

RESEARCH ARTICLE

EFFICIENCY VS. RESILIENCE: MODELLING, ZONING, AND COST TRADE-OFFS FOR WARD-LEVEL SOLID WASTE IN ONITSHA, ANAMBRA STATE, NIGERIA

Ezeakafor, S. C., Osakwe, O.T., Mmonwuba, N. C.

Department of Civil Engineering, Chukwuemeka Odumegwu Ojukwu University, Anambra State, Nigeria
*Corresponding Author Email: amaro4baya@yahoo.com

This is an open access article distributed under the Creative Commons Attribution License CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ARTICLE DETAILS

Article History:

Received 27 January 2026
Revised 20 February 2026
Accepted 25 February 2026
Available online 16 March 2026

ABSTRACT

The rapid urbanization and economic growth in Onitsha have worsened the level of municipal solid waste (MSW) production, exceeding the waste management capacity of the city. The current systems have caused environmental and health risks due to their inefficiency, and therefore, forecasting methods of various systems based on data are required. The nonlinear relationships between socio-economic, demographic and environmental conditions that prevail in traditional linear models as applied by municipal agencies have not been taken into consideration by them, leading to erroneous waste estimates and ineffective infrastructure planning. The study is expected to address (i) the estimation of the daily waste production at the ward level by utilizing field data and machine learning models, (ii) the prediction of the most important elements influencing the volume of waste, and (iii) the creation of the most effective service areas that optimally collect waste and its localization. Primary field data was obtained on 20 wards in Onitsha that was complemented by demographic and economic data of the Anambra State Waste Management Authority. Extreme Gradient Boosting (XGBoost) algorithm, Multiple Linear Regression (MLR), SHAP interpretability, and K-Means clustering were used. XGBoost showed better predictive power ($R^2 = 0.9984$), where population size, density, and commercial activity were found to be the dominating factors. Machine learning combined with demographic data will increase the accuracy of waste forecasting. Data-driven planning, deployment of more transfer stations, and reinforcement of spatial zoning ought to be a priority for policymakers in creating resilient and sustainable waste management in Onitsha.

KEYWORDS

Machine Learning, SHAP, Solid Waste, Waste Management, XGBoost, Regression

1. INTRODUCTION

The high rate of urbanization, population increase, and the growing levels of commercial activities in urban centers in Nigeria have enhanced the tremendous generation of solid waste in the cities, which has placed enormous strain on the available waste management facilities. An example of this problem is Onitsha, a populous commercial centre in the Anambra State, which generates huge quantities of Municipal Solid Waste (MSW) every day within households, markets and small to medium scale enterprises. Although important economically, the Solid Waste Management (SWM) system of the city is associated with unpredictable collection time, lack of disposal facilities, inefficient routes, and open dumping, which increase environmental and human health risks (Igwé and Evan, 2020; Okeke and Nwaiwu, 2025).

Empirical research in related urban environments has shown that solid waste production has a significant relationship with demographic and socioeconomic variables like population density, commercial intensity, household size, and consumption behaviour. Past studies in the Anambra State show that waste streams are organic in nature with a range of 4860 in the presence of leachate, methane gas, and vectors breeding environments in poorly managed states (Chukwunke et al., 2016; Nwakoby et al., 2020). The dumpsites have been observed to cause a contamination of groundwater with coliform bacteria and other risky parameters in Onitsha, which was poorly engineered, thus highlighting a system-wide infrastructural and regulatory failure (Chukwunke et al.,

2016).

Nevertheless, one of the limitations that has been mentioned in the studies is the continued lack of ward-level, data-resolved, and seasonally adjusted estimates of waste production. The current municipal records are not well organized, rough, and not suitable to predict space hotspots or loads in the future across the collection zones (Okeke et al., 2024). Also, most of the previous models utilize linear regressions, survey inferences, and deterministic models that do not support the nonlinear relationships between variables, which include: concurrent effects of population density, commercial index, and organic content on waste volumes. This gap constrains the structure of cost-effective waste management facilities, such as the ideal location of transfer stations, placing the collection trucks and the capacity of landfills.

The latest developments in the field of Machine Learning (ML) have revealed that the accuracy of prediction is better with complex environmental systems than classical methods (Ceylan et al., 2020). One of them, Extreme Gradient Boosting (XGBoost), has become well-known because it can learn nonlinear relationships, reduce overfitting by regularization, and learn missing or heterogeneous data (Chen and Guestrin, 2016). Experiments associated with using XGBoost in MSW intelligence have proven that it performs better in predicting waste production compared to traditional regressions and is thus suitable in high-variance waste streams typical of fast-urbanizing cities such as Onitsha.

Quick Response Code



Access this article online

Website:
www.jtin.com.my

DOI:
10.26480/jtin.01.2026.41.53

Moreover, the empirical reviews stress the importance of proper forecasting, which should be associated with the facility-scale design decisions, which are not much covered in the literature on Nigeria. It was found that there are limited works that link the output of the prediction with practical system design parameters, e.g., transfer-station capacity, routing clusters, or cost-efficient operational solutions (Okeke et al., 2024). Moreover, the application of Geographic Information Systems (GIS) and clustering algorithms like K-Means to optimize the zoning of spatial services has been investigated in other studies, but it has not been researched in the case of Onitsha (Hoque and Rahman, 2020). Therefore, the systemic burdens, lack of data resolution, nonlinear generation relationship, inadequate infrastructure planning, and non-sustained operational expenses explain why the application of an XGBoost-based modelling framework should incorporate variables of ward-level demographic, socio-economic, and environmental factors. This study will supplement the predictions with clustering and optimization analysis to support the evidence-based engineering of waste management facilities that increase the collection efficiency, travel distance reduction, and rational budgetary allocation.

The investigation, thus, will help the body of knowledge through (i) deriving high-resolution daily and annual MSW estimations of Onitsha, (ii) determining major determinants of waste generation based on interpretable Machine Learning methods, and (iii) creating facility-level infrastructure based on the quantitative projections of the load. The findings resonate with international demands of governance of urban waste based on data, and a viable solution to environmental sustainability, protection of human health, and preparedness of the cities to the circular economy in fast urbanizing cities in Africa (Ezeudu et al., 2021; Okeke and Nwaiwu, 2025).

In low- and middle-income settings, there is empirical evidence that solid-waste production is a joint function of socio-economic arrangement, city form, behaviour, and seasonality and that better prediction accuracy is obtained with the combination of different data and advanced analytics. Machine-learning-based models, in particular, Random Forest and Extreme Gradient Boosting, always exhibit better performance than traditional regression to forecast the household and city-level waste amounts, and the studies in Nigeria and Ghana also prove the topicality of demographic and economic covariates, including population density, income, household size, and commercial intensity. Spatial technologies were found to reinforce estimation by identifying the presence of hotspots and positioning collection logistics according to the geography (Hoque and Rahman, 2020), with time-series evidence showing strong seasonal variations related to climate, consumption cycles, and planners need to consider in service schedules. Similar evidence indicates that composition is important: organic fractions are dominant in most of the Nigerian streams, which is a sign of large composting possibilities and directed diversion. Demand-side levers, including behaviour and awareness, are shown to reduce generation and help improve segregation, and the system-dynamics and the view of a circular economy perspective state that policies that increase recycling levels, minimize material loops, and address uncertainty should be used (Emetere et al., 2021; Liu et al., 2021; Ezeudu et al., 2021).

Enlarging the area of Anambra State, empirical research records systemic, institutional and infrastructural bottlenecks that aggravate estimation and

service provision: long distances to official receptacles stimulate the use of informal collectors, underfunding, lax by-laws, and insufficient facilities are the primary impediments to formal systems, and streams of waste are generally organic and poorly segmented into Informal actors (scavengers) quantitatively decrease loads and reuse materials without being exposed to health risks and even thin local markets (Egbu and Okoroigwe, 2014; Igwe and Evan, 2020; Noiki et al., 2021; Igwe and Anaje, 2022). Broadcast campaigns to social media have been employed to create awareness and may even change attitudes, but compliance rates remain low without matching investments in bins, fleets, and reliable service (Azode and Ebeze, 2024; Okamgba and Nwafor, 2025). It has been demonstrated through optimization that engineering and statistical instruments (e.g., Response Surface Methodology) can be used to rationalize operations; however, continued improvement has to be based on continuous quality control and policy adherence (Okeke et al., 2024). These themes are reflected in national reviews: large per-capita generation, organic-heavy streams, open burning, and the use of poorly managed landfills make the air and groundwater hazardous, and the performance of the private sector strongly depends on the income level of neighbourhoods and the accessibility of roads (Okeke and Nwaiwu, 2025).

Hence, the literature can be summarized to have three implications to estimation and planning: (1) combine multi-source data (demographics, spatial layers, behavior, and composition) with advanced learners like XGBoost to learn nonlinearity and heterogeneity; (2) embed spatial-temporal structure, GIS to be used in siting and routing, and calendar-sensitive forecasts of seasonal/time-series features (Hoque and Rahman, 2020); and (3) couple predictive modelling with system reforms, source separation, circular This evidence is expected to support a machine-learning-based, data-driven technique of ward-level waste prediction and infrastructure control in Anambra/Onitsha sensitive to the local composition, seasonal, and service constraints.

1.1 Conceptual Framework

To have an efficient and sustainable waste management in a city, particularly in Onitsha, where population is rapidly increasing, the determinants of solid waste generation are necessary. This framework combines socio-economic, demographic, and environmental factors, including population density, business density, market count, the proportion of organic composition, and waste per capita, into a predictive model using the approach of Machine Learning (XGBoost) to estimate the waste loads, optimize zoning of services, and development of the infrastructure towards a better management result.

The model establishes that the population size, density and business endeavors are the main causes of solid waste generation and other variables such as organic content and per capita waste are a composite and behavioural dynamic. These predictors are inputted into the XGBoost algorithm which learns nonlinear interactions, and predicts quantities of waste at the ward level. The output of this model provides insight into the analysis of clustering and optimization to allocate resources efficiently and locate transfer stations as well as to plan operations in a cost-efficient manner (Chen and Guestrin, 2016; Okeke et al., 2024). The policy decisions based on machine learning predictions and the dedication to data-driven evidence is conceptually guaranteed to design adaptive, resilient waste systems in Onitsha (Ezeudu et al., 2021; Okeke and Nwaiwu, 2025).

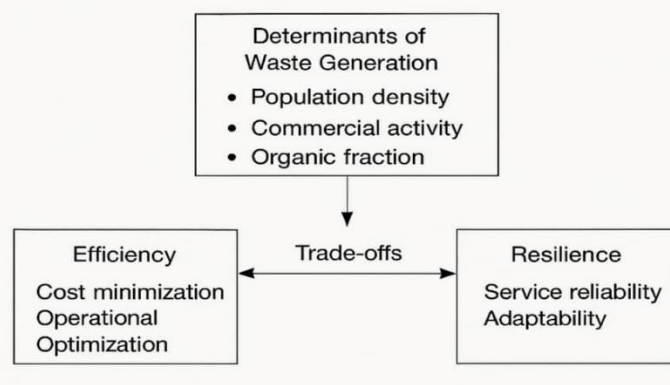


Figure 1: Conceptual Framework for the prediction of determinants of waste generation in Onitsha

Figure 1 defines the conceptual framework, which shows the determinants of waste production, and their influencing effects on solid waste management. Essentially, this framework shows that the combination of demographic, economic, and environmental determinants using sophisticated analytical models is more accurate in the forecasting of waste and the design of the system. By offering a predictive modeling as

a viable solution to practical decisions, it offers a guided channel through which policymakers can enhance the efficiency of waste services by minimizing operational expenses and promoting sustainable environmental regulation. The framework therefore links scientific modeling to real world infrastructure planning thus supporting data-based solution plans on circular economy pre-emption and sustainable

urban development in Anambra State and beyond..

2 MATERIAL AND METHODS

2.1 Research Design

The research design used in this study was quantitative and computational research design which combined data collection in the field and machine learning modelling to determine solid waste produced in 20 wards in Onitsha, Anambra State, Nigeria. The cross-sectional method was used and enabled the data to be gathered at one time point under a wide range of residential, commercial, and industrial environments. This design offered a record of mobility and spatial variability of waste production in the city. The methodology put an emphasis on the predictive powers, interpretability, and generalizability of expression of the developed models. The exploratory variables were considered to be the population density, commercial activity, and waste disposal practices, which have been repeatedly found as important factors influencing the generation of municipal solid waste (MSW) in urban developing environments. The aim of the study was to identify the strongest estimator of the waste loads at the ward level, as well as to be transparent, reproducible and policy-relevant, through a systematic comparison of statistical and machine learning methods, specifically Extreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016).

2.2 Study Area



Figure 2: Map of Onitsha, Anambra State (Source: <https://www.britannica.com/place/Onitsha-Nigeria>)

2.3 Population of the Study

The target group used in the study included the residents, business owners, and industrial operators within Onitsha who make direct contributions towards solid waste production. The population was stratified into three broad groups namely the residential, commercial and industrial categories that represented the full range of the sources of waste. They included households, market traders, small and medium-sized enterprises (SMEs), and manufacturers as they express diverse production and consumption patterns. Demographic characteristics, including the number of people in the household, the income level and occupation were taken into account as they are highly correlated with the waste generation behavior. Temporal variations in the generation of waste were included by the seasonal change in population density, especially during the season when the market activities are at the height. The study used a wide range of economic and demographic layers in that way the data were collected were representative, dynamic, and would be applicable to other fast urbanizing cities in Nigeria.

2.4 Sampling Techniques

Stratified random sampling method was used in order to ensure that the residential, commercial and industrial segments were equally represented. Random sampling in each stratum was done based on the random-number generator to discard selection bias. A proportionate sampling of 20 wards was done based on the population size of each area and intensity of generation of waste. Purposive sampling was further used to areas with high wastes like major markets and industrial areas where a lot of packaging and hazardous wastes are usually generated. The integration of stratified and purposive methods provided the coverage as well as capturing of outliers who could be some of the extreme cases of waste generation (Okeke et al., 2024). Such a two-pronged approach increased the strength and representativeness of the data to be used in

The study was carried out in Onitsha, a commercial centre that is densely populated in the state of Anambra, southeastern Nigeria. It is a city of a high trading level, a fast-populating city, and complicated issues of waste management (Igwe and Evan, 2020; Okeke and Nwaiwu, 2025). The waste management system at Onitsha experiences chronic ineffectiveness due to inconsistent collection times, inefficient disposal plant, and the strain of urbanization. The nature of the waste production and the rate is affected by the tropical climate of the city, which is characterized by wet and dry seasons. The spatial heterogeneity was taken into consideration and the city was divided into wards and sub-zones which signified the residential, commercial and industrial sectors. The markets, residential zones, industrial clusters, and public areas were sampled in order to provide data that would be representative and reflect the sectoral differences in waste composition. These stratifications are a mirror of the heterogeneity of wastes in a densely populated economically active setting like Onitsha.

Figure 2 shows the geographical position of Onitsha as part of southeastern Nigeria that is strategically located along the Niger River. This is where Onitsha is located and it is a key commercial and transportation center bridging the southern and the northern trade routes. Its high population and rapid urbanization, as well as closeness to the river, are largely responsible in waste generation and waste management issues, which highlights the significance of the city to the development of the environment as the target of spatial planning and the optimization of infrastructure.

model calibration and validation.

2.5 Data Collection Methods

2.5.1 Equipment and Materials

Field data collection was conducted through the use of scales, calibrated waste bins and data loggers which were used to measure and record data on daily waste quantities. Waste collection vehicles with GPS monitored collection routes, and digital data-capture devices (Smartphones and tablets) with tailored software were used in real-time recording and synchronization of data to a central database. This minimized the number of transcription mistakes and improved the accuracy of data (Creswell and Creswell, 2018). Having properly labeled sorting bins and sieves as a way of compositional analysis and sorting the waste into the following categories; organic, recyclable, non-recyclable, and hazardous. Samples were sorted and then weighed to ensure that there was no decomposition or contamination. The field staff followed the safety measures by wearing masks, gloves, and safety equipment on collection. Those approaches offered a systematic, technologically advanced, and verifiable way of measuring the solid waste production, which guarantees the credibility of the results later to be analyzed by machine learning and regression modelling. Digital tracking, on-site weighing, and sectoral disaggregation rendered a strong empirical basis to the development of the predictive models that were applied in this study.

2.6 Method of Data Analysis

In this section, there is the presentation of the Extreme Gradient Boosting (XGBoost) approach to assist in the proper estimation of the solid waste production within the wards in Onitsha. XGBoost has been found to be efficient in its computation, resilient against missing data, and capable of modeling complex nonlinear interactions between many predictors. It has a sequential boosting architecture that reduces the error of prediction, and

regulates overfitting in terms of regularization. These are the key benefits of XGBoost that can be used to model the heterogeneous waste patterns in fast-urbanizing areas.

2.6.1 The Extreme Gradient Boosting Method

XGBoost technique was chosen in this paper because it is effective in nonlinear relations, missing information, and multi-variable interactions. It builds the decision trees sequentially with the trees successively improving the residuals of the previous one, minimizing the loss function and shrinkage and penalization terms to control overfitting (Friedman, 2001; Dey et al., 2021).

The general XGBoost predictive function is expressed as:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

where:

\hat{y}_i = predicted solid waste for ward i (in kg/day);

x_i = vector of input features (population density, population size, commercial index, etc.);

f_k = individual regression tree;

K = number of trees;

F = space of regression trees.

The objective function minimized by XGBoost as expressed in equation (1) is:

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

where:

$l(y_i, \hat{y}_i)$ = loss function (mean squared error);

$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2$ = regularization term;

T = number of leaves in tree k ;

w = leaf weights;

γ and λ = complexity and regularization parameters, respectively (Chen and Guestrin, 2016).

The dependent variable was daily waste generation (kg/day), while independent variables included:

- Population size (Pop)
- Population density (Dens)
- Commercial index (CI)
- Number of markets (Mkt)
- Percent organic content (%Org)
- Per capita waste (PCW)

2.6.2 Model Specification and Regression Analysis

For comparison, a Multiple Linear Regression (MLR) model was specified as:

$$Y_i = \beta_0 + \beta_1 Pop_i + \beta_2 Dens_i + \beta_3 CI_i + \beta_4 Mkt_i + \beta_5 Org_i + \beta_6 PCW_i + \epsilon_i \quad (3)$$

where:

Y_i = daily waste generated (kg/day) in ward i ;

β_0 = intercept;

β_1, \dots, β_6 = regression coefficients;

ϵ_i = random error term.

This model estimated marginal effects of each predictor while XGBoost captured nonlinear interactions. Regression residual diagnostics included Durbin-Watson test for autocorrelation and Variance Inflation Factor (VIF) for multicollinearity (Gujarati and Porter, 2020).

2.6.3 Model Evaluation and Validation

Model validation followed 10-fold cross-validation, where data were partitioned into training (80%) and test (20%) sets. The model's predictive accuracy was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2), expressed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

Lower RMSE/MAE and higher R^2 values indicated stronger predictive performance (Hyland-Wood et al., 2023).

In addition, variable importance was assessed through SHapley Additive exPlanations (SHAP) values:

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f_{S \cup \{j\}}(x_S) - f_S(x_S)] \quad (7)$$

where:

ϕ_j = SHAP value of feature j ;

F = full feature set;

S = subset of features excluding j ;

$f_S(x_S)$ = model trained on subset S .

SHAP analysis using equation (7) quantified each variable's contribution to predictions (Lundberg and Lee, 2017).

2.6.4 Clustering and Optimization Analysis

To zone waste management services, K-Means clustering was applied using ward-level standardized indicators. The optimal number of clusters (K) was determined using the Elbow Method and Silhouette Coefficient:

$$S = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (8)$$

where:

a_i = mean intra-cluster distance for ward i ;

b_i = nearest-cluster distance for ward i .

Cluster outputs informed zoning of collection routes, siting of transfer stations, and landfill allocation. Cost optimization was modelled using equation (9):

$$\text{Minimize } C = \sum_{i=1}^n (c_t x_t + c_m x_m + c_l x_l) \quad (9)$$

where:

C = total system cost;

c_t, c_m, c_l = costs of transport, manpower, and landfill operations;

x_t, x_m, x_l = respective operational decision variables (Couto et al., 2021).

The methodology combined field data, advanced machine learning, and optimization analytics to estimate and design waste management systems in Onitsha. By integrating socio-economic, demographic, and commercial variables within the XGBoost framework, the study produced a robust, interpretable, and policy-relevant estimation model.

3. RESULTS AND DISCUSSION

Results integrate descriptive profiling, machine-learning prediction, and clustering to describe the generation of waste at the ward level in Onitsha. Descriptive indicate great variability among 20 wards, which is in line with dense, commercial urban form. XGBoost has almost perfect fit, and population size, density, and commercial intensity are the most important in SHAP. These drivers are supported by regression. The unsupervised analysis reveals three consistent service areas, which informs focused collection, facility location, and costed transfer-station alternatives to scalable, resilient operations.

3.1 Result of the Analysis

This section presents a descriptive and exploratory evaluation of wastes production and its socio-economic predictors within the twenty wards in Onitsha metropolis. The analysis incorporates demographic, commercial, and compositional measures to reveal how spatial differences and

patterns of solid waste production are affecting the results. Through a combination of a summary statistic, correlation analysis, and visualization, the study finds the key drivers of the daily generation of waste, including the population size, density, and commercial intensity. These understandings are the basis of the predictive and clustering analyses afterwards, which will be used to optimize the waste management system and resource distribution in the metropolis.

3.1.1 Descriptive and Exploratory Analysis of Waste Generation and Socio-Economic Drivers in Onitsha

Table 1: Descriptive Statistics of Daily Waste Generation and Socio-Economic Indicators across Wards in Onitsha

Variable	Value
Number of wards	20
Total daily waste (tonnes)	469.27
Mean daily waste (tonnes)	23.46
Standard deviation of daily waste (tonnes)	17.54
Mean per capita waste (kg/day)	0.76
Mean commercial index	0.57
Mean population density (per km ²)	7,445

Table 2 summary statistics indicate that the 20 wards in Onitsha produce approximately 469 tonnes of waste per day with an average of 23.5 tonnes per ward with a significant variation (SD = 17.5 tonnes), indicating that there is a lot of variation in the wards. The average per capita waste of 0.76 kg/day is consistent with the urban values of developing settings, in low- and middle-income nations, the daily per capita waste is between 0.45 kg and 1.1 kg on average (Karak et al., 2012). The average commercial index of 0.57 indicates a moderate level of commercial activity, whereas the average population density of about 7,445 people/km² highlights the high urban concentration in Onitsha. These trends suggest that population pressure and commercial intensity have a significant impact on spatial waste production, and ward-specific waste management plans are required.

3.1.2 Predictive Modelling of Ward-Level Waste Generation Using XGBoost and Regression Analysis

This section demonstrates how machine learning and statistical processes can predict the daily level of waste production in the ward of Onitsha. This is done using the XGBoost algorithm as well as the multiple regression model to evaluate the predictive performance, predictors that have the most significant impact, and compare the findings of the methodology analysis. The objectives still are to generate plausible data to be used in planning, infrastructure sizing, and evidenced-based decision-making to reach the desired sustainable waste management.

Table 1: Performance Metrics of the XGBoost Model for Predicting Daily Waste Generation in Onitsha

Metric	Value
RMSE	5,578.70
MAE	3,500.30
R ²	0.9984

The results of the model evaluation presented in Table 3 indicate that the XGBoost algorithm fits the data well to predict the daily waste generation in wards in Onitsha. Root mean square error (RMSE = 5,578.7) and mean absolute error (MAE = 3,500.3) are relatively low in comparison with the size of daily waste values (in tens of thousands of kilograms), which means that there is a small error in prediction. In addition, the coefficient of

determination ($R^2 = 0.9984$) indicates that the model accounts almost all the variation in the amount of waste generated in wards. These findings affirm the strength and predictive power of XGBoost, which makes it a viable instrument in predicting the level of waste at the ward and assisting in evidence-based planning of waste management infrastructure.

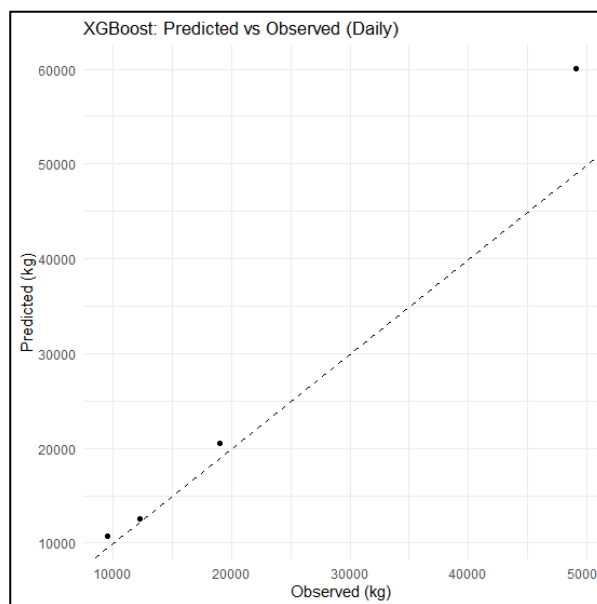
**Figure 3:** XGBoost Predicted vs. Observed Daily Waste Generation in Onitsha

Figure 3 is a scatterplot that compares the predicted and actual observed values of daily waste in wards using the XGBoost model. The majority of points are near the 45° reference line, which means that there is a high level of agreement between predictions and observations. However, there

are slight deviations in the upper range (more than 50,000 kg), which signifies a slight deviation of the wastes of wards with excessive daily results. However, the overall fit indicates that the model is quite useful in forecasting the underlying trend, which aligns with the high R^2 (0.9984) in

the performance metrics. The presence of similarities between the observed and predicted values is an indication that XGBoost is a valid

forecasting model when it comes to forecasting the quantity of waste that is produced in the ward in Onitsha. It means that a decision-maker can use the model to predict the conditions of future wastes and plan the

infrastructure (where to construct a transfer station, how big the landfill will be, or in what place to place a recycling hub). These small variations at the extreme upper tail underline the necessity to continue doing the calibration using new data, especially in the wards where there is high commercial activity such as Mainmarket and Ochanja where the alterations in waste production are more likely to occur.

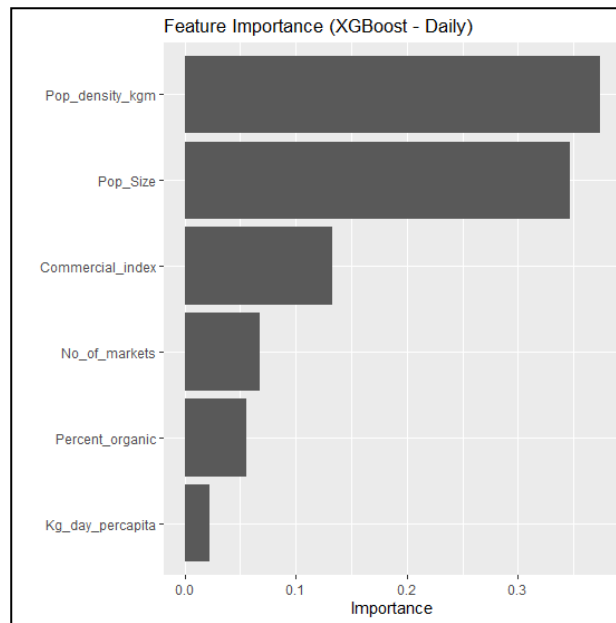


Figure 4: Feature Importance of Predictors in the XGBoost Model for Daily Waste Generation in Onitsha

The results of the feature importance in Figure 4 show that the two most significant predictors of daily waste generation are population density (Pop_density_kgm) and population size (Pop_Size), which have the largest weights in the XGBoost model. Commercial index also has a significant contribution, which is the influence of economic activity on the level of waste. Conversely, other variables like the number of markets, percent organic waste, and per capita waste generation have a relatively lower impact. This implies that although all the factors are relevant in waste

prediction, the demographic and spatial density measures are the most influential in the variation across wards. The results highlight that population-related pressures are the main factors that contribute to waste generation in Onitsha, and commercial activity is the secondary factor. This has powerful planning implications: the waste management plans must focus on the wards that are densely populated and have a high population, as well as commercial centers to handle the increase in the volume of waste.

Table 2: Feature Importance of Predictors in XGBoost Model for Daily Waste Generation

Feature	Gain	Cover	Frequency
Population density (Pop_density_kgm)	0.374	0.268	0.264
Population size (Pop_Size)	0.347	0.423	0.458
Commercial index	0.133	0.157	0.145
Number of markets (No_of_markets)	0.068	0.046	0.041
Percent organic (Percent_organic)	0.056	0.102	0.089
Per capita waste (Kg_day_percapita)	0.022	0.005	0.004

Table 4 results indicate that the two variables that have the greatest impact on the prediction of daily waste generation in Onitsha are population density (Gain = 0.374) and population size (Gain = 0.347). These two predictors prevail in all three measures of importance, gain,

cover, and frequency, and thus they are at the center of the waste volumes. The commercial index (Gain = 0.133) also shows significant impact, then the number of markets and percent organic content, and the least impact is made by per capita waste generation.

Table 3: Global SHAP Importance Values for Predictors

Feature	SHAP Value
Population size (Pop_Size)	4,514.43
Population density (Pop_density_kgm)	4,283.37
Commercial index	2,129.64
Number of markets (No_of_markets)	1,119.04
Percent organic (Percent_organic)	979.73
Per capita waste (Kg_day_percapita)	633.36

This conclusion is supported by the SHAP values in Table 5, which indicate that population size (4,514.43) and population density (4,283.37) have the highest contribution to the prediction variance, then commercial index (2,129.64). These findings indicate that the combination of population

concentration and economic activity is the main factor in waste generation, which is in line with the urban waste generation trends in developing cities.

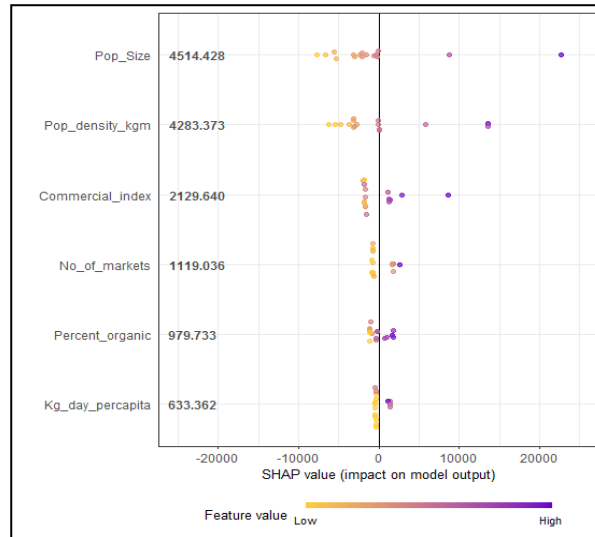


Figure 5: SHAP Summary Plot of Predictor Contributions to Daily Waste Generation in Onitsha

Figure 5 shows the SHAP summary plot that demonstrates the relative contribution of each predictor to the XGBoost model output of daily waste generation. The most influential are population size (SHAP = 4,514.43) and population density (SHAP = 4,283.37), where the higher the value, the higher the waste predictions. Commercial index (2,129.64) is also a significant contributor, as it indicates the contribution of economic activity to the increase in the volumes of waste. In the meantime, the markets (1,119.04) and percent organic waste (979.73) have moderate impacts, whereas the per capita waste generation (633.36) has the least impact. The color gradient shows that high values of features (purple) tend to be related to higher predicted waste, whereas low values (yellow) decrease predictions.

These findings highlight the fact that the most influential factors that determine waste production in Onitsha are demographic and economic factors, population concentration, population size, and commercial activity. This implies that the interventions of waste management must be strategically targeted at the densely populated and commercially active wards that produce disproportionate waste loads. The comparatively lesser role of per capita generation and organic fraction points to the fact that composition-based or household-level approaches might not be enough. Rather, sustainable solid waste management requires integrated policies that involve demographic planning, zoning of commercial hubs, and scaling of infrastructure.

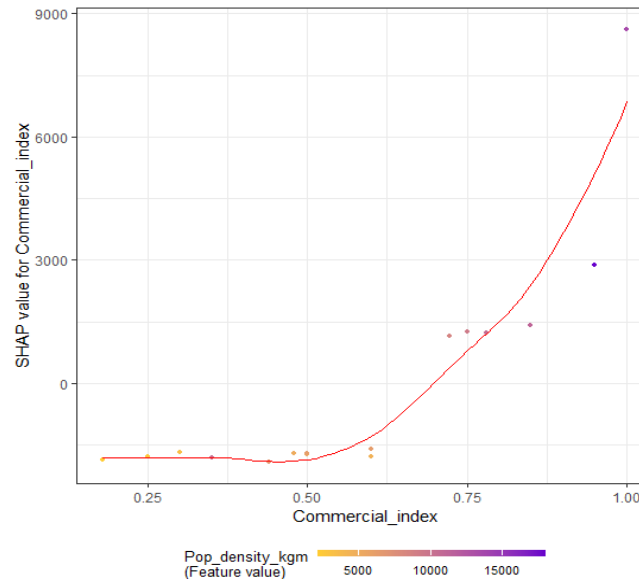


Figure 6: SHAP Dependence Plot for the Effect of Commercial Index on Daily Waste Generation in Onitsha

Figure 6 shows that the dependence plot of SHAP indicates that there is a strong nonlinear relationship between the commercial index and the daily waste generation. At low commercial index (less than 0.4) the SHAP effect is small or even negative, indicating a weak effect on waste volumes. Nevertheless, the influence of the commercial index on the forecasted waste output increases exponentially with the index, and the effect is even more pronounced above 0.75. The color gradient shows that this effect is even more pronounced in wards with higher population densities (purple dots), which demonstrates that dense, and highly commercialized areas are disproportionately increasing the amount of waste.

These results indicate that commercial activity is a critical source of waste production when it reaches a certain threshold level, particularly when it is coupled with high population density. It implies that the waste

management systems must focus on the distribution of resources to commercial centers (e.g., markets, trading centers) in highly populated wards. The policy interventions may involve special collection services in market areas, waste sorting and recycling centers in business streets, and dynamic scheduling of pickups based on the intensity of trading. The nonlinear impact of commercial activity can be identified by planners to predict the rapid increase in waste loads and plan proactively instead of reactively.

3.1.3 Result of Multiple Regression Analysis

The multiple linear regression model was fitted to examine the drivers of daily waste generation across wards in Onitsha. The model employed the demographic, commercial, and compositional predictors.

Table 4: Regression Coefficients for Predictors of Daily Waste Generation in Onitsha

Source of variation	Estimate	Std. Error	t-value	Pr(> t)	Significance
(Intercept)	2961.6	9885.54	0.3	0.769227	n.s.
Pop_Size	0.4	0.12	3.666	0.002848	**
Pop_density_kgm	1.1	0.24	4.502	0.000595	***
Commercial_index	1196.8	7619.94	0.157	0.877608	n.s.
No_of_markets	2736.4	548.79	4.986	0.000249	***
Percent_organic	-350.7	230.6	-1.521	0.152232	n.s.
Kg_day_percapita	17458.9	7192.69	2.427	0.030484	*

*Significance codes: *** $p < 0.001$; ** $p < 0.01$; $p < 0.05$; n.s. = not significant

Table 5: Model Fit Statistics

Statistic	Value
Residual Std. Error	1910
Multiple R-squared	0.9919
Adjusted R-squared	0.9881
F-statistic (df = 6, 13)	264.7
Overall model p-value	8.10E-13

Table 6 and Table 7 regression model show that the explanatory power ($R^2 = 0.992$; Adjusted $R^2 = 0.988$) is very high, which means that the selected predictors explain nearly all the variability in the daily waste generation. The small residual standard error also supports the strength of the model fit. The most significant predictors were population size ($p = 0.0028$) and population density ($p = 0.0006$) itself, which highlights the key role of the demographic pressure in waste volumes determination. Similarly, the markets ($p = 0.0002$) are extremely meaningful which demonstrates the importance of commercial centers in the generation of urban waste. The waste generation per capita ($p = 0.0305$) rate is also significant, and it implies the contributions of households to the waste load.

On the other hand, the commercial index and the per cent organic waste are not statistically significant meaning that there are no significant independent effects of the two in the presence of other predictors. It concludes that even though commercial intensity matters in models that are more generic (e.g., XGBoost, SHAP), it may be moderated, in linear regression, by population and market concentration. These findings uphold the duality of demographics as well as commercial activity and

suggest the subtle interplay of predictors.

The results obtained confirm that XGBoost is a highly accurate predictor ($R^2 = 0.9984$) and population density, population size and commercial activity are highly predictive. The recent standing of demographics and market concentration in waste outputs is also authenticated by regression models. All these strategies together form a solid evidence base of waste management planning in Onitsha. When combined with predictive accuracy and interpretability, the results may assist policymakers to predict future increases in wastes, design efficient service areas, and target the high-density and commercial wards to intervene.

3.1.4 Optimal Clustering Analysis for Waste Management Service Zoning in Onitsha

This section applied the clustering techniques to identify optimal service zones for waste management in Onitsha. Using the Elbow Method, Silhouette Analysis, and ward-level scatterplots, the study was able to determine the number of meaningful clusters based on waste generation patterns, socio-economic indicators, and demographic characteristics.

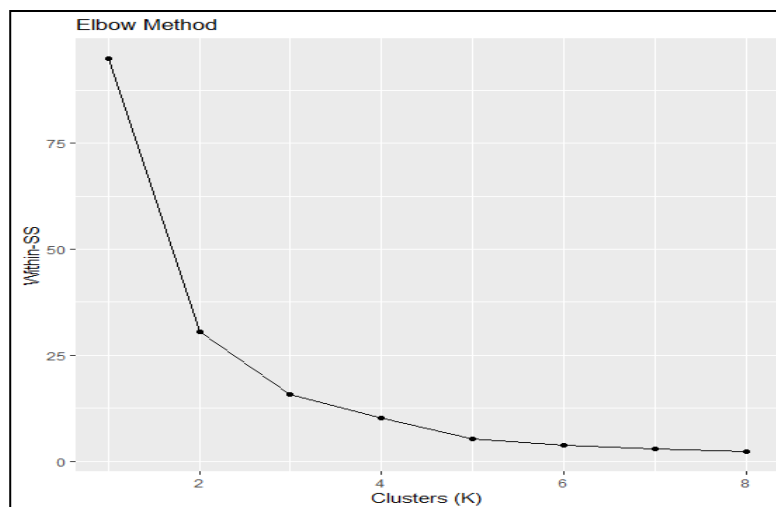
**Figure 7:** Elbow Method for Determining Optimal Clusters in Ward-Level Waste Generation Data

Figure 7 shows the elbow plot of the relationship between the number of clusters (K) and the within-cluster sum of squares (Within-SS). There is a sharp decrease in Within-SS between $K = 1$ and $K = 2$, and then the

decrease becomes smaller as K increases. The elbow is at $K = 3$, and further addition of clusters only brings slight gains in the reduction of within-cluster variation. This implies that there are three clusters that offer a

trade-off between the simplicity of the model and the explanatory power. The use of three clusters implies that the wards in Onitsha can be effectively divided into three different service areas according to the waste production patterns and socio-economic features. This has practical implications on waste management planning: the resources like collection

facilities, transfer stations, and transport routes can be optimized to these three zones, which will be more efficient and cost-effective. Selecting more than three clusters would probably complicate the system without providing significant benefits in representativeness.

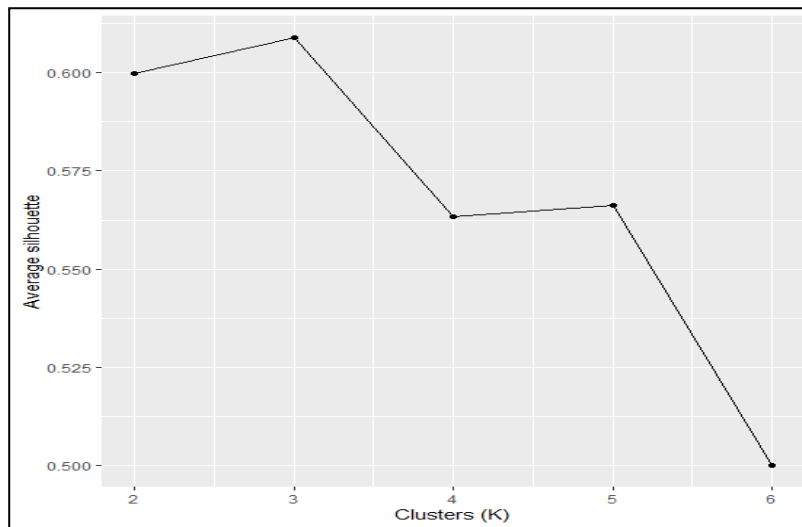


Figure 8: Silhouette Analysis for Optimal Cluster Validation in Ward-Level Waste Data

Figure 8 depicts the silhouette plot of the average silhouette width of the various cluster solutions ($K = 2$ to 6). The highest point is at $K = 3$, with an average silhouette width of more than 0.61, which means that three clusters give the most consistent and well-distinct grouping of wards. Cluster solutions of $K = 4$ or more give lower silhouette scores, indicating poor separation and less compact clusters. The result substantiates the elbow approach finding which held that three clusters are the optimal

forward-level waste production and socio-economic statistics. The three service areas are a compromise of interpretability versus statistical validity, where the wards all cluster together with high internal similarity, but are clearly different than other clusters. This provides an excellent platform of waste management planning, zoning and facility location in Onitsha to reduce logistical unproductiveness and expenses.



Figure 9: Ward-Level Clusters for Waste Management Service Zoning in Onitsha

Figure 9 presents a scatterplot of clustering at ward level to daily waste generation, population density, and population size. Three clusters are created with Cluster 1 (red) containing high-density, high-waste wards with a high number of patients; Cluster 2 (green) containing medium-density wards with moderate waste production; and Cluster 3 (blue) containing low-density wards with small populations and comparatively low waste output on a daily basis. The distance among the clusters shows apparent socio-demographic and spatial distributions of the waste production. These clusters provide viable base line in designing the areas of waste management services. The wards that generate the most waste are those in Cluster 1 and should be given the first priority by providing infrastructure such as transfer stations, high capacity collection fleets and frequent collection schedules. Cluster 2 wards could be served by medium-capacity facilities and Cluster 3 could be served by smaller capacity systems and flexible routing. This is the zoning plan that makes resource allocation as much as possible, cost of transport as least as possible, and

the solid waste management plans in Onitsha very cost-effective.

This is invariably found to give three clusters that provide the most suitable structure of waste zoning at the ward level and a simple and explanatory one. The intensive infrastructure is demanded in the high-waste, high-density ward and medium-scale systems can serve medium-density wards, and small, flexible facilities can serve the low-density wards. This information-based zoning system will give more detailed allocation of transfer stations, fleets, and collection schedules. The clustering method will enhance the effectiveness of the planning, the transportation costs will be reduced, and the general resiliency of the waste management system in Onitsha, as the statistical soundness and the operating needs will be balanced.

3.1.5 Cost Efficiency and Trade-Offs in Transfer Station Configurations

This section comprises of a comparative analysis of the annual costs in

Onitsha with different settings of transfer stations. The section will look at absolute expenditures and proportionate cost structure using detailed breakdowns of capital expenditure (CAPEX), operational expenditure

(OPEX), and transport costs. It will focus on how the number of facilities can be varied to impact the financial efficiency, long-term sustainability of the operations, and waste management.

Table 6: Annual Cost Breakdown for Waste Management Facility Scenarios							
Scenario	Total Cost (₦)	Transport Cost (₦)	OPEX (₦)	CAPEX (₦)	Transport Share (%)	OPEX Share (%)	CAPEX Share (%)
4 transfer stations	10.32 billion	4.14 billion	5.18 billion	1.00 billion	40.10%	50.20%	9.70%
6 transfer stations	10.82 billion	4.14 billion	5.18 billion	1.50 billion	38.30%	47.90%	

Table 8 results indicate that four transfer stations would cost about ₦10.32 billion per year, and six transfer stations would cost a little more ₦10.82 billion because of the increased CAPEX. Nonetheless, the increase in the number of stations to six decreases the share of transport costs (40.1 to 38.3) and OPEX share (50.2 to 47.9) and increases the capital investment burden (9.7 to 13.9). This implies a typical trade-off: the additional

facilities will decrease the dependence on transportation and enhance the efficiency of operations, yet will raise the initial expenditure on infrastructure. The decision between the two scenarios must be a balance between financial ability and the long-term requirement to decrease transportation dependency and enhance resilience in waste management logistics.

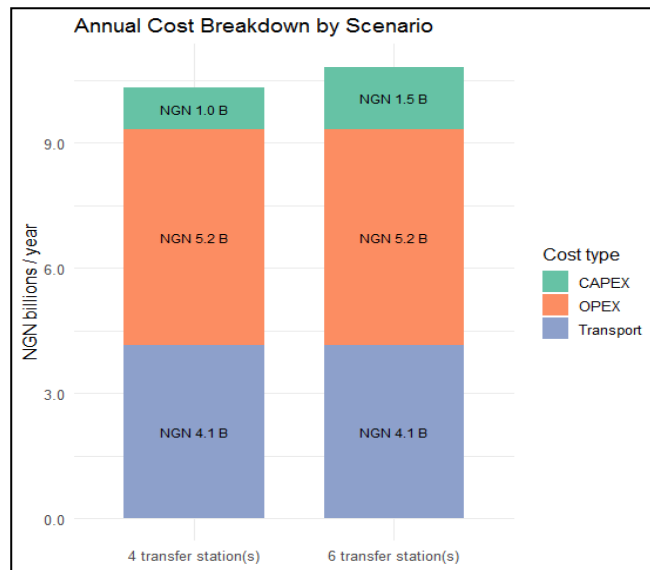


Figure 10: Annual Cost Breakdown for Waste Management under Alternative Facility Configurations

Figure 10 plot compares the annual cost of running four transfer stations and six transfer stations in Onitsha. In both cases, operational expenditure (OPEX) is the prevailing cost element (around ₦5.2 billion/year). The transport expenses do not change, and they are approximately 4.1 billion/year, whereas the capital expenditure (CAPEX) grows, as the system is extended to six stations, instead of four. Even though the overall annual expenditure increases slightly by 0.5 billion to 10.8 billion, the six-station system allocates the waste loads more effectively and may lead to

less strain on transport and operations in the long term despite the increased initial investments. The results indicate a strategic trade-off between fewer stations that reduce CAPEX but may overload operations and logistics, and more stations that raise costs slightly but enhance service reliability, redundancy, and resilience. The policymakers have to balance the fiscal limitations with the advantages of long-term efficiency, especially with the projected increase in waste of about 41 percent by 2035.

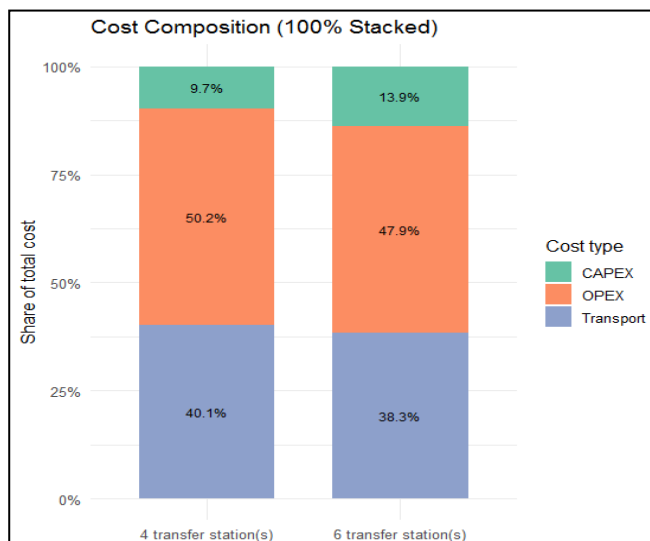


Figure 11: Cost Composition of Waste Management Scenarios: 4 vs. 6 Transfer Stations

Figure 11 shows the 100 percent stacked cost composition chart that indicates the relative proportions of Capital Expenditure (CAPEX), Operational Expenditure (OPEX), and transport costs in two scenarios of facility design. In the case of 4-station system, OPEX prevails at 50.2, transport (40.1) and CAPEX (9.7). The proportional burden of OPEX (47.9%) and transport (38.3%) decreases to 6-station system, and CAPEX increases to 13.9%. Therefore, the primary distinction between the two designs is that the cost balance between operating and logistics costs is shifted to capital investment. This observation underscores the fact that

more transfer stations would make the system more capital-intensive and operationally light, potentially improving the resilience of the system in the long term and reducing the risks of over-dependence on logistics. The policy makers are thus left with the choice of either focusing on short term affordability (low CAPEX with 4 stations) or long term sustainability and redundancy (high CAPEX with 6 stations). The cost structure also indicates that transport and OPEX are the primary sources of total expenditure and therefore are key areas of efficiency improvement irrespective of facility structure.

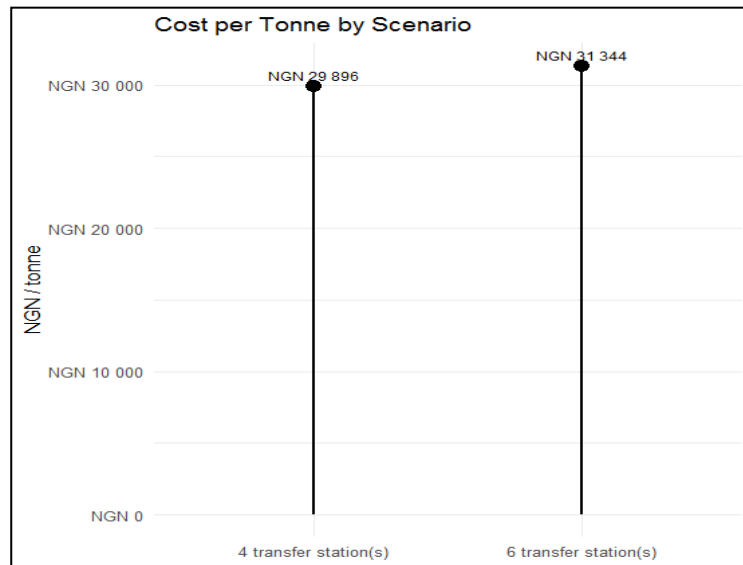


Figure 12: Cost per Tonne of Waste Managed under 4 vs. 6 Transfer Station Scenarios

Figure 12 is the outcome of the comparison of the unit cost of waste management per tonne in two facility configurations. The 4-station system has a slightly lower cost of 29,896/tonne and the 6-station system has a cost of 31,344/tonne. The difference is small, but it indicates that the expansion to six stations will lead to an increase in the per-tonne costs because of the increased capital expenditures with no corresponding decrease in OPEX or transport costs. The findings indicate that fewer

transfer stations (4) are more cost-efficient in terms of pure cost-per-tonne efficiency. Nevertheless, the 6-station design can offer operational resilience, redundancy, and better service equity across wards, which is worth the increased per-tonne cost in the long term. The policymakers thus have a trade-off between cost efficiency in the short term (₦/tonne) and system resilience and service coverage in the long term.

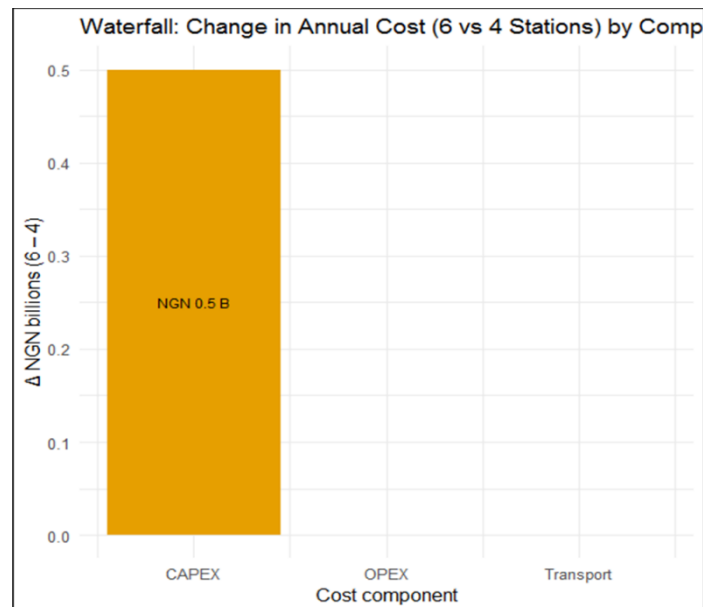


Figure 13: Waterfall Analysis of Cost Difference between 6 and 4 Transfer Stations

Figure 13 shows the waterfall chart that indicates the incremental cost difference between the 6-station and 4-station scenarios. The Capital Expenditure (CAPEX) is the only major contributor and it increases by about 0.5 billion every year when switching between 4 and 6 stations. There is no change in either Operational Expenditure (OPEX) or Transport Costs, indicating that the additional facilities are the main cause of increasing capital costs, but not significantly changing operating or logistics costs. This observation suggests that the increased total cost of the 6-station scenario is due to infrastructure investment and not operational or transport efficiencies. The decision-makers will have to

consider whether the incremental ₦0.5 billion in CAPEX is worth the possible long-term gains, including better service coverage, system redundancy, or fair access. In terms of short-term cost-control, 4 stations are more efficient, whereas in terms of resilience and equity, 6 stations can still be better despite the capital premium.

Overall, the results indicate that four transfer stations are cheaper in terms of annual costs and more efficient in terms of per-tonne, whereas six transfer stations are more redundant, equitable in terms of service, and flexible in terms of operations at the cost of increased capital

commitments. The decision between these two options, then, lies in the priorities of the policymakers to focus on short-term affordability or long-term system sustainability. Considering the estimated population and waste increase, it will be important to balance cost-effectiveness and sustainable infrastructure investment to guarantee effective waste management in Onitsha.

3.2 Discussion of Results

The results of the study indicate that the most important predictors of solid waste in Onitsha are population size, population density, and commercial activity. The XGBoost model also had an outstanding fit ($R^2 = 0.9984$), which was better than the Multiple Linear Regression ($R^2 = 0.992$), which highlights the benefit of machine learning in the nonlinear relationships between socio-economic variables. This is consistent with international and local literature that indicates the usefulness of ensemble learning models in predicting municipal solid waste, particularly in data-sparse settings (Ceylan et al., 2020). The prevalence of population-related factors substantiates the demographic-waste nexus that is reported in African urban settings, where the high rate of population growth and commercialization increases the amount of waste (Karak et al., 2012). The importance of the commercial index in the SHAP and feature importance plots suggests that the economic activity increases the rate of waste production when the commercial density reaches a certain threshold, which is in line with the results in Asian megacities (Afroz et al., 2011). Conversely, the smaller share of percent organic content and per capita waste indicates compositional stability among wards, which is supported by the findings that organic waste constitutes a stable majority in Nigerian waste streams (Chukwunneke et al., 2016; Nwakoby et al., 2020).

The regression analysis is a complement to the machine learning output, which supports the fact that high-density wards produce disproportionate waste loads. Nevertheless, the insignificance of the commercial index in the regression indicates the weakness of linear models in identifying the effects of interaction- additional justification of the use of XGBoost, which is more effective in identifying nonlinear interactions between demographic and economic variables.

The clustering analysis produced three optimum service regions with varying amounts of waste and demographics. This zoning is consistent with the stratified service model proposed, that further indicate that clustering allows improving logistical efficiency, reducing transportation costs, and aligning infrastructure with local needs (Okeke et al., 2024; Hoque and Rahman, 2020). The results also show that there are cost trade-offs between four and six designs of transfer stations: fewer stations save on capital expenditure, yet more stations enhance resilience, less reliance on transport and equity of services. It proves the statement that cost efficiency must be matched with long-term sustainability, as well as adaptive capacity (Ezeudu et al., 2021).

Thus, XGBoost, SHAP interpretability, and clustering analytics may be combined to provide an example of reproducible data-driven municipal planning. The findings to the policymakers imply that the waste management policies should be spatially diversified in which the areas with higher concentration of commercial wards should be prioritized in terms of intensive collection and infrastructural development. Furthermore, the insights point to the shift towards active, model-based planning with the assistance of machine learning tools rather than active, reactive waste collection that is also aligned with the request of deploying intelligent waste management in developing cities (UNEP, 2018).

4. CONCLUSION

This study was able to develop an integrated data-driven modelling system that incorporates Extreme Gradient Boosting (XGBoost), Multiple Linear Regression (MLR), SHapley Additive Explanations (SHAP), and K-Means clustering to approximate, explain and optimize solid waste production at the ward level in Onitsha. The results have shown that population density, population size, and commercial activity were the most significant predictors of daily waste production and explained nearly all the variance ($R^2 = 0.9984$). The clustering analysis has identified three optimal service areas according to varying demographic and waste generation factors, which provided an illustration of a spatially sound and cost-effective waste collection. The cost analysis has revealed that the four transfer-station design is the most economical in terms of capital investment, however, six transfer station design is more durable in the long run, more fair and flexible in their operations although it is a little more expensive. Those findings confirm that demographic concentration and business are the strongest predictors of waste generation, as well as that there is a need to apply data-driven, zone-specific management strategies.

The following recommendations were made based on the result of the study:

- To implement evidence-based policy, optimize collection routes, and design infrastructure based on ward-level waste intensities, the Anambra State Waste Management Authority (ASWAMA) should consider machine learning-based forecasting systems (e.g., XGBoost and SHAP).
- Waste service zoning should be based on the three-zone framework where high-density commercial wards are prioritized in terms of infrastructure and differentiated pricing mechanisms to promote fair cost recovery and sustainability.

Future studies ought to incorporate the time and behavioural aspects into the modelling framework by including seasonal changes, climate effects, and household behavioural data. The extension of this strategy to several cities in southeastern Nigeria would improve comparative knowledge and aid in harmonizing waste policy on a region-wide basis, based on data.

REFERENCES

- Abdallah, M., Alamri, S., and Alqahtani, A., 2020. Applications of artificial intelligence in solid waste management: A systematic review. *Environmental Science and Pollution Research*, 27(16), Pp. 19295–19312. <https://doi.org/10.1007/s11356-020-08390-5>
- Adeleke, O., Akinlabi, S. A., Jen, T. C., and Dunmade, I., 2022. Prediction of municipal solid waste generation: an investigation of the effect of clustering techniques and parameters on ANFIS model performance. *Environmental Technology*, 43(11), Pp. 1634–1647. https://ujcontent.uj.ac.za/view/pdfCoverPage?instCode=27UOJ_INS_TandfilePid=135790400007691anddownload=true#page=119
- Afroz, R., Hanaki, K., and Tudin, R., 2011. Factors Affecting Waste Generation: A Study in a Waste Management Program in Dhaka City, Bangladesh. *Environmental Monitoring and Assessment*, 179, Pp. 509–519. <https://doi.org/10.1007/s10661-010-1753-4>
- Anselin, L., 1995. Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27, Pp. 93–115. <http://dx.doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Azode, M. U., and Ebeze, U. V., 2024. Awareness and compliance of Anambra residents to media campaign on prevention of improper disposal of solid waste. *International Journal of Library and Information Technology*, 6, Pp. 1–15.
- Batagarawa, R. L., Saleh, D., Sani, J. E., Muazu, A. I., Ahmad, S., 2025. Effect of Seasons and Income Level on Solid Waste Generation Rates and Composition: A Case Study of Kaduna and Bauchi, Nigeria. *ATBU Journal of Science, Technology and Education*, [S.l.], v. 13, n. 1, p. 567–581, jun. 2025. ISSN 3093-0898. Available at: <<http://www.atbuftejoste.com.ng/index.php/joste/article/view/2428>>. Date accessed: 02 Nov. 2025.
- Ceylan, Z., Bulkan, S., and Elevli, S., 2020. Prediction of medical waste generation using SVR, GM (1, 1) and ARIMA models: a case study for megacity Istanbul. *Journal of Environmental Health Science and Engineering*, 18, Pp. 687–697.
- Chen, T., and Guestrin, C., 2016. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Pp. 785–794.
- Chukwunneke, J. L., Azaka, O. A., Chukwujike, I. C., and Sinebe, J. E. (2016). Optimization of waste management system in Anambra State: A case study of Ifite-Awka. *American Journal of Engineering Research*, 5(8), Pp. 138–151.
- Couto, L. C., Ferreira, J. A., and Gonçalves, G., 2021. Optimization of municipal solid waste collection using GIS and linear programming. *Sustainability*, 13(2), Pp. 489–503.
- Creswell, J.W. and Creswell, J.D., 2018. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage, Los Angeles.

- Egbu, A., and Okoroigwe, D., 2014. Informal Collection of Household Solid Waste in Three Towns of Anambra State, Nigeria. *PESD*, 8(2), Pp. 101–114. <https://doi.org/10.2478/pesd-2014-0028>
- Fadhullah, W., Imran, N. I. N., Ismail, S. N. S., Jaafar, M. H., and Abdullah, H., 2022. Household solid waste management practices and perceptions among residents in the East Coast of Malaysia. *BMC Public Health* 22, 1 (2022). <https://doi.org/10.1186/s12889-021-12274-7>
- Friedman, J. H., 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), Pp. 1189–1232.
- Getis, A. and Ord, J.K., 1992. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24, Pp. 189-206.
- Gujarati, D. N., and Porter, D. C., 2020. *Basic econometrics* (6th ed.). McGraw-Hill Education.
- Hoornweg, D., and Bhada-Tata, P., 2012. *What a Waste: A Global Review of Solid Waste Management*. World Bank Publications.
- Hoque, M. M. and Rahman, M. T. U., 2020. Landfill area estimation based on solid waste collection prediction using ANN model and final waste disposal options. *Journal of Cleaner Production*. <https://www.sciencedirect.com/science/article/pii/S0959652620304340>
- Igwe, P. U., and Anaje, C. E., 2022. Scavenging as a means of environmental management: A case study of Nsugbe solid waste dumpsite, Anambra State, Nigeria. *International Journal of Environment*, 11(1), Pp. 1–9. <https://doi.org/10.36632/ije/2022.11.1.1>
- Igwe, P. U., and Evan, E., 2020. An Assessment of Rural Solid Waste Management System in Aguata Local Government Area, Anambra State, Nigeria. *International Journal of Environment*, 9(1), Pp. 50–59. <https://doi.org/10.36632/ije/2020.9.1.4>
- Karak, T., Bhagat, R. M., and Bhattacharyya, P., 2012. Municipal solid waste generation, composition, and management: The world scenario. *Critical Reviews in Environmental Science and Technology*, 42(15), Pp. 1509–1630.
- Liu, B., Zhang, L., and Wang, Q., 2021. Demand gap analysis of municipal solid waste landfill in Beijing: Based on the municipal solid waste generation. *Waste Management*. <https://www.sciencedirect.com/science/article/pii/S0956756821000000>
- Lundberg, S. M., and Lee, S. I., 2017. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* (NIPS 30).
- Noiki, A. A., Afolalu, S. A., Yusuf, O. O., Emeteri, M. E., Ongbali, S. O., Oloyede, O. R., Joseph, O. O., and Banjo, S. O., 2021. Impact assessment of the current waste management practices in Nigeria. *IOP Conference Series: Materials Science and Engineering*, 1107, 012172. <https://doi.org/10.1088/1757-899X/1107/1/012172>
- Nwakoby, N. P., Okoye, N. J., and Chukwurah, D. C. J., 2020. Waste disposal and management in Awka Capital Territory, telescoping the roles of Anambra State Environmental Protection Agency (ANSEPA) 2012 - 2019. *International Journal of Academic Information Systems Research*, 4(12), Pp. 1–6.
- Okamgba, U. C., and Nwafor, G. U., 2025. Perception and attitude towards Anambra Broadcasting Service Facebook messages on solid waste management for sustainable development in Nigeria. *Journal of Environment, Climate, and Ecology*, 2(1), Pp. 29–39.
- Okeke, E. I., and Nwaiwu, N. E., 2025. Characterization of seasonally generated municipal solid wastes in Nnewi metropolis, Anambra State, Nigeria. *International Journal of Science and Research Archive*, 16(03), Pp. 423–430. <https://doi.org/10.30574/ijrsra.2025.16.3.2573>
- Okeke, S., Kharlamova, M., Didi, C., Nwobi, N., Oladipupo, O., and Osuagwu, F., 2024. Efficiency analysis of municipal solid wastes (MSW) and their environmental influences on groundwater resources in Onitsha, Anambra State, South-Eastern Nigeria. *Proceedings of the 9th North American Conference on Industrial Engineering and Operations Management*, Washington D.C., USA, June 4-6, 2024. IEOM Society International. <https://doi.org/10.46254/NA09.20240234>
- Shaamala, A., Yigitcanlar, T., Nili A., Nyandega, D., 2025. Machine learning applications for urban geospatial analysis: A review of urban and environmental studies. *Cities*, Volume 165,106139.
- UNEP., 2018. *Africa waste management outlook*. Nairobi: United Nations Environment Programme. Available at: <https://www.unep.org/ietc/resources/publication/africa-waste-management-outlook>

