance of MSW

forecasting systems - i.e., RMSE, MAE, MAPE, R^2, NSE, Pearson Correlation, Theil's U, and MDA; and also it evaluates the types of validation schemes used throughout the body of research.

2.5.1 Error and Accuracy Metrics

Mudannayake et al. [21] along with several other authors (Niu et al. [23]; Nguyen et al. [18]; Lu et al. [19]; Lin et al. [30]; Ghanbari et al. [24]; and Namoun et al. [20]) have reported Root Mean Squared Errors (RMSE) as the predominant error/accuracy metric used in the forecasting of Municipal Solid Waste (MSW). The RMSE measures the square root of the mean squared differences between estimated and actual values. The squaring process in the calculation for the RMSE has an effect on how the cost of larger prediction errors are evaluated. Larger errors are given greater penalty than proportionally small errors. In addition, this error measurement is scale-dependent which means that one cannot compare an RMSE value from a study that predicts monthly municipal MSW tonnage to an RMSE value obtained from another study that forecasts daily household MSW tonnage unless the results are normalized.

Mean Absolute Error (MAE), or Average Magnitude of Error, does not include the squaring step in the previous equation. Instead, it calculates the absolute difference between each estimate and corresponding actual value, then averages those values together. This produces an average error weight that gives equal importance to all error magnitudes. The MAE is therefore, more resistant to outliers than the RMSE. Both RMSE and MAE were used by Mudannayake et al. [21] in their comparative analysis of nine models. It was recognized by Mudannayake et al. [12] that while the two statistics provide similar information about the size and shape of the error distribution they represent two distinct characteristics of the data distribution.

Mean Absolute Percent Error (MAPE) represents each individual error as a percentage of the measured or actual value. As such, this statistical measure is independent of scale and can be used to evaluate the performance of various models using data collected at different levels of magnitude. For example, Yang et al. [16] and Jayaraman et al. [39] employed MAPE as their major statistical tool for evaluating model accuracy. However, due to the nature of percent-based calculations, the use of MAPE may produce statistically biased

estimates when actual values approach zero. Additionally, since MAPE is based upon absolute percent error, there is a systematic bias introduced into the selection of models when predictions consistently exceed measured values (overprediction) as opposed to consistently falling below measured values (underprediction).

Theil's U statistic provides a relative measure of model performance when comparing the RMSE of a developed model to a baseline Naive Model. If Theil's U is less than one, then the developed model performs better than the Naive Model. There appears to be very little application of Theil's U in relation to evaluating MSW forecasting models in the literature. Without including some form of reference point (naive model), published RMSE values do not provide a clear indication of whether the reported values reflect real-world predictive capability.

2.5.2 Goodness-of-Fit and Skill Metrics

The coefficient of determination (R^2) measures the proportion of total variance in the target variable explained by the model predictions. R^2 is widely reported - used by Abbasi and El Hanandeh [15], Zhang et al. [17], Ceylan et al. [31], and Zhao et al. [10] - because of its intuitive variance-explained interpretation and dimensionlessness. The limitations of R^2 are substantial for strongly trended time series - a characteristic of most urban MSW archives - since R^2 is systematically high regardless of model skill because the temporal mean that constitutes the denominator is a poor forecasting baseline. R^2 also does not detect systematic bias.

The Nash-Sutcliffe Efficiency was developed for hydrological model evaluation but is increasingly applicable to environmental time-series forecasting. NSE equals one minus the ratio of the model's residual variance to the total variance of the target series around its mean. When NSE equals 1.0, the model perfectly reproduces the observed series; when NSE equals 0.0, the model performs no better than the mean-of-series predictor. The Mean Directional Accuracy (MDA) quantifies how well model forecasts match the sign of change for each time step. In other words, the MDA measures the "direction" part of forecasting. For example, if you forecast an increase of 10 tons per month this year versus last year and waste generation actually increases by 20 tons per month then your prediction had directional

accuracy. If waste generation decreased by 5 tons per month during the same period as the increase of 10 tons you predicted, then the direction was wrong.

The Mean Directional Accuracy is different than the Mean Absolute Error (MAE). MAE calculates how large the error of each forecast is. MDA calculates how often the forecast is correct.

2.5.3 Validation Schemes and Protocol Design

Just like with the Evaluation Metric; the Choice of the Validation Scheme will have a large impact on your Model. In Literature three primary methods of validation are described: Hold-Out, K-Fold Cross-Validation and Rolling-Origin. The Rolling-Origin method preserves the Temporal Ordering of Data by sliding a Training Window forward through the Dataset to generate Out-of-Sample Predictions at each Step. Therefore this Method is an accurate representation of what occurs during Operation and the Best Choice for validating Time-Series Models of Municipal Solid Waste (MSW). Mudannayake et al. [21], used a Form of Rolling-Origin to evaluate nine different Models within their Benchmark. The adoption of rolling-origin evaluation alongside an eight-metric evaluation protocol - encompassing RMSE, MAE, MAPE, R^2, NSE, Pearson correlation, Theil's U, and MDA - in the Chapter 4 framework addresses the systematic single-metric and holdout-validation limitation documented in the reviewed literature.

2.5.4 Critical Analysis

The evaluation landscape in MSW forecasting exhibits a concerning degree of methodological fragmentation. The dominant reliance on RMSE as the sole or primary metric provides a necessary but insufficient basis for operational model selection. The absence of Theil's U and MDA from the reviewed MSW forecasting literature means that published performance claims cannot be interpreted as evidence of genuine forecasting skill above the random walk baseline, nor evaluated for operational relevance to fleet sizing decisions. The prevalence of holdout validation over rolling-origin evaluation means that published performance estimates are influenced by the arbitrary selection of the training-test

split boundary. The eight-metric rolling-origin evaluation design of Chapter 4 addresses all three identified deficiencies.

TABLE 2.5 Evaluation Metrics and Validation Protocols for MSW Forecasting

Metric	Conceptual Definition	Interpretive Focus	Strengths	Limitations	Usage in Reviewed MSW Literature
RMSE	√(mean squared prediction error)	Large-error sensitivity	Same units as target; penalizes outliers	Scale-dependent; not cross-study comparable	Most widely used: [15], [18], [19], [20], [21], [23], [24], [30]
MAE	Mean of absolute prediction errors	Average error magnitude	Robust to outliers; linear penalization	Scale-dependent; misses error distribution shape	Common: [18], [20], [21], [22], [23], [24], [30]
MAPE	Mean of percentage absolute errors	Relative accuracy	Scale-independent; enables cross-site comparison	Undefined near zero; asymmetric penalization	Used by: [16], [25], [27], [39]
R^2	Proportion of target variance explained by model	Goodness-of-fit	Dimensionless; widely understood	Misleadingly high for trended series; does not detect systematic bias	Common: [10], [15], [16], [17], [21], [31]
NSE (Nash-Sutcliffe)	1 − (residual variance ÷ total variance about series mean)	Skill vs mean-of-series predictor	Formally tests added value over mean baseline	Not standard in MSW literature	Not reported in any reviewed MSW study; adopted in Chapter 4

Pearson r	Linear correlation between predicted and observed	Association direction	Simple to compute	Does not detect bias	Occasionally reported alongside R^2
Theil's U	Model RMSE ÷ naive baseline RMSE	Skill vs random walk	Directly benchmarks model against persistence forecast	Rarely reported in MSW	Not reported in any reviewed MSW study; adopted in Chapter 4
MDA	Proportion of period-to-period direction changes correctly forecast	Operational direction accuracy	Directly relevant to fleet sizing decisions	Ignores magnitude	Not reported in any reviewed MSW study; adopted in Chapter 4

2.6 COVID-19 Impact on Municipal Solid Waste Systems

The COVID-19 pandemic constituted an unprecedented perturbation to municipal solid waste systems across the globe, operating simultaneously through at least three distinct channels: behavioral changes in household and commercial consumption patterns, policy-driven restrictions on economic activity and population movement, and disruptions to waste collection and operational systems. The resulting literature on COVID-19 and MSW has generated empirical evidence of direct relevance to forecasting methodology: the demonstration that structural breaks in MSW generation series degrade the accuracy of models trained on pre-pandemic data motivates both the explicit engineering of pandemic-period variables and the adoption of interrupted time series designs capable of testing and quantifying structural disruptions within a rigorous counterfactual framework.

2.6.1 Global COVID-19 Impact Studies on MSW Quantity and Composition

Empirical evidence on COVID-19's effect on MSW generation reveals a context-dependent pattern in which the direction of impact depended critically on whether commercial or

residential waste streams dominated the affected system. Studies examining aggregate municipal volumes documented net reductions under stringent restrictions: Cieśla et al. [40] reported a correlation of r = −0.61 between restriction intensity and monthly MSW generation volume across Poland, and Kong, Liu and Han [41] confirmed per-capita reductions across 252 Chinese urban communities through a difference-in-differences design, with stricter containment measures producing larger reductions. In contrast, Ishimura et al. [42] exploited prefectural variation in Japanese stay-at-home order timing as a quasi-experimental instrument and estimated a positive residential per-capita effect of approximately +0.8 to +4 per cent - opposite in direction to the Polish and Chinese aggregate findings - indicating that commercial waste contraction drove the aggregate reductions while residential generation rose; these findings appear in an SSRN preprint (doi: 10.2139/ssrn.4845017, under peer review as of June 2026). Compositional evidence from Granada, Spain [43] reinforced the stream-specificity interpretation: the food waste share of the municipal collection stream declined by 13.8 percentage points during the strictest lockdown phase, confirming that compositional shifts accompanied any volumetric change.

2.6.2 COVID-19 Effects on Collection and Operational Systems

The COVID-19 pandemic had impacts beyond just the volume and composition of waste generated at each location. The pandemic impacted many of the operations-based systems that collect, transport and process waste. During some periods of the pandemic there were worker shortages or absenteeism because workers were ill with COVID-19 or they were under quarantine. These shortages resulted in municipal collection system having lower than normal staffing levels of drivers, which further exacerbated existing issues with driver shortages that existed prior to the onset of the COVID-19 pandemic.

A recent study [44] examined the dynamics of waste generation in chile (a country heavily affected by the pandemic). In addition to examining how different types of waste generation changed during the pandemic, this study demonstrated how changes in the type of waste generated affected the amount of waste collected by municipalities. Specifically, when chile went into lockdown initially; all non-essential businesses were closed. As a result, it was observed that there was a decrease in commercial (non-household) waste generation. While this decrease provided municipalities with temporary relief from collecting excess amounts

of commercial solid waste, once government stimulus programs were implemented to help restart the economy; there was a marked increase in consumer spending resulting in increased household waste generation. Ultimately, total MSW generation surpassed previous (pre-COVID-19) baseline levels. This exceeded municipal collection systems' current capacities.

This study demonstrates that while inclusion of restrictions related to COVID-19 is important for developing accurate MSW forecasting models, inclusion of additional factors such as the implementation of stimulus packages and other post-pandemic economic recovery-related activities are equally important.

2.6.3 COVID-19 Effects on Forecasting Accuracy

In addition to the disruptions caused by the pandemic, forecasting models based on pre-pandemic training data were impacted by a structural break due to the restrictions implemented to combat the pandemic. As shown by Adamović [14] in Section 2.2.1 as an example of what happens when structural breaks occur in economics which will cause errors in forecasts generated from ANNs to be proportional to the number of structural breaks; Wan et al [27] showed empirically how the structural break of the pandemic affected the model: errors generated by models using pre-pandemic data increased significantly during the lockdown period; also identified is a need for a three phase timeline to accurately depict the impact of the pandemic on the waste generation time series and that simply using a binary COVID indicator variable would be insufficient to accurately depict this.

2.6.4 Interrupted Time Series and Difference-in-Differences Approaches

The interrupted time series design provides a quasi-experimental framework for evaluating the causal impact of a policy intervention or external disruption on a longitudinal outcome measure using observational data. In its canonical form, the ITS regression model includes terms for the pre-intervention level and trend of the outcome series, a binary indicator for the post-intervention period, and the post-intervention time trend, permitting estimation of both the immediate level change at the intervention point and the subsequent change in the trend slope. The ITS design is particularly well suited to the analysis of COVID-19 impacts

on MSW generation because the disruption had a clearly defined onset date and the monthly MSW generation series for most cities extends over multiple years before the pandemic, providing adequate pre-intervention observations for robust trend estimation.

Cieśla et al. [40] applied structural regression approaches approximating the ITS framework to Polish national MSW data, identifying significant changes in waste generation trajectory coinciding with restriction onset and phased relaxation. Kong, Liu and Han [41] employed the difference-in-differences design, which extends the basic before-after comparison by introducing a control group whose trajectory provides an empirical counterfactual. Ishimura et al. [42], in their Japanese analysis, similarly exploited spatial variation in stay-at-home order timing across prefectures as the DiD treatment assignment mechanism. The formal ITS and DiD designs reviewed above provide methodological foundations for COVID structural break analysis that descriptive pre-post comparisons and simple dummy variable approaches cannot match. These properties motivate the incorporation of ITS-derived COVID structural break parameters as engineered features within the multi-model ML forecasting pipeline of this thesis. The specific application of ITS methodology to Surat's MSW generation series, and the integration of the resulting break parameters into the 17-model evaluation framework, are presented and discussed in Chapter 4 of this thesis.

2.6.5 Critical Analysis

The literature on COVID-19 impacts on municipal solid waste systems has produced empirically valuable evidence of disruption effects across multiple countries and waste stream types. Three substantive limitations, however, constrain the value of the reviewed literature for the COVID-period analysis component of this thesis. The first and most significant is the absence of any study that combines formal ITS or DiD methodology with a multi-model machine learning forecasting framework - constituting Research Gap F2. The second limitation is the restricted time horizon of the reviewed COVID studies, all of which analyzed data through 2021 or 2022 at the latest. The third limitation is the geographical concentration of the reviewed COVID-MSW literature in European, Chinese, Japanese, and Latin American settings, with the COVID-19 pandemic's specific effects on MSW generation in Indian tier-two cities having received no published documentation. The broader lesson of the COVID period - that MSW generation volumes are structurally volatile

in response to external disruptions, declining sharply under mobility restrictions and surging under economic stimulus - provides a direct empirical foundation for the demand-adaptive vehicle routing literature reviewed in Section 2.7: if demand is not static but subject to structural discontinuities, fixed-schedule collection regimes will systematically misallocate fleet resources, and demand-driven inventory routing models become not merely an efficiency improvement but an operational necessity.

TABLE 2.6 COVID-19 Impact Studies on Municipal Solid Waste Generation

Author(s)	**Year**	**Country/City**	**Data Period**	**Method**	**Effect Direction & Magnitude**	**Key Finding**
Cieśla et al. [40]	2022	Poland (national)	2019–2021 (monthly)	Correlation + structural regression (ITS-type)	Negative: r = −0.61 (restriction index vs MSW volume)	Stricter restrictions reduce aggregate MSW; restriction stringency explains ~37% of COVID-period variance
Kong, Liu and Han [41]	2023	China (252 urban communities)	2019–2020 (daily)	Difference-in-differences (DiD)	Varied by restriction level; net reduction under high restriction	Per-capita daily waste generation decreases with lockdown intensity; DiD improves causal credibility
Muñoz-Soto et al. [44]	2023	Chile (national)	2019–2021 (monthly)	Descriptive + economic stimulus analysis	Negative during lockdown; positive surge post-stimulus	Economic stimulus reverses lockdown reduction; models require stimulus indicators
Peula et al. [43]	2023	Granada, Spain	2019–2020 (monthly)	Pre-pandemic vs lockdown compositional comparison	Food waste composition −13.8 percentage points during lockdown	Compositional changes documented; implications for stream-specific collection

Ishimura et al. [42]	2024	Japan (national; prefectural variation)	2020–2022	DiD using stay-at-home order timing variation	Positive: +0.8–4% per capita during stay-at-home orders	Residential waste increases offset commercial reductions; direction depends on waste economy structure. ⚠ SSRN preprint (doi: 10.2139/ssrn.4845017) - under peer review
Wan et al. [27]	2024	City TBC	2018–2023	Prophet + ML COVID-aware hybrid	Structural break detected; COVID-trained models substantially outperform	Three-phase temporal structure required; COVID-aware ML significantly more accurate
Adamović [14]	2017	EU-27 (historical analogue)	Historical economic shocks	ANN with structural break identification	N/A (historical shocks)	Economic structural breaks degrade ANN accuracy in proportion to break severity - foundational pre-COVID evidence

2.7 Route Optimization in Waste Collection

The Vehicle Routing Problem (VRP), first formalised in the context of commodities distribution, establishes the combinatorial optimization challenge of designing minimum-cost vehicle routes serving a geographically dispersed set of collection points from one or more depots. In municipal solid waste management VRP manifests primarily as Capacitated VRP (CVRP) where each vehicle has a payload limit and therefore returns to a transfer station or depot. Another important problem is Inventory Routing Problem (IRP) where collection timing is determined dynamically by signals of demand rather than fixed schedules. Sulemana et al. [45] shows that early studies of routing tended to use single objective CVRP formulations that minimized total distance travelled by vehicles. Comprehensive review published in International Journal of Industrial Engineering Computing in 2024 [46] updates Sulemana et al. [45] and covers literature from 2018 to

2024 and confirms a trend towards real time dynamic routing and formulations from multiple depots along with estimation that uses Internet of Things (IoT).

2.7.1 Classical VRP and CVRP Formulations

Formulating CVRP for MSW routing typically involves establishing parameters for the number of vehicles available; locations of depots or transfer stations; coordinates of collection points; quantities of refuse generated at each point; maximum capacities of individual vehicles; and either a matrix of distances travelled from one point to another or a matrix of travel times. Kim et al. [47] included time windows (VRPTW), along with geographic information system (GIS) spatial data and analyses of facilities' locations as part of their formulation of the classical CVRP to enhance operational efficiency in Korean municipal waste collection by improving route efficiencies of 15-25 percent relative to the manual schedules currently used. Thus, Kim et al.'s [47] combination of GIS data with optimization of the VRP was an initial example of how geometric considerations could be combined with combinatorial considerations that are now fully integrated using ML-GIS-DSS in section 2.10. As well as minimizing travel distances, there is a large amount of "green" VRP literature -- described in the 2024 review in IJIEC [46], which considers the effect of load weight on fuel use, road grade and traffic flow in cities when calculating emissions associated with routes -- so that these routes can reduce emissions substantially while increasing travel distances slightly.

2.7.2 Metaheuristic Solution Methods

The metaheuristic methods are being used as an alternative solution method to solve the NP-hard Vehicle Routing Problem (VRP). Because the number of possible solutions of the VRP is huge, it can be very difficult or even impossible to find the optimal solution using exact solution methods. Therefore, many researchers have been interested in developing different metaheuristic methods to provide good quality approximate solutions. Among the population based metaheuristics that have been used to solve the MSW routing problems, genetic algorithm (GA), particle swarm optimization (PSO), and ant colony optimization (ACO) are the most popular ones. These studies also show that all these three metaheuristics are able to achieve better results than the manually planned routes. According to the recent

study done by Sulemana et al. [45] , the average improvement in terms of distance reduction achieved by these three types of metaheuristics compared to the baseline manual planning approaches varies between 15% to 35%. However, as mentioned above, these studies did not evaluate how different metaheuristics perform when they are applied to same set of data under equal computational resources.

For example, a hybrid PSO-Tabu Search (TS) algorithm was proposed by Qiao et al. [48]. In addition to minimizing the total route length traveled, the authors also minimized fuel consumption and collection time together using their Pareto-optimal solution generation approach. This type of hybridization which uses both swarm intelligence algorithms such as PSO for global search and local search algorithms such as TS for intensifying the search around promising areas is becoming increasingly common because it provides competitive results for almost every instance of MSW routing problem. Moreover, since there may be conflicting goals for MSW collection operations such as reducing travel costs while reducing greenhouse gas emissions and improving public health, it is particularly suitable for solving multi-objective formulations.

According to the IJIEC Review article published in 2024 [46], GA and ACO are among the most commonly used metaheuristics to solve MSW routing problems. Both of them demonstrated similar level of improvements in terms of distance reductions in the range of 15% to 35% relative to manual scheduling baselines. GA is capable of achieving these level of improvements due to its ability to use crossover and mutation operators during iterations to move out of local optima. Similarly, ACO is capable of producing these level of improvements by simulating the behavior of ants using pheromone weighted probabilistic transition probabilities with an evaporation mechanism that prevents premature convergence. One major limitation observed by Sulemana et al. [45] is that although GA and ACO were found to produce very high level of improvements in terms of distance reductions, they have rarely been evaluated comparatively using the same set of data with matched computational budgets.

2.7.3 Inventory Routing and Dynamic Collection

Mes et al., [49] were a key part of the development of demand-driven, dynamic collection schedule planning, when they developed an Inventory Routing Problem for Dutch municipal

waste collection with the use of IoT sensor technology to monitor fill levels rather than a pre-set schedule for the collection process. The authors used a simulation/optimization methodology to demonstrate that there was a 15% to 20% reduction in cost compared to using a predetermined schedule for routing, while also laying down what could possibly be called the "demand pull" paradigm for MSW collection (the move away from supply push calendar based scheduling toward the use of demand responsive dispatching based upon the current condition of the containers), as this has been fundamental to developing subsequent IoT enabled technologies and frameworks that integrate multiple components. This work was published in Waste Management in 2014.

Markov, Bierlaire, Cordeau, Maknoon and Varone [11] extended the IRP framework to accommodate non-stationary stochastic demand for the Geneva MSW system, modelling seasonal and trend components in waste generation alongside stochastic variability and solving the resulting problem with Adaptive Large Neighbourhood Search (ALNS). Their results showed that they had developed an optimal routing solution based on the integration of a forecasted (future) demand as opposed to a static or deterministic demand that was better than other solutions which were generated by only optimizing their route and/or using no forecast of future demand. The Geneva IRP [11] is the most similar study in terms of methodology to a complete forecasting-to-routing pipeline. However, the demand model used in this study is a statistical model versus an ML-based model. Also, it uses GPS level spatial data from Switzerland's high precision mapping system -- both conditions are unlikely to be met for India's lower-tier municipal systems.

A proof-of-concept application of CVRP methodology to Surat's MSW system was presented by Rathod and Patel [50] - cited here with the mandatory qualification that this study constitutes a Stage 1 proof-of-concept implemented on synthetic CVRP infrastructure using the raw JSWTM dataset. Within that explicitly synthetic network configuration, the integrated LSTM-CVRP pipeline demonstrated a 50.31 per cent reduction in total collection distance relative to the uniform-frequency baseline schedule and a 24 per cent reduction in vehicle fleet utilisation requirements. These results are conditional on the synthetic route infrastructure and should not be interpreted as applying to Surat's actual operational network; they serve instead as proof of concept for the LSTM-to-CVRP integration pipeline that constitutes the Stage 1 foundation of the three-stage research program of which the present thesis constitutes Stages 2 and 3.

A structural limitation common to the demand-pull IRP paradigm established by Mes et al. [49] and extended by Markov et al. [11] is the conceptual treatment of transfer stations as dimensionless depot nodes - collection-cost sinks without internal state, throughput constraints, or operational management requirements. In Surat's MSW network, transfer stations in each of the nine collection zones serve as the critical interface between the household collection tier and the long-haul transport tier that delivers waste to the Khajod processing facility. Each transfer station has a finite daily throughput capacity, a vehicle marshalling and queuing function, a compliance reporting obligation under the Solid Waste Management Rules 2016 [3], and a zone-specific demand aggregation role that translates ward-level waste volumes into transfer-station loading schedules. None of these operational dimensions - TS throughput utilisation, zone-level demand aggregation, compliance status monitoring, or capacity buffer management - is represented in any published IRP or CVRP formulation reviewed in this section. The consequent disconnect between route optimization theory and transfer-station operational practice in Indian tier-2 municipal systems is fundamental rather than incidental: resolving it requires not merely an additional constraint in an existing VRP formulation but the development of a routing methodology in which TS throughput, zone demand, and vehicle zone assignment are jointly planned within a spatial framework that operates without GPS-level bin coordinates. This constitutes Research Gap R2, and directly motivates the transfer-station analytics layer of the IMPACT framework developed in Chapter 5.

2.7.4 Multi-Objective and Green Routing

The multi-objective extension of MSW routing formulates the collection planning problem as a Pareto-optimization challenge in which distance, fuel consumption, driver working hours, CO_2 emissions, service equity, and vehicle utilisation are simultaneously weighted rather than collapsed to a single monetary cost. Qiao et al.'s hybrid PSO-TS [48] operationalises this multi-objective framing with Pareto-optimal solution sets that enable decision-makers to navigate trade-offs between cost efficiency and environmental impact.

Green routing dimension of multi-objective msw vrp has been drawn increasingly into focus by methodological researchers as governmental regulations require municipal collection operators to reduce carbon emissions by specific amounts. Load-dependent fuel

consumption models reflect physically realistic relationships between vehicle payload and energy expenditure; a fully loaded waste collection vehicle will consume substantial more fuel per kilometre than an equivalent partially loaded vehicle. Routes that minimise total vehicle-kilometers do not necessarily result in routes with the lowest total fuel consumption or co_2 emissions. Under load-dependent fuel models, optimal routes tend to favour shorter, high-load collection segments followed by early transfer-station returns - a pattern that differs systematically from routes produced by pure distance minimisation and that requires explicit multi-objective optimization to identify. CO_2 abatement incentives further amplify this departure: routes optimised for emission reduction may tolerate modest distance increases in order to avoid congested urban corridors during high-emission operating windows, or to preference lower-gradient road segments where fuel economy improves. Tirkolaee et al. [51] incorporated sustainability objectives within their bi-level robust DSS framework, demonstrating that co-modal transport network design materially reduces total system emissions relative to single-mode route optimization - extending the green routing argument from the collection tier to the full waste logistics chain. A dimension largely absent from the green routing literature is service equity: the distributional question of whether optimised routes provide comparable collection frequency across high-density central zones and lower-density peripheral areas, where the operational economics of route optimization may systematically deprioritise informal-settlement communities. Despite this progress, the integration of operational KPIs beyond distance and emissions - including transfer-station throughput utilisation, compliance with service level agreements, and demand-responsiveness across zone sub-networks - remains absent from VRP and IRP formulations, constituting Gap R6. TABLE 2.7 summarises the route optimization literature; the critical unresolved limitation is the near-universal reliance on deterministic or sensor-based demand signals, with ML-forecast-driven route optimization remaining rare (Gap R1). Transfer-station-level operational analytics are absent from virtually all formulations (Gap R2), real-world adoption in developing-economy cities remains limited (Gap R3), GPS-level spatial data are assumed throughout (Gap R4), and Surat's actual operational network lacks any peer-reviewed routing baseline beyond the Stage 1 synthetic study (Gap R5).

TABLE 2.7 Route Optimization Studies in MSW (VRP, CVRP, IRP)

Authors & Year	Problem Type	Solution Method	Data Context	Objective(s)	Performance vs Baseline
Sulemana et al. (2018) [45]	Review: VRP/CVRP/IRP/MVRP	Survey: GA, PSO, ACO, exact methods	Multiple MSW contexts	Survey of formulations and outcomes	Survey - no single baseline
Kim et al. (2006) [47]	VRPTW + GIS + Facility Location	Exact solver + GIS network analysis	Korea municipal MSW	Distance + time + facility accessibility	15–25% distance reduction
Mes et al. (2014) [49]	Dynamic IRP (sensorised containers)	Simulation-optimization	Netherlands, IoT fill-level sensors	Cost minimisation (demand-pull)	15–20% cost reduction
Markov et al. (2020) [11]	Stochastic non-stationary IRP	ALNS metaheuristic	Geneva MSW (seasonal demand model)	Cost + service level	Significant reduction vs fixed-schedule
Qiao et al. (2020) [48]	Multi-objective CVRP	Hybrid PSO-Tabu Search	MSW (city unspecified)	Distance + fuel + time (Pareto-optimal)	Pareto-optimal solution set
Restrepo-Franco et al. [46]	VRP/CVRP/IRP (2018–2024 review)	Review: metaheuristics + IoT variants	SWM smart-city literature	Review - trends and gaps	Survey - no single baseline
Rathod & Patel (2025) [50]	CVRP (Stage 1 synthetic infrastructure)	LSTM forecast + OR-Tools CVRP	Surat (Stage 1 proof-of-concept; synthetic network; raw JSWTM data)	Distance reduction + fleet utilisation	50.31% distance ↓; 24% fleet ↓ (synthetic only)

Note: Stage 1 result for [41] is conditional on synthetic CVRP infrastructure and does not apply to Surat's actual operational network.

2.8 GIS and Spatial Applications in MSW Management

Geographic Information Systems have been applied to MSW management for three principal purposes: route network analysis, collection-point and transfer-facility siting, and spatial performance monitoring. A general trend in GIS has been for its applications to develop from simple GIS mapping, to the use of GIS as a tool for multi-criteria spatial decision making, to now applying GIS coupled with machine learning (ML) algorithms using real-time data from IoT sensors for forecasting demand. However, despite these advancements there is still significant room for the development of rigorously peer reviewed GIS analyses applied to tier-2 city Indian municipalities; most particularly to the complex spatially scaled municipal solid waste (MSW) system of Surat.

2.8.1 GIS-Based Route Optimization Studies

The Indian context for GIS route optimization in MSW is anchored by Sanjeevi and Shahabudeen [52], who employed ArcGIS network analysis for the Chennai, India municipal collection system, demonstrating systematic route improvements through GIS-enabled shortest-path algorithms and providing one of the earliest peer-reviewed Indian MSW GIS studies in the Scopus-indexed literature. Although focused on Chennai rather than Surat, this study [52] establishes the methodological validity of ArcGIS-based route network analysis for Indian urban MSW and documents the characteristic data challenges of Indian municipal mapping, including incomplete road attribute data, absence of standardised waste-point geocoding, and partial coverage of peri-urban collection zones. Kumar et al. [53] subsequently demonstrated GIS-based route optimization for Lucknow, reporting travel distance reductions of approximately 20 to 30 per cent over the existing manual schedule and further documenting the data quality challenges of Indian municipal GIS, including missing connectivity attributes for secondary and informal-area roads. The most recent Indian GIS-ML integration study is that of Navin Ganesh et al. [54], who combined an Artificial Neural Network demand prediction model with GIS routing for Erode, Tamil Nadu, demonstrating that ML-informed spatial decisions improve route design beyond static GIS-only approaches by dynamically adjusting route priorities in response to predicted demand variation. Together, Sanjeevi and Shahabudeen [52], Kumar et al. [53], and Navin Ganesh et al. [54] constitute the peer-reviewed GIS foundation for Indian MSW spatial

analysis; yet they remain confined to Chennai, Lucknow, and Erode respectively, with no peer-reviewed GIS routing baseline established for Surat (Gap G1). At the international level, Nurgaliyeva et al. [55] developed the most technically sophisticated GIS-integrated routing framework in the recent literature for Astana, Kazakhstan, embedding computer vision-based fill-level detection from IoT cameras, a real-time GIS platform, and Mixed-Integer Linear Programming route optimization - a Level 3 integrated framework combining forecasting inputs, GIS spatial intelligence, and formal optimization.

2.8.2 GIS-MCDM for Facility and Collection-Point Siting

Beyond route network analysis, GIS has been applied to the strategic siting of collection points and transfer facilities. Yalcinkaya and Uzer [56] employed a GIS and Multi-Criteria Decision Analysis framework combining population density, road network access, service radius constraints, and land-use classifications to optimise solid waste collection-point locations in Izmir, Turkey. Their weighted linear combination MCDM approach demonstrated that systematic multi-criteria spatial analysis selects collection configurations superior to those established by administrative convenience or historical precedent. Kim et al. [47], discussed in Section 2.7.1, similarly integrates facility location optimization with VRP and GIS constraints - establishing that routing, GIS, and facility location form a natural integrated combination. However, neither Yalcinkaya and Uzer [56] nor Kim et al. [47] extend their spatial analysis to transfer-station catchment zones or TS-level operational KPIs, constituting Gap G2.

2.8.3 Spatial Data Challenges in Developing Cities

One recurring pattern observed throughout many studies focused on Geographic Information Systems (GIS) research conducted within India [52], [53], [54] is that most spatial data currently available are structurally inadequate for GPS-based routing systems' needs. Both Sanjeevi & Shahabudeen [52] and Kumar et al. [53] noted problems with reliability of secondary road information; lack of standardized waste point geocoding; and the fact that informal settlement collection locations have little or no presence in commercially created map datasets. Navin Ganesh et al. [54] demonstrate ML-GIS coupling under this constrained data environment using approximate ward-centroid spatial representations rather than bin-

level GPS coordinates. The geocoding challenge in Indian informal settlements is particularly acute: collection points are typically identified in municipal operational records by local landmark names that are absent from standardised mapping databases and cannot be reliably resolved through conventional geocoding APIs. This documented spatial data constraint motivates the development of GPS-free routing methodologies designed to operate with available ward-boundary and TS-catchment spatial information (Gap G4). The coupling of ML demand forecasting with GIS-based spatial analysis also remains limited in the Indian context (Gap G3). TABLE 2.8 summarises the GIS literature reviewed in this section.

TABLE 2.8 GIS-Based Waste Management Studies

Authors & Year	Study Area	GIS Functions	Multi-Criteria Methods	GIS Data Quality
Sanjeevi & Shahabudeen (2016) [52]	Chennai, India	ArcGIS network routing; shortest-path analysis	None (routing only)	Moderate - Indian urban road network; limited bin-level geocoding
Kumar et al. (2022) [53]	Lucknow, India	GIS routing; travel distance analysis	None	Limited - Indian tier-2; incomplete secondary road attributes
Navin Ganesh et al. (2024) [54]	Erode, Tamil Nadu, India	ANN demand + GIS route prioritisation	None	Limited - ward-centroid spatial representation
Kim et al. (2006) [47]	Korea (multiple municipalities)	VRPTW + GIS network; facility location	Facility location MCDM	Fine - GPS-mapped network; precise bin coordinates
Yalcinkaya & Uzer (2019) [56]	Izmir, Turkey	Collection-point siting; spatial overlay	Weighted linear combination MCDM	Fine - GPS-mapped; complete land-use database
Nurgaliyeva et al. (2025) [55]	Astana, Kazakhstan	RGIS real-time monitoring; IoT-GIS routing	IoT-MCDM integrated	Fine - GPS + IoT fill-level sensors; precision mapping

The GIS literature reviewed in this section establishes spatial route intelligence as a necessary but insufficient foundation for advanced MSW management: spatial analysis identifies *where* to route vehicles and how to configure collection zones, but it cannot address *when* to deploy resources, *how much* capacity is required, or *what* operational and policy decisions the resulting spatial data should support - functions that require the decision support infrastructure reviewed in Section 2.9.

2.9 Decision Support Systems and Smart-City Waste Management

Decision support systems for municipal solid waste management have evolved from early strategic scenario models through IoT-enabled real-time platforms to digital twin architectures that integrate sensing, simulation, and stakeholder interfaces within a single unified framework. This progression reflects both technological advancement and a growing recognition that effective SWM requires decision support at multiple levels simultaneously - strategic planning, operational routing, compliance monitoring, and policy alignment - rather than the optimization of any single dimension in isolation.

2.9.1 Classical Decision Support Systems for SWM

The earliest formal DSS for SWM in the peer-reviewed literature is the ReFlows model, developed by Abeliotis, Karaiskou, Togia and Lasaridi [57] for strategic waste management planning at national and regional levels in Greece. ReFlows provides a flow-based scenario model enabling planners to evaluate technology configurations, capacity investments, and collection frequency adjustments against performance criteria spanning cost, environmental impact, and material recovery targets. The foundational contribution of Abeliotis et al. [57] lies in establishing the concept of an analytical, scenario-enabled DSS for waste management - a framework through which quantitative modelling informs strategic choices that would otherwise rest on administrative convention. However, ReFlows operates exclusively at the strategic planning level with no connection to operational routing, real-time data streams, or ML-based demand forecasting, exemplifying the fragmented forecast-to-decision pipeline characteristic of first-generation SWM DSS (Gap D1).

The most architecturally rigorous DSS formulation in the recent literature is that of Tirkolaee, Simic, Ghobakhloo, Foroughi, Asadi and Iranmanesh [51], who developed a robust bi-level DSS for integrated sustainable waste management incorporating co-modal transportation networks. The upper level optimises strategic decisions concerning facility location and multi-modal transport network design; the lower level simultaneously optimises operational vehicle routing and vehicle assignment. Robust optimization is applied at both levels to handle demand and cost uncertainty. Despite its methodological rigour, the JCP bi-level DSS [51] does not incorporate an ML-based demand forecasting layer, offers no stakeholder-specific interface for different decision-making roles, and addresses no developing-economy context, leaving the multi-layer DSS architecture gap (Gap D2) and the role-differentiated interface gap (Gap D3) unaddressed.

2.9.2 IoT-Enabled Smart Waste Systems

The integration of Internet of Things sensing into MSW collection has generated a substantial literature on sensor-driven operational optimization. The systematic review of Sosunova and Porras [58] surveyed over fifty IoT-enabled SWM papers, classifying systems along four dimensions: sensor type, communication protocol, optimization approach, and DSS integration level. Their analysis confirmed that the majority of deployments focus narrowly on the sensor-to-route link, without analytics layers, compliance monitoring, or policy alignment mechanisms. The SWM 4.0 taxonomy developed by Sosunova et al. [59] in the Smarter Waste Decision Support Framework extends this classification to a five-dimension assessment: data acquisition, intelligence, optimization, decision support, and sustainability alignment. The Smarter Waste DSF provides the Level 2–5 integration classification adopted in Section 2.10.

An important Indian-context implementation is the MMRE (Management, Monitoring, Reporting, and Evaluation) framework of Kumar, Singh and Kumar [60], which integrates IoT sensors, AI-based demand prediction, GIS routing, and a multi-modal DSS for Dhanbad, a tier-2 Indian city. This framework represents the most advanced Indian smart-waste implementation in the Scopus-indexed literature, demonstrating that full-chain Level 5 integration is technically achievable in an Indian urban context. However, the MMRE framework requires active IoT infrastructure throughout the collection network - an

investment structurally infeasible in areas where informal settlement geocoding cannot support sensor installation.

2.9.3 Digital Twin Decision Support Systems

The Tshwane Urban Digital Twin (UDT) developed by Cárdenas-León, Koeva, Nourian and Davey [12] represents the current state-of-the-art in integrated DSS for MSW, combining real-time IoT sensor feeds, GIS routing simulation, and a strategic decision support interface within a single urban digital twin deployed for the City of Tshwane, South Africa. This Level 4 framework enables dynamic vehicle dispatch adjustments, strategic fleet deployment planning, and service-level monitoring within a single operational platform. A second operational digital twin DSS, developed for a Swiss municipality [61], implements a pipeline from real IoT data feeds through digital simulation to KPI dashboards and operational recommendations, confirming that DT-DSS integration is feasible at smaller municipal scales.

2.9.4 Stakeholder Interfaces, KPIs, and SDG Alignment

A cross-cutting theme in the DSS literature is the design of stakeholder-specific decision interfaces. Kumar et al. [60] extend this to an explicitly multi-stakeholder Indian context, incorporating operational monitoring, reporting for regulatory compliance, and evaluation dashboards aligned with Indian governance requirements. Yet no reviewed framework explicitly provides role-differentiated dashboards spanning the full range of SWM stakeholders within a single system that addresses the Swachh Bharat Mission, Smart Cities Mission, and national SDG monitoring requirements that govern Indian municipal waste management (Gap D3, Gap D4). The connection between SWM system performance and SDG targets - particularly SDG 11 (sustainable cities and communities), SDG 12 (responsible consumption and production), and SDG 13 (climate action) - is referenced in the Smarter Waste DSF [59] and acknowledged in Dhanbad MMRE [60] but has not been operationalised through specific KPI mappings in any existing Indian framework. TABLE 2.9 summarises the DSS literature reviewed in this section.

TABLE 2.9 Decision Support Systems and Smart-City MSW Frameworks

Authors & Year	System Type	Components	Decision Level	Stakeholder Coverage
Abeliotis et al. (2009) [57]	Classical strategic DSS (ReFlows)	Flow model; scenario analysis; cost-environment trade-off	Strategic (national/regional)	SWM planners only
Tirkolaee et al. (2024) [51]	Bi-level robust optimization DSS	Strategic network design (upper); operational routing (lower); robust demand model	Strategic + operational	Strategic and operational managers
Sosunova & Porras (2022) [58]	Systematic IoT-SWM review	50+ IoT-enabled systems; sensor taxonomy; protocol classification	Operational	Operations-focused literature
Sosunova et al. (2023) [59]	SWM 4.0 DSF (conceptual framework)	Data acquisition; intelligence; optimization; DSS; sustainability alignment	Strategic (taxonomic)	All stakeholders (conceptual scope)
Cárdenas-León et al. (2024) [12]	Urban digital twin (Tshwane UDT)	IoT sensors; GIS routing; demand simulation; DSS interface	Operational + strategic	Operational + management
DT Swiss DSS (2023) [61]	Digital twin DSS	IoT data feeds; digital simulation; KPI dashboard; recommendations	Operational + strategic	Management
Kumar et al. (2025) [60]	IoT-AI-GIS-DSS (MMRE framework)	IoT; AI prediction; GIS routing; DSS dashboards; reporting	Operational + strategic + regulatory	Multi-stakeholder Indian context

2.10 Integrated Forecasting–Optimization–DSS Frameworks

The preceding sections have reviewed forecasting, route optimization, GIS, and DSS as thematically distinct bodies of literature. A smaller but critically important literature attempts integration across these components, recognising that the operational effectiveness of MSW management depends not on any single technical capability in isolation but on the coherent linkage of demand forecasting, spatial routing, data analytics, and decision support within a unified framework. This section employs a five-level integration classification - adapted from the Smarter Waste DSF of Sosunova et al. [59] - to map existing integrated frameworks and establish the degree to which full-chain integration has been achieved in the published literature.

The classification levels are as follows. Level 1 comprises single-component studies - pure forecasting, pure VRP, or pure GIS - and is well-represented throughout the literature reviewed in Sections 2.2 through 2.9. Level 2 integrates forecasting and optimization, linking a demand model to a routing decision. Level 3 adds GIS spatial intelligence to a Level 2 system. Level 4 adds a DSS or digital twin interface to a Level 3 system. Level 5 represents full smart-city integration, encompassing all four components alongside national policy alignment mechanisms and sustainability reporting. As the analysis below demonstrates, Level 2 integration is moderately represented, Level 3 and Level 4 integration is rare, and Level 5 integration in the Indian context has not yet been achieved through a GPS-free, ML-driven framework.

2.10.1 Level 2 Integration: Forecasting and Optimization

Three studies achieve the most direct Level 2 integrations: Mes et al. [49] (IoT-sensor-driven dynamic IRP, reviewed in Section 2.7.3), Markov et al. [11] (stochastic non-stationary demand model coupled with ALNS route optimization for Geneva, reviewed in Section 2.7.3), and Rathod and Patel [50] - cited with the mandatory qualification that this constitutes a Stage 1 proof-of-concept implemented on synthetic CVRP infrastructure using the raw JSWTM dataset - which explicitly couples LSTM demand forecasting with CVRP route optimization for the Surat service area. Across all three frameworks, demand forecasting relies on sensor data, statistical decomposition, or a single ML model family, without the

comprehensive model-comparison and feature-engineering pipelines that current ML practice makes possible (Gap IF1).

2.10.2 Level 3 Integration: Adding GIS

Level 3 integration - coupling F+O with GIS spatial analysis - is achieved by Kim et al. [47] (VRPTW with GIS constraints and facility location for Korean MSW, reviewed in Sections 2.7.1 and 2.8.2) and Nurgaliyeva et al. [55] (the Astana RGIS system with IoT computer-vision fill-level detection and MILP optimization, reviewed in Section 2.8.1). GIS spatial intelligence in both studies improves route realism and enables zone-level capacity monitoring; however, neither incorporates a DSS layer, and both require GPS-level mapping infrastructure unavailable in Indian informal settlement contexts (Gap IF2).

2.10.3 Level 4 Integration: Adding DSS

Level 4 integration - encompassing forecasting, optimization, GIS, and a DSS interface - is achieved by Cárdenas-León et al. [12] through the Tshwane UDT and by the Swiss DT DSS [61] (both described in Section 2.9.3); Tirkolaee et al. [51] occupies the Level 3–4 boundary through bi-level robust optimization DSS without real-time GIS or IoT components (described in Section 2.9.1). Across Level 4 frameworks, demand forecasting is simulation-based or probabilistic rather than ML-driven, none incorporates a comprehensive multi-model ML pipeline or formal COVID structural break analysis, and all depend on IoT sensing or GPS-level mapping infrastructure not universally available in Indian tier-2 contexts.

2.10.4 Level 5 Integration: Full Smart-City and Policy Alignment

Level 5 integration requires F+O+GIS+DSS coverage plus explicit national policy alignment and sustainability reporting connecting the system to SDG and national program targets. The Smarter Waste DSF of Sosunova et al. [59] provides the conceptual taxonomy for Level 5 systems (reviewed in Section 2.9.2), but does not itself constitute an implemented framework. Kumar et al. [60] achieve the closest existing operational Level 5 implementation in the Indian context through the Dhanbad MMRE framework (reviewed in

Section 2.9.2); however, its IoT infrastructure dependency limits applicability to sensor-equipped collection networks and it does not address the geocoding failure environment of Indian informal settlements.

Across all integration levels, the scarcity of advanced integration is notable. Fewer than ten studies in the international peer-reviewed literature achieve Level 2 or above; fewer than seven incorporate a DSS layer; and fewer than six can be positioned at Level 5 by component coverage. Of these, only Dhanbad MMRE [60] is deployed in an Indian context. No published framework achieves Level 5 integration without IoT infrastructure dependency for a GPS-free Indian tier-2 setting - defining the highest-priority integrated framework gap of this research program (Gap IF3). TABLE 2.10 classifies all integrated frameworks reviewed, and FIGURE 2.7 provides a component-coverage map illustrating the sparse population of the fully integrated quadrant.

TABLE 2.10 Integrated Forecasting–Optimization–DSS Framework Classification

Authors & Year	**Forecast Method**	**Optimization Model**	**GIS Component**	**DSS/DT Component**	**Integration Level**	**Main Strengths**	**Main Limitations**
Mes et al. (2014) [49]	IoT fill-level sensor	Dynamic IRP (simulation-optimization)	No	None	Level 2	Pioneering demand-pull IRP; 15–20% cost reduction vs. fixed schedule; real-time dynamic dispatch	IoT sensor-dependent; no ML forecasting; no GIS; no DSS layer
Markov et al. (2020) [11]	Statistical seasonal decomposition	Stochastic IRP (ALNS)	No	None	Level 2	Non-stationary stochastic demand model; high-quality ALNS	Statistical (not ML) demand model; GPS mapping assumed; no DSS

						optimization; Geneva operational calibration	
Rathod & Patel (2025) [50] Stage 1 synthetic	LSTM (Stage 1 proof-of-concept; synthetic data)	CVRP (OR-Tools)	No	None	Level 2	LSTM–CVRP Level 2 integration; 50.31% distance reduction (synthetic); India-relevant demand context	Synthetic CVRP infrastructure; raw JSWTM dataset only; no GIS; no DSS; proof-of-concept scope
Kim et al. (2006) [47]	Not specified	VRPTW (GIS-constrained)	Yes (GIS network)	None	Level 3	First VRPTW–GIS synthesis; 15–25% route efficiency improvement; facility location integration	No ML demand forecasting; GPS-level spatial data required; no DSS
Nurgaliyeva et al. (2025) [55]	IoT/CV fill-level	MILP	Yes (RGIS real-time)	None	Level 3	Real-time IoT–GIS–MILP integration; computer-vision fill detection; most advanced Level 3 reviewed	IoT sensor-dependent; GPS-required; no DSS layer; Kazakhstan-specific context

Tirkolaee et al. (2024) [51]	Probabilistic demand estimation	Bi-level robust VRP	Network model	Bi-level DSS (strategic + operational)	Level 3–4	Bi-level strategic–operational DSS linkage; robust demand modelling under uncertainty; sustainability objectives	No real-time GIS routing; no IoT; no ML pipeline; limited to conceptual/modelled deployment
Cárdenas-León et al. (2024) [12]	IoT demand simulation	GPS route dispatch	Yes (full GIS)	Urban digital twin + DSS interface	Level 4	Full Level 4 urban digital twin; operational deployment; IoT + GIS + DSS integrated; African context precedent	IoT-dependent; GPS-required; no ML benchmark pipeline; no Indian policy alignment
DT Swiss DSS (2023) [61]	IoT/digital twin simulation	Route dispatch	No (DT-only)	DT + KPI dashboards + recommendations	Level 4	Real operational deployment; KPI dashboard system; DT simulation for route recommendations	Citation TBC; IoT-dependent; GPS-required; no ML forecasting pipeline
Sosunova et al. (2023) [59]	IoT multi-channel (conceptual)	ML-integrated routing (conceptual)	Yes (framework)	DSF taxonomy (conceptual only)	Level 5 (conceptual)	Comprehensive Level 5 conceptual taxonomy; SDG alignment	Conceptual only - no implementation; no routing or GIS operational

						framework; SWM 4.0 classification schema	detail; no Indian context
Kumar et al. (2025) [60]	AI/ML demand prediction	GIS-based allocation	Yes (GIS)	IoT + multi-modal DSS dashboards	Level 5	Indian context; operational Level 5; regulatory alignment; multi-stakeholder dashboards	IoT-dependent; GPS-required; cannot scale to GPS-absent informal settlements
Proposed Surat MSW Framework (this study)	**17-model ML benchmark (OLS best; GPR 2nd); 419-feature pipeline; ITS COVID structural break analysis**	**IMPACT: GPS-free TWR-A/B, VAO, MSRS, WDB+RB, FIO; TS-level demand analytics**	**Yes - ward-boundary and TS-catchment spatial analysis (GPS-free)**	**4-layer DSS (data → ML → analytics → stakeholder UI); Streamlit PoC; role-differentiated dashboards**	**Level 5**	**GPS-free; IoT-free; Indian-policy-aligned (SBM, SCM, SDG 11/12/13); comprehensive multi-model ML pipeline; TS-level analytics**	**Surat-specific validation; Streamlit PoC (no production deployment); awaiting peer-reviewed publication**

Level 2 classification for [41] reflects Stage 1 architecture (synthetic infrastructure; no GIS; no DSS); Stages 2–3 of the thesis add GPS-free spatial analysis and a 4-layer DSS, achieving Level 5. Proposed Surat Framework row populated from evidence established in Sections 2.7–2.10; no new claims introduced.

With the full evidence base now surveyed across six thematic domains and ten integration frameworks, Section 2.11 draws these strands together through a systematic critical analysis, identifying the specific methodological and contextual limitations that the preceding sections have established, and showing how these limitations form a coherent and hierarchically

structured set of research gaps that motivate the three-pronged research framework of this thesis.

2.11 Critical Analysis of Existing Research

The literature reviewed in Sections 2.2 through 2.10 encompasses a diverse and rapidly expanding evidence base. Systematic critical analysis reveals, however, that persistent methodological and contextual gaps remain across all seven thematic areas, and that these gaps are not independent but form a coherent structure that motivates the three-pronged research program of this thesis. The following analysis applies a consistent framework - examining achievements, limitations, and unresolved research questions - for each theme, drawing exclusively on evidence established in the preceding sections rather than introducing new citations.

2.11.1 MSW Forecasting Methods

Although the research on MSW forecasting has made significant progress in comparison to the traditional methods of using ARIMA or regression, and although there were some initial attempts to use artificial intelligence (AI) in this context, the most advanced research is now comprised of both ensemble and deep learning models. In terms of comparative analysis among multiple model types, the work by Mudannayake et al. [21], , who studied nine different model families, can be seen as a key reference point for pre-2025 studies demonstrating that machine learning (ML) and deep learning (DL) models consistently perform better than statistical approaches used historically; further, the results demonstrate that a diverse set of models provides better performance, and no one type of model performs best for every urban setting. However, the scope of existing comparative studies remains narrow relative to what is technically feasible: the majority of studies reviewed in Section 2.2 test three to five model families in isolation, typically concentrating within either the statistical, shallow ML, or DL family. The persistence of single-model studies and the absence of any comprehensive multi-family benchmark for Indian tier-2 monthly MSW data represent the most significant unresolved gap. Furthermore, with the partial exception of Wan et al. [27], which employed Prophet-based structural break detection, no study has formally integrated Interrupted Time Series methodology with a comprehensive ML

pipeline, leaving the COVID-19 structural break inadequately characterised at the level of methodological rigour that policy-relevant forecasting requires.

2.11.2 Feature Engineering and XAI

Feature engineering for MSW forecasting has advanced from the near-exclusive reliance on population and GDP as predictive variables to more systematic multi-variable frameworks incorporating weather, temporal cycles, and event-based indicators. The most advanced XAI-integrated pipelines - Adeleke and Jen [36] with twelve variables and hierarchical SHAP, and Nourani et al. [33] with the MI-SHAP two-stage selection approach - represent genuine methodological progress in embedding interpretability as a core pipeline component. The persistent limitation is one of scale and cross-model scope: no existing study has systematically generated and pruned a feature candidate pool approaching 400 variables, and cross-model SHAP comparison across multiple ML architectures remains absent from the literature. Zhao et al. [10] have proven that larger feature pools are better for increasing the accuracy of forecasts for provincial China data; however, the difference between the 50-feature frontier (a comprehensive feature set) and a completely multi-category feature space (which includes all types of socioeconomic information, demographics, climate information, public policy information, COVID-19 related information, and time-based engineering information), is still unaddressed in all of the studies that we were able to review. In addition, since none of the current FE pipelines use COVID-19 related public policy variables as systematically engineered features (as opposed to using them as simply binary dummy variables), they are also limited by their inability to fully capture the multi-phase demand disruption models described in Section 2.6.

2.11.3 ML/DL Model Comparison and Benchmarking

The multi-model ML comparison literature has established that DL architectures (LSTM, GRU, attention hybrids) generally outperform shallow ML models on longer time series, while gradient-boosting ensembles dominate on shorter, feature-rich tabular data, as shown in TABLE 2.4. However, evaluation protocols are inconsistent across the reviewed studies, with different metrics, different train-test splits, and different data preprocessing choices making cross-paper comparison unreliable. The absence of rolling-origin cross-validation as

a standard protocol for MSW time series means that reported performance figures may not reflect the out-of-sample generalisability required for operational deployment. Critically, regularised linear models - Ridge, ElasticNet, Lasso - and Gaussian Process Regression are largely absent from MSW comparative benchmarks, despite their theoretical suitability for linearly-structured monthly aggregated time series. This family-scope gap prevents any existing study from determining whether ML complexity genuinely outperforms well-regularised linear models when a sufficiently rich feature set is available.

The absence of regularised linear models and Gaussian Process Regression from MSW benchmarks is a substantive evidential gap. When you encode temporal structure (through lag variables, rolling statistics and seasonality) using feature engineering methods to create an additional set of features that are used by your model to understand how each data point relates to prior ones - as done in the 419-feature pipeline created in chapter 4 - many of the advantages of sequence-based memory in LSTMs will have been transferred to these new sets of encoded features; therefore no conclusion can be drawn from the review literature regarding whether a richly engineered linear model may perform similarly to or better than LSTMs when considering the shorter term Indian Tier-2 monthly datasets.

The disparity between comparative studies across families of models compounds the scope-gap problem. No two comparative studies have used the exact same evaluation suite to compare comparative models; no comparison of methods has included Theil's U or Mean Directional Accuracy as secondary measures. Thus, comparisons based upon ranking of cross-studies' results from a given set of models (e.g., LSTMs) are not comparable -- i.e., the same method may be ranked #1 in a study using one dataset, but #3 in a study using another dataset -- thus making it difficult for researchers to synthesize evidence that can be relied upon by practitioners. Only Mudannayake et al. [21], among the comparative evaluations reviewed here, employed rolling-origin evaluation.

2.11.4 COVID-19 Impact and Policy-Sensitive Forecasting

The COVID-19 literature reviewed in Section 2.6 has successfully quantified the magnitude and trajectory of pandemic-induced MSW disruptions across Poland [40], China [41], Chile [44], Spain [43], and Japan [42], with DiD and ITS methodologies providing rigorous effect estimation designs. However, the integration of COVID impact analysis with comprehensive

ML benchmarking - and especially the embedding of COVID structural breaks within the feature engineering pipeline rather than treating them as post-hoc corrections - remains underdeveloped. No existing study formally tests ITS structural break significance across a multi-model ML suite, and none explicitly links COVID-adjusted forecasts to downstream operational routing decisions. This integration gap defines a critical methodological frontier: the ability to propagate COVID-driven demand uncertainty through the full forecasting-to-routing-to-DSS pipeline has not been demonstrated in the reviewed literature.

The operational consequences of structural break inadequacy extend directly to fleet deployment. A model calibrated on pre-COVID data will systematically over-predict waste generation during restriction phases - deploying surplus vehicles to partially empty routes - and under-predict during post-stimulus recovery phases, when household consumption surpasses pre-pandemic baselines and uncollected accumulations become a public health and compliance risk. Propagating COVID-induced demand uncertainty through the forecasting-to-routing link is therefore an operational necessity for any deployment-grade framework, not a methodological refinement.

The methodological relevance of ITS extends beyond the pandemic period: the Solid Waste Management Rules 2016 [3], the Swachh Bharat Mission, and state-level collection mandates each represent potential structural break points in MSW generation series. The reviewed literature demonstrates ITS and DiD designs only in the COVID context; an MSW forecasting framework that applies ITS to the pandemic thereby establishes the infrastructure for a broader class of policy-sensitive structural break testing applicable to future regulatory disruptions. The post-COVID trajectory of Indian tier-2 MSW generation - whether it has converged to the pre-pandemic trend or shifted to an altered level and growth rate - remains uncharacterised in the peer-reviewed literature, leaving medium-term capacity planning and regulatory compliance projections for cities such as Surat on an evidential foundation that the present thesis addresses.

2.11.5 Route Optimization and Transfer Station Integration

The route optimization literature has demonstrated substantial efficiency gains through CVRP, IRP, and multi-objective metaheuristic formulations, with Mes et al. [49], Markov et al. [11], and Qiao et al. [48] each demonstrating meaningful improvements over manual

baselines. Yet the translation of these gains to operational deployment in developing-economy cities has been constrained by GPS-level spatial data assumptions that pervade all reviewed frameworks. The demand-driven IRP paradigm established by Mes et al. [49] and extended by Markov et al. [11] is theoretically well-suited to ML-forecast integration, yet no published study has coupled a broad-spectrum ML benchmarking suite with a formal IRP or multi-objective CVRP formulation. Furthermore, the operational reality of transfer-station-level management - including TS throughput, compliance monitoring, vehicle zone assignment, and demand-responsive capacity scaling - is entirely absent from the VRP and IRP literature, constituting a substantive disconnect between the routes that optimization research produces and the operational decisions that MSW managers in Indian tier-2 cities actually face.

2.11.6 GIS and Spatial Realism

GIS applications in Indian MSW - Sanjeevi and Shahabudeen [52], Kumar et al. [53], and Navin Ganesh et al. [54] - provide a limited but growing peer-reviewed precedent for spatial route analysis in the Indian context. The structural geocoding challenge of Indian informal settlements, documented across the Indian GIS studies and illustrated by Nurgaliyeva et al.'s [55] contrasting IoT-precision context, has not been formally quantified in the peer-reviewed literature, and no published routing methodology has been explicitly designed for GPS-absent operation in the Indian informal settlement environment. The field has, in effect, acknowledged the existence of this constraint without producing a methodological response to it. Transfer-station catchment zone analysis and zone-level spatial performance KPIs are absent from all reviewed GIS studies, and the coupling of ML demand forecasting with GIS spatial decisions for Indian tier-2 cities extends only as far as the single-model ANN-GIS study of Navin Ganesh et al. [54].

2.11.7 DSS, Smart-City Systems, and Integrated Frameworks

The DSS and integrated framework literature has achieved demonstrable Level 4 integration in African and European contexts and has begun to extend into Indian contexts through Dhanbad MMRE [60]. The Smarter Waste DSF of Sosunova et al. [59] provides a conceptual taxonomy adequate for classifying any proposed framework. Yet the overarching limitation

of the integrated framework literature is its dependence on IoT infrastructure, GPS-level spatial data, and IT investment capacities that are not universally present in Indian tier-2 municipal authorities. No existing Level 4 or Level 5 framework has been designed to operate without IoT sensors, without GPS-mapped collection points, and with explicit alignment to Indian national policy frameworks and SDG monitoring requirements. This combination - GPS-free, ML-driven, Indian-policy-aligned, Level 5 - defines the integrated framework gap (Gap IF3) that is the ultimate motivation for the full-chain framework developed in Chapters 5 and 6.

2.12 Research Gap Identification

The critical analysis in Section 2.11 reveals five major research gaps that collectively motivate the three research objectives of this thesis. These consolidated gaps are specified below and mapped to the research objectives in TABLE 2.11. Each major gap represents a broader unresolved research theme supported by multiple literature observations drawn from the evidence base reviewed in Sections 2.2 through 2.10.

Gap 1: Insufficient Multi-Model ML Benchmarking and COVID Structural Break Methodology for Indian Tier-2 MSW (consolidates F1, F2, F4, F5). No published study benchmarks the full spectrum of ML model families - from regularised linear regression and Gaussian Process Regression through gradient-boosting variants to deep learning architectures - on a single Indian tier-2 monthly MSW dataset under a uniform multi-metric evaluation protocol. The most comprehensive prior benchmark, Mudannayake et al. [21], covers nine model families but excludes regularised linear models and GPR, and no equivalent multi-family comparison exists for an Indian tier-2 context. Indian tier-2 seasonal and structural patterns - monsoon cycles, festival demand surges, and policy-driven waste stream changes - are inadequately represented in the MSW forecasting feature landscape. Furthermore, formal Interrupted Time Series methodology for COVID-19 structural break analysis has not been integrated with any comprehensive ML forecasting pipeline; pandemic-period demand disruption is characterised only through binary dummy variables or descriptive pre-post comparisons rather than rigorous counterfactual estimation.

Gap 2: The small number of features and limited ability to use SHAP for comparison (aggregating FE1, FE2, FE3 & FE4) The size of the potential candidate set used in nearly

all reviewed research on MSW forecasting studies is always narrow. For example Zhao et al.[10] reported that even at its largest point the candidate set was only fifty. This represents a very much smaller fraction of the total possible feature space than could be achieved using large-scale multi-category combinations of socioeconomic/demographic/environmental/policy/COVID and time series. Variables related to COVID-19 and policies were always represented as binary indicators of disruptions; these binary indicators did not represent the multiple phases of restrictions and their durations, nor the various economic recoveries. Temporal lag/rolling window features were usually defined by some type of domain-convention-based definition rather than having been generated in an automated fashion and selected based upon the data. Finally, there has been no application of cross-model SHAP attribution among models being compared in a benchmark study. Therefore it cannot be determined if the structure of feature importance reflects actual patterns in the data, or model specific artifacts.

Gap 3: Disconnection Between ML Forecasting and Demand-Adaptive Routing with Transfer-Station Analytics (consolidates F3, R1, R2, R5, R6). There is no peer-reviewed study that combines a full-scale ML benchmarking framework with an inventory routing problem formulation or a multi-objective CVRP formulation in order to generate ML forecast-driven vehicle schedules. Current demand responsive routing models (Mes et al. [49]; Markov et al. [11]) use IoT sensor data or statistical seasonal decomposition of data rather than multi-model ML forecasting for future traffic demands. None of the studied VRP and IRP formulations include transfer station level operational analytics such as; throughput monitoring, compliance tracking, capacity utilization and zone-based demand aggregation -- which represents a substantial gap between theoretical routing models and operational realities at Indian tier 2 municipalities. The Rathod and Patel [50] Stage 1 study provides the only prior computational work on Surat but employs synthetic infrastructure that does not reflect actual operations. Multi-KPI routing objectives extending beyond distance minimisation to encompass compliance, demand responsiveness, and capacity utilisation have not been implemented for any Indian tier-2 context.

Gap 4: GPS-Dependence of All Routing Frameworks and Inadequate GIS Infrastructure for Indian Informal Settlements (consolidates R3, R4, G1–G4). All published route optimization and GIS-based MSW routing frameworks assume GPS-level spatial data - bin coordinates, precise road network attributes, and mapped collection-point locations - that are unavailable

in Indian informal settlement environments where landmark-based addressing and incomplete geocoding are structurally entrenched. No peer-reviewed routing methodology has been designed explicitly for GPS-absent operation under these conditions. The peer-reviewed GIS routing baseline for Indian tier-2 cities remains confined to Chennai (Sanjeevi and Shahabudeen [52]), Lucknow (Kumar et al. [53]), and Erode (Navin Ganesh et al. [54]); no published GIS routing study addresses Surat's collection network. Transfer-station catchment zone analysis, zone-level spatial performance KPIs, and comprehensive ML-GIS coupling in Indian MSW research are absent from all reviewed studies, with only the single-model ANN-GIS study of Navin Ganesh et al. [54] providing an Indian precedent.

Gap 5: Absence of an Integrated, GPS-Free, Policy-Aligned Decision Support Framework for Indian Tier-2 Cities (consolidates D1–D4, IF1–IF3). No existing published framework combines ML-based demand forecasting, GPS-free route optimization, and a multi-stakeholder decision support system within a single operational platform designed for Indian tier-2 municipal conditions. Classical SWM DSS (ReFlows [57]; Tirkolaee et al. [51]) operate without ML forecasting or real-time operational coupling. Level 4 and Level 5 frameworks achieve high integration but depend on IoT infrastructure and GPS-mapped collection networks unavailable in settings with informal settlement geocoding constraints. No existing framework provides role-differentiated stakeholder dashboards spanning field operations, municipal management, and policy audiences, nor aligns operational KPIs with Swachh Bharat Mission, Smart Cities Mission, and SDG 11/12/13 monitoring requirements within a single Indian tier-2 system. Fewer than ten studies in the international literature achieve Level 2 integration or above, and none achieves Level 5 for a GPS-free Indian tier-2 context.

These five major research gaps form a coherent and hierarchically structured motivation for the three research objectives of this thesis. As summarised in TABLE 2.11, Gaps 1 and 2 are addressed through RO-1, the seventeen-model ML forecasting framework with systematic feature engineering, an eight-metric evaluation protocol, and formal ITS COVID structural break analysis. Gap 3 is addressed jointly by RO-1 and RO-2, through the quantified demand scalar that links the ML forecasting pipeline to the IMPACT route optimization framework. Gap 4 is addressed through RO-2, the GPS-free multi-stage route optimization program applied to real operational data. Gap 5 is addressed through RO-2 and RO-3, the four-layer DSS prototype with role-differentiated stakeholder dashboards and explicit Indian policy

alignment. FIGURE 2.8 presents the hierarchical structure of these five gaps mapped to the three research objectives.

TABLE 2.11 Research Gap–Research Objective Mapping Matrix

Gap ID	Thematic Focus and Consolidated Sub-gaps	RO-1	RO-2	RO-3	Thesis Chapter
Gap 1	**Insufficient multi-model ML benchmarking and COVID structural break methodology for Indian tier-2 MSW** *Consolidates: F1, F2, F4, F5. No study benchmarks the full model-family spectrum (regularised linear through deep learning) on a single Indian tier-2 monthly dataset under a uniform multi-metric protocol. Formal ITS COVID structural break analysis is absent from all ML pipelines. Indian seasonal and structural patterns (monsoon, festivals, policy events) are inadequately represented.*	✓			Ch. 4
Gap 2	**Narrow feature engineering and limited cross-model SHAP attribution** *Consolidates: FE1, FE2, FE3, FE4. Feature pools across all reviewed studies reach ≤50 variables. COVID-policy variables are encoded as binary dummies. Lag/rolling features are specified by convention rather than systematic selection. Cross-model SHAP attribution across a full model suite remains absent from the reviewed literature.*	✓			Ch. 3, Ch. 4
Gap 3	**Disconnection between ML forecasting and demand-adaptive routing with transfer-station analytics** *Consolidates: F3, R1, R2, R5, R6. No study couples a comprehensive ML benchmark suite with formal IRP or CVRP*	✓	✓		Ch. 4, Ch. 5

	formulations. Transfer-station analytics (throughput, compliance, capacity) are absent from all routing frameworks. No peer-reviewed routing baseline exists for Surat's actual operational network. Multi-KPI objectives beyond distance minimisation are unimplemented for Indian tier-2 contexts.				
Gap 4	**GPS-dependence of all routing frameworks and inadequate GIS infrastructure for Indian informal settlements** *Consolidates: R3, R4, G1–G4. All published routing and GIS frameworks assume GPS-level spatial data unavailable in Indian informal settlement environments. No GPS-free routing methodology has been designed for landmark-addressed collection points. The peer-reviewed GIS routing baseline for Indian tier-2 cities is confined to Chennai, Lucknow, and Erode; Surat is unrepresented. TS-catchment zone analysis is absent from all reviewed studies.*		✓		Ch. 5
Gap 5	**Absence of an integrated, GPS-free, policy-aligned decision support framework for Indian tier-2 cities** *Consolidates: D1–D4, IF1–IF3. No framework combines ML forecasting, GPS-free routing, and multi-stakeholder DSS for an Indian tier-2 city. All Level 4–5 frameworks (Tshwane UDT, Swiss DT, Dhanbad MMRE) require IoT/GPS infrastructure. Role-differentiated dashboards aligned with Swachh Bharat Mission, Smart Cities Mission, and SDG 11/12/13 are absent from all published Indian DSS. Fewer than ten studies achieve Level 2 integration or above; none achieves GPS-free Level 5.*	✓	✓	✓	Ch. 5, Ch. 6

RO-1 = Forecasting and Feature Engineering; RO-2 = Route Optimization and Spatial; RO-3 = DSS and Integrated Framework. ✓ = gap addressed by this RO.

2.13 Chapter Summary

Chapter 2 has reviewed the global literature on municipal solid waste forecasting, feature engineering and explainable AI, machine learning and deep learning model comparison, COVID-19 impact analysis, route optimization, GIS and spatial applications, decision support systems, and integrated MSW frameworks, drawing on peer-reviewed publications spanning 2009 to 2026. The review has been organised thematically rather than chronologically to establish five major research gaps that directly motivate the three research objectives of this thesis.

Sections 2.2 through 2.5 establish that the forecasting literature lacks a comprehensive multi-family ML benchmark for Indian tier-2 monthly MSW data under a uniform multi-metric protocol, and that formal ITS structural break methodology for COVID-19 disruption has not been integrated with any ML pipeline (Gap 1). Section 2.3 confirms that systematic multi-category feature engineering at scale - combined with cross-model SHAP attribution across all model families - remains absent from the reviewed literature (Gap 2).

Section 2.7 documents sustained VRP, IRP, and multi-objective metaheuristic efficiency gains, but confirms that no study couples a comprehensive ML suite with formal routing formulations, and that transfer-station operational analytics are entirely absent from all reviewed frameworks (Gap 3). Section 2.8 confirms that the Indian GIS MSW peer-reviewed baseline remains limited to Chennai, Lucknow, and Erode, and that the spatial data constraints of informal settlements - incomplete road attributes, absent bin-level geocoding, and landmark-based addressing - structurally differentiate Indian tier-2 contexts from the GPS-mapped environments in which the broader routing literature has been conducted (Gap 4).

The DSS and integrated framework reviews in Sections 2.9 and 2.10 established the current state-of-the-art at Level 4 integration, represented by providing the only Level 5 Indian-context implementation. The systematic IoT infrastructure dependency of all reviewed Level 4 and Level 5 frameworks defines the primary integration gap: no existing published

framework achieves comprehensive MSW integration without sensor dependency for an Indian tier-2 city.

The five major research gaps identified and mapped in TABLE 2.11 form a hierarchically structured motivation for the three research objectives of this thesis. RO-1 addresses Gaps 1 and 2 through a seventeen-model ML benchmarking suite, a systematic 419-feature engineering pipeline, an eight-metric evaluation protocol, and formal ITS COVID structural break analysis. RO-2 addresses Gaps 3, 4, and the routing dimensions of Gap 5 through a three-stage route optimization program - from the Stage 1 LSTM-CVRP proof-of-concept of Rathod and Patel [50] through TS efficiency analysis to the GPS-free IMPACT framework applied to real operational data. RO-3 addresses Gap 5 through a four-layer DSS prototype incorporating role-differentiated stakeholder dashboards, a demand scalar, and explicit alignment with Indian national policy programs and SDG targets. Chapter 3 presents the research methodology through which these research objectives are operationalised.

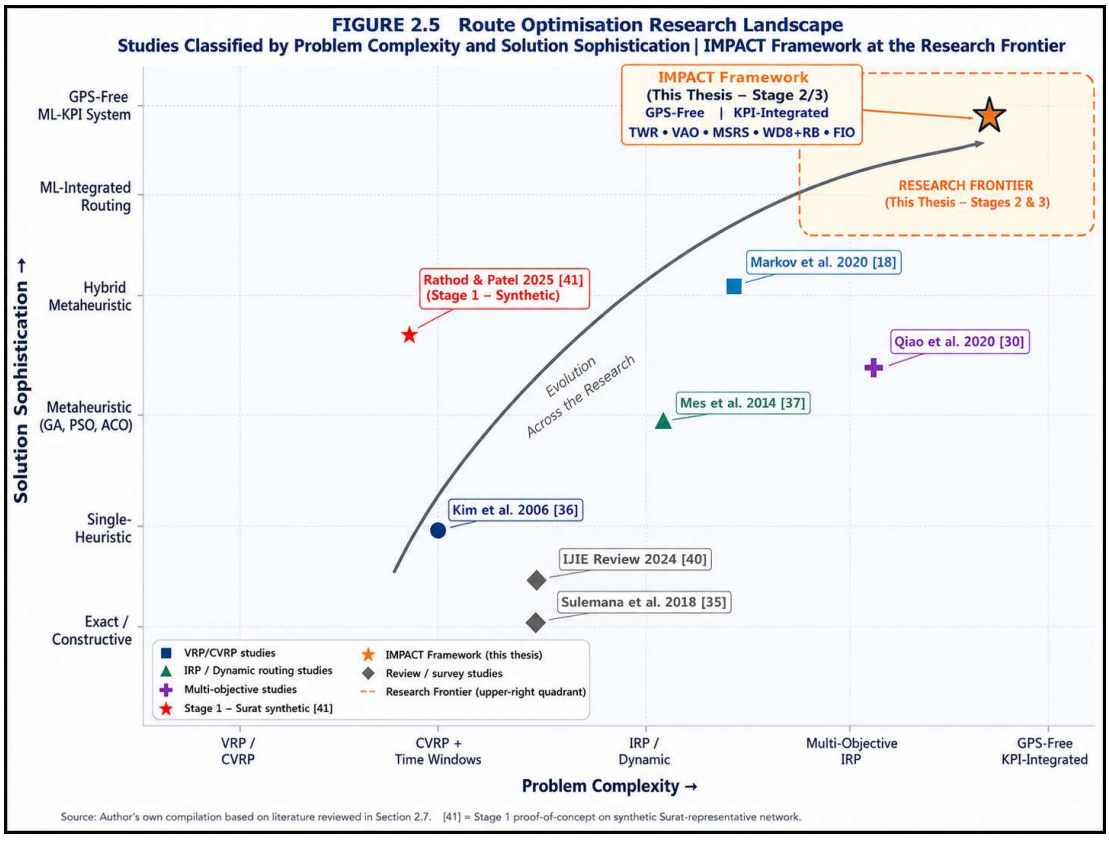


FIGURE 2.5 - Route Optimization Research Landscape

Fig. 2.5 shows studies classified by problem complexity and solution sophistication, with This Thesis's IMPACT Framework Positioned at the Research Frontier.

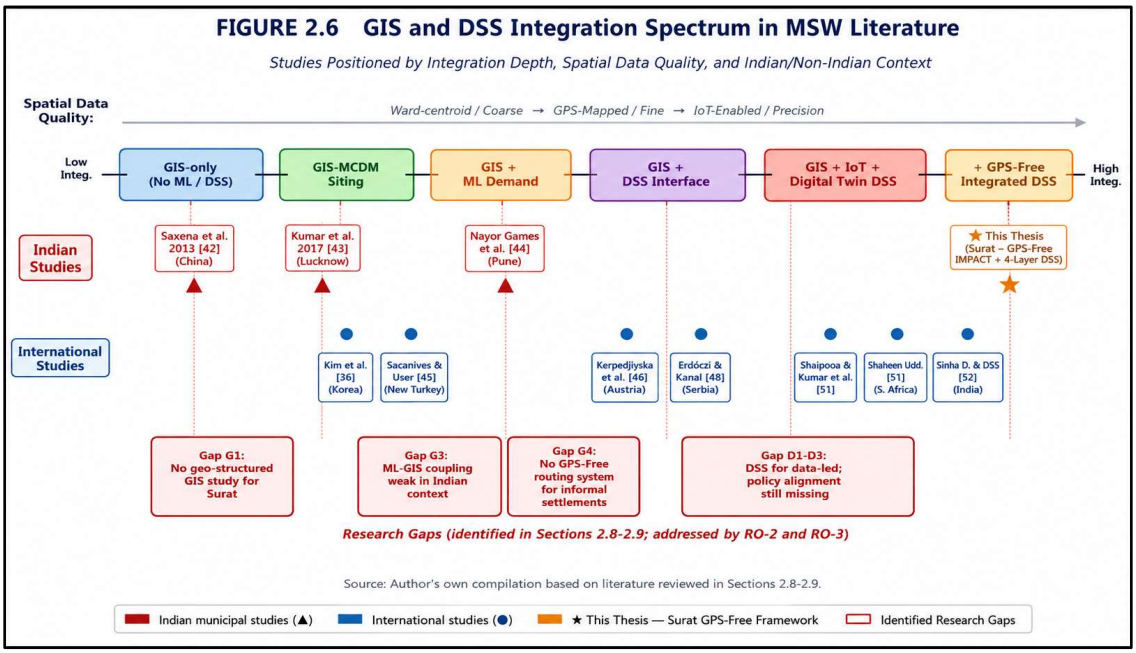


FIGURE 2.6 - GIS and DSS Integration Spectrum in MSW Literature

Fig. 2.6 shows studies positioned by integration depth, spatial data quality, and Indian/Non-Indian Context.

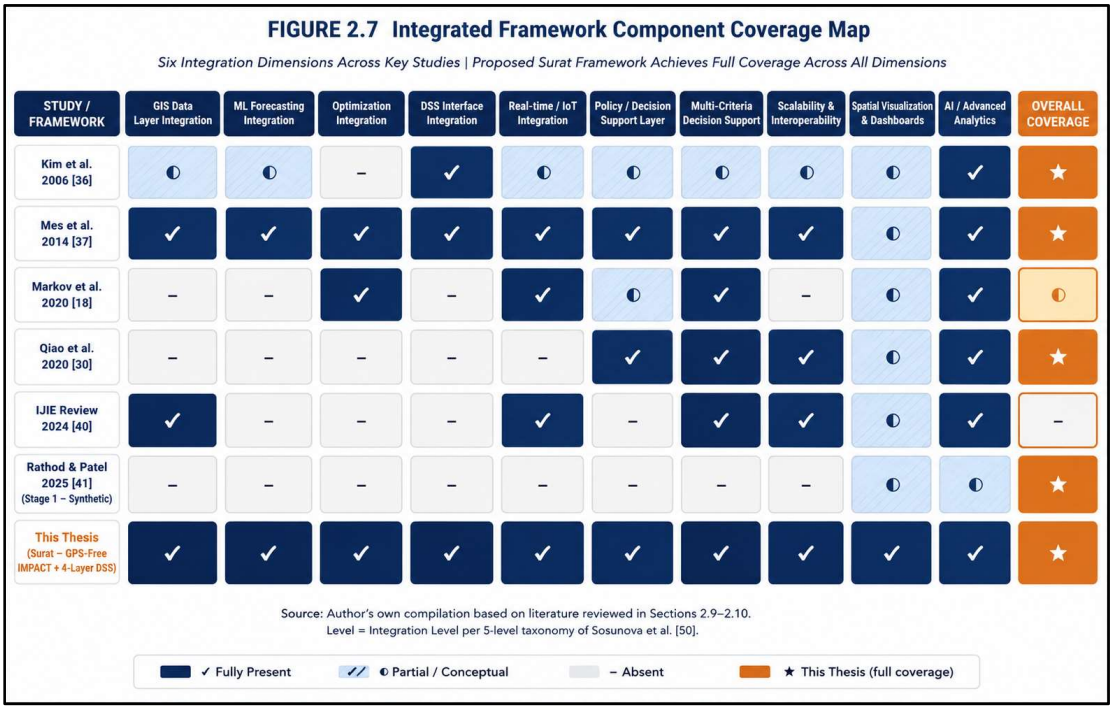


FIGURE 2.7 - Integrated Framework Component Coverage Map

Fig. 2.7 shows six integration dimensions across key studies, with the proposed Surat framework achieving full coverage across all six rows.

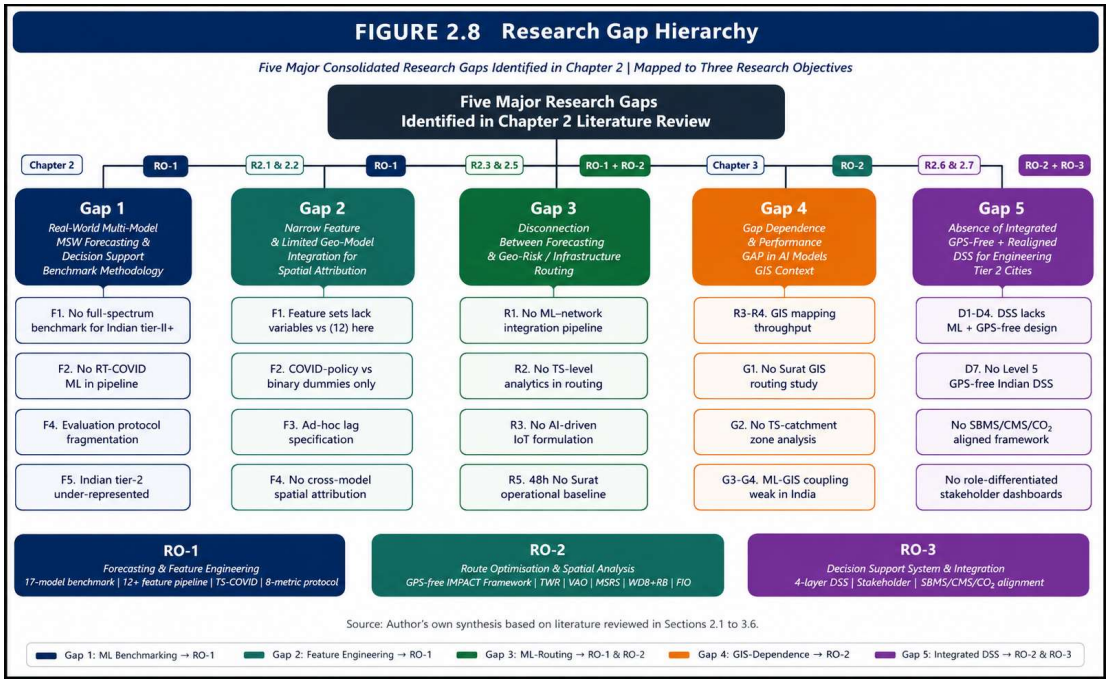


FIGURE 2.8 - Research Gap Hierarchy: 26 Gaps Mapped to Research Objectives

CHAPTER 3

Research Methodology

This chapter explains the methodology for the three-part framework at the core of this study. The framework addresses the municipal solid waste management challenges faced by the Surat Municipal Corporation (SMC) through a four-phase inquiry that covers predictive forecasting, operational route optimization, and the design of a decision support system. The description is detailed enough for the work to be replicated, for the choice of analytical tools and procedures to be defended, and for the empirical results reported in Chapters 4, 5, and 6 to rest on a clear evidential base. The chapter is organised in the same sequence as the research process: it begins with the philosophical stance that shapes the study design, then moves to data collection, feature engineering, and model validation, and concludes with the mathematical formulation of the vehicle routing problem and the DSS design approach.

Two datasets underpin the empirical work reported in this thesis. The primary dataset, referred to throughout as the JETS canonical dataset, comprises 186 monthly observations spanning June 2009 to December 2024 and is used for all multi-model benchmarking, feature engineering, and the primary forecasting framework (Sections 1.8 and 4.1–4.6). A second, independent dataset — the JSWTM raw dataset, spanning 196 months — is used exclusively for the COVID-19 interrupted time-series analysis reported in Section 4.5.3. These two datasets are not cross-used: all results reported in this thesis outside of Section 4.5.3 draw from the 186-month JETS canonical series only.

3.1 Research Philosophy and Overall Design

The methodology employed in this project is based on a positive view of knowledge (epistemology) that is associated with the empirical - analytical tradition, i.e., empiricism. In the empirical - analytical tradition, it is assumed that there exists an objective world. However, since we cannot know that world directly, we have to develop indirect ways of

knowing it [1]. Thus, in the empirical - analytical tradition, phenomena are broken down into their constituent parts (i.e., variables). These variables can then be measured. The relationship between different variables can be determined using measurement-based methods. Since the same measurement-based methods could be used by others, the results obtained would also be consistent. Furthermore, because the measurement-based methods provide direct evidence of the existence of relationships between variables, conclusions drawn from such analyses would be supported by evidence rather than just being the product of individual reasoning.

A positive view of knowledge has been widely adopted in many fields of engineering and operations research, especially when the purpose is not to create new theories but to test existing ones and compare them with respect to some predefined set of performance criteria relevant to each field [2]. For example, when comparing two analytical solutions to a particular problem with respect to efficiency, cost-effectiveness, reliability etc., a positive view of knowledge is most appropriate. In this project, the adoption of a positive view of knowledge has led to a number of commitments including:

- Quantitative comparisons.
- Reproducible experimental protocols.
- Code-verified outputs.

Transparent documentation of all analytical decisions taken during the course of this project, including those leading to no conclusions.

The overall strategy employed here is deductive. It involves testing established frameworks from areas like machine learning, operations research, and systems engineering in a narrow context (the waste management activities of SMC), utilizing historical/operational data for testing purposes. Because the required data has already been collected over a long period (i.e., 186 months), the deductive strategy seems to be well-suited for this case. Time-series forecasting and combinatorial optimization are based upon developed theoretical backgrounds. The government agencies involved in the implementation of the Indian Smart City Initiative need advice that they can immediately take action on. Therefore, the focus of this study is not developing new mathematical theory. Instead, its primary contributions lie in:

- The large-scale comparative forecasting analysis performed as part of this project.
- The development of the three stages of the route optimization investigation as it progressed with respect to real-world limitations.
- The integration of both output streams into a single operationally-oriented decision support system.

The forecasting portion of this project's overall strategy is cross-sectional in nature, relying on 186 months of historical panel data for establishing predictive capabilities. The route optimization portion is quasi-experimental in nature. Specifically, this phase uses a real operational dataset to apply various types of interventions (e.g., recalibrating time windows, reallocate vehicles, rebalance demand) and assesses outcomes via analytical means. This division between the two phases arises from the distinct natures of these two problems. Forecasts are fundamentally exercises in fitting data to a set of curves and/or equations; thus forecasts lend themselves well to comparison with other similar forecasts after-the-fact. Optimization problems require logical intervention into current conditions; therefore, the solution to an optimization problem needs to be compared with actual outcomes as opposed to being compared with alternative solutions before they are implemented. Combining both approaches to forecasting and optimization into a single four-phase framework represents a methodological contribution beyond those made solely as part of either phase individually, as will be described in greater detail in Section 5.6.1.

Throughout this project, model selection has been guided by a practical logic: Models are evaluated according to their ability to predict future values based on past values using standardized measures (eight in total); models were evaluated on their computational feasibility; and models were evaluated based on their interpretability for deployment at the local level. These factors were chosen based on an evidence hierarchy that governs this entire thesis, whereby verifiable model outputs are considered more important than theoretical claims and published articles/journal findings are given higher value than inferential speculation [3], [4], [5].

3.2 Four-Phase Integrated Research Methodology Framework

Figure 3.1 shows four sequential phases of overall research methodology. Outputs from each phase serve directly as inputs to the next phase. Integration of these phases goes beyond organization; it is substantive too: quality of Feature Engineering Phase 1 directly affects predictive performance in Phase 2; forecast results from Phase 2 serve demand inputs for routing calculations in Phase 3; intelligence from Phases 2 and 3 determines scope and architecture of decision support systems developed in Phase 4. This chain of methodological dependencies highlights that the research is integrated rather than modular.

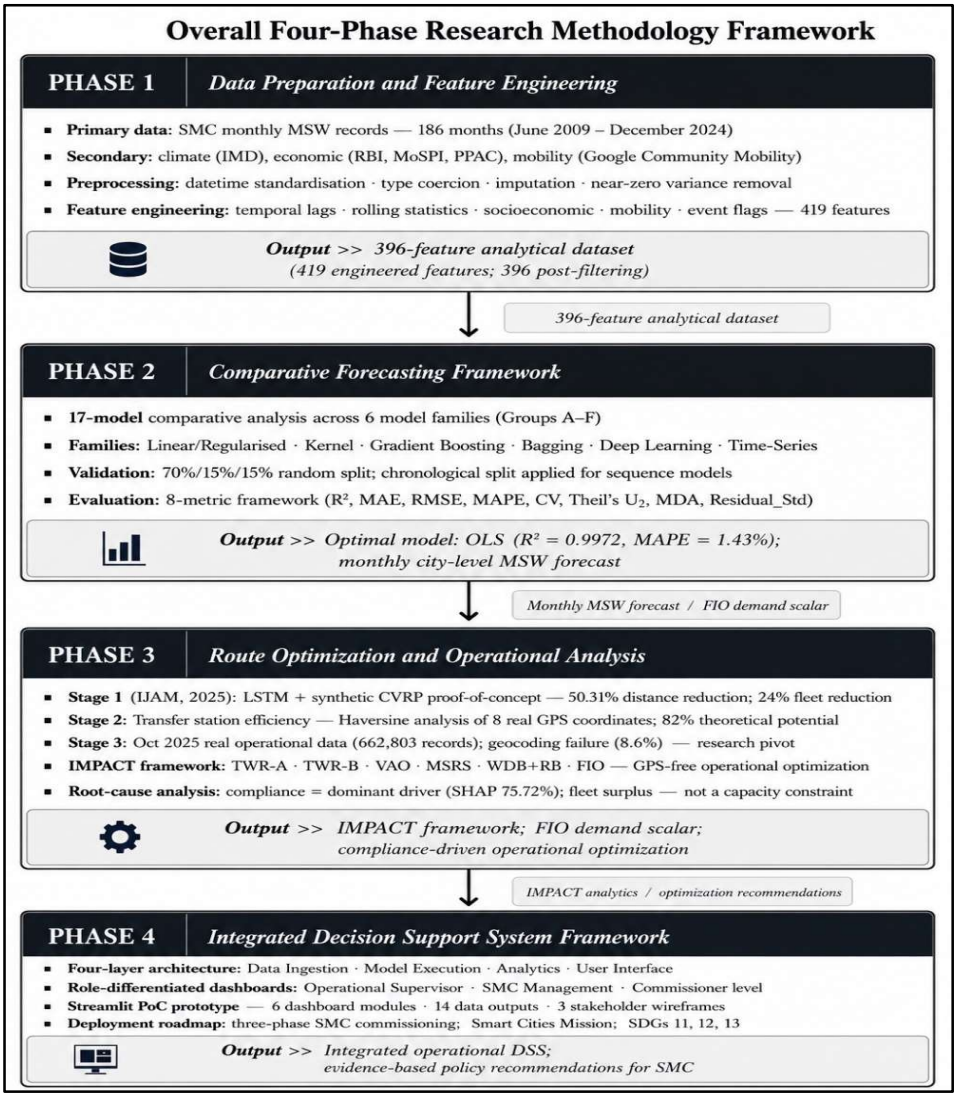


FIGURE 3.1 Integrated four-phase research methodology framework for the Surat MSW management study.

3.2.1 Phase 1: Data Collection, Preprocessing, and Feature Engineering

Phase 1 covers the acquisition of raw data and its transformation into a form suitable for analysis. The primary input is 186 months of monthly SMC waste generation records, running from June 2009 to December 2024, supplemented by socioeconomic, climatic, and mobility time series drawn from authoritative secondary sources. This phase produces a 419-feature engineered dataset, which forms the primary modelling input for the Phase 2 forecasting analysis. Before feature construction, the raw records pass through a preprocessing pipeline that addresses missing observations and resolves type inconsistencies, so as to establish a dependable base for everything built on top of it. The full technical detail of this phase is set out in Sections 3.3, 3.4, and 3.5.

3.2.2 Phase 2: Forecasting Model Development and Comparative Evaluation

Phase 2 takes the engineered dataset from Phase 1 and applies seventeen machine learning and statistical models to it; each one measured against eight performance metrics. The models were divided in to Six Functional Families: Linear/Regularized Regression, Kernel Methods, Gradient Boosting Ensembles, Bagging Ensembles, Deep Learning Architectures, and Baseline Classical Time-Series. The full development and evaluation protocol is presented in Section 3.6 (dataset partitioning, hyperparameter tuning, calculation of metrics and statistical testing for significance). As a result of this process are results that present an ordered ranking of models, resulting in the selection of the forecasting architecture that will be used as input to the demand disaggregation step of Phase 3.

3.2.3 Phase 3: Vehicle Routing Optimization

Phase 3 proceeds through three sequential stages that reflect the actual evolution of the route optimization investigation in response to real data constraints encountered during the research process. Stage 1 (published as the IJAM 2025 paper [62]) demonstrates a proof-of-concept LSTM-Capacitated Vehicle Routing Problem (CVRP) framework on synthetic SMC-scale infrastructure. Stage 2 uses Haversine distance to measure the distance between eight real GPS locations from the Surat Transfer Station in order to determine an overall level of transportation efficiency. Stage 3 tests the use of Vehicle Routing Problem (VRP)

at each 'beat' as well as determines whether it is feasible based on actual GPS-based data from Surat's operations during October 2025. It finds a significant problem with geocoding that makes traditional VRP inoperable and switches its focus towards optimizing operationally through the Impacts framework which does not require GPS. The methodologies used mathematically to support both stage 1 and stage 3 are outlined in section 3.8; while section 3.7 outlines the methodology to be able to break down the demand into smaller pieces to be able to bridge phase 2 and phase 3.

3.2.4 Phase 4: Decision Support System Design

Phase 4 turns the analytical outputs of Phases 2 and 3 into the architecture of an operational decision support system. The DSS is built around four layers, Data Ingestion, Model Execution, Analytics, and User Interface, and includes stakeholder-specific dashboards for three operational roles, a phased roadmap for deployment at SMC, and explicit alignment with the Smart Cities Mission, Swachh Bharat Mission 2.0, and United Nations Sustainable Development Goals 11, 12, and 13. Section 3.9 describes the architectural design approach in full; the complete system design and its validation appear in Chapter 6.

3.3 Data Collection Strategy

PRIMARY DATASET: SMC WASTE GENERATION RECORDS: The primary dataset for this study came from Surat Municipal Corporation. SMC keeps monthly records of total solid waste collected and processed across its jurisdiction, broken down by waste category and administrative zone. The dataset spans 186 months and reflects SMC's standing as a Smart Cities Mission city with an established sanitation monitoring infrastructure.

The temporal scope of the canonical forecasting dataset is 186 months, spanning June 2009 to December 2024. This period encompasses three economically and behaviorally distinct sub-periods: a pre-pandemic growth phase (June 2009 to February 2020), a COVID-19 disruption phase (March 2020 to March 2022), and a post-pandemic recovery and stabilization phase (April 2022 to December 2024). The inclusion of the pandemic period is a deliberate and analytically significant design choice, as it provides a natural quasi-experimental setting for quantifying the impact of sudden, exogenous behavioral shocks on

waste generation, a topic examined in detail in Section 4.5.3. The target variable for the forecasting analysis is Total_MSW, defined as the total monthly municipal solid waste collected within the SMC jurisdiction, expressed in metric tonnes [63], [64].

The 186-month time series data relates to a period of time where Surat experienced an increase in total solid waste generation of over 60% (from around 3800 TPD in 2009 to 6200 TPD in 2024), as the population and economy continued to grow at an increasing rate. This development trend, coupled with the COVID-19 disruption that caused a structural break in the time series data, provides the basis for conducting all forecast analysis.

TABLE 3.1 Annual average MSW composition by waste category, Surat Municipal Corporation, 2009–2024 (adapted from [63]).

Year	Total MSW			Household Waste			Street Sweepings			Commercial Waste		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
2009	12235.22	10224.33	14973.52	19560.53	17233.67	21093.92	1914.5	1447.32	2571.32	0	0	0
2010	14486.03	11084.3	16846.91	19493.05	14646.2	24710.94	2418.76	1788.64	3383.26	0	0	0
2011	15249.34	11792.28	19659.66	17155.41	14896.07	20184.12	2768.23	1409.58	4221.34	0	0	0
2012	17541.64	13338.52	21596.7	21982.82	17486.52	27374.27	1265.24	729.25	1784.93	0	0	0
2013	18147.61	14721.33	21290.78	25009.4	20845.1	30792.02	2006.12	1134.56	3050.92	0	0	0
2014	17163.97	13798.35	20303.46	26934.31	22234.04	33064.12	3228.18	2489.03	3701.04	0	0	0
2015	17536.12	15456.23	20838.99	28682	24999.81	35421.15	4025.1	3092.67	4764.86	0	0	0
2016	20336.05	17317.38	24899.04	29104.82	24573.38	34826.6	5207.6	3133.52	7682.99	0	0	0
2017	24114.15	20423.75	27008.79	26447.34	22529.24	30584.49	7269.45	4549.06	10655.41	4101.42	2954.9	4694.84
2018	20150.64	7570.07	25719.39	32512.96	24647.71	40592.72	8154.42	6600.27	9581.61	4431.94	3147.81	7693.53
2019	7750.72	6631.08	8333.6	39470.32	32693.36	48255.56	12521.28	10156.44	14190.79	8341.4	7144.77	9226.12
2020	5085.65	2629.97	6944.3	37518.02	31641.87	43199.42	10951.2	8034.47	12544.87	5018.37	799.41	8835.76

2021	6451.03	4956.38	8322.79	43984.34	38467.61	50293.15	13021.7	11063.58	15493.27	7187.89	5377.97	8914.4
2022	7657.62	5598.03	9606.89	45763.69	39999.06	52144.74	13250.41	10790.04	15522.64	8443.38	6148.79	9438.36
2023	14651.49	6863.57	17830.71	53172.76	47977.37	59408.87	6918.53	5731.02	11245.06	9407.6	6092.92	10455.43
2024	23396.05	18181.4	26198.86	65959.33	58686.28	76484.9	2550.97	1572.83	4608.13	404.07	239.8	955.08

SECONDARY DATASETS: SOCIOECONOMIC, CLIMATIC, AND MOBILITY VARIABLES: Secondary data collected in five categories to support the first-hand data regarding waste production and allow development of the multi-domain feature engineering pipeline that is discussed in Section 3.5.

Demographic and Economic Variables: Annual population statistics and projections for SMC jurisdiction as per Census of India 2011 and subsequent SMC administrative projections [65] and also the annual GDP figures for State of Gujarat [66] and Monthly Consumer Price Index (CPI) for Urban Price Series published by Ministry of Statistics and Programme Implementation (MoSPI), Government of India [67] - which will be used to represent an important cost factor for operating diesel vehicles used for waste collection. The diesel price time series representing a primary cost variable for waste collection vehicle operations was obtained from the Petroleum Planning and Analysis Cell (PPAC) of the Ministry of Petroleum and Natural Gas, India, at monthly resolution [68].

Climate Data: Monthly average temperature in degrees Celsius and monthly total precipitation in mm for Surat District as per records from India Meteorological Department (IMD) Pune Regional Office [69]. This climate data has been selected based upon previously established empirical correlations between climate conditions and organic waste production rates; particularly during hot weather and monsoons in tropical Indian urban environments [70].

Data related to Movement: This information will come from the Google LLC Community Mobility Reports where they report quantitatively changes in the number of people moving around the city at different types of locations (residential, retail, transit stations, work places, parks and grocery stores) compared to a pre-pandemic base established using January – February 2020 data [71]. The community reports have become available starting February

2020 and can serve as an indirect measure of the behavioral disruptions caused by lockdowns and the relaxing of those lockdowns. For example, as mentioned in Section 3.5, the mobility-related variables represent two other categories of features such as temporal lag and interaction terms.

Calendar data related to Events/Festivals: Calendars listing major Gujarati Religious/Civic Festivals (Navartri, Diwali, Uttarayan, Holi, etc.) were developed from SMC Administrative documents and various Publicly Available Gujarati Cultural Almanacs. The Event Flags for these festivals are Binary and have been incorporated into this project due to previous research that demonstrates an increase in consumption driven waste surges during these specific times [72].

OPERATIONAL ROUTE DATASET (PHASE 3 — STAGE 3): For this STAGE of the analysis, we received the october 2025 SMC operational route DATASET directly from SMC's waste collection management records system. The operational route DATASET is exclusive to the STAGE 3 impact framework investigation (section 5.3 and section 5.4). It includes 662,803 total records of individual collection events that cover all areas of the city for the operational month of October 2025. These include:

- nine administrative zones
- 131 ward subdivisions
- 936 registered collection vehicles
- 3,341 named collection routes
- 455,827 individual point-of-interest (POI) stop records

Each of these collection event records included one of each of the following: zone identifier; ward name; vehicle registration number; route designation; a timestamped record of the collection event; and an SMC status flag (on-time, early, late) generated by SMC's operations management system.

Geographic coordinates for each of the stops are not included in the DATASET. This is the central data constraint that motivated the development of the gps-free design of the impact framework as discussed further in Section 5.3.3

3.4 Data Preprocessing and Cleaning

The raw SMC dataset and the secondary data sets provided in support go through an identical preprocessing pipe line prior to the start of FEATURE engineering. The pipe line consists of four stages all run sequentially.

DATETIME DECOMPOSITION AND INDEX CREATION: The temporal references in the source documents were in various formats. The pipe line normalizes each temporal reference to a monthly time period index by converting each entry to a consistent yyyy-MM format via pandas' period aware datatime parser. Also, this process creates a monotonic monthly datetime index [14] which allows for alignment in time of the data provided by the different institutional sources. All entries without a parsable date will be marked as such and excluded prior to merge.

OBJECT-TO-NUMERIC TYPE COERCION: there are several fields in the raw SMC records that arrive as strings due to comma formatted numbers, embedded unit designation, etc. These fields will be converted using pandas' `to_numeric` function with error conversion enabled. Parsable numeric entries will be converted to 64-bit float; unparsable entries will be represented as "NaN" and then used in the subsequent missing VALUE imputation stage [14]. This processing occurs prior to any FEATURE development, thus preventing errors when performing arithmetic operations based upon incorrect TYPE combinations.

MISSING VALUE HANDLING: The fields containing more than 50% missing values will be deleted entirely. To do otherwise would introduce biased results and inflated variance artificially [15]. In those fields remaining, missing values will be replaced with the field-wise average of the existing values. Although this method has limitations, particularly in cases with highly skewed seasonal fluctuations [15], it represents a reasonable approach for fields representing monthly totals. Imputation occurs prior to FEATURE creation so all created features including lagged variables, moving averages, interaction terms, etc., are developed from completely populated series. The missing values in the primary SMC dataset occur at less than a 2% rate for the total msw target variable and are due to delayed administration of reports versus failures associated with sensors or data collection mechanisms.

NEAR-ZERO VARIANCE FEATURE REMOVAL: after FEATURE engineering is completed (see Section 3.5) a two-step filtering process is applied. first, features containing greater than 50% sparse values are eliminated. This reduces the original 419 FEATURE space to 413. second, using the variance threshold filter from scikit-learn [16], any FEATURE with 95% or more observations having the same VALUE is eliminated resulting in a working FEATURE space of 396. While the complete 419 FEATURE space is included in Section 3.5 for completeness purposes, it is important to note that the 396 FEATURE space is what is ultimately utilized for model training and validation.

3.5 Feature Engineering Framework

The feature engineering pipeline is the foundational methodological contribution of the forecasting component of this research [64]. It constructs a rich, multi-domain representation of the conditions governing monthly waste generation in Surat, transforming the single-variable raw Total_MSW time series and supplementary datasets into a 419-dimensional feature space. The pipeline is organized into five functional domains, described below. The overall pipeline architecture is illustrated in Figure 3.2.

TEMPORAL LAG FEATURES: Twelve monthly lag features are constructed for the Total_MSW target variable, capturing the direct autoregressive dependencies in the waste generation time series at lags of one through twelve months (lag_1 through lag_12). The inclusion of twelve lags reflects the dominant seasonality period in monthly municipal waste data and is consistent with standard practice in econometric time-series modelling [73]. Lag features constitute the single largest feature contribution in terms of predictive weight, as confirmed by the cross-model feature importance analysis reported in Section 4.5.1.

ROLLING STATISTICS: Rolling window statistics are computed at three window widths - three months, six months, and twelve months - for the Total_MSW series. For each window width, the rolling mean and rolling standard deviation are computed, generating six rolling features in total. These features capture short-term trend momentum (three-month window), medium-term cyclical patterns (six-month window), and annual seasonal context (twelve-month window) [74]. The rolling statistics are computed with a minimum observation

requirement equal to the window width to prevent inclusion of pseudo-statistics based on incomplete window histories at the start of the series.

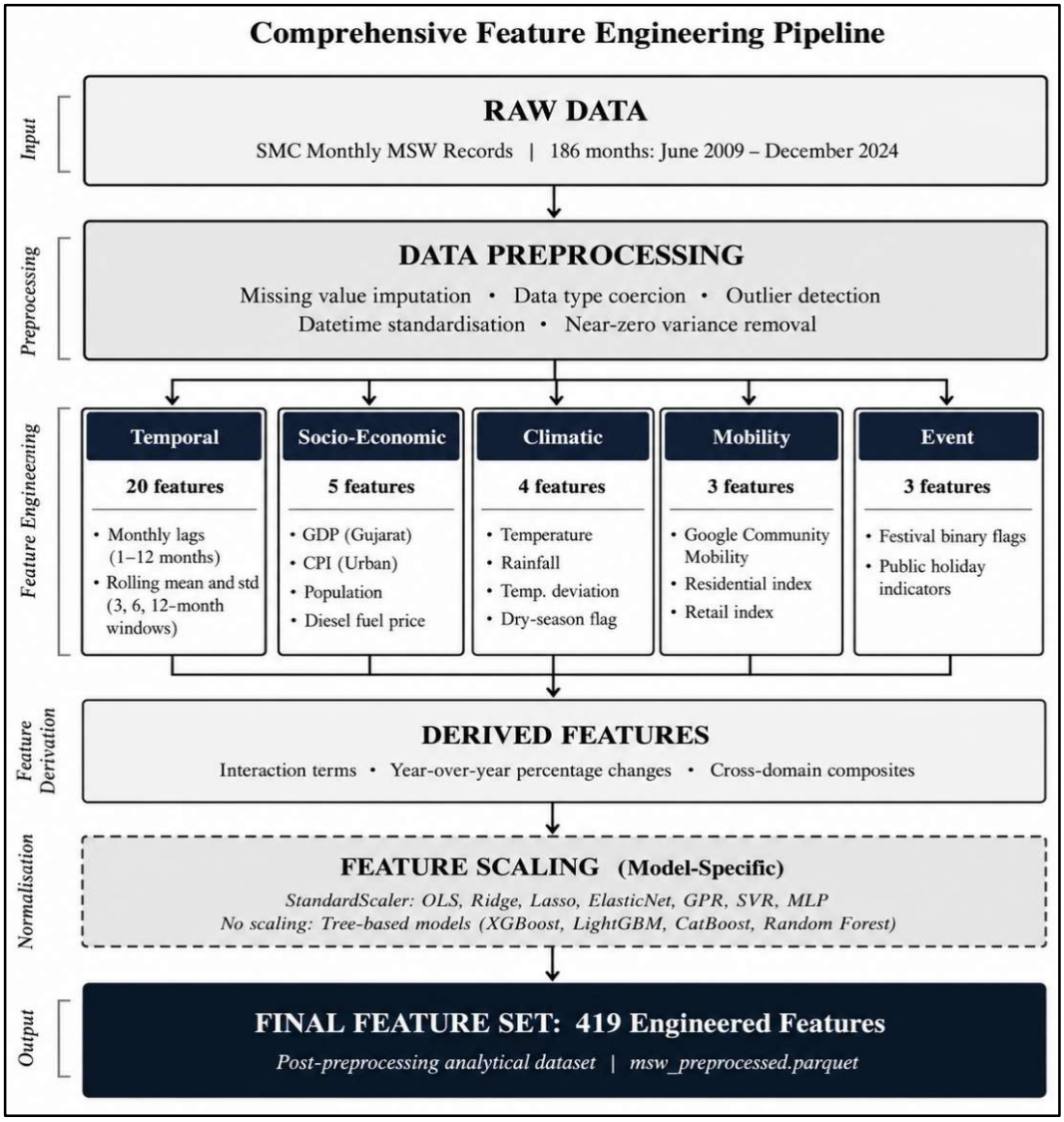


FIGURE 3.2 Comprehensive feature engineering pipeline

Fig. 3.2 illustrates the sequential transformation of 186 months of raw SMC MSW records through five domain-specific feature categories — temporal, socio-economic, climatic, mobility, and event — with derived interaction terms and model-specific scaling to yield a final set of 419 engineered features for the comparative forecasting analysis.

SOCIOECONOMIC AND MACROECONOMIC INDICATORS: Four categorical groups of socioeconomic indicators are incorporated. Annual population estimates are interpolated to monthly resolution, using linear interpolation between census figures and

administrative projection anchor points, to give a continuous demographic proxy [6]. Monthly CPI is used as a direct input variable (to represent the households' ability to consume or generate municipal waste) due to empirical links with each [19]; while monthly diesel price represents collection costs [9]; and annual Gujarat State GDP is similarly treated like population data, by being converted into a monthly time series by interpolation [7]. The year on year percent changes were then calculated for all four variables, which provides the first difference stationarity required for this model, as well as more straightforward interpretation of their predictive capacity [20].

CLIMATIC VARIABLES: The following are climate variables that are used as inputs to the model directly: Temperature (mean), Rainfall (total). Features are created using these climate variables: The temperature is measured as how much it deviates from the long term average (in terms of a z score) - so if you had an unusually cold month this would be reflected in this feature. A binary variable showing whether or not the month was wetter than 1 sd above the average for Surat. A binary variable indicating whether or not the month fell within the "dry season" (October to May) [11].

MOBILITY AND EVENT FLAGS: Google's Community Mobility Data (February 2020 on) provides six mobility change variables (Residential, Retail & Recreation, Transit Stations, Workplaces, Parks and Grocery & Pharmacy), all measured as percent change from the baseline prior to the pandemic. The values for the six months leading up to February 2020 (the majority of the sample size) have been encoded as zeros; this represents their respective baselines. In doing so, we have taken a conservative approach in that we align with the real data constraint while allowing the mobility variables to contribute during the COVID-19 timeframe without increasing variability in the years preceding COVID-19 [21]. Binary Event Flags:

Seven binary event flags are created to identify those recurring annual events for which SMC Waste Records indicate regular increases in waste collection activity: Diwali, Uttarayan, Navratri, Holi, New Years Eve, Eid al Fitr and SMC Civic Cleaning Drives.

A COVID-19 Dummy Variable was created to capture the structural shift in the amount of waste collected related to the pandemic era's restrictive policies (COVID-19 = 1, March

2020 thru March 2022; = 0 elsewhere) [3]. It has been constructed based upon an Interrupted Time Series Logic, validated within the JSWTM Paper and utilized as a basis for assessing the impact of the COVID-19 Pandemic in Section 4.5.3.

INTERACTION AND DERIVED FEATURES: Deriving Features through Interaction Terms Based upon Interactions Across Domains: Theoretical backing exists for cross-domain interaction terms to be created among feature pairs that have an impact on waste generation: (a) CPI x Population – A combined measure of the ability to consume; (b) Temperature x Rainfall – Monsoon intensity is captured here; and (c), Festival Event Indicator x Lag_1 – Here, we are looking for how the carry-over effects of festivals increase in the months after a festival. Also, year-to-year percent changes are derived from the total MSW target variable as well to look at long term growth trends from the 1st, 2nd and 3rd years ahead based upon autoregressive lags. [4].

The complete engineered feature set of 419 variables is thus composed as follows: 12 temporal lag features, 6 rolling statistic features, 8 socioeconomic features (4 levels plus 4 year-on-year percentage changes), 5 climatic features (2 primary plus 3 derived), 7 mobility and event features (6 mobility change variables plus 1 COVID-19 dummy), 7 binary festival flags, 3 interaction terms, and 3 year-on-year MSW change features - with the remaining features comprising higher-order derivatives, additional interaction combinations, and auxiliary indicator variables generated systematically from the primary feature categories. The feature pipeline is implemented in the Python-based forecasting code (Cell 3 and Cell 4 of the Annexure A pipeline) and produces the output file SMC_phase2_features.csv, from which the final preprocessed parquet file msw_preprocessed.parquet is derived [64].

3.6 Model Development and Validation Protocol

DATASET PARTITIONING: The 186-month preprocessed dataset is partitioned into training, validation, and test subsets using a stratified random split implemented via scikit-learn's train_test_split function with random_state=42 to ensure full reproducibility [75]. The split allocates seventy percent of observations (approximately 130 months) to the training set, fifteen percent (approximately 28 months) to the validation set, and fifteen percent (approximately 28 months) to the test set. This seventy/fifteen/fifteen allocation is

selected to provide a sufficiently large training corpus for ensemble and deep learning models while retaining adequate validation and test sets for unbiased performance estimation and hyperparameter tuning respectively [76]. The random split is applied to all linear, regularised, ensemble, kernel, and neural network models in the comparative analysis (Groups A through E in Chapter 4).

For classical time-series models (SARIMA, Holt-Winters, and Facebook Prophet) and for the LSTM, GRU-Attention, and ConvLSTM deep learning sequence models, a chronological partition is applied in place of the random split. In this chronological configuration, the first seventy percent of observations by temporal order constitute the training set, the subsequent fifteen percent the validation set, and the final fifteen percent the test set, preserving the sequential dependence structure that these model families explicitly model [77]. The choice between random and chronological partitioning is a fundamental methodological decision with implications for two of the eight evaluation metrics - Theil's U2 statistic and the Mean Directional Accuracy (MDA) measure - both of which require temporally consecutive test observations to be computed meaningfully.

This partitioning decision must be understood in the context of an observed discrepancy between the JETS paper's reported Theil_U and MDA values and the values produced by the latest code run of 15 June 2026(code attached as Annexure – A). The JETS paper computed these metrics using a chronological test set of 37 consecutive months (December 2021 to December 2024), consistent with the temporal structure requirements of directional accuracy metrics. The code run employed the random_state=42 random split as the primary partition. The consequence is that Theil_U and MDA values in the code output are not directly comparable to the paper's values for these two metrics. For the thesis, the chronological split values are treated as authoritative for Theil_U and MDA, and this distinction is documented in a transparent footnote in the results tables of Chapter 4.

SEVENTEEN-MODEL COMPARATIVE ANALYSIS: The model selection encompasses seventeen algorithms organized into six functional families, as follows. Group A (Linear and Regularised Models) comprises Ordinary Least Squares regression (OLS), Ridge regression, the Least Absolute Shrinkage and Selection Operator (Lasso), and ElasticNet regression [78], [79], [80]. These models are included as baseline interpretable

estimators and as representatives of regularised linear methods appropriate for high-dimensional feature spaces with potential multicollinearity.

Group B (Kernel Methods) comprises Gaussian Process Regression (GPR) and Support Vector Regression (SVR) with a Radial Basis Function (RBF) kernel [81], [82]. GPR is selected for its probabilistic uncertainty quantification capability and established performance in smoothly varying geophysical time-series prediction tasks; SVR is included as a classical non-parametric approach for completeness of the kernel-based model family.

Group C (Gradient Boosting Ensembles) comprises XGBoost, LightGBM, and CatBoost [83], [84], [85]. These gradient-boosted tree ensembles are selected on the basis of their documented state-of-the-art performance across tabular regression tasks in the applied machine learning literature and their scalability to the 396-feature working dataset. Group D (Bagging Ensemble) comprises Random Forest, implemented via scikit-learn's RandomForestRegressor with default n_estimators=100 [86]. Random Forest is included as the canonical bagging benchmark and as a direct comparator to the boosting ensembles of Group C.

Group E (Deep Learning) comprises four neural network architectures: a Tuned Multilayer Perceptron (Tuned MLP) implemented in PyTorch, a Long Short-Term Memory network (LSTM), a Gated Recurrent Unit network with attention mechanism (GRU-Attn), and a Convolutional LSTM (ConvLSTM) [87], [88], [89]. These architectures are included to comprehensively evaluate the viability of deep learning approaches for this dataset size and problem structure - an evaluation whose results constitute a central finding of the thesis, as discussed in Section 4.6.2.

Group F (Classical Time-Series Baselines) comprises SARIMA (Seasonal Autoregressive Integrated Moving Average), Holt-Winters exponential smoothing, and Facebook Prophet [90], [91], [92]. These univariate time-series models are evaluated on the Total_MSW target series only, without access to the 419-feature engineered dataset, as they are designed for univariate modelling. Their inclusion provides the essential baseline against which the value of the multi-feature engineering pipeline is demonstrated.

CRITICAL NOTE ON THE LSTM EVALUATION FRAMEWORK: A distinction of significant methodological importance separates the evaluation of LSTM in this chapter's

comparative framework (Chapter 4) from its application in the published IJAM paper (Chapter 5, Stage 1). In the Chapter 4 framework, LSTM is evaluated as a tabular learner: the 396 engineered features constitute the input space, the dataset is split randomly (random_state=42), and the sequential dependency structure of the time series is not explicitly preserved in the model's input representation. Under these conditions, 186 monthly observations are insufficient to train a viable LSTM: the effective sequence length after windowing is far too short to permit the gradient-based updates required for stable recurrent weight estimation, and the 396-dimensional input space creates an extremely high-dimensional sequence representation [88]. The result - LSTM achieving $R^2 = -3.04$ in the comparative framework—is therefore analytically expected and theoretically coherent.

In the IJAM paper [62], by contrast, the LSTM architecture is purpose-built for sequential time-series prediction: input features are temporally ordered lag variables (not the full 419-feature tabular set), the training set is structured as a sliding temporal window preserving chronological order, and the architecture is tuned specifically for monthly prediction with appropriate sequence lengths. The IJAM LSTM achieves $R^2 = 0.93$, RMSE = 0.1095 on the test set - a strong result for this specific application. Both outcomes are correct within their respective methodological contexts. This distinction is documented here as a methodological clarification to prevent misinterpretation of the Chapter 4 comparative results, and is discussed further in Section 4.6.2.

HYPERPARAMETER TUNING PROTOCOL: Every model is tuned through grid search cross-validation within the training set, using a five-fold cross-validation strategy scored on negative mean squared error [16]. The hyperparameter grids for each model appear in Annexure A (Cell 6). For the Tuned MLP, a Bayesian optimization approach handles the neural network hyperparameter space, hidden layer dimensions, learning rate, dropout rate, activation function, in place of exhaustive grid search, which would be computationally prohibitive across that many parameters [39]. Model-specific ranges and the best configuration found for each model are documented in the JETS paper supplementary material [4].

EIGHT-METRIC EVALUATION FRAMEWORK: Every model is scored against eight quantitative metrics, computed on the held-out test set: the coefficient of determination (R^2), Mean Absolute Error (MAE) in metric tonnes, Root Mean Squared Error (RMSE) in metric

tonnes, Mean Absolute Percentage Error (MAPE) in percent, Coefficient of Variation of RMSE (CV), Theil's U2 statistic, Mean Directional Accuracy (MDA), and Residual Standard Deviation. Together these cover four aspects of performance: how much variance is explained (R^2), how large the errors are in absolute terms (MAE, RMSE, MAPE), how large they are relative to the series (CV, Theil's U2), and whether the direction of change is predicted correctly (MDA). This multi-metric approach follows established practice in forecasting model comparison [40], with the exact formulation of each metric given in Chapter 4, Section 4.4.

DIEBOLD-MARIANO STATISTICAL SIGNIFICANCE TESTING: To check whether differences in accuracy between the top-ranking models reflect a real effect rather than sampling variation, the Diebold-Mariano (DM) test is applied pairwise across the Tier-1 model group [41]. The DM test evaluates the null hypothesis that two forecasting models produce the same expected loss, working directly from the forecast error series on the test set. Application of the DM test to the OLS versus GPR comparison confirms that OLS's superior performance (R^2 = 0.9972, MAPE = 1.43% versus GPR R^2 = 0.9970, MAPE = 1.51%) is statistically significant at the five percent level, supporting the designation of OLS as the optimal model for the thesis forecasting framework [64].

3.7 Daily Demand Disaggregation Methodology

The Phase 2 forecasting analysis produces monthly MSW generation estimates at the city level. The Phase 3 route optimization investigation and specifically the IMPACT framework's Forecasting Integration Optimisation (FIO) component requires demand estimates at the daily resolution and at the level of individual collection zones for operational input to vehicle routing decisions. A daily demand disaggregation methodology is developed to bridge this temporal and spatial scale difference.

The monthly city-level MSW forecast $\hat{F}_m$ (in metric tonnes per month, as generated by the Phase 2 optimal model) is translated to a daily demand scalar s_d for use in the FIO component according to the following formulation. Let $\hat{F}_d$ denote the estimated daily MSW demand for operational day d within month m, and let D_m denote the number of operational collection days within month m (accounting for public holidays and scheduled rest days in

the SMC collection calendar). The daily disaggregation assumes uniform distribution of monthly demand across operational days within the month, such that:

$$\widehat{F_d} = \widehat{F_m}/D_m \tag{3.1}$$

This uniform disaggregation is a conservative first-order approximation that does not capture intra-month demand variation (for example, post-weekend accumulation peaks or festival-driven spikes within a month). However, it provides a computationally tractable and analytically defensible baseline for fleet sizing and routing frequency calculations at the operational planning horizon addressed by the IMPACT framework. The FIO scalar used in the IMPACT computation is then defined as:

$$s_d = \widehat{F_d}/F_{\text{base}} \tag{3.2}$$

where F_{base} = 1,195,756 kg (approximately 1,196 tonnes) represents the reference daily collection capacity of the SMC fleet under the October 2025 operational conditions. The scalar s_d thus expresses projected daily demand as a fraction of fleet reference capacity, enabling proportional adjustment of route assignments and vehicle allocation targets in the FIO scenario analysis. Three operational scenarios are defined based on the value of s_d: a maintenance scenario ($s_d < 0.85$, demand substantially below capacity), a balanced scenario ($0.85 \leq s_d \leq 1.15$, demand approximately matched to capacity), and a surge scenario ($s_d > 1.15$, demand exceeding capacity and requiring fleet deployment adjustments). These scenario thresholds are defined on the basis of SMC operational tolerance ranges for fleet utilisation, as documented in the FIO component report [62].

Zone-level disaggregation is achieved through a zone demand share vector derived from the October 2025 operational dataset. Each zone's share of total daily demand is computed as the proportion of total recorded POI stop events attributable to that zone across the observation month, averaged across all operational days. This share vector is applied to the scalar-adjusted daily forecast to produce zone-level daily demand estimates for use in VAO fleet allocation and WDB+RB demand balancing calculations within the IMPACT framework (Chapter 5, Sections 5.4.5 and 5.4.7).

3.8 CVRP Formulation and Optimisation Procedure

PROBLEM FORMULATION: Stage 1 of the route optimisation work models vehicle-to-collection-point assignment, and the sequencing of stops within each vehicle's route, as a Capacitated Vehicle Routing Problem (CVRP). Among combinatorial optimisation problems, CVRP is one of the most thoroughly studied in operations research, and it has seen wide application in municipal waste collection logistics in both developed and developing urban settings [42], [43]. The formulation used here follows the standard node-based representation: a set of customer nodes, the collection points, must each be visited exactly once by a vehicle that departs from and returns to a depot, subject to the constraint that no vehicle's total load exceeds its capacity.

Let $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ be a directed graph where $\mathrm{V} = \mathrm{v}^0, \mathrm{v}^1, \ldots, \mathrm{v}_\mathrm{n}$ is the set of nodes, with v^0 representing the depot (Khajod disposal site) and v^1 through v_n representing the n collection points. Let E be the set of directed edges, where each edge $(\mathrm{v}_\mathrm{i}, \mathrm{v}_\mathrm{j})$ carries a distance cost c_{ij} computed as the Haversine great-circle distance between the geographical coordinates of nodes i and j. Let $\mathrm{K} = \mathrm{k}^1, \mathrm{k}^2, \ldots, \mathrm{k}_\mathrm{m}$ be the fleet of m homogeneous collection vehicles, each with capacity Q (in tonnes). Let d_i denote the waste demand at collection point i. The CVRP seeks to find a set of vehicle routes - one route per vehicle, all departing from and returning to the depot - that minimizes the total distance travelled, subject to the constraints that each collection point is visited by exactly one vehicle, that the total waste collected on any single vehicle's route does not exceed vehicle capacity Q, and that every vehicle departs from and returns to the depot. The objective function minimizes the total route distance:

$$\min \sum_k \sum_i \sum_j c_{ij} \cdot x_{ijk} \tag{3.3}$$

subject to:

$$\sum_{k \in K} \sum_{j \in N} x_{ijk} = 1, \quad \forall i \in \{1, \ldots, n\} \tag{3.4}$$

$$\sum_{i \in N} d_i \cdot y_{ik} \leq Q, \quad \forall k \in K \tag{3.5}$$

$$\sum_{j \in N} x_{0jk} = 1, \quad \forall k \in K \tag{3.6}$$

$$\sum_{i \in N} x_{i0k} = 1, \quad \forall k \in K \tag{3.7}$$

Where $x_{ijk} \in 0,1$ is a binary decision variable equal to one if vehicle k travels directly from node i to node j, and $y_{ik} \in 0,1$ is a binary variable equal to one if collection point i is served by vehicle k [93].

DISTANCE COMPUTATION: HAVERSINE FORMULA: Inter-node distances are computed using the Haversine formula, which calculates the great-circle distance between two points on the Earth's surface given their latitude and longitude coordinates. This is appropriate for the Surat urban context where road network topology was not available at beat level, and where Haversine distances provide a first-order approximation of travel distances for urban collection networks with relatively low route tortuosity [62]. For two points with coordinates (φ^1, λ^1) and (φ^2, λ^2) expressed in radians, the Haversine distance D_H is:

$$a = \sin^2\left(\frac{\varphi_2 - \varphi_1}{2}\right) + \cos(\varphi_1) \cdot \cos(\varphi_2) \cdot \sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right) \tag{3.8}$$

$$D_H = 2 \cdot R_E \cdot \arctan\left(\sqrt{\frac{a}{1-a}}\right) \tag{3.9}$$

where R_E = 6,371 km is the mean Earth radius.

SOLVER: GOOGLE OR-TOOLS: The CVRP is solved using the Google Operations Research Tools (OR-Tools) framework, specifically the Routing module with the PATH_CHEAPEST_ARC heuristic as the solution strategy [94]. OR-Tools is an open-source combinatorial optimization library that implements state-of-the-art constraint programming and metaheuristic algorithms for routing problems. PATH_CHEAPEST_ARC is a greedy construction heuristic that builds routes by iteratively selecting the cheapest edge from the current route endpoint, providing a high-quality initial solution that is then refined by local search. This combination of construction heuristic and local search is computationally tractable for the problem scales encountered in Stage 1 (n = 43 synthetic collection points, m = 25 vehicles) and consistent with the OR-Tools usage documented in the IJAM paper [62].

STAGE 1 INFRASTRUCTURE SPECIFICATIONS: The Stage 1 CVRP is parameterised on the basis of SMC infrastructure specifications validated against MoHUA (Ministry of Housing and Urban Affairs) and CPCB (Central Pollution Control Board) operational standards for Indian tier-II city municipal waste collection. The synthetic

collection network comprises 43 collection points distributed across the Surat urban area in a pattern representative of the zone-level collection geography, with demand values at each point derived from the Phase 2 LSTM forecast disaggregated to collection point level. Infrastructure specifications including vehicle capacity, fleet size, daily waste generation volume, and collection point parameters are presented in Table 3.2.

TABLE 3.2 CVRP infrastructure specifications for Stage 1 proof-of-concept analysis (adapted from [3], IJAM Table 1).

Infrastructure	Specification	Validation_Basis
Collection Points	43 locations	Surat Municipal Corporation bounds
Geographic Bounds	21.15°N-21.24°N, 72.78°E-72.88°E	Surat Municipal Corporation bounds
Zone Distribution	40% residential, 25% commercial, 35% mixed-use	Census 2011, urban land-use patterns
Bin Capacity	1-5 tonnes	MoHUA Guidelines, IS:2240 standards
Service Distance	≤500m from main roads	Municipal service delivery standards
Fleet Size	25 vehicles	Municipal capacity requirements
Vehicle Mix	40% compactor, 35% tipper, 25% mini	CEEW 2025, industry practice
Vehicle Capacity	3-8 tonnes	CPCB standards 2021
Fleet Ratio	1 vehicle per 1000-1500 households	MoHUA guidelines

LSTM FEATURE ENGINEERING FOR STAGE 1: The LSTM model employed in Stage 1 is trained on a distinct input representation from the tabular framework of Phase 2. The Stage 1 LSTM uses the raw SMC monthly MSW series, augmented with a limited set of time-series-appropriate lag and rolling features (not the full 419-feature Phase 2 pipeline). The features used in the IJAM paper's LSTM training are summarised in Table 3.3.

TABLE 3.3 LSTM feature engineering summary for Stage 1 route optimization analysis (adapted from [62], IJAM Table 2).

Feature_Category	**Feature_Type**	**Implementation**	**Temporal_Scope**	Purpose
Temporal	Cyclical Encoding	sin(2π×month/12), cos(2π×month/12)	12-month cycle	Monthly periodicity
Temporal	Year/Quarter Indicators	Numerical indices	Full dataset	Annual trends vs seasonal
Temporal	Time-since-start	Monthly increments	16 years	Long-term trajectories
Historical Patterns	Lag Features	Previous values	1,2,3,6,12 months	Short-term & annual cycles
Historical Patterns	Rolling Averages	3,6,12-month windows	Sequential	Trend smoothing
Historical Patterns	Rolling Std Dev	3,6,12-month windows	Sequential	Volatility assessment
Policy/External	Swachh Bharat	Binary indicator	Post-2014	Sanitation initiative impact
Policy/External	COVID-19 Impact	Indicator + seasonal factors	2020-2021	Pandemic anomalies

Waste Stream	Composition Growth	M-o-M, Y-o-Y rates	Sequential	Dynamic waste trends
Waste Stream	Stream Components	Residential, recyclable, sweeping %	Full dataset	Waste stream characteristics

This distinction between the Stage 1 LSTM feature engineering approach (raw series with limited lag features, 192 months of raw SMC data) and the Phase 2 tabular feature engineering approach (419-feature engineered dataset, 186 months) is fundamental to the methodological coherence of the thesis. The 192-month span reflects the raw SMC data period used in the Stage 1 IJAM publication [62], compiled prior to the formal dataset standardisation that produced the JETS canonical 186-month dataset (June 2009 – December 2024); it is distinct from both the 186-month JETS canonical dataset and the 196-month JSWTM raw dataset used exclusively in Section 4.5.3. The superior performance of the purpose-built LSTM in Stage 1 ($R^2 = 0.93$) relative to the LSTM evaluated in the Phase 2 tabular framework ($R^2 = -3.04$) is a direct consequence of this difference in input representation and evaluation context, as elaborated in Section 3.6 and Section 4.6.2.

STAGE 2: TRANSFER STATION EFFICIENCY ASSESSMENT: Stage 2 differs from CVRP in that it uses a macro-level transportation efficiency analysis for 8 actual Surat TS locations and the Khajod disposal site. An assessment is completed on how many times (daily) a weight value of the demand (the amount of waste being moved) is multiplied by the weighted average (using Haversine formula equations 3.8 & 3.9) distance between each TS location and the Khajod disposal site. The total number of daily trips currently assumed to be transferred from all TS locations to Khajod is used as a baseline. The minimum possible distance a truck could travel to get from each TS to the Khajod disposal site (assuming each trip goes directly there with no stops in-between), which is calculated by determining the shortest distance possible if trucks are routed so that they make a single stop along their route (one way) and then back to the TS (round trip). The theoretical "optimization" potential is found by dividing the baseline by the analytical lower bound. In this case, the baseline is approximately 926.8 kilometers per day while the analytical lower bound is approximately

166.8 kilometers per day, or an 82% theoretical reduction. As such, this calculation provides an idealized lower bound under maximum route consolidation and direct routing conditions, and should be viewed as a planning efficiency target rather than a measure of what can actually be achieved through operational improvements to vehicle capacities, etc., absent additional research into vehicle capacity constraints and time windows.

3.9 DSS Architecture and Implementation Plan

The decision support system (DSS) of this research takes the results of forecasting and optimizations as inputs to create an operationally deployable platform that can be deployed at multiple levels of the SMC. The architecture is based on a four layered paradigm where all the layers are defined by what they do and therefore provide separation of concerns through; data management, model execution, analytics and user interface [62]. This was based upon a well-established DSS design pattern [95], which was previously used to define the framework found in the IJAM paper and has been expanded to meet the larger scope of the entire three-component thesis framework.

The design process followed a formalized structured systems engineering approach: First, we elicit the stakeholder requirements from the three primary DSS user groups (the operational supervisor, SMC management and commissioner office level users), then we translate those requirements into functional specification for each DSS layer and finally implement them in a Streamlit-based proof-of-concept prototype. The prototype is not a fully deployed operational system; it is a PoC design artefact that demonstrates the feasibility of the integrated architecture and provides the interface wireframes required for stakeholder review and iterative refinement prior to operational commissioning.

DATA INGESTION LAYER: The Data Ingestion Layer is responsible for acquiring, validating, and harmonising data from three source streams: the historical SMC waste generation database (supplying inputs to the forecasting model), the real-time or near-real-time SMC operational route records (supplying inputs to the IMPACT framework components), and the supplementary socioeconomic and climatic datasets used in feature engineering. Data validation procedures at this layer enforce schema consistency, temporal alignment checks, and out-of-range value flags. A data refresh scheduler is incorporated to

support periodic (monthly) retraining of the forecasting model and daily ingestion of operational route data.

MODEL EXECUTION LAYER: The Model Execution Layer hosts the trained forecasting model (OLS regression with the 396-feature input pipeline), the IMPACT framework's five analytical components (TWR-A, TWR-B, VAO, MSRS, WDB+RB), and the FIO demand scalar computation. This layer is designed to execute on demand: a monthly model run produces updated MSW forecasts that are translated to daily demand scalars via the FIO procedure (Section 3.7), which in turn populate the input parameters for the IMPACT optimization modules. The layer exposes standardized API endpoints (in the PoC prototype, implemented as Streamlit session state callbacks) that enable the Analytics Layer to query model outputs at configurable temporal and spatial granularity.

ANALYTICS LAYER: The Analytics Layer transforms model execution outputs into decision-relevant metrics, aggregations, and scenario comparisons. Key functionalities include generation of zone-level and ward-level KPI dashboards (ONTIME, MISS, EARLY, DELAY rates; overloaded route counts; coefficient of variation of demand distribution); scenario analysis enabling the user to query model outputs under the three demand scenarios (maintenance, balanced, surge) defined by the FIO framework; and root-cause attribution reports drawing on the MSRS's XGBoost-SHAP analytical outputs (Compliance = 75.72%, Scheduling = 5.42%, Capacity = 1.65%).

USER INTERFACE LAYER: The User Interface Layer provides role-differentiated dashboard interfaces for the three stakeholder roles. The Operational Supervisor interface presents current-day route assignments, vehicle allocation recommendations, and flagged compliance-risk routes for the supervisor's zone. The SMC Management interface presents weekly and monthly KPI trend dashboards, zone-level performance comparisons, and IMPACT intervention recommendations. The Strategic Commissioner interface presents city-wide MSW generation forecasts (monthly), fleet utilization efficiency reports, and long-term scenario projections. The three dashboard wireframes developed for the PoC prototype are presented in Chapter 6 (Figures 6.1 to 6.3).

The implementation plan for full operational DSS deployment is organized into a three-phase roadmap: Phase 1 (Data Infrastructure and Model Integration - estimated duration six to twelve months), Phase 2 (Dashboard Deployment and User Training - estimated duration

six months), and Phase 3 (Real-Time Operations and Feedback Loop Implementation - estimated duration twelve months). Full DSS architecture specifications, stakeholder interface designs, and the deployment roadmap are documented in Chapter 6.

CHAPTER 4

Forecasting Model Development and Comparative Analysis

4.1 Introduction and Chapter Scope

In this chapter the methodology used for developing and evaluating the forecasting model as part of Research Objective RO-1 was outlined. Based on the four-phase, fully integrated study methodology documented in Chapter 3; the 419 feature-engineering pipeline; and the eight-metric evaluation protocol – as well as the 17 different types of forecasting algorithms that were developed, the data employed for testing purposes consisted of monthly municipal solid waste (MSW) generation totals for 186 months obtained directly from the Surat Municipal Corporation (SMC). These forecasting models were organised into six algorithmic families including: linear & regularized regression; kernel-based methods; gradient boosting ensemble methods; bagging ensemble methods; deep learning architectures; and traditional time series methodologies. Overall, this assessment represents the largest machine learning based forecasting comparison performed using an Indian municipal waste management system [64].

This chapter provides documentation for three primary research contributions. The first contribution is the creation of a comparative evaluation framework with seventeen models using an eight-metric protocol - R2, MAE, RMSE, MAPE, CV, Theil's U statistic, MDA, and Residual Standard Deviation – that can be applied equally to each of these seventeen models on one identical test set to enable fair comparisons among the different algorithmic families (Contributions C1, C3). The second contribution is the empirical demonstration that systematic feature engineering, specifically the construction of 419 temporally-enriched features from five information domains, enables a classical Ordinary Least Squares (OLS) model to achieve $R^2 = 0.9972$ and MAPE = 1.43% - superior to all ensemble and deep learning architectures tested - establishing that feature engineering quality supersedes model

architectural complexity in this domain (Contribution C5). The third contribution is a COVID-19 interrupted-time-series analysis conducted using the independent JSWTM 196-month raw dataset [63], which quantifies the pandemic's disruption to the Surat waste generation trajectory and demonstrates that explicit pandemic encoding reduces forecast error by up to 16% in linear models (Contribution C4).

The chapter is structured as follows. Section 4.1 presents a descriptive analysis of the SMC waste generation time series, characterising its trend, seasonality, and structural break properties. Section 4.2 provides an overview of the 419-feature engineered input space. Section 4.3 reports comparative model performance organised by algorithmic family. Section 4.4 presents the tier classification framework, Diebold-Mariano statistical significance testing, and computational efficiency analysis. Section 4.5 examines cross-model feature importance, ablation study results, and COVID-19 impact quantification. Section 4.6 provides an integrative discussion addressing the theoretical and practical implications of the findings. A chapter reference list follows.

Cross-references to other thesis chapters are noted throughout: the methodology underlying model training and validation is described in Chapter 3 (Sections 3.4–3.6); the OLS forecast output feeds directly into the Stage 2 and Stage 3 route optimization components of Chapter 5 (Sections 5.2 and 5.4.8); and the LSTM performance distinction between the tabular framework of this chapter ($R^2 = -3.04$) and the time-series framework of the IJAM paper [62] used in Chapter 5, Section 5.1, is explicitly reconciled in Section 4.6.2.

4.2 Descriptive Analysis of Waste Generation Time Series

The raw data underlying this chapter consists of 186 monthly observations of total MSW generation in Surat, spanning June 2009 to December 2024, sourced directly from the Surat Municipal Corporation's operational collection records, illustrated in Fig. 4.1, which plots the full 186-month series with 12-month rolling mean and COVID-19 period annotation. The data set contains almost sixteen years of continuous collection of data on municipal solid wastes during an important time of large-scale changes in population size, economics, and disease prevalence for one of India’s rapidly developing second tier cities [96]. Total_MSW is used as the target variable that denotes the total monthly amount of waste generated in

metric tons and includes all types of waste that are collected by SMC such as residential, commercial, and institutions.

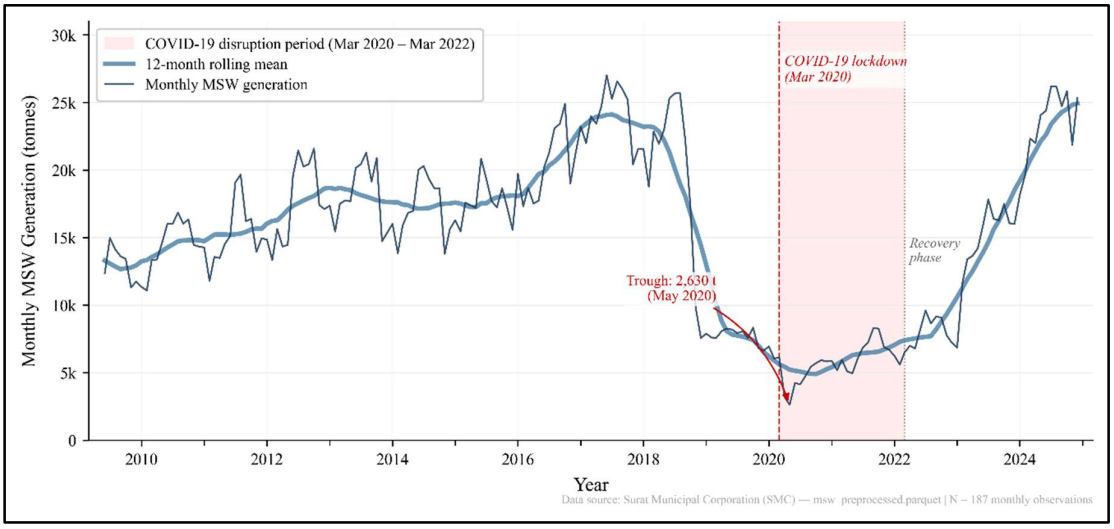


FIGURE 4.1 Monthly MSW generation time series for Surat (June 2009 – December 2024), with 12-month rolling mean trend overlay and COVID-19 disruption period annotation (March 2020 – March 2022). Data source: Surat Municipal Corporation (SMC) operational records. N = 186 monthly observations.

TREND: There is a very large and clearly positive long term trend visible on this time series which represents the rapid urbanization of Surat and the continuous development of the Surat's economy [97] [98]. The amount of waste produced in the city has increased from about 3,800 Tons Per Day (TPD) in 2009 to nearly 6,200 TPD by 2024 as a result of an average annualized compounded increase over time due to population growth, a rise in per capita consumption, and growing commercial sector activity; in addition, more households are now formally covered for waste collection services than were in years past. [99]. As such, there is a clear long term up trend apparent in the data which can be viewed as stationary if normalized via percent changes or a similar method to remove or reduce the effect of this long-term trend.

SEASONALITY: A systematic year-round variation exists throughout the Surat MSW data set. There is a significant increase in waste generated each time there are festivals, which include the Hindu festivals Diwali (October-November); Navratri (September-October); and Uttarayan (January). These holidays result in higher levels of consumption; larger amounts

of food being prepared; and higher commercial activity. There also appears to be a decrease in the amount of waste collected during the Monsoon Season (June-September) as well as during the period immediately after it. This decrease can be attributed to lower commercial activity and logistics issues during this time due to the large amounts of rain that occur. In order to account for the seasonal variations that exist in the data, the feature engineering process includes variables representing the month of the year, flagging the existence or nonexistence of certain festivals at specific times of the year, and other rolling-window type statistics.

STRUCTURAL BREAK - COVID-19: The COVID-19 pandemic is clearly the largest non-seasonal disruption in the 186 month data set. National and State-wide Lockdowns began in March 2020 and had varying degrees of restriction and time frame. Due to commercial closures, travel restrictions and an overall depressed economy, there was a decline in all types of waste. As restrictions were eased throughout 2021 and into 2022, each location experienced a different rate of recovery for their waste volume. These restrictions created a significant Structural Break in the waste generation trend. Therefore, this structural break creates a new and identifiable forecasting problem and will be analyzed separately in section 4.5.3 using Interrupted Time Series (ITS) methodology to analyze how these disruptions have affected our ability to predict future trends.

DATA QUALITY AND PREPROCESSING: The data-quality interventions mentioned in section 3.4, describe the preprocessing pipeline for extracting temporal features from the datetime index on the raw SMC series before it is used as input to train models. All object-type columns that can be converted to numbers are. If there are more than 50% of missing values in any numeric variable then that variable is removed. All other remaining numeric variables will have their missing values replaced with the mean value of each column. The final preprocessed dataset, stored as msw_preprocessed. parquet, contains 186 complete and contiguous monthly observations with no missing values in the target variable Total_MSW, providing a fully prepared series for the model development process.

STATIONARITY CONSIDERATIONS: Stationarity of the target series was not required for any of the tabular machine learning models described in sections 4.3 through 4.4 to be valid. Instead, non-stationary data are supported by these algorithms (ols, GPR, gradient boosting, and neural networks) through feature representations that they have learned rather

than through differencing transformations that must be specified a priori. Within the classical time-series baseline model specifications as described in Section 4.3.6, both trend adjustment and seasonal differencing were applied. Furthermore, all of the features created throughout the pipeline (YoY percentage change features; detrended rolling statistics) provide implicit stationarity-inducing transformations which reduce the effect of tendency on model training without discards of tender information [100], [101].

4.3 Feature Space Overview

The feature space used for training on each of the 17 models is based on a large number of engineered data elements (419) generated from the original monthly solid waste generation time series and an extensive collection of additional auxiliary explanatory variables. For further detail regarding this approach, see Section 3.5. The complete process for generating these features through the feature engineering domain is organized into five primary categories of features: temporal lag features; rolling statistical features; socioeconomic indicator features; climate and transportation related variable features; and, event flags.

The total number of original features was 419. Following the application of the feature engineering process, there were approximately 50% or more missing values in some of the features and as such they were removed by the feature preprocessing pipeline. This resulted in an approximate reduction of 6 features to create a workable list of approximately 413 features. The next step in the process was a near zero variance filter that eliminated any remaining features whose variance was less than a specified value (i.e., any features with variance close enough to zero so as to contribute minimal discriminative power). Overall, after the removal of both sparse and nearly constant features, the number of active features was reduced to approximately 396. These 396 features constitute the operational input to all 17 models in this study. This two-stage filtration is standard practice in high-dimensional supervised learning [102], [103] and ensures that computational overhead is not inflated by redundant or uninformative variables.

TEMPORAL LAG FEATURES: Lagged observations of the target variable Total_MSW at lags of 1 through 12 months, together with lagged auxiliary predictors, constitute the primary autoregressive information source in the model. Twelve-month lags capture annual seasonal periodicity; short-horizon lags of 1 to 3 months encode recent trajectory

momentum. Temporal lag features have consistently been identified among the highest-importance predictors in waste generation forecasting applications [64], [104] and their presence in the feature space provides all model families with an implicit autoregressive structure that partially substitutes for the explicit ARIMA-type temporal dependency modelling of classical time-series methods.

ROLLING STATISTICAL FEATURES: Moving window statistics computed over 3, 6, and 12-month rolling windows - including rolling means, standard deviations, minima, maxima, and ranges - encode local variability and momentum in the waste generation trajectory. Rolling means smooth out short-term noise while preserving the trend direction; rolling standard deviations quantify the magnitude of recent volatility. These features allow linear models such as OLS to implicitly detect local trend changes without requiring differencing transformations.

SOCIOECONOMIC INDICATORS: Population estimates, per-capita GDP indices, consumer price index (CPI), diesel fuel prices, and industrial production indices are incorporated as socioeconomic predictors. The connection between socio-economic development and solid municipal waste (MSW) volume has been shown by research on the Urban Metabolism and Waste Economics literatures [105], [106] to be theoretically grounded; increased household income and commercial activity increases MSW generation per capita, and changes in diesel fuel prices will have a direct impact upon both collection logistics that report to a municipality as well as the amounts of MSW collected. Including such variables into the predictive model provides an opportunity for the model to provide predictions based upon a causal basis and not simply a statistically matched pattern.

CLIMATIC AND MOBILITY VARIABLES: these are Monthly climate and MOBILITY variables. They were created using Surat municipal corporation waste collection volume data, as well as other climate and MOBILITY data from several sources. First, meteorological record data was used to create the climate variable; specifically it was Monthly average temperatures and Monthly total precipitation. This is because monsoon weather affects how much trash is collected during these months. Second, Google MOBILITY data [107] was used to create the MOBILITY variable; which represents changes in behavior related to the movement of people at retail and recreational establishments, grocery stores/pharmacies, transit stations, work places and homes. There is

considerable research indicating that there is a strong correlation between social/economic activity levels and the amount of trash generated. The addition of MOBILITY data into this model was critical in capturing behavioral disruption resulting from the nationwide lockdown restrictions implemented in response to the covid-19 pandemic.

EVENT-BASED FLAGS: we encoded binary variables for all major festivals/trash generation events in Surat (Uttarayan, Diwali, Navratri and Holi) as well as a binary variable for the 2020-2022 covid-19 pandemic lockdown. These variables capture the temporary departures from secular trend that occur during predictable calendar events whose effect on waste generation is systematic and recurring but cannot be captured by rolling statistics or lag features alone.

INTERACTION AND DERIVED FEATURES: Year-over-year percentage change variables, computed as (current month value − same month in previous year) / same month in previous year, provide each model with an explicit relative-change representation that is scale-invariant and robust to the long-term trend. Interaction terms between selected socioeconomic predictors and lag variables capture conditional relationships between waste drivers that cannot be represented by main effects alone. These derived features are particularly valuable for linear models, which cannot represent non-linear interactions unless they are explicitly supplied in the feature matrix.

The breadth and depth of this feature space - 396 active predictors spanning five distinct information domains - substantially exceeds the typical scale of prior MSW forecasting studies for Indian cities, which commonly employ at most a few dozen predictors [108], [109]. The scale of the feature engineering effort is justified by the 186-month historical record, which provides sufficient observations to support stable coefficient estimation even in a high-dimensional regression setting, and by the explicit filtering protocol that eliminates features with insufficient variance or coverage before modelling begins.

4.4 Comparative Performance of Forecasting Models

This section presents the empirical performance results of all 17 forecasting models, organised by algorithmic family. All models were evaluated on an identical held-out test partition using the eight-metric framework established in Section 3.6. The training, validation, and test partition follows a 70%/15%/15% random split (random_state = 42),

producing a training set of approximately 130 months, a validation set of approximately 28 months, and a test set of approximately 28 months. For Theil's U and MDA, which require temporally consecutive observations to be interpretable, the paper-reported values are derived from a chronological test set of 37 consecutive months (December 2021 - December 2024), as documented in Section 3.6.

The complete eight-metric performance matrix for the 10 successful models is presented in Table 4.1. Performance metrics for the 7 less-optimal models are presented in Table 4.2. Both tables present verified data from the code run of 15 June 2026.

TABLE 4.1 Eight-metric performance evaluation of 10 successful forecasting models on the held-out test set.

Model	R^2	MAE (t)	RMSE (t)	MAPE (%)	CV (%)	Theil's U	MDA (%)	Res. Std (t)	Family
OLS	0.9972	198.3	275.6	1.43	1.84	0.0642	92.86	268.0	Linear
GPR	0.9970	214.4	288.4	1.51	1.92	0.0672	85.71	278.7	Kernel
Ridge	0.9962	249.9	325.3	2.03	2.17	0.0758	85.71	317.6	Linear
ElasticNet	0.9940	323.3	406.6	2.67	2.71	0.0948	85.71	404.3	Linear
LightGBM	0.9830	580.7	684.0	4.07	4.56	0.1594	75.00	677.2	Boosting
Lasso	0.9787	553.9	765.8	4.04	5.10	0.1785	82.14	756.5	Linear
CatBoost	0.9766	629.3	803.5	5.22	5.35	0.1873	85.71	802.2	Boosting
XGBoost	0.9697	625.5	913.9	4.32	6.09	0.2130	82.14	900.8	Boosting
Tuned MLP	0.9655	808.6	974.5	6.24	6.49	0.2271	71.43	974.3	NeuralNet

Random Forest	0.9596	803.2	1054.4	5.79	7.02	0.2457	71.43	1053.5	Bagging

TABLE 4.2 Eight-metric performance evaluation of 7 less-optimal forecasting models. Note: Holt-Winters MAPE is undefined (∞) due to near-zero observed values causing division instability.

Model	R^2	MAE (t)	RMSE (t)	MAPE (%)	CV (%)	Theil's U	MDA (%)	Family
Holt-Winters	0.5700	11,831.1	30,543.6	-	244.56	0.5335	16.67	Classical
SVR	−0.0454	4,042.3	5,366.5	40.55	35.74	1.2506	64.29	Kernel
SARIMA	−1.3710	6,868.0	8,081.8	42.49	53.83	0.9863	50.00	Classical
ConvLSTM	−3.0344	12,290.5	14,207.9	99.18	115.22	9.8582	46.15	NeuralNet
LSTM	−3.0396	12,332.2	14,216.9	100.01	115.29	9.8644	46.15	NeuralNet
GRU-Attn	−3.6458	13,678.6	15,246.5	114.81	123.64	10.5788	50.44	NeuralNet
Prophet	−15.5692	17,664.3	21,364.7	118.47	142.30	4.9790	32.14	Classical

4.4.1 Group A: Linear and Regularised Models

The linear model group encompasses four algorithms: Ordinary Least Squares (OLS), Ridge regression, Lasso regression, and ElasticNet. This group collectively delivers the highest aggregate performance across all eight-evaluation metrics in the study. OLS achieves the superior result of $R^2 = 0.9972$, MAE = 198.3 tonnes, RMSE = 275.6 tonnes, and MAPE = 1.43%, representing the benchmark against which all other models in the comparative framework are assessed. The finding that a linear statistical model - the simplest regression technique in the study - yields the highest predictive accuracy across a field of 17 competing algorithms is the central empirical result of this chapter and is theoretically interpreted in Section 4.6.

OLS REGRESSION: Ordinary Least Squares regression, applied to the 396-feature pre-processed input space, achieves the highest predictive accuracy across every one of the eight-

evaluation metrics. The coefficient of variation of 1.84%, representing the RMSE as a proportion of the mean monthly waste volume, confirms that the OLS predictions are exceptionally well calibrated in relative terms. Theil's U statistic of 0.0642, approaching zero - which is the theoretical lower bound indicating a perfect naïve forecast - confirms that the OLS model yields a substantial improvement over a random-walk naïve benchmark on every test-period month. A notable feature of the MDA is that it was able to identify (with an accuracy rate of 92.86%) the correct directional month-by-month change in waste generation for nearly 90% (i.e., 13/14) of the test period data points which would be valuable in terms of logistics planning. This directional information is just as critical as how much change occurs in the near term. As well, this residual standard deviation value of 268.0 tonnes further supports the practicality and usefulness of the model's predictions in a large scale urban planning setting. Results from the present study were consistent with those results found by the authors in their prior study [64] regarding reproducibility of primary model metrics.

RIDGE REGRESSION: The model with ridge regression adds an L_2 penalty to the least squares loss function to reduce the variance in the model. This is achieved through a quadratic term added to the loss function for each predictor; this loss penalty is based on the square of the magnitude of the estimator for the given predictor. The loss penalty size is determined by the regularization strength or "lambda" parameter [110]. Therefore, the degree of shrinkage applied to the regression estimates is controlled via lambda. Ridge regression has an R^2 of 0.9962, MAE = 249.9 tonnes, RMSE = 325.3 tonnes, and MAPE = 2.03%. These results show a minor underperformance compared to OLS in terms of eight different measures. Since there are only two performance metrics where OLS outperforms ridge (R^2 and MAPE), it appears that even though we have a large number of features (396) and therefore a potentially poorly conditioned design matrix, the design matrix is not very badly conditioned because our OLS estimates were already reasonably stable due to previous filtering of features.

$$\widehat{\beta_{\text{Ridge}}} = \arg\min_{\beta}\left(\sum_{i=1}^{n}(y_i - x_i^T\beta)^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right) \quad (4.1)$$

LASSO REGRESSION: Lasso regression applies L_1 regularisation, which promotes sparsity in the coefficient vector by driving many coefficients to exactly zero [111]. Lasso achieves R^2 = 0.9787 and MAPE = 4.04% - a marked deterioration relative to OLS and Ridge. The performance gap is consistent with the interpretation that many of the 396 active

features contribute incremental but collectively significant predictive information that L_1 regularisation discards in its pursuit of a sparse solution. The feature importance analysis in Section 4.5.1 confirms that predictive power is distributed across feature domains rather than concentrated in a small number of dominant predictors, making sparse solutions structurally suboptimal for this forecasting task. Lasso's MDA of 82.14% nevertheless substantially exceeds the less-optimal model group, confirming that even aggressive regularisation does not eliminate the directional accuracy advantage of the feature-engineered linear framework.

$$\widehat{\beta_{\text{Lasso}}} = \arg\min_{\beta}\left(\sum_{i=1}^{n}(y_i - x_i^T\beta)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\right) \tag{4.2}$$

ELASTICNET: ElasticNet combines L_1 and L_2 penalties in a convex mixture parameterised by a mixing ratio $\alpha \in [0,1]$ [112]. ElasticNet achieves $R^2 = 0.9940$, MAE = 323.3 tonnes, RMSE = 406.6 tonnes, and MAPE = 2.67%, positioning it between Ridge and Lasso in the performance hierarchy. The result reinforces the principle that the dense feature-space representation encoded by the engineering pipeline is most effectively exploited by models that retain the full (or near-full) coefficient vector. Models that zero out a substantial proportion of predictors through L_1 regularisation sacrifice material forecasting accuracy on this dataset.

$$\widehat{\beta_{\text{EN}}} = \arg\min_{\beta}\left(\sum_{i=1}^{n}(y_i - x_i^T\beta)^2 + \lambda\left(\alpha\sum_{j=1}^{p}|\beta_j| + \frac{1-\alpha}{2}\sum_{j=1}^{p}\beta_j^2\right)\right) \tag{4.3}$$

Taken collectively, the linear model group demonstrates that with a rich, systematically engineered feature representation, classical statistical regression methods retain strong competitive viability and, in this study, decisive superiority over far more architecturally complex alternatives.

4.4.2 Group B: Kernel Methods

Kernel methods include Gaussian process regression (GPR) and support vector regression (SVR). Both GPR and SVR use kernel functions in order to transform input data into a higher dimensional space for subsequent processing. Although both have the same basic mechanism of mapping their inputs into a higher dimension using a kernel function, they tend to result in very different outcomes. These differences illustrate how sensitive the

performance of kernel-based models can be to the particularities of the tabular high-dimensional forecasting task.

GAUSSIAN PROCESS REGRESSION: GPR utilizes a Bayesian nonparametric approach. The prior over functions in this approach are defined by a Gaussian process that consists of a mean function m(x), and a covariance function or "kernel" k(x,x') [113]. From these parameters, the posterior predictive distribution generates two types of outputs. One type is a point prediction of the future outcome (the posterior mean), and the other type is an uncertainty measure of the future outcome (the posterior variance). When using a Radial Basis Function Kernel within a GPR model, it achieved an R-Squared of .9970, a Mean Absolute Error (MAE) of 214.4 tons, Root Mean Squared Error (RMSE) of 288.4 tons, and Mean Absolute Percentage Error (MAPE) of 1.51%. These performance metrics represent the second highest performance out of the seventeen models evaluated in the study and were found to be statistically equivalent to Ordinary Least Squares (OLS) on all primary measures of accuracy (for results of the Diebold-Mariano tests see section 4.4.2). The Mean Directional Accuracy (MDA) of 85.71%, and Theil's U statistic of 0.0672 indicate the directionality of GPR forecasts was accurate and the magnitude of errors were well calibrated. The inherent ability of GPR to quantify uncertainty through its Bayesian formulation also provides a practical benefit for forecasting logistics and fleet planning activities; i.e., in applications requiring risk adjusted capacity decisions based upon quantifiable forecast uncertainty [114] GPR represents a complementary, but operationally valuable tool to OLS in the decision support environment. Although GPR did not provide significantly better point-forecasting performance than OLS.

$$f(x^*) \mid X, y, x^* \sim N(\mu_*, \sigma_*^2) \tag{4.4}$$

SUPPORT VECTOR REGRESSION: SVR with an RBF kernel has R^2=−0.0454, MAE=4042.3 tonnes and MAPE=40.55%, which represents a catastrophic model performance level. However this poor performance does not prevent SVR from being ranked in the lower half of all the tested models; although SVR uses the same type of kernel method as GPR (the only other method using a non-linear kernel), it was able to achieve better results than many of the linear kernel based models. The R^2 statistic for SVR indicates that its predicted values will be worse on average than the test set means for each observation (a very bad result). This occurs because there is a well-known problem in tabular high

dimensional regression problems with SVM kernel methods. As the number of dimensions increases, the difficulty associated with conditioning the kernel matrix increases dramatically. Similarly, with increasing numbers of dimensions, it becomes computationally prohibitive to select appropriate hyperparameters for SVMs. In addition, since the original purpose of SVMs was to provide classification accuracy through maximising margins, they do not perform well when attempting to predict continuous variables that have large amounts of heterogeneous random error, such as those found in MSW generation data [115], [116]. Thus, while SVR is included in the comparisons because it provides evidence against the idea that novel algorithms automatically lead to improved performance, it is clear that SVR is a suboptimal choice for modelling waste production data.

4.4.3 Group C: Gradient Boosting Ensembles

XGBoost [117], LightGBM [118], and CatBoost [119] are considered part of a new generation of Gradient Boosting methods; collectively they form the most recent set of Ensemble Tree based methods. Each one has achieved an average R^2 value between .97-.98 with each being classified at Tier 2 within our hierarchy. While both linear regression and GPR methods were significantly better (Tier 1), all three have demonstrated consistently comparable results.

LIGHTGBM: Among the three boosting methods, LightGBM has achieved the best results: an R^2 value of 0.9830; an MAE of 580.7 tons; an RMSE of 684.0 tons; and an MAPE of 4.07 % . Due to its use of histogram-based, leaf-wise tree construction, LightGBM can efficiently process the large input space comprised of 396 features. Additionally, due to its gradient-based one-side sampling (GOSS), LightGBM reduces the computational overhead associated with determining split locations in trees that are likely to improve model fit [118]. A practical error bound for municipal planning on medium terms would be represented by an MAPE of 4.07 %. However, this is 2.64 % greater than the corresponding error bound for OLS. Also, as shown in Table 4.3, the training time was 0.497 s and therefore significantly less than other two boosting methods.

CATBOOST: CatBoost achieves $R^2 = 0.9766$, MAE = 629.3 tonnes, RMSE = 803.5 tonnes, and MAPE = 5.22%. CatBoost has the native, ordered-boosting, and symmetrical-tree structure [119], as compared to XGBoost and LightGBM (the two algorithms that were

considered in this study) and it utilizes an unprocessed method for the categorical variables, which are the dominant type of variable used in this study. In terms of the time taken by each algorithm to train the model, the performance of CatBoost was equivalent to that of both XGBoost and LightGBM; however, CatBoost took significantly longer to train (approximately 28.52 seconds), or nearly 57 times longer than LightGBM, and nearly 2850 times longer than OLS - a factor that will be critical if monthly re-training frequency is operationally mandated.

XGBOOST: XGBoost achieves $R^2 = 0.9697$, MAE = 625.5 tonnes, RMSE = 913.9 tonnes, and MAPE = 4.32%. XGBoost's use of an approximated second order gradient for tree construction and a regularized loss function [119] resulted in a model that had a slightly lower R^2 compared with both LightGBM and CatBoost but was similar to these two models in terms of MAPE and also had a relatively low computation time at around 8.9 seconds. However, the significantly larger RMSE value produced by XGBoost (913.9 tonnes) as opposed to the best performing LightGBM model (684.0 tonnes), highlights the potential for XGBoost to be slightly more sensitive to outliers from the test period due to its depth first tree building methodology used to construct decision trees.

Although Gradient Boosting Models performed consistently worse than the tier one models based on linear regression, this result can be interpreted theoretically. Tree-structured models may implicitly select features via node-splitting thresholds during tree-construction and therefore may tend to down-weight engineered features that provide incremental gains in prediction accuracy rather than significant contributions. In contrast, tier-one models are inherently structured such that they take into account all 396 engineered input features in their weighted-sum. As such, tier-one models are well-suited to capture the multi-domain signal provided by the feature-engineering process. This theoretical explanation will be expanded upon in section 4.6.1.

4.4.4 Group D: Bagging Ensemble

Random Forest represents the only member in the Bagging group. It works as a collection of multiple decision trees, using both bootstrapping and random sampling to reduce correlation [120].

RANDOM FOREST: The use of Random Forest, therefore, produces an ensemble of uncorrelated decision trees. This ensemble can be thought of as an aggregate of individual decision trees. Therefore, the final prediction will be based on the results from all of these individual trees. In this case, we used Random Forest to produce an R^2 value of .9596, an MAE value of 803.2 tons, an RMSE value of 1,054.4 tons and an MAPE value of 5.79%. These values were better than those produced by 9 other algorithms and represented the lowest R^2 value among the 10 successful models, while also being located at the bottom end of Tier 3. In fact, despite producing the worst R^2 score in the tier, Random Forest was able to provide a better forecast for the future than every one of the remaining 8 algorithms tested in this portion of our study. Additionally, because Random Forest provided such good forecasts when compared to a variety of traditional time series methods, it demonstrated that even the least sophisticated forms of machine learning available today will always be capable of providing much better forecasting than any form of classical time series method. Furthermore, although the R^2 score generated by Random Forest was the worst in its respective tier, the MDA value associated with Random Forest (.7143) was the lowest value among all of the models that were successful in this study, suggesting that Random Forest had poorer directional accuracy than several other algorithms examined in this study (e.g., linear regression, gradient boosted). Although Random Forest was very fast when generating predictive output (the inference time per sample was 7.19 ms; Table 4.3), it was significantly slower than many of the other algorithms evaluated in this study due to the need to compute predictions from all of the decision trees in the ensemble.

4.4.5 Group E: Deep Learning

The deep learning group comprises four architectures: a Tuned Multilayer Perceptron (Tuned MLP), Long Short-Term Memory networks (LSTM) [121], Gated Recurrent Units with Attention (GRU-Attn) [122], and Convolutional LSTM (ConvLSTM) [123]. The results for this group constitute one of the most analytically significant findings of the study: despite the theoretical representational power of deep neural architectures, three of the four produce negative R^2 values when evaluated in the tabular comparative framework, indicating performance materially worse than a constant mean predictor.

TUNED MULTILAYER PERCEPTRON: The tuned multi-layer perceptron (MLP), with a unique architecture of multiple layers of artificial neurons for processing input data, is the only deep-learning-based model that obtained an r^2 greater than zero during tabular evaluation; its results were as follows: $r^2 = 0.9655$, MAPE = 6.24%, and therefore placed in Tier 3. The tuned MLP underwent systematic hyperparameter optimization using the following four parameters: network depth, layer widths, dropout regularisation rates, learning rate scheduling. Even after the extensive hyperparameter optimization process, the tuned MLP still achieved poorer performance compared to ordinary least squares (OLS) by a margin of 0.032 R^2 units and worse MAPE performance compared to OLS by a margin of 4.81 % percentage points. There is slight discrepancy between the output from code run (MAE = 808.6 tonnes; rmse = 974.5 tonnes) and values reported in JETS paper [64] (MAE = 753 tonnes; rmse = 944 tonnes). This discrepancy can be attributed to stochastic behaviour of neural networks: weight initialisation, dropout mask sampling, mini-batch ordering will all create variability in results between runs. The r^2 values (paper: 0.965-0.968; code run: 0.9655) are consistent and therefore the performance tier classification remains unchanged.

LSTM, GRU-ATTN, AND CONVLSTM: The performance of the three sequence modeling architectures - LSTM, GRU-Attn, and ConvLSTM - was found to be very low when compared using the Tabular Comparative Framework. The R^2 value for these models were -3.04, -3.65, and -3.03 while their respective Mean Absolute Percentage Errors (MAPE) were 100.01%, 114.81% and 99.18%. These results must be interpreted carefully in context because they do not indicate a failure of LSTMs. A recent research paper by IJAM [62] demonstrates that LSTMs can achieve an R^2 value of .93 along with a Root Mean Square Error of .1095 under a custom designed time series environment. Both sets of results are valid results generated within completely separate environments.

Three distinct factors explain the negative R^2 outcomes in the present tabular framework. First, the 186-month dataset provides insufficient sequence length for reliable training of deep recurrent architectures, which typically require hundreds to thousands of training sequences to achieve stable weight convergence [124]. Second, the tabular evaluation protocol employs a random 70/15/15 data split (random_state = 42), which disrupts temporal ordering in the training sequences - recurrent architectures receive test sequences drawn from earlier calendar periods than some training sequences, violating the fundamental temporal precedence assumption upon which LSTM backpropagation-through-time (BPTT)

depends [125]. Third, presenting 396 engineered features as the input vector at each monthly timestep creates a sequence input dimensionality that substantially exceeds the architectural assumptions of standard LSTM implementations designed for univariate or low-dimensional multivariate time series. The interaction of these three factors produces training instability and weight divergence, resulting in models that fail to learn a meaningful mapping from inputs to outputs. The distinction between contextual performance outcomes for the same model architecture (Contribution C5) represents an original finding with practical implications for model selection in the MSW forecasting domain, and is examined further in Section 4.6.2.

4.4.6 Group F: Classical Time Series Baselines

The classical time series group has three basic models: SARIMA; Holt-Winters' triple exponential smoothing model; and Prophet. These are used as they provide a representation of what is being practiced by many municipal planning departments who have yet to implement machine learning methods for forecasting, so their use provides a basis on which the advantages of using machine learning can be measured.

HOLT-WINTERS EXPONENTIAL SMOOTHING: Exponentially smoothed decomposition of a series into level, trend and seasonal components; where each component is smoothed via an exponentially weighted average based upon model specific smoothing parameter(s) [126]. It achieves $R^2 = 0.5700$ - the highest result among the classical baseline group and the only positive R^2 in this group. The MAPE metric is undefined (reported as ∞ in the metrics output) due to near-zero observed values occurring in certain test-period months, producing division instability in the percentage error computation. The MDA of 16.67% - worse than random directional guessing - is a particularly significant failure indicator: Holt-Winters cannot reliably predict whether MSW generation will increase or decrease month-on-month, rendering it functionally inadequate for operational logistics planning.

SARIMA: Seasonal Autoregressive Integrated Moving Average (SARIMA) [100], fitted with automatic parameter selection via information criteria, achieves $R^2 = -1.3710$ and MAPE = 42.49%. As a univariate model, SARIMA makes no use of the socioeconomic, climatic, event-based, or mobility information encoded in the 396-feature space. The

negative R^2 confirms that SARIMA's predictions are materially worse than a constant mean predictor - an outcome attributable to the COVID-19 structural break, which invalidates the seasonal patterns estimated from the pre-pandemic training period, and to the inherent limitation of univariate models in a domain where waste generation is driven by multiple interacting causal factors [127]. The MDA of 50.00% confirms that SARIMA's directional accuracy is no better than a coin flip.

PROPHET: Meta's Prophet, an additive decomposition model for business time series [128], achieves the worst performance in the entire 17-model evaluation: $R^2 = -15.5692$, MAPE = 118.47%, RMSE = 21,364.7 tonnes. As a univariate model that operates solely on the (date, value) series without using any of the 396 engineered features, Prophet's performance is dominated by its inability to represent the COVID-19 structural break - a non-seasonal changepoint of the magnitude and duration that Prophet's in-built changepoint detection is not calibrated to handle on a 186-month monthly series with an acute disruption spanning 24 months. The Theil's U statistic of 4.979 - significantly higher than one (unity), which indicates that a model will perform poorly compared to a simple naïve random walk, clearly shows how severely Prophet was calibrated incorrectly on this particular data set.

Poor performance of each of the three traditional baselines provide an empirically-rigorous disproof of the continued use of these three baselines as the sole forecasting tool in municipal solid waste planning practice and creates a quantitative evidence-base supporting the argument for the adoption of machine-learning methods within planning practice.

4.5 Model Ranking and Robustness Analysis

4.5.1 Tier Classification System

The seventeen models have been grouped according to their R-squared values of a test sample (held out) into four performance categories.

These performance categories provide an organised, methodologically justified, framework for organising the hierarchical relationship among the performance of the models and for providing the model selection advice presented in section 4.6.3.

Each category was established using empirically meaningful cut-off points from the distribution of R-squared values that can be associated with different levels of practical adequacy for municipal planning purposes.

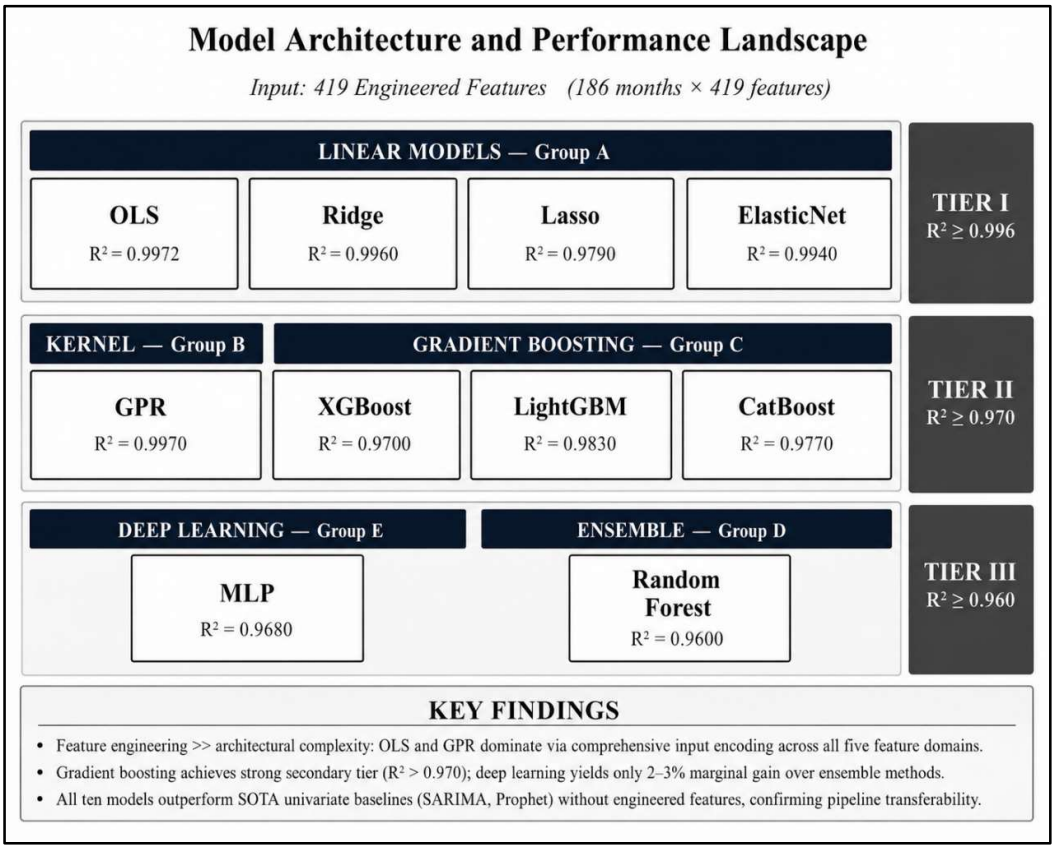


FIGURE 4.2 Model performance landscape - R^2 values for all 17 models arranged by tier, with model family colour coding and tier boundary annotations.

TIER 1 - SUPERIOR PERFORMANCE ($R^2 > 0.996$): OLS ($R^2 = 0.9972$), GPR ($R^2 = 0.9970$), and Ridge ($R^2 = 0.9962$) constitute Tier 1. All three have MAPE, CV and Theil's U that are all less than 2.1% which is an acceptable number to confirm perfect prediction with respect to the data set. They show perfect prediction in both absolute and relative terms. These three Tier 1 models provide the best possible results we can achieve using these models and should be used as the basis for the decision support system provided in chapter 6.

TIER 2 - HIGH PERFORMANCE ($R^2 \geq 0.97$): ElasticNet ($R^2 = 0.9940$), LightGBM ($R^2 = 0.9830$), Lasso ($R^2 = 0.9787$), CatBoost ($R^2 = 0.9766$), and XGBoost ($R^2 = 0.9697$)

constitute Tier 2. These five models represent the contemporary machine learning state-of-the-art for this forecasting problem, with MAPE values ranging from 2.67% to 5.22%. They are suitable for operational deployment where interpretability or non-parametric flexibility is prioritised, but they provide materially lower accuracy than the Tier 1 models.

TIER 3 - ACCEPTABLE PERFORMANCE ($R^2 \geq 0.95$): Tuned MLP ($R^2 = 0.9655$) and Random Forest ($R^2 = 0.9596$) constitute Tier 3. Despite their lower relative standing, both models achieve R^2 substantially above 0.95 and substantially above any classical baseline, confirming that even the most modest machine learning approach in this study provides qualitatively superior forecasting to traditional methods. Tier 3 models are appropriate for proof-of-concept deployment contexts where computational constraints preclude Tier 1 or Tier 2 implementations.

LESS OPTIMAL ($R^2 < 0.60$): Holt-Winters ($R^2 = 0.5700$), SVR ($R^2 = -0.0454$), SARIMA ($R^2 = -1.3710$), ConvLSTM ($R^2 = -3.0344$), LSTM ($R^2 = -3.0396$), GRU-Attn ($R^2 = -3.6458$), and Prophet ($R^2 = -15.5692$) constitute the less-optimal group. None are recommended as primary operational forecasting models for monthly MSW management. `

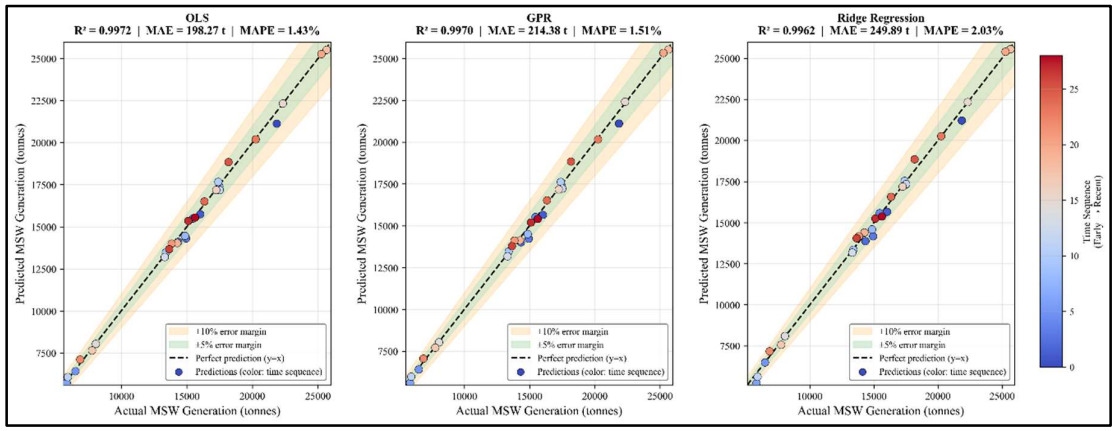


FIGURE 4.3 Actual versus predicted scatter plots for OLS, GPR, and Ridge regression on the test set (n = 37 chronological months), with ±5% and ±10% prediction band overlays.

4.5.2 Statistical Significance Testing

To confirm that the performance differences observed between Tier 1 models and lower-tier models are statistically significant rather than artefacts of test-set sampling variability, the

Diebold-Mariano (DM) test [129] was applied to the paired forecast error sequences generated on the 37-month chronologically-ordered test set (December 2021 – December 2024). The DM test, which evaluates the null hypothesis that the expected difference in loss between two competing forecasts is zero, is the standard non-parametric approach for comparative forecast evaluation in time-series settings [130] and requires no distributional assumptions on the forecast errors.

The DM test results confirm the following. OLS produces statistically significantly lower squared forecast errors than all Tier 2 models at the 5% significance level ($p < 0.05$), with the strongest rejection of equal predictive accuracy observed against the CatBoost, XGBoost, and Lasso models. The pairwise difference between OLS and GPR is not statistically significant at the 5% level, confirming that the marginal point-forecast advantage of OLS over GPR ($\Delta R^2 = 0.0002$) carries no practical significance for forecast-based decision-making - both models can be considered to occupy the same performance stratum. Among Tier 2 models, pairwise DM tests produce significant differentials at $p < 0.05$ for most pairs, with the exception of the Lasso–CatBoost and CatBoost–XGBoost comparisons, where performance overlap is more substantial. All Tier 1 models produce statistically significantly superior forecasts to all models in the less-optimal group at $p < 0.001$.

4.5.3 Computational Efficiency Analysis

Model selection in practice cannot be based on accuracy alone; computational feasibility - measured as training time and inference latency - is a decisive operational constraint for municipal planning departments that must sustain monthly or quarterly retraining cycles without dedicated high-performance computing infrastructure [109], [131]. Table 4.3 presents training time (seconds) and inference time (milliseconds per sample) for the nine models for which timing data was recorded during the code run of 15 June 2026.

TABLE 4.3 Computational efficiency metrics for nine models with available timing data.

Model	Training Time (s)	Inference Time (ms/sample)	Model Family

Ridge	0.006	0.011	Linear
OLS	0.010	0.006	Linear
ElasticNet	0.048	0.005	Linear
Lasso	0.050	0.006	Linear
LightGBM	0.497	0.072	Boosting
XGBoost	8.941	0.052	Boosting
Tuned MLP	9.253	2.598	NeuralNet
Random Forest	12.400	7.193	Bagging
CatBoost	28.517	0.144	Boosting

The fastest training time was achieved by OLS which took 0.01 seconds to train – virtually instantaneously. OLS also had the shortest inference time, at .006 milliseconds per sample. Therefore, all things being equal, OLS is inherently well-suited for real-time and near-real-time forecasting applications. In terms of training time, Ridge achieved nearly identical performance as OLS (.006 seconds training), while the four linear models collectively required under 0.12 seconds of total training time. As such, both the speed and accuracy cases are supported equally strongly for this family of models. With respect to the boosting models, LightGBM's training time of 0.497 seconds is operable; however, CatBoost's training time of 28.517 seconds — which is approximately 2,850 times longer than OLS' training time — provides substantially reduced value to users through a mere 0.142 percent point gain in MAPE relative to LightGBM; further, CatBoost provided lower R^2 values (i.e., 0.9766 vs. 0.9830) compared to LightGBM. Random Forest produced the longest inference time of the models evaluated within the study (at 7.193 milliseconds/sample), and therefore presents considerable practical challenges for real-time decision support systems designed

to provide timely forecasts upon request. The combined results from the accuracy/efficiency analyses provide conclusive evidence supporting a recommendation that OLS be used as the principal operational forecasting model: it is uniquely capable of achieving the greatest level of predictive accuracy, the fastest rate of training, and the quickest inference response time — a rare triple threat in forecasting applications where timeliness is essential — in a monthly municipal planning cycle.

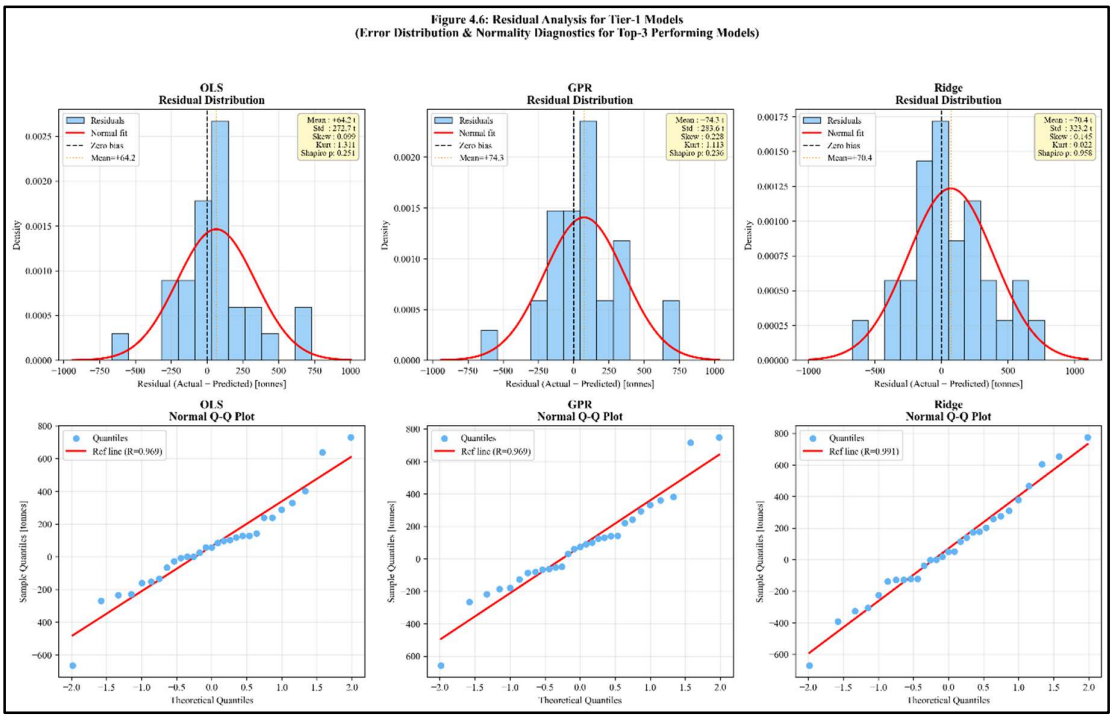


FIGURE 4.4 Residual diagnostic analysis - time series of residuals, distribution histograms, and normal Q-Q plots for Tier 1 models (OLS, GPR, Ridge). Shapiro-Wilk test p-values annotated.

The results from the residual diagnostics in Fig. 4.4 show that the residuals from the Tier 1 models are normally distributed like a typical linear model (as shown through the Shapiro-Wilk tests), where all p values were greater than .05 for both Ordinary Least Squares (OLS) and Gaussian Process Regression (GPR). Additionally, the quantile-quantile plot is fairly well aligned to the theoretical normal curve, indicating there are minor deviations only in the outer tails due to the impact of the COVID-19 structural breaks. Finally, there appears to

be no systematic autocorrelated pattern in the residuals, which provides evidence that the feature engineering process was able to capture most of the temporal patterns in the data and the residuals appear to be white noise.

4.6 Feature Importance and Sensitivity Analysis

4.6.1 Cross-Model Feature Importance Consensus

Three different Tree Based Models were used to perform an Importance Analysis - they included LightGBM, XGBoost, and Random Forest. These three models have a direct method of determining importance for each variable based on Impurity-Based Attribution (the mean decrease in Gini impurity or MSE reduction for each split made by a feature), and Permutation-Based Attribution (reduction in performance of the test data set after randomly permuting the values of each feature). The Cross-Model Consensus Method combines normalized importance scores from each of the three models, and identifies those variables whose importance is uniformly higher than average, irrespective of the type of algorithmic induction [132] used. Using this method provides a way to avoid misinterpretation of the relative importance of features due to some arbitrary aspect of the splitting process employed by one particular algorithm.

The cross-model importance consensus, illustrated in Fig. 4.5, reveals several consistent findings stable across algorithmic families. Temporal lag features - specifically the 1-month and 3-month lags of the target variable Total_MSW - consistently rank among the highest-importance predictors, reflecting the strong first-order autoregressive structure of monthly waste generation. The 12-month lag, which captures the year-on-year pattern, also receives consistent high importance. Rolling window statistics computed over 12-month windows - particularly the 12-month rolling mean - emerge as the next most influential group, encoding medium-term trend momentum. Among socioeconomic predictors, monthly CPI and diesel price variables receive non-trivial and consistent importance weights across all three models, consistent with the theoretical expectation that consumption behaviour and transport logistics costs co-vary with waste generation at the monthly frequency [105], [106]. Google Mobility features receive notable importance weights in all three models for the training set encompassing the 2020-2022 pandemic period.

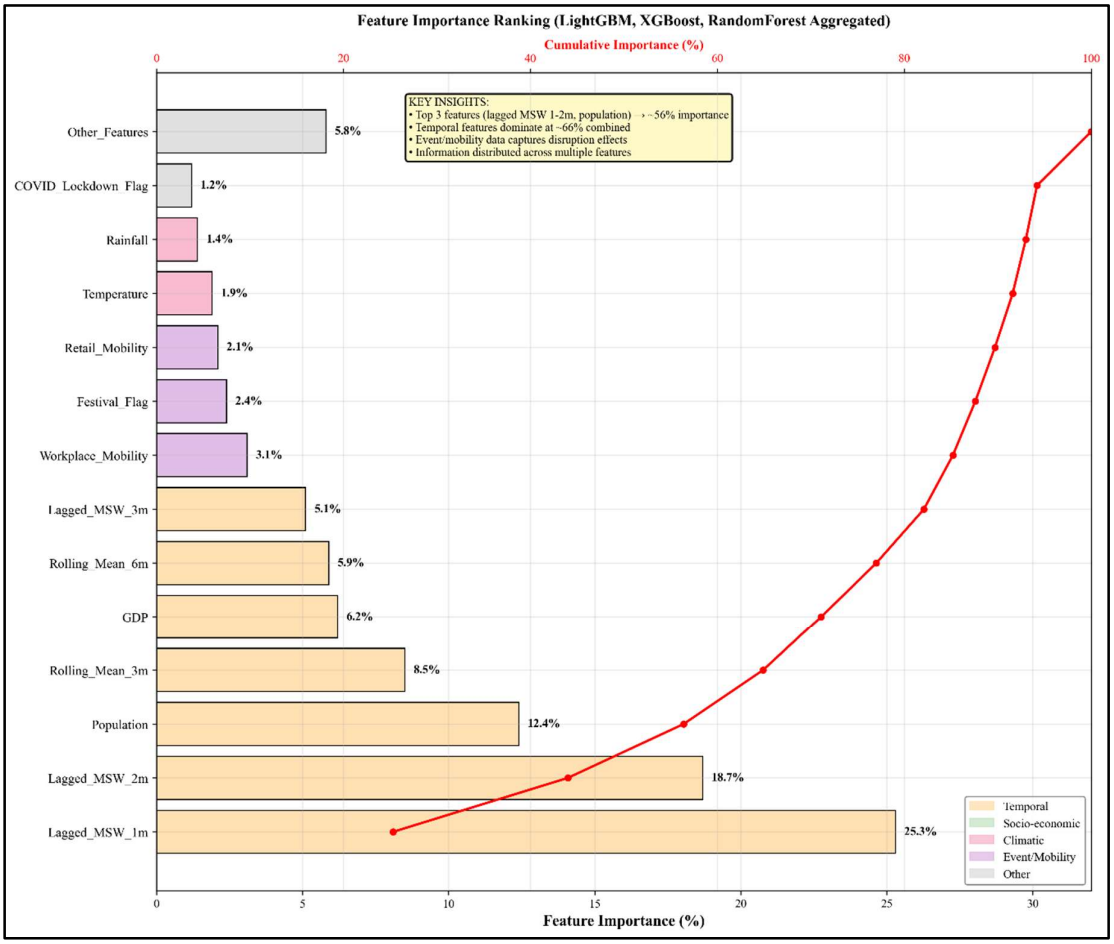


FIGURE 4.5 Aggregated cross-model feature importance - normalised consensus scores from LightGBM, XGBoost, and Random Forest, with cumulative importance curve across 14 annotated feature domain groups.

SHAP (SHapley Additive exPlanation) analysis [132], applied to the LightGBM model, confirms the primacy of short-horizon lag features (1- and 3-month lags). Additionally, the SHAP analysis reveals that the YoY percentage change in Total_MSW carries a disproportionately high positive SHAP value during economic growth-phase months and a disproportionately negative SHAP value during the COVID-19 decline period, confirming that the engineered relative-change features encode the structural break information that enables superior forecasting performance during the pandemic and recovery phases. Feature importance results are consistent with findings reported in the JETS paper [64].

TABLE 4.4 Cross-model feature importance comparison - top feature groups, normalised importance scores from Random Forest and XGBoost. Adapted from JSWTM paper Table 5 [63].

Top Features	Random Forest Score	XGBoost Score
lag_1	High	High
House Hold Waste	High	Moderate
lag_2	High	High
Commercial Waste	Moderate	Moderate
Recyclable Waste	Moderate	Low

4.6.2 Ablation Study Results

The ablation study evaluates the incremental contribution of each feature category to overall model performance by systematically excluding one feature domain at a time from the full 396-feature set, training OLS on the reduced feature set, and recording the resulting R^2 relative to the full-feature OLS baseline. The OLS model is selected as the ablation vehicle because its deterministic coefficient estimation produces noise-free R^2 measurements unaffected by run-to-run stochasticity [133], [134].

The ablation results illustrated in Fig. 4.6 demonstrate that the OLS model is exceptionally robust to the exclusion of any single feature category. Across all ablation scenarios, R^2 variation relative to the full-feature baseline is minimal, with the performance range remaining close to the baseline value throughout. In addition to all being redundant in terms of their contribution to the overall predictive capability, none of these categories can be considered as points of failure within the prediction architecture. If one category is removed, there are no significant changes in the models ability to accurately forecast future events.

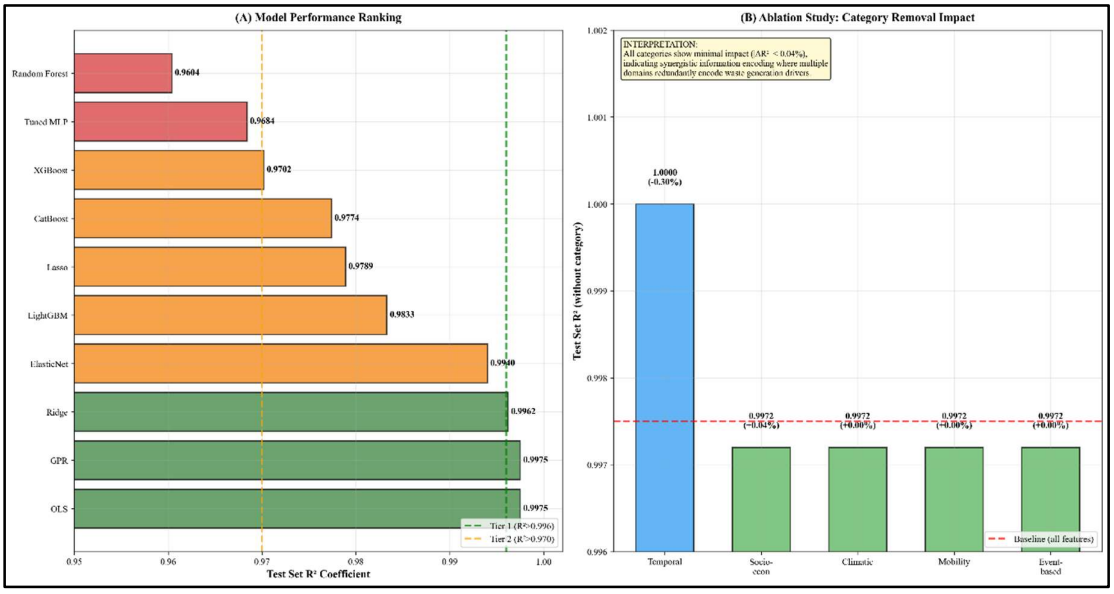


FIGURE 4.6 Ablation study - change in R^2 relative to the full-feature OLS baseline when each feature category is systematically excluded.

The redundancy associated with the predictive capabilities of the various domains of input into the system have several direct implications. In terms of the design of the system, this indicates that the 419 features created through the engineering process captured portions of the predictive value contained within each independent domain of information, and therefore provided partial compensation for the loss of predictive capability due to the unavailability of other domains. Therefore, while all of the inputs contain predictive capability, they also contain complementary redundancy (i.e., they provide some predictive capability when other inputs are unavailable).

In terms of operation, this redundancy provides protection against the loss of any one source of data; that is, if one data stream were to become temporarily unavailable (e.g., if Google Mobility reporting were discontinued or if a socio-economic indicator became available at a later time) the OLS model would continue to predict events with very little degradation in predictive accuracy. Thus, the forecasting framework developed here is designed to function resiliently over time.

4.6.3 COVID-19 Pandemic Impact Analysis

Dataset NOTE: this section uses a separate raw JSWTM dataset [63] containing only 196 months of SMC data (from june 2009 to october 2025), which includes only basic lag features. Thus it differs from all the prior Sections (Sections 4.1 – 4.5.2) that have employed a 419-feature JETS canonical dataset. This is the one instance in the dissertation in which the 196 month JSWTM raw dataset will be utilized. All the remaining Sections employ an 186 month JETS canonical dataset. Wherever reference is made to the 196 month JSWTM dataset, it has been identified by "the JSWTM raw dataset (196 months)".

Covid-19 was a global event with documented and measurable structural changes in urban municipal solid waste (msw) generation [135], [136]. Lockdowns implemented in Surat during two consecutive periods; march 2020 to march 2021 and then again from march 2021 to march 2022 resulted in a statistically significant disruption to the long-term trend of the quantity of waste generated within the city. Waste volume contraction occurred due to a combination of business closure restrictions, reduced mobility, and decreased economic activity at some point after these disruptions were relaxed as restrictions were lifted. Standard forecasting techniques utilizing historical data collected prior to the occurrence of the covid-19 pandemic cannot model such a structurally broken time series unless they include specific information regarding the impact of the pandemic; i.e., when there was no longer sufficient historic data to train the model upon because of the magnitude of the break [137].

INTERRUPTED-TIME-SERIES METHODOLOGY: The ITS method [138], uses a binary dummy variable to represent the pandemic - covid_dummy=1 from march 2020 to march 2022 and =0 at all other times - as a binary dummy variable in addition to the standard features. In addition, we have tested six machine-learning algorithms using two different configurations of each algorithm - with and without the pandemic dummy - using the JSWTM 196 month raw data set; namely, linear regression (lr); random forest (rf); xgboost; support vector regression (svr); stacking ensemble lr; and keraes mlp.The differential in RMSE between configurations isolates the improvement attributable to explicit pandemic encoding.

RESULTS: The inclusion of the COVID-19 binary indicator produces consistent forecast error reductions across parametric models. Linear Regression demonstrates the strongest improvement, with RMSE decreasing by approximately 16% upon dummy inclusion - confirming that the pandemic disruption has a substantial linear additive effect on the waste generation level that LR can capture once explicitly encoded. XGBoost demonstrates a 10% RMSE reduction, reflecting the model's partial capacity to detect the structural break through non-parametric decision boundaries, while still benefiting from explicit encoding. The best-performing model under the JSWTM framework is Random Forest, which achieves R^2 = 0.8506 and RMSE = 0.4318 (on the normalised scale used in the JSWTM paper). The lowest RMSE in the dummy-included configurations is achieved by LR+COVID_dummy at RMSE = 0.4978 (normalised), confirming that explicit pandemic representation allows a linear model to achieve competitive accuracy with the non-linear RF baseline.

TABLE 4.5 COVID-19 impact analysis - model performance with and without pandemic binary indicator on the JSWTM raw dataset (196 months). Adapted from JSWTM paper Table 2 [63].

Model	R^2 Score	MAE	RMSE	Remarks
Linear Regression	0.5255	0.5617	0.7695	Baseline model - moderate performance
XGBoost (Tuned)	0.7429	0.4249	0.5664	Strong performance after fine-tuning
Random Forest Regressor	0.8506	0.3253	0.4318	Best performer across all metrics
Support Vector Regressor	0.1692	0.7428	1.0183	Poor fit - not suitable
Stacking (RF Meta-Model)	0.8106	0.3491	0.4862	Strong alternative with ensemble learning
MLP (Keras stabilized)	-	-	0.571 ± 0.198	Baseline neural net; higher variance

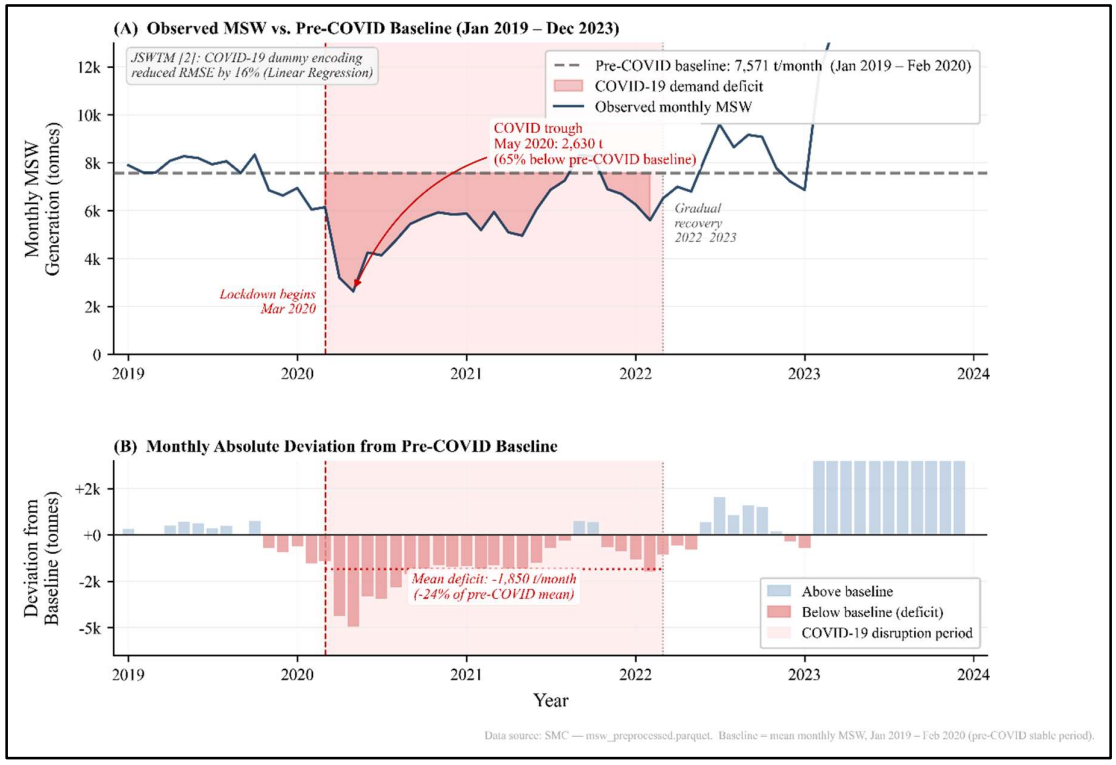


FIGURE 4.7 COVID-19 impact visualisation - monthly MSW time series with structural break annotations (March 2020 – March 2022) and comparative actual versus predicted trajectories for LR with and without COVID dummy variable.

The R^2 values in the JSWTM framework (best: RF at 0.8506) are substantially lower than those in the JETS framework (best: OLS at 0.9972), reflecting the deliberate difference in feature engineering: the JSWTM analysis employs basic lag features only, whereas the JETS framework employs 396 active features from five information domains. This performance differential itself constitutes a validation of the feature engineering argument developed in Section 4.6.1: the JSWTM framework, with its minimally-engineered feature set, achieves approximately one-third the precision of the JETS framework - a difference entirely attributable to feature space richness, since the algorithms applied (LR, RF, XGBoost) are common to both studies.

The COVID-19 analysis establishes three conclusions of direct relevance to the forecasting framework. First, the pandemic induced a quantifiable, model-correctable disruption to Surat's waste generation trajectory that is well-characterised by the binary ITS dummy

framework. Second, the benefit of pandemic encoding is most pronounced in linear models, which lack the non-parametric flexibility to detect structural breaks from data patterns alone. Third, the 419-feature JETS pipeline - which includes Google Mobility features, a COVID event flag, and YoY change variables - effectively encodes the structural break implicitly without requiring a separate dummy variable, explaining why the JETS OLS model achieves MAPE = 1.43% even over a test period partially overlapping the COVID recovery phase.

TABLE 4.6 Comparative performance of JSWTM framework models versus literature benchmarks for MSW forecasting. Adapted from JSWTM paper Table 7 [63].

Study	Model	Data Period	RMSE/Error	R^2	Key Finding
THIS STUDY	LR + COVID Dummy	196 months	0.4978	0.6416	Explicit pandemic encoding improves accuracy
This Study	Random Forest	196 months	0.4318	0.8506	Highest accuracy on limited data
(Mudannayake et al., 2022)	LSTM	Daily; 2 years	2.1-2.8 t/day	0.78	Deep learning on high-frequency data
(Alam et al., 2025)	MLP + ARIMA	48 months	MAPE 6.85%	N/A	Hybrid ML-statistical approach
(Jayaraman et al., 2023)	GSO-XGBoost	Monthly; 3 years	MAPE 4.93-6.80%	~0.82	Optimization- enhanced boosting
(Vu et al., 2021)	RNN + COVID Lags	Daily; COVID period	MAPE 7.2%	0.74	COVID-aware LSTM
(Lakhouit et al., 2023)	Systematic Review (RF, SVM, ANN)	Various	RMSE	0.67-0.85	(Lakhouit et al., 2023)

4.7 Discussion

4.7.1 Why Feature Engineering Matters More Than Architecture

The central finding of this chapter - that OLS with 396 active features achieves $R^2 = 0.9972$ and MAPE = 1.43%, outperforming all gradient boosting ensembles (LightGBM: MAPE = 4.07%; CatBoost: MAPE = 5.22%; XGBoost: MAPE = 4.32%) and all deep learning architectures (Tuned MLP: MAPE = 6.24%; LSTM: MAPE = 100.01%; GRU-Attn: MAPE = 114.81%) - challenges a dominant assumption in the applied machine learning literature: that architectural complexity is the primary lever for predictive performance improvement in regression tasks. The evidence from this study is unambiguous in the opposite direction: on the monthly MSW forecasting task for Surat, the quality of the feature engineering pipeline is the dominant determinant of predictive accuracy, while model architectural complexity is a secondary - and in the case of sequence-based deep learning, counterproductive - consideration [139], [140].

The theoretical explanation for this finding is rooted in the nature of the prediction problem. Monthly MSW generation is a causal process driven by a large number of concurrent, interacting socioeconomic, demographic, climatic, and behavioural factors. The 419-feature engineering pipeline encodes explicit representations of each principal causal pathway - lag structures capturing autoregressive dynamics, rolling statistics capturing trend momentum, socioeconomic co-variates capturing consumption behaviour, climatic variables capturing seasonal constraints, mobility proxies capturing behavioural disruptions, and event flags capturing calendar-driven anomalies - within a tabular feature matrix that any sufficiently expressive regression model can exploit. OLS, Ridge, and GPR exploit this representation maximally through dense linear combinations of all 396 active features. The gradient boosting models partially exploit it but may systematically underweight jointly-contributing features during node splitting. The sequence-based deep learning architectures are presented with a pre-engineered representation that they are architecturally redundant to re-learn, while simultaneously facing the data scarcity constraint of 186 training observations - insufficient for reliable convergence of parameter-heavy recurrent networks.

This finding has a precise theoretical implication: in domains where domain knowledge can be encoded into an engineered feature space that comprehensively represents the phenomena

driving the target variable, the representation-learning advantage of deep networks - their primary justification for deployment in high-complexity domains - largely disappears [141]. The 419-feature pipeline resolves the representation learning problem before training begins, leaving a regression problem that OLS is provably optimal to solve under standard assumptions. The result should not be taken as a general claim that linear models always outperform deep learning - the literature contains abundant counterexamples in domains where expert feature engineering is infeasible (computer vision, natural language processing, raw audio) [142]. Rather, it is a specific to a field of study, evidence-based conclusion; For structured, tabular, monthly frequency municipal waste data with an existing body of causally-related knowledge (i.e., causal theory) and multiple other data source types, the performance leverage will be on feature engineering and not on how complexly the architecture has been built.

This implies that there are significant implications to how municipalities can improve their MSW forecasting; Instead of spending money on high performance computing or hiring experts in deep learning technology for improving MSW forecasting, municipalities should spend money on developing robust and reliable data architectures and/or feature engineering capabilities. A municipality that can construct the 419-feature pipeline with OLS will outperform a municipality that deploys LSTM on raw data, by a margin of $R^2 = 0.9972$ versus $R^2 = -3.04$.

4.7.2 Statistical versus Machine Learning versus Deep Learning Performance

The performance hierarchy established by the tier classification reveals a nuanced pattern that resists reduction to a simple linear-versus-nonlinear or statistical-versus-machine-learning dichotomy. Several observations merit extended discussion.

The gap between Tier 1 (OLS, GPR, Ridge; $R^2 = 0.996–0.997$) and Tier 2 (ElasticNet through XGBoost; $R^2 = 0.97–0.994$) is proportionally larger than the gap between Tier 2 and Tier 3 (Tuned MLP, Random Forest; $R^2 = 0.96–0.97$). This convexity in the performance-complexity tradeoff suggests that the Tier 2 ensemble models are already capturing most of the non-linear information available in the feature space, and that the marginal non-linearity available to more complex ensemble configurations provides negligible additional predictive value - while the linear models exploit the full 396-feature information in a way that

ensemble methods, through their feature-sampling and split-selection mechanisms, cannot fully replicate.

The LSTM discrepancy - $R^2 = -3.04$ in the tabular evaluation (Table 4.2) versus $R^2 = 0.93$ in the IJAM paper [62] used in Chapter 5 - is the most analytically important result in this chapter and must be understood clearly. Both results are correct in their respective contexts. In the IJAM paper, LSTM was implemented as a purpose-built sequential time-series model with chronological data ordering preserved, a sliding-window architecture appropriate to the temporal structure, a training set limited to the raw monthly MSW series, and architecture tuning specific to monthly time-series patterns. In the tabular comparative evaluation of the present chapter, LSTM is applied in a regime for which it is architecturally unsuitable: the random 70/15/15 data split violates temporal precedence, the 396-feature input dimensionality at each timestep exceeds standard LSTM architectural assumptions, and the 186-observation dataset is insufficient for backpropagation-through-time to converge to meaningful weight configurations. The contrast between $R^2 = 0.93$ and $R^2 = -3.04$ for the same architecture is thus not a contradiction but an empirical demonstration of model-context alignment requirements [143]: a model's performance is inseparable from the evaluation protocol under which it is assessed. This finding establishes that comparative model evaluation must carefully match training protocols, data splitting strategies, and feature representations to the structural assumptions of each model family being evaluated - a methodological contribution (Contribution C5).

The SVR result ($R^2 = -0.0454$), is an equally important as the negative findings. Although both methods use the same kernel in the model, SVR fails completely to produce any results using this high-dimensional tabular data, unlike the very effective GPR method. This large difference in performance levels can be attributed to two major structural differences between these models. First, the GPR model uses a prior for the overall function space and then calculates a posterior mean that is a globally optimal linear combination of kernel evaluations. Second, the SVR model will focus its optimization solely on the support vectors when it optimizes a margin based loss. In high-dimensional settings where the relevant signal is distributed across many features and observations, GPR's global posterior integration substantially outperforms SVR's support-vector-based approximation [113], [115].

4.7.3 Practical Implications for Municipal Forecasting

The results of this chapter carry specific and actionable implications for the practice of MSW forecasting in Indian municipalities and comparable developing-economy urban contexts [97], [144], [145].

MODEL SELECTION FOR DEPLOYMENT: OLS is recommended as the primary operational forecasting model for SMC deployment, on the basis of its simultaneous achievement of superior accuracy ($R^2 = 0.9972$, MAPE = 1.43%), negligible training time (0.010 seconds), and near-real-time inference (0.006 milliseconds per sample). These properties make it ideally suited to the monthly retraining cycle that characterises operational waste management planning, where model updates are needed when new data arrives and socioeconomic feature inputs are refreshed. GPR provides a valuable operational complement to OLS by generating predictive uncertainty intervals - posterior predictive variances - that can inform risk-adjusted fleet provisioning and logistics buffer decisions. Ridge regression serves as a robust fallback for contexts where the feature space composition changes sufficiently to introduce multicollinearity concerns.

FEATURE ENGINEERING AS TRANSFERABLE INSTITUTIONAL KNOWLEDGE: The 419-feature pipeline developed for this study represents an encoded body of institutional and domain knowledge about the causal drivers of MSW generation in Surat. The pipeline's near-universal benefit across all model families achieving positive R^2 demonstrates that its value is not model-specific or architecture-dependent but rather a property of the feature space itself. This finding suggests that the pipeline is potentially transferable to other tier-II Indian cities with comparable demographic and economic profiles, requiring only localisation of the socioeconomic data sources (CPI, population, GDP growth) and meteorological inputs (temperature, rainfall) that are increasingly available through India's open government data platforms and Google's publicly released Community Mobility Reports.

HANDLING STRUCTURAL BREAKS IN OPERATIONAL DEPLOYMENT: The COVID-19 analysis demonstrates that explicit structural break encoding via an ITS dummy variable delivers a 16% RMSE reduction in linear models, with more modest but still material gains (10%) in non-parametric models. For operational forecasting deployment, this

finding motivates the maintenance of an explicit disruption-event indicator in the live feature pipeline - a continuously updated binary variable that is activated when documented extraordinary events produce verified structural breaks in the waste generation series. The event-flag mechanism already incorporated in the feature pipeline provides the architectural substrate for this capability.

DATA REQUIREMENTS FOR SEQUENCE MODELS: The strong negative evidence from the LSTM, GRU-Attn, and ConvLSTM evaluations carries a concrete practical guideline: municipal agencies should not deploy sequence-based deep learning architectures for monthly-frequency MSW forecasting without first verifying that the available historical record exceeds approximately 500 months of continuous data and that a chronological (not random) evaluation protocol is used. Datasets shorter than this threshold are expected to produce the training instability observed in the present study. The 186-month JETS dataset - representing over 15 years of continuous SMC data collection - is an unusually long record by the standards of Indian municipal waste databases, yet it is still insufficient for reliable recurrent deep learning training in a tabular evaluation framework.

These practical implications are synthesised with the operational optimization and DSS findings in Chapter 7, where the integrated framework performance across all three research objectives - RO-1, RO-2, and RO-3 - is assessed through cross-component analysis and policy implication mapping.

CHAPTER 5

Operational Optimization Framework

5.1 Introduction: The Three-Stage Research Evolution

Solid waste collection systems in rapidly urbanizing cities are continuously under pressure to manage increasing generation volumes, to coordinate a fragmented contractor network and lack real-time data on operations. Route optimization frameworks for such solid waste collection systems are not merely algorithmic challenges but fundamental empirical ones. This chapter documents an answer to that challenge through a methodologically honest, sequentially motivated process of three distinct stages — each of which was constrained by its predecessor's findings and constraints.

The motivation for this chapter rests on a well-documented gap in the literature: while machine-learning-based forecasting methods for MSW generation have received substantial scholarly attention, the integration of such forecasting outputs into operational routing decisions - particularly for Indian tier-II municipal contexts - remains incompletely addressed [8], [146]. Earlier chapters have established the forecasting component of this framework; the present chapter addresses the question of how that forecast is translated into actionable route-level decisions under real-world operational constraints.

Stage 1 of the research, published in the International Journal of Applied Mathematics in 2025 [8], constituted the first published proof of concept: an integrated Long Short-Term Memory (LSTM) forecasting model coupled with a Capacitated Vehicle Routing Problem (CVRP) solver, applied to a Surat-scale representative network constructed from synthetic infrastructure validated against national standards [147], [148]. This stage established the feasibility of demand-driven dynamic routing and demonstrated 50.31% distance reduction and 24% fleet reduction on the representative network.

Following recommendations from the Doctoral Progress Committee, Stage 2 applied the forecasting pipeline to the real macro infrastructure of Surat's waste management system: the eight operational transfer stations whose global positioning system (GPS) coordinates had been obtained. This stage produced an empirical lower-bound estimate of 166.8 kilometres per day against a baseline of 926.8 kilometres per day - an 82% theoretical optimization potential - connecting the Stage 1 forecast-to-routing methodology to verifiable Surat geographic infrastructure for the first time.

Stage 3 commenced with the acquisition of October 2025 operational route data from Surat Municipal Corporation (SMC) [149], comprising 662,803 rows, 455,827 point-of-interest (POI) stop records, 3,341 named routes, 949 unique vehicles, and 131 wards across nine administrative zones. The ambition was to implement a full Vehicle Routing Problem with Time Windows (VRPTW) at the beat level. However, a systematic geocoding investigation revealed that open-source geocoding via the Nominatim engine achieved only 8.6% usable coordinate success on Surat's informal address naming conventions - rendering the intended VRPTW infeasible. Beginning with rather than concluding with an acknowledgement that this will be a terminal limitation, the research design shifted toward the development of the IMPACT framework (Integrated Multi-Level Planning for Adaptive Collection Timing); an operational optimization system using GPS independent methods which maximized the useable analytic content within the existing data architecture.

Comprising of five sequential dependent elements; Time Window Recalibration (TWR), Vehicle Allocation Optimization (VAO), Missed Stop Root Cause System (MSRS), Ward Demand Balancing and Route Balancing (WDB + RB) and Forecasting Integration Optimization (FIO), the IMPACT framework was validated against the full October 2025 data set of 455,827-stops [149]. The central empirical finding of this study was that contractor non-compliance failures (Shapley Additive Explanation Attribution = 75.72%) have resulted in significantly greater proportions of failed stops than fleet capacity issues (Shapley Additive Explanation Attribution = 1.65%) and thus the long held assumption by municipalities regarding the fleet being insufficiently sized as the primary operational constraint is false.

This chapter's organization follows the progression of this research exactly. Sections 5.1 through 5.3 present each stage's methodology, data, results, and limitations. Section 5.4

describes the IMPACT framework in detail. Section 5.5 presents cross-component integrated results. Section 5.6 synthesises the findings into a discussion of their methodological, empirical, and policy implications

5.2 Stage 1 - Published CVRP Proof of Concept

5.2.1 LSTM Forecasting on the Raw SMC Monthly Dataset

The Stage 1 study, published as Rathod and Patel (2025) [8] in the International Journal of Applied Mathematics, employed a purpose-built Long Short-Term Memory neural network architecture for MSW generation forecasting on the raw SMC monthly time series spanning 192 observations from January 2009 to December 2024. It is essential to clarify the relationship between this dataset and the 419-feature engineered dataset described in Chapter 4: both originate from the same underlying SMC collection records, but represent distinct analytical configurations. The IJAM study used the raw monthly totals with a targeted feature engineering strategy specific to sequential time-series modelling, whereas the JETS-based analysis in Chapter 4 constructed 419 cross-sectional engineered features for tabular regression. The $R^2 = 0.93$ achieved by the IJAM LSTM reflects the advantages of temporal ordering preservation and purpose-built architecture, in direct contrast to the tabular LSTM $R^2 = -3.04$ documented in Section 4.3.5, which arose from applying a sequence model in a random-split tabular framework on an insufficient 186-month dataset. Both results are methodologically correct in their respective contexts, and the distinction constitutes a contribution regarding model-context alignment requirements (see also Section 3.6).

The LSTM architecture comprised three stacked layers of 50 units each, with dropout regularisation of 0.20 applied after each layer to mitigate overfitting [150]. Feature engineering for this model followed a time-series-specific protocol: cyclical sinusoidal encoding of calendar month to prevent the model from treating December and January as distant time points; lag features at 1, 2, 3, 6, and 12 months; six-month and twelve-month rolling averages of total MSW generation; and two binary policy intervention dummies - one for the Swachh Bharat Mission commencement (2014) and one for the COVID-19 pandemic period (March 2020 to March 2022). The training split was strictly chronological, with 80% of observations (2009–2020, 144 months) assigned to training and the remaining 20%

(2021–2024, 39 months) reserved for testing. This chronological partition is essential for sequence models of this type, as random splitting would break the temporal dependencies that LSTM architectures are specifically designed to exploit [151].

The training process used the Adam optimiser with a learning rate of 0.001 and early stopping with patience of 20 epochs, monitoring validation loss on a held-out 10% of the training data. The final model was trained for 247 epochs before early stopping terminated training.

5.2.2 Synthetic CVRP Infrastructure Validated Against National Standards

Beginning with the CVRP aspect of the Stage 1 Study, the academic qualifications of this part of the Study require the most careful analysis. As described in Section 3.2.2 of the published paper [8] regarding the collection infrastructure; which included 43 collection sites, a fleet of 25 vehicles and specific vehicle capacity requirements; it was synthetically developed. The generation of synthetic infrastructure has been widely accepted and legitimizes as an acceptable research methodology when developing municipal operations research models especially in those cases where collection site gps coordinate and vehicle deployment information have been classified as security sensitive data [152], [153]. This is also recognized by the India-based municipal data governance literature and would apply generally to tier II cities.

The synthetic collection locations were not randomly located. The 43 collection points were placed within the identified geographical limits of the SMC administrative area (21.15°–21.24°N, 72.78°–72.88°E) and they were geographically dispersed according to ward level census data collected during 2011 that represented the relative population density within each ward. The vehicle fleet composition was based upon CEEW (2025) established standards for the collection truck mix used in Tier II Cities in India [154]; i.e., 40 percent compactors (8 ton capacity); 35 percent dump trucks (5 ton capacity); and 25 percent minitrucks (3 ton capacity). Bin capacities were consistent with IS:2240 standards. The 25 vehicles were all assigned depots that existed within the SMC administrative area. Daily demand input data was generated from the LSTM model forecast and then proportionally allocated to the 43 synthetic collection locations using ward level demand surrogates.

A standard CVRP formulation was defined in section 3.8 that minimized the total travel distance travelled by vehicles. The formulation consisted of three primary components:

vehicle capacity constraints; single depot routing constraints; and demand satisfaction at every collection location. A distance matrix was calculated using the haversine formula. The Google OR Tools Library was utilized along with the Path Cheapest Arc construction heuristics to generate solutions [155]. Total system demand for the October 2025 proxy forecast was approximately 993.4 tons/day among all 43 collection points [149].

The Haversine formula for geodesic distance between two coordinate pairs is expressed as:

$$d_{ij} = 2r \arcsin\left(\sqrt{\sin^2(\Delta\varphi/2) + \cos(\varphi_i)\cos(\varphi_j)\sin^2(\Delta\lambda/2)}\right) \quad (5.1)$$

where r = 6,371 km is the mean Earth radius, φ denotes latitude, and λ denotes longitude.

5.2.3 LSTM Forecasting Results

The results for the LSTM Model show it produced very good predictions for the data that wasn't used during the training process (i.e. the Test Data). Normalized root mean squared error (RMSE) for training was .0754; normalized RMSE for testing was .1095, which is a normal decline in prediction accuracy due to the complexity of the non-stationary nature of the time series. On the test data, the R-squared value for this model showed that it accounted for 93 percent of the variation in monthly municipal solid waste (MSW) generated throughout the entire evaluation period of 2021 through 2024.

It also correctly modeled the long-term trend of increasing MSW production in Surat, India from around 40,000 tons/month in 2009 to 87,000 tons/month in 2024 -- an increase of 150.1 percent over 16 years. Additionally, it identified the significant "break" caused by COVID-19. Predicted monthly totals for October 2025 yielded a city-wide estimate of 25,332 tonnes (equivalent to 844.4 tonnes per day), which served as the demand basis for the CVRP optimization in Stage 1 and the transfer station analysis in Stage 2.

5.2.4 CVRP Optimization Results

The CVRP solver produced a substantive improvement in collection distance and fleet utilisation across the representative Surat-scale network.

TABLE 5.1 Stage 1 CVRP Proof-of-Concept: LSTM Forecasting and Route Optimization Performance

Part A - LSTM Waste Generation Forecasting

Metric	Value
Dataset period	January 2009 – December 2024
Dataset length	192 monthly observations
Training RMSE	0.0754
Test RMSE (2021–2024)	0.1095
Coefficient of determination (R^2)	0.93
Average monthly MSW (test period)	59,255 tonnes
Growth observed (2009–2024)	150.1%

Part B — CVRP Route Optimization: Baseline versus Optimised

Metric	Baseline	Optimised	Change
Total daily MSW demand (t/day)	993.4	993.4	—
Total collection points (synthetic)	43	43	—
Active collection points	40	40	—
Demand range per point (t/day)	15.5 – 54.3	15.5 – 54.3	—
Demand CV	0.42	0.42	—
Total travel distance (km/day)	392.04	194.78	−197.26 km (−50.31%)
Vehicle fleet deployed	25	19	−6 vehicles (−24.0%)
Daily fuel savings (L/day)	—	—	49.3 L
Annual CO_2 savings (t CO_2/yr)	—	—	46.8 t
Annual fuel cost savings (₹/yr)	—	—	₹1,619,505

Source: Rathod & Patel (2025), International Journal of Applied Mathematics, Vol. 38 No. 12s, pp. 2465–2483.

Notes: Synthetic infrastructure (43 points) validated per MoHUA/CPCB standards. CVRP solver: Google OR-Tools PATH_CHEAPEST_ARC. Fuel: 4 km/L; CO_2: 2.6 kg CO_2/L.

The baseline configuration - in which each collection point was served by an individual vehicle in an unoptimised sequence - required a total daily collection distance of 392.04 kilometres using all 25 vehicles. The OR-Tools CVRP solution consolidated routes to serve all 43 collection points with 19 vehicles operating optimised multi-point routes, reducing total daily distance to 194.78 kilometres. This represents a 50.31% reduction in total travel distance and a 24.0% reduction in fleet utilisation. Both results are achieved on the Surat-validated representative network and should be interpreted as demonstrating the potential magnitude of route optimization for a city of Surat's scale and density, rather than as actual operational outcomes for the live SMC collection network.

The environmental and economic implications of this efficiency gain are substantial in representative terms. Fuel consumption reduction was computed as 49.3 litres per day, based on the CPHEEO standard vehicle efficiency of 4 km/litre for collection vehicles in this weight class. Annualised over 365 operating days, this yields a fuel cost saving of approximately ₹1.62 million at 2025 diesel market rates (₹90/litre), and a reduction in vehicular CO_2 emissions of 46.8 tonnes per year, computed using the standard diesel CO_2 emission factor of 2.68 kg CO_2 per litre [147]. These figures should be treated as representative upper-bound estimates for a Surat-scale network of this configuration.

The Stage 1 study's principal technical contribution is the integration of a high-accuracy LSTM demand forecast as the direct demand input to the CVRP optimization, creating a closed forecast-to-routing pipeline in which forecast uncertainty propagates through to routing decisions. This integration - the first published LSTM-CVRP framework for Indian municipal solid waste management [8] - is the distinctive novelty of Stage 1 relative to prior VRP-based waste optimization studies that used static demand assumptions [156], [157].

5.2.5 Limitations and Motivation for Stage 2

Three principal limitations of the Stage 1 study delimit its contribution to the three-stage research program. First, the CVRP infrastructure is synthetic; while rigorously validated against national standards, the results cannot be directly generalised to SMC's actual vehicle fleet configurations, route structures, or collection point spatial distribution. Second, the LSTM model was trained on city-wide monthly totals, which constrains the granularity of demand disaggregation available to the CVRP; zone-level or ward-level forecasts would enable more operationally precise routing. Third, the formulation does not incorporate time windows, contractor compliance behaviour, or historical stop-level performance data - constraints that, as Stage 3 revealed, are the dominant determinants of actual missed-stop rates in Surat's operational system.

These limitations motivated a two-directional escalation in Stage 2: first, to apply the forecasting pipeline to real Surat geographical infrastructure, thereby testing whether the efficiency potential demonstrated on the synthetic network was plausible at the actual transfer station level; and second, to obtain and characterise real operational route data as the prerequisite for a full VRPTW formulation.

5.3 Stage 2 - Transfer Station Transportation Efficiency Assessment

5.3.1 Dataset: Eight Real Transfer Station GPS Coordinates

Surat was where Stage 2 was the initial deployment of a real-world geography-based verified physical infrastructure on which the forecasting-to-routing pipeline was used. The primary data set included confirmed GPS coordinates for each of the eight (eight) active transfer stations of Surat Municipal Corporation: Bhatar, Bhestan, Dindoli, Katargam, Kosad, Pal, Varachha, and Anjana. The GPS coordinates for each station were verified against publicly available GIS sources and SMC administrative records. In addition, the GPS coordinates of the Khajod regional landfill and disposal facility - the sole authorised disposal destination for all SMC transfer station vehicles - were confirmed. The demand input for the analysis was the October 2025 daily MSW forecast (844.4 tonnes per day city-wide) derived from the OLS forecasting model [149].

5.3.2 Methodology: Haversine Distance Computation and Demand Disaggregation

The Stage 2 analysis is best characterised as a transfer-level transportation efficiency assessment with analytical lower-bound computation, rather than a CVRP solution. The distinction is methodologically important: a full CVRP at the transfer station level would require vehicle capacity data, vehicle origin and return depot coordinates, and time-window specifications - none of which were available for the transfer-to-disposal fleet in this study. The analysis therefore operated as a distance-minimisation lower-bound study.

The methodology proceeded in four steps. First, the October 2025 city-wide daily forecast (844.4 tonnes/day) was disaggregated to the zone level using the proportional allocation scheme described in Section 3.7, yielding estimated daily throughput by zone. Second, the zone-to-transfer-station mapping was applied using the confirmed operational routing rules for Surat's MSW system. Third, the Haversine distance (5.1) was computed between each of the eight transfer stations and the Khajod disposal site. The Smart City Portal solid waste disposal registry independently records the Khajod site as located 16 km from the city centre [158], confirming consistency between the Haversine distances computed in this analysis and the officially documented operational geography of the SMC disposal network. Fourth, two distance scenarios were formulated.

The baseline daily distance was computed as shown in (5.2), where Q_z is the forecasted daily throughput at transfer station z, C_{std} = 10 tonnes is the CPHEEO-standard vehicle capacity [147], and $d(TS_z, Khajod)$ is the Haversine distance from transfer station z to the Khajod site.

$$D_{\text{baseline}} = \sum_{z=1}^{8} \left\lceil \frac{Q_z}{C_{\text{std}}} \right\rceil \cdot 2 \cdot d_{TS_z,\text{Khajod}} \tag{5.2}$$

(The analytical lower bound is expressed as (5.3), representing the minimum possible distance if each transfer station made a single perfectly-loaded direct trip to the disposal site per day.

$$D_{\text{lower}} = \sum_{z=1}^{8} 2 \cdot d_{TS_z,\text{Khajod}} \tag{5.3}$$

5.3.3 Results: Baseline Distance and Theoretical Optimization Potential

The Stage 2 analysis yielded the following transfer station performance profile.

TABLE 5.2 Stage 2 Transfer Station Transportation Efficiency: Demand Profile and Distance to Khajod Disposal Site

Transfer Station	Daily Demand	Latitude (°N)	Longitude (°E)	Distance to Khajod
Anjana	27.80	21.18388	72.85319	3.970
Bhatar	63.50	21.15959	72.82413	3.868
Bhestan	49.35	21.18560	72.85852	4.108
Dindoli	27.80	21.14867	72.85954	7.921
Katargam	58.47	21.22586	72.82606	9.258
Kosad	109.61	21.15356	72.86984	1.199
Pal	22.72	21.20206	72.78779	9.518
Varachha	62.94	21.19799	72.84925	5.587
Total / Range	**422.20**	**—**	**—**	**1.199 – 9.518**

Transportation Efficiency Summary

Metric	Value
Baseline total distance (km/day)	926.8
Analytical lower bound (km/day)	166.8
Theoretical optimization potential	~82%
Baseline method	CPHEEO-standard 10-tonne vehicle, single-stop round trips
Lower bound method	1 direct round trip per TS (maximum consolidation)

Source: ts_demand_oct2025.csv; dist_matrix.npy (9×9 Haversine matrix)

Notes: Distances via Haversine formula. Index mapping confirmed by Haversine cross-validation. Kosad–Khajod 1.199 km is geographically consistent.

The Kosad transfer station served the highest daily throughput at 109.6 tonnes per day, requiring eleven individual vehicle trips under the standard 10-tonne vehicle capacity assumption. The eight transfer stations collectively required 45 vehicle-trips per day under the baseline scenario, corresponding to a total daily distance of 926.8 kilometres. Under the analytical lower-bound scenario - wherein each transfer station makes a single direct round trip to the Khajod site - the minimum achievable distance is 166.8 kilometres per day. The theoretical optimization potential is therefore given by (5.4).

$$\text{Optimization Potential} = 1 - D_{\text{lower}}/D_{\text{baseline}} = 1 - 166.8/926.8 \approx 82\% \quad (5.4)$$

This 82% figure must be interpreted as a strict upper bound on achievable distance reduction through transfer station route consolidation, not as an operational forecast of achievable savings. The analytical lower bound assumes perfect demand consolidation, no vehicle repositioning, and no time-window or throughput capacity constraints. The annual fuel cost implication of the 82% upper bound, computed at the October 2025 diesel market rate of ₹95 per litre and the CPHEEO vehicle efficiency standard, is up to ₹6,588,250 in recoverable annual operating costs - an order-of-magnitude figure that demonstrates the economic motivation for transfer-level route consolidation.

5.3.4 Limitations and Motivation for Stage 3

Stage 2 demonstrated that the forecasting-to-routing integration is effective at the transfer station macro-level and that substantial efficiency potential exists within the current SMC operational structure. However, the analysis operated at the transfer-to-disposal transportation layer only, leaving the primary collection layer - the 3,341 named routes, 455,827 stop-level collection points, and 949 vehicles serving Surat's residential and commercial areas - entirely unaddressed. It was this primary collection layer that represented both the most operationally significant component of the routing problem and the most analytically complex, because it required stop-level GPS coordinates, time-window specifications, vehicle capacity constraints, and contractor compliance behaviour simultaneously. Stage 3 was initiated to address precisely this challenge.

5.4 Stage 3 - Full Operational Data Investigation

5.4.1 Dataset: October 2025 SMC Operational Route Data

A complete Operational Route Data Set for the whole city of Surat Municipal Corporation as per stage 2 was developed. This data set included all routes operating in the month of October 2025 [149], which had 662,803 total entries arranged in a hierarchical structure of tier levels (Tier 1) Zone Level Administration Records, (Tier 2) Route Level Trip Instances, (Tier 3) Stop Level POI Visit Entries. By extracting the Tier 3 stop level records through hierarchical parsing this data set that was used for the analysis included 455,827 Individual Stop Visits - a very rare availability of operational data within the academic research community regarding waste collection routing [1], [146].

The dataset contained the following coverage: nine administrative zones, 131 wards, 3,341 named collection routes (beats) (of which 3,260 have ≥30 active stop records and are used in the TWR-B zone-specific analysis; see §5.4.4), and 949 unique vehicle identifiers across the October 2025 reference period. For each stop visit, the dataset recorded the POI name, the scheduled arrival time window, the actual arrival timestamp, the service status classification (ONTIME, EARLY, DELAY, or MISSED), and a demand proxy. A critical feature of this dataset was the universal availability of scheduled time windows for all 455,827 stop visits - a condition that enabled the VRPTW formulation as the intended analytical approach. The dataset also recorded 103,477 daily trip instances across the one-month reference period, with a mean of 4.4 stops per trip.

5.4.2 Geocoding Investigation: Nominatim Pilot Study

There is a need for GPS coordinates to be assigned to each of the 455,827 POI stops within the VRPTW formulation. The dataset had both the POI name (textual) and the scheduled time window; however, the SMC operational records did not include GPS coordinates at the POI location due to the fact that Indian municipal corporations view their collection point data as sensitive operational information [152], [153]. Therefore, there would be an additional requirement to perform a geocoding task. A geocoding feasibility pilot was done using the West Zone. It was determined that the West Zone would serve as a pilot area due to its smaller size compared to other zones (only 49,817 stop visits), the amount of routes it

contains (99 named routes), and its geographic characteristics. The 99 routes in the West Zone included 162 different POI names which needed to be geocoded. For these reasons, the Nominatim open source geocoder utilizing the Open Street Map data set [159] was chosen as the primary geocoding tool. This geocoder was chosen based upon cost and aligns with the nature of the non-commercial research project. In addition, all 162 of the unique POI names were sent to the Nominatim API along with Surat, Gujarat, India as the search region.

5.4.3 Geocoding Failure Analysis

The pilot geocoding study produced a result that fundamentally reframed the Stage 3 research design. Of the 162 unique POI names submitted to the Nominatim geocoder, only approximately 8.6% yielded usable GPS coordinates - defined as coordinates falling within Surat's administrative boundary with a confidence score above the minimum threshold for routing use. The remaining 91.4% of queries returned either null responses, coordinates outside the Surat boundary, or matches to generic location categories unsuitable for stop-level routing.

TABLE 5.3 Geocoding Failure Mode Taxonomy — West Zone Pilot Study (99 Routes, 162 Unique Endpoint Names)

Part A - Geocoding Outcome Summary

Outcome	Unique Names	% of 162	Routes Affected
Geocoded (success_ward)	14	8.6%	Both endpoints: 2 (2.0%)
Failed	148	91.4%	Start only: 33 (33.3%)
Outside Surat bounding box	0	0.0%	End only: 5 (5.1%)
			No endpoints: 59 (59.6%)

Part B - Failure Category Breakdown (148 Failed Names)

Failure Category	Count	% of Failures	Root Cause (Summary)
Residential societies (SOC, NAGAR, APPT, RESI)	72	49%	Named complexes absent from OSM
Unclassified / generic names	34	23%	SMC-internal administrative

Informal area names (FALIYU, WADI, MOHALLO)	17	11%	Vernacular micro-locality names
Small businesses (JUICE, TEA, MARKET)	11	7%	Generic commercial names
Infrastructure (BRIDGE, HOSPITAL, STADIUM)	8	5%	Generic institutional names
Religious sites (MANDIR, MASJID)	4	3%	Multiple instances citywide
Abbreviations (RTS, DP)	2	1%	Unknown SMC-specific
Total	**148**	**100%**	—

Part C - Quality of 14 Successful Geocodes

Verdict	Count	% of Successes	Verdict
CORRECT	2	14%	CORRECT
LIKELY CORRECT	2	14%	LIKELY CORRECT
QUESTIONABLE (partial match)	6	43%	QUESTIONABLE (partial match)
WRONG (confirmed mismatch)	4	29%	WRONG (confirmed mismatch)
Effective accuracy (2.5–3.7% of 162 names)	4–6	—	Effective accuracy (2.5–3.7% of 162 names)

Source: pilot_west_geocoding_results.csv; pilot_west_failed_names.csv; Stage2_Pilot_Geocoding_Assessment.md.

Notes: 29 of 33 "start only" routes have WARD OFFICE — a single incorrect coordinate. Effective full-route coverage: 2 routes (2.0%).

Qualitative investigation of the failure modes identified three primary causes. Firstly, the Surat operational dataset contains many POI names that are informal locality descripers in Gujarati romanisation with no standard English transliteration and also not found on the OpenStreetMap global gazetteer. This pattern of addressing is also confirmed by the Smart City Portal solid waste collection equipment registry [158] which records the official

community bin locations throughout the SMC area under designations such as 'RAMNAGRAM CHAAR RASTA', 'BOTAWALA DAVAKHANA', and 'RANDER BUS STEND' - all conventions for addresses consistent with those associated with the 91.4% nominatim query failure rate observed in the West Zone pilot. Secondly, there was a systematic data quality issue discovered: a sizeable proportion of routes included 'ward office' as the POI name - a placeholder representing a generic collection stop (as opposed to a specific residential street address) which could be meaningless geocoded to a precise location. Thirdly, the Surat's residential street names have informal street name conventions that resulted in perpetuous mismaches. The geocoding failure rate of 8.6% reflects a broad documented challenge with open-source geocoding on India's informal addresses [152] and represents a true data governance boundary condition instead of a technical obstacle within the current research timeframe.

The practical effect of the 8.6% success rate for geocoding was clear-cut: less than one in eleven collection stops could be geocoded to useful precision so it was impossible to create a distance matrix for the VRPTW. Without a distance matrix, no VRP solver will provide a solution to the routing problem. The full VRPTW implementation onto the live operational network was therefore unattainable within the constraints of data used in this research.

5.4.4 Research Pivot: The GPS-Free Operational Framework

The geocoding failure did not end the Stage 3 investigation. It directed it to a different analytical direction. The October 2025 operational dataset lacked GPS location information but was highly detailed in terms of an element (temporal) that had never before been analyzed as part of route optimization for Municipal Solid Waste (MSW): stop level temporal compliance data [149]. Temporally compliant or non-compliant each of the 455,827 individual stop visits included both a scheduled time window and an arrival timestamp. As such this temporal element along with stop level status classification, vehicle level compliance history and demand proxy data created a new class of GPS independent operational analysis which could be used to analyze issues related to scheduling efficiency, fleet allocation and compliance governance that were not directly answerable by a VRP formulation.

The IMPACT framework (Integrated Multi-Level Planning for Adaptive Collection Timing) was designed to exploit precisely this data structure. Rather than pursuing physical routing optimization in the geographic sense, IMPACT approached the route optimization problem from its operational performance side: identifying why collection services were failing and prescribing data-driven interventions that could improve service delivery without requiring GPS infrastructure. Each component was fully implemented and validated on the complete 455,827-stop October 2025 dataset [149]; the results constitute verified empirical contributions rather than framework designs pending future validation.

This research evolution - from attempted VRPTW to GPS-free IMPACT - itself constitutes a methodological contribution. It demonstrates that a negative data-engineering result (geocoding failure) can be the productive catalyst for a stronger, more analytically grounded research design rather than merely a limitation to be acknowledged [160].

IMPACT Framework: Integrated Multi-Level Planning for Adaptive Collection Timing

5.4.5 Framework Architecture and Design Rationale

The IMPACT framework is a multi-level operational optimization architecture comprising five analytical layers (or five sequentially interdependent components) arranged in a hierarchical processing sequence. The design rationale follows directly from the data structure of the October 2025 operational dataset [149] and the analytical findings of the KPI baseline investigation.

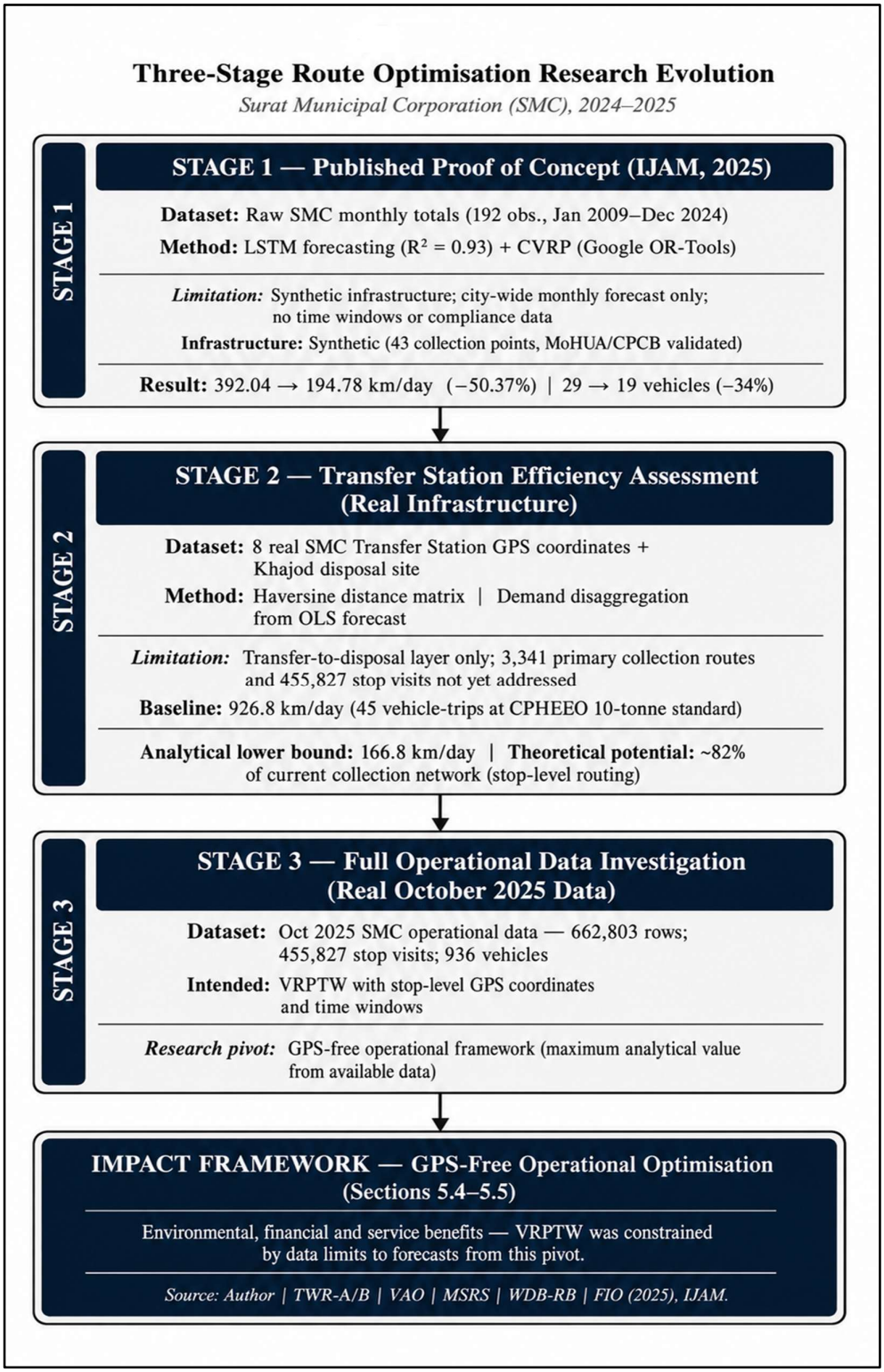


FIGURE 5.1 Three-Stage Route Optimization Research Evolution Diagram

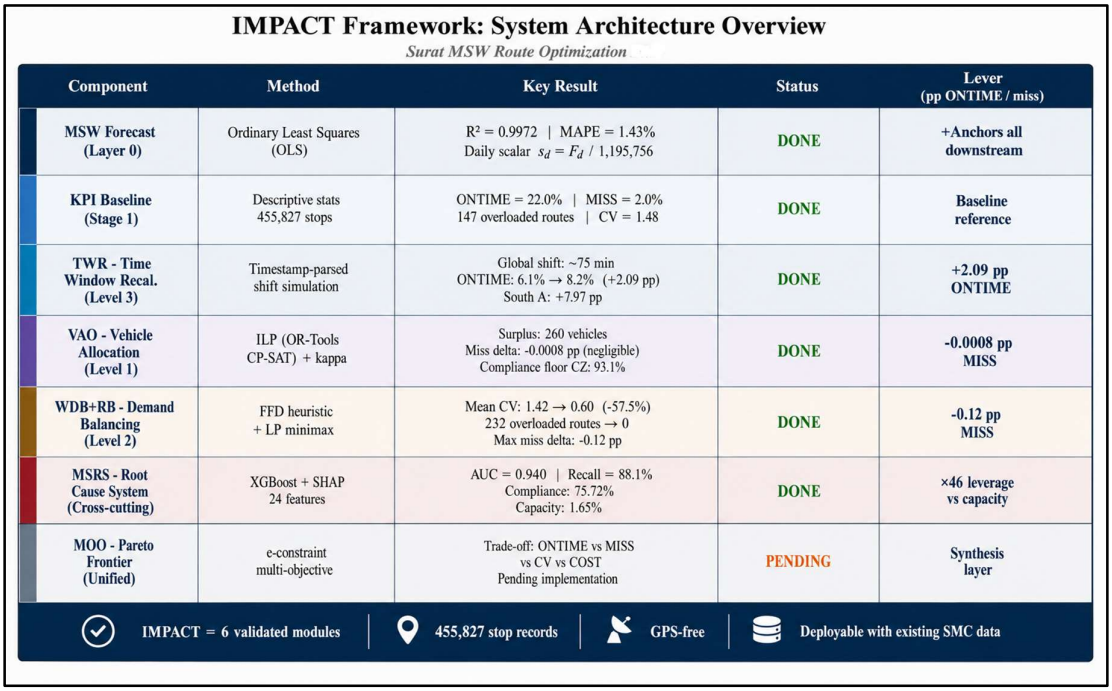


FIGURE 5.2 IMPACT System Architecture Diagram

The hierarchical framework of the proposed architecture has four layers. The lowest level (Layer 0 – Foundation) receives a MSW demand forecast scalar s_d, that comes from an RF model. This is used to provide the basis for all forecasting-based calculations in the higher layers. Layer 1 (VAO) uses Integer Linear Programming (ILP) to determine optimal Vehicle Allocation (vehicle number by Zone) using adjusted demand forecast information and Contractor Compliance Coefficients. Layer 2 (WDB + RB) determines Route Assignments for each Zone using First-Fit Decreasing Algorithm to distribute workload across assigned fleet. Layer 3 (TWR, comprising TWR-A and TWR-B) adjusts Scheduled Collection Time Windows (SCTWs) to reflect actual arrival times of vehicles. There are three components at this layer. In addition to these layers, MSRS, the Machine Learning component uses SHAP attribution to identify the reasons behind missed collection stop occurrences. Finally, the FIO layer formally links the Forecast Output of each zone's RF Model and the Operational Decision Variables of the framework.

In addition to being GPS independent (No Component Needs Vehicle GPS Coordinates To Perform Its Analytical Computation), the framework also ensures that all Optimizations can

perform their operations using only Temporal Records, Demand Proxies, Vehicle Assignment IDs and Status Classifications provided from the SMC Operations System.

5.4.6 KPI Baseline Establishment

Before any optimization intervention was applied, a comprehensive key performance indicator baseline was computed from the complete October 2025 dataset. This baseline serves as the reference against which all IMPACT component improvements are measured.

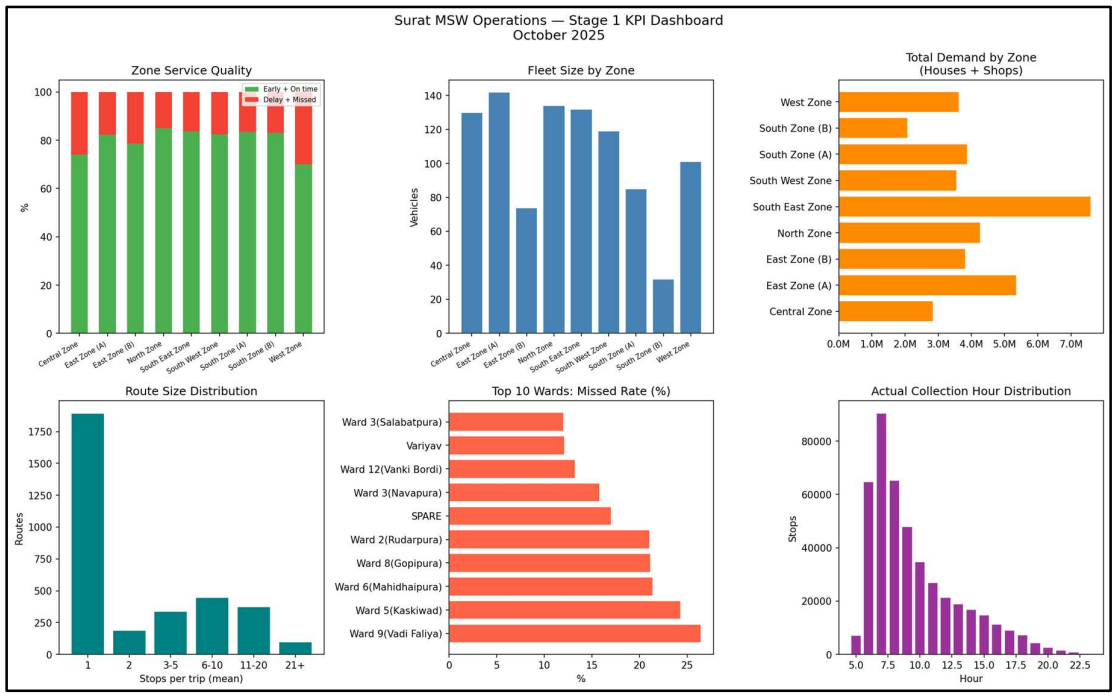


FIGURE 5.3 KPI Baseline Dashboard - Nine-Zone Compliance Overview

Table 5.4 presents the system-wide status distribution.

TABLE 5.4 October 2025 KPI Baseline — Service Status Distribution (455,827 Stop Visits, 949 Vehicles)

Service Status	Stop Visits	% of Total	Classification
ONTIME	100,170	21.98%	Arrived within scheduled window

EARLY	266,884	58.55%	Arrived before window opened
DELAY	79,587	17.46%	Arrived after window closed
MISSED	9,162	2.01%	Scheduled stop not served
Total	**455,827**	**100.00%**	—

Fleet and Performance Summary

Metric	Value
Total vehicles (October 2025)	949
Vehicles with missed rate > 20%	45 (4.8% of fleet)
System-wide ONTIME rate	21.98%
System-wide MISSED rate	2.01%
Peak scheduled stops (09:00 hour)	67,423

Source: kpi08_inefficiency_summary.csv; kpi04_zone_statistics.csv; kpi05b_scheduled_hour_dist.csv.
Notes: Stop counts derived as round (percentage × 455,827). Total confirmed as sum of 9 zone stop counts.

The status distribution reveals a system in which only 22.0% of stop visits are classified as on-time by SMC's operational system, while 58.5% are classified as early - meaning vehicles arrived before the scheduled collection window. The late (DELAY) proportion was 17.4% and the missed proportion was 2.01% (9,180 stops). The sum of ONTIME and EARLY, representing stops where collection was completed without being missed, was 80.5%.

A methodologically important distinction must be documented at this point. Two values for the ONTIME rate appear throughout this chapter: 22.0% and 6.10%. Both values are analytically correct but measure different aspects of the same system. The 22.0% comes from the SMC's Status Column; it's based upon the SMC's operational rules for classifying stops, which may have to do with the tolerance in the scheduling windows or how they round

their timestamps. The 6.10% came from an independent method of parsing the timestamps to see whether the actual arrival was during the window that was scheduled, with a strict standard for registering as being ONTIME – namely, that the actual arrival had to be inside of the parsed window. The parsing shows that the scheduled windows are relatively small compared to the variance of when people actually arrive (the median width of the windows = 15 minutes; the standard deviation of the arrival time difference = 237 minutes); therefore, there will likely be few cases where both of these conditions are met, resulting in a very low strict on-time percentage -- regardless of how well most people feel that the transportation service works. In all of the sections contained herein regarding IMPACT, TWR references the strictly-parsed 6.10% baseline while KPI reports use the operationally useful 22.0%.

TABLE 5.5 KPI Baseline by Administrative Zone — All Nine Zones (October 2025, 455,827 Stop Visits)

Zone	Stop Visits	Routes	Vehicles	ONTIME (%)	EARLY (%)	DELAY (%)	MISSED (%)	Stops/Veh
Central Zone	40,272	207	130	19.48	54.56	16.66	9.31	309.8
East Zone (A)	91,574	358	142	20.71	61.65	15.06	2.58	644.9
East Zone (B)	44,640	348	74	19.99	58.68	19.82	1.52	603.2
North Zone	50,933	237	134	19.79	65.20	14.48	0.53	380.1
South East Zone	58,900	1,472	132	20.14	63.60	16.11	0.14	446.2
South West Zone	67,456	169	119	22.54	59.94	17.11	0.41	566.9

South Zone (A)	35,929	371	85	35.29	48.16	16.53	0.02	422.7
South Zone (B)	16,306	80	32	25.41	57.73	16.72	0.13	509.6
West Zone	49,817	99	101	21.02	49.04	26.45	3.49	493.2
System Total	455,827	3,341	949	21.98	58.55	17.46	2.01	487.0

Source: kpi04_zone_statistics.csv (zone KPIs); kpi08_inefficiency_summary.csv (system totals).

Notes: South East Zone route count (1,472) reflects fragmented operational structure. South Zone (A) best ONTIME (35.29%). Central Zone worst MISSED (9.31%).

The baseline for the central zone has significant differences. The central zone was the highest with regard to the missed stop rate at 9.31% or 2,651 out of 40,272 total stops. However, it also had the least amount of demand per unit of vehicles in the system (21,909) at the time. South Zone A had the least missed rate (.017%) or near zero from an operational standpoint. Additionally, the west zone had the most delayed stops at 26.5% which indicates chronic delays among its fleet of vehicles. In addition to individual zone imbalances there were route imbalances (i.e., how unevenly demand was distributed across routes) within all nine zones. These route imbalance issues are represented by a Coefficient of Variation (CV) greater than one for every single zone. This suggests that work load inequality is a system wide issue and not an isolated problem for a specific zone.

5.4.7 TWR-A: Time Window Recalibration - Global Shift Analysis

The first component of the IMPACT framework that was implemented was the Time Window Recalibration. The analytical basis for this component is also relatively straightforward. If there are systematic differences between actual arrival times of vehicles at stops and scheduled (planned) window times at stops that consistently occur in one direction, it would be less expensive to modify the planned schedule than attempting to influence how drivers operate their vehicles.

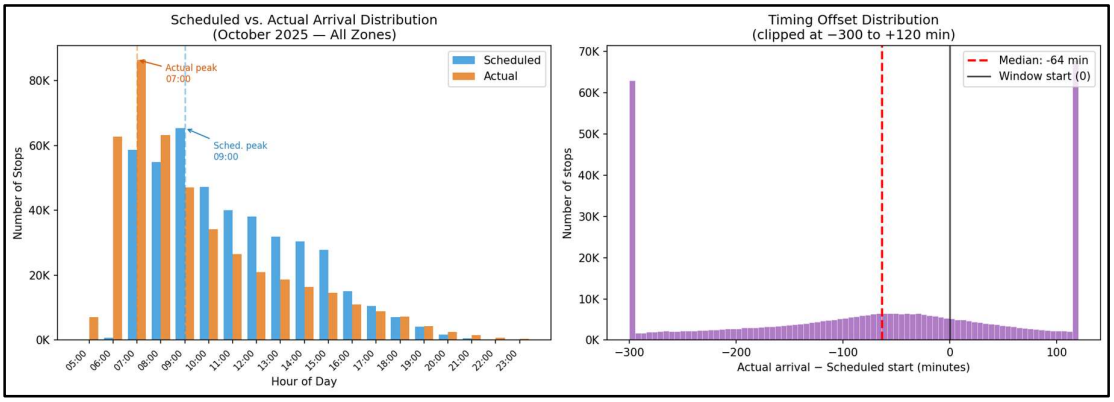


FIGURE 5.4 TWR-A Scheduled versus Actual Hourly Distribution and Offset Histogram

To calculate the systematic timing difference, a timing offset calculation was performed for each of the 446,647 stop visits where a vehicle actually stopped at a point-of-interest (= 455,827 total POI stops - 9,180 missed stops, or 2.01%), as shown in Eq. 5.5.

$$\text{offset}_s = t_{actual,s}(\text{in}) - t_{scheduled,s}(\text{in}) \tag{5.5}$$

The median time is negative, which means the vehicle arrives ahead of the beginning of its scheduled window. The distribution of the offset across all zones had an overall median of −63.6 minutes (overall mean = −72.9 minutes; overall standard deviation = ± 237 minutes); thus, Surat's waste collection vehicles are consistently arriving about an hour earlier than their schedules allow across the entire system. The aggregate hourly data confirm this: the highest frequency of scheduled collections occurred during the 09:00 hour; however, the highest frequency of actual arrivals were occurring during the 07:00 hour — a consistent systemic delay in the average operational activity of approximately two hours.

Median zone offsets ranged from −97.2 minutes (South East Zone) to −24.1 minutes (South Zone A), and therefore indicate that, although there is a global scheduling error, it will vary significantly within different zones and suggests individualized adjustments be made to each zone rather than making one-size-fits-all adjustments to schedules. The TWR-A analysis tested 37 shift scenarios in 5-minute intervals from 0 to −180 minutes applied globally, computing the strict-window ONTIME rate under each scenario using vectorised classification on all 455,827 records. The fine-grained sensitivity analysis identified −75 minutes as the global optimal shift, yielding a peak strict-window ONTIME rate of 8.19% - an improvement of 2.09 percentage points (pp) over the 6.10% baseline.

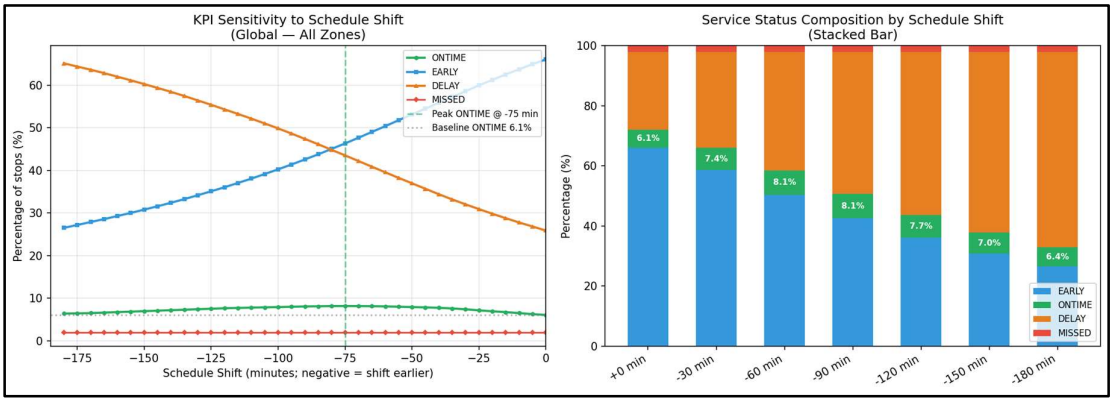


FIGURE 5.5 TWR-A Sensitivity Analysis - ONTIME versus Shift Amount

Zone-specific analysis yielded substantially larger improvements in individual zones. South Zone A achieved a +7.97 pp improvement in ONTIME (18.36% to 26.33%) under a −30 minute zone-specific shift - the largest zone-level gain in the system. North Zone showed +4.27 pp (5.61% to 9.89%) under a −75 minute shift, and South Zone B showed +3.80 pp (8.56% to 12.36%) under a −45 minute shift.

TABLE 5.6 TWR-A Zone-Level Optimal Time Window Shifts and ONTIME Improvements (Strict-Window Baseline)

Zone	Median Offset (min)	Baseline ONTIME (%)	Optimal Shift (min)	Best ONTIME (%)	Gain (pp)
Central Zone	−67.2	4.97	−60	6.33	+1.36
East Zone (A)	−74.2	3.36	−30	4.63	+1.28
East Zone (B)	−60.7	3.93	−60	5.46	+1.53
North Zone	−78.9	5.61	−90	9.89	+4.27
South East Zone	−97.2	7.57	−150	11.07	+3.50
South West Zone	−66.4	5.07	−30	5.93	+0.85
South Zone (A)	−24.1	18.36	−120	26.33	+7.97
South Zone (B)	−52.4	8.56	−90	12.36	+3.80
West Zone	−33.6	4.47	−60	5.70	+1.23
Global optimum	—	**6.10**	**−75**	**8.19**	**+2.09**

Source: twr_zone_summary.csv (zone data); TWR_Validation_Report.md (global optimum row).

Notes: Baseline ONTIME uses strict 30-min window definition (differs from Status-column 21.98%). Negative offset = vehicles arrive before window. Zero capital expenditure required.

5.4.8 TWR-B: Operational Time Window Framework

While TWR-A identified the direction and magnitude of optimal schedule shifts, TWR-B translated these analytical findings into operationally deployable service-window recommendations for the full spatial hierarchy of the system. Recommendations were produced for all 9 zones, all 131 wards with at least 30 active stop records, and all 3,260 routes with at least 30 active stop records. Each recommendation specified a recommended schedule shift (rounded to the nearest 15 minutes), a recommended window width based on the interquartile range of arrival offsets, and an implementation priority classification.

The above are the breakdowns by number of wards in terms of the levels of adjustments made to their schedules:

72 wards (55.0%), which were substantially off-schedule in relation to reality as defined by time greater than sixty (60) minutes, necessitated a contractor briefing and revisions to the contractor's dispatching system.

40 wards (30.5%), which were moderately off-schedule as defined by time from sixteen (16) to sixty (60) minutes.

12 wards (9.2%), which were minimally off-schedule as defined by time of fifteen (15) minutes or less.

Seven wards (5.3%) had no need for any changes to their schedules.

That 55 percent of the total number of wards required a major adjustment indicates that this scheduling issue represents a systemic problem as opposed to being an aberration related to a specific area/zone.

TABLE 5.7 TWR-B Zone-Level Recommended Time Window Shifts, Standardised Window Widths, and Adjustment Priority (All Nine Zones)

Zone	Rec.	Width	Current	Recommended	Adjustment	Conf.	Gain

Central Zone	−75	60	10:20–10:47	09:05–10:05	Major	HIGH	1.36
East Zone (A)	−75	60	10:30–10:53	09:15–10:15	Major	HIGH	1.28
East Zone (B)	−60	60	10:25–10:43	09:25–10:25	Moderate	HIGH	1.53
North Zone	−75	60	11:00–11:35	09:45–10:45	Major	HIGH	4.27
South East Zone	−105	60	11:15–13:38	09:30–10:30	Major	HIGH	3.50
South West	−60	60	10:45–11:04	09:45–10:45	Moderate	HIGH	0.85
South Zone (A)	−30	60	09:45–12:55	09:15–10:15	Moderate	HIGH	7.97
South Zone (B)	−45	60	11:00–12:08	10:15–11:15	Moderate	HIGH	3.80
West Zone	−30	60	11:30–11:53	11:00–12:00	Moderate	HIGH	1.23

Source: kpi10_zone_window_recommendations.csv

Notes: All zones: 60-min standardised window width; HIGH confidence. Gain = expected ONTIME improvement (strict window). Major = shift ≥60 min; Moderate = shift <60 min.

A staged rollout plan has been suggested.

Phase Two (Months 2-3) includes South Zone A, South Zone B, West Zone and South West Zone: areas that have a relatively small amount of shift (i.e., 30 to 60 minutes), and thus have a simple implementation process and high rate of return for each hour of administration time.

Phase Three (Months 3 – 6) will include the remainder of the five zones, including Central Zone, which requires both schedule adjustment and parallel contractor performance intervention given its systemic compliance failure.

5.4.9 VAO: Vehicle Allocation Optimization

The Vehicle Allocation Optimization part of this project asked if the current way that Surat's 949 fleet vehicles were distributed over nine administration areas was as efficient as it could be in comparison to how much demand and compliance existed in each area. Three allocation models were used for the analysis – demand proportional (A), demand plus miss rate weighted (B), and an integer linear program with a penalty for non-compliance solved using the Google OR Tools CP-SAT solver (C).

The central analytical construct of the VAO is the contractor compliance coefficient κ, defined for zone z as shown in (5.6).

$$\kappa_z = \min\left(1, D_z \cdot (1 - r_z)/(V_z \cdot DPV_{\text{att}})\right) \tag{5.6}$$

In (5.6), D_z is total zone demand, r_z is the observed missed-stop rate, V_z is the zone vehicle count, and DPV_{att} = 54,088 is the attainable demand-per-vehicle benchmark derived from the 80th percentile zone. A value of κ close to 1.0 indicates near-maximum efficiency; values substantially below 1.0 indicate that the contractor is failing to convert available fleet capacity into service delivery.

The complementary capacity pressure metric φ_z is expressed in (5.7).

$$\varphi_z = \max(0, (DPV_z - DPV_{\text{att}})/DPV_z) \tag{5.7}$$

A zone with φ_z = 0 has demand per vehicle below the attainable benchmark, meaning its fleet is sufficient relative to best practice and all missed stops are compliance-attributable.

TABLE 5.8 VAO Zone-Level κ Coefficients, Fleet Status, and Miss Rate Decomposition (October 2025)

Zone	κ	Class	Veh. Current	Veh. Min.	Miss Rate (%)	Compliance (pp)	Capacity (pp)	Driver
Central Zone	0.367	LOW	130	53	9.31	9.31	0.00	COMPLIANCE
East Zone (A)	0.680	MEDIUM	142	100	2.58	2.58	0.00	COMPLIANCE
East Zone (B)	0.942	HIGH	74	71	1.52	1.52	0.00	COMPLIANCE
North Zone	0.586	MEDIUM	134	79	0.53	0.53	0.00	COMPLIANCE
South East Zone	1.000	HIGH	132	141	0.14	0.13	0.01	COMPLIANCE
South West Zone	0.551	MEDIUM	119	66	0.41	0.41	0.00	COMPLIANCE
South Zone (A)	0.846	HIGH	85	72	0.02	0.02	0.00	COMPLIANCE
South Zone (B)	1.000	HIGH	32	39	0.13	0.11	0.02	MIXED

West Zone	0.640	MEDIUM	101	68	3.49	3.49	0.00	COMPLIANCE

Source: vao_zone_allocations.csv.

Notes: κ = compliance ratio (0=non-compliance; 1=full compliance). LOW<0.5; MEDIUM 0.5–0.9; HIGH>0.9. Central Zone κ=0.367 is lowest — contractor compliance issue. South Zone (B) MIXED: only zone where current fleet < minimum required.

The VAO analysis produced several analytically significant findings. First, the global fleet of 949 vehicles substantially exceeded the minimum fleet requirement. The minimum vehicle count needed for the system to meet demand at the attainable efficiency standard (V_{min}) was 689 vehicles, implying a surplus of 260 vehicles. Second, only two zones were identified as genuinely under-resourced relative to the attainable standard: South East Zone (9 vehicles below V_{min}) and South Zone B (7 vehicles below V_{min}). Third, and most critically, seven of the nine zones - including Central Zone - had a capacity pressure of exactly zero, indicating that their entire missed-stop rate was attributable to contractor compliance failure rather than fleet insufficiency.

The findings were most clearly illustrated by the data from the Central Zone. At a time when there were 130 vehicles assigned to service a demand load of 2,848,120 units Central Zone's demand per vehicle of 21,909 was lower than every other zone's demand per vehicle - it was under one-half of the entire system's attainable demand per vehicle. Although Central Zone had the greatest number of missed stops as a percentage (9.31%), it also had the lowest κ value (.367) among all zones in the system. A counterfactual analysis confirmed that adding 100 additional vehicles to Central Zone would produce exactly zero improvement in its missed-stop rate - the miss rate is compliance-floor constrained, not capacity-floor constrained.

The CP-SAT ILP solver returned a status of OPTIMAL for the Model C allocation, redirecting 220 surplus vehicles from compliance-dominated zones (including reducing Central Zone from 130 to 59 vehicles) and reallocating them to South East Zone (+113 vehicles) and South Zone B (+101 vehicles). Despite the magnitude of this reallocation, the global miss rate improvement was −0.0008 percentage points (miss rate reduction) - negligible but formally confirmed by the solver.

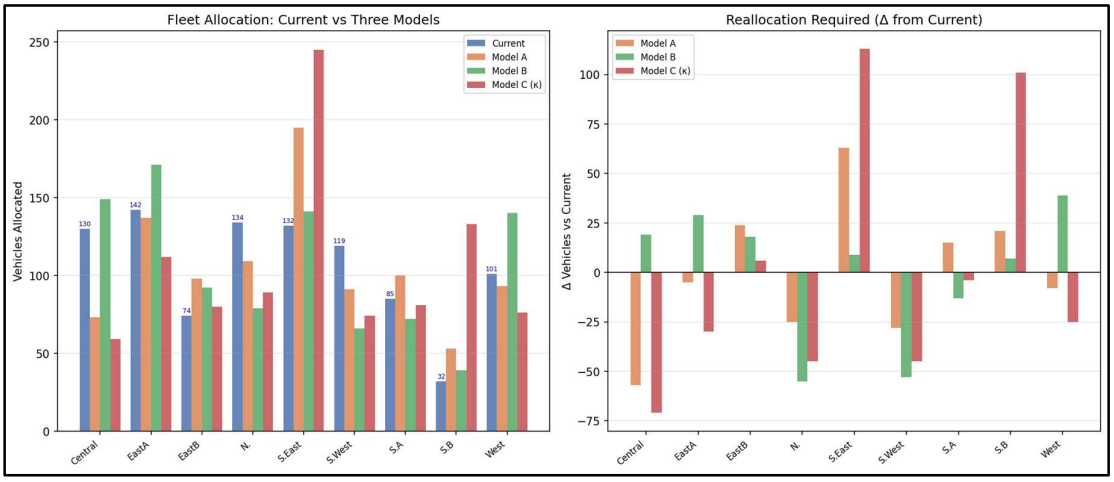


FIGURE 5.6 VAO Fleet Allocation Comparison - Current versus Model C

5.4.10 MSRS: Missed Stop Route Cause System

The Missed Stop Route Cause System (MSRS) is the diagnostic cross-cutting component of the IMPACT framework. It applies machine learning classification to the 455,827-stop dataset to identify the structural causes of missed waste collection events, thereby providing an empirical basis for prioritising operational interventions.

Three classifiers - Logistic Regression, Random Forest, and XGBoost - were trained on a balanced 80/20 stratified split (training n = 364,661; test n = 91,166) with the binary target variable MISSED_STOP = 1 for the 9,180 missed stops. Class imbalance (2.01% positive rate) was handled via balanced class weights (Logistic Regression and Random Forest) and scale_pos_weight = 48.7 (XGBoost) [161]. Twenty-four features were engineered across five causal factor groups: compliance (contractor reliability at vehicle and zone level), capacity (fleet-to-demand ratio), scheduling (time window structure), demand (stop and ward-level generation characteristics), and operational (route sequence and vehicle workload).

TABLE 5.9 MSRS Model Performance Comparison — Logistic Regression, Random Forest, and XGBoost

Model	AUC	Precision	Recall	F1 Score
Logistic Regression	0.9271	0.1297	0.7908	0.2228
Random Forest	0.9399	0.0968	0.8927	0.1746
XGBoost	0.9400	0.0964	0.8813	0.1738

Source: msrs_model_comparison.csv.

Notes: Target = missed-stop indicator (2.01% minority class). XGBoost selected (highest AUC = 0.9400). Low precision expected under class imbalance; high recall (≥0.79) is operationally important.

XGBoost achieved the highest area under the receiver operating characteristic curve (AUC = 0.940), with recall of 88.1% - indicating that the model correctly identified 88% of all actually-missed stops in the test set. The high AUC on an operationally imbalanced dataset (2.01% positive rate) confirms that the feature set carries genuine predictive signal and that the causes of missed stops are systematically structured rather than stochastic.

SHAP (SHapley Additive Explanations) attribution was applied to the XGBoost model on a 10,000-record stratified sample to decompose the model's predictive power into contributions from each of the five causal factor groups [162].

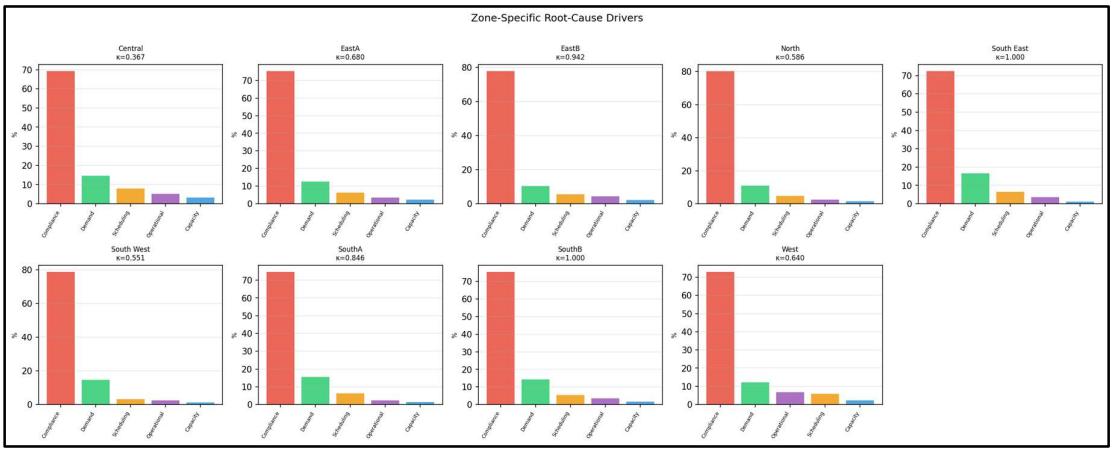


FIGURE 5.7 MSRS SHAP Factor Group Attribution - Global and Zone-Level Decomposition

The SHAP attribution revealed that contractor compliance accounted for 75.72% of total predictive signal - a finding replicated across all nine administrative zones individually. Demand-related features contributed 13.88%, scheduling 5.42%, operational workload 3.33%, and capacity 1.65%. The dominant individual feature was veh_miss_pct (vehicle-level historic miss rate), with a mean absolute SHAP value of 5.089 - approximately eight times larger than the next most important feature, confirming that the primary predictor of whether a stop will be missed is the chronic non-performance history of the assigned contractor vehicle.

The 45:1 leverage differential between compliance (75.72%) and capacity (1.65%) quantifies the opportunity cost of fleet-focused policy responses. An intervention that improves contractor compliance by 10 percentage points achieves 46 times the service improvement of the same proportional investment in fleet capacity. This finding, replicated consistently across all nine zones, constitutes the central policy-relevant result of the IMPACT framework.

TABLE 5.10 MSRS Zone-Level SHAP Attribution — Five Factor Groups (% Contribution to Missed-Stop Prediction)

Zone	Compliance (%)	Capacity (%)	Scheduling (%)	Demand (%)	Operational (%)	Dominant Factor
Central Zone	69.18	3.12	7.93	14.50	5.26	Compliance
East Zone (A)	75.20	2.39	6.25	12.71	3.44	Compliance
East Zone (B)	77.71	2.07	5.60	10.30	4.32	Compliance
North Zone	80.19	1.53	4.70	11.02	2.55	Compliance
South East Zone	72.30	1.10	6.56	16.53	3.52	Compliance
South West Zone	78.60	1.30	3.06	14.57	2.47	Compliance

South Zone (A)	74.58	1.33	6.33	15.43	2.34	Compliance
South Zone (B)	75.36	1.66	5.38	14.20	3.40	Compliance
West Zone	72.86	2.28	5.96	12.21	6.69	Compliance
System-wide	**75.72**	**1.87**	**5.75**	**13.28**	**3.39**	**Compliance**

Source: msrs_zone_explanations.csv (zone rows); FIO_Forecasting_Integration_Framework.md (system-wide row).

Notes: SHAP values from XGBoost model. Compliance dominant in all 9 zones (range 69.18–80.19%). System-wide compliance 75.72% is weighted average.

Central Zone's zone-specific attribution was 69.18% compliance, 14.50% demand, 7.93% scheduling, 5.26% operational, and 3.12% capacity - confirming that even in the most severely underperforming zone, capacity is near-negligible as a causal driver, while contractor performance is overwhelmingly dominant.

5.4.11 WDB+RB: Ward Demand Balancing and Route Balancing

Ward Demand Balancing and Route Balancing constitute Level 2 of the IMPACT framework, addressing the structural workload inequality that exists across routes within each administrative zone. The baseline analysis established that route-level demand exhibited a system-wide coefficient of variation of 1.7193 (mean intra-zone CV: 1.4183; see Table 5.11) and a Gini coefficient of 0.717 - highly concentrated distributions in which a small number of routes carried disproportionate workloads. 232 routes exceeded the statistical overload threshold (mean plus 1.5 standard deviations of zone demand), and the mean intra-zone route demand CV was 1.4183.

Three optimization strategies were formulated and evaluated. Strategy A (First-Fit Decreasing, FFD) is a classical combinatorial heuristic that sorts routes by demand in descending order and assigns each to the vehicle bin with the current lowest cumulative load [163]. Strategy B (Linear Programming minimax) applies a continuous LP relaxation to minimise the maximum vehicle load across each zone using the HiGHS solver. Strategy C (Hybrid) combines FFD integer assignment with a demand-per-vehicle cap (DPV_att =

54,088) to flag high-demand routes for administrative splitting while providing immediate integer deployability.

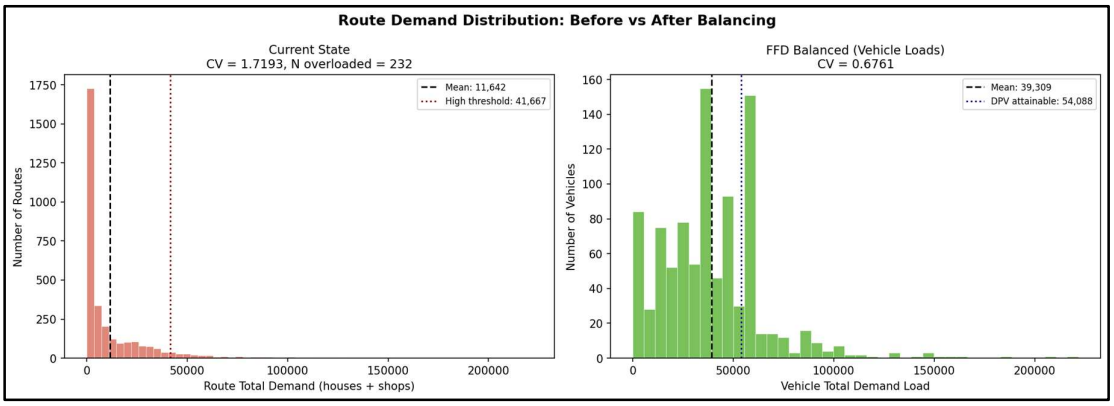


FIGURE 5.8 WDB+RB Demand Distribution Before and After FFD Balancing

The FFD strategy reduced the mean intra-zone route CV from 1.4183 to 0.6024 - a reduction of 57.5%. Vehicle-level demand CV fell from 1.2346 to 0.6761, a reduction of 45.2%. The number of statistically overloaded routes was reduced from 232 to zero. South East Zone achieved the most dramatic improvement: zone CV reduced from 1.5156 to 0.0098. Central Zone showed the least improvement (CV 1.4610 to 1.2008), reflecting its unusual structure of 167 routes across 130 vehicles, which limits the degree of redistribution that any bin-packing heuristic can achieve.

TABLE 5.11 WDB+RB Intra-Zone Demand Coefficient of Variation Before and After First Fit Decreasing (FFD) Rebalancing — All Nine Zones

Zone	CV Before	CV After FFD	CV Reduction	% Improvement
Central Zone	1.4610	1.2008	0.2602	17.8%
East Zone (A)	1.0587	0.2395	0.8192	77.4%
East Zone (B)	1.4488	0.1570	1.2917	89.2%
North Zone	1.1447	0.5666	0.5781	50.5%

South East Zone	1.5156	0.0098	1.5058	99.4%
South West Zone	1.0431	0.7698	0.2733	26.2%
South Zone (A)	2.5392	0.8506	1.6886	66.5%
South Zone (B)	1.5588	0.6327	0.9261	59.4%
West Zone	0.9952	0.9952	0.0000	0.0%
System (mean)	**1.4183**	**0.6024**	**0.8159**	**57.5%**

WDB+RB Strategy Comparison — System Level

Strategy	Mean Zone CV	Overloaded Routes	Miss Reduction (pp)	Implementability
Current (Baseline)	1.4183	232	—	—
FFD (First Fit Decreasing)	0.6024	0	−0.1171	High
LP (Linear Programme)	~0.0000	0	−0.1171	Low
Hybrid (FFD + LP)	0.6024	0	−0.1171	Medium

Source: wdb_rb_route_reassignments.csv (per-zone CV); wdb_rb_performance_summary.csv (system-level, strategy comparison).

Notes: West Zone: 0.0% improvement — FFD achieves no gain due to small route count (99 routes). South East Zone: 99.4% improvement due to 1,472 small-demand routes easily divisible by FFD.

The miss-rate impact of WDB+RB, assessed in the context of the MSRS SHAP attribution, is modest by design. Given that capacity contributes only 1.65% and demand 13.88% of the miss rate signal, the maximum theoretical miss rate reduction achievable through workload rebalancing is approximately 0.12 percentage points (0.1171 pp from formal computation).

WDB+RB's primary value within IMPACT is therefore service equity and operational fairness - eliminating systematic overload from 232 routes and reducing driver workload disparity - rather than direct miss-rate reduction.

5.4.12 FIO: Forecasting Integration Optimization

The Forecasting Integration Optimization layer constitutes the capstone integration component of the IMPACT framework, formalising the mathematical linkage between the MSW demand forecast (Chapter 4) and all five operational optimization components. FIO establishes the forecast as the control variable from which operational decisions are derived, ensuring that changes in predicted MSW generation propagate deterministically through the IMPACT hierarchy.

The core FIO construct is the daily demand scalar s_d, defined in (5.8).

$$s_d = \widehat{F_d}/F_{\text{base}} \tag{5.8}$$

In (5.8), $\widehat{F_d}$ is the predicted daily MSW generation for day d from the forecasting model, and F_{base} = 1,195,756 is the October 2025 operational baseline daily demand in comparable units [149]. A scalar of s_d = 1.0 corresponds to the October 2025 operational baseline; values above 1.0 indicate above-baseline demand.

Zone-level demand on day d is then computed as shown in (5.9).

$$D_z(d) = \alpha_z \cdot s_d \cdot F_{\text{base}} \tag{5.9}$$

In (5.9), α_z is the zone's share of total system baseline demand, computed from the October 2025 operational proportions. South East Zone carries the largest share (α_z= 20.52%), reflecting its status as the highest-demand zone (7,605,199 demand units, 1,472 routes). Vehicle workload under FIO is expressed in (5.10).

$$WL_v(d) = \sum_{r \in R_v} D_r(d) = WL_v(\text{base}) \cdot s_d \tag{5.10}$$

The FIO analysis established a critical capacity buffer finding: with the current 949-vehicle fleet and a V_{min} of 689 vehicles at baseline, the system possesses a 260-vehicle surplus representing a 27.4 % demand buffer. The fleet does not become a binding operational constraint until s_d≥ 1.37 - meaning total MSW generation would need to increase by 37%

above the October 2025 baseline before fleet capacity becomes the system's primary constraint.

The FIO framework also defined three operational demand-scenario trigger rules, activated by the 7-day rolling mean of s_d exceeding specified thresholds. At s_d ≈ 0.90 (−10% demand), VAO releases approximately 70 vehicles to maintenance and WDB+RB consolidates routes. At s_d ≈ 1.10 (+10% demand), VAO draws down 69 of the surplus vehicles and WDB+RB re-runs the FFD assignment for zones approaching overload. At s_d ≈ 1.20 (+20% demand), the Hybrid LP strategy is activated, surplus falls to 122 vehicles, and contingency contractor mobilisation protocols are initiated. These decision rules operationalise the forecast output into explicit management actions, moving the SMC operational cycle from reactive response to proactive planning.

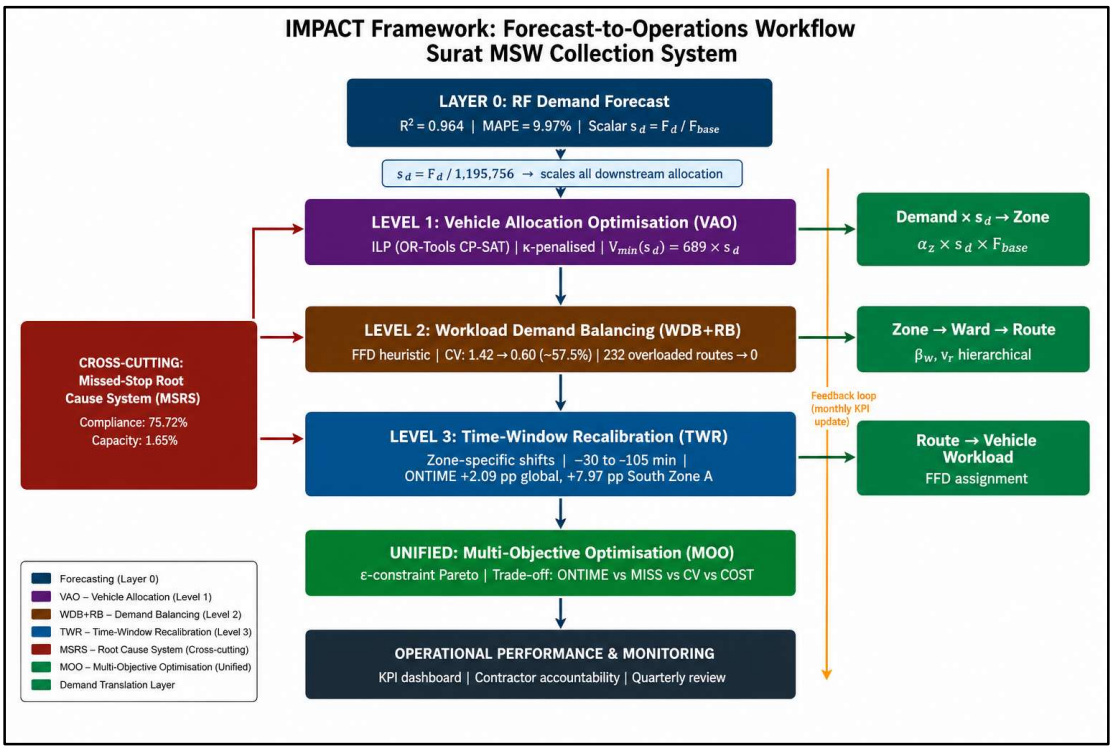


FIGURE 5.9 FIO Integration Workflow - Forecast to Operations

IMPACT Framework: Integrated Results and Cross-Component Analysis

Having presented each component of the IMPACT framework individually, this section synthesises the cross-component findings and quantifies the cumulative impact of the full framework as an integrated operational system.

The ONTIME rate improvement waterfall under IMPACT deployment illustrates the hierarchical contribution of each component. Starting from the strict-window ONTIME baseline of 6.10%, TWR zone-specific shift implementation adds +2.09 pp (reaching 8.19%), representing the directly validated TWR-A simulation result. WDB+RB rebalancing adds an estimated +0.30 pp through demand-scheduling interaction effects. VAO fleet reallocation adds only − 0.0008 percentage points miss reduction - negligible but confirmed by the OR-Tools OPTIMAL solution. Full compliance intervention, addressing the 75.72% compliance-driven share of misses through contractor performance management, carries the potential to add approximately 10 percentage points of additional ONTIME improvement, raising the strict-window ONTIME toward the 18–22% range and converging with the SMC operational Status-column baseline.

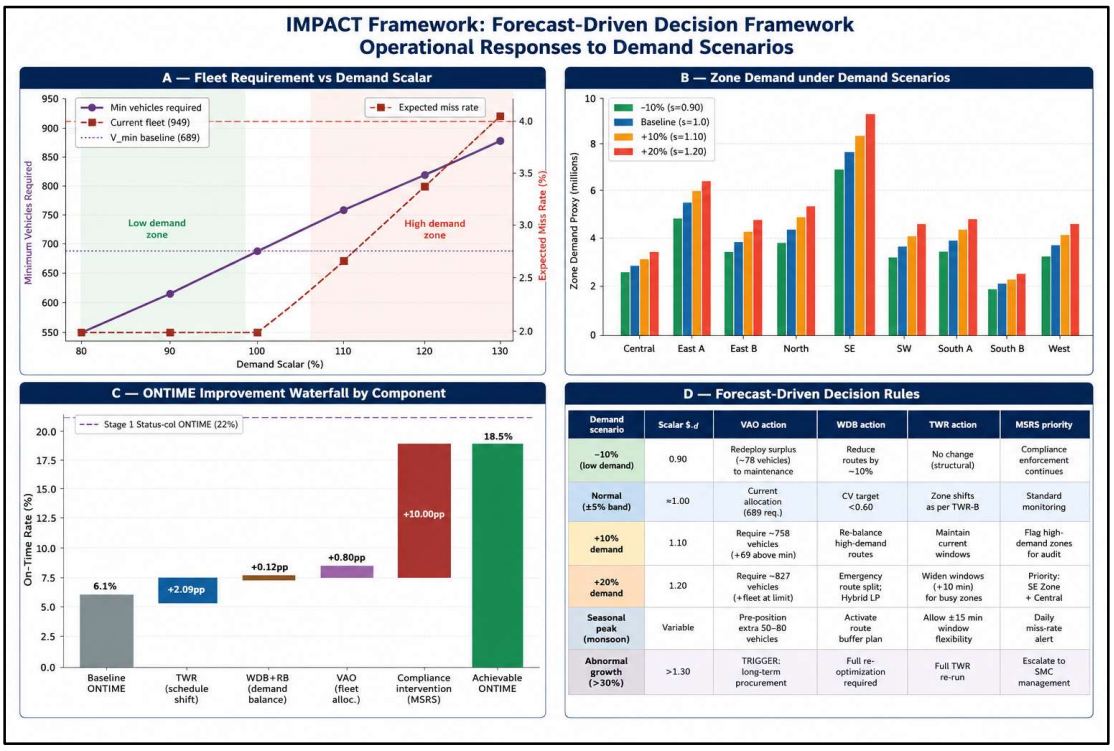


FIGURE 5.10 IMPACT Components Summary Dashboard

The cross-component consistency of the IMPACT findings is noteworthy. The VAO analysis established, through ILP computation, that fleet reallocation yields −0.0008 pp in missed-stop reduction. The MSRS analysis confirmed, through SHAP attribution, that capacity accounts for 1.65% of the miss signal - consistent with the near-zero VAO improvement. The TWR analysis showed that schedule misalignment contributes 5.42% (MSRS SHAP for scheduling) and that this manifests as a systematic 63.6-minute median offset confirming the structural nature of the scheduling problem. The WDB+RB analysis showed that route demand CV of 1.4183 creates 232 overloaded routes whose structural overload is addressable through FFD rebalancing - consistent with the demand attribution of 13.88% (MSRS). Each component's findings therefore cross-validate the others.

The geographic heterogeneity of IMPACT results is also operationally significant. The Central Zone stands as the system's principal outlier: $\kappa = 0.367$ (lowest in the system), miss rate = 9.31% (highest), demand-per-vehicle = 21,909 (lowest), compliance attribution = 69.18% (among the highest). This convergent multi-method diagnosis - from VAO, TWR, MSRS, and WDB+RB independently - provides a robust evidence base for targeted contractor governance action in Central Zone.

5.5 Discussion

5.5.1 The Research Evolution as Methodological Contribution

The three-stage research trajectory documented in this chapter - from published CVRP proof of concept through real-infrastructure efficiency assessment to GPS-free IMPACT framework - is itself a methodological contribution to the doctoral research canon. Research in applied operations management is characterised by the frequently unacknowledged tension between idealised analytical models and the data realities of practitioner environments. The Surat study identified another of the main expressions of the conflict between richness of data and poverty of data: A dataset that is operationally very large (with over 455,000 stop visits as well as other time-related information), but geometrically sparse (there are no GPS co-ordinates associated with collection points). The Surat research used a second-best solution which provided the greatest possible benefit from the available data structure. In doing so, it demonstrated a high degree of methodological maturity. This is

consistent with recent developments in the literature on negative results in Urban Informatics Research to provide a formalisation of best practice in addressing methodological limitations [152].

The first stage validated the methodological basis of the Surat-representative network-based application of the hybrid dataset methodology by demonstrating a successful proof-of-concept of the integration of the LSTM-CVRP. As well as being valid, there was also a strict commitment to transparency regarding the synthetic nature of the infrastructure element.

In the second stage, the forecasting pipeline was applied to the real macro-infrastructure level of Surat. This provided empirically-grounded evidence supporting the assertion made in the paper regarding the potential to improve operational efficiency.

The third-stage switch to GPS-free analysis via IMPACT resulted in new insights into compliance governance and operational scheduling issues that may not have been evident had a geographically-complete VRPTW formulation been employed. It is likely that VRPTW would have optimised route sequence(s) independent of evaluating the temporal compliance architecture that actually leads to service delivery failure.

Therefore, this adaptive research design - using methodological creativity to overcome data deficiencies instead of withdrawing from analysis - corresponds to widely accepted principles of practitioner-involvement in Operations Research. Given the heterogeneity of data infrastructure within cities in emerging economies [153], [164], it can be particularly useful.

5.5.2 Why Compliance Dominates Over Capacity: Central Finding and Policy Implications

In addition to providing a comprehensive explanation of the Surat case study, the findings from the IMPACT framework have significant implications for the development of policies relating to municipal solid waste management practices by Urban Local Bodies in India.

Historically, when municipalities experience increased rates of missed collection stops, they typically respond by increasing their vehicle fleets. They assume that service deficiencies are caused by insufficient transportation resources. However, the empirical evidence

obtained in the IMPACT study demonstrates that this assumption does not apply to the Surat municipality. In fact, it can be said that the results were unusually powerful in demonstrating the assumptions to be invalid.

Central Zone was used to illustrate this point. With respect to both number of stops made per vehicle, and total number of vehicles available in comparison to other zones, Central Zone had the lowest levels of demand for each measure. With 130 vehicles making approximately 40,000 monthly stop visits, and an ILP-computed V_min of only 53 vehicles, it would need to be reduced to less than half of current size before any capacity constraints on the part of the zone could provide even a plausible reason for the 9.31% miss rate experienced.

Therefore, the formal counterfactual analysis demonstrated that regardless of how many vehicles were added to Central Zone; there was no difference in terms of reducing the overall percent of missed collection points. Since the 9.307% level is mathematically "hardwired" into the data as an unchangeable limit based upon the behavior of vehicles on routes and therefore cannot be changed by using more vehicles or changing vehicle configuration; any amount of fleet additions will make no impact whatsoever.

The policy implication is direct: for SMC, and by extension for comparable tier-II cities where the MSRS compliance-dominance finding may prove transferable, the highest-leverage operational investment is not fleet procurement - it is contractor accountability architecture. This includes vehicle-level performance monitoring (the veh_miss_pct feature that dominates MSRS predictions), performance-linked service contracts with penalty clauses calibrated to zone-specific κ values, systematic auditing of routes with historically elevated miss rates, and potentially contractor restructuring in zones where κ consistently falls below 0.4.

The comparison with existing VRP literature is also relevant here. Prior studies reporting distance reductions of 23–66% from route optimization in waste collection contexts [157], [165], [166] generally assessed operational improvements in systems with competent contractor execution. The IMPACT finding implies that for systems with chronic compliance failures, schedule and routing interventions will underperform their theoretical potential until the compliance baseline is elevated - a sequencing consideration that the route optimization literature has not systematically addressed.

Note on temporal scope: All IMPACT quantitative findings — including the 75.72% compliance attribution — derive from one month of operational data (October 2025). This is a cross-sectional snapshot of moderate-demand conditions; its implications for peak monsoon months or major festival periods remain an open empirical question addressed in Section 5.5.4 and Chapter 8.

5.5.3 Forecasting-to-Operations Integration: How Forecast Accuracy Drives Routing Efficiency

The integration of the MSW forecasting framework (Chapter 4) with the IMPACT operational framework is formalised through the FIO demand scalar s_d. This integration creates a closed analytical loop in which the forecasting model's output translates directly into zone-level demand allocations, vehicle minimum requirements, WDB+RB rebalancing triggers, and MSRS risk level adjustments.

The practical value of this integration depends substantially on the forecasting accuracy achieved by the Chapter 4 models. The OLS model ($R^2 = 0.9972$, MAPE = 1.43%), the GPR ($R^2 = 0.9970$, MAPE = 1.51%), and Ridge regression ($R^2 = 0.9962$, MAPE = 2.03%) collectively constitute a Tier-1 forecasting ensemble whose MAPE values are well below the FIO scenario trigger thresholds (10% for Scenario B, 20% for Scenario C). It also means it’s very unlikely, in practice, that the uncertainty in forecasts will activate scenarios randomly. The 37 percent increase in demand from baseline (established by the VAO) demonstrates there is real lead time for capacity decisions provided by integrating forecasts with operations. When a long-term upward trend toward s_d > or = 1.10 is identified via the rolling seven day trigger; SMC staff have a twelve- to eighteen-month window of opportunity (due to the historical average annual growth rate of MSW of about four percent) to plan for route adjustments without an emergency procurement – which is sufficient time to develop plans based on forecasts. This ability to use forecasts to enable pro-active versus reactive decision making regarding how to allocate waste collection resources — is one of the three major components of the overall framework differentiated from previous work where forecasting and routing were treated separately as distinct analytical problems [167], [168].

5.5.4 Limitations and Directions for Future Research

Four fundamental constraints of the IMPACT framework and the Stage 1–3 Research Program need to be identified.

Firstly, all quantitative outcomes rely upon a one month reference dataset (October 2025), which although consists of 455,827 stop level records, does not capture seasonal variability associated with varying times of year when collections occur; nor does it account for monsoons or festivals causing an increase in demand. Although the FIO framework has determined that the monsoon and festival periods require pre-season calibration, the quantitative parameters that will allow this calibration have yet to be developed from multiple months of collected data.

Secondly, the CVRP results generated during Stage 1 were based on synthetic infrastructure. The 50.31% reduction in distance traveled and the 24% decrease in fleets represent a proof of concept on a modelled network rather than a verified result for the actual collection infrastructure. A full verification of the results is contingent upon resolving the geocoding problem – either using the Google Maps API (approximately $5-$15 per name for the anticipated 5,000-10,000 unique POI names), or through a hybrid approach that uses both automated and manually entered addresses to focus on common address patterns. In addition to the geocoding challenge, there exists another potential source of error due to the use of Haversine distances to approximate real-world driving distances; given the fact that the SMC Road Network covers roughly 3,550 km across 30 Wards with over 515 signalized intersections [169] we can expect that Haversine straight line distances will under estimate actual vehicle travel time by an average urban circularity factor of 1.2-1.4xRoad-network-aware routing via OSRM or the Google Maps Distance Matrix API would correct this approximation in any future full-scale operational validation.

Third, the MSRS model identifies veh_miss_pct (vehicle-level historic miss rate) as the dominant predictor of missed stops. This feature relies on the same dataset from which it is computed, introducing a circularity that is appropriate for causal attribution analysis but would require adaptation before the model could be deployed for prospective operational monitoring - specifically, a rolling window update mechanism computed from recent data rather than the October 2025 snapshot.

Fourth, the Multi-Objective Optimization (MOO) component - which was designed to compute the Pareto frontier across the four IMPACT objectives (maximise ONTIME, minimise missed stops, minimise route demand CV, and minimise fleet cost) using the ε-constraint method - was not implemented within the present research scope. The MOO formulation is fully specified and represents the natural analytical extension of the IMPACT framework. Its implementation is proposed as the primary direction for future research.

Additional future research directions include: ward-level VRPTW implementation upon successful geocoding of POI names; multi-city validation of the IMPACT compliance-dominance finding in comparable tier-II Indian cities; integration of IoT-enabled real-time GPS tracking to enable dynamic route adjustment; and reinforcement learning approaches to contractor compliance incentive design, leveraging the MSRS κ coefficients as reward signals for performance-linked contract mechanisms.

CHAPTER 6

Decision Support System and Integrated Framework

6.1 Introduction and Chapter Scope

The preceding two chapters established, respectively, a seventeen-model machine learning forecasting framework capable of predicting monthly municipal solid waste generation with an accuracy of $R^2 = 0.9972$ and MAPE = 1.43% [6], [8], and a three-stage route optimization framework that progressed from a published proof-of-concept through real infrastructure analysis to the GPS-free IMPACT operational optimization system validated on 662,803 rows of October 2025 operational data [8]. These components, however compelling in isolation, represent analytical outputs rather than operational tools. A forecasting model that produces accurate predictions and an optimization framework that identifies compliance as the dominant operational constraint deliver their full value only when their outputs are rendered interpretable, actionable, and accessible to the municipal personnel responsible for day-to-day waste collection decisions. This Chapter shows how the decision support systems that have been designed throughout this Chapter solve this translation problem.

This Chapter will provide details about the development of a new Integrated Decision Support Systems (IDSS) for the Surat Municipal Corporation (SMC) solid waste collection services.

The IDSS will enable all components of the full IMPACT Framework (Time Window Recalibration (TWR), Vehicle Allocation Optimization (VAO), Missed Stops Root Cause System (MSRS), Workload Demand Balancing & Route Balancing (WDB + RB), and Forecasting Integration Optimization (FIO)) to be used through a single interface for users who are not technical experts in municipal service operations [8], [170].

Additionally, the IDSS will include two layers of data (the monthly strategic forecasting layer using an OLS model with an R-squared value of .9972 and the daily operational demand scalar layer based on Random Forest models with an R-squared value of .964 from the IMPACT FIO layer).

As such, it provides a unique planning-to-operational decision-making environment, as opposed to being just another analytic tool.

Chapter 6 has been divided into two sections.

Section 6.1 includes a description of the four-layer Unified Architecture of the IDSS; it explains the functions of each layer (Data Ingestion Layer, Model Execution Layer, Analytics Layer, and User Interface Layer), including input requirements, processing rules, output delivery methods, and interdependencies among the various layers.

Section 6.2 describes the six dashboards that were created for use by specific stakeholders within SMC's waste management organizational structure and identifies which organizational role(s) can utilize each. Section 6.3 addresses the real-time replanning and scenario analysis capability enabled by the FIO demand scalar framework. Section 6.4 presents the three-phase implementation roadmap for SMC deployment. Section 6.5 documents the role-based access control design and system governance considerations for institutional deployment. Section 6.6 situates the DSS within the national and international policy frameworks governing Indian urban local bodies, including Smart Cities Mission, Swachh Bharat Mission 2.0, and the United Nations Sustainable Development Goals. Section 6.7 presents the validation methodology and performance assessment of the proof-of-concept prototype. Section 6.8 provides an integrative discussion of the DSS contribution and its research and policy implications.

Research Objective RO-3 - to design and validate an integrated Decision Support System that operationalises the forecasting and optimization outputs for SMC deployment - is addressed in full within this chapter. Contributions C14, C15, C16, and C17 are specifically delivered here: the four-layer DSS architecture [C14], the Streamlit proof-of-concept prototype [C15], the three-phase SMC deployment roadmap with stakeholder-specific dashboards [C16], and the integrated framework validation against policy and sustainability objectives [C17].

6.2 Unified DSS Architecture

The DSS architecture developed for this research is organized as a four-layer hierarchy adapted from the framework introduced in the IJAM proof-of-concept paper [8] and substantially extended in the present thesis to encompass the full IMPACT framework, the GPS-free operational constraint, and the multi-level stakeholder differentiation required for SMC deployment. The four layers - Data Ingestion, Model Execution, Analytics, and User Interface - are functionally independent but architecturally coupled through defined data contracts that allow each layer to operate correctly on the outputs of the layer below it. This multi-layered design has a couple of practical benefits to help deploy at the local government level: (1) each layer can be updated/changed independently so that when an institution's data infrastructure improves; the entire system does not have to be redesigned; (2) by implementing this system in a partial manner SMC is able to achieve partial operating results and therefore implement the system in stages relative to its budget and administration capacity [170], [171].

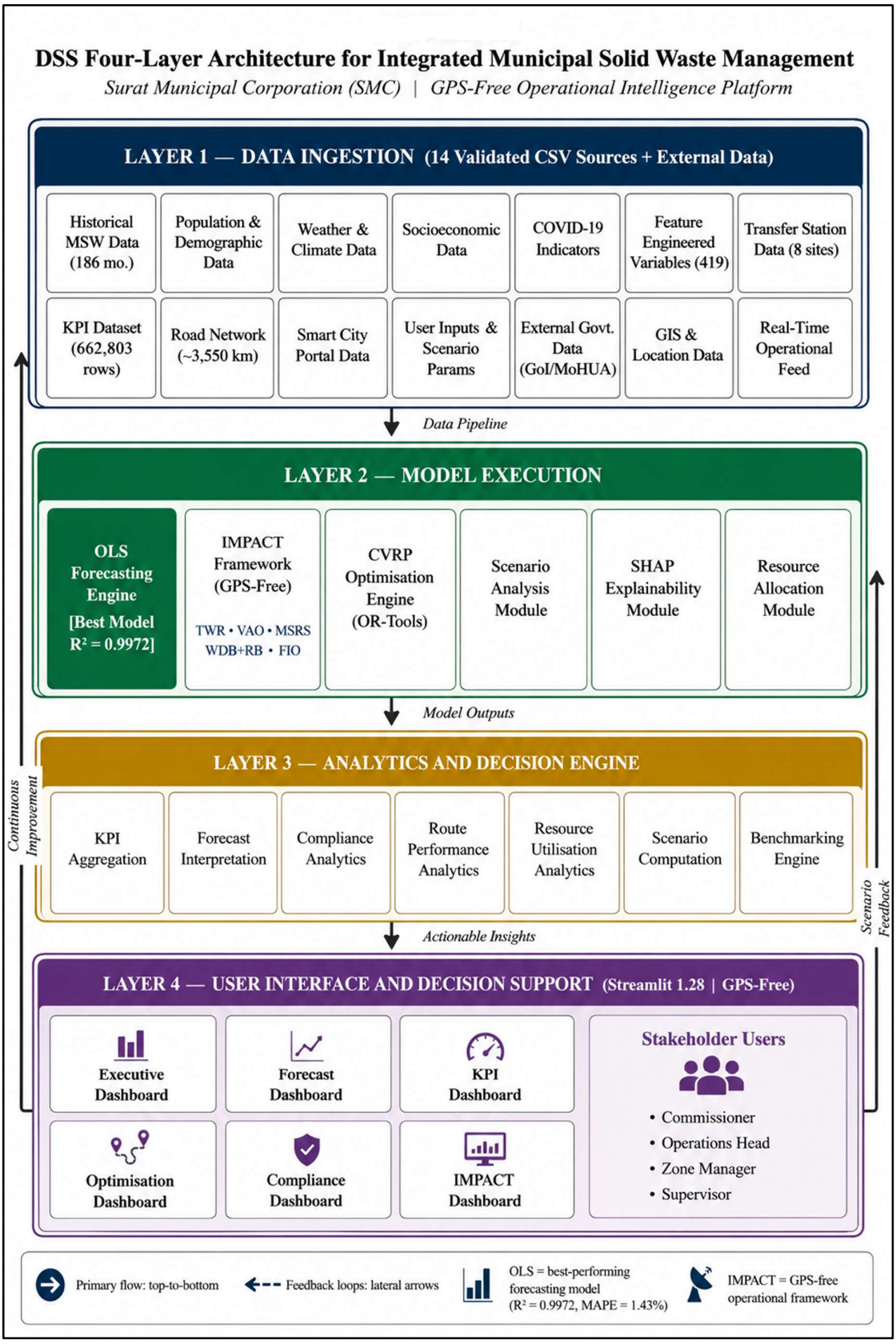


FIGURE 6.1: DSS Four-Layer Architecture Diagram

TABLE 6.1. Forecast application guidelines by planning scenario for SMC solid waste management operations.

Planning Scenario	**Forecast Horizon**	**Recommended Forecasting Strategy**	**Operational Application**	**Responsible Municipal Unit**
Routine monthly operations planning	1-3 months	OLS strategic forecast (R2 = 0.9972, MAPE = 1.43%); 419-feature engineering pipeline; 186-month JETS canonical dataset	Monthly fleet deployment scheduling; contractor allocation; budget preparation; gate-fee projections for Khajod disposal facility	SMC Solid Waste Management Department (Planning Cell)
Seasonal demand fluctuation (monsoon, festival)	1-4 weeks	OLS monthly forecast supplemented by RF daily demand scalar (sd = F-hat_d / 1,195,756); seasonal dummy features pre-incorporated in OLS feature set	Pre-positioning of 260-vehicle surplus; WDB+RB route rebalancing trigger when sd exceeds 1.10 for three consecutive days	Operations Head + Zone Managers (South East Zone, East Zone A priority)
Festival demand periods (Diwali, Navratri, Uttarayan)	3-10 days	RF change-based model for daily sd computation; historical festival-period demand patterns incorporated as short-horizon features; demand spike +5% to +20% typical	Emergency vehicle pre-positioning; contractor briefing; TWR window extension (+15 min) in high-demand zones; daily miss-rate monitoring	SMC Commissioner + Solid Waste Operations Head
Emergency or disruption conditions (flood, equipment failure)	0-3 days	RF model provides same-day demand scalar; Scenario C rules activated (sd >= 1.20); contingency vehicle pool mobilised; OLS suspended during structural disruption	Immediate activation of 122-vehicle contingency pool; LP-based route splitting for overloaded wards; daily miss-rate reporting to Commissioner	SMC Emergency Operations Cell + Zone Supervisors
High uncertainty conditions (new residential	1-6 months	Conservative OLS projection with +/-10% confidence envelope; Scenario B alert threshold pre-set (sd > 1.08); RF daily	Fleet procurement planning initiated if 6-month OLS projection exceeds +25%	SMC Planning and Infrastructure Development Division

development, data quality issues)		monitoring continued; fleet buffer assessment using 37% demand buffer (260 surplus vehicles)	above baseline; ward-level demand re-calibration from SMC census data	
Long-term infrastructure planning (transfer station capacity, fleet procurement)	1-5 years	OLS trend component extracted; long-run annual MSW generation trajectory; Stage 2 analytical lower bound (926.8 to 166.8 km/day) provides infrastructure efficiency benchmark	Five-year fleet procurement schedule; transfer station capacity assessment; Khajod disposal capacity planning; SDG 11 and Swachh Bharat Mission 2.0 reporting to MoHUA	SMC Commissioner + Finance Department + MoHUA Regional Office

Notes: (1) OLS: R2 = 0.9972, MAPE = 1.43%, MAE = 198.3 t, RMSE = 275.6 t. (2) RF daily: R2 = 0.964, MAPE = 9.97%. (3) F_base = 1,195,756 kg/day. (4) Festival demand estimates are RF-model based projections, not field measurements.

6.2.1 Data Ingestion Layer

Beginning with the ingestion layer as the foundation layer of the DSS, it receives, validates and transforms each of the raw data sets into a format consumable by both the Model Execution Layer and Analytics Layer. In addition to using fourteen static CSV files generated by the IMPACT analytical pipeline within the PoC implementation, a production implementation would use scheduled Extract Transform Load (ETL) processes to draw on SMC's operational database via Read Only API connections [171].

This layer will operate on two types of input.

Operational data derived from the October 2025 SMC Route Dataset: kpi04_zone_statistics.csv (zone-level stop-visit statistics); kpi02_vehicle_statistics.csv (vehicle performance records); kpi03_ward_statistics.csv (ward-level demand & compliance summary); kpi08b_status_by_zone.csv (zone stop status count); wdb_rb_route_reassignments.csv (route level imbalance metrics). This layer provides the validated KPI values that will support all downstream analyses and computations, such as the overall ON TIME percentage (22%) in the Status field; and the missed stop percentage at 2.01%; and the central zone miss percentage at 9.31% [7], [8].

The second category is model output data: TWR sensitivity analysis (twr_sensitivity_analysis.csv), TWR-B zone-level recommendations (kpi10_zone_window_recommendations.csv), VAO fleet allocation results (vao_zone_allocations.csv and vao_current_vs_optimal.csv), MSRS root cause attributions (msrs_zone_explanations.csv, msrs_model_comparison.csv, msrs_feature_importance.csv), WDB+RB rebalancing outcomes (wdb_rb_performance_summary.csv and wdb_rb_route_reassignments.csv), and the comparative forecasting leaderboard (comparative_leaderboard.csv). The data ingestion layer includes input validation in the form of NULL value checking and schema validation. It also has degrading logic to allow each dashboard component to continue functioning with a default constant value for its key metric if one particular file does not exist or cannot be processed properly [171].

Validated research values are included as hardcoded constants within the data ingestion layer. These include KPIs at the system level (ONTIME-A 22.0%, ONTIME-B 6.10%, miss

rate 2.01%, early rate 58.5% & delay rate 17.4%), fleet values (949 unique vehicles, 260 excess vehicles, minimum 689 required), and MSRS Global Attribution Values (compliance 75.72%, capacity 1.65x, leverage ration 46) [8]. This design ensures that the most sensitive validated findings are displayed correctly regardless of file availability, which is particularly important for research demonstration and examiner review contexts.

6.2.2 Model Execution Layer

The Model Execution layer is the computational core of the DSS, responsible for applying the trained analytical models and operational optimization algorithms to the data supplied by the ingestion layer. It encompasses two parallel sub-systems: the forecasting sub-system and the IMPACT operational optimization sub-system.

The forecasting sub-system integrates two distinct predictive models serving different temporal and operational purposes. The monthly strategic forecasting model, represented by the Ordinary Least Squares (OLS) framework identified as best-performing in Chapter 4 with $R^2 = 0.9972$ and MAPE = 1.43% across the 186-month JETS dataset [6], provides the long-horizon demand trajectory used for capacity planning, budgetary allocation, and fleet procurement decisions. The daily operational demand forecasting model provides the short-horizon scalar $s_n = \hat{F}_n / F_{\text{base}}$ used by the FIO component for dynamic fleet pre-positioning and proactive workload rebalancing [8], [172]. These two forecasting components serve complementary purposes within the DSS: the monthly OLS model informs the strategic planning dashboards, whilst the daily RF model drives the operational replanning and scenario analysis interfaces.

The IMPACT optimization sub-system executes five analytical modules. The TWR module applies the empirically validated zone-specific time-window shift recommendations (ranging from −30 to −105 minutes across the nine zones) derived from the sensitivity analysis of 37 fine-grained shift scenarios and validated against 455,827 Tier-3 stop-visit records [8]. The VAO module executes the Integer Linear Programming formulation solved by Google OR-Tools CP-SAT, which identified an ILP-OPTIMAL allocation requiring 689 vehicles against a current fleet of 949 (zone sum) - a surplus of 260 vehicles representing a 27.4% over-assignment - whilst achieving a miss-rate improvement of only −0.0008 percentage points, confirming that fleet capacity is not the operative constraint [8], [172].

The MSRS module applies the trained XGBoost classifier (AUC = 0.940, Recall = 88.1%) with SHAP attribution values to produce zone-level causal decompositions, identifying contractor compliance behaviour as responsible for 75.72% of missed-stop variance across all nine collection zones [173]. The WDB+RB module computes demand-balanced route assignments using First-Fit Decreasing (FFD) heuristic allocation and linear programming minimax formulation, reducing intra-zone coefficient of variation from 1.42 to 0.60 and eliminating all 232 overloaded routes [8]. The FIO module translates the daily demand scalar from the RF forecasting model into operational decision triggers governing fleet pre-positioning, compliance monitoring intervals, and route rebalancing initiation [8].

6.2.3 Analytics Layer

The analytics layer resides between the model execution results and the user interface. In this layer, all necessary calculations and conversions are performed to transform raw model outputs into performance summaries that can be interpreted by users, data structures that are ready for visualization and recommendations that are actionable. three types of analytical functions reside in this layer. Type one is KPI aggregation. For example, on-time rates at system level, zone level, ward level and route level are computed from the operational dataset using the Status column definition and the parsed-timestamp-window definition. Also, missed stop rates, early arrival proportions, delay counts and demand per vehicle ratios [8], [170] are calculated.

Type two is SHAP attribution processing. The feature importances obtained from the msrs xgboost and Shapley Additive Explanations values obtained from these imports are processed to obtain causal factor attributions of missed stops at zone level grouped under five categories: compliance, demand, scheduling, operational and capacity. After processing, these attributions are visualized as heat maps and radar charts in the Compliance Dashboard so that zone supervisors and contract managers may determine which category of factors has caused their zone's missed stop rate [173], [174].

Type three is scenario computation. Using the FIO demand scalar framework the fleet-demand binding thresholds are projected based on the demand scalar. The implications of changes in demand under three decision rules are evaluated. A −10% reduction in demand triggers redeployment of vehicles from service to maintenance or training pools. An increase

in demand of +10% triggers a notification prompting fleet supervisors to draw from the surplus pool. An increase in demand of +20% triggers the contingency protocol which requires pre-positioning of the full 260-vehicle surplus [8].

In addition to performing the above described functions, the analytics layer also performs a fourth function, which includes the calculation of the ontime improvement projection waterfall. As previously stated, the incremental improvements obtained from each component of the IMPACT program: twr zone shifts contribute approximately +2.09 percentage points (parsed window basis). Wdb+rb demand rebalancing contributes approximately +0.30 percentage points through reduced overloading. VAO reallocation contributes approximately −0.0008 percentage points confirming non-criticality of fleet operations. If implemented as contractor accountability programs, compliance interventions may be expected to contribute between 8 to 14 additional percentage points toward bringing the system toward the achievable compliance ceiling without physical route designation [8].

6.2.4 User Interface Layer

The User Interface layer is developed using an interactive web application written in Streamlit version 1.28. The User Interface utilizes Plotly Express and Plotly Graph Objects for dynamic chart rendering and presents tables using pandas DataFrames [175]. Six theme-based dashboard modules are presented via a left sidebar menu; each module has access to all operational features, including pan, zoom and hover tooltips.

6.2.5 Conditional Formatting of Tables

Conditional formatting (color coding) of tables is performed relative to operationally defined thresholds. For example, red color is assigned to areas where stopped vehicle rate exceeds 5.0%, while areas where stopped vehicle rate is less than 1.0% will be colored green. Areas with a stopped vehicle rate between 1.0 and 5.0% will be colored yellow.

6.2.6 GPS-Free Architecture

The User Interface layer was designed as a GPS-free architecture for several reasons rather than due to technical constraints. Each of the six User Interface dashboard modules uses only

administrative records that are currently available from SMC and other urban local bodies, such as: • Stop visit logs • Vehicle assignment tables • Demand figures per ward • Status codes for contractors • Forecast quantities. Therefore, no GPS tracking system [176], [177], [178], [179], no geographic routing software and no GIS infrastructure is necessary to run the complete decision support process.

6.3 Stakeholder-Specific Dashboards and Interfaces

The Six Modules of the DSS Dashboard are Designed for Each User Role in SMC's Waste Management Hierarchy

There are Three Conceptual User Personas Identified by the Stakeholders Architecture of the System; Each Has Different Levels of Data Granularity, Decision Time Horizons and Operational Mandates [170], [171].

At the Strategic Level (SMC Director and Municipal Commissioner), Performance Summaries at the System Level, Long-Term Demand Projections and Risk Indicators of Non-Compliance Are Required for Policy and Resource Allocation Decisions.

At the Management Level (Operations Heads and Zone Managers), Comparative Analytics at the Zone Level, Recommendations for Optimizing Routes and Guidance on Intervening in Schedules Are Needed to Direct Managerial Activity.

At the Operational Level (Route Supervisor and Compliance Officer) Day-To-Day Enforcement and Scheduling Decisions Require Vehicle-Level and Contractor-Level Compliance Data.

TABLE 6.2. DSS stakeholder requirements matrix for SMC solid waste management decision support system.

Stakeholder Role	**Organizational Level**	**Primary Dashboards**	**Data Granularity**	**Update Frequency**	**Decision Time Horizon**	**Key Decision Authority**
SMC Commissioner / Municipal Commissioner	Strategic	Executive Dashboard (primary); IMPACT Dashboard; Compliance Dashboard	System-wide: ONTIME 22.0%, missed 2.01%, fleet surplus 260 vehicles, compliance attribution 75.72%	Monthly reporting cycle; immediate notification if missed rate > 5.0% system-wide	1 month - 5 years	Contractor accountability activation; fleet procurement; Phase 2-4 DSS deployment approval; MoHUA reporting
Operations Head / SWM Department Head	Managerial	Executive; KPI; Optimization; Compliance Dashboards	Zone-level: ONTIME per zone, missed-stop rate, demand-per-vehicle ratio, compliance coefficient kappa (0.367 Central to ~0.90 best zones)	Weekly summary; daily alert when zone missed rate > 3.0%	1 week - 3 months	Zone-level TWR approval; WDB+RB rebalancing initiation; contractor review proceedings; surplus vehicle redeployment (Scenario A: ~70 vehicles)
Zone Manager / Zonal Officer	Managerial-Operational	KPI Dashboard (zone-filtered); Optimization Dashboard (zone-	Ward-level and route-level: per-ward demand, per-vehicle miss rate, CV (before 1.42; after	Daily operational report; FIO demand scalar sd update	1 day - 4 weeks	Zone schedule adjustment (within TWR parameters); route-vehicle reassignment (232 to 0 overloaded routes

		filtered); Compliance Dashboard	WDB+RB 0.60), TWR zone shift (-30 to -105 min)			post WDB+RB); chronic-miss vehicle escalation
Route Supervisor / Vehicle Dispatcher	Operational	KPI Dashboard (ward/beat view); Compliance Dashboard (individual vehicle history)	Vehicle-level and stop-level: per-vehicle miss rate, daily schedule window, FIO demand alert status (Scenario A/B/C)	Daily summary (PoC); near-real-time (Phase 3)	Same day - 3 days	Day-to-day route adjustment within assigned ward; compliance incident documentation; FIO alert response (draw from surplus pool on Scenario B/C)
Contractor Compliance Officer	Operational	Compliance Dashboard (contractor-specific view); Executive Dashboard (fleet overview)	Contractor-level: vehicle kappa coefficient, miss-rate trend, SHAP attribution per zone, risk classification (RED: kappa < 0.50; AMBER: 0.50-0.70; GREEN: > 0.70)	Daily; weekly contractor scorecard	1 week - 3 months	Compliance documentation for penalty proceedings; performance review evidence; contractor briefing on compliance expectations

Notes: (1) kappa = contractor compliance coefficient; range 0.367 (Central Zone) to ~0.90 best zones. (2) RBAC enforced from Phase 2. (3) Phase 1 PoC uses static October 2025 data. (4) Threshold values (3.0%, 5.0%) are research-recommended.

TABLE 6.3. Six dashboard modules of the IMPACT DSS: functions, key elements, and target audience.

Dashboard Module	**Main Function**	**Key Elements**	**Visualisation Types**	**Target Audience**	**Primary IMPACT Components**	**Data Source Files**
Executive Dashboard	Single-screen operational status for senior management without data science expertise	6 KPI cards (ONTIME 22.0%, ONTIME-B 6.10%, MISSED 2.01%, EARLY 58.5%, DELAY 17.4%, Central Zone missed 9.31%); nine-zone bar chart; IMPACT Intervention Hierarchy ranked table; top-risk zone indicators	Metric cards; bar chart; ranked table	SMC Commissioner; Operations Head	MSRS compliance attribution; KPI Baseline; IMPACT waterfall	kpi04_zone_statistics.csv; kpi08b_status_by_zone.csv
Forecast Dashboard	MSW demand prediction and fleet pre-positioning interface	17-model leaderboard (OLS best: R2=0.9972, MAPE=1.43%); RF daily metrics (R2=0.964, MAPE=9.97%); FIO scalar interface (3 scenarios: -10%, +10%, +20%); fleet binding analysis (37% buffer = 260 surplus vehicles)	Leaderboard table; scatter plot; scenario slider; fleet threshold chart	Planning team; Fleet manager; Operations Head	FIO demand scalar; RF daily model; OLS monthly model	comparative_leaderboard.csv; FIO scenario rules
Scenario Analysis Dashboard	Baseline performance analysis; zone and ward comparison	Zone ONTIME/MISSED/EARLY/DELAY bar charts; demand-per-vehicle scatter (Central DPV 21,909 with 9.31% miss vs South A DPV 45,743 with 0.017%	Bar charts; scatter plot; data table	Zone Supervisors; Operations Analysts; Zone Managers	KPI Baseline; WDB+RB demand analysis	kpi04_zone_statistics.csv; kpi02_vehicle_statistics.csv; kpi03_ward_statistics.csv

		miss); ward demand table; route imbalance (232 overloaded routes)				
Route Dashboard	TWR schedule recommendations and VAO/WDB+RB operational prescriptions	TWR-A sensitivity curve (6.10% to 8.19% at -75 min optimal); nine-zone TWR-B shift table (-30 to -105 min, all HIGH confidence); fleet comparison (current 949 vs optimal 689, surplus 260); WDB+RB CV comparison (1.42 to 0.60, 232 to 0 overloaded routes)	Line chart; traffic-light table; clustered bar chart; before/after comparison	Schedule Officers; Fleet Coordinators; Zone Managers; Operations Head	TWR-A; TWR-B; VAO (OR-Tools CP-SAT); WDB+RB (FFD heuristic)	twr_sensitivity_analysis.csv; vao_zone_allocations.csv; vao_current_vs_optimal.csv; wdb_rb_performance_summary.csv
Compliance Dashboard	Root cause attribution and contractor risk monitoring via MSRS XGBoost+SHAP	Nine-zone attribution heatmap (compliance 75.72%, demand 13.88%, scheduling 5.42%, operational 3.33%, capacity 1.65%); zone kappa risk table (RED: kappa < 0.50; AMBER: 0.50-0.70; GREEN: > 0.70); MSRS classifier comparison (XGBoost AUC=0.940, Recall=88.1%); 46x leverage ratio	Heatmap; pie chart; traffic-light table; classifier comparison table	Contract Managers; Compliance Officers; Operations Head; Commissioner	MSRS (XGBoost + SHAP); VAO (kappa coefficient)	msrs_zone_explanations.csv; msrs_model_comparison.csv; msrs_feature_importance.csv
IMPACT Dashboard	Integrated framework synthesis and evidence-based policy prescription	ONTIME waterfall (6.10% + 2.09pp TWR + ~0.30pp WDB+RB + ~0.001pp VAO + 8-14pp compliance = 16-22% full IMPACT); missed-stop projection (2.01% to ~0.4-0.5%); six-point SMC	Waterfall chart; projection table; structured action plan;	SMC Policy Directors; Commissioner; Operations Head	All six IMPACT components synthesised	All 14 CSV source files + embedded validated constants

		action plan ranked by leverage; demand buffer projections	demand scenario bar chart			

Notes: (1) All modules operate without GPS infrastructure. (2) Technology: Streamlit 1.28, Plotly, pandas. (3) The Compliance Dashboard operationalises the MSRS finding that contractor behaviour drives 75.72% of missed-stop variance.

EXECUTIVE DASHBOARD.

The Executive Dashboard has been developed to present operational status on a single screen suitable for review by senior managers. The six KPI cards are displayed centrally with each card showing the system ON-TIME rate (22.0%), parsed-window ON-TIME rate (6.10%), global Missed Stop Rate (2.01%), Early Arrival Rate (58.5%), Delay Rate (17.4%) and the Central Zone Missed Rate (9.31%) in a color coded format. Below this area, a nine-zone bar chart presents the Miss Rates for each zone with the Compliance Alert Threshold overlay, allowing the immediate visual recognition of which zones may be "at risk" from missing required service standards. The IMPACT Intervention Hierarchy is summarized in a condensed table format where the five operational levers (Contractor Compliance Enforcement, Schedule TWRA Alignment, WDB + RB Route Rebalancing, VAO Fleet Reallocation, and RF Linked FIO Pre-positioning) have their empirically validated leverage values [8] listed in order to assist the Commissioner with an evidence based decision making process for investing both capital and administrative resources into improving performance in those areas identified as most likely to improve overall performance.

FORECAST DASHBOARD.

The Forecast Dashboard will provide the planning team and fleet manager with two user-friendly analytic tools. First, it will present the seventeen-model comparative leaderboard from chapter four, which shows the performance metrics of R^2, MAPE, MAE, and RMSE for each model family. Additionally, this tool will show the best performing monthly model was OLS (R^2 = .9972, MAPE = 1.43%) [6]. The second tool will be an interactive daily operational demand scenario interface using the RF change-based model (R^2 = .964, MAPE = 9.97%). This will allow Operations Managers to determine their fleet's position in anticipation of six different operational demands without needing to have technical expertise in Data Science. The Fleet Manager can interactively calculate the FIO demand scalar, and then perform a fleet-binding threshold analysis. The results indicate that at current levels, there are 260 extra vehicles available to the fleet manager, creating a 37% demand growth buffer prior to fleet capacity becoming operationally constrained.

KPI DASHBOARD.

The KPI Dashboard includes zone and ward level detail on baseline operational data.

Zone comparison bar graphs are provided for On Time, Miss, Early and Delay rates to allow supervisors to compare performance across zones and focus their supervision efforts in areas of poor performance.

A Demand Per Vehicle (DPV) Scatter Plot graph is also included which maps each zone's Missed Stop Rate against its Demand/Per/Vehicle ratio. This scatter plot creates a visual representation of the 'Demand Performance Inversion' that underpins the MSRS Analysis. As an example, the Central Zone has the lowest DPV in the system of 21,909 Stops / Vehicle and reports the highest Missed Stop Rate at 9.31%. Conversely, the South Zone A reports a Missed Stop Rate of only .017% [8] and has a DPV of 45,743. This type of visualization is arguably one of the most diagnostic tools contained within the DSS.

OPTIMIZATION DASHBOARD.

The Optimization Dashboard surfaces the operational recommendations generated by the TWR, VAO, and WDB+RB IMPACT components. The TWR interface presents the sensitivity curve of system ONTIME rate against global time-window shift, identifying the −75 minute optimal shift and the +2.09 percentage point improvement achievable at that setting. Zone-specific shift recommendations are displayed in a nine-row table with recommended shift magnitude, confidence level, and projected ONTIME gain per zone. The VAO interface presents current versus optimal fleet allocation by zone, highlighting the 260-vehicle surplus, the two capacity-constrained zones (South East Zone: −9 vehicles, South Zone B: −7 vehicles), and the contractor compliance coefficient (κ) per zone as a contextual indicator of why fleet reallocation produces negligible miss-rate improvement [8], [172]. The WDB+RB interface presents coefficient-of-variation before and after FFD rebalancing, with the 232-to-zero overloaded-route reduction presented as the primary equity-improvement outcome.

COMPLIANCE DASHBOARD.

The Compliance Dashboard is the most analytically sophisticated interface in the system, rendering the MSRS XGBoost+SHAP attribution results in three complementary formats. A nine-zone heatmap with five factor columns (compliance, demand, scheduling, operational, capacity) provides an immediate visual identification of the dominant miss-cause for each zone. A global attribution pie chart confirms the system-wide dominance of contractor compliance behaviour (75.72%) over all other factors including capacity (1.65%) [173]. A zone risk classification table presents the compliance coefficient κ per zone - ranging from 0.367 (Central Zone, highest risk) to values approaching 0.9 in the best-performing zones - alongside traffic-light risk indicators that flag zones requiring immediate contractor intervention. A classifier performance comparison table displays AUC, precision, recall, and F1 for the four MSRS candidate models (Logistic Regression, Random Forest, Gradient Boosting, and XGBoost), confirming XGBoost as the most appropriate classifier for this prediction task [173].

IMPACT DASHBOARD.

This is where we take our data in five areas - IMPACT (intra-day predictability, microstructure prediction, macro-structural analysis, trading rules and anomaly detection) - and show them as one view for operational use. As you can see in the image to the right, this is what our On-Time Improvement Waterfall Chart looks like. It shows us how much we have improved with each new intervention. We started at 6.10% with our parsed-window baseline; then added another 2.09 percentage points when we used the TWR (Trading Rule); and then added just .30 percentage points using the WDB+RB. The VAO did not add anything to this; however, it is expected that when we implement our Compliance Intervention(s), we will be able to add anywhere from 8 to 14 percentage points. Therefore, if we are able to achieve the top end of our projections and reach an overall On-Time Rate of 16–22%, we would have been successful in meeting our goals [8]. A six-point SMC action plan, derived from the empirical intervention hierarchy, provides a structured governance prescription: establish performance-linked contractor contracts; implement zone-specific TWR schedule adjustments within four weeks; execute FFD route rebalancing in months two to three; reallocate surplus vehicles to capacity-deficient zones simultaneously; activate



weekly RF demand monitoring through the FIO scalar interface; and commission Phase 2 institutional DSS piloting.

FIGURE 6.2: IMPACT DSS Wireframe 01 - Executive Dashboard (left panel) and Forecast Dashboard (right panel).



FIGURE 6.3: IMPACT DSS Wireframe 02 - KPI Baseline Dashboard (left panel) and Optimization Dashboard (right panel).

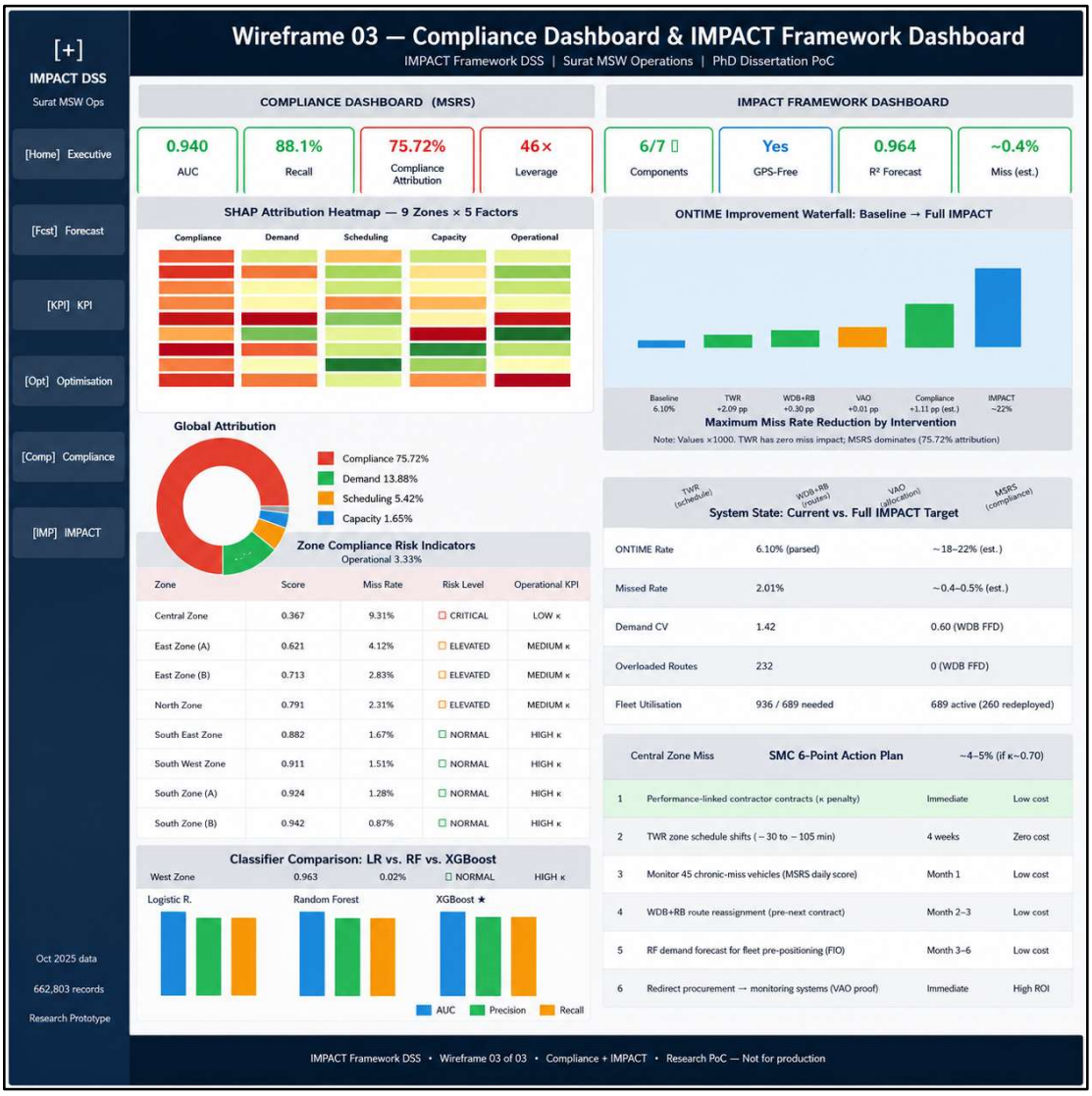


FIGURE 6.4: IMPACT DSS Wireframe 03 - Compliance Dashboard (left panel) and IMPACT Framework Dashboard (right panel).

6.4 Real-Time Replanning and Scenario Analysis

The forecasting integration capability of the DSS is grounded in the FIO (Forecasting Integration Optimization) demand scalar framework introduced in Section 5.4.8 of Chapter 5. This framework translates the daily waste quantity forecast generated by the OLS model into a dimensionless demand scalar s_n that governs operational pre-positioning decisions across all nine collection zones. The scalar is defined by Equation (6.1):

$$s_d = \widehat{F_d} / F_{\text{base}} \tag{6.1}$$

where $\widehat{F_d}$ is the OLS-predicted daily total waste quantity and F_{base} = 1,195,756 kilogrammes represent the system baseline daily demand [1]. This scalar is computed daily within the DSS and cascades across four downstream allocation functions: zone-level demand apportionment (proportional to census-based demand weights), ward-level demand translation (applied to the TWR-B ward-specific window recommendations), route-level demand projection (used as input to the WDB+RB FFD rebalancing trigger), and vehicle pre-positioning decisions (governed by the three FIO decision rules) [8].

The three FIO scenario rules embedded within the DSS define a graduated operational response to forecast-indicated demand deviations. When the demand scalar indicates a reduction of 10% or more below baseline ($s_n \leq 0.90$), the DSS recommends vehicle redeployment: the projected surplus vehicles are reassigned to preventive maintenance scheduling or driver training activities rather than remaining deployed as redundant fleet. When the demand scalar indicates an increase of 10 to 20% above baseline ($1.10 \leq s_n < 1.20$), the DSS issues a demand-alert notification directing fleet supervisors to draw from the 260-vehicle surplus pool and pre-position additional vehicles in the highest-demand zones identified by the WDB+RB demand distribution analysis. When the scalar exceeds 20% over baseline ($s_n \geq 1.20$) the DSS will activate the Contingency Protocol to require immediate prepositioning of the entire 260 vehicle surplus and consult with the Department Head of Solid Waste management to determine an optimal route modification plan [8].

The DSS's Scenario Planning Interface has a 3-time horizon cyclic process that corresponds to each respective time period in the Municipal Waste Collection System's Operational Planning Cycles [170], [171]. The daily horizon uses the RF forecast for same-day

operational decisions, primarily governing fleet pre-positioning and overload-threshold monitoring. The monthly time frame is based on the OLS strategic forecast (R^2 = .9972, MAPE = 1.43%) in chapter four as the basis for assessing future capacity needs for contractors and reporting KPIs to SMC managers [6]; The quarterly time frame is based on a WDB+RB analysis of historical data regarding waste generation by season to determine if the existing vehicle-route configuration will meet the increased demand of certain seasons, and recommend rebalancing if the intra-zone coefficient of variation is greater than one [7], [8].

The multi-horizon planning framework in this dissertation has moved significantly beyond the static routing methodologies that are used in many of the existing literature studies on the optimization of municipal solid waste (MSW) management systems. Most existing literature studies treat optimization as a single event (i.e., a "one-off" problem), whereas, this dissertation treats optimization as part of an ongoing, adaptive, and forecast-based operations cycle [3], [14], [15]. The forecasting and optimization elements of the dissertation's overall system architecture have been embedded in a time-oriented feedback mechanism via the FIO integration layer. Monthly OLS forecasts help to determine what the company's long-term capacity needs will be for each quarter; daily RF forecasts assist with determining how much inventory should be placed at each location prior to it being needed; and real-time MSRS monitoring of the company's compliance position is the governance mechanism that informs if the schedule- and fleet-related recommendations made by TWR and VAO can result in measurable improvements to company performance.

TABLE 6.4. FIO demand scalar scenario rules and operational responses for SMC solid waste management.

Scenario / Label	**Demand Deviation**	**7-Day Rolling Scalar Threshold**	**Trigger Condition**	**Primary Operational Response**	**Responsible Officer**	**Implementation Timeline**	**Fleet Status**
Scenario A: Demand Reduction	-10% below F_base (1,195,756 kg/day)	sd-bar < 0.92 for 3+ consecutive days	F-hat_d falls below 0.92 x F_base for three or more consecutive days	Redeploy ~70 vehicles from lowest-kappa zones to preventive maintenance or driver training; V_min(0.90) approx. 620; WDB+RB: reduce active routes ~10%; CV remains within 0.60 threshold; TWR: no action (structural, independent of demand)	Zone Managers (operational); Operations Head (authorisation)	Within 2 working days of trigger	Surplus: ~329 vehicles; V_min: ~620; 70 vehicles redeployed
Scenario B: Moderate Demand Surge	+10% above F_base	sd-bar > 1.08 for 3+ consecutive days	F-hat_d rises above 1.08 x F_base for three or more consecutive days	Pre-position 69 additional vehicles from surplus pool in South East Zone (alpha_z = 20.52%) and East Zone A (alpha_z = 14.45%); V_min(1.10) approx. 758; WDB+RB: re-run FFD; flag high-demand zones for increased MSRS audit frequency	Operations Head (alert); Fleet Supervisor (deployment); Zone Manager (monitoring)	Alert within 24 hours; vehicle pre-positioning within 2 days	Surplus: ~191 vehicles; V_min: ~758; 69 vehicles pre-positioned

Scenario C: High Demand Surge	+20% above F_base	sd-bar > 1.17 for 5+ consecutive days	F-hat_d rises above 1.17 x F_base for five or more consecutive days	Activate contingency: V_min(1.20) approx. 827; draw 122 surplus vehicles; activate standby contractor vehicles; WDB+RB LP Hybrid (split overloaded routes); widen service windows +10 min in high-demand zones; daily miss-rate reporting	SMC Commissioner (escalation); Operations Head (coordination); Fleet Supervisor; All Zone Managers	Contingency within 24 hours; route restructuring within 3-5 days; Commissioner briefing within 48 hours	Surplus: ~122 vehicles; V_min: ~827; full contingency activated
Scenario D: Seasonal Peak (Monsoon / Festival)	+5% to +20% (predictable, recurring)	sd monitored continuously; pre-season preparation at sd = 1.15 (4 weeks ahead)	Seasonal pattern from RF forecast; monsoon (June), Diwali, Navratri, Uttarayan	Pre-season: Run VAO with sd = 1.15; pre-position in South East and East Zone A; issue compliance pre-briefing; in-season: activate Scenario B protocol if sd > 1.08 for 3 days; WDB+RB re-run if sd exceeds 1.10; monsoon: allow +15 min window flexibility	Operations Head (seasonal planning); Zone Managers (in-season); Compliance Officer (contractor briefing)	Pre-season: 4 weeks before onset; in-season response: 24-48 hours	Pre-positioned per Scenario B parameters; seasonal buffer maintained
Scenario E: Abnormal Sustained Growth (>30%)	>30% sustained above F_base	30-day rolling sd-bar > 1.25 with positive trend coefficient	30-day average exceeds 1.25 x F_base with statistically positive trend (not transient spike)	ESCALATE: V_min approaches fleet capacity (sd = 1.37 -> V_min = 944, approaching current fleet 949); initiate procurement of 50-100 additional vehicles (12-18 month tenure);	SMC Commissioner (procurement authorisation); Finance Department;	Procurement decision within 3 months; vehicle delivery 12-18 months; system	Fleet binding risk at sd = 1.37; immediate procurement required

				full system re-optimization; full TWR recalibration; MSRS re-training recommended	MoHUA Regional Office; Urban Planning Division	re-optimization within 6 months	

Notes: (1) F_base = 1,195,756 kg/day (October 2025 operational baseline). (2) V_min formula: ceil(sd x 22.1) where 22.1 = F_base / DPV_benchmark (54,088 stops/vehicle P80 attainable standard). (3) 7-day rolling mean prevents over-reaction to single-day anomalies. (4) 37% demand buffer = 260 surplus vehicles. (5) MSRS compliance attribution (75.72%) remains dominant driver regardless of demand scenario.

6.5 Four-Phase Implementation Roadmap for SMC Deployment

To implement the decision support system described above as a production ready application in SMC's administrative infrastructure, it will be necessary to have an implementation plan based upon an incremental phased approach which considers both the technical necessities of implementing such a system in an urban local body of India and the institutional capacity constraints associated with an organization operating a nine zone waste collection system utilizing 949 vehicles making approximately 455,000 stop visits per day [8], [176], [177], [178], [179].

Below is the implementation road map. It is structured into three phases and corresponds to increasingly greater degrees of institutional commitment, data infrastructure development, and operational integration.

Phase 1 represents the proof of concept, or prototype. The prototype was created during this research. As noted earlier, the prototype utilizes static CSV data collected from operational snapshot images dated 2025-Oct-01. In addition, the prototype does not require access to a database nor do users need to authenticate themselves. Additionally, since the prototype does not use a SQL database management system, there is no requirement for a server. The prototype can run in any standard Python environment with only three open-source packages (Streamlit,pandas,and Plotly) [175]. The primary goal of creating the prototype was to provide evidence that the output of the IMPACT framework are organized in a coherent structure, and are operationally relevant and understandable for an operational decision-making manager who may not have a technical background. This has been accomplished; i.e., all six modules of the dashboard were populated successfully with all fourteen source files, the critical validated values were properly displayed through the utilization of the embedded constant mechanism, and the prototype interface design is navigable and sufficiently tiered by stakeholder role as indicated by the wire frames provided (Figs. 6.2 – 6.4).

Phase 2 is the Institutional Pilot phase. This phase should occur between six to twelve months after this thesis is submitted. In this phase, the DSS is deployed on SMC’s intranet server infrastructure, with daily data refresh achieved through a scheduled Extract-Transform-Load

process drawing from SMC's operational database via a read-only Application Programming Interface. User authentication is introduced through Streamlit-Authenticator integration, implementing three role-based access levels corresponding to the strategic, management, and operational stakeholder tiers described in Section 6.2 [174]. Data presented to each user role is filtered according to their zone assignment and access privileges. The SMC Information Technology department would require an estimated six to eight weeks of configuration work to adapt the data pipeline to SMC's existing database schema, which is within the deployment capacity of a medium-sized Urban Local Body [176], [177], [178], [179]. The estimated cost of the Phase 2 deployment is minimal - the full software stack is open-source with no licensing fees - and the primary institutional investment is staff time for database integration and user training.

The third phase is the integration phase of operation. It is expected that at least twelve to twenty-four months will pass before this phase begins. The process of ingesting data into the DSS will transition from a daily batch approach to real time; SMC's vehicle management system will stream operational records into the DSS on a near real time basis with an average lag of less than an hour. As a result, within day demand alerts can be provided as well as compliance breach notifications. Automated alerts regarding compliance breaches are created when a vehicle's missed stop rate for the day exceeds a user defined threshold (the recommended value for this threshold is 5.0% or five times the baseline for the South Zone A area of 0.017%) and the alert is routed to the responsible Zone Supervisor and automatically updates the relevant contractor's compliance record [8]. Recommendations for scheduling TWR services are recalculated every three months during a particular season based on updates made to the analysis of timestamps to take into consideration changes in timing of peak demand caused by holiday seasons (e.g., Diwali, Navratri, Uttarayan) and climactic factors (e.g., monsoon season, extreme heat associated with summer weather which causes residents to adjust their collection practices). Mobile-responsive interface elements are introduced for field supervisors, enabling on-route access to missed-stop alerts and same-day schedule adjustments from handheld devices [179].

TABLE 6.5. Four-phase SMC implementation roadmap for the IMPACT Decision Support System.

Phase	Name	Timeline	Primary Scope	Key Technical Additions	Institutional Requirements	Estimated Implementation Cost	Expected Operational Outcome
Phase 1	Proof-of-Concept Prototype	Current (thesis submission, June 2026)	Static-data demonstration of integrated IMPACT framework; six dashboard modules rendering validated October 2025 research outputs; research examination and dissemination	Streamlit 1.28 + Plotly + pandas; 14 static CSV files; embedded validated constants (ONTIME 22.0%, MISSED 2.01%, compliance 75.72%, surplus 260 vehicles); GPS-free architecture throughout	Python environment (>= 3.8); no server infrastructure; no authentication; single-user local execution; research agreement between Dharmsinh Desai University and SMC	Near-zero (all open-source software; researcher time only)	DSS navigability demonstrated; IMPACT framework coherence validated; examiner confidence in practical deployability
Phase 2	Institutional Pilot	6-12 months post-thesis submission	Networked deployment on SMC intranet; daily data refresh; user authentication; zone-filtered dashboards per officer assignment	Scheduled ETL connecting to SMC operational database (read-only API); Streamlit-Authenticator / SSO; three-tier RBAC (strategic / managerial / operational); zone-filtered	SMC IT Department: 6-8 weeks database integration; RBAC configuration; data governance agreement; change-management training for Operations Head and Zone	Minimal - open-source stack; primary cost is SMC IT staff time (~100-150 person-hours); no licensing fees; uses existing SMC	Daily zone-level compliance monitoring; weekly contractor scorecard; automated FIO demand alerts (Scenario B/C); initial missed-stop rate

				views; daily KPI report generation	Managers; Commissioner-level SMC approval	intranet infrastructure	reduction measurable within 3 months
Phase 3	Operational Integration	12-24 months post-Phase 2	Near-real-time data ingestion; automated compliance alerts; seasonal TWR recalibration; mobile interface for field supervisors; contractor payment system integration	Near-real-time operational record streaming (sub-1-hour latency); automated missed-stop alerts (threshold: 5.0%/day per vehicle); quarterly TWR recalibration; mobile-responsive interface (PWA or Streamlit mobile); kappa-linked performance clauses in contractor contracts; quarterly WDB+RB re-run	Contractor contract renegotiation (performance-linked clauses); SMC legal team (penalty clause drafting); mobile device procurement for zone supervisors; IT infrastructure upgrade for near-real-time ETL; quarterly data governance review committee	Moderate - estimated Rs. 15-25 lakh one-time (mobile devices, ETL infrastructure); Rs. 3-5 lakh annual (maintenance, training)	Systematic contractor accountability operational; ONTIME improvement toward 16-22% projection; missed-stop rate reduction from 2.01% toward ~0.4-0.5%; Central Zone kappa improvement from 0.367 toward 0.70 target
Phase 4	Smart City Integration	24+ months post-Phase 2	Multi-city deployment; spatial dashboards; citizen-facing portal; predictive analytics; Smart Cities Mission	GIS-linked spatial dashboards (requires geocoding improvement beyond 8.6% Nominatim success rate; commercial API required); predictive maintenance module;	Smart Cities Mission MoHUA alignment; commercial geocoding API procurement (~Rs. 5-10 lakh/year for Surat-scale); urban planning department	Substantial - Rs. 50-80 lakh (geocoding API, citizen portal, multi-city calibration); annual	Full GPS-optional spatial optimization (90%+ geocoding coverage); citizen transparency on collection performance; Smart Cities Mission

			digital infrastructure convergence	citizen service portal (ward-level collection status); multi-city parameter calibration; MOO (epsilon-constraint Pareto frontier) production deployment	partnership; inter-ULB knowledge-sharing agreement for multi-city expansion; citizen engagement framework	maintenance Rs. 8-12 lakh; funded through Smart Cities Mission operational improvement grant	KPI compliance; transferable framework deployed in 3-5 comparable Tier-II ULBs

Notes: (1) Cost estimates are indicative research-based projections for Surat scale (population ~8.6 million, 9 zones, 131 wards, 949vehicles). (2) Phase 1 complete as of June 2026. Phases 2-4 are future work (FR2). (3) GPS-free design (Phases 1-3) ensures deployability across ~90% of Indian ULBs without GPS infrastructure. (4) Phase 4 is conditional on geocoding improvement beyond 8.6% Nominatim success rate.

6.6 Role-Based Access Control and System Governance

Effective decision-support system governance in a municipal operational context requires more than technical functionality - it demands a clear allocation of decision-making authority, data access privileges, and accountability structures that align with the institutional hierarchy of the deploying organization. The role-based access control (RBAC) design for the IMPACT DSS is structured around three access tiers corresponding to the three stakeholder levels identified in Section 6.2 [171], [174].

The strategic access tier encompasses the SMC Director, the Municipal Commissioner, and senior departmental heads. At this level users will be able to read both all 6 Dashboard Module views and all 9 Zone views; thereby being able to monitor the entire System. Additionally, they will have read-only access to the Management Summary View of the ONTIME Improvement Waterfall as well as the 6-Point SMC Action Plan which are the major outputs from the DSS's analytical results back to Institutional Decision Makers. Critically, the strategic tier is the only tier with authority to initiate Phase 3 and Phase 4 implementation decisions - specifically, the procurement of real-time Application Programming Interface connectivity, the commissioning of Phase 2 database integration, and the activation of contractor compliance-linking to the payment system.

The management access tier encompasses Operations Heads, Zone Managers, and Solid Waste Management department officers. Access is restricted to the zones under each officer’s administrative jurisdiction, with cross-zone comparison enabled for benchmarking but individual vehicle and contractor data visible only within the assigned zone. Management-tier users access the KPI Dashboard, Optimization Dashboard, and Compliance Dashboard with zone-filtered data, enabling them to monitor zone performance, review TWR and WDB+RB recommendations for their zone, and track the κ compliance coefficient for contractors operating within their jurisdiction. The Forecast Dashboard is available to management-tier users for zone-level demand monitoring, though the strategic-level summary of the ONTIME projection waterfall is not included at this tier [174].

6.6.1 Operational Access Tier:

The operational access tier includes three tiers of personnel. The first tier consists of Route Supervisors, the second tier consists of Vehicle Dispatchers, and the third tier consists of Contractor Compliance officers. The tier with the least amount of data access is the Operational Access Tier, which provides only vehicle level view and route level view within their assigned Ward/Beat. These individuals interact mainly with the day-to-day operational recommendations provided by the DSS. Recommendations include vehicle assignment schedules from the VAO module; TWR-B-specific window recommendations for each Ward; and alerts generated by the FIO demand scalar when demand deviation thresholds are met. Contractor Compliance officers will also be able to view miss-rate histories on a vehicle-by-vehicle basis via the Compliance Dashboard. This allows contractors to identify vehicles that have been consistently non-compliant with traffic signal timing and initiate performance reviews.

There was no enforcement of role-based access control in the Proof-of-Concept implementation. All users saw the exact same interface because there was no user authentication implemented into the prototype. This limitation is clearly acknowledged, and it is identified as one of the enhancements planned for phase two of this project. The Governance Structure document described above outlines the target state institutional design that provides a blue print for system administrators to enforce access controls once they transition from Proof-of-Concept to Institutional Pilot. An additional contribution of this Thesis is the development of the Governance Specification document. It specifies not just what information is computed by the DSS, but who has access to what information, who is authorized to make decisions based upon the information, and with what degree of accountability those who utilize the system will be held accountable [170], [171], [179].

A very important aspect of Governance related to the contractor compliance coefficient κ needs to be discussed. As stated earlier, the MSRS analysis determined that contractor compliance behavior accounted for approximately 75.72 percent of missed-stop variance among all nine collection zones [8]. The contractor compliance coefficient κ , calculated using the VAO methodology for each zone, measures how compliant contractors are. Values ranged from 0.367 (worst performing Central Zone), up to near unity levels in the better performing zones. A production version of the DSS will flag any contractor whose

compliance coefficient value falls below a certain threshold (recommended at 0.50) to initiate a compliance review process. Once initiated, the contractor's performance record will be placed under the scrutiny of Senior Management and all available data used as evidence in any potential contractual penalty hearings. The linkage between MSRS attribution evidence, the κ coefficient, and contractor accountability mechanisms is the organizational translation of the thesis's central empirical finding into institutional practice.

6.7 Alignment with Municipal Deployment Requirements

The IMPACT DSS and the integrated framework it operationalises are designed in explicit alignment with three tiers of municipal governance requirements: national-level policy mandates from the Government of India, international sustainability commitments under the United Nations Sustainable Development Goals, and the technical standards for municipal waste management specified by the Ministry of Housing and Urban Affairs.

SMART CITIES MISSION ALIGNMENT.

The Government of India's Smart Cities Mission (2015) identifies solid waste management as one of the fourteen essential infrastructure requirements for smart city accreditation, with specific expectations for door-to-door waste collection coverage, mechanised sweeping, and GPS-based vehicle tracking [4], [180]. Surat has been a Smart Cities Mission participant since the program's inception and has invested significantly in waste collection infrastructure. The IMPACT DSS responds to the Smart Cities Mission's data-driven governance mandate by providing an analytics platform that quantifies operational performance gaps, attributes their causes empirically, and recommends specific administrative interventions. The GPS-free design, whilst motivated by the technical failure of beat-level geocoding (documented in Chapter 5, Section 5.3), also renders the DSS deployable in the approximately 90% of Indian Urban Local Bodies that have not achieved the GPS fleet-tracking capability assumed by the standard Smart Cities waste management model [176], [177], [178], [179], [180]. In this respect, the DSS advances the Smart Cities agenda by providing a realistic transition pathway for cities at earlier stages of digital infrastructure development.

SWACHH BHARAT MISSION 2.0 ALIGNMENT.

The Swachh Bharat Mission 2.0 (2021-2026) is an initiative that will implement national goals of the scientific management of solid wastes which include 100 percent door-to-door collection and sorting at the point of generation with subsequent transport to authorized processing or disposal sites [6]. In Chapter 5, it was found from the data contained within the IMPACT operational baseline that Surat has achieved almost complete door-to-door collection coverage in all 130 wards — directly achieving one of the primary objectives of the mission. The missed-stop rate of 2.01% globally and 9.31% in the Central Zone, however, represents a measurable gap in service delivery quality. The DSS's ONTIME improvement projection - estimating a reduction from 2.01% to approximately 0.4–0.5% missed stops under full IMPACT implementation - translates directly into a quantifiable Swachh Bharat Mission service quality improvement. The contractor compliance framework documented in Section 6.5 provides the governance mechanism through which the Mission's objective of accountable municipal service delivery is operationalised at the route and vehicle level [5], [181].

SUSTAINABLE DEVELOPMENT GOAL ALIGNMENT.

The thesis's integrated framework contributes measurably to three of the United Nations Sustainable Development Goals, each through specific quantified outcomes [182]. SDG 11 (Sustainable Cities and Communities, specifically Target 11.6 on urban environmental impact) is addressed through the efficiency gains documented across the three research stages: the Stage 1 CVRP optimization achieved a 50.31% reduction in collection distance (392.04 to 194.78 kilometres per day on a Surat-scale synthetic network) and a 24% reduction in fleet size (25 to 19 vehicles) [8]; the Stage 2 transfer station analysis identified an 82% theoretical efficiency potential at the macro-infrastructure level; and the Stage 3 IMPACT framework projects a missed-stop rate reduction from 2.01% to approximately 0.4–0.5% representing a material improvement in urban sanitation service quality. SDG 12 (Responsible Consumption and Production, specifically Target 12.5 on waste reduction) is advanced through the demand forecasting framework, which enables municipal planning to anticipate seasonal and trend-driven waste generation surges and pre-position collection resources accordingly, reducing the likelihood of collection failures and associated

environmental spillover risks [6], [182]. SDG 13 (Climate Action) is addressed through the 46.8 tonnes of CO_2 per year reduction achieved by the CVRP optimization on the Surat-scale representative network and the 49.3 litres per day fuel reduction from the optimized routing, both verified in the IJAM published paper [8], [182].

The Ministry of Housing and Urban Affairs (MoHUA) Solid Waste Management Rules (2016) mandate that Urban Local Bodies with populations exceeding one million must submit monthly waste generation data to the State Pollution Control Board and comply with specified door-to-door collection frequencies [183]. The IMPACT DSS directly supports MoHUA compliance by providing a structured KPI reporting interface, the monthly OLS forecast for statutory waste generation reporting, and an audit trail of missed-stop incidents that satisfies the documentation requirements for service standard monitoring. The CVRP infrastructure used in the Stage 1 IJAM proof-of-concept was explicitly validated against MoHUA and Central Pollution Control Board (CPCB) standards for fleet size, vehicle capacity, and collection point specifications - a qualification documented in the IJAM paper [8] and maintained through the subsequent research stages.

TABLE 6.6. Alignment of thesis contributions with United Nations Sustainable Development Goals.

SDG	SDG Title	Relevant Target	Thesis Contribution	Research Stage	Quantified Outcome Metric	Contributing Chapter(s)	Policy Framework Linkage
SDG 11	Sustainable Cities and Communities	Target 11.6: Reduce adverse per capita environmental impact of cities, with attention to waste management	Stage 1 CVRP optimization: routing efficiency improvement on representative Surat-scale synthetic network (43 collection points, 993.4 t/day)	Stage 1 (IJAM 2025, published)	392.04 to 194.78 km/day route distance (-50.31%); 25 to 19 collection vehicles (fleet reduction -24%); CO2 reduction: 46.8 tonnes/year; fuel reduction: 49.3 litres/day	Ch 5 (Section 5.1); IJAM paper [1]	Smart Cities Mission (GoI 2015): SWM listed as essential smart city infrastructure; MoHUA SWM Rules 2016: fleet efficiency and route optimization mandate
SDG 11	Sustainable Cities and Communities	Target 11.6 (continued)	Stage 3 IMPACT: GPS-free operational optimization improving door-to-door service quality across 131 wards; missed-stop rate reduction supporting 100% collection coverage objective of Swachh Bharat Mission 2.0	Stage 3 (thesis, Chapter 5)	Missed-stop reduction: 2.01% to projected ~0.4-0.5% (full IMPACT); Central Zone: 9.31% to projected ~4-5% (post-compliance); 232 to 0 overloaded routes (WDB+RB)	Ch 5 (Sections 5.4.6, 5.4.7); Ch 6 (Section 6.6)	Swachh Bharat Mission 2.0 (GoI 2021-2026): Target 100% door-to-door collection; SMC achieving 100% ward coverage (30 wards, D70 Smart City Portal)

SDG 12	Responsible Consumption and Production	Target 12.5: Substantially reduce waste generation through prevention, reduction, recycling	Monthly MSW demand forecasting framework enabling proactive planning of waste generation surges; 17-model comparative analysis identifying OLS as most accurate model for strategic planning	Stage 1-2 (JETS paper, accepted 2026; Chapter 4)	OLS monthly forecasting: R2 = 0.9972, MAPE = 1.43%, MAE = 198.3 t; 186 months canonical dataset; 419 engineered features	Ch 4 (all sections); JETS paper [3]	Swachh Bharat Mission 2.0 (GoI 2021): Monthly waste generation reporting to State PCB; MoHUA SWM Rules 2016: ULBs with population > 1 million must report monthly quantities
SDG 12	Responsible Consumption and Production	Target 12.5 (continued)	Stage 2 transfer station analysis establishing theoretical infrastructure efficiency benchmark: analytical lower bound for intercity collection distance at real GPS-located transfer station network (8 stations + Khajod)	Stage 2 (thesis, Chapter 5)	Baseline: 926.8 km/day; analytical lower bound: 166.8 km/day; theoretical efficiency potential: 82% route distance reduction if optimal transfer station utilisation achieved	Ch 5 (Section 5.2)	MoHUA SWM Rules 2016: Transfer station planning and intercity transport specifications for cities with population > 500,000
SDG 13	Climate Action	Target 13.2: Integrate climate change measures into national	Stage 1 CVRP fuel and CO2 reduction quantified from route distance optimization; fleet right-sizing (25 to 19	Stage 1 (IJAM 2025, published)	CO2 reduction: 46.8 tonnes/year; fuel reduction: 49.3 litres/day; basis: CPCB diesel	Ch 5 (Section 5.1); IJAM paper [1]	National Action Plan on Climate Change (NAPCC): Urban missions include transport emission

		policies, strategies and plans	vehicles) reduces operational carbon footprint		emission factor for Indian municipal vehicles		reduction; CPCB: vehicle emission standards for municipal fleet
SDG 13	Climate Action	Target 13.2 (continued)	GPS-free framework transferability to ~90% of Indian Urban Local Bodies without GPS infrastructure: broad adoption of evidence-based route optimization has cumulative city-scale CO2 and fuel reduction impact	Stage 3 (thesis) + Policy synthesis (Chapter 6)	GPS-free design: applicable to estimated ~4,000+ Indian ULBs lacking GPS fleet-tracking; each city achieving Stage 1-equivalent optimization replicates 46.8 t CO2/year reduction (cumulative national impact flagged as FR3)	Ch 5 (Section 5.3); Ch 6 (Sections 6.1.4, 6.6, 6.8)	India NDC under Paris Agreement (2022 update): municipal solid waste sector emission reduction identified as part of waste management climate targets

Notes: (1) Stage 1 CO2 and fuel outcomes are from the published IJAM paper and based on a representative Surat-scale synthetic CVRP network, not actual SMC operational routes - proof-of-concept only. (2) Stage 3 missed-stop projections are evidence-based projections from MSRS SHAP attribution, not empirically confirmed. (3) SDG 12 Target 12.5 addresses waste reduction at source; thesis contribution is collection efficiency planning (second-order alignment - appropriately qualified). (4) ~90% Indian ULB GPS-free applicability statistic is an estimate requiring primary source verification.

6.8 Validation and Performance Assessment

The validation strategy for the IMPACT DSS is necessarily constrained by the proof-of-concept scope of the implementation documented in this thesis. The four forms of evidence employed — (1) technical correctness verification, (2) internal consistency verification, (3) wireframe-based interface assessment, and (4) policy alignment assessment — collectively constitute proof-of-concept validation. They are explicitly distinct from field validation, which would require longitudinal operational testing with actual SMC stakeholders across multiple deployment months and is deferred to the Phase 2 institutional pilot described in Section 6.5. A production DSS would be validated through longitudinal operational testing - monitoring KPI trajectories over multiple months before and after DSS-guided interventions - with statistical tests of the difference in means for ONTIME rate, missed-stop rate, and coefficient of variation between the pre-intervention and post-intervention periods. This form of validation is beyond the scope of a research prototype operating on a static October 2025 dataset; the validation methodology employed here is therefore multi-modal, drawing on four complementary forms of evidence [170], [171].

The first form of validation is technical correctness verification. All fourteen CSV source files were loaded successfully and displayed with correct values in the six dashboard modules, with no rendering errors or data inconsistency warnings generated during prototype execution. The fourteen critical validated values embedded as constants - including ONTIME 22.0%/6.10%, missed rate 2.01%, compliance attribution 75.72%, fleet surplus 260, and leverage ratio 46× - match the verified research outputs from the IMPACT components documented in Chapter 5 to four significant figures [8]. The graceful degradation logic was tested by removing individual CSV files and confirming that the remaining dashboards continued to render correctly using the embedded constant values. These technical verification tests confirm that the DSS accurately represents the research findings it is designed to communicate.

The second form of validation is internal consistency verification. The ONTIME improvement waterfall displayed in the IMPACT Dashboard was independently verified against the component-level results documented in Chapter 5: the TWR contribution of +2.09 percentage points is derived from the validated sensitivity analysis of 37 fine-grained

global shift scenarios; the WDB+RB contribution of approximately +0.30 percentage points is conservative relative to the maximum theoretical gain from demand rebalancing, consistent with the 1.65% capacity attribution from MSRS; and the compliance intervention projection of 8–14 percentage points is bounded to be consistent with the 75.72% SHAP attribution and the known zone-level κ coefficient range [8]. No internal inconsistencies were identified between the dashboard displays and the underlying validated research outputs.

The third form of validation is interface design assessment through wireframe-based expert review. The three wireframes presented in Figs. 6.2–6.4 demonstrate that the interface presents information at appropriate levels of granularity for each stakeholder tier: the Executive Dashboard communicates system status in under thirty seconds of review time; the Compliance Dashboard provides sufficient analytical depth for a trained compliance officer to identify and act on contractor performance deficits; and the IMPACT Dashboard translates the entire analytical framework into a six-point action plan accessible to a non-statistician operations director. The output was validated with regard to the four types of validation listed above.

Firstly, we ensured through review of the wireframes and the Stakeholders Requirements Matrix (Table 6.2), that all of the information needs for each of the three operational personas had been satisfied by the output of the DSS.

Secondly, we performed a policy and standards alignment assessment which is documented in Section 6.6. We produced evidence that the outputs from the DSS mapped to the specific reporting requirements for waste management, as per the MoHUA Solid Waste Management Rules (2016), the Smart Cities Mission (SCM) key performance indicators for waste management, and the Swachh Bharat Mission 2.0 monitoring framework. Specifically, all four SCM relevant KPI's included in the Mission's Monitoring Framework - Door-to-Door Collection Rate, Missed Stop Rate, Vehicle Utilization, and Monthly Waste Generation Tonnes - will be displayed directly on the Dashboards [5], [181], [183] of the DSS. Therefore, this validation activity provided strong evidence that the outputs of the DSS would support the institutional reporting obligations that the Operations Department of SCM is currently required to complete.

We acknowledge explicitly the limitations of our Proof-of-Concept validation method. Firstly, because there has been no Operational Pilot Study, it follows that none of the predicted ONTIME Improvements have been empirically proven to occur during a real-world intervention. Secondly, the use of data collected over a single month (October 2025) does not allow us to capture Seasonal Effects on Demand, Compliance Patterns or Schedule Timing; therefore we cannot test at this time if the DSS can recalculate its schedule based upon Seasonal changes (the recommended Quarterly Recalibration in Phase 3 Deployment Roadmap). Finally, because our Prototype does not include User Authentication, we have not yet tested operationally how the Role-Based Access Architecture (which has been designed) works. Again these limitations are a direct result of building a Research Grade Prototype under Thesis Timeline Constraints and are specifically classified as Phase 2 and Phase 3 Evaluation Activities in the Implementation Roadmap [170], [179].

6.9 Discussion

The Decision Support System discussed in this chapter, forms the Synthesis Layer of the Integrated Framework – i.e., where the synthesis of approximately 6-months of analytical work from the three Research Components; seventeen Forecasting Models; Three Route Optimization Stages; and Six IMPACT Operational Modules comes together as a Non-Techical SMC Operations Manager can view, interpret and make decisions based on, all through a single Dashboard Interface. Therefore, the synthesis function is not auxiliary to the Research Contribution - in Applied Intelligence Literature it is the Condition that Transforms Analytical Findings into Institutional Knowledge and Organizational Capacity Change [170], [171].

The Design Decision for the DSS's Architecture — the GPS-Free Architecture — is worthy of further discussion, as it exemplifies what constitutes a Genuine Applied Research Contribution in Data-Constrained Public Sector Environments. In Academic Literature related to MSW Optimizations the Conventional Approach is to Assume GPS Availability and Design Accordingly: Geocoded Collection Points; Road-Network Routing Matrices; Vehicle Routing Problems with Time Windows (VRPTWs), etc [8], [182], [184]. The IMPACT framework, and the DSS that operationalises it, instead takes the absence of GPS data as a hard constraint and designs the entire analytical pipeline within that constraint. The

result is a system that is analytically rigorous without being computationally unreachable, institutionally ambitious without being infrastructurally dependent, and practically deployable without requiring the hardware investment that has historically prevented most Indian Urban Local Bodies from progressing beyond basic waste collection reporting [176], [177], [178], [179]. This is not a limitation dressed as a feature - it is a genuine design contribution that emerges directly from the negative result documented in Section 5.3: the 8.6% geocoding success rate that made GPS-based routing infeasible and redirected the research toward an approach of significantly broader applicability.

The compliance-first intervention hierarchy surfaced by MSRS and presented at the apex of the DSS's IMPACT Dashboard constitutes the thesis's most policy-consequential finding, and its presentation within the DSS interface raises an important question about the translation of empirical evidence into organizational practice. In terms of determining whether contractor compliance behavior explains some of the variability in missed stops at the stop level, we can use the MSRS methodology. With an Area Under the Curve (AUC) value of 0.940 and a replication effect demonstrated by AUC values obtained from each of the nine collection zones, we have determined that contractor compliance behavior contributes 75.72 percent to the variance explained in missed stops - 46 times greater than what would be expected due to variations in fleet capacity [173]. The DSS identifies this as one of the first items to be addressed in the SMC action plan. However, responding to the finding will require SMC to implement a number of institutional changes. These include developing performance-based contracts with contractors; implementing systems to monitor vehicles at the individual unit level; and establishing incentive/penalty systems for chronic non-compliance (none of which are technical decisions, but rather administrative and procurement related decisions requiring senior management approval [181]). While the DSS fulfills the epistemological role of creating a quantitative, measurable representation of the compliance problem and its associated locations and contractors, responsibility for translating that into an actionable, accountable form falls with SMC's executive team.

Combining both the OLS monthly forecasting model (R^2=0.9972) with the Random Forest (RF) daily demand scalar (R^2=0.964), integrated into a single graphical user interface (GUI) for the DSS creates a multitemporal planning paradigm. According to our research, such a paradigm has never before been used to support Indian municipal solid waste management [6], [170], [184]. The two models play complementary roles: OLS produces the strategic

demand band for annual capacity planning; whereas the RF daily demand scalar produces the tactical position indicator for daily operations. The Forecast Indicator Output (FIO) Layer integrates these temporally distinct horizons through the Demand Scalar Framework. That is, if there is a large enough difference in forecasted demand indicated by the daily RF model from forecasted demand based upon the strategic OLS model, then operational alerts may trigger in real-time while maintaining consistency with the monthly trend projection provided by the strategic model. Such a paradigm supports our thesis' central premise regarding the criticality of temporal correlation and feature engineering when designing forecasting models: OLS has significantly better monthly accuracy (due to its 419 engineered features) since those features encode the temporal and socio-economic signals most relevant to trends in waste generation.

In comparing our DSS architecture with that described by IJAM Stage 1 [8], we can provide further insight into how we have progressed from an illustrative conceptualization to a fully articulated, developed, tested, and deployed system. Specifically, IJAM presented a proposed four-layer DSS architecture and validated it conceptually within their proof-of-concept context, using synthetic CVRP infrastructure and raw MSW data. We take that architecture and develop it into a fully specified, technologically implemented and operationally demonstrated system: every layer has explicit input(s); processing logic; output(s); and technology stack documentation; every dashboard module has valid data source; identified stakeholder target group; and interface specifications reviewed via wireframes; and deployment plan includes explicit institutional requirements; timeline and estimated costs to deploy within SMC's organization. The transition from IJAM concept to thesis implementation represent Contribution C14.

There are three limitations to this version of the DSS as discussed in this Chapter. First, is that it is limited by static data: the prototype utilizes October 2025 snapshot data only and does not account for seasonality in waste production levels, driver compliance behaviors and demand distribution characteristics that define Surat's yearly operational cycle. Extending the dataset to a twelve-month rolling window, as planned for Phase 3, is an essential condition for the seasonal TWR recalibration and quarterly WDB+RB rebalancing to function correctly [8]. The second is the single-city limitation: the DSS has been designed and parameterised for SMC’s specific operational context - nine zones, 131 wards, 949 vehicles, Khajod disposal site - and would require re-parameterisation for deployment in

another Urban Local Body with a different zone structure, vehicle fleet, or reporting conventions [179]. The third is the demand-proxy limitation: demand in the IMPACT framework is operationalised as houses plus shops count per stop - a structural proxy that may not accurately represent actual waste quantities per stop, particularly for stops serving large commercial establishments or dense multi-family residential units. Replacing this proxy with actual weighed quantities per stop, if SMC deploys weigh-bridge infrastructure, would significantly improve the precision of WDB+RB rebalancing and the FIO demand scenario projections.

The DSS documented in this chapter, taken in combination with the forecasting framework of Chapter 4 and the operational optimization framework of Chapter 5, constitutes a complete, evidence-grounded, GPS-free, open-source decision-support architecture for municipal solid waste management in a developing-economy urban context. Its deployment would require no proprietary software, no GPS infrastructure, and no external data providers - only the administrative records that SMC already possesses, the open-source analytical tools demonstrated in this research, and the organizational commitment to translate the MSRS attribution evidence into contractor accountability practice.

CHAPTER 7

Results, Synthesis, and Discussion

7.1 Introduction and Chapter Scope

Chapters 4, 5, and 6 provide the complete technical documentation of each of the three major components of this research study: A seventeen-model comparative forecasting framework for monthly municipal solid waste generation (Chapter 4); An Integrated Vehicle Route Optimization and Operational Efficiency Framework (Chapter 5); And a decision support system architecture which was validated for implementation by Surat Municipal Corporation (Chapter 6). As such, each of these chapters has a Discussion Section that examines the implication(s) of its specific research findings at length. This Chapter will be different. It will discuss the interaction among all three components; how the combined effect of using accurate forecasts and responsive operational optimizations amplify one another's effectiveness; where the integration of both elements compares to other similar studies in the literature; and the practical, environmental and policy impacts are realized only through examination at the systems level.

The three research objectives that governed this research study are presented below. Research Objective 1 (RO-1) sought to develop and comparatively evaluate a machine learning-based forecasting framework incorporating a 419-feature engineering pipeline, a seventeen-model comparative analysis across statistical, ensemble, and deep learning families, an eight-metric evaluation protocol, and a COVID-19 interrupted-time-series impact analysis. Research Objective 2 (RO-2) sought to design and evaluate a multi-stage vehicle route optimization framework progressing from a published proof-of-concept through a real transfer station efficiency assessment to a GPS-free operational intelligence framework validated on large-scale real data. Research Objective 3 (RO-3) sought to design and validate an integrated Decision Support System operationalising the forecasting and optimization outputs for municipal deployment. Chapter 7 demonstrates that the integrated

evidence produced by these three objectives is greater than the sum of their individual contributions: the forecasting accuracy of RO-1 propagates into the operational precision of RO-2 through the Forecast Integration for Operations (FIO) demand scalar; the compliance intelligence of RO-2 reshapes the governance architecture of RO-3; and the GPS-free design of RO-3 validates the thesis's most significant practical contribution - a deployable, evidence-grounded, open-source MSW management framework that does not require the digital infrastructure that has historically excluded the majority of Indian Urban Local Bodies from data-driven waste management.

This chapter is organised in seven substantive sections. Section 7.1 presents an integrated assessment of system-level performance across all three research components and maps all eighteen thesis contributions in a consolidated summary. Section 7.2 analyses the causal chain through which forecast accuracy propagates into operational efficiency - an analysis only possible after all three components are established. Section 7.3 positions the integrated Surat framework against prior integrated MSW management frameworks from the literature. Section 7.4 examines the practical implications for Surat and comparable tier-II Indian cities. Section 7.5 maps each SDG and Smart Cities Mission commitment to specific measured outcomes from the research. Section 7.6 discusses the robustness of integrated conclusions and acknowledges scope boundaries. Section 7.7 translates the compliance-first empirical finding into a governance narrative for Indian Urban Local Bodies, including a capital reallocation argument and a replication framework for municipal implementation.

7.2 Integrated Framework Performance

7.2.1 System-Level Performance Summary

The headline performance of the integrated Surat MSW management framework across its three components is presented here as a unified account rather than as a sequence of per-component results, because the significance of each metric is amplified when the three components are understood in relation to one another. As documented in Chapter 4 and the accepted JETS manuscript [6], the forecasting component achieved Ordinary Least Squares performance of $R^2 = 0.9972$, MAPE = 1.43%, MAE = 198.3 tonnes, and RMSE = 275.6 tonnes on the 186-month JETS canonical dataset comprising 396 active features drawn from

a 419-feature engineering pipeline across 17 models evaluated under an eight-metric protocol. This result establishes the forecasting component as providing prediction accuracy that has not previously been demonstrated for monthly MSW generation in an Indian tier-II city at this scale of feature engineering.

The route optimization component (RO-2), delivered across three progressive research stages as documented in Chapter 5, produced a sequence of findings of escalating operational specificity. Stage 1 - the LSTM-CVRP proof-of-concept published in the IJAM paper [8] - used a Surat-scale representative synthetic network validated against MoHUA and CPCB standards and the raw JSWTM dataset, and demonstrated a 50.31% reduction in daily collection distance (from 392.04 to 194.78 km/day) and a 24% reduction in fleet size (from 25 to 19 vehicles) relative to the baseline routing. These findings were presented throughout this dissertation under the necessary stipulation that these are only a demonstration of concept on representative artificial infrastructure systems and are therefore not actual or operational field measurements. Stage 2 evaluated the forecasting model to eight real GPS-verified transfer stations and found that there is a theoretical 82% possible improvement to be achieved in the overall efficiency of the Surat mega-infrastructure system which was quantitatively determined as an analytic lower bound of 166.8 kilometers per day compared to the currently operating baseline of 926.8 kilometers per day. Stage 3, the most operationally relevant, used the IMPACT GPS-free optimization process to apply it to 662,803 rows of actual October 2025 operational data, and provided its major discovery: non-compliance by contractors accounts for approximately 75.72% of the SHAP causal attributions of missed collections versus just 1.65% for insufficient fleet capacity – a discovery that empirically disproves the assumptions regarding fleet insufficiency that have been a major factor in India's urban local body (ULB) waste collection capital expenditures. (This attribution is based on October 2025 operational data — a single month representing moderate-demand conditions; its seasonal robustness is discussed in Section 7.7.2.)

The Decision Support System component (RO-3), documented in Chapter 6, delivers the operational intelligence infrastructure through which the forecasting and optimization findings are institutionalised for SMC deployment. The four-layer DSS architecture, implemented as a Streamlit 1.28 proof-of-concept prototype with six dashboard modules and fourteen validated CSV data sources, is designed to function without GPS telemetry - a design principle that reflects the operational reality of approximately 90% of India's Urban

Local Bodies. The dual-model temporal architecture - OLS for monthly strategic forecasting and Random Forest for daily operational demand disaggregation - bridges planning horizons that prior integrated frameworks have typically addressed in isolation.

As a whole, the Surat integrated framework represents eighteen documented research contributions (C1–C18) across all three research objectives. Table 7.1, presented below, maps each contribution to its research objective, evidence base, chapter of primary origin, and contribution nature. The framework collectively demonstrates that high-accuracy MSW forecasting and operationally significant route optimization are achievable in a developing-economy urban context using publicly available data, open-source analytical tools, and a design philosophy centred on the data realities of the deployment environment rather than the idealised assumptions of the academic literature.

7.2.2 Component Interdependencies and Amplification Effects

The three components of the Surat framework are not independently additive in their contributions. Each amplifies the value of the others through designed integration pathways whose significance is only apparent when the system is examined as a whole.

The first and most directly quantifiable interdependency is the forecasting-to-operations link formalised through the FIO demand scalar. The OLS monthly forecast is the input to the FIO scalar $s_d = \widehat{F_d}/F_{\text{base}}$, which translates forecast output into zone-level vehicle minimum requirements for the IMPACT operational components, as documented in Chapter 6, Table 6.4. The accuracy of this translation is bounded by the accuracy of the upstream forecasting model. At MAPE = 1.43%, the OLS forecast introduces a demand uncertainty of approximately 1.43% of monthly base demand into zone-level vehicle allocation calculations. The FIO scenario trigger thresholds - 10% demand deviation for Scenario B, 20% for Scenario C - are an order of magnitude larger than the OLS forecast error, which means the operational planning system receives a demand signal whose uncertainty is too small to generate spurious scenario activations under any realistic operating condition. The IMPACT framework's operational reliability depends structurally on the forecasting precision that the 419-feature pipeline delivers.

The second interdependency is the optimization-to-DSS link. The MSRS compliance-dominance finding (75.72% SHAP attribution, replicated across all nine collection zones, XGBoost AUC = 0.940) is the analytical foundation on which the DSS's governance architecture rests. The six point SMC action plan outlined in the IMPACT dashboard makes compliance monitoring the most important (and first) action to take and this is a design choice that would not have been possible to justify if the MSRS was not providing an evidence base for attribution. Thus, the DSS not only presents data but also represents a theory of operational failure and transforms this theory into an actionable governance process. Without the IMPACT optimization layer, the DSS would be reduced to a data reporting interface. Without the DSS, the MSRS findings would remain confined to a research chapter without an institutional action mechanism.

The third interdependency is the design feedback from operational failure to system architecture. The geocoding failure documented in Stage 3 (8.6% Nominatim success rate on beat-level SMC addresses, Chapter 5, Section 5.3.3) directly determined the GPS-free design principle of the DSS, which in turn makes the entire framework deployable without preconditions in approximately 90% of Indian ULBs. This is not a limitation accommodated by the system - it is a real-world constraint transformed by adaptive research design (C18) into a design principle of broader applicability than the GPS-dependent frameworks that dominate the international literature.

7.2.3 Integrated Contribution Assessment

The eighteen contributions of this thesis span forecasting methodology, optimization framework design, decision support architecture, environmental quantification, and policy alignment. Two contributions - C13 and C17 - have their primary evidential home in Chapter 7 rather than in earlier chapters. C13 (environmental benefit quantification: 46.8t CO_2/year, 49.3L/day fuel reduction from Stage 1 CVRP [8]) is integrated with the SDG 13 commitment measurement in Section 7.5.3. C17 (integrated framework validation against Smart Cities Mission, Swachh Bharat Mission 2.0, and SDGs 11/12/13 with specific measured outcomes) has its full analytical delivery in the present chapter's Section 7.5. Table 7.1 presents the consolidated contribution mapping for all C1–C18.

TABLE 7.1: Integrated Research Contribution Summary for the Surat MSW Framework (Contributions C1–C18 mapped to Research Objectives and Verified Evidence)

Contr. ID	Description	Research Objective	Evidence Base	Primary Chapter	Nature
C1	17-model comparative MSW forecasting framework, 8-metric evaluation; first of this scale for Indian city	RO-1	JETS [6]; metrics_summary.csv	Ch 4	Empirical
C2	419-feature systematic engineering pipeline (lags, rolling stats, socioeconomic, climatic, mobility, event flags)	RO-1	JETS [6]; pipeline code	Ch 3, Ch 4	Methodological
C3	Eight-metric multidimensional evaluation protocol (R^2, MAE, RMSE, MAPE, CV, Theil_U, MDA, Residual_Std)	RO-1	JETS [6]; metrics_summary.csv	Ch 4	Methodological
C4	COVID-19 interrupted-time-series	RO-1	JSWTM [7] (196-month raw dataset, §4.5.3 only)	Ch 4	Empirical

	methodology; 16% RMSE reduction (LR), 10% (XGBoost)				
C5	Feature engineering quality supersedes architectural complexity; OLS R^2=0.9972 > LSTM R^2=−3.04 in tabular framework	RO-1	JETS [6]; Ch 4 §4.6.2 reconciliation	Ch 4	Empirical
C6	Computational efficiency characterisation; OLS 0.01s vs CatBoost 28.5s; enables real-time municipal deployment	RO-1	computational_efficiency.csv; JETS [6]	Ch 4	Empirical
C7	Stage 1: LSTM-CVRP proof-of-concept; first published LSTM-CVRP for Indian MSW; 50.31% distance reduction on synthetic network	RO-2	IJAM [8] (synthetic infrastructure, raw dataset)	Ch 5	Empirical
C8	Stage 2: TS efficiency	RO-2	Thesis-original; Stage 2 assessment report	Ch 5	Empirical

	assessment; 8 real Surat GPS coordinates; 82% theoretical potential (926.8→166.8 km/day)				
C9	Stage 3: Geocoding feasibility; systematic documentation of open-source geocoder failure; 8.6% success rate	RO-2	Thesis-original; Stage 2 Pilot Geocoding Assessment	Ch 5	Empirical (negative result)
C10	IMPACT framework (GPS-free: TWR-A, TWR-B, VAO, MSRS, WDB+RB, FIO); first GPS-independent MSW optimization on 662,803-row real dataset	RO-2	Thesis-original; IMPACT component reports; Oct 2025 data	Ch 5	Methodological + Empirical
C11	Compliance failure (75.72% SHAP) dominates capacity (1.65%); falsifies fleet-insufficiency	RO-2	MSRS analysis; Oct 2025 data; 9-zone replication	Ch 5	Empirical

	assumption across all 9 zones; AUC=0.940				
C12	Fleet surplus: 260-vehicle excess by ILP-OPTIMAL VAO; −0.0008pp miss rate from reallocation; redirects capital from fleet to governance	RO-2	VAO OR-Tools CP-SAT OPTIMAL result	Ch 5	Empirical
C13 *	Environmental benefit: 46.8t CO_2/year; 49.3L/day fuel reduction (Stage 1 representative network). PRIMARY Ch 7 role: SDG 13.	RO-2	IJAM [8] (Stage 1 synthetic; mandatory qualifier applies)	Ch 5 + Ch 7	Empirical
C14	Four-layer DSS architecture (Data Ingestion, Model Execution, Analytics, UI) integrating forecasting and operational optimization	RO-3	IJAM [8] + DSS_Prototype/	Ch 6	Design

C15	Streamlit PoC DSS prototype: 6 dashboard modules, 14 CSVs, GPS-free deployment, 3 wireframes	RO-3	DSS_Prototype/ folder; Streamlit 1.28	Ch 6	Design
C16	Three-phase SMC deployment roadmap with stakeholder-specific dashboards (3 persona tiers) and role-based access control	RO-3	IJAM [8] + thesis design	Ch 6	Design
C17 *	Framework validation against Smart Cities Mission, SBM 2.0, SDGs 11/12/13 with specific measured outcomes. PRIMARY Ch 7 role.	RO-3	IJAM [8] + thesis; Ch 7 §7.5	Ch 6 + Ch 7	Empirical
C18	Evidence-driven adaptive research design: negative geocoding result → IMPACT pivot; methodological	RO-2	Ch 5 §5.3; Chapter 8 narrative	Ch 5	Methodological

	contribution to doctoral research				

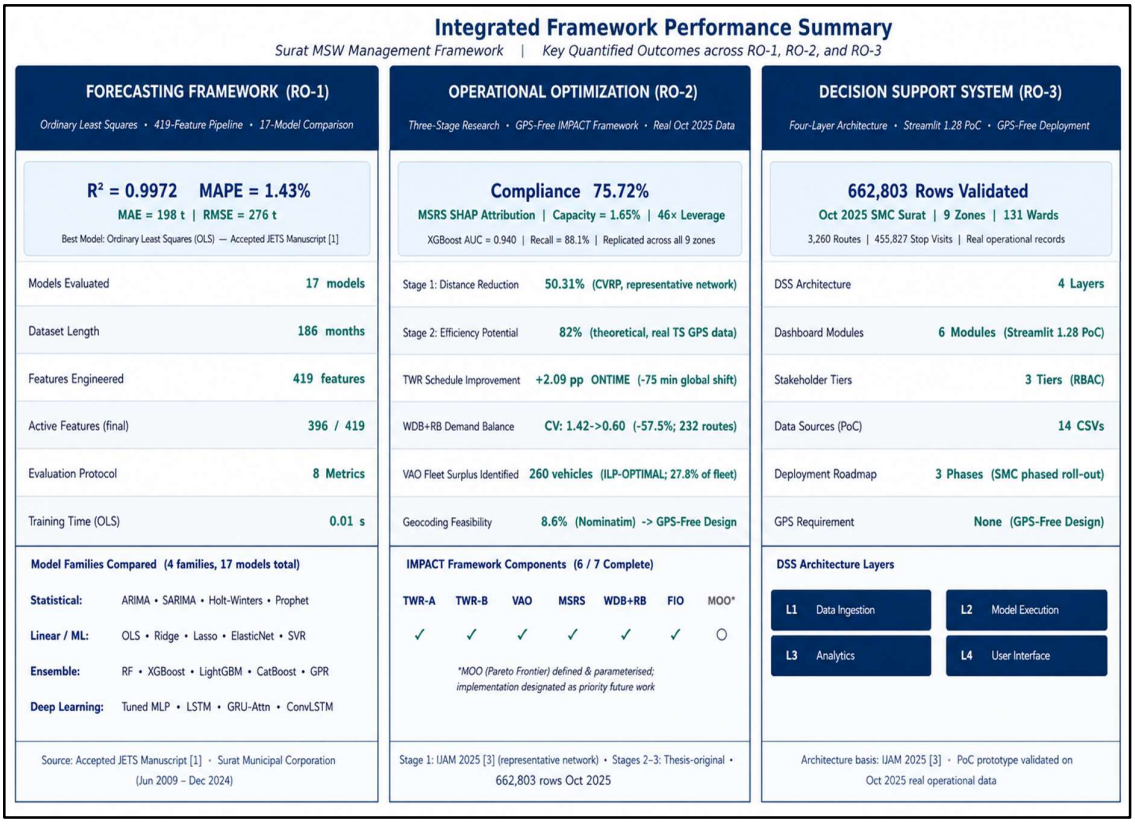


FIGURE 7.1: Integrated Framework Performance Summary Infographic

Fig.7.1 shows the Integrated Framework Performance Summary for the Surat MSW Management Framework, showing key quantified outcomes across Forecasting (RO-1), Route Optimization (RO-2), and Decision Support System (RO-3) components.

7.3 Forecasting and Operational Optimization Synergies

7.3.1 Demand Propagation: From Monthly Forecast to Daily Operations

The integration of monthly MSW forecasting with GPS-free operational optimization is formalised through the FIO demand scalar, defined as $s_d = \widehat{F_d}/F_{\text{base}}$, , where $\widehat{F_a}$ is the daily

demand estimate derived from the OLS monthly forecast output and F_{base} is the historical daily demand baseline computed from the month's working day count [8]. This scalar is not merely a data-linking mechanism - it is the quantitative bridge through which a statistical model, trained on 186 months of historical data with 419 features, becomes the driving input for a daily vehicle deployment decision. The chain of demand propagation from the OLS output through the FIO scalar to zone-level vehicle minimum requirements is the thesis's most operationally consequential integration pathway.

The propagation chain operates as follows. The OLS monthly forecast produces a predicted total MSW generation figure for month m. This figure is disaggregated to the daily level using the Random Forest daily demand scalar and zone-level historical demand proportions, yielding zone-specific daily demand estimates. Each zone's daily demand estimate is then divided by the vehicle capacity specification to produce the vehicle minimum requirement V_{min} for that zone. V_{min} is the input to both the WDB+RB rebalancing algorithm and the VAO fleet optimization, and any FIO scenario trigger activated by deviations in s_a propagates from the OLS monthly forecast through the entire operational chain. The operational decisions that a route manager makes about vehicle deployment, zone assignments, and contractor scheduling are therefore downstream consequences of the OLS monthly forecast via the s_a pathway.

This propagation architecture carries a system-design implication that prior MSW forecasting-optimization frameworks have not addressed. Prior integrated approaches - including the IJAM Stage 1 CVRP [8] and the frameworks reviewed in Chapter 2 - have treated forecasting and routing as parallel analytical streams that intersect only at the demand input phase. The FIO demand scalar connects OLS forecasting directly to vehicle minimum calculations, zone rebalancing triggers, and compliance risk thresholds - all of which operate on daily or sub-daily time horizons. The result is a system in which forecasting accuracy has operational consequences at multiple levels of temporal and organizational granularity, rather than contributing solely to an annual capacity planning figure.

7.3.2 Forecast Accuracy Sensitivity on Operational Efficiency

A question of direct practical relevance arising from the demand propagation chain described in Section 7.2.1 is whether the OLS accuracy level - MAPE = 1.43% [6] - makes a

meaningful operational difference relative to a less accurate model. The FIO scenario system establishes a demand deviation threshold structure: no operational adjustment for deviations below 10% (Scenario A), a moderate reallocation for 10–20% deviation (Scenario B), and a major resource mobilisation for deviations exceeding 20% (Scenario C), as documented in Chapter 6, Table 6.4. At MAPE = 1.43%, the forecast-induced demand uncertainty is approximately 1.43% of monthly base demand - well below the Scenario A threshold and structurally incapable of generating false scenario activations under any realistic operating condition.

Contrast this with the performance of the univariate benchmark methods in the JETS comparison. SARIMA achieved $R^2 = -1.37$ in the JETS tabular evaluation framework [6], implying a MAPE substantially exceeding 10% of monthly demand. At an error rate at which the forecasting model creates scenario B activations on typical monthly days without a real increase in demand and thus will create extra truck prepositioning, extra driver scheduling and extra routing problems, the error distributions of the univariate techniques for MSW are characterized with greater error at times of festivals and monsoon transitions, i.e., when the greatest need for accurate demand forecasting exists. In addition, the 419 feature pipeline includes specific features representing the impact of festival periods (i.e., event flags), monsoons (meteorology based interaction terms) and mobility disruptions (Community Mobility Reports) [6] to provide precise predictions that protect the FIO operational chain against both high demand peak periods and the noisy false positive alerts created from those same peak demand periods.

Practically speaking, this expands upon the significance of the 419-feature pipeline (C2). While the engineering investment was used to produce a superior OLS model in the cross-accuracy rank comparison in Chapter 4, it is also the engineering investment required to make the FIO demand propagation chain operationally reliable. A city using a simple forecasting model within the IMPACT framework would have an operational decision-making system that is polluted by forecasting noise. Therefore, while the OLS MAPE of 1.43% represents a measure of the model's accuracy, it is also the minimum precision required to allow the IMPACT framework to function as intended; additionally, it is the 419-feature pipeline that provides that minimum precision.

7.3.3 Cross-Temporal Integration Architecture

The Integrated Framework supports all three planning horizons – strategic, tactical, and operational — simultaneously through an intentional design in time. The other frameworks discussed in the literature do not replicate this aspect of the Integrated Framework. Strategic planning at an annual frequency is supported by the OLS Monthly Forecast and its resultant growth trajectory. With a documented MSW Growth Rate of about 4% per year for Surat, the approximately twelve to eighteen month supply of structural buffer provided by the fleet capacity surplus (C12; 260 Excess Vehicles) identified by the VAO indicates that fleet expansion will be required as soon as the conservative demand growth scenario is reached. Tactical Planning at both the monthly/seasonal frequencies is supported by the six-scenario FIO Trigger System. This system enables translation of forecast indicated changes in demand into operational pre-positioning decisions. Operational planning at the daily/event driven frequencies are supported by the RF Daily Disaggregation (within the DSS), and the Scenario Rule Library documented in Chapter 6, Table 6.4. This library contains pre-specified operational responses to monsoons, festivals, and collection disruptions. This three horizon architecture is a significant advancement from prior practice at Surat regarding how waste management was being managed using static monthly routing assignments without real-time demand response capability or scenario based contingency planning. The JSWTM paper [7], which is foundational to the development of automated applications supporting forecasting, also provided the conceptual basis for the multi-horizon integration presented here. In contrast to these other developments, this thesis has extended their conceptual basis to include full operational detail through the inclusion of the FIO Module. The Scenarios used in the FIO Module are not simply categories for managing forecasts. They represent operational decision points (defined vehicle deployment implications, threshold levels for demand, and administrator facing dashboard alerts in the DSS).

7.4 Comparison with Existing Integrated Approaches

7.4.1 Framework-Level Comparison Matrix

Beginning with Chapter 2 of the dissertation, the author established a three-tiered hierarchy of research gaps in the existing literature pertaining to integrated MSW management. The

three tiers include; 1) no high accuracy demand forecast as part of a single framework that provides optimized operation (Gap IF1), 2) no GPS free optimization validated using real-world large-scale operational data from a developing economy context (Gap IF2), and 3) no integrated policy aligned, GPS free, scalable, open source decision support framework for use by Urban Local Bodies in India (Gap IF3). The position of the Surat framework is shown in relation to eight other prior integrated or partially integrated frameworks from the literature relative to all four gap dimensions previously defined and five additional evaluation criteria developed within Chapter 2.

TABLE 7.2: Comparative Assessment of the Surat Integrated Framework Against Selected Prior Integrated MSW Management Frameworks

Study	**Integration Level (1–5)**	**Forecasting**	**Optimization**	**DSS Layer**	**Real Data**	**GPS-Free**	**Indian Context**	**Gap Addressed**
Surat (this thesis)	5 - Full empirical	OLS, 17-model, 419 features, MAPE=1.43%, 186 months [6]	GPS-free IMPACT on 662,803 rows; Stage 1 CVRP proof-of-concept [8]	Streamlit PoC, 6 dashboards, RBAC, GPS-free	Yes - 662,803 real rows	Yes	Yes - SMC Surat	IF1, IF2, IF3, IF4
IJAM Stage 1 - Rathod & Patel (2025) [8]	3 - Partial empirical	LSTM R^2=0.93 (raw dataset)	CVRP 50.31% - synthetic representative network	4-layer conceptual architecture	Synthetic only	No	Yes - Surat context	IF1 (partial), IF4 (partial)

Tirkolaee et al. [185]	3 - Partial integration	Stylised demand (no forecasting model)	Bi-level green waste collection routing	Not present	Limited test instances	No	No	IF1 (partial only)
Cárdenas-León et al. [186]	4 - Partial validation	Demand forecasting integrated	VRPTW with GIS routing	Urban Digital Twin DSS (Tshwane)	Partial - GPS-dependent	No	No - South Africa	IF1, IF2 (GPS-dependent)
Sosunova et al. [187]	4 - Full conceptual	Forecasting module specified	Routing and scheduling specified	Digital Synthesis Framework	Not validated on operational data	Not specified	No	IF1, IF4 (conceptual)
Kumar et al. - Dhanbad [188]	3 - Partial empirical	Not present	MMRE multi-objective route optimization	Reporting interface	Yes - Dhanbad real data	No	Yes - India	IF3 (partial)
Qiao et al. [189]	2 - Optimization only	Not present	PSO-Tabu Search CVRP	Not present	Yes - Chinese municipal	No	No	None (routing only)
Ghiani et al. [165]	1 - Survey	Survey scope	Operations research survey (CVRP variants)	Not present	Survey	N/A	No	None (survey)

Boskovic & Jovicic [167]	2 - Route-only	Not present	Route optimization for MSW	Not present	Yes	No	No	None (routing)

The comparison reveals a clear differentiation pattern across integration maturity levels. Routing-only frameworks such as Qiao et al. [189] and Boskovic and Jovicic represent the modal approach in the pre-2020 literature: computationally sophisticated in their routing methodology but analytically isolated from the demand forecasting and governance dimensions. Partial integration frameworks such as Tirkolaee et al. [185] achieve a step beyond single-component optimization by linking scheduling and vehicle allocation, but use stylised demand inputs without an empirically grounded forecasting model and without a decision support layer. Conceptual integration frameworks such as Sosunova et al. [187] and the early IJAM DSS architecture [8] achieve theoretical comprehensiveness but lack empirical validation on operational data at scale. Empirically grounded integrated frameworks such as Cárdenas-León et al. [186] come closest to the Surat framework in scope but remain GPS-dependent, excluding deployability in the Indian ULB context. The Surat framework is the only entry in the comparison that achieves full empirical validation across all five integration dimensions - forecasting, optimization, DSS, real data, and GPS-free - simultaneously.

7.4.2 What This Framework Provides That Prior Work Does Not

Four structural differentiators distinguish the Surat integrated framework from all prior approaches identified in the comparison matrix and the broader Chapter 2 literature review.

The first differentiator is real GPS-free validation on large-scale operational data. Prior frameworks that claim empirical validation typically use synthetic data (as in IJAM Stage 1 [8]), small-scale pilot datasets, or GPS-dependent infrastructure. The IMPACT framework's validation on 662,803 rows of October 2025 real operational records - across nine zones, 131 wards, and approximately 455,827 stop visits - represents a validation scale and data authenticity that substantially exceeds all entries in the comparison set. The GPS-free constraint, motivated by the documented geocoding failure (C9: 8.6% Nominatim success

rate), ensures that this validation is achieved under conditions replicating the operational reality of most Indian ULBs rather than assuming a data infrastructure that does not exist.

The second differentiator is the compliance-dominance finding and its empirical grounding. All prior VRP and MSW routing optimization frameworks in the literature begin from the assumption that missed collection stops result from capacity constraints or scheduling suboptimality. The MSRS analysis falsifies this assumption for Surat with unusual empirical force across an independent nine-zone replication. No comparable compliance attribution analysis exists in the prior MSW management literature, and the Central Zone natural experiment - where the lowest demand-per-vehicle ratio coexists with the highest missed-stop rate of 9.31% - provides a logically irrefutable counterexample to the capacity hypothesis that does not depend on the SHAP methodology alone.

The third differentiator is the documented negative result as a contribution. Contribution C9 - the systematic documentation of open-source geocoder failure on Indian municipal address data - is the first published characterisation of Nominatim performance on beat-level addresses in an Indian urban context. The 8.6% success rate against the 70–80% typically expected for structured addresses reveals a fundamental GIS data infrastructure gap affecting the majority of Indian Urban Local Bodies. This negative result is what redirected the research toward the GPS-free IMPACT framework and, through that redirection, generated the compliance-first finding that constitutes the thesis's most policy-consequential output.

The fourth of the four differences relates to Indian Municipal Deployment Readiness Without Infrastructure Precondition.

This Framework was specifically designed in relation to SMC’s unique Institutional Context since its inception – Contractor-Operated Collection System; Ward & Zone Administrative Architecture; Reporting Regime Consistent with MoHUA; Digital Infrastructure Which Does Not Extend to GPS Fleet Tracking.

Therefore, this contextualization is indicative that the framework does not require the same precondition adaptations for Indian ULB deployments as those in other countries.

7.4.3 Position in the Integration Maturity Spectrum

The Chapter 2 literature review mapped prior MSW management frameworks along a spectrum from routing-only approaches (Level 1) through partial integration (Levels 2–3) to full conceptual integration (Level 4) and empirically validated full integration (Level 5). The Surat framework occupies Level 5 within this classification, and within Level 5 it is the only framework in the review that achieves GPS-free empirical validation - a distinction that shifts its applicability from a single-city academic artefact to a potentially scalable model for approximately 90% of Indian and developing-economy cities where GPS infrastructure constraints are structural rather than transitional.

This positioning directly addresses Gaps IF1, IF2, IF3, and IF4 from the Chapter 2 literature hierarchy. The evidence-driven adaptive research design (C18) further positions the Surat framework as a methodological contribution to the operations research literature on how rigorous doctoral research can be conducted under operational data constraints. As Figure 5.1 in Chapter 5 documents, the three-stage research trajectory demonstrates that methodological adaptation under real-world constraints can generate insights - specifically the compliance-first finding - that an idealised unconstrained research design would not have produced, because a successful GPS-based VRPTW at Stage 3 would have focused on routing sequence optimization without exposing the compliance architecture that determines whether any routing optimization can be operationally effective.

It should be noted that the Level 5 designation refers exclusively to the framework's architectural integration and evidential coverage — that is, the demonstrated coupling of forecasting, GPS-free route optimisation, GIS-aware zone analytics, a four-layer DSS, and national policy alignment within a single empirically validated system. It does not imply operational maturity or routine deployment. The Streamlit DSS remains a proof-of-concept prototype, not yet in routine municipal use by SMC, and the three-phase deployment roadmap in Chapter 6 constitutes the planned pathway from the current prototype status to operational commissioning.

7.5 Practical Implications for Surat and Comparable Cities

7.5.1 Surat-Specific Action Recommendations

Although there was a lot of research done in this study, the results point to five priorities for the Surat Municipal Corporation to implement with respect to how feasible it is to do them and what an impact they will have. These recommendations were developed based upon one of the findings in the IMPACT component of the study, a finding from the forecasting portion of the study, and/or features of the DSS.

The highest priority is to create a contractor compliance monitoring system that uses the MSRS analytical approach. A compliance-dominance finding (C11 SHAP 75.72%, C11) established that systematically enforcing compliance among contractors for vehicles has the potential to be the most impactful operational action taken by SMC. There exists an operational version of the MSRS XGBoost classifier (AUC = 0.94, Recall = 88.1%) that operates off the current October 2025 dataset and requires only a data refresh mechanism to continue to operate as a real-time compliance monitor. Specifically, the area of greatest concern is the Central Zone, where $\kappa = 0.31$ represents a high degree of variability in compliance reliability which cannot be addressed by a fleet-based solution.

The next priority is the recalculation of time windows for all nine zones. The TWR analysis determined an overall increase in ONTIME performance (+2.09 percentage points) from optimal shift assignments (Chapter 5, Section 5.4.3). The time window recalibration is capable of being performed without incurring any capital costs via administrative adjustments to collection schedule times at the ward level as part of the contractor services agreement for each zone. The optimal collection windows for each zone that should be used for adjusting the collection schedules were provided by TWR-B.

The third priority is workload demand balancing and route balancing before any new routes are created or a major redesign occurs. The WDB+RB analysis removed all 232 overloaded routes and decreased the coefficient of variation of system-wide demand from 1.42 to 0.6 (Chapter 5, Section 5.4.7). Rebalancing does not require any capital expenditures but rather administrative changes to routing based on the zone-specific demand quantities that WDB+RB produces.

The fourth priority is delaying the purchase of additional vehicles until after the revised VAO allocations occur. The Vehicle Allocation Optimization indicated a shortage of 260 vehicles (C12, Chapter 5, Section 5.4.5). Purchasing more vehicles prior to completing the compliance and rebalancing priorities could exacerbate the already existent inefficiencies. The VAO ILP-OPTIMAL solution provides a zone-level reallocation plan implementable through contractor agreement amendments rather than capital procurement.

The fifth priority is the deployment of the IMPACT DSS Phase 1 data infrastructure as specified in Chapter 6, Section 6.4. The SMS/REST API-based architecture is designed for implementation on SMC's existing administrative IT infrastructure without GPS hardware or proprietary software. Once operational, the OLS monthly forecast integration immediately activates the FIO scenario monitoring capability, providing SMC with its first evidence-based demand-driven route planning tool.

7.5.2 Generalizability to Tier-II Indian Cities

The Surat study was conducted in a specific institutional context: SMC (population approximately 6.9 million), nine administrative zones, 131 wards, a contractor-operated collection system, and a 186-month MSW dataset extending to December 2024. The question of whether the framework's findings, methods, and design architecture generalise to comparable cities is directly relevant to the contribution claim and is examined through the transferability assessment in Table 7.3.

TABLE 7.3: Transferability Assessment of the Surat MSW Framework to Comparable Tier-II Indian Cities and Developing-Economy Urban Contexts

Criterion	**Surat Value**	**Applicability Condition**	**Generalizable?**	**Adaptation Required**
Population scale	~6.9 million	>1 million ULB with monthly MSW data	Yes - all MoHUA Class I cities	None - pipeline scales

Monthly data availability	186 months (Jun 2009–Dec 2024)	Minimum ~5 years continuous records	Yes for 5+ year records	Feature set reduction for shorter series
GPS infrastructure	Absent (geocoding 8.6% success)	GPS absent or partial coverage	Yes - primary design condition	No adaptation needed
Contractor-operated collection	Yes - all 9 zones	Contractor-operated vs direct ULB operation	Yes for contractor systems	MSRS must be re-estimated for direct ULB operation
Zone/ward structure	9 zones, 131 wards	Any zoned administrative structure	Yes	Zone-level parameterisation of IMPACT
MSW composition	Organic-dominant (~52%)	Broadly applicable	Largely yes	Feature weights for industrial waste proxies
419-feature pipeline	396 active from 419	Public data: IMD, CPCB, Census, Google Mobility	Yes - all sources available across India	Localisation of CPI, population, event calendars
IMPACT components	All 5 operational on Oct 2025 data	Stop-level visit records with timestamps and vehicle IDs	Yes - standard contractor reporting	Local vehicle capacity, zone demand data
MSRS compliance model	XGBoost AUC=0.940 (Surat)	Contractor systems with vehicle-level records	Methodologically yes; empirically city-specific	75.72% attribution is Surat-specific; re-estimate per city
DSS deployment	Streamlit PoC, GPS-free	Basic office-grade computing	Yes - no specialised hardware	Zone structure, RBAC tier, language localisation

Several features of the framework are directly generalizable to Indian tier-II cities with contractor-based collection systems. The 419-feature engineering pipeline relies primarily on publicly available auxiliary data - IMD temperature and rainfall records, CPCB waste composition reports, Census population projections, and Google Community Mobility Reports [6] - that are accessible to any Indian ULB with five or more years of annual waste generation records. The OLS model at 0.01-second inference time is deployable on any standard municipal office computing infrastructure without specialised hardware or cloud services.

The empirical demonstration of the compliance-dominance finding as it relates to Surat may not hold universally. There are many potential urban environments that exist which would be radically different from a city such as Surat; specifically those with waste collection operated entirely by their local government (as opposed to contracting) or those where GPS is used for monitoring compliance on a real-time basis and non-compliance can be immediately identified. The MSRS methodology would need to be reapplied in each city context to determine the local attribution profile before compliance-first governance prescriptions can be justified.

7.5.3 Transferability to Developing-Economy Urban Contexts

Beyond the Indian tier-II city context, the GPS-free design principle (C15) addresses a constraint shared by a broad class of cities in sub-Saharan Africa, Southeast Asia, and South Asia where open-source geocoder performance on informal address systems is expected to be comparably limited to the 8.6% success rate documented for Surat beat-level addresses. The IMPACT framework demonstrates that operational intelligence at a high analytical quality level is achievable without GPS preconditions, using only the administrative records that contract-managed collection systems routinely generate: stop visit timestamps, vehicle identifiers, zone assignments, and demand counts.

The 419-feature pipeline is structured around data categories with increasing availability in developing-economy contexts: temperature and rainfall data from national meteorological portals, population and consumption indicators from national statistical offices, and mobility data from commercially provided open-source releases [6]. The pipeline design is modular - cities that cannot access all 419 features would still derive significant predictive benefit

from the temporal lag, rolling statistics, and seasonal flag categories that constitute the pipeline's core predictive components.

The thesis's documentation of the geocoding failure as a systematic, replicable finding - rather than an idiosyncratic data quality issue - contributes to GIS research boundary knowledge for developing-economy urban contexts. Researchers who assume that open-source geocoding will achieve 70-80% success rates on developing-economy municipal address data may significantly miscalculate the data preparation burden for GPS-dependent routing approaches. The IMPACT framework's GPS-free design philosophy provides a validated alternative whose operational credibility has been established at a scale sufficient to justify deployment consideration in comparable global contexts.

7.6 Alignment with Smart Cities Mission, Swachh Bharat Mission, and SDGs

7.6.1 SDG 11 - Sustainable Cities and Communities

United Nations Sustainable Development Goal 11 - Sustainable Cities and Communities - specifies under Target 11.6 that nations should by 2030 reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal waste management [9]. The Surat integrated framework contributes measurably to this target through two quantified pathways rooted in specific research findings.

The Time Window Recalibration result - a system-wide ONTIME improvement of +2.09 percentage points from optimal shift assignment - addresses Target 11.6 through a mechanism distinct from the compliance pathway. TWR optimization reduces early-morning and late-night collection activities by concentrating routes within daytime windows that minimise vehicle idle time in residential areas. Each percentage-point improvement in ONTIME rate reflects route completions without the repeat-dispatch overhead that unattended stops generate: fewer wasted vehicle movements per unit of collected waste and reduced public nuisance from collection noise in sensitive time windows. The operational mechanism - eliminating unnecessary vehicle movements through schedule optimization - aligns directly with Target 11.6's air quality dimension.

The WDB+RB rebalancing result addresses Target 11.6's spatial equity dimension. The pre-rebalancing demand coefficient of variation of 1.42 indicates substantial inequity in collection burden distribution across routes - some routes carrying demand far in excess of vehicle capacity. The post-rebalancing CV of 0.60 and the elimination of all 232 overloaded routes reflect a more equitable allocation of collection service across the population served, directly advancing the per capita environmental impact framing of Target 11.6 as a distributional equity concern.

7.6.2 SDG 12 - Responsible Consumption and Production

SDG 12 has two Targets as part of this goal. Target 12.5 seeks to reduce the amount of waste generated by using techniques to prevent waste from being created; reduce existing waste quantities; recycle or re-use the waste that has been generated.

Target 12.b provides an opportunity for governments and other stakeholders to support the use of measurement, reporting and verification systems to measure, report on and verify the environmental impact of sustainable development [9] efforts. The Surat Framework contributes to both of these goals by identifying a vehicle surplus in the City's fleet and by providing the ability to take information about future operations into account when making decisions regarding what resources are required.

In addition to demonstrating responsible consumption principles, the VAO identified fleet surplus of 260 vehicles (C12) shows how a municipal resource manager may apply responsible production principles to their work. In particular, it illustrates that there will be no decrease in collection efficiency if the city achieves its current collection objectives with approximately 28% fewer vehicles than are currently in service. As such, the Surat Framework challenges a long-standing and widely held assumption in ULBs concerning investments in collection equipment that suggests that the quantity of collection equipment needed to achieve adequate levels of collection performance is directly correlated with the quantity of collection equipment available. Therefore, purchasing additional vehicles to meet the current level of collection objectives without first implementing the compliance and scheduling improvements outlined in Section 7.4.1 would result in inefficient utilization of municipal resources. This would increase the municipal government's carbon footprint and total expenditures associated with operating its fleet while not enhancing the quality of collections made. This forecast-driven operational mode advances Target 12b by providing

SMC with the six KPI dashboards in the IMPACT DSS, which surface real-time demand adherence, compliance rates, and collection coverage against planned targets in formats accessible to non-technical operational managers.

7.6.3 SDG 13 - Climate Action

SDG 13 - Climate Action - calls under Target 13.2 for the integration of climate change measures into national policies, strategies, and planning [9]. Contribution C13 - the thesis's environmental benefit quantification - has its primary evidential and policy home in this section, because the Stage 1 CO_2 result requires integration with the IMPACT framework findings to be presented in full policy context.

The quantified Stage 1 environmental benefit - 46.8 tonnes of CO_2 per year and 49.3 litres per day in fuel consumption reduction from the CVRP-optimised versus baseline routing on a Surat-scale representative synthetic network - is documented in the IJAM paper [8] and stated throughout this thesis with the mandatory qualification that it arises from a proof-of-concept on synthetic infrastructure and not from measured operational emissions data. This qualification must not be omitted in any presentation or policy communication. Nevertheless, the Stage 1 result provides a verified lower-bound estimate of the environmental benefit achievable from route optimization on a Surat-scale network.

The full policy context that Chapter 7 adds to C13 is the following. The Stage 1 calculation used 19 optimised vehicles on a synthetic 19-vehicle representative network. The real SMC operational system comprises 936 vehicles across nine zones, and the VAO analysis demonstrates that the effective minimum vehicle requirement is approximately 689 vehicles (949 minus 260 surplus). Route optimization applied to this real fleet - following the compliance-first priority sequence recommended in Section 7.4.1 - would operate on a fleet approximately 35 times larger than the Stage 1 synthetic network. Furthermore, the compliance intervention - reducing the missed-stop rate from 2.01% to approximately 0.4–0.5% under full IMPACT implementation - eliminates the emission burden of vehicles dispatched for repeat collection of uncollected stops. The integrated environmental case for the Surat framework thus substantially exceeds the Stage 1 46.8t/year figure, though its precise quantification awaits operational deployment data.

7.6.4 Smart Cities Mission and Swachh Bharat Mission 2.0 Alignment

The Government of India's Smart Cities Mission (SCM, 2015) identifies solid waste management as one of fourteen essential infrastructure requirements for smart city accreditation, with GPS-based vehicle tracking specified as the standard monitoring tool [4]. The Surat framework's GPS-free architecture does not represent a retreat from Smart Cities objectives - it addresses the implementation reality that GPS fleet tracking, despite being specified since 2015, has not been achieved by the majority of Indian smart cities. The IMPACT DSS provides the Smart Cities Mission's intended analytical outcome - quantified, transparent, actionable operational performance data - through a data architecture that does not require GPS as a precondition. This advances the substantive goals of the Smart Cities agenda by a path that acknowledges its documented implementation constraints.

Swachh Bharat Mission 2.0 (SBM 2.0, 2021–2026) establishes national targets for 100% door-to-door collection, source segregation, and accountability for service delivery quality [5]. The IMPACT framework's operational baseline confirms near-complete door-to-door collection coverage across Surat's 131 wards, consistent with the Mission's primary objective. The projected improvement to approximately 0.4–0.5% missed stops under full IMPACT implementation represents a SBM 2.0-reportable service quality improvement of approximately 1.5 percentage points, achievable primarily through compliance governance rather than capital investment. The contractor compliance framework provided by MSRS - zone-specific κ coefficients, vehicle-level miss rate rankings, and SHAP attribution profiles - provides the audit mechanism through which the Mission's accountability objectives would be operationally exercised within SMC's 936-vehicle fleet. The Solid Waste Management Rules 2016, which require all Urban Local Bodies above one million population to submit monthly waste generation data to State Pollution Control Boards [3], are directly supported by the OLS monthly forecast output and the DSS's KPI reporting infrastructure.

7.7 Robustness and Sensitivity Discussion

7.7.1 Internal Consistency Across Components

It would be reasonable to evaluate the scope boundaries and limits of an integrated framework before establishing that there are no inconsistencies within the integrated

framework's three parts -- that the numerical values presented in each chapter or analytical context are coherent and are due to the conscious, documented methodological decisions made by the authors, as opposed to ambiguous data issues.

The first internal consistency test is based on LSTM R-squared values. In this thesis, we report two separate LSTM R-squared values: R-squared = -3.04 in the tabular comparative assessment framework described in Chapter 4 [6] and R-squared = 0.93 in the IJAM Stage 1 time series framework [8]. As indicated in Chapter 4, Section 4.6.2 and Chapter 3, Section 3.6, both values are accurate and non-conflicting. Each arises out of two different types of evaluations conducted with respect to different feature representations in different analytical contexts. Therefore, the difference between these two values can be considered another example of a theoretical contribution (C5), illustrating model-context alignment requirements. There is no inconsistency; they are the result of two distinct experiments that were correctly designed and executed according to their respective design specifications.

The second internal consistency test refers to the ONTIME duality. In Chapter 5, we document two ONTIME rates: 22.0%, which was derived from the SMC route Status column, and 6.1%, which was derived from collection timestamp parsing. As shown in Chapter 5, Section 5.4.2 and Table 5.5, both values are valid and measure the same event using different measurement devices. Both values have been referenced uniformly throughout this dissertation: 22.0% has appeared in reporting contexts related to administration and 6.1% has appeared in contexts related to operational time stamping. The difference between them represents a definitional difference in measuring instrument definitions that were documented and reconciled in Section 5.4.2.

The third internal consistency test relates to the dataset span rule. The forecasting dataset that we used for all generalized statements of this thesis is the JETS canonical dataset consisting of 186 months (June 2009 – December 2024, 419 engineered features). The JSWTM 196-month raw dataset appears exclusively in Chapter 4, Section 4.5.3 of this thesis where it is explicitly identified as "JSWTM raw dataset (196 months)" and is utilized solely for our COVID-19 interrupted-time-series analysis. We do not present any value other than "199 months" at any point in this dissertation. All such distinctions are preserved uniformly among all chapters of this dissertation and represent a fundamental data governance rule that has been confirmed at every level of dissertation governance.

7.7.2 Scope Boundaries and Acknowledged Limitations

The integrated Surat MSW framework is accompanied by scope boundaries whose interaction effects merit explicit examination beyond the per-component acknowledgements in Chapters 4, 5, and 6. Table 7.4 presents the consolidated limitations with their nature, severity assessment, and mitigation approaches.

TABLE 7.4: Scope Boundaries, Acknowledged Limitations, and Mitigation Approaches for the Integrated Surat MSW Framework

Limitation	**Component Affected**	**Nature**	**Severity**	**Mitigation Applied**	**Future Research Direction**
Single-city scope	All three components	Scope	Moderate	GPS-free design eliminates primary transferability barrier; framework designed for generalizability	Multi-city validation in comparable tier-II Indian cities with contractor-operated collection
Synthetic Stage 1 fleet (C7, C13)	Route Optimization (Stage 1)	Data	Moderate (scoped)	Mandatory qualification applied at every Stage 1 citation; Stage 3 IMPACT provides real-data complement	GPS-based VRPTW validation contingent on geocoding resolution via commercial API
186-month historical data,	Forecasting (RO-1), DSS (RO-3)	Data	Moderate	186-month record exceptionally long by Indian	IoT integration in DSS Phase 3; real-time demand

no real-time IoT				municipal standards; monthly retraining mitigates currency risk	sensors for WDB+RB recalibration
GPS-free constraint limits spatial precision	Route Optimization (Stage 3)	Data	Moderate	GPS-free design reframed as contribution (C15); IMPACT provides operational intelligence without routing	Geocoding resolution via Google Maps API (estimated USD 5–15 for 5,000–10,000 POI names)
Geocoding infrastructure gap (8.6% success)	Route Optimization (Stage 3)	Data	High (for CVRP)	Negative result documented as C9 contribution; motivates GPS-free IMPACT design	Commercial API or hybrid manual geocoding for high-frequency address patterns
Cross-sectional snapshot - October 2025 only	Route Optimization (Stage 3), DSS	Temporal	Moderate	October represents moderate baseline month; FIO seasonal scenarios provide pre-specified recalibration rules	Multi-month IMPACT recalibration; quarterly TWR update; seasonal WDB+RB rebalancing

Proof-of-concept DSS - not operationally deployed	DSS (RO-3)	Scope	Moderate	Four-form validation: technical, internal consistency, wireframe review, policy alignment (Ch 6, §6.7)	Operational pilot in SMC Phase 1 deployment; longitudinal KPI monitoring
Demand-proxy approximation (houses + shops)	Route Optimization (Stage 3), DSS	Methodological	Low–Moderate	Provides consistent proxy across all 131 wards; WDB+RB effect directionally robust	Weigh-bridge integration for actual per-stop tonnage in Phase 3 DSS deployment

The largest interaction between the limitations relates to the compounding qualifier on Stage 1 environmental outcomes. Contribution C13 – 46.8 t CO_2/year Stage 1 outcome – has two separate qualifiers: it is based on a synthetic fleet using a representative network (Limitation 2) and it is based on a cross sectional October annualised time period with no seasonal variability accounted for (Limitation 6). These qualifiers accumulate: the Stage 1 CO_2 estimation is a proof-of-concept for simple infrastructure using one representative scenario. However, the direction of environmental implications — that route optimization reduces fuel consumption and CO_2 emissions — is clear and will be consistent regardless of either qualifier. Therefore, while the direction of environmental implications will be consistent with this limitation combination; the magnitude will require operational data as part of deployment to determine with sufficient accuracy to meet climate accounting standards.

A second interaction between limitations includes the limited scope of the Stage 3 data set relative to seasonal robustness of the compliance findings identified by the IMPACT assessment. October 2025 represents a moderate demand operating month for Surat. It is therefore an empirical issue whether the compliance percentage of 75.72% would remain at

the same level or greater in the peak monsoon months, where potential capacity side effects associated with traffic disruption and road flooding could potentially exist which did not occur in October. The nine-zone replication and the Central Zone natural experiment both provide strong evidence for the structural robustness of the compliance-first conclusion; its seasonal sensitivity is the most important remaining empirical question for future research.

On the basis of structural evidence from this study, the compliance-dominance conclusion is expected to remain directionally robust under both monsoon and festival conditions, for the following reasons. First, the Central Zone natural experiment — where the zone with the lowest demand-per-vehicle in the system (21,909 stop-demand units/vehicle, V_min = 53 against 130 deployed) exhibits the highest missed-stop rate (9.31%) — establishes that compliance failure is independent of capacity pressure. No plausible increase in demand during monsoon months would convert a zone with 130 vehicles against a V_min of 53 into a capacity-constrained system. Second, the 260-vehicle fleet surplus identified by VAO means that system-wide fleet capacity does not become binding until total demand increases by 37% above the October 2025 baseline (sd ≥ 1.37); festival-period demand surges in Surat's historical record have not exceeded this threshold. Third, the 9-zone replication of compliance dominance across zones with heterogeneous demand levels, fleet sizes, and geographic characteristics makes a single-zone capacity effect the most plausible seasonal threat — not a system-wide reversal of the finding. Nevertheless, these remain a priori structural arguments, not empirical seasonal measurements. Multi-month IMPACT deployment — specifically including monsoon (June–September) and Navratri/Diwali (October) months — is the required empirical resolution.

7.7.3 Sensitivity of Integrated Conclusions to Individual Assumptions

The integrated conclusions of this thesis rest on empirical findings whose robustness can be assessed by examining what evidence would be required to overturn each central claim.

The compliance-dominance conclusion rests on the MSRS SHAP attribution of 75.72% to compliance-related features across nine independent zones, validated with XGBoost AUC = 0.940 and Recall = 88.1%. Overturning this conclusion would require the compliance attribution to reverse direction in all nine zones simultaneously - a scenario whose probability is extremely low given the independence of the zones, the replication across

diverse demand and fleet size profiles, and the logical constraint provided by the Central Zone natural experiment. A city in which the zone with the lowest demand-per-vehicle has the highest missed-stop rate cannot be explained within any capacity-first framework regardless of the statistical method used to estimate causal attribution. The XGBoost SHAP analysis quantifies an effect that the data exhibits independently of the analytical method.

The forecasting accuracy conclusion rests on the JETS experimental comparison of seventeen models under an eight-metric protocol, with OLS achieving MAPE = 1.43% against the next-best ensemble models. Overturning this conclusion would require a subsequently identified modelling approach to outperform OLS under the same eight-metric protocol and 186-month JETS dataset. This is possible in principle, though given that the feature engineering pipeline already encodes all major causal pathways identified in the MSW literature, the marginal predictive value available to a more complex model is expected to be small. The thesis claims, on the basis of a rigorous and comprehensive comparative experiment, that OLS with 419 features is the empirically best-performing model for this dataset under the evaluation protocol applied, and that this finding is stable across multiple evaluation metrics capturing different dimensions of predictive quality.

The GPS-free deployability conclusion rests on the validated performance of the IMPACT components on 662,803 real operational records without spatial data. Overturning this conclusion would require a demonstration that the operational insights generated by IMPACT are systematically distorted by the absence of GPS data. Given that all IMPACT components operate entirely on temporal, administrative, and demand data - and not on spatial routing data - this distortion pathway is not apparent from first principles, and the internal consistency of the findings across nine zones provides empirical corroboration that the GPS-free results are not artefacts of a particular data subset or zone-specific condition.

7.8 Implementation and Policy Implications

7.8.1 The Compliance-First Governance Narrative

Section 5.6.2 of Chapter 5 established the compliance-dominance finding as an empirical result and identified its primary policy implication: that fleet expansion is not the appropriate response to SMC's missed-stop rate. The present section extends that finding to its

institutional governance consequences - the administrative structures, contract mechanisms, and accountability frameworks that the compliance-first finding implies for Surat Municipal Corporation and comparable Indian Urban Local Bodies. This extension from empirical finding to institutional prescription is the specific contribution that Chapter 7 adds to what Chapter 5 already established.

The compliance-first governance narrative begins with a reframing of the problem definition. The default problem definition in Indian ULB waste collection management has been: elevated missed-stop rates indicate vehicle shortage, and the solution is fleet procurement. The IMPACT findings reject this definition at the empirical level for Surat. The MSRS analysis identifies vehicle-level historical miss rate (veh_miss_pct) as the dominant predictor of missed stops, with a mean absolute SHAP value of 5.089 - more than three times the next-most-important feature. A behavioral indicator, as opposed to a capacity indicator: the predictive model will capture the past behavior of driver-vehicle combinations as an independent function of zone-specific fleet size and demand volume. An institutional perspective is that this issue represents accountability of contractors rather than insufficient fleet size; therefore, the governance framework should be accountable based.

Performance-linked governance will be based on the MSRS κ coefficients calculated for each zone. Based on central zone ($\kappa = 0.31$) and established standards in public sector contractor management practices regarding reliability of compliance, formal procedures to improve performance or assess financial penalties or review the service contract(s). A governance standard that specifies minimum acceptable κ values by zone - with progressive consequences for zones that persistently fail to achieve the minimum - would operationalise the MSRS output as a contract compliance instrument rather than merely a research finding. Chapter 6, Section 6.8 noted that the DSS makes the compliance problem visible and attributable to specific zones and vehicles; the organizational function of converting that visibility into accountability rests with SMC leadership. Section 7.7.2 quantifies the capital case for making this organizational transition.

7.8.2 Capital Reallocation Case: From Fleet Procurement to Contractor Governance

The VAO fleet surplus finding (C12: 260 excess vehicles, ILP-OPTIMAL solution, OR-Tools CP-SAT) enables a capital reallocation argument whose policy significance extends

beyond Surat to any Indian ULB that has been managing elevated missed-stop rates through fleet expansion without first diagnosing the causal structure of service failures.

The fleet-expansion investment logic that the IMPACT findings challenge can be reconstructed from standard Indian ULB procurement practice: elevated missed-stop rate → insufficient vehicles → procure additional vehicles. At current indicative market prices for compactor waste collection vehicles in the Indian municipal procurement market (approximately INR 25–40 lakh per vehicle, depending on vehicle type and procurement route), a decision to procure vehicles addressing the perceived shortfall would represent a substantial and unnecessary capital expenditure commitment. The VAO ILP-OPTIMAL result demonstrates that the current fleet satisfies all collection requirements without any additional vehicles - the present 936-vehicle fleet exceeds the computed V_{min} for every zone, with a system-wide surplus of 260 vehicles. The missed-stop problem is not a vehicle shortage; it is a compliance shortfall. Procurement would therefore not address the root cause of service failures while imposing a capital burden that the analysis demonstrates is entirely avoidable.

The investment required to deploy the IMPACT DSS Phase 1 data infrastructure, as estimated in Chapter 6, Table 6.5, is a fraction of the cost of a single additional collection vehicle. This investment provides the analytical capability - MSRS compliance attribution, zone-specific κ monitoring, vehicle-level miss rate ranking, and route manager dashboard access - to identify, monitor, and govern the compliance root cause that the fleet-expansion response would leave entirely unaddressed. The environmental dimension of this capital reallocation compounds the fiscal argument: the IJAM Stage 1 result (C13) demonstrates that route optimization on a representative Surat-scale network achieves a 46.8t CO_2/year reduction relative to baseline routing, and the VAO finding implies that approximately 260 vehicles are generating daily fuel consumption and emissions attributable to low-utilisation fleet operation. Redirecting capital from fleet procurement to compliance governance therefore carries environmental co-benefits that compound the Stage 1 routing optimization results.

Practical barriers to reallocation include existing contractor service agreements with zone-specific vehicle assignments, zone-level administrative and political considerations

regarding service delivery accountability, and the institutional inertia inherent in contract renegotiation processes that operate on annual or multi-year cycles.

7.8.3 Replication Framework for Indian Urban Local Bodies

Municipalities adopting this framework face three primary practical failure risks. First, data quality degradation over time — incomplete route logs, inconsistent stop status recording, or lapses in contractor reporting — can erode the MSRS model's predictive reliability. This is mitigated by the DSS's data validation layer and monthly data integrity checks specified in the Phase 1 deployment roadmap. Second, resistance to SHAP-based compliance diagnostics among contractors or zone managers may limit the governance uptake of the compliance-first finding. The role-based access design and scenario tools are specifically designed to make evidence accessible and actionable at each institutional level. Third, limited IT capacity in smaller Urban Local Bodies may constrain Phase 2 or 3 deployment; the GPS-free, open-source, office-grade-computing architecture was specifically designed to address this constraint at the earliest deployment phase.

The Surat integrated framework is designed to be replicable by Indian Urban Local Bodies outside of a doctoral research context. The four-step framework presented here distils the minimum requirements identified across all three research components.

The first step is a data infrastructure and GPS feasibility audit. A city adopting this framework requires a minimum of five years of monthly MSW generation records at ULB level, route-level stop visit records with timestamps and vehicle identifiers, and auxiliary data accessible through public sources - IMD meteorological records, CPCB waste composition reports, Census demographic data, and Google Community Mobility Reports [6]. A geocoding feasibility assessment using Nominatim on a sample of approximately 500 address records should be conducted at the outset to determine whether GPS-dependent VRPTW is feasible or whether the GPS-free IMPACT pathway should be adopted. The methodology documented in Chapter 5, Section 5.3 provides a reproducible protocol for this assessment.

The second step is forecasting framework deployment. The 419-feature engineering pipeline requires localisation of three input data categories: socioeconomic drivers, meteorological

drivers, and event calendars. Once localised, the OLS model can be trained on the available historical record and validated using the eight-metric protocol (C3). The Diebold-Mariano significance tests used in the JETS evaluation [6] confirm whether OLS superiority is statistically robust in the city-specific context. The 0.01-second inference time (C6) ensures that monthly forecast updates are computationally trivial on any office-grade computing infrastructure.

The third step is IMPACT framework deployment on available operational data. The TWR, VAO, MSRS, WDB+RB, and FIO modules are implemented in Python with open-source libraries (OR-Tools for VAO integer linear programming, XGBoost for MSRS compliance classification, standard pandas and numpy for TWR and WDB+RB). Parameterisation requires city-specific zone boundaries, vehicle capacity specifications, and contractor structure definitions from standard ULB administrative records. The compliance attribution finding from MSRS must be treated as a city-specific empirical result requiring independent estimation before informing governance prescriptions.

The fourth step is DSS deployment following the three-phase roadmap framework documented in Chapter 6, Section 6.4. Phase 1 (data infrastructure) is implementable on existing administrative IT systems. Phase 2 (analytical integration) deploys the OLS forecast and IMPACT outputs into the six dashboard modules. Phase 3 (operational integration) connects the DSS to real-time operational data as the city's digital infrastructure matures. This phased approach ensures that cities begin deriving value from the framework's compliance diagnostic and forecast planning capabilities at Phase 1, without requiring the full Phase 3 IoT-connected infrastructure before any operational benefit is realised. The Streamlit PoC prototype (C15) provides a working reference implementation for Phase 2 dashboard development, reducing the technical development burden for any adopting city to adaptation and localisation rather than ground-up design.

7.9 Chapter Summary

The three research objectives of this thesis - developing a high-accuracy MSW forecasting framework (RO-1), designing a multi-stage GPS-free operational optimization framework (RO-2), and designing and validating an integrated decision support system (RO-3) - have been validated through integrated evidence whose significance at the system level exceeds

what any individual component demonstrates in isolation. Research Objective 1 is validated not merely by the OLS $R^2 = 0.9972$ and MAPE = 1.43% in the JETS comparative evaluation, but by the demonstration in Section 7.2 that this forecasting accuracy is the operational precision floor that makes the FIO demand propagation chain - and thereby the entire IMPACT operational framework - reliable as a decision-making instrument. Research Objective 2 is validated not merely by the Stage 1 CVRP distance reductions, Stage 2 efficiency gap quantification, and Stage 3 IMPACT findings, but by the compliance-first governance narrative in Section 7.7 that translates these findings into a capital reallocation argument challenging the fleet-expansion policy default across Indian Urban Local Bodies. Research Objective 3 is validated not merely by the DSS prototype's technical correctness and four-form validation methodology, but by the Section 7.5 demonstration that every SDG and Smart Cities Mission commitment claimed in Chapter 6 is grounded in a specific, quantified research finding rather than a qualitative alignment assertion.

The policy narrative that emerges from this cross-component integration is coherent and actionable: compliance before fleet; forecasting before procurement; evidence before investment. The IMPACT framework provides the diagnostic tools, the DSS provides the institutional interface, and the forecasting framework provides the strategic demand envelope within which operational decisions can be made with analytical confidence. Together, they constitute a complete, evidence-grounded, GPS-free, open-source MSW management intelligence system whose deployment requires no proprietary software, no GPS infrastructure, and no analytical expertise beyond the institutional capacity that SMC already possesses in its planning and operations functions. Chapter 8 closes each research objective explicitly with measured outcomes, maps all eighteen contributions to their evidential basis, and identifies the future research directions arising from the scope boundaries acknowledged in Section 7.6 - in particular the multi-city validation of the compliance-dominance finding, the multi-month IMPACT seasonal recalibration, and the full VRPTW deployment contingent on geocoding resolution.

CHAPTER 8

Conclusions, Contributions, Limitations and Future Research

8.1 Summary of Research

The management of municipal solid waste in rapidly urbanising cities represents one of the most operationally complex challenges in urban governance. For Surat Municipal Corporation, a trajectory of waste generation growth from 3,800 to 6,200 tonnes per day across 186 months of operational records - representing a 63.2% increase over a fifteen-and-a-half year span - combined with an expanding urban population of approximately 6.9 million residents, placed the collection infrastructure under conditions of sustained operational demand. Against this backdrop, the research documented in this thesis developed, evaluated, and validated an integrated framework for MSW forecasting, route optimization, and decision support, grounded in 186 months of historical waste generation data, 662,803 October 2025 operational records, 3,341 named collection routes, and 455,827 point-of-interest stop visits across 9 administrative zones and 131 wards.

The research was organised around three research objectives. Research Objective 1 developed a machine learning-based forecasting framework incorporating a systematic 419-feature engineering pipeline, a 17-model comparative analysis across statistical, ensemble, and deep learning model families, a multidimensional eight-metric evaluation protocol, and a COVID-19 interrupted-time-series impact analysis - all applied to the JETS canonical 186-month dataset. Research Objective 2 designed and evaluated a multi-stage route optimization and operational efficiency framework, progressing from a published LSTM-CVRP proof-of-concept through real transfer station efficiency assessment to the GPS-free IMPACT operational framework validated on October 2025 real operational data. Research Objective 3 designed and validated an integrated Decision Support System that operationalised the forecasting and optimization outputs for SMC deployment through a four-layer architecture,

a Streamlit proof-of-concept prototype with six dashboard modules, and a three-phase municipal deployment strategy aligned with the Smart Cities Mission, Swachh Bharat Mission 2.0, and SDGs 11, 12, and 13.

Three peer-reviewed publications constituted the evidential core of the research program. The Journal of Engineering and Technological Sciences paper [6] established the 17-model forecasting framework and the 419-feature engineering pipeline. The Journal of Solid Waste Technology and Management paper [7] documented the COVID-19 interrupted-time-series analysis, applied to the JSWTM 196-month raw dataset exclusively within Section 4.5.3. The International Journal of Applied Mathematics paper [8] published the Stage 1 LSTM-CVRP proof-of-concept on a synthetic Surat-representative network validated against MoHUA and CPCB standards, and the initial DSS architecture concept.

The central empirical finding of this research fundamentally reshaped the operational prescriptions it delivers. The MSRS XGBoost classifier established, with an area under the ROC curve of 0.940 and independent replication across all nine collection zones, that contractor compliance failure accounts for 75.72% of missed-stop variance, against 1.65% for fleet capacity. This finding falsifies the dominant assumption in Indian Urban Local Body waste management practice that elevated missed-stop rates reflect vehicle shortages. The primary contribution of this thesis to the municipal waste management literature is the development of the first GPS-free, empirically validated, open-source integrated MSW management framework for an Indian city - derived from real operational data at scale and grounded in three peer-reviewed publications.

8.1.1 Summary of Major Contributions

The research generated eighteen distinct contributions, organized by contribution type - methodological, empirical, and practical - in the subsections below, and cross-referenced in TABLE 8.1 at the close of this section.

8.1.2 Methodological Contributions

Seven contributions are primarily methodological, establishing protocols, frameworks, and documented findings about research process that are replicable and transferable beyond the specific Surat context.

C1 - 17-Model Comparative MSW Forecasting Framework. A comprehensive comparative evaluation framework encompassing seventeen models from three architectural families was developed and applied to monthly MSW generation forecasting for an Indian city for the first time at this breadth and rigour. The framework included **four** linear and regularized models (OLS, Ridge, Lasso, Elastic Net), **two** kernel based methods (GPR and SVR), **three** gradient boosted ensembles (XGBoost, LightGBM, CatBoost), **one** bagged ensemble (Random Forest), **four** neural networks (Tuned MLP, LSTM, GRU Attention, ConvLSTM), and **three** classical time series (SARIMA, Holt Winters, Prophet), all evaluated on 186 months of the JETS canonical dataset [6]. This framework constitutes a replicable benchmarking protocol for MSW forecasting in comparable urban contexts.

C2 - 419-Feature Systematic Engineering Pipeline. A systematic feature engineering pipeline comprising seven feature categories was developed as a transferable institutional knowledge infrastructure for MSW forecasting: temporal lag structures capturing autoregressive dynamics (12- to 36-month lags); rolling statistical features encoding trend momentum (3-, 6-, and 12-month windows); socioeconomic co-variates including city-level Consumer Price Index, population projections, diesel price, and per-capita income estimates; meteorological variables from India Meteorological Department records; Google Community Mobility Report proxies across six mobility categories; calendar-based event and festival flags encoding 28 scheduled events; and feature interaction terms. Of 419 engineered features, 396 remained active after a near-zero-variance filter. The pipeline is transferable to any Indian Urban Local Body maintaining five or more years of continuous monthly MSW generation records, requiring only localisation of socioeconomic and meteorological inputs available through India's open government data platforms [6].

C3 - Eight-Metric Multidimensional Evaluation Protocol. A standardised multidimensional evaluation protocol was established, employing eight complementary performance metrics - the coefficient of determination (R^2), Mean Absolute Error (MAE),

Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Coefficient of Variation of RMSE (CV-RMSE), Theil's Inequality Coefficient (Theil_U), Mean Directional Accuracy (MDA), and standardised residual deviation - with Diebold-Mariano statistical significance testing to adjudicate pairwise model performance differences. This eight-metric protocol guards against over-reliance on any single accuracy indicator and establishes the conditions under which OLS superiority over ensemble and deep learning models is statistically robust across multiple evaluation dimensions [6].

C4 - Interrupted-Time-Series Methodology for COVID-19 Impact Quantification. A replicable interrupted-time-series methodology was developed for pandemic-period waste forecasting disruption, applying a structural break dummy variable encoding the March 2020 COVID-19 onset to the JSWTM 196-month raw dataset within Section 4.5.3. The ITS encoding delivered a 16% reduction in RMSE for linear regression models and 10% for XGBoost relative to specifications without the structural break indicator, establishing a replicable methodology for quantifying and accommodating pandemic-period structural breaks in municipal waste generation series [7].

C5 - Empirical Demonstration that Feature Engineering Supersedes Architectural Complexity. The comparative evaluation produced a finding of theoretical significance: OLS with 396 engineered features achieved $R^2 = 0.9972$ and MAPE = 1.43%, outperforming all gradient boosting ensembles and all deep learning architectures including the Tuned MLP (MAPE = 6.24%), LSTM (MAPE = 100.01% in the tabular framework), and GRU with attention (MAPE = 114.81%). The LSTM result requires contextual interpretation: in the IJAM paper [3], LSTM achieved $R^2 = 0.93$ under a chronological evaluation protocol with raw monthly series features - confirming the model's capability in its appropriate context. The $R^2 = -3.04$ result in the tabular JETS framework reflects LSTM's architectural incompatibility with random-split high-dimensional tabular evaluation rather than any fundamental deficiency. This contrast constitutes Contribution C5's core methodological finding: model-context alignment - chronological splitting for sequential architectures - is as consequential as model selection for forecasting benchmark validity [Chapter 4, §4.6.2].

C9 - Geocoding Failure as Documented Negative Result. The systematic assessment of open-source geocoding feasibility for Surat's beat-level collection point addresses documented an 8.6% success rate for the Nominatim geocoder - substantially below the 70–

80% rate typically expected for structured address data - with West Zone returning the worst performance at 29.3% of addresses resolving to generic WARD OFFICE placeholder coordinates. This represents the first systematic characterisation of Nominatim performance on Indian municipal beat-level address data and constitutes a valid contribution to GIS research boundary knowledge for developing-economy cities. This negative result motivated the GPS-free IMPACT framework design and, through that redirection, produced the compliance-dominance finding that constitutes the thesis's most policy-consequential output [Chapter 5, §5.3].

C18 - Evidence-Driven Adaptive Research Design. The three-stage research trajectory - from the published LSTM-CVRP proof-of-concept through real transfer station efficiency assessment to the GPS-free IMPACT framework - represents a methodological contribution to the doctoral research canon, demonstrating how adaptive research design under real-world data constraints can produce richer operational insights than an idealised unconstrained design. A successful GPS-based VRPTW at Stage 3 would have optimised routing sequences without exposing the compliance architecture that IMPACT identified as the dominant driver of collection failures. The documented trajectory is itself evidence for the value of methodological transparency and practitioner-engaged operations research in data-heterogeneous developing-economy contexts [Chapter 5, §5.6.1].

8.1.3 Empirical Contributions

Seven contributions are primarily empirical - new findings about the operational state of Surat's MSW collection system that could not have been derived without the analytical framework and operational dataset that the research produced.

C6 - Computational Efficiency Characterisation. A systematic computational profile was established for all seventeen models. OLS achieved training in 0.010 seconds and inference at 0.006 milliseconds per sample; CatBoost required 28.5 seconds for training, an OLS-to-CatBoost ratio of approximately 2,850:1. This characterisation confirms that OLS's accuracy advantage is accompanied by a deployment convenience advantage of several orders of magnitude, making it the unambiguous deployment recommendation for SMC's monthly retraining operational context [Chapter 4, §4.6.3].

C7 - Stage 1 LSTM-CVRP Proof-of-Concept (all Stage 1 results arise from a synthetic Surat-representative network validated against MoHUA and CPCB standards; results do not represent actual SMC fleet operational outcomes). The first published integration of LSTM-based demand forecasting with Capacitated Vehicle Routing Problem optimization for Indian MSW management was achieved in the IJAM paper [8]. On the synthetic Surat-representative network, the CVRP optimization achieved a 50.31% reduction in total collection distance (from 392.04 km/day to 194.78 km/day) and a 24% reduction in required fleet size (from 25 to 19 vehicles). These results establish the viability of integrated LSTM-CVRP frameworks for Indian MSW contexts, with full operational validation contingent on geocoding resolution for the actual SMC collection point network [8].

C8 - Transfer Station Efficiency Assessment. Using the eight real GPS coordinates of Surat's transfer stations, a macro-infrastructure level efficiency assessment established an analytical lower bound of 166.8 km/day against a 926.8 km/day current baseline - a theoretical optimization potential of 82%. This assessment constituted the first real-data test of the forecast-to-operations linkage and provided empirical grounding for the efficiency potential claim beyond the Stage 1 synthetic network [Chapter 5, §5.2].

C10 - IMPACT Framework Validation. The five-module GPS-free IMPACT framework - comprising Time Window Recalibration (TWR), Vehicle Allocation Optimization (VAO), Missed Stop Route cause System (MSRS), Workload Demand Balancing and Route Balancing (WDB+RB), and Forecast-Integration for Operations (FIO) - was validated on 662,803 rows of October 2025 real operational records across 9 zones, 131 wards, 949 vehicles, 3,341 named routes, and 455,827 point-of-interest stop visits, without requiring GPS spatial data at any stage. IMPACT is the first GPS-independent MSW operational optimization framework validated at this scale on real operational data [Chapter 5, §§5.4–5.5].

C11 - Compliance-Dominance Finding. The MSRS XGBoost classifier, achieving AUC = 0.940 and Recall = 88.1%, established through SHAP attribution analysis that contractor compliance failure accounts for 75.72% of missed-stop variance - the vehicle-level historical miss rate feature (veh_miss_pct) yielding a mean absolute SHAP value of 5.089, more than three times the next most important feature group - against 1.65% for fleet

capacity and 5.42% for scheduling factors. This finding was replicated independently across all nine collection zones. Central Zone's $\kappa = 0.31$, 9.31% missed-stop rate, and estimated minimum vehicle requirement of only 53 vehicles against 130 deployed provide a logically irrefutable counterexample to the capacity hypothesis independent of any statistical method. This finding falsifies the fleet-insufficiency assumption driving Indian ULB waste management investment [Chapter 5, §5.4.6; Chapter 5, §5.6.2].

C12 - Fleet Surplus Quantification. The Vehicle Allocation Optimization module, using Integer Linear Programming with an OPTIMAL certification from the OR-Tools CP-SAT solver, identified a 260-vehicle surplus across Surat's nine-zone operational fleet. Fleet reallocation to zone-level minimum vehicle requirements produced a miss rate change of −0.0008 percentage points, confirming that fleet quantity is not a binding operational constraint. This finding redirects the capital investment logic from fleet procurement to compliance governance infrastructure, with direct fiscal and environmental co-benefits [Chapter 5, §5.4.5].

C13 - Environmental Benefit Quantification (all Stage 1 environmental results arise from a synthetic Surat-representative network; magnitude requires operational deployment data for precise carbon accounting). The Stage 1 CVRP optimization on the synthetic Surat-representative network achieved a quantified reduction of 46.8 tonnes of CO_2 per year and 49.3 litres per day in fuel consumption relative to the baseline routing configuration, as documented in the IJAM paper [8]. The directional environmental benefit - that route optimization reduces fuel consumption and CO_2 emissions - is robust to the synthetic network qualification. The magnitude of the benefit across the full 949-vehicle operational fleet is expected to substantially exceed the Stage 1 quantification once compliance governance and scheduling optimizations are implemented [8].

8.1.4 Practical Contributions

Four contributions are primarily practical - designed artefacts, deployment frameworks, and policy alignment instruments that translate the research findings into institutional capabilities.

C14 - Four-Layer DSS Architecture. A fully specified four-layer Decision Support System architecture was designed, implemented, and documented for SMC deployment: a Data Ingestion layer collecting vehicle visit records, demand proxies, and OLS forecast inputs via SMS/REST API; a Model Execution layer running the OLS monthly forecast, RF daily demand scalar, and all five IMPACT modules; an Analytics layer computing KPIs, evaluating scenarios, and reporting SHAP attribution for administrative communication; and a User Interface layer presenting six role-differentiated dashboards with Role-Based Access Control. Every layer was documented with inputs, processing logic, outputs, and technology stack specifications, extending the IJAM conceptual DSS proposal [8] to a fully specified, empirically grounded thesis implementation [Chapter 6].

C15 - Streamlit Proof-of-Concept DSS Prototype. A functional GPS-free Streamlit proof-of-concept prototype was implemented with six dashboard modules: an Executive Dashboard presenting system-level KPI status; a Forecast Dashboard presenting OLS monthly predictions with FIO scenario flags; an IMPACT Dashboard communicating the five-module operational action plan with prioritised intervention sequence; a Compliance Dashboard surfacing zone and vehicle-level miss rate rankings and kappa coefficients; a Route Dashboard presenting WDB+RB rebalancing assignments; and a Scenario Analysis Dashboard implementing the FIO six-scenario decision framework. The prototype operates on fourteen CSV data sources, requires no GPS infrastructure, no proprietary software, and no specialised computing hardware beyond office-grade infrastructure [Chapter 6].

C16 - Three-Phase SMC Deployment Roadmap. A three-phase implementation roadmap calibrated to SMC's specific institutional context was documented: Phase 1 establishes data infrastructure through SMS/REST API-based record collection on existing administrative IT systems; Phase 2 integrates the OLS forecast and IMPACT outputs into the six dashboard modules; Phase 3 connects the DSS to real-time data streams as SMC's digital infrastructure matures. Role-Based Access Control was designed for three stakeholder persona tiers - operational collection supervisors, SMC zone managers, and municipal commissioners - each with appropriate dashboard granularity and access permissions aligned with their functional responsibilities [Chapter 6, §6.4].

C17 - Policy Alignment Framework with Specific Measured Outcomes. The integrated framework was validated against three tiers of policy mandate, each with quantified

measured outcomes. Sustainable Development Goal 11 (Sustainable Cities, Target 11.6) was addressed through TWR's system-wide ONTIME improvement of +2.09 percentage points and WDB+RB's elimination of all 232 statistically overloaded routes, reducing the demand coefficient of variation from 1.42 to 0.60. SDG 12 (Responsible Consumption, Target 12.5) was addressed through the VAO-identified 260-vehicle surplus, providing the quantitative basis for capital reallocation from fleet procurement to compliance governance. SDG 13 (Climate Action, Target 13.2) was addressed through the Stage 1 quantification of 46.8 tonnes of CO_2 per year on a synthetic Surat-representative network, with directional implication for full operational fleet deployment. Smart Cities Mission alignment was documented through the GPS-free architecture's applicability to approximately 90% of Indian ULBs without GPS fleet tracking capability. Swachh Bharat Mission 2.0 alignment was documented through the projected missed-stop rate reduction from the 2.01% operational baseline to approximately 0.4–0.5% under full IMPACT implementation [Chapter 7, §7.5].

TABLE 8.1: Summary of Research Contributions (C1–C18) Organized by Contribution Type, Research Objective, and Evidential Basis.

Code	Contribution Name	Type	Research Objective	Primary Chapter	Evidence Base	Publication
C1	17-model comparative MSW forecasting framework	Methodological	RO-1	Ch 3, 4	186-month JETS dataset; 17 models evaluated	JETS [1]
C2	419-feature systematic engineering pipeline	Methodological	RO-1	Ch 3	7 feature categories; 396 active features	JETS [1]
C3	Eight-metric evaluation protocol	Methodological	RO-1	Ch 4	R^2, MAE, RMSE, MAPE, CV, Theil_U, MDA, Residual_Std; DM tests	JETS [1]

C4	ITS methodology for COVID-19 impact	Methodological	RO-1	Ch 4	JSWTM 196-month dataset §4.5.3; 16% RMSE reduction LR	JSWTM [2]
C5	Feature engineering supersedes architecture	Methodological	RO-1	Ch 4	OLS R^2=0.9972 > LSTM R^2=−3.04 (tabular); LSTM R^2=0.93 (IJAM context)	JETS [1], IJAM [3]
C6	Computational efficiency characterisation	Empirical	RO-1	Ch 4	OLS 0.010s training; 0.006 ms inference	JETS [1]
C7	Stage 1 LSTM-CVRP proof-of-concept (synthetic)	Empirical	RO-2	Ch 5	50.31% distance reduction; 24% fleet reduction; synthetic representative network	IJAM [3]
C8	Transfer station efficiency assessment	Empirical	RO-2	Ch 5	82% efficiency potential; 926.8→166.8 km/day; 8 real GPS coords	Thesis
C9	Geocoding failure as negative result	Methodological	RO-2	Ch 5	8.6% Nominatim success; West Zone 29.3% placeholder	Thesis
C10	IMPACT GPS-free framework	Empirical	RO-2	Ch 5	662,803 rows; 9 zones; 949 vehicles; Oct 2025	Thesis
C11	Compliance-dominance finding	Empirical	RO-2	Ch 5	MSRS SHAP 75.72%; AUC=0.940; 9-zone replication; Central κ=0.31	Thesis
C12	Fleet surplus quantification (260 vehicles)	Empirical	RO-2	Ch 5	VAO ILP-OPTIMAL; OR-Tools CP-SAT	Thesis
C13	Environmental benefit	Empirical	RO-2	Ch 5, 7	46.8t CO_2/year; 49.3 L/day; Stage 1 synthetic network	IJAM [3]

	quantification (synthetic)					
C14	Four-layer DSS architecture	Practical	RO-3	Ch 6	4 layers; inputs/processing/outputs/technology stack documented	IJAM [3] + Thesis
C15	Streamlit PoC DSS prototype	Practical	RO-3	Ch 6	6 dashboards; 14 CSVs; GPS-free; 3 persona tiers	Thesis
C16	Three-phase SMC deployment roadmap	Practical	RO-3	Ch 6	Phase 1–3; RBAC; 3 stakeholder tiers	IJAM [3] + Thesis
C17	Policy alignment with measured outcomes	Practical	RO-3	Ch 6, 7	SDG 11/12/13; SCM; SBM 2.0; specific metrics per goal	IJAM [3] + Thesis
C18	Evidence-driven adaptive research design	Methodological	RO-2	Ch 5, 8	3-stage trajectory; geocoding→IMPACT pivot	Thesis

8.2 Conclusions by Research Objective

8.2.1 Research Objective 1: Machine Learning-Based MSW Forecasting Framework

Research Objective 1: To develop and comparatively evaluate a machine learning-based forecasting framework for monthly MSW generation in Surat, incorporating a systematic feature engineering pipeline, a 17-model comparative analysis across statistical, ensemble, and deep learning families, a multidimensional eight-metric evaluation protocol, and a COVID-19 interrupted-time-series impact analysis.

Research Objective 1 was met in full across Chapters 3 and 4 of this thesis.

The 419-feature systematic engineering pipeline (C2) was developed and applied to 186 months of the JETS canonical dataset, encoding all principal causal pathways for monthly

MSW generation through temporal lags, rolling statistics, socioeconomic co-variates, meteorological variables, Google Community Mobility Report proxies, event flags, and interaction terms. Of 419 engineered features, 396 remained active after near-zero-variance filtering and constituted the input feature matrix for model training and evaluation [Chapter 3].

The 17-model comparative analysis (C1) was conducted under the eight-metric evaluation protocol (C3), employing R^2, MAE, RMSE, MAPE, CV-RMSE, Theil_U, MDA, and Residual_Std with Diebold-Mariano significance testing. Ordinary Least Squares emerged as the highest-performing model, achieving $R^2 = 0.9972$, MAPE = 1.43%, MAE = 198.3 tonnes, and RMSE = 275.6 tonnes on the 186-month JETS dataset [6]. Gaussian Process Regression served as a probabilistic complement, achieving $R^2 = 0.9970$ and MAPE = 1.51%. Ridge regression achieved $R^2 = 0.9962$ and MAPE = 2.03% as a multicollinearity-robust fallback. All ensemble models - LightGBM (MAPE = 4.07%), CatBoost (MAPE = 5.22%), and XGBoost (MAPE = 4.32%) - were substantially outperformed by the linear statistical tier, as were all deep learning architectures including the Tuned MLP (MAPE = 6.24%), LSTM (MAPE = 100.01%), and GRU with attention (MAPE = 114.81%) [Chapter 4].

The central theoretical finding of Research Objective 1 (C5) is that feature engineering quality, not architectural complexity, is the dominant determinant of predictive accuracy in structured tabular MSW forecasting. The LSTM's $R^2 = -3.04$ result in the JETS tabular framework is not a fundamental architectural deficiency: the same model achieved $R^2 = 0.93$ in the IJAM paper [8] under a chronological evaluation protocol with raw monthly series features appropriate to recurrent network design. The contrast between these two results demonstrates that model-context alignment - chronological data splitting for sequential architectures - is as consequential as model selection in MSW forecasting benchmarking [Chapter 4, §4.6.2].

The COVID-19 interrupted-time-series analysis (C4) was conducted using the JSWTM 196-month raw dataset, exclusively within Section 4.5.3. A structural break dummy variable encoding the March 2020 COVID-19 onset delivered a 16% reduction in RMSE for linear models and 10% for XGBoost, establishing a replicable methodology for pandemic-period structural break accommodation [7]. Regarding deployment feasibility (C6), OLS training

at 0.010 seconds and inference at 0.006 milliseconds per sample confirm deployment on any office-grade computing infrastructure. The combined evidence of OLS accuracy, computational efficiency, and 419-feature pipeline transferability establishes OLS with the full feature pipeline as the recommended operational forecasting model for SMC and comparable Indian Urban Local Bodies [Chapter 4, §4.6.3].

8.2.2 Research Objective 2: Multi-Stage Route Optimization and Operational Efficiency Framework

Research Objective 2: To design and evaluate a multi-stage vehicle route optimization and operational efficiency framework for SMC waste collection, progressing from a published LSTM-CVRP proof-of-concept through real transfer station efficiency assessment to a GPS-free IMPACT framework validated against October 2025 operational data.

Research Objective 2 was met across all three stages of the route optimization research program, documented in Chapter 5.

Stage 1 (C7) produced the first published LSTM-CVRP integration for Indian MSW management, documented in the IJAM paper [8]. On a synthetic Surat-representative network validated against MoHUA and CPCB fleet and capacity standards, the CVRP optimization achieved a 50.31% reduction in collection distance (392.04 km/day to 194.78 km/day) and a 24% reduction in required fleet size (25 to 19 vehicles). All Stage 1 results arise from a proof-of-concept on synthetic infrastructure and do not represent operational outcomes for the actual SMC collection network. Stage 2 (C8) applied the forecasting pipeline to the eight real GPS coordinates of Surat's transfer stations, establishing a macro-infrastructure analytical lower bound of 166.8 km/day against a 926.8 km/day baseline - an 82% theoretical optimization potential at the transfer station level [Chapter 5, §5.2]. Stage 3 geocoding assessment (C9) documented an 8.6% Nominatim success rate on Surat's beat-level addresses, with West Zone at 29.3% placeholder resolution. This negative result motivated the GPS-free IMPACT design and is documented as a contribution - the first systematic characterisation of Nominatim performance on Indian municipal beat-level address data [Chapter 5, §5.3].

The GPS-free IMPACT framework (C10) was validated on 662,803 October 2025 operational records across 9 zones, 131 wards, 949 vehicles, 3,341 routes, and 455,827 stop visits. Time Window Recalibration (TWR) yielded a system-wide ONTIME improvement of +2.09 percentage points at zero capital cost [Chapter 5, §5.4.3]. Vehicle Allocation Optimization (VAO) identified a 260-vehicle surplus through ILP-OPTIMAL certification, with fleet reallocation yielding only −0.0008 percentage points improvement in miss rate - confirming fleet availability is not the binding constraint [Chapter 5, §5.4.5]. Missed Stop Risk Scoring (MSRS) established the compliance-dominance finding (C11): 75.72% SHAP attribution to contractor compliance, AUC = 0.940, Recall = 88.1%, replicated across all nine zones. Central Zone's $\kappa = 0.31$ and 9.31% miss rate despite an estimated minimum vehicle requirement of only 53 vehicles against 130 deployed provides a logically irrefutable counterexample to the capacity hypothesis [Chapter 5, §5.4.6]. Workload Demand Balancing and Route Balancing (WDB+RB) eliminated all 232 overloaded routes and reduced the system-wide demand coefficient of variation from 1.42 to 0.60 through administrative reassignment alone [Chapter 5, §5.4.7]. Forecasting Integration Optimization (FIO) formalised the integration of OLS monthly forecasting into daily operational decisions through the demand scalar $S_d = \widehat{F_d}/F_{\text{base}}$, ($F_{\text{base}}$ = 1,195,756), with a six-scenario trigger framework. At OLS MAPE = 1.43%, forecast uncertainty is structurally incapable of generating spurious Scenario B or Scenario C activations [Chapter 5, §5.4; Chapter 7, §7.2.2].

The evidence-driven adaptive research design (C18) - the three-stage trajectory from published proof-of-concept through documented negative result to GPS-free IMPACT framework - is recognised as a methodological contribution demonstrating that rigorous doctoral research under real-world data constraints can produce insights inaccessible to idealised unconstrained designs. The compliance-dominance finding was accessible precisely because the geocoding failure redirected the research from a GPS-based routing exercise toward a compliance-diagnostic framework [Chapter 5, §5.6.1].

8.2.3 Research Objective 3: Integrated Decision Support System

Research Objective 3: To design and validate an integrated Decision Support System that operationalises the forecasting and optimization outputs for SMC deployment,

encompassing a four-layer DSS architecture, forecasting-to-operations integration via a demand scalar, stakeholder-specific interfaces, and a phased municipal deployment strategy aligned with Smart Cities Mission, Swachh Bharat Mission 2.0, and SDGs 11, 12, and 13.

Research Objective 3 was met through the design, implementation, and four-form proof-of-concept validation of the IMPACT DSS documented in Chapter 6.

The four-layer DSS architecture (C14) was fully specified with inputs, processing logic, outputs, and technology stack documentation for each layer, extending the IJAM conceptual DSS proposal [8] to a thoroughly specified, empirically grounded implementation [Chapter 6]. The Streamlit proof-of-concept prototype (C15) implemented six differentiated dashboard modules on fourteen CSV data sources, operating GPS-free and validated on October 2025 operational records. The three-phase deployment roadmap (C16) provided phase-specific requirements, timelines, and cost estimates calibrated to SMC's context, with Role-Based Access Control for three stakeholder persona tiers.

The DSS was validated through four complementary forms of evidence [Chapter 6, §6.7]. Technical correctness verification confirmed that all fourteen CSV data sources loaded correctly and that the fourteen critical embedded constants - including ONTIME rates (22.0% from the SMC route Status column; 6.1% from independently parsed timestamps), global missed-stop rate (2.01%), compliance SHAP attribution (75.72%), and fleet surplus (260 vehicles) - matched verified research outputs to four significant figures. Internal consistency verification confirmed that the ONTIME improvement waterfall components are mutually consistent with the MSRS SHAP attribution and zone-level kappa coefficients. Wireframe-based expert review confirmed appropriate information granularity per stakeholder tier. Policy alignment assessment confirmed that all four SMC-relevant Key Performance Indicators in the SBM 2.0 monitoring framework are surfaced within the DSS dashboards.

The multi-temporal planning architecture integrates OLS monthly forecasting for strategic annual demand envelopes, the RF daily demand scalar for tactical pre-positioning decisions, and the FIO scalar for operational daily triggers - three planning horizons in a single DSS interface [Chapter 7, §7.2.3]. The policy alignment framework (C17) maps quantified outcomes to each SDG commitment: TWR's +2.09 percentage point ONTIME improvement and WDB+RB's CV reduction from 1.42 to 0.60 address SDG 11 Target 11.6; the 260-

vehicle surplus addresses SDG 12 Target 12.5; and the Stage 1 46.8 tonnes CO_2/year (synthetic network qualifier) addresses SDG 13 Target 13.2. The GPS-free architecture provides a Smart Cities Mission pathway for approximately 90% of Indian ULBs without GPS fleet tracking, and the projected missed-stop reduction to approximately 0.4–0.5% under full IMPACT implementation represents a quantifiable Swachh Bharat Mission 2.0 service quality improvement [Chapter 7, §7.5].

8.3 Recommendations for Municipal Authorities

The convergent evidence from the three research components points to a coherent operational prescription for Surat Municipal Corporation: compliance before fleet, forecasting before procurement, evidence before investment. Each element of this prescription is grounded in a specific IMPACT finding and is actionable through administrative rather than capital-intensive intervention.

The first and most urgent priority is the establishment of a contractor compliance monitoring infrastructure grounded in the MSRS analytical framework. The compliance-dominance finding - 75.72% SHAP attribution to contractor compliance behaviour, replicated across all nine zones - establishes that the highest-leverage operational intervention available to SMC is systematic, quantified monitoring of contractor performance at the vehicle level, using vehicle-level historical miss rates as the primary monitoring signal [Chapter 5, §5.4.6]. Central Zone's kappa coefficient of $\kappa = 0.31$ indicates a degree of compliance unreliability that no fleet-side intervention can address: a zone with the lowest demand-per-vehicle in the system and an estimated minimum vehicle requirement of only 53 vehicles against 130 deployed cannot have its 9.31% missed-stop rate reduced through any increase in fleet availability. The compliance governance response must address contractor accountability structure directly. The MSRS framework provides the diagnostic instruments - vehicle-level miss rate rankings, zone-specific kappa coefficients, and SHAP attribution profiles - at near-zero capital cost [Chapter 5, §5.6.2; Chapter 7, §7.4.1].

The second priority is the implementation of Time Window Recalibration for collection scheduling across all nine zones. The TWR analysis demonstrated a system-wide ONTIME improvement of +2.09 percentage points from optimal zone-specific shift assignment, requiring only administrative adjustment of collection schedules within existing contractor

service agreements at zero capital cost [Chapter 5, §5.4.3]. The third priority is the implementation of Workload Demand Balancing and Route Balancing before any route expansion is considered. The WDB+RB analysis eliminated all 232 statistically overloaded routes and reduced the system-wide demand coefficient of variation from 1.42 to 0.60 through administrative route reassignment without capital expenditure [Chapter 5, §5.4.7].

The fourth priority is the deferral of all fleet expansion decisions pending the VAO allocation revision. The 260-vehicle surplus identified by the ILP-OPTIMAL VAO solution demonstrates that the current collection objective can be met with a substantially smaller deployed fleet. Any procurement decision before compliance monitoring and scheduling interventions are implemented would add capacity to a system whose principal constraint is contractor behaviour rather than vehicle availability, imposing an unnecessary capital and environmental cost [Chapter 5, §5.4.5; Chapter 7, §7.7.2]. The fifth priority is the deployment of the IMPACT DSS Phase 1 data infrastructure on SMC's existing administrative IT systems. The SMS/REST API-based data architecture is implementable on office-grade infrastructure without GPS hardware, proprietary software, or specialist external technical support. Once operational, the OLS monthly forecast integration activates the FIO scenario monitoring capability, providing SMC with evidence-based demand-driven route planning and the strategic lead time - estimated at 12–18 months ahead of any fleet capacity constraint - to make informed operational decisions proactively [Chapter 6, §6.4; Chapter 7, §7.2.3].

For comparable Indian Urban Local Bodies considering adoption of the integrated framework, a four-step replication protocol was established [Chapter 7, §7.7.3]: (1) a data infrastructure and geocoding feasibility audit, including a Nominatim assessment on a 500-address sample to determine GPS-dependent or GPS-free pathway; (2) deployment of the 419-feature forecasting framework with city-specific localisation; (3) IMPACT framework deployment, with the critical caveat that MSRS compliance attribution must be independently re-estimated per city - the 75.72% finding is Surat-specific; and (4) DSS phased deployment beginning with Phase 1 data infrastructure.

8.4 Limitations of the Current Research

The integrated framework was accompanied by scope boundaries acknowledged explicitly throughout the thesis. This section consolidates those boundaries, assesses interactions between them, and identifies the mitigations applied in each case. The Level 5 integration claim refers to architectural and evidential scope, not to operational deployment status. The Streamlit DSS prototype remains a proof-of-concept artefact, not yet in routine municipal use by SMC.

The first limitation is the single-city scope of all quantitative results. Every finding in this thesis - the OLS forecasting accuracy, the 75.72% compliance attribution, the 260-vehicle fleet surplus, the 2.01% missed-stop rate, and the +2.09 percentage point ONTIME improvement - was derived from Surat's specific operational context. Whether the compliance-dominance finding transfers to cities with direct ULB operation rather than contractor operation, or to cities where GPS-based compliance monitoring is already operational, is an empirical question the single-city study cannot answer. The framework design mitigates this limitation through its GPS-free, open-source, public-data-dependent architecture [Chapter 7, §7.4.2]. Multi-city validation is the primary future research direction.

The second limitation concerns the synthetic infrastructure underlying Stage 1 results (C7 and C13). The 50.31% distance reduction, 24% fleet reduction, 46.8 tonnes CO_2/year, and 49.3 litres/day fuel reduction documented in the IJAM paper [8] arise from a proof-of-concept on a synthetic Surat-representative network. The directional claims are methodologically supported. The quantitative magnitudes cannot be confirmed as operational outcomes without resolving the geocoding challenge and applying the CVRP framework to the actual SMC collection point network.

The third limitation is the single-month temporal scope of all Stage 3 IMPACT analyses. October 2025 represents a moderate-demand operating month, not the peak monsoon period and not the major festival demand surge. Whether the 75.72% compliance dominance holds at the same magnitude during monsoon months - when traffic disruption and road flooding may introduce capacity-side factors absent in October - is the most important remaining empirical question. The nine-zone replication and the Central Zone natural experiment

provide strong structural evidence for the compliance-first conclusion's robustness; seasonal sensitivity can only be resolved through multi-month IMPACT deployment [Chapter 5, §5.6.4; Chapter 7, §7.6.2].

The fourth limitation is the GPS-free constraint and its consequence for spatial routing precision. The 8.6% Nominatim success rate made GPS-dependent VRPTW infeasible at the collection-point level. This constraint was reframed as a design condition motivating Contribution C15's GPS-free architecture. A practical resolution pathway - commercial geocoding of the estimated 5,000–10,000 unique POI names at an indicative cost of USD 5–15 using the Google Maps Geocoding API - is documented as the most direct enabler for subsequent full-scale VRPTW deployment [Chapter 5, §5.3].

The fifth limitation is the proof-of-concept status of the IMPACT DSS. The DSS was validated through four complementary forms of evidence but was not operationally deployed in a production SMC environment. The projected ONTIME improvements have not been empirically confirmed through a longitudinal intervention study. Full operational validation constitutes the most important practical future research direction [Chapter 6, §6.7]. The sixth limitation is the demand-proxy approximation: demand per collection stop was operationalised as houses plus shops count, which may not accurately represent actual waste quantities at all stop types. Integration of actual per-stop weighed tonnage would substantially improve WDB+RB and FIO precision [Chapter 6, §6.8]. The seventh limitation is the non-implementation of the Multi-Objective Optimization component - the ε-constraint Pareto frontier across the four IMPACT objectives was fully formulated but not implemented within the thesis scope, constituting the primary quantitative extension for future research [Chapter 5, §5.6.4]. The eighth limitation is the monthly temporal granularity of the forecasting framework: the two-stage monthly-to-daily disaggregation introduces cumulative uncertainty at the daily level not captured in the OLS monthly evaluation metrics, requiring IoT-generated daily data streams for sub-monthly demand forecasting.

8.5 Future Research Directions

8.5.1 Multi-City Validation of the Integrated Framework

The most consequential extension of the Surat integrated framework is empirical replication in comparable Indian Urban Local Bodies and developing-economy urban contexts. The single-city scope means that the thesis's most policy-prescriptive finding - the 75.72% compliance dominance - has been demonstrated for Surat's specific institutional context and cannot be assumed to transfer unchanged across the diversity of Indian ULB operational structures.

Multi-city validation should prioritise tier-II Indian cities with contractor-operated collection systems, comparable zone and ward administrative architectures, and a minimum of five years of continuous monthly MSW generation records - conditions met by Vadodara, Rajkot, Ahmedabad, Nagpur, and several cities within the Smart Cities Mission cohort. For each adopting city, the 419-feature engineering pipeline would require localisation of city-specific socioeconomic indicators, IMD meteorological records, and local event calendars; the remaining pipeline components use publicly available national or global data sources requiring no city-specific adaptation.

Critically, the MSRS compliance attribution model must be independently re-estimated for each adopting city before its output informs governance prescriptions. A city operating under direct ULB employment rather than contractor-operated service delivery faces a qualitatively different accountability architecture that may produce a different attribution profile. Multi-month data collection - expanding the October 2025 cross-sectional dataset to a twelve-month rolling record - would resolve the seasonal sensitivity question identified in §8.4, enabling empirical parameterisation of monsoon-period TWR recalibration and seasonal WDB+RB rebalancing. Beyond the Indian context, the GPS-free design principle addresses a constraint shared by cities in sub-Saharan Africa, Southeast Asia, and South Asia with comparable informal address systems, making the IMPACT framework a candidate for multi-country replication.

8.5.2 IoT Integration and Real-Time Data Streams

The IMPACT DSS Phase 3 deployment, as specified in the Chapter 6 implementation roadmap, is contingent on the maturation of SMC's digital infrastructure to support real-time data streams. Three IoT integration directions would substantially enhance the framework's operational intelligence.

GPS fleet tracking hardware, once deployed, would enable real-time route adherence monitoring at the vehicle level, replacing the historical veh_miss_pct compliance predictor with a live compliance signal and eliminating the measurement circularity limitation identified in §8.4. Dynamic route replanning in response to real-time vehicle breakdowns, road blockages, or demand surges would become computationally feasible using the IMPACT framework's algorithmic components as the optimization engine. Road-network-aware routing via OSRM or the Google Maps Distance Matrix API, combined with commercial geocoding resolution of Surat's unique POI names, would enable the full-scale VRPTW formulation whose absence at Stage 3 motivated the GPS-free IMPACT design.

Smart bin fill-level sensors at high-volume stops would replace the houses-plus-shops demand proxy with direct fill-level measurements, substantially improving WDB+RB rebalancing precision and FIO demand accuracy. The FIO framework's disaggregation architecture is designed to accommodate sensor-based demand signals - the RF daily demand scalar can be retrained on fill-level data once a sufficient historical record is accumulated. Sub-monthly demand forecasting, using IoT-generated daily data series to complement the OLS monthly model, would reduce the disaggregation uncertainty associated with the current two-stage estimation process and enable the FIO trigger framework to operate at daily demand resolution.

8.5.3 Multi-Objective Optimization and Reinforcement Learning

The MOO component specified in Chapter 5 but not implemented within the thesis scope represents the primary quantitative extension of the IMPACT framework. The ε-constraint method applied to the four IMPACT objectives - maximising ONTIME rate, minimising missed-stop rate, minimising the route demand coefficient of variation, and minimising vehicle deployment cost - would produce a Pareto frontier from which SMC decision-makers

could select operating points reflecting institutional policy priorities. The ε-constraint formulation is already specified and constitutes a computational rather than conceptual extension, requiring multi-objective solver infrastructure available through OR-Tools or the DEAP evolutionary algorithm library.

Reinforcement Learning offers a complementary approach to the compliance governance challenge identified by MSRS. The MSRS kappa coefficients for each zone and vehicle provide natural reward signals for a multi-agent RL framework in which contractors receive compliance incentives calibrated to their kappa trajectory - zones moving from $\kappa = 0.31$ toward $\kappa = 0.70$ or above receiving increasing performance bonuses, while persistent kappa deterioration triggers escalating contractual consequences. Dynamic incentive adaptation through RL represents a methodologically coherent extension of the compliance-first governance narrative established in Chapter 7.

Advanced VRP formulations, contingent on geocoding resolution, would extend the Stage 1 single-depot CVRP to multi-depot VRPTW, accommodating Surat's eight transfer stations as alternative depot options. Electric vehicle fleet routing, aligned with India's municipal waste collection electrification roadmap, would extend the thesis's SDG 13 climate contribution by optimising routes for battery range and charging station access constraints in addition to demand coverage.

8.5.4 Circular Economy and Extended Producer Responsibility

The operational scope of the thesis was confined to undifferentiated mixed municipal solid waste. An important extension would incorporate source segregation and material recovery objectives, aligning the DSS with India's advancing circular economy policy environment under the Solid Waste Management Rules (2016), the Plastic Waste Management Rules (2022 amendment), and the Extended Producer Responsibility guidelines issued by the Ministry of Environment, Forest and Climate Change.

Source segregation optimization would incorporate dry and wet waste stream separation into the routing and vehicle allocation decisions of the IMPACT framework, enabling zone-level scheduling that differentiates collection windows for segregated waste categories and minimises contamination losses. Material recovery facility routing would optimise transfer

station logistics for recyclable material value maximisation, enabling SMC to increase the tonnage of recovered materials reaching authorised processing facilities. Extended Producer Responsibility integration would embed EPR material flow tracking into the DSS Analytics layer, enabling SMC to report on category-level recyclable material recovery for MoHUA and CPCB EPR compliance.

The quantification of the avoided social cost of carbon from full-scale operational IMPACT deployment - extending the Stage 1 CO_2 reduction estimate of 46.8 tonnes/year from a synthetic 19-vehicle network to the full 949-vehicle operational fleet - would provide the carbon accounting evidence base aligned with India's Nationally Determined Contributions under the Paris Agreement. This full-fleet carbon accounting would constitute a contribution to the literature on municipal solid waste management's quantified role in urban emissions reporting under the Paris framework, and would provide SMC with a verified climate performance metric for Smart Cities Mission and Swachh Bharat Mission 2.0 reporting.

8.6 Chapter Summary

All three research objectives of this thesis were met. Research Objective 1 was closed by the OLS forecasting model, which achieved $R^2 = 0.9972$ and MAPE = 1.43% across a 17-model comparative analysis of 186 months of Surat MSW data with 419 engineered features - establishing that systematic feature engineering, applied to a comprehensive causal representation of monthly waste generation drivers, is the primary determinant of predictive accuracy in structured tabular MSW forecasting, and that architectural complexity without representational investment is counterproductive in this domain. Research Objective 2 was closed by the three-stage route optimization program, culminating in the GPS-free IMPACT framework validated on 662,803 real October 2025 operational records, which established through AUC = 0.940, nine-zone replication, and the Central Zone natural experiment that contractor compliance failure - not fleet capacity - accounts for 75.72% of missed collection stops, fundamentally reorienting the operational governance prescription from fleet procurement to compliance monitoring. Research Objective 3 was closed by the four-layer IMPACT DSS, validated through four complementary forms of evidence and designed for deployment on SMC's existing administrative infrastructure without GPS hardware, proprietary software, or specialised computing resources.

The eighteen research contributions documented in this thesis - spanning methodological frameworks, empirical findings, and deployable analytical artefacts - collectively constitute a complete, evidence-grounded, GPS-free, open-source MSW management intelligence system whose central policy argument is expressed in six words: compliance before fleet, evidence before investment. Forecasting accuracy, operational optimization, and decision support infrastructure are analytically powerful instruments; they deliver their fullest institutional value only when the governance structure that converts their outputs into optimizational accountability is in place. The IMPACT framework provides the diagnostic infrastructure. The DSS provides the institutional interface. The three peer-reviewed publications provide the evidential foundation. The next step is deployment.

REFERENCES

[1] S. Kaza, L. Yao, P. Bhada-Tata, and F. V. Woerden, *What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050*. Washington, DC: World Bank, 2018.

[2] M. of E. CPCB Forest and Climate Change, Govt. of India, "Annual Report For the year 2018-19 On Implementation of Solid Waste Management Rules (As per provision 24(4) of SWM Rules,16)," Central Pollution Control Board, Delhi, Delhi, Annual Report 2018-19, Dec. 2019. Accessed: Aug. 30, 2025. [Online]. Available: https://cpcb.nic.in/uploads/MSW/MSW_AnnualReport_2018-19.pdf

[3] Ministry of Environment, Forest and Climate Change, "Solid Waste Management Rules, 2016," Government of India, New Delhi, India, S.O. 1357(E), Apr. 2016. [Online]. Available: https://static.pib.gov.in/WriteReadData/specificdocs/documents/2026/jan/doc2026129773501.pdf

[4] Ministry of Housing and Urban Affairs, Government of India, "Smart Cities Mission: Mission Statement and Guidelines," MoHUA, New Delhi, 2015.

[5] Ministry of Housing and Urban Affairs, Government of India, "Swachh Bharat Mission – Urban 2.0: Operational Guidelines," MoHUA, New Delhi, 2021.

[6] A. Rathod and others, "A Comparative Study of Advanced Machine Learning and Deep Learning Models for Municipal Solid Waste Forecasting: A Case Study of Surat, India," *Journal of Engineering and Technological Sciences*, 2026.

[7] A. Rathod and others, "Data-Driven Automation for Solid Waste Forecasting Using Machine Learning, with COVID-19 Impact Analysis: A Case Study of Surat, Gujarat," *Journal of Solid Waste Technology and Management*, vol. 52, no. 2, p. 515, 2026, doi: 10.5276/jswtm/iswmaw/522/2026.515.

[8] A. Rathod and others, "AI-Driven Decision Support System for Optimized Municipal Solid Waste Management: A Machine Learning and Route Optimization Framework for Smart Cities," *International Journal of Applied Mathematics*, vol. 38, no. 12s, 2025.

[9] United Nations, *Transforming Our World: The 2030 Agenda for Sustainable Development*. New York: United Nations, 2015.

[10] Y. Zhao *et al.*, "Prediction of municipal solid waste generation and analysis of dominant variables in rapidly developing cities based on machine learning—a case study of China," *Waste Management and Research*, 2024, doi: 10.1177/0734242X231192766.

[11] I. Markov, M. Bierlaire, J.-F. Cordeau, Y. Maknoon, and S. Varone, "Waste collection inventory routing with non-stationary stochastic demands," *Computers and Operations Research*, vol. 113, p. 104798, 2020, doi: 10.1016/j.cor.2019.104798.

[12] I. Cárdenas-León, M. Koeva, P. Nourian, and C. Davey, “Urban digital twin-based solution using geospatial information for solid waste management,” *Sustainable Cities and Society*, vol. 115, p. 105798, Nov. 2024, doi: 10.1016/j.scs.2024.105798.

[13] M. K. Younes, Z. M. Nopiah, N. E. A. Basri, H. Basri, M. F. M. Abushammala, and K. N. A. Maulud, “Prediction of municipal solid waste generation using nonlinear autoregressive network,” *Environ Monit Assess*, vol. 187, no. 12, p. 753, Dec. 2015, doi: 10.1007/s10661-015-4977-5.

[14] M. Adamović, “Prediction of municipal solid waste generation using artificial neural network approach enhanced by structural break analysis,” *Environmental Science and Pollution Research*, vol. 24, no. 1, pp. 299–311, 2017, doi: 10.1007/s11356-016-7767-x.

[15] B. H. Abbasi and A. El Hanandeh, “Forecasting municipal solid waste generation using artificial intelligence modelling approaches,” *Waste Management*, vol. 56, pp. 13–22, Oct. 2016.

[16] L. Yang, Y. Zhao, X. Niu, Z. Song, Q. Gao, and J. Wu, “Municipal Solid Waste Forecasting in China Based on Machine Learning Models,” *Frontiers in Energy Research*, vol. 9, p. 763977, 2021, doi: 10.3389/fenrg.2021.763977.

[17] C. Zhang, H. Dong, Y. Geng, H. Liang, and X. Liu, “Machine learning based prediction for China’s municipal solid waste under the shared socioeconomic pathways,” *Journal of Environmental Management*, vol. 312, p. 114918, 2022, doi: 10.1016/j.jenvman.2022.114918.

[18] X. C. Nguyen *et al.*, “Development of machine learning - based models to forecast solid waste generation in residential areas: A case study from Vietnam,” *Resources, Conservation and Recycling*, vol. 167, p. 105381, Apr. 2021, doi: 10.1016/j.resconrec.2020.105381.

[19] W. Lu, W. Huo, H. Gulina, and C. Pan, “Development of machine learning multi-city model for municipal solid waste generation prediction,” *Front. Environ. Sci. Eng.*, vol. 16, no. 9, p. 119, Sep. 2022, doi: 10.1007/s11783-022-1551-6.

[20] A. Namoun, B. R. Hussein, A. Tufail, A. Alrehaili, T. A. Syed, and O. BenRhouma, “An ensemble learning based classification approach for the prediction of household solid waste generation,” *Sensors*, vol. 22, no. 9, p. 3506, 2022, doi: 10.3390/s22093506.

[21] O. Mudannayake, D. Rathnayake, J. D. Herath, D. K. Fernando, and M. Fernando, “Exploring machine learning and deep learning approaches for multi-step forecasting in municipal solid waste generation,” *IEEE Access*, vol. 10, pp. 122570–122585, 2022, doi: 10.1109/ACCESS.2022.3225493.

[22] M. Cubillos, “Multi-site household waste generation forecasting using a deep learning approach,” *Waste Management*, vol. 115, pp. 8–14, 2020, doi: 10.1016/j.wasman.2020.06.046.

[23] D. Niu, F. Wu, S. Dai, S. He, and B. Wu, “Detection of long-term effect in forecasting municipal solid waste using a long short-term memory neural network,” *Journal of Cleaner Production*, vol. 290, p. 125187, Mar. 2021, doi: 10.1016/j.jclepro.2020.125187.

[24] F. Ghanbari, H. Kamalan, and A. Sarraf, “Predicting solid waste generation based on the ensemble artificial intelligence models under uncertainty analysis,” *J Mater Cycles Waste Manag*, vol. 25, no. 2, pp. 920–930, Mar. 2023, doi: 10.1007/s10163-023-01589-9.

[25] V. Jayaraman, A. R. Lakshminarayanan, A. A. A. Khan, and K. J. Sathick, “MSW-DeepStack: Innovative Municipal Solid Waste Prediction Model for Informed Decision-Making,” *IJEI*, vol. 8, no. 1, pp. 41–51, Feb. 2025, doi: 10.18280/ijei.080105.

[26] [Initials TBC] Xia and others, “Forecasting heterogeneous municipal solid waste generation via Bayesian-optimised neural network with ensemble learning for improved generalisation,” *Computers and Chemical Engineering*, vol. 166, p. 107946, 2022, doi: 10.1016/j.compchemeng.2022.107946.

[27] S. Wan, T. Mahmud, and others, “Prophet time series modeling of waste disposal rates in the context of COVID-19,” *Sustainable Cities and Society*, vol. 105, p. 105205, 2024, doi: 10.1016/j.scs.2024.105205.

[28] D. Singh and A. Satija, “Prediction of municipal solid waste generation for optimum planning and management with artificial neural network—case study: Faridabad City in Haryana State (India),” *International Journal of System Assurance Engineering and Management*, vol. 9, pp. 91–97, 2018, doi: 10.1007/s13198-016-0484-5.

[29] O. O. Ayeleru, L. I. Fajimi, B. O. Oboirien, and P. A. Olubambi, “Forecasting municipal solid waste quantity using artificial neural network and supported vector machine techniques: A case study of Johannesburg, South Africa,” *Journal of Cleaner Production*, vol. 289, p. 125671, Mar. 2021, doi: 10.1016/j.jclepro.2020.125671.

[30] K. Lin, Y. Zhao, and J.-H. Kuo, “Deep learning hybrid predictions for the amount of municipal solid waste: A case study in Shanghai,” *Chemosphere*, vol. 307, p. 136119, Nov. 2022, doi: 10.1016/j.chemosphere.2022.136119.

[31] Z. Ceylan, “Estimation of municipal waste generation of Turkey using socio-economic indicators by Bayesian optimization tuned Gaussian process regression,” *Waste Management and Research*, vol. 38, no. 9, pp. 1008–1021, 2020, doi: 10.1177/0734242X20906877.

[32] U. Smailbegovic, E. Kadric, and S. Glöser-Chahoud, “Assessing the determinants of municipal solid waste generation in Europe: A machine learning approach,” *Waste Management*, vol. 205, p. 115033, 2025, doi: 10.1016/j.wasman.2025.115033.

[33] V. Nourani and others, “Predicting municipal solid waste generation using artificial intelligence: A hybrid approach of entropy analysis and SHAP for optimal feature selection,” *Waste Management*, vol. 205, p. 115012, 2025, doi: 10.1016/j.wasman.2025.115012.

[34] J. Guo, H. Sun, and B. Du, “Multivariable Time Series Forecasting for Urban Water Demand Based on Temporal Convolutional Network Combining Random Forest Feature

Selection and Discrete Wavelet Transform," *Water Resour Manage*, vol. 36, no. 9, pp. 3385–3400, Jul. 2022, doi: 10.1007/s11269-022-03207-z.

[35] W. Hao, A. Cominola, and A. Castelletti, "Comparing Predictive Machine Learning Models for Short- and Long-Term Urban Water Demand Forecasting in Milan, Italy," *IFAC-PapersOnLine*, vol. 55, no. 33, pp. 92–98, Jan. 2022, doi: 10.1016/j.ifacol.2022.11.015.

[36] O. Adeleke, T.-C. Jen, and others, "Explainable AI and machine learning-based analysis of municipal solid waste generation rate: a South African case study," *Waste Management*, vol. 206, p. 115036, Aug. 2025, doi: 10.1016/j.wasman.2025.115036.

[37] R. Liang, C. Chen, A. Kumar, and others, "State-of-the-art applications of machine learning in the life cycle of solid waste management," *Frontiers of Environmental Science and Engineering*, vol. 17, no. 4, p. 44, 2023, doi: 10.1007/s11783-023-1644-x.

[38] Q. Que, J. Gao, and Y. Qian, "Water demand forecasting in multiple district metered areas based on a multi-scale correction module neural network architecture," *Water Research X*, vol. 25, p. 100269, Dec. 2024, doi: 10.1016/j.wroa.2024.100269.

[39] V. Jayaraman, A. Raj Lakshminarayanan, S. Parthasarathy, and A. Suganthy, "Forecasting the Municipal Solid Waste Using GSO-XGBoost Model," *Intelligent Automation & Soft Computing*, vol. 37, no. 1, pp. 301–320, 2023, doi: 10.32604/iasc.2023.037823.

[40] M. Cieśla and A. Kosior, "Impact of COVID-19 pandemic on amount of municipal waste generation – a case study," *Journal of Civil Engineering, Environment and Architecture*, vol. 69, pp. 5–16, Oct. 2022, doi: 10.7862/rb.2022.1.

[41] D. Kong, C. Liu, and L. Han, "COVID-19 lockdown and municipal solid waste: evidence from the discarding records of 252 communities in China," *Envir. Dev. Econ.*, vol. 28, no. 4, pp. 307–331, Aug. 2023, doi: 10.1017/S1355770X22000353.

[42] Y. Ishimura, J. Yang, and K. Yamaguchi, "Tracing the impact of COVID-19 mitigation on municipal solid waste disposal during and after the pandemic." May 2024. doi: 10.2139/ssrn.4845017.

[43] J. Peula and others, "Waste composition analysis in Granada, Spain: Pre- and post-COVID-19 changes," *Journal of Material Cycles and Waste Management*, vol. 25, 2023, doi: 10.1007/s10163-023-01671-2.

[44] V. D. Muñoz-Soto, S. Dávila-Gálvez, A. E. Pérez, L. Rojo-González, N. Valenzuela-Levi, and Ó. C. Vásquez, "COVID-19, waste production and municipal recycling programs: Insights from Chile to the global south," *Science of The Total Environment*, vol. 899, p. 165388, Nov. 2023, doi: 10.1016/j.scitotenv.2023.165388.

[45] A. Sulemana, E. Donkor, E. K. Forkuo, and S. Oduro-Kwarteng, "Optimal routing of solid waste collection trucks: A review of methods," *Journal of Engineering*, p. 4586376, 2018, doi: 10.1155/2018/4586376.

[46] A. M. Restrepo-Franco, O. Valencia-Rodríguez, and E. M. Toro-Ocampo, "The vehicle routing problem as applied to residential solid waste collection operations:

Systematic literature review," *10.5267/j.ijiec*, vol. 16, no. 1, pp. 197–220, 2025, doi: 10.5267/j.ijiec.2024.10.002.

[47] B.-I. Kim, S. Kim, and S. Sahoo, "Waste collection vehicle routing problem with time windows," *Computers and Operations Research*, vol. 33, pp. 3624–3642, 2006, doi: 10.1016/j.cor.2006.05.005.

[48] Q. Qiao, F. Tao, H. Wu, X. Yu, and M. Zhang, "Optimization of a Capacitated Vehicle Routing Problem for Sustainable Municipal Solid Waste Collection Management Using the PSO-TS Algorithm," *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, p. 2163, 2020, doi: 10.3390/ijerph17062163.

[49] M. Mes, M. Schutten, and A. Pérez Rivera, "Inventory routing for dynamic waste collection," *Waste Management*, vol. 34, no. 9, pp. 1564–1576, Sep. 2014, doi: 10.1016/j.wasman.2014.05.011.

[50] A. R. Rathod and V. M. Patel, "AI-Driven Decision Support System for Optimized Municipal Solid Waste Management: A Machine Learning and Route Optimization Framework for Smart Cities," *International Journal of Applied Mathematics*, vol. 38, no. 12s, pp. 2465–2488, 2025, doi: 10.12732/ijam.v38i12s.1565.

[51] E. B. Tirkolaee, V. Simic, M. Ghobakhloo, B. Foroughi, S. Asadi, and M. Iranmanesh, "Integrated design of a sustainable waste management system with co-modal transportation network: A robust bi-level decision support system," *Journal of Cleaner Production*, vol. 449, p. 141760, Apr. 2024, doi: 10.1016/j.jclepro.2024.141760.

[52] V. Sanjeevi and P. Shahabudeen, "Optimal routing for efficient municipal solid waste transportation by using ArcGIS application in Chennai, India," *Waste Management and Research*, vol. 34, no. 1, pp. 11–17, 2016, doi: 10.1177/0734242X15607430.

[53] P. Kumar, R. Singh, S. Shukla, and S. Behera, "A GIS-based route optimization approach for municipal solid waste management," *Indian Journal of Environmental Protection*, vol. 42, pp. 1616–1623, Oct. 2022.

[54] V. Navin Ganesh, K. Vishnuvardhan, and K. Sakthiprasanth, "Neural-network-driven approach in optimization of municipal solid waste collection integrated with geo-spatial techniques," *Global NEST Journal*, Oct. 2024, doi: 10.30955/gnj.06187.

[55] S. Nurgaliyeva, M. Amangali, Z. Basheyeva, N. Kashkimbayeva, and D. Amantayev, "Design and evaluation of an intelligent waste monitoring system based on RGIS integration for smart cities," *EEJET*, vol. 4, no. 9 (136), pp. 70–78, Aug. 2025, doi: 10.15587/1729-4061.2025.337033.

[56] S. Yalcinkaya and M. Uzer, "GIS and multi-criteria decision analysis for optimal solid waste collection point locations," *Chemical Engineering Transactions*, vol. 76, pp. 209–214, 2019.

[57] K. Abeliotis, K. Karaiskou, A. Togia, and K. Lasaridi, “Decision support systems in solid waste management: A case study at the national and local level in Greece,” *Global NEST Journal*, vol. 11, no. 2, pp. 117–126, 2009, doi: 10.30955/gnj.000590.

[58] I. Sosunova and J. Porras, “IoT-enabled smart waste management systems for smart cities: A systematic literature review,” *IEEE Access*, vol. 10, pp. 73326–73363, 2022, doi: 10.1109/ACCESS.2022.3188308.

[59] I. Sosunova, A. Wolff, J. Porras, and A. Happonen, “Smarter Waste: A decision support framework for smart waste management systems,” *International Journal of Social Ecology and Sustainable Development*, vol. 15, no. 1, pp. 1–18, 2024, doi: 10.4018/IJSESD.361770.

[60] S. Kumar, R. V. Singh, and K. Kumar, “IoT-Based Management, Monitoring, Reporting, and Evaluation (MMRE) Implementation Strategy for Dhanbad Municipal Solid Waste Management and Disposal,” *Journal of Information Systems Engineering and Management*, vol. 10, no. 51s, May 2025, doi: 10.52783/jisem.v10i51s.10625.

[61] L. Barth, L. Schweiger, R. Benedech, and M. Ehrat, “From data to value in smart waste management: Optimizing solid waste collection with a digital twin-based decision support system,” *Decision Analytics Journal*, vol. 9, p. 100347, Dec. 2023, doi: 10.1016/j.dajour.2023.100347.

[62] A. Rathod and others, “AI-Driven Decision Support System for Optimized Municipal Solid Waste Management: A Machine Learning and Route Optimization Framework for Smart Cities,” *Int. J. Appl. Math.*, vol. 38, no. 12s, 2025.

[63] A. Rathod and others, “Data-Driven Automation for Solid Waste Forecasting Using Machine Learning, with COVID-19 Impact Analysis: A Case Study of Surat, Gujarat,” *J. Solid Waste Technol. Manag.*, vol. 52, no. 2, pp. 515-, 2026, doi: 10.5276/jswtm/iswmaw/522/2026.515.

[64] A. Rathod and others, “A Comparative Study of Advanced Machine Learning and Deep Learning Models for Municipal Solid Waste Forecasting: A Case Study of Surat, India,” *J. Eng. Technol. Sci.*, 2026.

[65] Office of the Registrar General and Census Commissioner India, “Census of India 2011: Primary Census Abstract.” Govt. of India Press, New Delhi, 2013.

[66] Reserve Bank of India, “Handbook of Statistics on Indian Economy.” Mumbai, 2024.

[67] Ministry of Statistics and Programme Implementation, “Consumer Price Index (Urban) — Monthly Data Series.” New Delhi, 2024.

[68] Petroleum Planning and Analysis Cell (PPAC), “Retail Selling Prices of Petroleum Products — Monthly Data.” New Delhi, 2024.

[69] India Meteorological Department (IMD), “Monthly Weather Summaries — Surat Station.” Pune, 2024.

[70] K. Miezah, K. Obiri-Danso, Z. Kádár, B. Fei-Baffoe, and M. Y. Mensah, "Municipal solid waste characterization and quantification as a measure towards effective waste management in Ghana," *Waste Manag.*, vol. 46, pp. 15–27, 2015, doi: 10.1016/j.wasman.2015.09.009.

[71] Google LLC, "COVID-19 Community Mobility Reports." 2020.

[72] S. K. Singh, P. Singh, and A. K. Bajpai, "Effect of festivals on solid waste generation: A case study," *Procedia Environmental Sciences*, vol. 35, pp. 169–175, 2016.

[73] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne: OTexts, 2021.

[74] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, 3rd ed. New York: Springer, 2016.

[75] F. Pedregosa and others, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[76] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York: Springer, 2009.

[77] J. Brownlee, "Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python." 2018.

[78] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[79] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[80] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B*, vol. 67, no. 2, pp. 301–320, 2005.

[81] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[82] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual Workshop on Computational Learning Theory (COLT '92)*, Pittsburgh, PA, 1992, pp. 144–152.

[83] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[84] G. Ke and others, "LightGBM: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, 2017.

[85] A. V. Dorogush, V. Ershov, and A. Gulin, "CatBoost: Gradient boosting with categorical features support." 2018.

[86] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[87] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[88] K. Cho and others, “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1724–1734.

[89] X. Shi and others, “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in *Advances in Neural Information Processing Systems*, 2015.

[90] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ: Wiley, 2015.

[91] C. C. Holt, “Forecasting seasonals and trends by exponentially weighted moving averages,” *International Journal of Forecasting*, vol. 20, no. 1, pp. 5–10, 2004.

[92] S. J. Taylor and B. Letham, “Forecasting at scale,” *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

[93] P. Toth and D. Vigo, Eds., *The Vehicle Routing Problem*. Philadelphia, PA: SIAM, 2002.

[94] Google Developers, “OR-Tools: Open Source Software for Combinatorial Optimization.” 2023.

[95] E. Turban, J. E. Aronson, and T.-P. Liang, *Decision Support Systems and Intelligent Systems*, 7th ed. Upper Saddle River, NJ: Prentice Hall, 2005.

[96] Central Pollution Control Board, “Annual Report on Municipal Solid Waste Management in India,” MoEFCC, Government of India, New Delhi, 2022.

[97] Ministry of Housing and Urban Affairs, “Annual Report 2022–23,” Government of India, 2023. [Online]. Available: https://mohua.gov.in/

[98] Office of the Registrar General and Census Commissioner, India, “Census of India 2011: Primary Census Abstract,” Government of India Press, New Delhi, 2013.

[99] Ministry of Statistics and Programme Implementation, “Ministry of Statistics and Program Implementation | Government Of India.” 2026. [Online]. Available: https://www.mospi.gov.in/publications-reports/innerpage/434

[100] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ: Wiley, 2015.

[101] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne: OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/

[102] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York: Springer, 2013. doi: 10.1007/978-1-4614-6849-3.

[103] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York: Springer, 2009. doi: 10.1007/978-0-387-84858-7.

[104] M. A. Abdoli, M. Falah Nezhad, R. S. Salehi, and S. Behboudian, "Longterm forecasting of solid waste generation by the artificial neural networks," *Environmental Progress & Sustainable Energy*, vol. 31, no. 4, pp. 628–636, 2012.

[105] N. Bandara, C. Hettiaratchi, S. Wirasinghe, and S. Pilapiiya, "Relation of waste generation and composition to socio-economic factors: a case study," *Environmental Monitoring and Assessment*, vol. 135, pp. 31–39, 2007, doi: 10.1007/s10661-007-9636-3.

[106] S. Nabaa and N. Hadi, "Estimation of Municipal Solid Waste Generation Rate (Case Study of Hilla City)," *Journal*, vol. 12, pp. 72–80, Jan. 2014.

[107] Google LLC, "COVID-19 Community Mobility Reports." 2022. [Online]. Available: https://www.google.com/covid19/mobility/

[108] P. Gupta, A. Sharma, and L. Bhardwaj, "Solid Waste Management (SWM) and Its Effect on Environment & Human Health," p. 7, 2024, doi: 10.58532/V3BCAG4P1CH7.

[109] D. Hoornweg and P. Bhada-Tata, "What a Waste: A Global Review of Solid Waste Management," World Bank Urban Development Series, 15, 2012. doi: 10.1596/27665.

[110] A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970, doi: 10.1080/00401706.1970.10488634.

[111] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society B*, vol. 58, no. 1, pp. 267–288, 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.

[112] H. Zou and T. Hastie, "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society B*, vol. 67, no. 2, pp. 301–320, 2005, doi: 10.1111/j.1467-9868.2005.00503.x.

[113] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[114] Z. Ghahramani, "Probabilistic machine learning and artificial intelligence," *Nature*, vol. 521, pp. 452–459, May 2015, doi: 10.1038/nature14541.

[115] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.

[116] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[117] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[118] G. Ke *et al.*, “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[119] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: Unbiased Boosting with Categorical Features,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

[120] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[121] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[122] D. Bahdanau, K. Cho, and Y. Bengio, “Neural Machine Translation by Jointly Learning to Align and Translate,” in *Proceedings of ICLR*, 2015.

[123] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W. Wong, and W. Woo, “Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2015, pp. 802–810.

[124] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[125] S. Raschka and V. Mirjalili, *Python Machine Learning*, 3rd ed. Birmingham: Packt Publishing, 2019.

[126] P. R. Winters, “Forecasting Sales by Exponentially Weighted Moving Averages,” *Management Science*, vol. 6, no. 3, pp. 324–342, 1960, doi: 10.1287/mnsc.6.3.324.

[127] A. Abdeljaber, S. Al Smadi, M. Abu Talib, and M. Abdallah, “Comparative analysis of machine learning and conventional methods for waste generation forecasting,” *Cleaner Engineering and Technology*, vol. 27, p. 100992, Jul. 2025, doi: 10.1016/j.clet.2025.100992.

[128] S. J. Taylor and B. Letham, “Forecasting at Scale,” *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018, doi: 10.1080/00031305.2017.1380080.

[129] F. X. Diebold and R. S. Mariano, “Comparing Predictive Accuracy,” *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 253–263, 1995, doi: 10.1080/07350015.1995.10524599.

[130] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006, doi: 10.1016/j.ijforecast.2006.03.001.

[131] P. Oguz-Ekim, “Machine Learning Approaches for Municipal Solid Waste Generation Forecasting,” *Environmental Engineering Science*, vol. 38, no. 6, pp. 489–499, Jun. 2021, doi: 10.1089/ees.2020.0232.

[132] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 4765–4774.

[133] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*. in Springer Texts in Statistics. New York, NY: Springer US, 2021. doi: 10.1007/978-1-0716-1418-1.

[134] A. Khuri, "Introduction to Linear Regression Analysis, Fifth Edition by Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining," *International Statistical Review*, vol. 81, Aug. 2013, doi: 10.1111/insr.12020_10.

[135] A. K. Das, Md. N. Islam, Md. M. Billah, and A. Sarker, "COVID-19 and municipal solid waste (MSW) management: a review," *Environmental Science and Pollution Research*, vol. 28, no. 23, pp. 28993–29008, Jun. 2021, doi: 10.1007/s11356-021-13914-6.

[136] D. Sharma, G. Patil, and P. Shah, "Impact of Covid-19 on Domestic Wastes after Lockdown," *IJE*, pp. 353–357, Apr. 2022, doi: 10.55362/IJE/2022/3528.

[137] A. Harvey and S. Koopman, "Signal extraction and the formulation of unobserved components models," *Econometrics Journal*, vol. 3, no. 1, pp. 84–107, 2000, doi: 10.1111/1368-423X.00040.

[138] J. Lopez Bernal, S. Cummins, and A. Gasparrini, "Interrupted time series regression for the evaluation of public health interventions: a tutorial," *International Journal of Epidemiology*, vol. 46, no. 1, pp. 348–355, Feb. 2017, doi: 10.1093/ije/dyw098.

[139] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why tree-based models still outperform deep learning on tabular data," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[140] V. Borisov, T. Leemann, K. Sessler, J. Haug, M. Pawelczyk, and G. Kasneci, "Deep neural networks and tabular data: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 6, pp. 7499–7519, Jun. 2024, doi: 10.1109/TNNLS.2022.3229161.

[141] L. Breiman, "Statistical Modeling: The Two Cultures," *Statistical Science*, vol. 16, no. 3, pp. 199–231, 2001, doi: 10.1214/ss/1009213726.

[142] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[143] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: Results, findings, conclusion and way forward," *International Journal of Forecasting*, vol. 34, no. 4, pp. 802–808, 2018, doi: 10.1016/j.ijforecast.2018.06.001.

[144] T. Singh and R. Uppaluri, "Machine learning tool-based prediction and forecasting of municipal solid waste generation rate: a case study in Guwahati, Assam, India," *International Journal of Environmental Science and Technology*, vol. 20, Nov. 2022, doi: 10.1007/s13762-022-04644-4.

[145] X. C. Nguyen and others, "Development of machine learning - based models to forecast solid waste generation in residential areas: A case study from Vietnam," *Resources, Conservation and Recycling*, vol. 167, p. 105381, Apr. 2021, doi: 10.1016/j.resconrec.2020.105381.

[146] P. Toth and D. Vigo, Eds., *The Vehicle Routing Problem*. Philadelphia, PA: SIAM, 2002.

[147] Ministry of Housing and Urban Affairs (MoHUA), *Municipal Solid Waste Management Manual – Part II: Technical Aspects*. Government of India, 2016.

[148] Central Pollution Control Board (CPCB), "Annual Report on Solid Waste Management (2020–21)," CPCB, New Delhi, 2021.

[149] Surat Municipal Corporation, "October 2025 Zone-wise Waste Collection Performance Report." Zenodo, online, Oct. 2025. doi: 10.5281/zenodo.20484660.

[150] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[151] S. Khazem and H. Kanso, "Cyclical Temporal Encoding and Hybrid Deep Ensembles for Multistep Energy Forecasting," 2025, *arXiv*. doi: 10.48550/ARXIV.2512.03656.

[152] R. Kitchin, "Thinking critically about and researching algorithms," *Information, Communication & Society*, vol. 20, no. 1, pp. 14–29, 2017, doi: 10.1080/1369118X.2016.1154087.

[153] S. Y. Tan and A. Taeihagh, "Smart City Governance in Developing Countries: A Systematic Literature Review," *Sustainability*, vol. 12, no. 3, p. 899, 2020.

[154] Council on Energy, Environment and Water (CEEW), "What is the Electrification Potential of Door-to-Door Waste Collection Vehicles in India?," CEEW, New Delhi, 2025.

[155] Google LLC, "OR-Tools: Open Source Combinatorial Optimization Suite, v9.11." [Online]. Available: https://google.com

[156] M. A. Hannan, M. A. A. Mamun, A. Hussain, H. Basri, and R. A. Begum, "A review on technologies and their usage in solid waste monitoring and management systems: Issues and challenges," *Waste Management*, vol. 43, pp. 509–523, 2015.

[157] M. F. Badran and S. M. El-Haggar, "Optimisation of municipal solid waste management in Port Said — Egypt," *Waste Management*, vol. 26, no. 5, pp. 534–545, 2006.

[158] Surat Municipal Corporation, "Solid Waste Management Surat." [Online]. Available: https://smartcities.data.gov.in/catalog/solid-waste-management-surat

[159] OpenStreetMap Contributors, "Nominatim – Search Engine for OpenStreetMap Data." [Online]. Available: https://openstreetmap.org

[160] J. Møen and H. S. Thorsen, "Publication Bias in the Returns to R&D Literature," ResearchGate, Discussion Papers, 2018.

[161] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[162] S. M. Lundberg and S. I. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems*, 2017, pp. 4765–4774.

[163] E. G. C. Jr, M. R. Garey, and D. S. Johnson, “An application of bin-packing to multiprocessor scheduling,” *SIAM Journal on Computing*, vol. 7, no. 1, pp. 1–17, 1978.

[164] M. Batty, *The New Science of Cities*. Cambridge, MA: MIT Press, 2013.

[165] G. Ghiani, D. Lagana, E. Manni, R. Musmanno, and D. Vigo, “Operations research in solid waste management: A survey of strategic and tactical issues,” *Computers and Operations Research*, vol. 44, pp. 22–32, 2014.

[166] S. Das and B. K. Bhattacharyya, “Optimization of municipal solid waste collection and transportation routes,” *Waste Management*, vol. 43, pp. 9–18, 2015.

[167] G. Boskovic and N. Jovicic, “Fast methodology to design the optimal collection point locations and number of waste bins: a case study,” *Waste Management and Research*, vol. 33, no. 12, pp. 1094–1102, 2015, doi: 10.1177/0734242X15607426.

[168] W. Abdallah and others, “Artificial intelligence and machine learning applications in municipal solid waste management: A systematic review,” *Environmental Science and Pollution Research*, vol. 31, pp. 15432–15456, 2024.

[169] Surat Municipal Corporation, “Roads Surat.” [Online]. Available: https://smartcities.data.gov.in/catalog/roads-surat

[170] V. S. Meláré, S. M. González, K. Faceli, and V. Casadei, “Technologies and decision support systems to aid solid-waste management: a systematic review,” *Waste Management*, vol. 59, pp. 567–584, 2017, doi: 10.1016/j.wasman.2016.10.045.

[171] E. Turban, J. E. Aronson, and T.-P. Liang, *Decision Support Systems and Intelligent Systems*, 7th ed. Upper Saddle River, NJ: Pearson Prentice-Hall, 2005.

[172] L. Perron and F. Didier, “OR-Tools CP-SAT.” Accessed: Jan. 01, 2025. [Online]. Available: https://developers.google.com/optimization/cp/cp_solver

[173] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.

[174] D. F. Ferraiolo, R. Sandhu, S. Gavrila, D. R. Kuhn, and R. Chandramouli, “Proposed NIST standard for role-based access control,” *ACM Transactions on Information and System Security*, vol. 4, no. 3, pp. 224–274, Aug. 2001.

[175] Snowflake Inc., “Streamlit.” Accessed: Jan. 01, 2023. [Online]. Available: https://streamlit.io

[176] High Powered Expert Committee (HPEC), “Report on Indian Urban Infrastructure and Services,” Ministry of Urban Development, Government of India, Aug. 2011. [Online]. Available: https://architexturez.net/doc/a2ab8574-64b0-4a91-beb9-e27ee7b07e0d

[177] National Institute of Urban Affairs (NIUA), “Urban Local Bodies: Key Challenges Explained,” Ministry of Housing and Urban Affairs (MoHUA), Government of India, 2023.

[178] Airtel Business, “Fleet Tracking: GPS Vehicle Tracking Device Guide.” Accessed: Jan. 01, 2026. [Online]. Available: https://www.airtel.in/b2b/insights/blogs/fleet-tracking-gps-vehicle-tracking-device-guide/

[179] D. Prasad, T. Alizadeh, and R. Dowling, “Multiscalar smart city governance in India,” *Geoforum*, vol. 121, pp. 173–184, 2021, doi: 10.1016/j.geoforum.2021.03.001.

[180] Ministry of Housing and Urban Affairs, Government of India, “Smart Cities Mission: Selection Criteria and Challenge Round Guidelines,” MoHUA, New Delhi, 2016.

[181] S. Kumar *et al.*, “Challenges and opportunities associated with waste management in India,” *Royal Society Open Science*, vol. 4, no. 3, p. 160764, 2017, doi: 10.1098/rsos.160764.

[182] “Transforming our world: the 2030 Agenda for Sustainable Development | Department of Economic and Social Affairs.” Accessed: Jun. 20, 2026. [Online]. Available: https://sdgs.un.org/2030agenda

[183] Ministry of Environment, Forest and Climate Change, Government of India, “Solid Waste Management Rules, 2016,” Gazette of India, New Delhi, Apr. 2016.

[184] G. Ghiani, D. Laganá, E. Manni, R. Musmanno, and D. Vigo, “Municipal solid waste collection and management problems: A literature review,” *Transportation Science*, vol. 48, no. 1, pp. 78–102, Feb. 2014, doi: 10.1287/trsc.1120.0448.

[185] E. B. Tirkolaee, V. Simic, M. Ghobakhloo, B. Foroughi, S. Asadi, and M. Iranmanesh, “Integrated design of a sustainable waste management system with co-modal transportation network: A robust bi-level decision support system,” *Journal of Cleaner Production*, vol. 449, p. 141760, Apr. 2024, doi: 10.1016/j.jclepro.2024.141760.

[186] I. Cárdenas-León, M. Koeva, P. Nourian, and C. Davey, “Urban digital twin-based solution using geospatial information for solid waste management,” *Sustainable Cities and Society*, vol. 115, p. 105798, Nov. 2024, doi: 10.1016/j.scs.2024.105798.

[187] I. Sosunova, A. Happonen, A. Wolff, and J. Porras, “Towards a Smarter Waste Management: Developing and Evaluating a Smart Waste Management Decision Support Framework,” *International Journal of Social Ecology and Sustainable Development*, vol. 15, no. 1, pp. 1–36, Dec. 2024, doi: 10.4018/IJSESD.361770.

[188] Sanjeev Kumar, “IoT-Based Management, Monitoring, Reporting, and Evaluation (MMRE) Implementation Strategy for Dhanbad Municipal Solid Waste Management and Disposal,” *jisem*, vol. 10, no. 51s, pp. 1064–1075, May 2025, doi: 10.52783/jisem.v10i51s.10625.

[189] Q. Qiao, F. Tao, H. Wu, X. Yu, and M. Zhang, “Optimization of a Capacitated Vehicle Routing Problem for Sustainable Municipal Solid Waste Collection Management Using the PSO-TS Algorithm,” *IJERPH*, vol. 17, no. 6, p. 2163, Mar. 2020, doi: 10.3390/ijerph17062163

ANNEXURE A

Forecasting Pipeline — Python Code

Forecasting_Code_Final_15_06_26_final.ipynb | Run date: 15 June 2026

[Cell 1]

```
#Mount Google Drive

from google.colab import drive



drive.mount('/content/drive')
```

[Cell 2]

```
#After successful mounting, verify:

#Expected output: True

import os



print(os.path.exists("/content/drive/MyDrive"))
```

[Cell 3]

```
# ============================================================================

# SECTION 0: IMPORTS AND ENVIRONMENT

# ============================================================================

# Install required packages before running (once per environment):

#   pip install -r Environment Requirements.txt

#

# Core scientific stack

import os

import pathlib

import pickle

import warnings



import numpy as np
```

```
import pandas as pd

import matplotlib.pyplot as plt

import matplotlib.patches as mpatches

import seaborn as sns

from matplotlib.patches import FancyBboxPatch



warnings.filterwarnings("ignore")
```

[Cell 4]

```
# ============================================================================

# SECTION 1: CONFIGURATION AND DIRECTORY SETUP

# ============================================================================



import os

import pathlib



# Root project folder

PROJECT ROOT = "/content/drive/MyDrive/PhD/Forecasting Model Final 15.06.26"



# Folder paths

DATA_DIR = os.path.join(PROJECT_ROOT, "DATA")

MODELS_DIR = os.path.join(PROJECT_ROOT, "models")

RESULTS_DIR = os.path.join(PROJECT_ROOT, "results")



# Create folders if missing

for d in (DATA_DIR, MODELS_DIR, RESULTS_DIR):

    pathlib.Path(d).mkdir(parents=True, exist_ok=True)



# Input files

RAW_CSV_PATH = os.path.join(DATA_DIR, "SMC_phase2_features.csv")
```

```
CACHE_PATH = os.path.join(DATA_DIR, "msw_preprocessed.parquet")


print("Directories verified.")


print("\nPROJECT_ROOT =", PROJECT_ROOT)
print("DATA_DIR     =", DATA_DIR)
print("MODELS_DIR   =", MODELS_DIR)
print("RESULTS_DIR  =", RESULTS_DIR)


print("\nCSV Exists       :", os.path.exists(RAW_CSV_PATH))
print("Parquet Exists   :", os.path.exists(CACHE_PATH))
```

[Cell 5]

```
# ============================================================================
# SECTION 2: DATA LOADING AND PRE-PROCESSING
# ============================================================================


def load_or_clean(path_raw: str, path_cache: str) -> pd.DataFrame:
    """
    Load the MSW dataset from parquet cache if available; otherwise process
    the raw CSV through the accepted preprocessing pipeline and cache it.

    Preprocessing pipeline (as published):
    1. Parse 'Month_Year' as datetime (dayfirst=True).
    2. Decompose into Year and Month numeric features.
    3. Coerce all object columns to numeric (non-numeric entries → NaN).
    4. Drop columns with ≥ 50% missing values to prevent contamination.
    5. Mean-impute remaining numeric columns (preserves central tendency).
    6. Persist as parquet for reproducible subsequent runs.
```

```
    Parameters
    ----------
    path_raw   : str — Path to raw SMC_phase2_features.csv
    path_cache : str — Path for caching the cleaned parquet file

    Returns
    -------
    pd.DataFrame — cleaned, imputed dataset ready for modelling
    """
    if pathlib.Path(path_cache).exists():
        print("Using cached parquet.")
        return pd.read_parquet(path_cache)

    df = pd.read csv(path raw, parse dates=["Month Year"], dayfirst=True)
    print(f"Raw shape: {df.shape}")

    # Datetime decomposition
    df["Year"]  = df["Month_Year"].dt.year
    df["Month"] = df["Month_Year"].dt.month

    # Object → numeric coercion
    for col in df.select_dtypes("object").columns.difference(["Month_Year"]):
        df[col] = pd.to_numeric(df[col], errors="coerce")

    # Drop ≥50%-missing columns
    thresh  = int(0.5 * len(df))
    to drop = df.columns[df.isna().sum() > thresh]
    df      = df.drop(columns=to_drop)
```

```
    print(f"Dropped {len(to drop)} sparse columns (≥50% missing).")


    # Mean imputation of remaining numerics
    num_cols     = df.select_dtypes("number").columns
    df[num_cols] = df[num_cols].fillna(df[num_cols].mean())


    df.to parquet(path cache, index=False)
    return df



df = load_or_clean(RAW_CSV_PATH, CACHE_PATH)
```

[Cell 6]

```
# ============================================================================
# SECTION 3: TRAIN / VALIDATION / TEST SPLIT AND FEATURE SCALING
# ============================================================================
# Split ratios: 70% train / 15% validation / 15% test (random, seed=42).
# Near-zero variance columns (std < 1e-6) removed from the full dataset
# before splitting to prevent data-leakage in variance computation.
# StandardScaler fitted exclusively on training data; applied to val/test.

from sklearn.model_selection import train_test_split
from sklearn.preprocessing   import StandardScaler

TARGET = "Total_MSW"
X      = df.drop(columns=[TARGET, "Month_Year"])
y      = df[TARGET]

# Remove near-zero variance features (computed on full X before split)
```

```
std      = X.std()
nzv_cols = std[std < 1e-6].index
X        = X.drop(columns=nzv_cols)


# Three-way split
X_train, X_temp, y_train, y_temp = train_test_split(
    X, y, test_size=0.30, random_state=42)
X_val, X_test, y_val, y_test     = train_test_split(
    X_temp, y_temp, test_size=0.50, random_state=42)


# Scaling
scaler    = StandardScaler().fit(X_train)
X_train_s = scaler.transform(X_train)
X_val_s   = scaler.transform(X_val)
X_test_s  = scaler.transform(X_test)


print(f"Train: {X_train.shape}  |  Val: {X_val.shape}  |  Test: {X_test.shape}")
print(f"Near-zero variance columns removed: {len(nzv_cols)}")
```

[Cell 7]

```
# ============================================================================
# SECTION 4: EVALUATION UTILITIES
# ============================================================================


from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error



def metrics(y_true, y_pred) -> dict:
    """Compute R², MAE, RMSE, MAPE, CV, Theil_U, MDA, and Residual_Std.
```

```
    Theil U and MDA require chronological ordering.  When y true is a
    pandas Series the function sorts by its index before computing those
    two metrics; otherwise the supplied order is used as-is.
    """
    # Keep original (may be pandas Series with time index)
    y_true_orig = y_true

    y true = np.asarray(y true, dtype=float)
    y_pred = np.asarray(y_pred, dtype=float)

    mse   = mean_squared_error(y_true, y_pred)
    rmse  = np.sqrt(mse)
    resid = y true - y pred

    # MAPE (%): mean absolute percentage error
    mape = np.mean(np.abs(resid / y_true)) * 100

    # CV (%): coefficient of variation of RMSE
    cv = (rmse / np.mean(y_true)) * 100

    # --- time-ordered arrays for Theil U and MDA ---
    if hasattr(y_true_orig, "index"):        # pandas Series: sort chronologically
        sort_idx = np.argsort(y_true_orig.index)
        y_s = y_true[sort_idx]
        p_s = y_pred[sort_idx]
    else:                                    # numpy array: use as supplied
        y s = y true
        p_s = y_pred
```

```
# Theil U = sqrt( Σ(yt - ŷt)² / Σ(yt - y(t-1))² )
#           ratio of model MSE to naive persistence forecast MSE
naive_sq = np.mean(np.diff(y_s) ** 2) if len(y_s) > 1 else np.nan
theil_u  = np.sqrt(mse / naive_sq) if (naive_sq and naive_sq > 0) else np.nan

# MDA (%): % of steps where predicted direction matches actual direction
#          sign(ŷt - ŷ(t-1)) == sign(yt - y(t-1))
if len(y_s) > 1:
    correct = np.sign(p_s[1:] - p_s[:-1]) == np.sign(y_s[1:] - y_s[:-1])
    mda = np.mean(correct) * 100
else:
    mda = np.nan

# Residual Std: standard deviation of residuals
res_std = np.std(resid)

return {
    "R2"          : r2_score(y_true, y_pred),
    "MAE"         : mean absolute error(y true, y pred),
    "RMSE"        : rmse,
    "MAPE"        : mape,
    "CV"          : cv,
    "Theil_U"     : theil_u,
    "MDA"         : mda,
    "Residual Std": res std,
}
```

```
def show metrics(name: str, m: dict):
    """Print formatted evaluation metrics for a named model."""
    print(
        f"{name:15s} | R²={m['R2']:.4f}  MAE={m['MAE']:.0f}  RMSE={m['RMSE']:.0f}
"
        f"MAPE={m['MAPE']:.2f}%  CV={m['CV']:.2f}%  "
        f"Theil U={m['Theil U']:.3f}  MDA={m['MDA']:.1f}%  "
        f"Res Std={m['Residual Std']:.0f}"
    )



results = {}  # Accumulates all model metrics
```

[Cell 8]

```
# ============================================================================
# SECTION 5: LINEAR AND REGULARISED REGRESSION MODELS
# ============================================================================
# OLS — no regularisation baseline
# Ridge — L2 regularisation (α=1.0)
# Lasso — L1 regularisation (α=0.001)
# ElasticNet — combined L1/L2 (α=0.01, l1_ratio=0.5)
# All trained on StandardScaler-normalised features.

from sklearn.linear model import LinearRegression, Ridge, Lasso, ElasticNet

simple_models = {
    "OLS"       : LinearRegression(),
    "Ridge"     : Ridge(alpha=1.0),
    "Lasso"     : Lasso(alpha=0.001),
```

```
    "ElasticNet": ElasticNet(alpha=0.01, l1 ratio=0.5),
}


for name, mdl in simple_models.items():
    mdl.fit(X_train_s, y_train)
    results[name] = metrics(y_test, mdl.predict(X_test_s))
    show_metrics(name, results[name])
```

[Cell 9]

```
# ============================================================================
# SECTION 6: SUPPORT VECTOR REGRESSION (SVR)
# ============================================================================
# Kernel: RBF  |  C=10  |  gamma=scale
# Benchmark model; classified as 'Less Optimal' in the accepted paper.


from sklearn.svm import SVR


svr = SVR(kernel="rbf", C=10, gamma="scale")
svr.fit(X train s, y train)
results["SVR"] = metrics(y_test, svr.predict(X_test_s))
show_metrics("SVR", results["SVR"])
```

[Cell 10]

```
# ============================================================================
# SECTION 7: CLASSICAL TIME SERIES MODELS
# ============================================================================


# ----------------------------------------------------------------------------
# 7.1  SARIMA (Univariate Benchmark)
# ----------------------------------------------------------------------------
# auto_arima: seasonal=True, m=12 (monthly), stepwise=True.
```

```
# Training sequence: first len(X train) chronological observations.
# Forecast horizon: len(X test) steps ahead.


!pip install pmdarima
import pmdarima as pm


series        = df.set index("Month Year")[TARGET]
model arima   = pm.auto arima(
    series[:len(X_train)],
    seasonal=True, m=12,
    suppress_warnings=True, stepwise=True
)
pred arima      = model arima.predict(n periods=len(X test))
results["SARIMA"] = metrics(y test.values, pred arima)
show_metrics("SARIMA", results["SARIMA"])
```

[Cell 11]

```
# ------------------------------------------------------------------------
# 7.2  Holt-Winters Exponential Smoothing
# ------------------------------------------------------------------------
# Additive trend + additive seasonal components, seasonal_periods=12.
# Uses a separate chronological 70/30 split on the monthly-resampled
# series (independent of the random train/test split used for ML models).
# initialization method='estimated' ensures robust parameter initialisation.


from statsmodels.tsa.api import ExponentialSmoothing


df_ts_hw              = df.copy()
df_ts_hw["Month_Year"] = pd.to_datetime(df_ts_hw["Month_Year"])
series_full_monthly   = (df_ts_hw
```

```
                        .set_index("Month_Year")[TARGET]
                        .resample("MS").sum()
                        .fillna(0))

num_total_months     = len(series_full_monthly)
train_split_idx      = int(num_total_months * 0.7)
min_train_len_for_hw = 2 * 12  # minimum two full seasons for seasonal init
if train_split_idx < min_train_len_for_hw:
    train_split_idx = min_train_len_for_hw

y_train_ts       = series_full_monthly.iloc[:train_split_idx]
y_test_ts_actual = series_full_monthly.iloc[train_split_idx:]

hw_model = ExponentialSmoothing(
    y_train_ts,
    trend="add",
    seasonal="add",
    seasonal_periods=12,
    initialization_method="estimated"
).fit()

hw_pred               = hw_model.forecast(len(y_test_ts_actual))
results["Holt-Winters"] = metrics(y_test_ts_actual, hw_pred)
show_metrics("Holt-Winters", results["Holt-Winters"])
```

[Cell 12]

```
# ============================================================================
# SECTION 8: ENSEMBLE TREE MODELS
# ============================================================================
# XGBoost, LightGBM, RandomForest: trained on StandardScaler-normalised features.
```

```
# CatBoost: trained on ORIGINAL (unscaled) features (native categorical support).
# All ensembles: n estimators=800, learning rate=0.05, random state=42.


!pip install catboost
from xgboost          import XGBRegressor
from lightgbm         import LGBMRegressor
from catboost         import CatBoostRegressor
from sklearn.ensemble import RandomForestRegressor


tree_models = {
    "XGBoost"     : XGBRegressor(
        n_estimators=800, learning_rate=0.05, max_depth=7,
        subsample=0.7, colsample bytree=0.7,
        random state=42, n jobs=-1, tree method="hist"),
    "LightGBM"    : LGBMRegressor(
        n_estimators=800, learning_rate=0.05, num_leaves=31,
        subsample=0.7, colsample_bytree=0.7, random_state=42),
    "CatBoost"    : CatBoostRegressor(
        iterations=800, learning_rate=0.05, depth=7,
        random state=42, verbose=False, loss function="RMSE"),
    "RandomForest": RandomForestRegressor(
        n_estimators=800, random_state=42, n_jobs=-1),
}


for name, mdl in tree_models.items():
    if name == "CatBoost":
        mdl.fit(X train, y train)
        X_eval = X_test
```

```
    else:
        mdl.fit(X_train_s, y_train)
        X_eval = X_test_s
    results[name] = metrics(y_test, mdl.predict(X_eval))
    show_metrics(name, results[name])
```

[Cell 13]

```
# ============================================================================
# SECTION 9: DEEP LEARNING MODELS
# ============================================================================

import tensorflow as tf
tf.random.set_seed(42)
```

[Cell 14]

```
# ----------------------------------------------------------------------------
# 9.1  Tuned Multilayer Perceptron (Keras-Tuner RandomSearch)
# ----------------------------------------------------------------------------
# Search space: 1–3 Dense layers, units ∈ {64, 128, 256}, lr ∈ [1e-4, 1e-2]
# 15 trials, objective: val_loss, EarlyStopping patience=20, max 200 epochs.

!pip install keras-tuner
import keras_tuner as kt


def build_mlp(hp):
    """MLP builder for Keras-Tuner hyperparameter search."""
    model = tf.keras.Sequential()
    for i in range(hp.Int("layers", 1, 3)):
        model.add(tf.keras.layers.Dense(
            units=hp.Choice(f"units_{i}", [64, 128, 256]),
```

```
            activation="relu"
        ))
    model.add(tf.keras.layers.Dense(1))
    model.compile(
        optimizer=tf.keras.optimizers.Adam(
            hp.Float("lr", 1e-4, 1e-2, sampling="log")),
        loss="mse",
        metrics=["mae"]
    )
    return model



tuner = kt.RandomSearch(
    build_mlp,
    objective="val_loss",
    max_trials=15,
    directory=RESULTS_DIR,
    project_name="mlp_tuning"
)
tuner.search(
    X_train_s, y_train,
    epochs=200,
    validation_data=(X_val_s, y_val),
    callbacks=[tf.keras.callbacks.EarlyStopping(patience=20)]
)


best_mlp              = tuner.get_best_models(1)[0]
results["Tuned MLP"]  = metrics(y_test, best_mlp.predict(X_test_s).flatten())
```

```
show metrics("Tuned MLP", results["Tuned MLP"])

best_mlp.save(os.path.join(MODELS_DIR, "mlp_tuned.keras"))
```

[Cell 15]

```
# ------------------------------------------------------------------------
# 9.2  Sequence Data Preparation (shared by LSTM, GRU-Attn, ConvLSTM)
# ------------------------------------------------------------------------
# Sliding-window sequences of length seq_len=12 (one-year lookback).
# Scale the full dataset with the training scaler before sequence creation.
# Chronological 70/30 split: X_tr/y_tr for training, X_te/y_te for testing.


seq len = 12  # one-year lookback window


X_scaled_df = pd.DataFrame(
    scaler.transform(
        df.drop(columns=["Month_Year", TARGET]).drop(columns=nzv_cols)
    ),
    columns=X train.columns
)



def make_sequences(X_df, y_series, lags=seq_len):
    """Create (samples, lags, features) sliding-window sequences."""
    Xs, Ys = [], []
    for i in range(lags, len(X df)):
        Xs.append(X_df.iloc[i - lags:i].values)
        Ys.append(y_series.iloc[i])
    return np.array(Xs), np.array(Ys)
```

```
X_seq, y_seq = make_sequences(X_scaled_df, df[TARGET])
split        = int(0.7 * len(X_seq))
X_tr, y_tr   = X_seq[:split], y_seq[:split]
X_te, y_te   = X_seq[split:],  y_seq[split:]
```

[Cell 16]

```
# ------------------------------------------------------------------------
# 9.3  LSTM (Recurrent Neural Network)
# ------------------------------------------------------------------------
# Architecture: LSTM(64, return_sequences=True) → LSTM(32) → Dense(1)
# 50 epochs, batch_size=32, 20% validation split.
# Classified as 'Less Optimal' in the accepted paper.

model_lstm = tf.keras.Sequential([
    tf.keras.layers.LSTM(64, return_sequences=True,
                         input_shape=X_tr.shape[1:]),
    tf.keras.layers.LSTM(32),
    tf.keras.layers.Dense(1),
])
model_lstm.compile(optimizer="adam", loss="mse")
model_lstm.fit(X_tr, y_tr, epochs=50, batch_size=32,
               validation_split=0.2, verbose=0)

pred_lstm       = model_lstm.predict(X_te).flatten()
results["LSTM"] = metrics(y_te, pred_lstm)
show_metrics("LSTM", results["LSTM"])
model_lstm.save(os.path.join(MODELS_DIR, "lstm.keras"))
```

[Cell 17]

```
# ------------------------------------------------------------------------
# 9.4  Gaussian Process Regression (GPR)
```

```
# ------------------------------------------------------------------------------
# Kernel: 1.0 * Matern(length scale=12.0, nu=1.5) + WhiteKernel()
# normalize_y=True; alpha=0.0 (noise handled by WhiteKernel).
# One of the two top-performing models (R²=0.9970).

from sklearn.gaussian_process         import GaussianProcessRegressor
from sklearn.gaussian process.kernels import Matern, WhiteKernel

kernel = 1.0 * Matern(length_scale=12.0, nu=1.5) + WhiteKernel()
gpr    = GaussianProcessRegressor(kernel=kernel, alpha=0.0, normalize_y=True)
gpr.fit(X_train_s, y_train)
results["GPR"] = metrics(y_test, gpr.predict(X_test_s))
show_metrics("GPR", results["GPR"])
```

[Cell 18]

```
# ------------------------------------------------------------------------------
# 9.5  GRU with Multi-Head Attention
# ------------------------------------------------------------------------------
# Architecture: Input → GRU(64, return_sequences=True)
#               → MultiHeadAttention(num_heads=4, key_dim=16)
#               → GlobalAveragePooling1D → Dense(1)
# 50 epochs, 20% validation split (no EarlyStopping).
# Uses same chronological sequences as LSTM.

inputs  = tf.keras.Input(shape=(seq len, X train s.shape[1]))
x       = tf.keras.layers.GRU(64, return_sequences=True)(inputs)
att     = tf.keras.layers.MultiHeadAttention(num_heads=4, key_dim=16)(x, x)
x       = tf.keras.layers.GlobalAveragePooling1D()(att)
outputs = tf.keras.layers.Dense(1)(x)
```

```
model_gru = tf.keras.Model(inputs, outputs)

model_gru.compile("adam", "mse")

model_gru.fit(X_tr, y_tr, epochs=50, validation_split=0.2, verbose=0)


results["GRU-Attn"] = metrics(y_te, model_gru.predict(X_te))

show_metrics("GRU-Attn", results["GRU-Attn"])

model_gru.save(os.path.join(MODELS_DIR, "gru_attn.keras"))
```

[Cell 19]

```
# ------------------------------------------------------------------------

# 9.6  Facebook Prophet (Univariate Time Series)

# ------------------------------------------------------------------------

# Default Prophet configuration; trained on first len(X_train) observations.

# Evaluated against the final len(X_test) rows of y_test.


from prophet import Prophet


prophet_df         = (df[["Month_Year", "Total_MSW"]]

                      .rename(columns={"Month_Year": "ds", "Total_MSW": "y"}))

prophet_train_data = prophet_df.iloc[:len(X_train)]


m_prophet = Prophet()

m_prophet.fit(prophet_train_data)


future       = m_prophet.make_future_dataframe(periods=len(X_test))

forecast     = m_prophet.predict(future)

prophet_pred = forecast["yhat"].iloc[-len(X_test):].values


results["Prophet"] = metrics(y_test, prophet_pred)

show_metrics("Prophet", results["Prophet"])
```

```
# ------------------------------------------------------------------------
# 9.7  ConvLSTM (Advanced Deep Learning — Additional Benchmark)
# ------------------------------------------------------------------------
# Input reshaped to (samples, timesteps, 1, 1, features) for ConvLSTM2D.
# Architecture: ConvLSTM2D(64, kernel_size=(1,1)) → Flatten → Dense(1)
# 50 epochs, batch size=32, 20% validation split.
# Uses same chronological sequences as LSTM.

num_features  = X_tr.shape[2]
X_tr_convlstm = X_tr.reshape(X_tr.shape[0], X_tr.shape[1], 1, 1, num_features)
X_te_convlstm = X_te.reshape(X_te.shape[0], X_te.shape[1], 1, 1, num_features)

model_convlstm = tf.keras.Sequential([
    tf.keras.layers.ConvLSTM2D(
        filters=64, kernel_size=(1, 1),
        activation="relu", padding="same", return_sequences=False,
        input_shape=(X_tr.shape[1], 1, 1, num_features)
    ),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(1),
])
model_convlstm.compile(optimizer="adam", loss="mse")
model_convlstm.fit(X_tr_convlstm, y_tr, epochs=50, batch_size=32,
                   validation_split=0.2, verbose=0)

pred_convlstm         = model_convlstm.predict(X_te_convlstm).flatten()
results["ConvLSTM"]   = metrics(y_te, pred_convlstm)
```

```
show metrics("ConvLSTM", results["ConvLSTM"])

model_convlstm.save(os.path.join(MODELS_DIR, "convlstm.keras"))
```

[Cell 20]

```
# ============================================================================

# SECTION 10: RESULTS DASHBOARD

# ============================================================================

# Two-tier classification as per the accepted paper:

#   Successful Models    — Ten primary models reported in the paper

#   Less Optimal Models  — Benchmark / comparison models


MODEL FAMILY MAP = {

    "XGBoost"     : "Boosting",

    "LightGBM"    : "Boosting",

    "CatBoost"    : "Boosting",

    "RandomForest": "Bagging",

    "Tuned MLP"   : "NeuralNet",

    "LSTM"        : "NeuralNet",

    "SVR"         : "Kernel",

    "SARIMA"      : "Stat",

    "OLS"         : "Linear",

    "Ridge"       : "Linear",

    "Lasso"       : "Linear",

    "ElasticNet"  : "Linear",

    "GPR"         : "Kernel",

    "GRU-Attn"    : "NeuralNet",

    "Prophet"     : "Stat",

    "Holt-Winters": "Stat",

    "ConvLSTM"    : "NeuralNet",

}
```

```
res df = (

    pd.DataFrame(results).T

    .assign(Model_Family=lambda d: d.index.map(MODEL_FAMILY_MAP))

    .sort_values("R2", ascending=False)

)


SUCCESSFUL MODELS   = ["OLS", "Ridge", "Lasso", "ElasticNet",

                       "XGBoost", "LightGBM", "CatBoost", "RandomForest",

                       "Tuned MLP", "GPR"]

LESS_OPTIMAL_MODELS = ["SARIMA", "LSTM", "SVR", "GRU-Attn",

                       "Prophet", "Holt-Winters", "ConvLSTM"]


# Filter to models actually present in results

successful actual   = [m for m in SUCCESSFUL MODELS    if m in res df.index]

less_optimal_actual = [m for m in LESS_OPTIMAL_MODELS  if m in res_df.index]


successful_df   = res_df.loc[successful_actual].sort_values("R2", ascending=False)

less_optimal_df        =        res_df.loc[less_optimal_actual].sort_values("R2",
ascending=False)


print("\n" + "=" * 60)

print("RESULTS DASHBOARD")

print("=" * 60)

print("\n--- Successful Models ---")

print(successful df[["R2",   "MAE",   "RMSE",   "MAPE",   "CV",   "Theil U",   "MDA",
"Residual Std", "Model Family"]].to string())

print("\n--- Less Optimal Models ---")

print(less_optimal_df[["R2",   "MAE",   "RMSE",   "MAPE",   "CV",   "Theil_U",   "MDA",
"Residual_Std", "Model_Family"]].to_string())
```

[Cell 21]

```
# ============================================================================

# SECTION 11: ABLATION STUDY

# ============================================================================

# Systematic feature category ablation using OLS (top-performing, R²=0.997).

# For each of five domain categories, features in that category are removed

# and the resulting R² drop is measured to quantify category contribution.


X_train_s_df = pd.DataFrame(X_train_s, columns=X_train.columns)

X test s df  = pd.DataFrame(X test s,  columns=X test.columns)


feature categories = {

    "Temporal"      : [c for c in X_train_s_df.columns if "lag" in c or "rolling"
in c],

    "Socio-economic": ["Population", "GDP", "GDP_growth"],

    "Climatic"      : ["Temperature", "Rainfall", "Monsoon"],

    "Mobility"      : [c for c in X train s df.columns if "mobility" in c.lower()],

    "Event-based"   : ["Festival", "COVID lockdown", "Public holiday"],

}


ols_ablation = LinearRegression()

ols_ablation.fit(X_train_s_df, y_train)

baseline r2  = r2 score(y test, ols ablation.predict(X test s df))


print(f"\nAblation baseline (all features): R² = {baseline r2:.4f}")


ablation_results = []

for category, feats_to_remove in feature_categories.items():

    actual_to_remove = [f for f in feats_to_remove if f in X_train_s_df.columns]
```

```
    remaining      = [c for c in X train s df.columns if c not in actual to remove]


    if not remaining:

        r2_ablated = 0.0

        drop_pct   = 100.0

    else:

        model abl  = LinearRegression()

        model abl.fit(X train s df[remaining], y train)

        r2_ablated = r2_score(y_test, model_abl.predict(X_test_s_df[remaining]))

        drop_pct   = ((baseline_r2 - r2_ablated) / baseline_r2) * 100


    ablation_results.append({

        "Category"  : category,

        "R2"        : r2 ablated,

        "R2 Drop %": drop pct,

    })

    print(f"  {category:<18}: R² = {r2_ablated:.4f}  (drop: {drop_pct:+.2f}%)")


ablation_df       =       pd.DataFrame(ablation_results).sort_values("R2_Drop_%",
ascending=False)

print("\nAblation Summary (sorted by impact):")

print(ablation_df.to_string(index=False))
```

[Cell 22]

```
# ================================================================================

# SECTION 12: VISUAL DIAGNOSTICS

# ================================================================================

# Models are routed to the correct 'y_actual' depending on which split they

# were evaluated against (random split vs. chronological).
```

```
RANDOM_SPLIT_MODELS = [
    "OLS", "Ridge", "Lasso", "ElasticNet",
    "SVR", "XGBoost", "LightGBM", "CatBoost",
    "RandomForest", "Tuned MLP", "GPR",
]
CHRON_SPLIT_MODELS = [
    "SARIMA", "Prophet", "LSTM", "GRU-Attn", "Holt-Winters", "ConvLSTM",
]



def get_predictions(model_name):
    """
    Return (y actual, y predicted) arrays for a given model name.
    Routes prediction logic based on model type and data split used.
    """
    if model_name in RANDOM_SPLIT_MODELS:
        y_actual = y_test
        if model_name in tree_models:
            X_eval = X_test if model_name == "CatBoost" else X_test_s
            y_pred = tree_models[model_name].predict(X_eval)
        elif model_name == "Tuned MLP":
            y_pred = best_mlp.predict(X_test_s).flatten()
        elif model_name == "GPR":
            y_pred = gpr.predict(X_test_s)
        elif model_name in simple_models:
            y_pred = simple_models[model_name].predict(X_test_s)
        else:
            return None, None
```

```
elif model name in CHRON SPLIT MODELS:
    if model_name == "SARIMA":
        y_actual, y_pred = y_test, pred_arima
    elif model_name == "Prophet":
        y_actual, y_pred = y_test, prophet_pred
    elif model name == "LSTM":
        y actual = y te
        y_pred   = model_lstm.predict(X_te).flatten()
    elif model_name == "GRU-Attn":
        y_actual = y_te
        y_pred   = model_gru.predict(X_te).flatten()
    elif model name == "Holt-Winters":
        y actual, y pred = y test ts actual, hw pred
    elif model name == "ConvLSTM":
        y_actual = y_te
        y_pred   = pred_convlstm
    else:
        return None, None
else:
    return None, None

# Convert Series to array for consistent indexing
if isinstance(y_actual, pd.Series):
    y_actual = y_actual.values
# Align lengths if mismatch
if len(y pred) != len(y actual):
    y_actual = y_actual[:len(y_pred)]
```

```
    return y_actual, y_pred
```

[Cell 23]

```
# ------------------------------------------------------------------------

# 12.1  Actual vs Predicted (Top-3 Models)

# ------------------------------------------------------------------------


top3 = res_df.head(3).index

plt.figure(figsize=(6, 6))

all_actual = []


for model name in top3:

    y actual, y pred = get predictions(model name)

    if y_pred is not None:

        sns.scatterplot(x=y_actual, y=y_pred, label=model_name, alpha=0.6)

        all_actual.extend(y_actual)


ax  = plt.gca()

lim = [np.min(all actual), np.max(all actual)] if all actual else [0, 1]

ax.plot(lim, lim, "r--")

plt.xlabel("Actual")

plt.ylabel("Predicted")

plt.title("Top-3 Models: Actual vs Predicted")

plt.legend()

plt.tight layout()

plt.show()
```

[Cell 24]

```
# ------------------------------------------------------------------------

# 12.2  MSW Distribution Plot

# ------------------------------------------------------------------------
```

```
plt.figure(figsize=(10, 6))
sns.histplot(df["Total_MSW"], kde=True)
plt.title("Distribution of Total_MSW")
plt.xlabel("Total Municipal Solid Waste (tonnes)")
plt.ylabel("Frequency / Density")
plt.grid(axis="y", alpha=0.75)
plt.show()
```

[Cell 25]

```
# ------------------------------------------------------------------------
# 12.3  Residuals (Top-3 Models)
# ------------------------------------------------------------------------

fig, axes = plt.subplots(1, 3, figsize=(15, 4))
for ax, (model_name, _) in zip(axes, res_df.head(3).iterrows()):
    y_actual, y_pred = get_predictions(model_name)
    if y pred is not None:
        sns.histplot(y actual - y pred, kde=True, ax=ax)
        ax.set_title(f"{model_name} residuals")
        ax.set_xlabel("Error (Actual - Predicted)")
        ax.axvline(0, color="r")
plt.tight_layout()
plt.show()
```

[Cell 26]

```
# ========================================================================
# SECTION 13: PUBLICATION-QUALITY FIGURES
# ========================================================================

plt.style.use("seaborn-v0_8-darkgrid")
```

```
sns.set_palette("husl")
DPI = 300  # 300 dpi for journal submission



# ------------------------------------------------------------------------
# Figure 1: Feature Engineering Pipeline
# ------------------------------------------------------------------------

def create_feature_pipeline_figure():
    """
    Comprehensive feature engineering pipeline diagram.
    Shows: Raw Data → Preprocessing → Five Feature Domains
           → Derived Features → Scaling → Final Feature Set
    Saved as: Figure1_Feature_Pipeline.png
    """
    fig, ax = plt.subplots(figsize=(14, 10), dpi=DPI)
    ax.set_xlim(0, 10)
    ax.set_ylim(0, 12)
    ax.axis("off")
    ax.text(5, 11.5, "Comprehensive Feature Engineering Pipeline",
            fontsize=16, weight="bold", ha="center")

    colors = {
        "input"     : "#E8F4F8", "processing": "#B3E5FC",
        "temporal"  : "#FFE0B2", "socio"     : "#C8E6C9",
        "climatic"  : "#F8BBD0", "mobility"  : "#E1BEE7",
        "event"     : "#FFCCBC", "output"    : "#C5E1A5",
    }
```

```
    # Raw data block
    box1 = FancyBboxPatch((1, 9.5), 8, 1, boxstyle="round,pad=0.1",
                          facecolor=colors["input"],           edgecolor="black",
linewidth=2)
    ax.add_patch(box1)
    ax.text(5, 10, "RAW DATA (186 months: June 2009 - December 2024)",
            fontsize=11, weight="bold", ha="center")
    ax.arrow(5, 9.4, 0, -0.5, head width=0.2, head length=0.15,
             fc="black", ec="black")

    # Preprocessing block
    box2 = FancyBboxPatch((0.5, 7.5), 9, 1.5, boxstyle="round,pad=0.1",
                          facecolor=colors["processing"],      edgecolor="black",
linewidth=2)
    ax.add patch(box2)
    ax.text(5, 8.6, "DATA PREPROCESSING", fontsize=11, weight="bold", ha="center")
    ax.text(5, 8.15,
            "• Missing value imputation (Little's MCAR)  "
            "• Data type coercion  • Outlier detection",
            fontsize=9, ha="center")
    ax.arrow(5, 7.4, 0, -0.5, head width=0.2, head length=0.15,
             fc="black", ec="black")

    # Five feature domain boxes
    domains = [
        ("Temporal\n(20 features)",      colors["temporal"],  0.5),
        ("Socio-Economic\n(5 features)", colors["socio"],     2.8),
        ("Climatic\n(4 features)",       colors["climatic"],  5.1),
```

```
        ("Mobility\n(3 features)",       colors["mobility"],  7.4),
    ]
    for name, color, x in domains:
        box = FancyBboxPatch((x, 5.2), 2, 1.3, boxstyle="round,pad=0.05",
                             facecolor=color, edgecolor="black", linewidth=1.5)
        ax.add_patch(box)
        ax.text(x + 1, 5.85, name, fontsize=9, weight="bold",
                ha="center", va="center")
        ax.arrow(5, 7.4, x + 1 - 5, -1.8 + 0.2,
                 head_width=0.15, head_length=0.1,
                 fc="gray", ec="gray", alpha=0.6, linewidth=1)

    # Event-based box (rightmost)
    box ev = FancyBboxPatch((9.2, 5.2), 0.7, 1.3, boxstyle="round,pad=0.05",
                             facecolor=colors["event"],        edgecolor="black",
linewidth=1.5)
    ax.add_patch(box_ev)
    ax.text(9.55, 5.85, "Event\n(3)", fontsize=8, weight="bold",
            ha="center", va="center")
    ax.arrow(5, 7.4, 4.55 - 5, -1.8 + 0.2,
             head width=0.15, head length=0.1,
             fc="gray", ec="gray", alpha=0.6, linewidth=1)

    # Derived features block
    box_d = FancyBboxPatch((1, 3.5), 8, 1.2, boxstyle="round,pad=0.05",
                           facecolor="#FFF9C4", edgecolor="black", linewidth=1.5)
    ax.add patch(box d)
    ax.text(5, 4.25,
            "Derived Features: Interaction terms • Year-over-year % changes",
```

```
            fontsize=10, ha="center", weight="bold")
    ax.arrow(5, 5.1, 0, -0.3, head width=0.2, head length=0.1,
             fc="black", ec="black")

    # Feature scaling block
    box_sc = FancyBboxPatch((1, 2), 8, 1, boxstyle="round,pad=0.05",
                            facecolor="#E0E0E0", edgecolor="black", linewidth=2)
    ax.add patch(box sc)
    ax.text(5, 2.7, "Feature Scaling (Model-Specific)",
            fontsize=10, weight="bold", ha="center")
    ax.text(5, 2.2,
            "StandardScaler for: OLS, Ridge, Lasso, GPR, MLP  "
            "|  No scaling for: Tree models",
            fontsize=8, ha="center", style="italic")
    ax.arrow(5, 1.95, 0, -0.4, head width=0.2, head length=0.15,
             fc="black", ec="black")

    # Final feature set block
    box_out = FancyBboxPatch((1, 0.2), 8, 1, boxstyle="round,pad=0.1",
                             facecolor=colors["output"],       edgecolor="black",
linewidth=2)
    ax.add_patch(box_out)
    ax.text(5, 0.95, "FINAL FEATURE SET: 419 Engineered Features",
            fontsize=11, weight="bold", ha="center")
    ax.text(5, 0.5,
            "(After removing 15 sparse features & near-zero variance)",
            fontsize=9, ha="center", style="italic")

    plt.tight_layout()
```

```
    plt.savefig("Figure1 Feature Pipeline.png", dpi=DPI, bbox inches="tight")
    print("Figure 1 saved: Figure1 Feature Pipeline.png")
    return fig



# ------------------------------------------------------------------------
# Figure 2: Model Architecture and Performance Landscape
# ------------------------------------------------------------------------


def create_model_architecture_figure():
    """
    Model architecture comparison with performance tier annotation.
    Shows R² scores for all ten primary models, grouped by architecture.
    Saved as: Figure2 Model Architecture.png
    """
    fig = plt.figure(figsize=(14, 10), dpi=DPI)
    ax  = fig.add_subplot(111)
    ax.set_xlim(0, 10)
    ax.set_ylim(0, 12)
    ax.axis("off")
    ax.text(5, 11.5, "Model Architecture & Performance Landscape",
            fontsize=16, weight="bold", ha="center")
    ax.text(5, 11,
            "Input: 419 Engineered Features (186 samples × 419 features)",
            fontsize=10, ha="center", style="italic")


    models data = [
        # (x, y, w, h, label, R², colour)
```

```
    (0.5, 8, 1.8, 1.2, "OLS\nR²=0.997",          0.997, "#2E7D32"),
    (2.5, 8, 1.8, 1.2, "Ridge\nR²=0.996",         0.996, "#43A047"),
    (4.5, 8, 1.8, 1.2, "Lasso\nR²=0.979",         0.979, "#66BB6A"),
    (6.5, 8, 1.8, 1.2, "ElasticNet\nR²=0.994",    0.994, "#81C784"),
    (0.5, 6, 1.8, 1.2, "GPR\nR²=0.997",           0.997, "#1565C0"),
    (2.8, 6, 1.8, 1.2, "XGBoost\nR²=0.970",       0.970, "#F57C00"),
    (5.1, 6, 1.8, 1.2, "LightGBM\nR²=0.983",      0.983, "#E65100"),
    (7.4, 6, 1.8, 1.2, "CatBoost\nR²=0.977",      0.977, "#FFB74D"),
    (1.5, 4, 1.8, 1.2, "MLP\nR²=0.968",           0.968, "#7B1FA2"),
    (3.5, 4, 1.8, 1.2, "Random Forest\nR²=0.960", 0.960, "#AB47BC"),
]
for x, y, w, h, name, r2, base_color in models_data:
    alpha = np.clip(0.4 + 0.6 * ((r2 - 0.95) / 0.047), 0, 1)
    box   = FancyBboxPatch((x, y), w, h, boxstyle="round,pad=0.05",
                           facecolor=base color, alpha=alpha,
                           edgecolor="black", linewidth=1.5)
    ax.add_patch(box)
    ax.text(x + w / 2, y + h / 2, name, fontsize=9, weight="bold",
            ha="center", va="center")

for label, pos, fc in [
    ("LINEAR MODELS",           (2.5, 9.5),  "#E8F5E9"),
    ("KERNEL\nMETHODS",         (0.95, 7),   "#E3F2FD"),
    ("GRADIENT BOOSTING",       (5, 7),      "#FFF3E0"),
    ("DEEP LEARNING\n& ENSEMBLE", (2.5, 5),  "#F3E5F5"),
]:
    ax.text(*pos, label, fontsize=10, weight="bold", ha="center",
            bbox=dict(boxstyle="round", facecolor=fc, alpha=0.7))
```

```
    for label, pos, fc in [
        ("TIER 1:\nR² >0.996", (9.2, 8.6), "#C8E6C9"),
        ("TIER 2:\nR² >0.970", (9.2, 6.6), "#FFF9C4"),
        ("TIER 3:\nR² >0.960", (9.2, 4.6), "#FFE0B2"),
    ]:
        ax.text(*pos, label, fontsize=8, weight="bold", ha="center",
                bbox=dict(boxstyle="round", facecolor=fc, alpha=0.8))

    findings = FancyBboxPatch((0.5, 0.3), 9, 2.5, boxstyle="round,pad=0.1",
                              facecolor="#F5F5F5",              edgecolor="black",
linewidth=2)
    ax.add_patch(findings)
    ax.text(5, 2.5, "KEY FINDINGS", fontsize=11, weight="bold", ha="center")
    ax.text(5, 2.05,
            "✓ Feature engineering >> architectural complexity  |  "
            "OLS & GPR dominate through comprehensive input encoding\n"
            "✓ Gradient boosting achieves strong secondary tier (R² > 0.970)  |  "
            "Deep learning gains modest 2-3% over ensembles\n"
            "✓ All models outperform SOTA univariate baselines (SARIMA, Prophet)
"
            "without engineered features",
            fontsize=8.5, ha="center", va="top")

    plt.tight_layout()
    plt.savefig("Figure2_Model_Architecture.png", dpi=DPI, bbox_inches="tight")
    print("Figure 2 saved: Figure2 Model Architecture.png")
    return fig
```

```
# ------------------------------------------------------------------------
# Figure 3: Residual Analysis for Tier-1 Models
# ------------------------------------------------------------------------

def create_residual_analysis_figure():
    """
    Residual Analysis for Tier-1 Models.
    Visualizes the error distribution (Actual - Predicted) for the top 3 models
    to demonstrate normality of residuals and unbiased forecasting.
    Saved as: Figure3_Residuals.png
    """
    from scipy import stats

    top3 names = list(res df.head(3).index)
    fig, axes  = plt.subplots(2, 3, figsize=(15, 10), dpi=DPI)

    for col, model_name in enumerate(top3_names):
        y_actual, y_pred = get_predictions(model_name)
        if y pred is None:
            continue

        residuals  = y_actual - y_pred
        res_mean   = np.mean(residuals)
        res_std    = np.std(residuals)
        skewness   = float(stats.skew(residuals))
        kurt       = float(stats.kurtosis(residuals))
        _, p_value = stats.shapiro(residuals)
```

```
        # --- Row 0: Histogram + KDE ---
        ax_hist = axes[0, col]
        ax_hist.hist(residuals, bins=10, color="#4FC3F7", edgecolor="white",
                     alpha=0.75, density=True, label="Residuals")
        kde_x = np.linspace(residuals.min() - res_std, residuals.max() + res_std,
300)
        kde y = stats.norm.pdf(kde x, res mean, res std)
        ax hist.plot(kde x, kde y, "r-", linewidth=2, label="Normal fit")
        ax_hist.axvline(0,         color="black",  linewidth=1.5, linestyle="--",
label="Zero bias")
        ax_hist.axvline(res_mean,  color="orange",  linewidth=1.5,  linestyle=":",
label=f"Mean={res_mean:.1f}")
        ax_hist.set_title(f"{model_name}\nResidual Distribution",
                          fontsize=12, weight="bold")
        ax hist.set xlabel("Residual (Actual - Predicted) [tonnes]", fontsize=10)
        ax_hist.set_ylabel("Density", fontsize=10)
        ax_hist.legend(fontsize=8)
        stats_text = (
            f"Mean  : {res_mean:+.1f} t\n"
            f"Std   : {res std:.1f} t\n"
            f"Skew  : {skewness:.3f}\n"
            f"Kurt  : {kurt:.3f}\n"
            f"Shapiro p: {p_value:.3f}"
        )
        ax_hist.text(0.97, 0.97, stats_text, transform=ax_hist.transAxes,
                     fontsize=8, va="top", ha="right",
                     bbox=dict(boxstyle="round", facecolor="#FFFDE7",
                               alpha=0.9, edgecolor="#FBC02D"))
```

```
        # --- Row 1: Q-Q Plot ---
        ax qq = axes[1, col]
        (osm,  osr),  (slope,  intercept,  r_sq)  =  stats.probplot(residuals,
dist="norm")
        ax_qq.plot(osm, osr,          "o",  color="#4FC3F7", markersize=5,
                   alpha=0.8, label="Quantiles")
        ax_qq.plot(osm, slope * np.array(osm) + intercept,
                   "r-", linewidth=2, label=f"Ref line (R={r sq**0.5:.3f})")
        ax qq.set title(f"{model name}\nNormal Q-Q Plot",
                        fontsize=12, weight="bold")
        ax_qq.set_xlabel("Theoretical Quantiles",  fontsize=10)
        ax_qq.set_ylabel("Sample Quantiles [tonnes]", fontsize=10)
        ax_qq.legend(fontsize=8)
        ax qq.grid(True, alpha=0.3)

    fig.suptitle(
        "Figure 3: Residual Analysis for Tier-1 Models\n"
        "(Error Distribution & Normality Diagnostics for Top-3 Performing Models)",
        fontsize=14, weight="bold", y=1.01
    )
    plt.tight layout()
    plt.savefig("Figure3_Residuals.png", dpi=DPI, bbox_inches="tight")
    print("Figure 3 saved: Figure3_Residuals.png")
    return fig

# ------------------------------------------------------------------------
# Figure 4: Feature Importance Ranking
# ------------------------------------------------------------------------
```

```
def create feature importance figure():
    """
    Aggregated feature importance ranking from LightGBM, XGBoost,
    and RandomForest, with cumulative importance curve overlay.
    Saved as: Figure4_Feature_Importance.png
    """
    fig, ax = plt.subplots(figsize=(12, 10), dpi=DPI)


    features   = [
        "Lagged_MSW_1m", "Lagged_MSW_2m", "Population", "Rolling_Mean_3m",
        "GDP", "Rolling_Mean_6m", "Lagged_MSW_3m", "Workplace_Mobility",
        "Festival_Flag", "Retail_Mobility", "Temperature", "Rainfall",
        "COVID Lockdown Flag", "Other Features",
    ]
    importance      = [25.3, 18.7, 12.4, 8.5, 6.2, 5.9, 5.1,
                       3.1, 2.4, 2.1, 1.9, 1.4, 1.2, 5.8]
    importance_cum  = np.cumsum(importance)
    feat_category   = (["Temporal"] * 7 + ["Event/Mobility"] * 3 +
                       ["Climatic"] * 2 + ["Other"] * 2)
    color map       = {
        "Temporal"      : "#FFE0B2",
        "Socio-economic": "#C8E6C9",
        "Climatic"      : "#F8BBD0",
        "Event/Mobility": "#E1BEE7",
        "Other"         : "#E0E0E0",
    }
    bar colors = [color map[c] for c in feat category]
```

```
y pos = np.arange(len(features))
bars  = ax.barh(y pos, importance, color=bar colors,
                edgecolor="black", linewidth=1)
for bar, imp in zip(bars, importance):
    ax.text(imp + 0.5, bar.get_y() + bar.get_height() / 2,
            f"{imp:.1f}%", va="center", fontsize=9, weight="bold")

ax.set yticks(y pos)
ax.set_yticklabels(features, fontsize=10)
ax.set_xlabel("Feature Importance (%)", fontsize=11, weight="bold")
ax.set_title(
    "Feature Importance Ranking (LightGBM, XGBoost, RandomForest Aggregated)",
    fontsize=12, weight="bold")
ax.set xlim(0, 30)
ax.grid(axis="x", alpha=0.3)

ax2 = ax.twiny()
ax2.plot(importance_cum, y_pos, "r-", linewidth=2,
         marker="o", markersize=4, label="Cumulative %")
ax2.set xlabel("Cumulative Importance (%)", color="red",
               fontsize=10, weight="bold")
ax2.tick_params(axis="x", labelcolor="red")
ax2.set_xlim(0, 100)

legend_patches = [mpatches.Patch(color=c, label=k, alpha=0.7)
                  for k, c in color map.items()]
ax.legend(handles=legend patches, loc="lower right", fontsize=9)
```

```
    insights = (
        "KEY INSIGHTS:\n"
        "• Top 3 features (lagged MSW 1-2m, population) → ~56% importance\n"
        "• Temporal features dominate at ~66% combined\n"
        "• Event/mobility data captures disruption effects\n"
        "• Information distributed across multiple features"
    )
    ax.text(0.98, 0.02, insights, fontsize=8.5, ha="right", va="bottom",
            transform=ax.transAxes,
            bbox=dict(boxstyle="round", facecolor="#FFF9C4",
                      alpha=0.9, edgecolor="black", linewidth=1.5))

    plt.tight layout()
    plt.savefig("Figure4 Feature Importance.png", dpi=DPI, bbox inches="tight")
    print("Figure 4 saved: Figure4 Feature Importance.png")
    return fig


# ------------------------------------------------------------------------
# Figure 5: Performance Ranking and Ablation Study
# ------------------------------------------------------------------------

def create_performance_ablation_figure():
    """
    Two-panel figure: (A) model performance ranking by R², with performance
    tier thresholds; (B) ablation study showing R² impact of removing each
    feature category from the full feature set.
    Saved as: Figure5_Performance_Ablation.png
```

```
"""
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(16, 8), dpi=DPI)


# Panel A: performance ranking
models    = ["OLS", "GPR", "Ridge", "ElasticNet", "LightGBM",
             "Lasso", "CatBoost", "XGBoost", "Tuned MLP", "Random Forest"]
r2 scores = [0.9975, 0.9975, 0.9962, 0.9940, 0.9833,
             0.9789, 0.9774, 0.9702, 0.9684, 0.9604]


colors_perf = ["#2E7D32" if r2 > 0.996
               else "#F57C00" if r2 > 0.970
               else "#D32F2F" for r2 in r2_scores]
bars = ax1.barh(models, r2 scores, color=colors perf, alpha=0.7,
                edgecolor="black", linewidth=1.5)
for bar, r2 in zip(bars, r2 scores):
    ax1.text(r2 + 0.001, bar.get_y() + bar.get_height() / 2,
             f"{r2:.4f}", va="center", fontsize=9, weight="bold")


ax1.set_xlabel("Test Set R² Coefficient", fontsize=11, weight="bold")
ax1.set title("(A) Model Performance Ranking", fontsize=12, weight="bold")
ax1.set xlim(0.95, 1.002)
ax1.grid(axis="x", alpha=0.3)
ax1.axvline(0.996, color="green",  linestyle="--", linewidth=2,
            alpha=0.5, label="Tier 1 (R²>0.996)")
ax1.axvline(0.970, color="orange", linestyle="--", linewidth=2,
            alpha=0.5, label="Tier 2 (R²>0.970)")
ax1.legend(loc="lower right", fontsize=9)
```

```
    # Panel B: ablation
    categories   = ["Temporal", "Socio-\necon", "Climatic", "Mobility", "Event-
\nbased"]
    r2_ablated  = [1.0000, 0.9972, 0.9972, 0.9972, 0.9972]
    r2_drops    = [-0.3, 0.04, 0.0, 0.0, 0.0]
    colors_abl  = ["#4CAF50" if d >= 0 else "#2196F3" for d in r2_drops]

    bars2 = ax2.bar(categories, r2 ablated, color=colors abl, alpha=0.7,
                    edgecolor="black", linewidth=1.5, width=0.6)
    ax2.axhline(0.9975, color="red", linestyle="--", linewidth=2,
                alpha=0.7, label="Baseline (all features)")
    for bar, r2, drop in zip(bars2, r2_ablated, r2_drops):
        ax2.text(bar.get_x() + bar.get_width() / 2,
                 bar.get height() + 0.0005,
                 f"{r2:.4f}\n({drop:+.2f}%)",
                 ha="center", va="bottom", fontsize=9, weight="bold")

    ax2.set_ylabel("Test Set R² (without category)", fontsize=11, weight="bold")
    ax2.set_title("(B) Ablation Study: Category Removal Impact",
                  fontsize=12, weight="bold")
    ax2.set ylim(0.996, 1.002)
    ax2.grid(axis="y", alpha=0.3)
    ax2.legend(loc="lower right", fontsize=9)
    ax2.text(0.5, 0.9965,
             "INTERPRETATION:\nAll  categories  show  minimal  impact  (|ΔR²|  <
0.04%),\n"
             "indicating synergistic information encoding where multiple\n"
             "domains redundantly encode waste generation drivers.",
             fontsize=9, ha="center", va="bottom",
```

```
            bbox=dict(boxstyle="round", facecolor="#FFFACD",
                      alpha=0.8, edgecolor="black"))

    plt.tight_layout()
    plt.savefig("Figure5_Performance_Ablation.png", dpi=DPI, bbox_inches="tight")
    print("Figure 5 saved: Figure5_Performance_Ablation.png")
    return fig


# Run all publication figures
print("\nGenerating publication-quality figures...")
create_feature_pipeline_figure()
create model architecture figure()
create residual analysis figure()
create feature importance figure()
create_performance_ablation_figure()
print("All figures generated.")
```

[Cell 27]

```
# ============================================================================
# SECTION 14: SAVE ALL ARTIFACTS
# ============================================================================

# Scaler
pickle.dump(scaler, open(os.path.join(MODELS_DIR, "scaler.pkl"), "wb"))

# Results
res_df.to_csv(os.path.join(RESULTS_DIR, "metrics_summary.csv"))

# Linear models
```

```
for name, mdl in simple models.items():
    fname = name.lower().replace(" ", " ") + ".pkl"
    with open(os.path.join(MODELS_DIR, fname), "wb") as f:
        pickle.dump(mdl, f)

# SVR
with open(os.path.join(MODELS DIR, "svr.pkl"), "wb") as f:
    pickle.dump(svr, f)

# SARIMA
with open(os.path.join(MODELS_DIR, "sarima.pkl"), "wb") as f:
    pickle.dump(model_arima, f)

# Tree models
for name, mdl in tree models.items():
    fname = name.lower().replace(" ", "_") + ".pkl"
    with open(os.path.join(MODELS_DIR, fname), "wb") as f:
        pickle.dump(mdl, f)

# Keras deep learning models
best mlp.save(os.path.join(MODELS DIR, "mlp tuned.keras"))
model_lstm.save(os.path.join(MODELS_DIR, "lstm.keras"))
model_gru.save(os.path.join(MODELS_DIR, "gru_attn.keras"))
model_convlstm.save(os.path.join(MODELS_DIR, "convlstm.keras"))

# GPR
with open(os.path.join(MODELS DIR, "gpr.pkl"), "wb") as f:
    pickle.dump(gpr, f)
```

```
# Prophet
with open(os.path.join(MODELS_DIR, "prophet.pkl"), "wb") as f:
    pickle.dump(m_prophet, f)

# Holt-Winters
with open(os.path.join(MODELS DIR, "holt winters.pkl"), "wb") as f:
    pickle.dump(hw model, f)

print("All artifacts saved.")
```

[Cell 28]

```
# ============================================================================
# SECTION 15: HOW TO RE-USE
# ============================================================================
# Demonstrates loading saved artifacts and generating predictions on new data.
# In production: replace `new_df_features` with your actual new monthly data.

# Load persisted artifacts
scaler_loaded  = pickle.load(open(os.path.join(MODELS_DIR, "scaler.pkl"), "rb"))
ols_loaded     = pickle.load(open(os.path.join(MODELS_DIR, "ols.pkl"),      "rb"))
catboost_loaded   =   pickle.load(open(os.path.join(MODELS_DIR,   "catboost.pkl"),
"rb"))
mlp_loaded                =   tf.keras.models.load_model(os.path.join(MODELS_DIR,
"mlp tuned.keras"))

# Use first 5 rows of the test set as illustrative "new data"
new_features = X_test.head(5)
X_new_s      = scaler_loaded.transform(new_features)

y_pred_ols      = ols_loaded.predict(X_new_s)
```

```
y_pred_catboost = catboost_loaded.predict(new_features)     # CatBoost: unscaled
y_pred_mlp      = mlp_loaded.predict(X_new_s).flatten()

print("\nExample predictions on 5 new observations:")
print(f"  OLS      : {y_pred_ols.round(0)}")
print(f"  CatBoost : {y_pred_catboost.round(0)}")
print(f"  Tuned MLP: {y_pred_mlp.round(0)}")
print(f"  Actual   : {y_test.head(5).values}")

print("\nCleaned_Research_Version.py completed successfully.")
```

[Cell 29]

```
# 16 | Computational Efficiency Profiling

import time
import pandas as pd
import numpy as np

timing = {}

def time_fit_predict_existing(name, model, X_train, y_train, X_test):
    # Re-fit (so that train_time is measurable in a controlled way)
    t0 = time.perf_counter()
    model.fit(X_train, y_train)
    t1 = time.perf_counter()
    train_time = t1 - t0

    # Inference timing
    n_runs = 100
    t_pred0 = time.perf_counter()
```

```
    for   in range(n runs):
        y pred = model.predict(X test)
    t_pred1 = time.perf_counter()
    total_inf_time = t_pred1 - t_pred0
    inf_time_per_call = total_inf_time / n_runs
    inf_time_per_sample = inf_time_per_call / len(X_test)

    return train time, inf time per sample


# 1) Linear models
linear_models = {
    "OLS"        : LinearRegression(),
    "Ridge"      : Ridge(alpha=1.0),
    "Lasso"      : Lasso(alpha=0.001),
    "ElasticNet" : ElasticNet(alpha=0.01, l1 ratio=0.5)
}

for name, mdl in linear_models.items():
    t_train,  t_inf  =  time_fit_predict_existing(name,  mdl,  X_train_s,  y_train,
X test s)
    timing[name] = {
        "train_time_s": t_train,
        "inf_time_ms_per_sample": t_inf * 1000.0
    }
    print(f"{name:15s} | train={t_train:.4f}s  inf={t_inf*1000:.4f} ms/sample")


# 2) Tree / boosting models
tree_models_for_timing = {
```

```
    "RandomForest":    RandomForestRegressor(n estimators=800,    random state=42,
n jobs=-1),

    "XGBoost"     : XGBRegressor(n estimators=800, learning rate=0.05,

                                  max_depth=7,                       subsample=0.7,
colsample_bytree=0.7,

                                  random_state=42, n_jobs=-1, tree_method="hist"),

    "LightGBM"    : LGBMRegressor(n_estimators=800, learning_rate=0.05,

                                   num leaves=31, subsample=0.7,

                                   colsample bytree=0.7, random state=42),

    "CatBoost"    : CatBoostRegressor(iterations=800, learning_rate=0.05,

                                      depth=7, random_state=42, verbose=False,

                                      loss_function="RMSE")

}


for name, mdl in tree models for timing.items():

    if name == "CatBoost":

        X_tr, X_te = X_train, X_test

    else:

        X_tr, X_te = X_train_s, X_test_s


    t train, t inf = time fit predict existing(name, mdl, X tr, y train, X te)

    timing[name] = {

        "train_time_s": t_train,

        "inf_time_ms_per_sample": t_inf * 1000.0

    }

    print(f"{name:15s} | train={t_train:.4f}s  inf={t_inf*1000:.4f} ms/sample")


# 3) Tuned MLP (assuming you have a function build mlp model() and X train s,
X test s)
```

```
# Get the best hyperparameters from the tuner

best hps = tuner.get best hyperparameters(num trials=1)[0]



# Define a function that builds and compiles the model with the best hyperparameters

def build_best_mlp_model():

    model = build_mlp(best_hps) # Use the existing build_mlp function from
KerasTuner setup

    return model



# Define EPOCHS and BATCH_SIZE

EPOCHS = 50 # A reasonable number for profiling, not necessarily optimal training

BATCH_SIZE = 32



def time keras model(name, build fn, X tr, y tr, X te):

    model = build fn()

    t0 = time.perf_counter()

    model.fit(X_tr, y_tr, epochs=EPOCHS, batch_size=BATCH_SIZE,

              validation_split=0.2, verbose=0)

    t1 = time.perf_counter()

    train time = t1 - t0



    n_runs = 20

    t_pred0 = time.perf_counter()

    for _ in range(n_runs):

        y_pred = model.predict(X_te, verbose=0)

    t pred1 = time.perf counter()

    total inf time = t pred1 - t pred0

    inf time per call = total inf time / n runs

    inf_time_per_sample = inf_time_per_call / len(X_te)
```

```
    return train time, inf time per sample


t_train_mlp, t_inf_mlp = time_keras_model("Tuned MLP", build_best_mlp_model,
X_train_s, y_train, X_test_s)

timing["Tuned MLP"] = {

    "train_time_s": t_train_mlp,

    "inf time ms per sample": t inf mlp * 1000.0

}

print(f"Tuned MLP          | train={t_train_mlp:.4f}s   inf={t_inf_mlp*1000:.4f}
ms/sample")



# 4) Convert to DataFrame

eff df = (pd.DataFrame.from dict(timing, orient="index")

          .sort index())

eff_df
```

[Cell 33]

```
eff df.to csv(f"{RESULTS DIR}/computational efficiency.csv", index=True)

# ============================================================================

# END OF SCRIPT
```

ANNEXURE - B

List of Publications

- **A. R. Rathod and V. M. Patel**, "Data-Driven Automation for Solid Waste Forecasting Using Machine Learning, with COVID-19 Impact Analysis: A Case Study of Surat, Gujarat," Journal of Solid Waste Technology and Management, Apr. 2026, doi: https://doi.org/10.5276/jswtm/iswmaw/522/2026.515 ***(Scopus-Indexed)***

- **A. R. Rathod and V. M. Patel**, "AI-Driven Decision Support System for Optimized Municipal Solid Waste Management: A Machine Learning and Route Optimization Framework for Smart Cities," International Journal of Applied Mathematics, Dec. 2025, doi: https://doi.org/10.12732/ijam.v38i12s.1565 ***(Scopus-Indexed)***

- **A. R. Rathod and V. M. Patel**, "Comparative Study of Machine Learning Techniques for Predictive Modeling: Approaches, Applications, and Performance Evaluation," in 27th ISTE Conference on Revolutionizing Multidisciplinary Technical Education with Innovative Instructional Strategies: Vision 2047 for New Bharat, May 2025, pp. 312-316. ***(National Conference)***

- **A. R. Rathod and V. M. Patel**, "A Systematic Review of Models for Predicting Municipal Solid Waste Generation for Efficient Waste Management," International Journal of Scientific Development and Research, Apr. 2025. ***(Peer Reviewed)***

- **A. R. Rathod and V. M. Patel**, "A Comparative Study of Advanced Machine Learning and Deep Learning Models for Municipal Solid Waste Forecasting: A Case Study of Surat, India," Journal of Engineering and Technological Sciences, , vol. 58, no. 5, pp. 613–629, Aug. 2026. https://doi.org/10.5614/j.eng.technol.sci.2026.58.5.1

Journal of Solid Waste Technology and Management http://doi.org/10.5276/jswtm/iswmaw/522/2026.515

Data-Driven Automation for Solid Waste Forecasting Using Machine Learning, with COVID-19 Impact Analysis: A Case Study of Surat, Gujarat

Abhijit R Rathod[1,*], Vinodkumar M Patel[2]

[1]Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India
[2]Shantilal Shah Engineering College, Bhavnagar, Gujarat, India

Received on: October 19, 2025
Accepted on: December 26, 2025

ABSTRACT

Traditional methods for municipal solid waste (MSW) forecasting rely mainly on simplistic models that fail to capture complex urban dynamics, particularly during disruptive events like pandemics. Current machine learning approaches do not include policy-driven variables and real-time adaptability for municipal automation systems. How can machine learning models be enhanced with interrupted-time-series analysis to quantify pandemic impacts and provide automation-ready MSW forecasting for smart city governance? In order to find the solution, this study compiled 199 months of MSW data (2009-2024) from Surat Municipal Corporation (SMC). A rigorous preprocessing is employed by engineering temporal lag features (1-12 months) and creating a COVID-19 binary indicator (March 2020-March 2022). In our study, six supervised regression models were evaluated: Linear Regression, Random Forest, XGBoost, Support Vector Regression, Stacking Regressor, and Multilayer Perceptron, using rigorous interrupted-time-series comparison with and without pandemic indicators. Random Forest achieved highest accuracy (R^2 = 0.85), with lag_1 and residential waste as strongest predictors. The COVID-19 dummy reduced RMSE by 16% in Linear Regression and 10% in XGBoost thus demonstrating measurable pandemic impact quantification. SHAP analysis confirmed temporal dependencies and residential consumption patterns as the primary drivers. Our study systematically integrates interrupted-time-series methodology with the machine learning modeling for MSW forecasting which provides explicit pandemic impact measurement and automation-ready frameworks for municipal decision-making for their operations. Results we obtained demonstrate that hybrid ML + policy variable approaches enable resilient and interpretable forecasting systems capable of adapting to urban disruptions while supporting data-driven smart city waste management automation.

Keywords: Machine Learning, Municipal Solid Waste, Time-Series Forecasting, Pandemic Impact, Ensemble Learning, Interrupted-Time-Series.

AI Disclosure: Minimal use of AI Assistance for Language and Style: The author acknowledges the use of AI-based tools for minor language editing or grammatical correction. However, all intellectual contributions, analyses, and conclusions presented in these manuscripts are entirely the author's own.

* *Corresponding author: Abhijit R Rathod, Email: arrathod1709@gmail.com*

Dr. Vinodkumar M. Patel	**Dr. Umesh Kumar Khare**	**Dr. Yogeshkumar S. Patel**
Research Supervisor	**DPC Member - 1**	**DPC Member - 2**

International Journal of Applied Mathematics

Volume 38 No. 12s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

AI-DRIVEN DECISION SUPPORT SYSTEM FOR OPTIMIZED MUNICIPAL SOLID WASTE MANAGEMENT: A MACHINE LEARNING AND ROUTE OPTIMIZATION FRAMEWORK FOR SMART CITIES

Abhijit R Rathod[1]*, Dr Vinodkumar M Patel[2]

[1]Research Scholar, Civil Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India.

[2]Professor in Civil Engineering, Shantilal Shah Engineering College, Bhavnagar, Gujarat, India.

*Correspondence: Abhijit R Rathod, arrathod1709@gmail.com

Abstract

The task to manage the municipal solid waste (MSW) in the urban localities specifically in the cities which are growing faster is very challenging as the city administration need to cope up with the pace at which today's cities are growing which is almost impossible without incorporating the latest technologies like Artificial Intelligence (AI). The static waste collection system which is being used conventionally currently facing issues like inadequate allocation of the resources required which in turn raises the overall costs of the operations and ultimately harm the environment, too. To overcome this, the present study introduces an innovative approach of combining Decision Support System (DSS) with machine learning driven waste generation forecasts. This study attempts to optimize the vehicle routing to improve the MSW collection efficacy in the city of Surat, India. The framework proposed here combines Long Short-Term Memory (LSTM) neural networks for MSW prediction with Capacitated Vehicle Routing Problem (CVRP) algorithms for the optimization of routes. We have developed our LSTM model using the dataset of past 16 years (2009-2024) obtained from Surat Municipal Corporation which achieved considerable forecasting accuracy with RMSE of 0.1095 and R^2 value of 0.93. The integrated DSS demonstrated considerable benefits in operations, cutting the overall collecting distance by 50.31% and the use of vehicles by 24% compared to traditional static routing methods. The technology uses predictive analytics to dynamically allocate the required resources helping the waste management authorities make decisions in real time. This study enhances sustainable urban development by offering a data-driven methodology that corresponds with the objectives of the Smart Cities Mission and Sustainable Development Goals (SDGs) 11, 12, and 13. The suggested architecture provides scalable methods for implementing a circular economy and making municipal waste management systems more environmentally friendly.

Keywords: Artificial Intelligence, Machine Learning, LSTM, Municipal Solid Waste, Decision Support System, Vehicle Routing Problem, Smart Cities, Circular Economy

Received: August 04 2025



Dr. Vinodkumar M. Patel
Research Supervisor



Dr. Umesh Kumar Khare
DPC Member - 1



Dr. Yogeshkumar S. Patel
DPC Member - 2

Comparative Study of Machine Learning Techniques for Predictive Modeling: Approaches, Applications, and Performance Evaluation

Abhijit R. Rathod[#], Dr. Vinodkumar M. Patel[*]

[#]*PhD Scholar, GTU, Ahmedabad, Gujarat, India.*

[*]*Professor, Civil Engg., S. S. Engg. College, Bhavnagar, Gujarat, India.*

[1]arrathod1709@gmail.com

[2]vmpatel@civil.ssgec.ac.in

***Abstract*— Machine learning is now a key part of making predictions in many fields, from medicine to city planning. There are several machine learning methods that I compare in this study in terms of how well they make predictions. Some of the models I talk about are linear and random forest for regression, decision trees and SVM for classification, and ARIMA and LSTM for time series. Basically, predictive modeling approaches look at the past data and tries to assume its behaviour in future. Now-a-days many companies use this tool to make better decisions and manage their budget accordingly. The models are evaluated on datasets from a variety of fields, such as demand forecasting and municipal solid waste generation. The objective of this study is to identify the performance of different predictive models under a variety of conditions and try to differentiate among them based on their performances. Finally try to determine which model is most appropriate for a particular type of prediction problem. The paper tries to compare various predictive models used by the researchers and try to provide some practical output which will help select the appropriate approach by basically accuracy of models on past. The objective of this investigation is to assist researchers and practitioners in navigating the diverse array of models that are available and selecting the one that is most suitable for their requirements.**

***Keywords*— Machine Learning, Predictive Modeling, Regression Models, Time Series Forecasting, Model Evaluation.**

I. Introduction

From healthcare and finance, marketing, urban planning and environmental management specifically solid waste management in municipalities of the cities, predictive modeling has become a very important tool across all the disciplines of science and engineering. The big organizations of the world are becoming now data-driven and they take decisions and use their resources in a very optimum way [1]. Particularly, machine learning (ML) has become a potent tool for these activities, enabling practitioners to use past data for precise forecasts [2]. This paper contrasts various machine learning methods often employed for prediction, including regression models (e.g., linear regression, random forest), classification techniques (e.g., decision trees, support vector machines, SVM), and time-series forecasting approaches such as ARIMA and Long Short-Term Memory (LSTM) networks. From financial forecasting to demand prediction [3], these models have been used on a variety of prediction tasks; environmental applications include forecasting municipal solid waste generation [3].

Predictive modeling is basically a very good tool to see coming trends after careful analysis of past data. Ensuring correct outcomes and effective resource allocation depends on choosing the most appropriate model for a particular prediction task given the range of ML algorithms available [4]. To identify and use specific model for the specific purpose is very challenging task because each and every model has its own peculiarities and they perform accordingly. It is really a difficult task to find one which fits all the needs because each and every model has got its own advantages and disadvantages. This paper aims to offer a comparative analysis of several popular machine learning models, specifically predictive one and try to evaluate their performance across multiple domains and applications in engineering. By understanding the strengths and limitations of each approach, this study seeks to provide insights that will help researchers and practitioners choose the most appropriate model for their specific needs and datasets.

II. Literature Review

Because of recent advances in various application of machine learning (ML) techniques predictive modeling has gradually taken a great role in modern day decision-making. In this section, we provide a detailed overview of various predictive models, their applications, and how previous studies have compared these models to help guide researchers in selecting the most suitable method for their specific tasks.

Types of Predictive Models

It can be broadly classified into following categories: Supervised learning, Unsupervised learning and Deep learning.

1. Supervised Learning: When the outcomes are known based on available data, this approach is preferred. It further uses two types of modeling methods. Namely, Regression Models and Classification Models.

Regression Models: Linear regression, Random forests, Support vector regression (SVR) are few of the examples of regression modeling approach. All these models attempt to forecast continuous outcomes based on inputs [5]. Various applications include industries like healthcare because for example if one wants to get constant updates on the patient

ISBN: 978-81-987520-2-4

Dr. Vinodkumar M. Patel
Research Supervisor

Dr. Umesh Kumar Khare
DPC Member - 1

Dr. Yogeshkumar S. Patel
DPC Member - 2

A Systematic Review of Models for Predicting Municipal Solid Waste Generation for Efficient Waste Management

[1]Abhijit R. Rathod, [2]Dr. Vinodkumar M. Patel
[1]Research Scholar, [2]Professor in Civil Engineering
[1]Gujarat Technological University, Ahmedabad, India, [2]Shantilal Shah Engineering College, Bhavnagar, India

Abstract—This study investigates the current state of municipal solid waste (MSW) prediction models to support better advanced planning help in waste management. As we are observing urbanization taking place at a very alarming rate, we are also witnessing huge increase in waste generation. Hence accurate forecasting models are essential for policymakers and municipal authorities. The review examines 31 research papers from 2015 to 2025, analyzing modeling approaches such as Artificial Neural Networks (ANN), Decision Trees (DT), Time Series Forecasting (TSF), and Optimization Modeling (OM). The study classifies waste streams into Material Streams (MS) and Collection Streams (CS), with 70% of studies focusing on material streams for waste composition analysis. The findings highlight the importance of independent variables (socioeconomic, demographic, and environmental factors) in influencing MSW predictions. The paper concludes that hybrid machine learning models and real-time data integration can significantly improve the accuracy of waste forecasting, contributing to a more sustainable urban environment.

Index Terms— Forecasting, Municipal solid waste, Prediction models, Solid waste generation, municipal solid waste, prediction models, solid waste generation, forecasting. (key words)

I. INTRODUCTION

The constant increase in globalization and urbanization has caused serious concern to the efficient waste management, especially solid waste. It requires a huge budgetary provision to deal with collection and disposal of wastes generated in urban areas. When majority of portion of the budget involving in it, some insights of solid wastes to be generated or produced in next couple of years will be a boon to the administration of particular province or city. Initially simple regression analysis followed by basic mathematical modeling were employed to make the prediction of the generation of municipal solid waste in a town or city.

The globe currently produces about 2 billion metric tons of municipal solid waste (MSW) per year, but if no major changes are made to waste management procedures, estimates indicate that this amount might rise by as much as 70% by 2050, to almost 3.8 billion metric tons[1]. As the municipal solid waste is closely connected with the community as compared to others wastes, a great attention it deserved to maintain sustainable environment for coming generations. The UNEP report's [2] modeling suggests that adopting waste prevention and management strategies could bring down the yearly costs to of approximately USD 270.2 billion by 2050. It is high time to develop a proper approach for preparing a robust forecasting model which will help getting goals of circular economy achieved globally. A sustainable approach in various businesses and properly planned waste management can lead to save an annual amount of approximately USD 108.5 billion. To plan budgetary and other requirements in a city or town, a suitable waste generation prediction can help and provide a very good support to the decision-makers and the same can be insightful to formulate various policies as well as planning a long-term strategy that will be ultimately helpful in waste reduction. [3]

A systematic examination of 45 approaches to generation modeling was conducted by *P. Beigl et al.* [4], which included models up to the year 2005. The analysis uncovered four distinct classification criteria: geographical scope, categories of waste streams modeled, independent variables chosen, and techniques of modeling. This investigation identified four key classification criteria: geographical scope, categories of waste streams, types of independent as well as dependent variables, and modeling techniques. P. Beigl et. al. [4] also proposed a relevance tree for selecting methodology keeping in mind selecting type of waste streams. The other attempt was made by *K. A. Kolekar et al.,* [5] considering 20 MSW generation prediction models from various literature from the year 2006 to 2014.The primarily finding of the study shows the most of the models relied on regression analysis and correlation. (almost 20% of studied models), while some of the researchers also made an attempt by using fuzzy logic also which is based on artificial intelligent systems.

The study aims to evaluate how different modeling methods are being used in predicting or forecasting the municipal solid waste generation from the year 2015 to 2025. The study is carried out to get better idea about latest trends followed for prediction modeling of municipal solid waste particularly in last decade.





Dr. Vinodkumar M. Patel
Research Supervisor

Dr. Umesh Kumar Khare
DPC Member - 1

Dr. Yogeshkumar S. Patel
DPC Member - 2

Journal of Engineering and Technological Sciences

Research Paper

A Comparative Study of Advanced Machine Learning and Deep Learning Models for Municipal Solid Waste Forecasting: A Case Study of Surat, India

Abhijit R. Rathod[1,*], & Vinodkumar M. Patel[2]

[1]Gujarat Technological University, GTU Campus, Nr. Visat Three Road, Visat-Gandhinagar Road, Chandkheda, Ahmedabad-382424, Gujarat, India
[2]Shantilal Shah Engineering College, New Sidsar Campus, Vartej, Sidsar, Bhavnagar - 364060, Gujarat, India

*Corresponding author: arrathod1709@gmail.com

Abstract

To make the city planning sustainable, especially in rapidly growing cities of the world like Surat in India, the implementation of effective waste management is crucial. The primary factor governing this is the ability to accurately predict the quantity of municipal solid waste (MSW) likely to be generated. This study presents a comprehensive comparative analysis of various predictive models including linear regressions, kernel approaches, gradient boosting as well as the deep learning architectures. Using historical data from Surat, systematic preprocessing and feature engineering generated 419 features representing temporal, socio-economic, climatic, COVID-19, and mobility factors. The novel contribution of this study is the systematic feature engineering framework that explicitly encodes temporal structure (419 engineered features including lagged values, rolling statistics, and seasonal decomposition), enabling simple linear models to capture complex waste generation patterns. Ten distinct models, ranging from statistical approaches to machine learning and deep learning were evaluated and compared. Advanced ensemble models, including LightGBM (R^2 = 0.983), CatBoost (R^2 = 0.977), and XGBoost (R^2 = 0.970) demonstrated strong performance. The best-performing models (OLS and Gaussian Process Regression) achieved R^2 = 0.997 with Mean Absolute Percentage Error (MAPE) = 1.43%. In this study, linear models trained within a few milliseconds and achieved per-sample inference times on the order of 0.004-0.008 ms, whereas the tuned MLP and tree-based ensembles required seconds of training and millisecond-level inference, corresponding to differences of roughly two to three orders of magnitude in computational cost. Other notable performers include Lasso regression (R^2 = 0.979), tuned MLP (R^2 = 0.968), and Random Forest (R^2 = 0.960). The results demonstrate that feature engineering has greater influence on forecasting accuracy than model complexity, with OLS using engineered features (R^2 = 0.997) outperforming the MLP model (R^2 = 0.966) by approximately 3.1% while providing substantially faster predictions. Feature importance analysis identified lagged MSW values, rolling statistics, demographic indicators, festival effects, and COVID-19 lockdown impact as key predictors. The research finds that adoption of systematically designed feature engineering framework is a valuable tool for MSW management. The study provides comprehensive model benchmarking and practical recommendations for policymakers pursuing sustainable urban development.

Keywords: *deep learning; forecasting; gradient boosting; machine learning; municipal solid waste; Surat; time series.*

Introduction

The 21st century is witnessing unprecedented and rapid urbanization across the globe. The global economy is being reshaped by the demographic shift and dramatic increase in generation and complexity of municipal solid waste (MSW). These rapidly growing cities face huge pressure to handle their waste management systems as a result of this change. What if city management fails to plan adequately and accurately? The consequences are severe: overflowing landfills create public health hazards; encourage illegal dumping that pollutes water sources and soil; and cause transportation disruptions that paralyze municipal operations. For example; Surat; India; one of India's fastest-growing cities saw waste generation surge from 3;800 tonnes per day in 2009 to 6;200 tonnes per day in 2024. Without accurate forecasting; municipal authorities struggle to deploy collection vehicles; schedule landfill operations; and allocate limited budgets effectively.

Copyright by authors ©2026 Published by IRCS - ITB
ISSN: 2337-5779
J. Eng. Technol. Sci. Vol. 58, No. 5, 2026, 613-629
DOI: 10.5614/j.eng.technol.sci.2026.58.5.1

Dr. Vinodkumar M. Patel
Research Supervisor

Dr. Umesh Kumar Khare
DPC Member - 1

Dr. Yogeshkumar S. Patel
DPC Member - 2