

SHORT-TERM FORECASTING OF HOUSEHOLD WASTE GENERATION IN ALGIERS: STATISTICAL AND DEEP LEARNING APPROACHES

*** Nadjib BRAHIMI:**

PhD -Ecole Nationale Supérieure de Statistique et d'Economie Appliquée (ENSSEA)

Email : brahimi.nadjiballah@enssea.edu.dz

Brahim TAOUSSI:

Professor – Ecole Supérieure de Commerce (ESC), Koléa, Algeria

Email : b_tauoussi@esc-alger.dz

Reçu le: 08/10/2025 Accepté le : 19/11/2025 Publication en ligne le: 09/12/2025

ABSTRACT:

Algiers, the largest city in Algeria, is home to over 4 million inhabitants and serves as the country's vibrant economic hub. As one of the most densely populated cities, it faces significant challenges in managing household waste efficiently. Accurate forecasting of household waste generation is essential to address this issue, as it facilitates better resource allocation, optimizes waste collection schedules, and ensures adequate infrastructure to handle increasing waste volumes. Additionally, accurate forecasts inform decision-making for the planning and development of recycling facilities and landfills. They also quantify environmental risks, such as pollution and greenhouse gas emissions, thereby promoting sustainable waste management practices. Statistical time series techniques, such as Seasonal Autoregressive Integrated Moving Averages (SARIMA), are well-suited for short-term predictions due to their simplicity, interpretability, and minimal data requirements. However, these methods may fail to capture complex nonlinear dependencies in waste generation patterns. In contrast, deep learning models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks excel at identifying intricate patterns in data. Despite their proficiency, RNN and LSTM models demand large datasets, careful hyperparameter tuning, and are less interpretable. In this study, we analyzed daily household waste weights recorded in Algiers between January 2022 and October 2024. Descriptive analysis revealed that waste generation is positively skewed, with a mean of 2112 tons and a standard deviation of 266 tons. Unusually high values were observed during the annual Eid al-Adha festivities, while Fridays consistently showed lower waste volumes compared to other weekdays. Using the 80/20 split for training and testing, we evaluated model performance with four goodness-of-fit metrics. Our findings demonstrated that the SARIMA model, with a seasonal component of seven days, outperformed the LSTM and RNN models in predicting daily household waste generation in Algiers. Future work could explore the integration of hybrid models that combine the strengths of statistical and deep learning techniques to improve accuracy and address the limitations of individual approaches.

* Nadjib BRAHIMI

Keywords: Household waste generation, Short-term forecasting, Seasonal ARIMA, Deep learning models, Algeria

I. INTRODUCTION

Ce The growing demand for electricity in Algeria, driven by industrial expansion and population growth, necessitates accurate and reliable forecasting models to ensure efficient energy management and distribution. Traditional methods of predicting electricity consumption have often fallen short in capturing the intricate patterns and variability inherent in energy usage data. With advancements in artificial intelligence and machine learning, particularly in Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, there is an unprecedented opportunity to enhance the accuracy of electricity consumption forecasts.

Effective waste management is a cornerstone of sustainable urban development, especially in rapidly growing cities. Algiers, the capital of Algeria, is no exception. With a population exceeding 4 million, the city generates substantial volumes of household waste daily, putting immense pressure on its waste management systems. Poorly managed waste leads to numerous social, economic, and environmental issues, including inefficient resource allocation, public health concerns, and environmental degradation. For cities like Algiers, forecasting household waste generation is critical to tackling these challenges. Accurate predictions enable authorities to optimize collection schedules, allocate resources efficiently, and plan infrastructure investments, such as recycling facilities and landfill sites. Beyond logistics, waste forecasting contributes to reducing environmental risks, such as pollution and greenhouse gas emissions, aligning with broader global goals for sustainable urban living.

Despite its importance, forecasting household waste is inherently complex due to the influence of various factors, including population density, socio-economic trends, cultural practices, and seasonal variations. In Algiers, waste generation exhibits unique patterns, such as increased waste volumes during religious festivities like Eid al-Adha and reduced levels on Fridays, reflecting cultural and social behaviors. Capturing these patterns requires robust and adaptive forecasting models capable of handling dynamic and nonlinear data.

Traditional statistical models, such as Seasonal Autoregressive Integrated Moving Averages (SARIMA), have been widely applied for short-term forecasting tasks. These models are favored for their simplicity, interpretability, and low data requirements, making them suitable for scenarios with limited computational resources. However, their linear structure often limits their ability to capture the complex dependencies and nonlinearities found in household waste data.

Recent advances in artificial intelligence have introduced deep learning models, such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, which are designed to handle sequential and time-dependent data. These models have shown remarkable success in various forecasting applications, including energy consumption, financial markets, and weather prediction. Their strength lies in their ability to uncover hidden patterns and intricate relationships in large datasets. However, these models require substantial amounts of data, demand computational power, and lack the interpretability that

statistical models provide. As a result, there is often a trade-off between the accuracy of deep learning methods and the simplicity and transparency of traditional statistical approaches.

This study examines the short-term forecasting of household waste generation in Algiers using both statistical and deep learning approaches. Using daily waste data collected from 2022 to 2024, we implemented and compared the performance of SARIMA, RNN, and LSTM models. Through extensive analysis, we evaluated each model's ability to capture the temporal and seasonal patterns in waste generation and assessed their strengths and limitations. The results demonstrated that SARIMA, with its seven-day seasonal component, outperformed the RNN and LSTM models in terms of forecasting accuracy. This highlights the suitability of SARIMA for short-term applications where interpretability and simplicity are critical.

Furthermore, our findings underscore the potential of hybrid modeling approaches that combine the strengths of statistical and deep learning methods. By integrating the interpretability of SARIMA with the pattern-recognition capabilities of RNNs and LSTMs, future models could achieve superior performance while addressing the limitations of individual methods.

This paper contributes to the field of waste management by providing a comprehensive evaluation of forecasting techniques for household waste generation in Algiers. It also lays the groundwork for future studies that seek to develop hybrid solutions capable of meeting the growing demands of urban waste management systems.

The structure of the paper includes Sections 2 discussing the literature review, whereas section 3 addresses the methodology of the utilized forecasting techniques. Section 4 describes the household waste data. Section 5 discusses the results. Finally, Section 6 provides a summary of the key findings.

II. LITERATURE REVIEW

Antanasijevic et al. [1] highlighted the challenge of municipal waste generation (MWG) data gaps in developing countries, which hinder effective waste management. Their study tested two Artificial Neural Network (ANN) models, Backpropagation and General Regression Neural Network using data from 26 European countries. The models employed sustainability indicators including gross domestic product, domestic material consumption, and resource productivity as inputs, with MWG as the output. The general regression neural network model outperformed backpropagation in accuracy, training time, and stability, particularly in data-scarce regions like Bulgaria and Serbia.

Sama Azadi et al. [2] compared ANN and Multiple Linear Regression (MLR) for predicting mean Seasonal Municipal Solid Waste Generation (SMSWG) in 20 cities in Fars Province, Iran. Using population, waste collection frequency, maximum seasonal temperature, and altitude as predictors, ANN significantly outperformed MLR, demonstrating superior accuracy in capturing non-linear waste patterns. The study concluded that ANN offers a cost-effective alternative for SMSWG prediction in regions with limited historical data.

Kannangara et al. [3] emphasized the need for accurate municipal solid waste (MSW) forecasts for efficient waste management. Their study developed predictive models using ANN and decision trees, incorporating socio-economic and demographic data from 220 municipalities in Ontario, Canada. ANN outperformed decision trees, explaining 72% of MSW variance. However, paper diversion models were less effective due to missing behavioral factors.

Johnson et al. [4] addressed short-term MSW forecasting challenges in urban areas, considering population activity, seasonal trends, and external factors. Using a Gradient Boosting Regression Model trained on New York City's sanitation data (2005–2011), the model achieved high accuracy ($R^2 > 0.88$) across 232 geographic sections and three waste streams (refuse, paper, and metal/glass/plastic). The study demonstrated the importance of high-resolution spatiotemporal data in predicting MSW fluctuations and optimizing sanitation operations.

Cubillos [5] explored deep learning for household waste forecasting, applying a multi-site Long Short-Term Memory (LSTM) model to weekly waste data from Herning, Denmark (2011–2018). Unlike individual household models, the multi-site LSTM approach improved forecasting accuracy by 28% and significantly outperformed ARIMA models, effectively capturing complex non-linear relationships in waste generation.

III. METHODS AND MATERIALS

III.1. SARIMA

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is a powerful time series forecasting technique that extends ARIMA by incorporating seasonality. SARIMA is represented as $(p, d, q) \times (P, D, Q, s)$, includes autoregressive (p), differencing (d), and moving average (q) components, alongside their seasonal counterparts (P, D, Q) for periodic fluctuations of length s. The autoregressive (AR) component captures relationships between past and present values, while the moving average (MA) component models dependencies between an observation and past forecast errors. The integration (I) ensures stationarity by differencing the data. Originally introduced by Box and Jenkins, the SARIMA model is formulated as [6], [7]

$$\Phi(B^s)\phi(B)(1-B)^d(1-B^s)^DW_t = \Theta(B^s)\theta(B)\varepsilon_t + \delta \quad (1)$$

The household waste at time t , denoted as W_t , is modeled using a Seasonal Autoregressive Integrated Moving Average (SARIMA) process. Equation (1) consists of the following key components [6], [8]:

- Trend differencing $(1-B)^d$ removes non-seasonal trends by differencing the data d times to ensure stationarity.

- Seasonal differencing $(1 - B^s)^D$ eliminates seasonal patterns by differencing at a seasonal lag s (e.g., $s = 7$ for weekly seasonality).
- Autoregressive terms: $\phi(B)$ models short-term dependencies in the time series, whereas $\Phi(B^s)$ captures seasonal dependencies over lag s .
- Moving Average terms: $\theta(B)$ accounts for short-term shock effects, and $\Theta(B^s)$ represents seasonal shock propagation over lag s .
- Error term: ε_t which represents white noise disturbances in the model.
- Intercept or bias: δ a constant term that adjusts the mean level of the series.

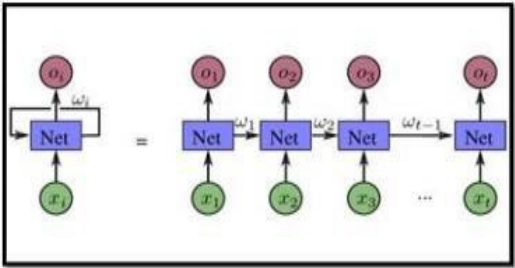
This model effectively captures both short-term variations and seasonal patterns in household waste data, ensuring accurate forecasting.

The Box-Jenkins methodology for SARIMA modeling consists of four key steps: model identification, parameter estimation, diagnostic checking, and short-term forecasting. Identification determines appropriate values of p , q , P , and Q using correlograms and Bayesian Information Criterion (BIC). Parameters are then estimated, followed by diagnostic checks to assess model fit. To prevent overfitting, BIC is used to select the most parsimonious model, balancing goodness-of-fit with model complexity. Forecast accuracy is evaluated using metrics such as Mean Absolute Error (MAE), Coefficient of Determination (R^2), and Root Mean Squared Error (RMSE) [9].

III.2. RECURRENT NEURAL NETWORKS (RNN)

Recurrent neural networks, are a category of DNN that make use of previous out-puts as inputs while having hidden states, being the only neural network provid-ing a memory from internal storage which is designed to memorize se- quences of data and enable the model to recall it and recognize patterns such as numeri-cal times series data .The Back propagation Through Time (BPTT) algorithm is used for the RNN training in order to calculate the gradi-ents and propagation of error and multiply it with weights in all training steps. The main weakness of the traditional RNN model when handling long term de- pendencies is the gradient vanishing and exploding issues. So a new model has been proposed to over-come RNNs failures which is the long short term memory (LSTM) model. [10].

Figure 01 RNN Networks.



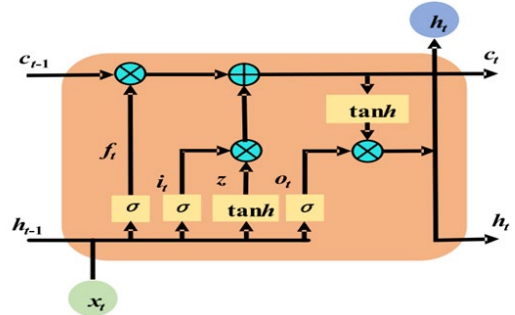
III.3. LSTM

In large datasets with complex, nonlinear relationships, traditional statistical models often struggle to capture intricate patterns, making deep learning techniques a more effective alternative. Long Short-Term Memory (LSTM), a specialized type of recurrent neural network (RNN), excels in processing sequential data where the order of observations is crucial. Widely used in natural language processing and time series forecasting, LSTM overcomes limitations of standard RNNs by efficiently capturing long-term dependencies while mitigating issues like vanishing gradients during training. Its architecture consists of four key components as indicated in Figure :

- (1) the input gate, which determines how much new information updates the memory cell.
- (2) the forget gate, which regulates the retention or removal of past information.
- (3) the output gate, which controls how much of the stored memory influences the current output.
- (4) the memory unit, which enables stable information transfer over long sequences.

By maintaining a structured flow of information, LSTM models can effectively learn and predict complex temporal patterns, making them a powerful tool for wind speed forecasting and other time-dependent applications [7], [8], [9], [11].

Figure 02: LSTM Networks.



III.3. PERFORMANCE PARAMETERS TO MEASURE FORECASTING ACCURACY

In the field of regression analysis, the evaluation of model performance is critical to understanding the accuracy and reliability of predictions. Various metrics are employed to assess how well a regression model fits the data. Among the most used metrics are Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). These metrics provide different perspectives on the performance of regression models, enabling us to evaluate and compare models effectively [2], [7], [8].

- RMSE measures the distance between the data points and the regression line, and it is represented as:

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

- MAPE represents the average absolute percentage difference between actual and predicted values., and it is denoted as:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (3)$$

- MAE represents the mean absolute difference between the actual and predicted values in the dataset, and it is denoted as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

- R^2 represents the proportion of variance in the dependent variable that can be explained by the independent variable

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

IV. DATA DESCRIPTION

These are the quantities of household waste and similar waste "household waste" generated per day by the 50 municipalities of the wilaya and Algiers during the period from January 1, 2022 to October 31, 2024 received at the level of CET HAMICI "technical office center of the wilaya of Algiers". In this part, we will try to take a preliminary look at the nature of Household Waste Data in the study sample, by using statistical methods to analyze the different parameters during the study period.

Table 01 : Summary statistics of measured household waste data.

Statistic	Household Waste
Count	1035.00
Mean	2111.89
Std	266.02
Min	1372.00
25%	1965.50
50%	2095.00
75%	2220.00
Max	4668.00

Table 1 states that the dataset contains 1035 observations representing daily household waste generation. The recorded waste values range between a minimum value of 1372 tonnes, and a maximum of 4668 tonnes. It exhibits a mean of 2111.89 tonnes, with a standard deviation of 266.02 tonnes. Before proceeding with further data analysis, we must ensure that the data is normalized to ensure suitability with RNN and LSTM. data normalization is typically performed using Min-Max scaling, which is given by the following equation [8], [9]:

$$W_N = \frac{W_t - W_{min}}{W_{max} - W_{min}} \tag{6}$$

Before normalization the range of household waste values range between 1372 to 4668 tonnes. Now after normalization we can observe that the data range is 0.0 - 1.0.

V. RESULTS AND DISCUSSION

To assess the forecasting accuracy of SARIMA(p, d, q) × (P, D, Q)[s], RNN, and LSTM models for household waste prediction, a structured evaluation process was followed. The dataset was split into training (80%) and testing (20%) subsets, with models trained on the former and tested on the latter. The training dataset comprised daily household waste measurements from January 1, 2022, to April 7, 2024, with total of 828 records. Predictions were then generated for the test period spanning April 8, 2024, to October 31, 2024, covering 207 records. Model performance was assessed based on forecasting precision, ensuring a robust comparison of statistical and deep learning approaches.

V.1. SARIMA

Time series analysis requires data to be stationary, meaning its statistical properties, such as mean and variance, remain constant over time. A visual inspection of household waste observations in Algiers, shown in Figure 1, suggests stationarity. To formally verify

this, we applied the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests [7]. As presented in Table 2, both ADF and PP tests yielded p-values below the 5% significance level, leading us to reject the null hypothesis of a unit root, confirming stationarity. Meanwhile, the KPSS test produced a p-value above 5%, indicating that we fail to reject its null hypothesis, further supporting stationarity in the data.

Figure 01 : Household waste data in Algiers between 2022 and 2024.

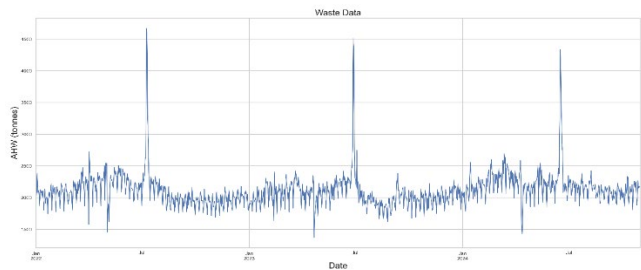


Table 02 : p-values of the ADF, PP, and KPSS tests.

Stationarity tests	ADF test	PP test	KPSS test
p-value	0.000	0.000	0.1

To analyze seasonality, we applied additive seasonal decomposition, revealing a clear weekly pattern. This method separates the household waste time series into trend, seasonality, and residual components, as shown in Figure 2.

Figure 2 Additive seasonality decomposition.

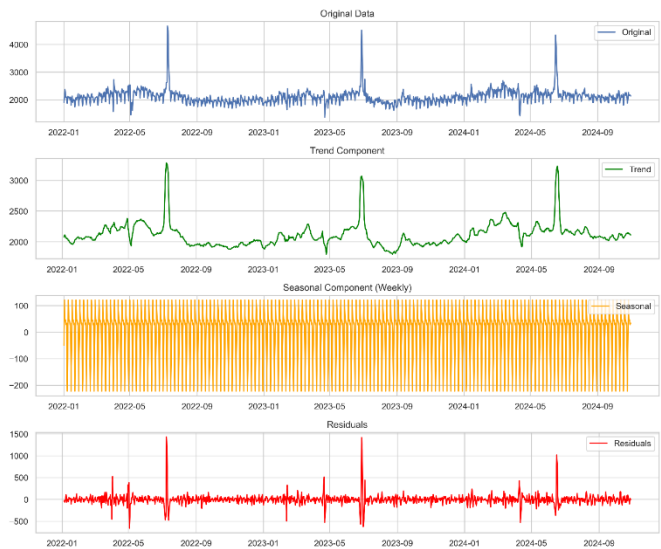
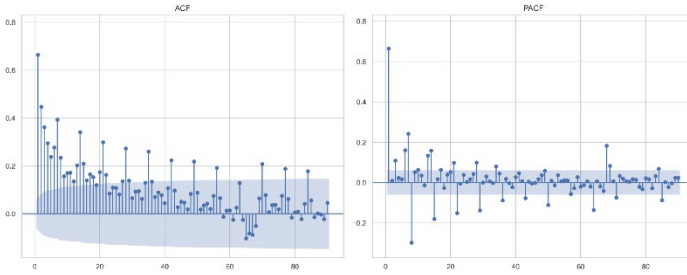


Figure 3 presents the ACF and PACF of household waste data. The ACF measures correlation between current and past values. It shows gradual declines rather than an abrupt drop, with noticeable spikes at lags 7, 14, 21, ..., indicating a weekly pattern. The PACF exhibits a sharp drop after lag 1, suggesting an AR(1) process where waste generation primarily depends on the previous day. A spike at lag 7 further reinforces weekly seasonality. The optimal SARIMA orders (p, q, P, Q) were selected by minimizing BIC with seasonality set to True and d=0. The SARIMA(1,0,0)(2,0,0)[7] model emerged as the best fit for describing household waste patterns in Algiers, as summarized in Table 3.

Table 03 : Optimal model characteristics.

Model	Log likelihood	AIC	BIC
SARIMA(1,0,0) (2,0,0)[7]	-6810.519	13631.038	13655.749

Figure 03 : Autocorrelation and Partial Autocorrelation plots.



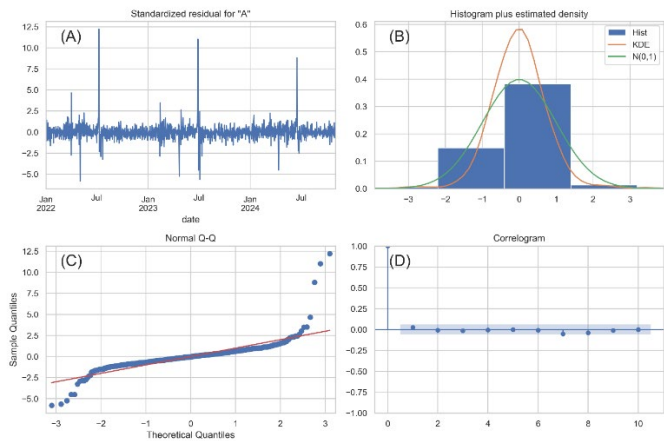
Equation (1) expresses the SARIMA(1,0,0)(2,0,0)[7] model. All coefficients were found significant at 5% threshold.

$$W_t = 2111.884_{(0.00)} + 0.6829 W_{t-1(0.00)} + 0.2974 W_{t-7(0.00)} + 0.274 W_{t-14(0.00)} + \varepsilon_t \quad (7)$$

The residuals analysis phase involves examining the characteristics of SARIMA residuals. A key criterion for an effective forecasting model is that residuals should exhibit properties resemble to white noise, including a mean close to zero, relatively constant variance, and minimal serial correlation.

Figure 4 provides a comprehensive diagnostic analysis of the SARIMA(1,0,0)(2,0,0)[7] residuals. Figure 4 (A) presents the residuals over time, assessing their stationarity and randomness. Figure 4 (B) displays the histogram and kernel density estimation (KDE) of standardized residuals, overlaid with a standardized Gaussian density curve, facilitating the evaluation of normality. Figure 4 (C) illustrates the quantile-quantile (Q-Q) plot, comparing the quantiles of observed residuals against a theoretical Gaussian distribution. Additionally, Figure 4 (D) shows the correlogram (ACF plot) of residuals to detect any remaining autocorrelation. The overall analysis suggests that the residuals exhibit desirable properties, including stationarity, a mean close to zero, an approximately Gaussian distribution, and minimal autocorrelation, indicating that the SARIMA model provides a reasonable fit to the data.

Figure 04 : SARIMA(1,0,0)(2,0,0)[7] diagnostic analysis.



After validating the SARIMA(1,0,0)(2,0,0)[7] model, we assessed its forecasting performance by predicting household waste from April 8, 2024, to October 31, 2024, covering a total of 207 records. Forecasting was conducted using two approaches: simultaneous forecasting and rolling forecasting. In the simultaneous approach, all 207 records were predicted at once, whereas in the rolling forecasting approach, predictions were generated sequentially, one step at a time. The results, illustrated in Figure 5, indicate that the rolling forecast method significantly improved accuracy. As shown in Figure 5 (A), rolling forecasting yielded predictions that were closer to actual values. Table 4 further supports this observation by summarizing the descriptive statistics of forecasting errors for both methods in Figure 5 (B). The rolling forecast approach exhibited a narrower error interval $[-233.52, 743.11]$, a lower mean error of 23.12, and reduced variability of 92.87, compared to the simultaneous approach, which had a wider error interval $[-716.53, 2233.49]$, a higher mean error of 65.15, and greater variability of 277.39. These results confirm that the rolling forecasting method enhances the predictive performance of the SARIMA model.

Figure 05 : (A) Simultaneous and Rolling forecast of household waste from April 8, 2024, to October 31, 2024, (B) Forecasting errors.

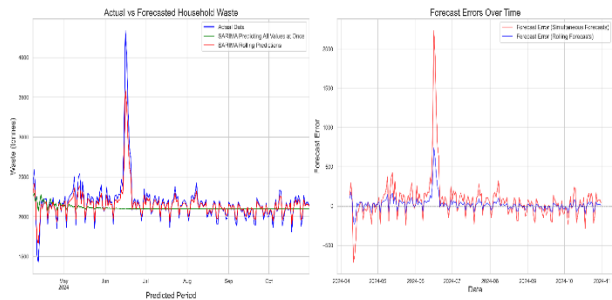


Table 04 : The descriptive statistics of the forecasting errors.

Forecasting error	min	Mean	Standard deviation	Median	Max
Simultaneous forecast	-716.53	65.15	277.39	55.78	2233.49
Rolling forecast	-233.52	23.12	92.87	18.24	743.11

The goodness-of-fit metrics for the SARIMA(1,0,0)(2,0,0)[7] model were evaluated using RMSE, MAE, MAPE, and R² as shown in Table 5. The rolling forecast approach demonstrated significantly better accuracy, with lower RMSE (95.48 tonnes) and MAE (52.69 tonnes) compared to the simultaneous forecast (RMSE: 284.28 tonnes, MAE: 150.42 tonnes). Additionally, the rolling forecast achieved a lower MAPE of 2.26% and a substantially higher R² of 88.13%, indicating a stronger predictive performance and better alignment with actual household waste data.

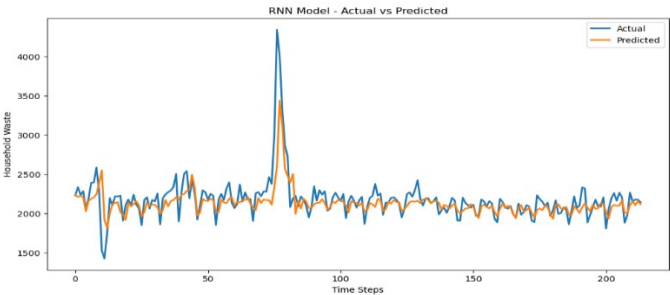
Table 05 : Goodness-of-fit metrics for simultaneous and rolling forecasting approaches.

Goodness-of-fit metrics	RMSE (tonnes)	MAE (tonnes)	MAPE (%)	R ² (%)
Simultaneous forecast	284.28	150.42	6.45	5.16
Rolling forecast	95.48	52.69	2.26	88.13

V.2 RNN MODEL RESULTS

The amount of training data, network structure, hyperparameters, and optimizers were all taken into consideration when designing the RNN model. The optimal number of epochs was established to be 200 with a batch size of 1, ReLU activation function. Utilizing performance statistics, the optimal RNN neural network structure was identified. For all operational RNN models, the mean square error (MSE) was employed as the loss function. In addition, the “Adam” optimizer was utilized to update the network weights in successive iterations based on the training data.

Figure 06 : The prediction results of the time series models on the actual values. RNN model.



RNN model's performance metrics are as follows:

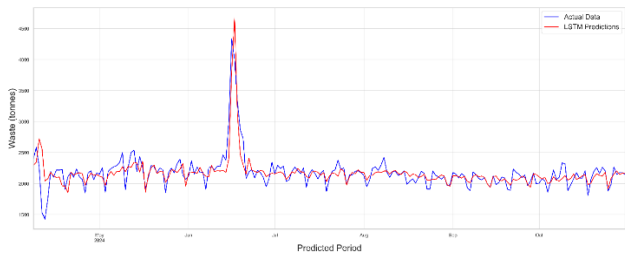
- Mean Absolute Error (MAE): 117.12
- Root Mean Squared Error (RMSE): 191.20
- R^2 Score: 0.5114 (indicating moderate explanatory power)
- Mean Absolute Percentage Error (MAPE): 19.39%

A MAPE of 19.39% suggests that, on average, the RNN model's predictions deviate by 19.39% from the actual values. Lower values are preferable, with <10% considered highly accurate and 10-20% as good accuracy in many practical scenarios.

V.3. LSTM

The LSTM model was implemented as a single-step forecasting tool with one hidden layer of 50 units. The model was trained over 200 epochs using the Adam optimizer and the ReLU activation function. A sequence length of 7-time steps was chosen to reflect seasonality, meaning the model used the previous 7 observations to predict the next step. The dataset was scaled using MinMax Scaler. The model structure consisted of an LSTM layer followed by a Dense output layer, and it was evaluated using Mean Squared Error (MSE) and Mean Absolute Error (MAE). The Keras implementation was developed within TensorFlow in Python 3.9. The forecasting results are presented in Figure 7. The LSTM model demonstrated a MAE of 102.19 tonnes and RMSE of 161.42 tonnes, indicating moderate prediction accuracy. The MAPE of 4.70% suggests a relatively low percentage deviation from actual values. Additionally, the R^2 value of 0.66 implies that the model captures 66% of the variance in household waste data, highlighting a reasonable but improvable fit.

Figure 07 : LSTM predictions of household waste from April 8, 2024, to October 31, 2024.



V.4. OVERALL COMPARISON

Figure 8 presents the household waste predictions from April 8, 2024, to October 31, 2024, using SARIMA, RNN, and LSTM models. The goodness-of-fit metrics, summarized in Table 6 and Figure 9, indicate that SARIMA outperforms the other models, achieving the lowest RMSE (95.48 tonnes), MAE (52.69 tonnes), and MAPE (2.26%), along with the highest R^2 (88.13%). The LSTM model demonstrates moderate performance, with RMSE, MAE, and MAPE increasing by 41%, 48%, and 52%, respectively, while R^2

decreases by 33% compared to SARIMA. The RNN model performs the worst, with RMSE and MAE more than doubling, MAPE deteriorating by 88%, and R^2 dropping by 72%. While deep learning models show some predictive capability, SARIMA coupled with Rolling forecast principle remains the most accurate and reliable approach for this dataset.

Figure 08 : SARIMA, RNN, and LSTM predictions of household waste from April 8, 2024, to October 31, 2024.

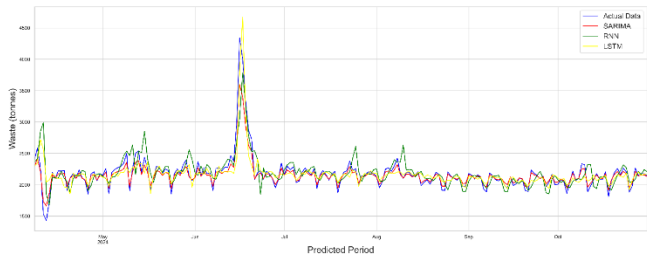
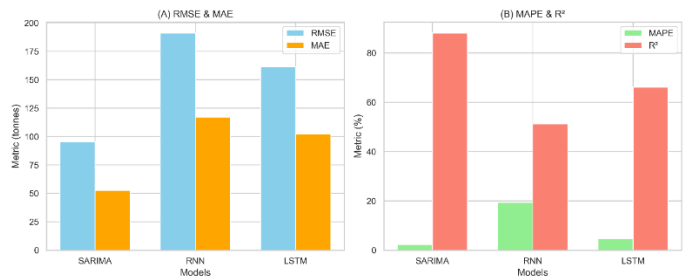


Table 06 : Household waste forecasting comparison across the SARIMA, RNN, and LSTM models.

Model	SARIMA	RNN	LSTM
RMSE (tonnes)	95.48	191.20	161.42
MAE (tonnes)	52.69	117.11	102.18
MAPE (%)	2.26	19.39	4.69
R^2 (%)	88.13	51.14	66.09
RMSE improvement (%)	0%	-50%	-41%
MAE improvement (%)	0%	-55%	-48%
MAPE improvement (%)	0%	-88%	-52%
R^2 improvement (%)	0%	72%	33%

Figure 09 : SARIMA, RNN, and LSTM Performance under different goodness-of-fit metrics.



VI. CONCLUSION

This study explored short-term forecasting of household waste generation in Algiers using both statistical and deep learning approaches. Daily waste data collected between 2022 and 2024 was analyzed to evaluate the performance of three models: SARIMA, RNN, and LSTM. Our analysis highlighted the unique seasonal and cultural patterns in waste generation, such as higher volumes during Eid al-Adha and consistently lower levels on Fridays.

The results demonstrated that the SARIMA model with rolling forecasts, with its seven-day seasonal component, outperformed both RNN and LSTM models in terms of forecasting accuracy. SARIMA’s simplicity, interpretability, and ability to capture periodic patterns made it particularly well-suited for this application. In contrast, the RNN and LSTM models, while capable of modeling complex nonlinear dependencies, were less effective due to the relatively limited size of the dataset and the need for extensive hyperparameter tuning. These findings underscore the continued relevance of traditional statistical models for applications where data is limited, patterns are periodic, and interpretability is essential. However, the limitations of SARIMA in capturing nonlinear trends highlight the potential benefits of hybrid approaches that integrate statistical and deep learning models. Combining the strengths of SARIMA’s interpretability with the pattern recognition capabilities of RNN and LSTM could enhance forecasting accuracy, particularly for complex or larger datasets.

In conclusion, this study provides valuable insights into waste forecasting methodologies and their applicability in urban contexts. Future research could focus on hybrid modeling approaches and incorporate external variables, such as weather conditions, economic factors, or policy changes, to further improve forecasting performance and support sustainable waste management practices in Algiers and similar urban environments.

VII. REFERENCES

- [1] D. Antanasijević, V. Pocajt, I. Popović, N. Redžić, and M. Ristić, “The forecasting of municipal waste generation using artificial neural networks and sustainability indicators,” *Sustain. Sci.*, vol. 8, no. 1, pp. 37–46, 2013, doi: 10.1007/s11625-012-0161-9.
- [2] S. Azadi and A. Karimi-Jashni, “Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of Fars province, Iran,” *Waste Manag.*, vol. 48, pp. 14–23, 2016, doi: 10.1016/j.wasman.2015.09.034.
- [3] M. Kannangara, R. Dua, L. Ahmadi, and F. Bensebaa, “Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches,” *Waste Manag.*, vol. 74, pp. 3–15, 2018, doi: 10.1016/j.wasman.2017.11.057.
- [4] N. E. Johnson et al., “Patterns of waste generation: A gradient boosting model for short-term waste prediction in New York City,” *Waste Manag.*, vol. 62, pp. 3–11, 2017, doi: 10.1016/j.wasman.2017.01.037.
- [5] M. Cubillos, “Multi-site household waste generation forecasting using a deep learning approach,” *Waste Manag.*, vol. 115, pp. 8–14, 2020, doi: 10.1016/j.wasman.2020.06.046.
- [6] D. N. Gujarati and B. Econometrics, *Basic Econometrics*. 2004. doi: 10.1126/science.1186874.
- [7] B. Taoussi, S. M. Boudia, and F. S. Mazouni, “Wind Speed Forecasting Using Univariate and Multivariate Time Series Models,” *Stoch. Environ. Res. Risk Assess.*, vol. 39, no. February 2025, pp. 547–579, 2024, doi: <https://doi.org/10.1007/s00477-024-02881-2>.
- [8] B. Taoussi, M. A. Damani, and S. Mohammed Boudia, “Seasonal ARIMA and LSTM Models for Wind Speed Prediction at El-Oued Region, Algeria,” *PAIS 2024 - Proc. 6th Int. Conf. Pattern Anal. Intell. Syst.*, pp. 1–8, 2024, doi: 10.1109/PAIS62114.2024.10541280.
- [9] M. PEIXEIRO, *Time Series Forecasting in Python. SHELTER ISLAND*, 2022.
- [10] K. Zouaidia, S. Ghanemi, M. Saber, and R. Lamine, “Hybrid intelligent framework for one-day ahead wind speed forecasting,” *Neural Comput. Appl.*, vol. 5, 2021, doi: 10.1007/s00521-021-06255-5.