2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)

# Waste Forecasting and Route Optimization for Smart Waste Collection System

Kwan-Hua Sim*
Faculty of Engineering, Computing & Science
Swinburne University of Technology Sarawak
Kuching, Malaysia
khsim@swinburne.edu.my

Felix Vincent[2]
Faculty of Engineering, Computing & Science
Swinburne University of Technology Sarawak
Kuching, Malaysia

Sing-Hong Lei[3]
Faculty of Engineering, Computing & Science
Swinburne University of Technology Sarawak
Kuching, Malaysia

Hudyjaya Siswoyo Jo[4]
Faculty of Engineering, Computing & Science
Swinburne University of Technology Sarawak
Kuching, Malaysia

Patrick Then[5]
Faculty of Engineering, Computing & Science
Swinburne University of Technology Sarawak
Kuching, Malaysia

***Abstract*—Efficient municipal solid waste (MSW) management is crucial for addressing environmental and operational challenges, particularly in urbanizing in developing country. This study presents an integrative framework combining waste weight forecasting with optimized route planning to enhance waste collection efficiency. An adjusted trend-seasonal index model captures dynamic seasonal variations, improving forecasting reliability. This model is paired with a modified Dijkstra's algorithm that optimizes routes by incorporating key factors such as predicted waste weight, road congestion, and fuel consumption. The proposed demonstrates its capability to reduce resource consumption and operational inefficiencies. By aligning predictive modeling with route optimization, the framework offers a solution for sustainable waste management. This research contributes to advancing digital solutions for resource-efficient and environmentally responsible waste collection systems.**

***Keywords-component; Waste collection; weight forecasting; collection route optimization***

## I. Introduction

Waste management remains a critical global issue, exerting significant impacts on public health, environmental sustainability, and economic resilience. In developing nations such as Malaysia, the challenges of municipal solid waste management have been exacerbated by rapid urbanization and population growth. Addressing these issues requires innovative and efficient strategies to improve waste collection processes, ensuring long-term environmental sustainability.

Malaysia generates an average of 38,207 tons of waste daily, equivalent to approximately 1.17 kg per person, primarily due to rapid urban development and rising living standards [1]. This substantial waste output is largely driven by accelerated urban development and an increasing standard of living [1]. The rising volume of waste has placed immense pressure on Malaysia's waste management infrastructure, which struggles to cope due to outdated documentation practices and inefficient collection systems [2].

Traditional waste collection practices are often plagued by inefficiencies, including poorly optimized collection routes and a lack of reliable forecasting tools. These limitations result in the overuse of essential resources such as fuel and labor, thereby driving up operational costs and reducing overall efficiency [3]. Furthermore, existing collection planning methods frequently rely on the assumption of a consistent volume of municipal solid waste (MSW) and uniform collection capacity. In reality, MSW volumes are highly dynamic, influenced by factors such as seasonal fluctuations, population growth, and shifting consumption [4].

This research addresses two pivotal challenges in waste management. First, traditional waste forecasting models often overlook the impact of dynamic seasonal variations, leading to inaccuracies in predicting waste generation. Forecasting is critical for optimizing resource allocation, as it allows waste management companies to align their operations more effectively with actual demand. Second, the existing route planning approaches, which primarily consider distance as the key factor, often lead to inefficiencies. These routing algorithms typically neglect other influential variables such as waste weight and road congestion, resulting in suboptimal utilization of resources and waste collection insufficiency.

To address these challenges, this study proposes the development of a robust MSW weight forecasting model that accounts for dynamic seasonal variations, thereby improving the accuracy and reliability of waste generation predictions. Building on these predictions, the study aims to optimize waste collection routes by enhancing Dijkstra's algorithm to incorporate additional parameters such as predicted waste weight, road congestion, and other weight factors.

This study is conducted in collaboration with a major local waste management company in Kuching city, Trienekens (Sarawak) Sdn. Bhd. The proposed smart waste collection system, as illustrated in Fig. 1, integrates an advanced data collection module comprising RFID tags, weight sensors, and

979-8-3315-0976-7/25/$31.00 ©2025 IEEE

location sensors installed on compactor trucks. These sensors facilitate the acquisition of critical operational data during waste collection activities. This study is conducted in a geographical location with a relatively small population of approximately 300,000 and without the influence of four distinct phases of the annual climate cycle. Additionally, as the entire waste collection routine occurs during the early hours of the morning, data related to factors such as weather conditions and complex traffic control are not collected as they are excluded from the scope of this study.

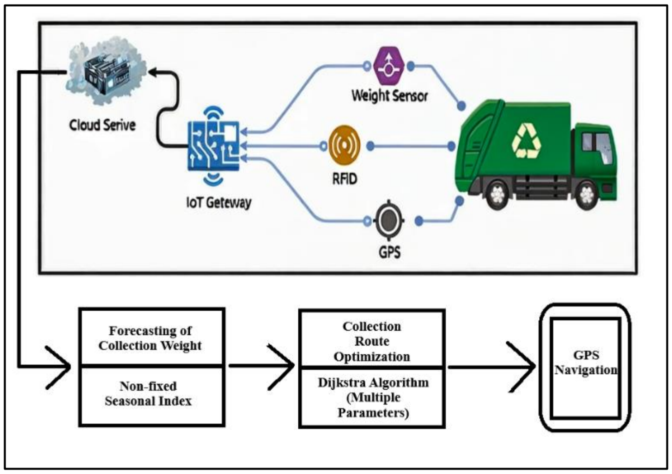


Figure 1. Smart Waste Collection System

All the relevant data collected from the compactor tracks will subsequently be utilized as historical data to predict the weight of waste for future collections. Based on the forecasted weight, along with additional parameters such as distance and travel duration influenced by congestion, Dijkstra algorithm is employed to perform collection route optimization by prioritizing on overall fuel consumption. This integrative approach seeks to create more efficient and adaptive routing strategies, enabling waste collection systems to proactively respond to the fluctuating demands of urban waste management.

This paper presents the interim findings and results of study done on forecasting the weight of MSW and evaluating the effectiveness of collection route optimization. The aim of this phase is to evaluate the viability and performance of these two modules before proceeding to the next phase, which involves real-world trial in practical and operational settings.

## II. Literature Review

Effective MSW management hinges on the implementation of reliable forecasting models and optimized collection routes. Recent advancements in time series analysis, mathematical modeling, and computational technologies offer promising solutions to overcome these challenges, enabling more precise waste generation forecasts and efficient route optimization.

Time series models, such as the Seasonal Autoregressive Integrated Moving Average (SARIMA), have proven effective in capturing seasonal and non-stationary trends in MSW data. SARIMA has been used for month-level predictions, successfully accounting for short-term seasonal variations. In contrast, the GM (1,1) model from Grey System theory is particularly advantageous for long-term forecasting, especially when sample data is limited, making it well-suited for annual MSW predictions[5]. Additionally, integrating Grey Relational Analysis (GRA) improves forecasting accuracy by synthesizing the outputs of SARIMA and GM (1,1), enhancing medium-term predictions through the identification of correlations across models [6].

In addition to statistical approaches, machine learning techniques, particularly Long Short-Term Memory (LSTM) networks, have gained significant attention for time-series forecasting across various domains. Empirical studies have been conducted to compare the performance of SARIMA and LSTM in time series forecasting, highlighting the strengths and limitations of both models [7][8]. While SARIMA excels at capturing linear and seasonal patterns, LSTM's ability to adapt to nonlinear trends makes it particularly well-suited for dynamic and complex datasets.

Optimizing waste collection routes is integral to translating forecasted waste volumes into practical and efficient waste management solutions, with numerous methodologies designed to mitigate operational inefficiencies and environmental impacts. Among these methodologies, Dijkstra’s algorithm has emerged as the widely adopted graph-based optimization technique approach [9]. Hossain (2020) introduced a dynamic route optimization model in which garbage trucks adaptively adjust their paths based on real-time factors such as distance, bin status, and road congestion [10].

Building on this model, another study incorporated fuel consumption as a critical factor in optimizing collection routes [11]. By modeling phases such as driving, collection, and waste disposal, the study quantified fuel utilization based on vehicle load and route characteristics, enhancing the efficiency and sustainability of waste management operations. It further employed random forest (RF) modeling, successfully capturing up to 97% of the data variance for $CO_2$ emissions. The RF models demonstrated effectiveness in identifying the most influential variables for each prediction model, with fuel consumption emerging as the most significant factor in predicting emissions of $CO_2$, CO, and $NO_x$. The findings highlight the potential of RF ensemble models to improve the accuracy of vehicular emission predictions, particularly in urban areas of developing countries, where precise modeling is essential for informed environmental management [11].

The complexity of optimizing multiple parameters in waste collection has led to the development of multi-objective optimization models to address the intricate trade-offs involved [12]. In a recent study, a mathematical model was introduced for the collection and transportation of solid waste, targeting three primary objectives: minimizing total travel distance, balancing workloads among collection teams, and reducing waste leakage. The model utilized a multi-objective 0-1 integer nonlinear programming framework, which was solved using a multi-objective evolutionary algorithm. This methodology effectively balanced competing priorities, showcasing its applicability to the environmental, social, and economic dimensions of waste management, and providing a robust framework for operational efficiency and sustainability [13].

## III. SMART WASTE COLLECTION SYSTEM

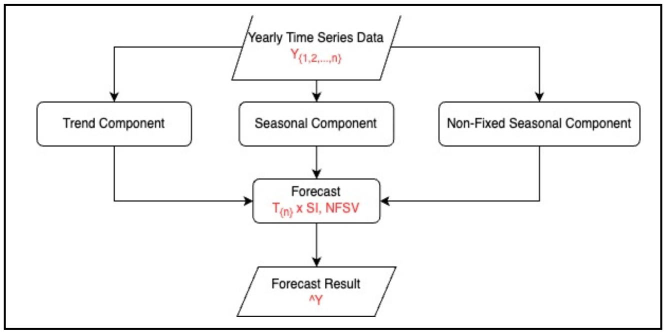


Figure 2. Trend-seasonal index model

Fig. 2 illustrates the workflow of the adjusted trend-seasonal index model used to forecast weight of MSW. It is divided into three components: Trend, Seasonal, and Non-Fixed Seasonal Component. The time series data is split into yearly segment T_({1,2,...,i,...,n}), the trend value for the latest year, Tn, is then utilized to forecast the MSW weight for the upcoming cycle. Additionally, the series of yearly trend values T_({1,2,...,i,...,n}) is used in the Seasonal component to obtain the detrended data. By obtaining the latest trend value $T_n$, the seasonal indices SI, and the non-fixed season variation, forecast results can be generated. For periods affected by non-fixed festive seasons, the forecasted weight is replaced with the calculated variation NFSV.

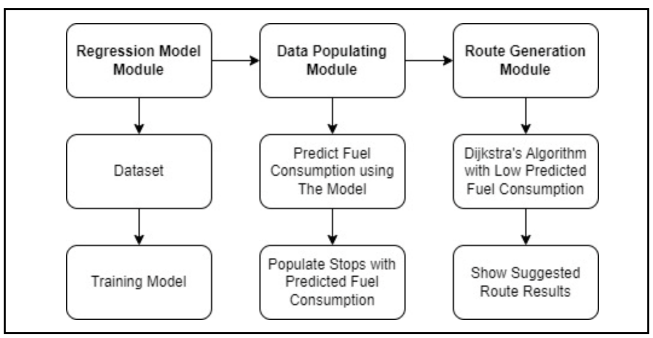


Figure 3. Waste collection route optimisation

The waste collection routes optimization is structured into three key modules, each contributing to the overall goal of minimizing fuel consumption. As outlined in Figure 3, the process begins with the Regression Model Module, in which a predictive model is developed using Random Forest techniques. This model is based on critical variables, such as distance, road congestion, and waste weight. Constructed using historical waste collection data, it serves as the foundation for predicting fuel consumption across various routes, providing valuable insights into the efficiency of different collection paths.

Once the Random Forest model was trained and validated, the Data Populating Module was utilized to predict fuel consumption for new datasets. These datasets, containing specific area codes, included information on distance, road congestion, and waste weight. Relevant features were extracted from these datasets, and the trained Random Forest model was applied to predict fuel consumption based on these input variables. This step allowed for the validation of the model's predictions against actual operational data, ensuring its ability to generalize effectively across different areas and conditions. The resulting predictions were then populated into a dataset, which served as the input for the subsequent steps of the optimization process.

Finally, in the Route Generation Module, the optimized fuel consumption data is fed into the Dijkstra's algorithm. Unlike traditional routing algorithms that prioritize the shortest distance, this modified algorithm generates routes that prioritize minimizing fuel consumption. It incorporates multiple variables that affect fuel consumption, including distance, road congestion, and MSW weight. By considering these factors, the algorithm generates more efficient routes that reduce fuel usage, aligning with the goal of optimizing waste collection operations for both cost-effectiveness and sustainability.

In this context, the waste collection areas and routes were represented as a graph, where nodes corresponded to collection points and edges represented the routes between these points. The edges were assigned weights that indicated either distances or predicted fuel consumption, depending on the variable being optimized. Dijkstra's algorithm was then employed to determine the optimal path that visits all collection points while minimizing fuel consumption. This approach considered various factors affecting route efficiency, such as distance, road congestion, and waste weight, allowing for the generation of more fuel-efficient collection routes.

## IV. RESULT & DISCUSSION

The MSW weight data utilized for MSW weight forecasting in this research is obtained from Trienekens Sdn. Bhd., a local waste management company based in Kuching, Malaysia. The data covers a period from 2012 to 2022 and includes the weekly weight of MSW collected. Since the collection frequency per week varies between 0 and 3, the MSW weight data have been aggregated into weekly data points for comparison against conventional trend-seasonal index model, LSTM and SARIMA model. Consequently, each area comprises a total of 574 data points. Meanwhile, another dataset comprises over 1,800 records of detailed daily waste collection data was also obtained, and this dataset contains key variables such as location, time and waste weight in tons and kilograms.

TABLE I. ERROR METRICS COMPARISON FOR MHH5

| | MSE | RMSE | MAE | MAPE |
|---|---|---|---|---|
| **Adjusted Trend-Seasonal Index Model** | 6,410,455.67 | 2,531.89 | 1,708.64 | 10.00% |
| **Conventional Trend-Seasonal Index Model** | 6,638,092.59 | 2,576.45 | 1,869.79 | 10.71% |
| **LSTM Model** | 5,837,345.01 | 2,416.06 | 1737.72 | 10.76% |
| **SARIMA Model** | 5,261,165.75 | 2,293.72 | 1620.80 | 9.50% |

TABLE II. ERROR METRICS COMPARISON FOR MHH6

| | MSE | RMSE | MAE | MAPE |
|---|---|---|---|---|
| **Adjusted Trend-Seasonal Index Model** | 4,140,854.98 | 2,034.91 | 1,501.90 | 9.27% |
| **Conventional Trend-Seasonal Index Model** | 13,357,771.49 | 3,654.83 | 3,237.85 | 19.14% |
| **LSTM Model** | 4,068 ,470.20 | 2,017.04 | 1,475.92 | 9.48% |
| **SARIMA Model** | 5,570,000.16 | 2,360.08 | 1,768.69 | 11.12% |

TABLE III. ERROR METRICS COMPARISON FOR MHH7

| | MSE | RMSE | MAE | MAPE |
|---|---|---|---|---|
| **Adjusted Trend-Seasonal Index Model** | 4,592,361.55 | 2,142.98 | 1,483.93 | 9.67% |
| **Conventional Trend-Seasonal Index Model** | 4,138,272.00 | 2,034.27 | 1,373.10 | 9.13% |
| **LSTM Model** | 3,618,807.42 | 1,902.32 | 1,270.39 | 8.78% |
| **SARIMA Model** | 6,847,903.79 | 2,616.85 | 2,130.25 | 13.46% |

In the MHH5 (Table I), the SARIMA model demonstrates the lowest error metrics, outperforming the other models. It is followed by the adjusted trend-seasonal index model, which also shows strong predictive performance. A t-statistic analysis reveals no significant differences for the year 2019, while the t-test values for 2021 and 2022 are lower compared to those observed in area MHH6. As a result, the full potential of the adjusted trend-seasonal index model is not realized in MHH5. Nonetheless, this model still outperforms both the conventional trend-seasonal index model and the LSTM model, indicating its relative efficacy in forecasting within this context.

In the MHH6, the non-fixed festive season exhibits the most pronounced differences, as reflected in the t-statistics from 2019 to 2022, which reveal very low p-values and high t-test values. This variability underscores the impact of such periods on forecasting accuracy. As shown in Table II, the adjusted trend-seasonal index model achieves the best performance in forecasting MSW weight, outperforming the LSTM and SARIMA models. These results highlight that the stability of MSW weight during non-fixed festive seasons plays a critical role in enhancing the performance of the adjusted trend-seasonal index model, making it particularly suited for addressing the complexities associated with variable seasonal patterns.

In the MHH7, the years 2019 and 2021 exhibit no significant differences, as indicated by p-values exceeding 0.05. Similarly, for 2020 and 2022, the t-test values remain below 3, suggesting that the differences during the non-fixed festive season are not substantial. Table III shows that the adjusted trend-seasonal index model ranks third in performance. However, it is noteworthy that the Mean Absolute Percentage Error (MAPE) of the proposed model remains below 10%, which is indicative of good forecasting accuracy despite the model's relative ranking.

Overall, the Adjusted Trend-Seasonal Index Model demonstrates robust performance, achieving a Mean Absolute Percentage Error (MAPE) of approximately 10% or lower across all three areas. These results indicate the model's effectiveness in addressing the challenges associated with non-fixed festive seasons. However, its accuracy is significantly influenced by the stability of seasonal weight variations. Substantial fluctuations in these variations may adversely affect the model's predictive performance, highlighting the need for further adjustments or complementary approaches in highly variable contexts.

TABLE IV. FUEL CONSUMPTION AFTER ROUTE OPTIMIZATION

| Collection Area | Previous Average Fuel Consumption (litters) | Optimized Fuel Consumption (litters) | Percentage Reduction (%) |
|---|---|---|---|
| **AREA1** | 56.58 | 54.52 | 3.64% |
| **AREA2** | 52.92 | 37.54 | 29.06% |
| **AREA3** | 51.63 | 51.31 | 0.62% |
| **AREA4** | 49.25 | 23.94 | 51.39% |
| **AREA5** | 37.72 | 27.28 | 27.66% |

These findings in Table IV underscore the effectiveness of integrating predictive modelling with algorithmic route optimization for fuel efficiency. It aimed to assess the effectiveness of integrating predictive modelling with route optimization algorithms for reducing fuel consumption in waste collection operations. In AREA1, the optimized route consumed 54.52 liters of fuel, representing a 3.64% reduction from the previous average consumption of 56.58 litres. Similarly, in AREA2, fuel consumption was reduced to 37.54 liters, which corresponds to a 29.06% reduction from the initial average of 52.92 liters. For AREA3, the optimized route achieved a minor reduction in fuel consumption, lowering it to 51.31 liters, which is a 0.62% decrease from the prior average of 51.63 liters. In AREA4, the optimized route resulted in substantial fuel savings, consuming only 23.94 liters, marking a significant 51.39% reduction from the previous average of 49.25 liters. Finally, in AREA5, the optimized route reduced fuel consumption to 27.28 liters, representing a 27.66% decrease from the previous average of 37.72 liters.

## V. CONCLUSIONS

This study highlights the value of combining dynamic weight forecasting with optimized routing to improve MSW management. The adjusted trend-seasonal index model enhanced forecasting accuracy, while the modified Dijkstra's algorithm reduced fuel consumption by incorporating variables like waste weight and road congestion. These results demonstrate a data driven approach to improve the efficiency in waste collection systems. Accordingly, the proposed waste forecasting approach incorporated with non-fixed seasonal indices and the multi-parameter route optimization, are ready for the next phase of real-world trial and evaluation. Although this study does not include other factors such as weather conditions and advanced traffic controls, future work could easily integrate additional factors into the existing framework. By expanding the waste collection system to include these additional elements, the improved waste management system can achieve greater adaptability and resilience in diverse operational environments.

## ACKNOWLEDGMENT

This research is funded by Translational Research Grants (TRG) from Sarawak Digital Economy Corporation Bhd (SDEC)

## REFERENCES

[1] Ng, KS, Yeoh, L, Iacovidou, E, Ghani, W & Yamaguchi, A, Towards Sustainable Municipal Solid Waste Management in Malaysia, 2023

[2] Mohammad, S, Ahmad, Z, & Rahman, H, Challenges and Opportunities in Modernizing Waste Management Practices in Malaysia. International Journal of Waste Management and Environmental Impact, 45(3), pp. 223-235, 2021

[3] Zainu, ZA & Songip, AR, 'Policies, Challenges and Strategies for Municipal Waste Management Malaysia', Journal of Science, Technology and Innovation Policy, vol. 3, no. 1, pp. 10-14, 2021

[4] Guo, R., Liu, H.M., Sun, H.H., Wang, D., Yu, H., Do Rosario Alves, D. and Yao, L., Forecasting of municipal solid waste generation in China based on an optimized grey multiple regression model. Journal of Material Cycles and Waste Management, 24(6), pp.2314-2327, 2022

[5] Curceac, S, Ternynck, C, Ouarda, TB, Chebana, F & Niang, SD, 'Short-term air temperature forecasting using nonparametric functional data analysis and SARMA models', Environmental Modelling & Software, vol. 111, pp. 394-408, 2019

[6] Tang, T, Jiang, W, Zhang, H, Nie, J, Xiong, Z, Wu, X & Feng, W, 'GM(1,1) based improved seasonal index model for monthly electricity consumption forecasting', Energy, vol. 252, 2022/08/01/, p. 124041, 2022

[7] Kumar Dubey, A, Kumar, A, García-Díaz, V, Kumar Sharma, A & Kanhaiya, K, 'Study and analysis of SARIMA and LSTM in forecasting time series data', Sustainable Energy Technologies and Assessments, vol. 47, 2021/10/01/, p. 101474, 2021

[8] Dubey, A.K., Kumar, A., García-Díaz, V., Sharma, A.K. and Kanhaiya, K., 2021. Study and analysis of SARIMA and LSTM in forecasting time series data. Sustainable Energy Technologies and Assessments, 47, p.101474

[9] Liu, L.S., Lin, J.F., Yao, J.X., He, D.W., Zheng, J.S., Huang, J. and Shi, P., Path planning for smart car based on Dijkstra algorithm and dynamic window approach. Wireless Communications and Mobile Computing, 2021(1), p.8881684, 2021

[10] Hossain, M A, Ahmedy, I, Zaid Harith, M Z, Idris, M, Tey, K S, Md. Noor, R & Yusoff, S, 'Route Optimization by using Dijkstra's Algorithm for the Waste Management System', 10.1145/3388176.3388186, 2020

[11] Hassan, M.A, Salem, H, Bailek, N & Kisi, O, 'Random Forest Ensemble-Based Predictions of On-Road Vehicular Emissions and Fuel Consumption in Developing Urban Areas', Sustainability, 15, 1503, 2023

[12] Abdallah, M., Hamdan, S. and Shabib, A., A multi-objective optimization model for strategic waste management master plans. Journal of Cleaner Production, 284, p.124714, 2021

[13] Zhou, J, Zhang, M & Wu, S, 'Multi-Objective Vehicle Routing Problem for Waste Classification and Collection with Sustainable Concerns: The Case of Shanghai City', Sustainability, MDPI, 2022