Expert Systems

WILEY

Expert Systems

ORIGINAL ARTICLE OPEN ACCESS

CAPTURE—Computational Analysis and Predictive Techniques for Urban Resource Efficiency

Marzia Canzaniello[1] | Stefano Izzo[1] | Diletta Chiaro[1] | Antonella Longo[2] | Francesco Piccialli[1]

[1]Department of Mathematics and Applications "R. Caccioppoli", University of Naples Federico II, Naples, Italy | [2]Department of Innovation Engineering, Ecotekne Centre Pal. O – S.P. 6, University of Salento, Monteroni, Italy

Correspondence: Marzia Canzaniello (marzia.canzaniello@unina.it)

Received: 17 June 2024 | **Revised:** 25 September 2024 | **Accepted:** 11 October 2024

Funding: This work was supported by NextGenerationEU (CN_00000013) (CUP: E63C22000980007).

Keywords: data fusion | deep learning | forecasting | graph neural networks | urban sustainability | waste management

ABSTRACT

Municipal waste management (MWM) poses significant challenges in the context of rapid urbanisation and population growth. Accurate forecasting of waste production is crucial for designing sustainable waste management strategies. However, traditional forecasting methods often struggle to capture the complexities of waste generation dynamics. This paper proposes a novel methodology leveraging deep learning techniques to forecast municipal waste production. By harnessing the power of deep neural networks, our approach transcends the limitations of conventional models, providing more accurate and impactful predictions. We integrate heterogeneous data sources, including demographic and territorial information, into a comprehensive graph representation of municipalities. Graph Neural Networks are then employed to extract intricate spatial and temporal patterns from the graph structure. Empirical validation through a case study in the Apulia region demonstrates the effectiveness of our methodology in furnishing accurate forecasts for waste production. Our framework is adaptable and scalable, making it suitable for application across diverse geographical areas. This research contributes to advancing waste management practices by providing stakeholders with actionable insights for informed decision-making.

1 | Introduction

Municipal waste management (MWM) stands at the forefront of global challenges, particularly in the face of rapid urbanisation and increasing population growth. With more people living in cities, the sheer volume of waste generated presents a formidable task for municipalities worldwide. Addressing this challenge requires not only efficient waste collection and disposal but also proactive planning and forecasting to anticipate and meet the evolving demands of waste management.

Effective forecasting of waste production plays a pivotal role in this endeavour, serving as a cornerstone for designing and implementing sustainable waste management strategies. By accurately predicting future waste generation trends, authorities can optimise resource allocation, streamline logistical operations, and minimise environmental impact. Moreover, timely and precise forecasts empower decision-makers to proactively address emerging challenges, such as waste overflows, landfill shortages, and environmental pollution, thus safeguarding public health and ecological integrity.

However, traditional forecasting methods in MWM often fall short in capturing the intricacies of waste generation dynamics. Statistical models, while widely used, may struggle to accommodate the complex temporal patterns inherent in waste production data. Moreover, the inherent variability and sparsity of waste data pose additional challenges, potentially leading to suboptimal forecasts and decision-making.

Given these constraints, there arises an urgent demand for innovative methodologies that can mitigate the drawbacks associated with conventional forecasting techniques. This is where Deep Learning (DL) steps in—a cutting-edge approach that has transformed numerous domains by harnessing the capabilities of artificial neural networks to learn complex patterns and extract insights from large datasets. In the realm of waste management, the application of DL holds immense promise, offering the potential to unlock hidden patterns, identify non-linear relationships, and enhance the accuracy of waste production.

This study proposes a novel methodology for forecasting municipal waste production, harnessing the transformative capabilities of DL techniques. Deep neural networks offer a paradigm shift in forecasting methodologies, providing a robust framework to tackle the complexities inherent in waste generation dynamics. By exploring the complex patterns ingrained in historical waste production data, our methodology strives to surpass the constraints of conventional forecasting models and provide predictions that are not only more accurate but also more impactful. This enhanced precision in predictions translates into a more informed waste management strategy and tailors the economic burden associated with waste disposal efforts.

At the heart of our approach lies a meticulously crafted methodology that seamlessly integrates heterogeneous data sources, including demographic information sourced from the Italian National Institute of Statistics (ISTAT), income data from the Italian Ministry of Economy and Finance (MEF), and territory data from the Apulia Region. By synthesising these disparate data streams, we construct a comprehensive graph representation wherein Apulian municipalities serve as nodes, each endowed with rich feature sets (i.e., population demographics, spending habits, income data, and geographic characteristics). The edges of this graph are weighted based on similarity metrics derived from the fused data, capturing nuanced interrelations between municipalities.

Central to our approach is the application of Graph Neural Networks (GNNs), a cutting-edge methodology adept at processing graph-structured data. By exploiting the interconnectedness of municipalities encoded within the graph, GNNs enable the extraction of intricate spatial and temporal patterns that transcend traditional modelling paradigms.

In this paper, we present a comprehensive framework encompassing data fusion, graph representation, and GNN-based modelling for forecasting municipal waste production. While our case study focuses on the Apulia region, it is important to note that, given data availability, our framework is adaptable to any urban or regional context. Our methodology is meticulously designed to be adaptable and scalable, making it suitable for application across diverse geographical areas and accommodating varying levels of data availability. Through a detailed case study, we provide empirical evidence of the effectiveness of our approach in furnishing accurate forecasts for municipal waste production, thereby underscoring its potential to revolutionise waste management practices. By advancing the frontier of forecasting in MWM, our research endeavours to empower stakeholders with the insights and tools needed to navigate the complexities of urban waste management in the 21st century.

Overall, the main contributions of this paper can be listed as follows:

- A novel approach to forecasting municipal waste production leveraging advanced DL techniques.
- A comprehensive methodology that leverages heterogeneous data sources, including demographic information, to construct a graph representation of municipalities.
- A strong empirical validation of the proposed methodology through a real case study, wherein its effectiveness in providing accurate forecasts for municipal waste production is assessed.

This paper unfolds in: Section 2 covers related work; Section 3 details our methodology; Section 4 presents the implementation details, experimental results, and analysis. Finally, Section 5 draws conclusions.

2 | Background and Related Works

Over the last few decades, the realm of MWM has witnessed notable transformations. These advancements stem from heightened data accessibility and the adoption of sophisticated Machine Learning (ML) techniques. Designing a reliable forecasting method for predicting municipal waste production is crucial for several reasons, directly impacting both operational efficiency and environmental sustainability (Ferronato 2019). First, accurate waste forecasts enable local authorities to optimise resource allocation for waste collection and processing, thereby reducing operational costs and enhancing service efficiency. By anticipating future waste generation patterns, municipalities can tailor their waste management infrastructure investments—such as recycling facilities, waste-to-energy plants, and landfill sites—to meet upcoming needs without overinvestment or resource wastage (Christensen 2010). From an environmental policy perspective, understanding waste production trends is essential for developing effective waste reduction and recycling programs. Data-driven insights from waste forecasting can inform targeted initiatives, such as community recycling campaigns or waste reduction incentives, tailored to specific waste generation patterns identified through predictive analysis. In this section, we undertake a short review of pertinent research endeavours, with a primary emphasis on the development of waste generation forecast models. We trace their evolution from conventional statistical methods to the utilisation of advanced ML and DL approaches.

2.1 | Waste Generation Forecasting Models

The forecasting of waste generation is a critical task that has garnered increasing attention. The complexity of waste generation patterns, influenced by factors such as population growth, urbanisation, and socioeconomic changes, poses a challenge for traditional forecasting methods. In response to this, recent studies have begun to apply advanced ML techniques to improve the accuracy of waste generation forecasts.

For example, in Ahmed et al. (2022) the authors utilised DL models, including 1D Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM), to forecast the status of municipal waste in smart bins, showcasing the potential of these models in handling the spatiotemporal variability of waste data. Similarly, Srivastava and Jha proposed a multi-model approach that integrates demographic and socioeconomic factors with LSTM, Autoregressive Integrated Moving Average (ARIMA), and incremental increase models to enhance the predictive power of their waste generation forecasts (Srivastava and Jha 2023). Their work emphasises the importance of considering a wide range of factors that influence waste generation rates.

Moreover, predictive analyses using optimised neural network models have been explored to further refine the forecasting process. These studies contribute to a growing body of literature that seeks to apply cutting-edge ML techniques to the challenges of MWM (Kakad 2022).

In addition to the aforementioned studies, several recent works have investigated modern and smart technologies for efficient waste disposal and management (Mukherjee et al. 2021), advances in artificial intelligence applications in solid waste management (Ihsanullah et al. 2022), and the role of artificial intelligence in waste management in smart cities (Fang et al. 2023). Moreover, research focusing on modelling and prediction of regional municipal solid waste generation and diversion in Canada (Kannangara et al. 2018), patterns of waste generation in New York City (Johnson et al. 2017), and insights into regional differences in predictions of municipal solid waste generation rates using artificial neural networks (Wu et al. 2020) has contributed valuable insights to the field.

2.2 | Graph Neural Networks in MWM

GNNs have emerged as a powerful tool for analysing graph-structured data, offering versatile capabilities for modelling complex relationships and dependencies among interconnected entities. Recent advancements in waste management research have seen a growing interest in the application of GNNs to address various challenges in the domain. These sophisticated models present a promising avenue for enhancing forecasting accuracy and capturing nuanced patterns.

Recent research has demonstrated the effectiveness of GNNs in various waste management applications. For instance, the study by Hembert et al. (2024) explores the use of GNNs to assess sensor integrity for nuclear waste monitoring. By modelling sensor data as a graph and employing GNNs, the researchers were able to detect anomalies and ensure the reliability of monitoring systems effectively. Similarly, Mauliska et al. (2023) applied Spatial Temporal Graph Neural Networks (STGNNs) to forecast data time series of river pollution waste content in Probolinggo. By capturing spatial and temporal dependencies within the data through graph structures, STGNNs demonstrated promising results in accurately predicting pollution levels, thereby aiding in proactive pollution management strategies.

Furthermore, in their critical review, Gatti, Barbierato, and Pozzi (2024) explored the role of classic and machine-learning-based algorithms, including GNNs, for smart bin collection in greener smart cities. By analysing the effectiveness of different algorithms, including GNNs, in optimising waste collection routes and schedules, the study highlights the potential of GNNs in enhancing the efficiency and sustainability of waste management practices in urban environments.

In light of these developments, our study aims to contribute to the field by presenting a novel methodology that leverages DL for municipal waste production forecasting. Our approach is informed by the advancements in time series forecasting and tailored to address the unique challenges of MWM. By integrating state-of-the-art DL algorithms with advanced preprocessing techniques, we aspire to provide a robust and adaptable solution for municipalities worldwide.

3 | Methodology

Our methodology, as illustrated in Figure 1, adopts a comprehensive approach that integrates various strategies. It combines data fusion techniques with GNNs, along with temporal elements like LSTM and RNN layers, to forecast municipal waste generation. This section provides an overview of our methodology, outlining its key components. These include data preprocessing (Section 3.1), where time series data extracted from the Apulia region are clustered and decomposed, graph construction (Section 3.1.3), utilising data from other sources namely ISTAT, MEF, and Regional Technical Maps, and model architecture (Section 3.2), which merges temporal and graphical components to generate predictions. Further elaboration on each aspect follows, detailing the sequential steps involved in our predictive approach for waste production.

3.1 | Data Preprocessing

The first task in our data preprocessing involves the aggregation and integration of heterogeneous data sources relevant to municipal waste production forecasting. These sources encompass diverse datasets, including historical waste production records, demographic information sourced from national agencies, and geographical data.

Efforts are made to harmonise these disparate data streams, ensuring compatibility and consistency across datasets. Data integration facilitates a holistic understanding of the underlying factors influencing waste generation dynamics, thereby enhancing the predictive capacity of the subsequent modelling framework.

3.1.1 | Time Series Clustering

In this subsection, we explore the process of time series clustering, a pivotal stage in our methodology aimed at grouping municipalities based on the temporal patterns exhibited by their waste production data. To accomplish this, we employ the `TimeSeriesKMeans` algorithm (TSK-means) (Huang

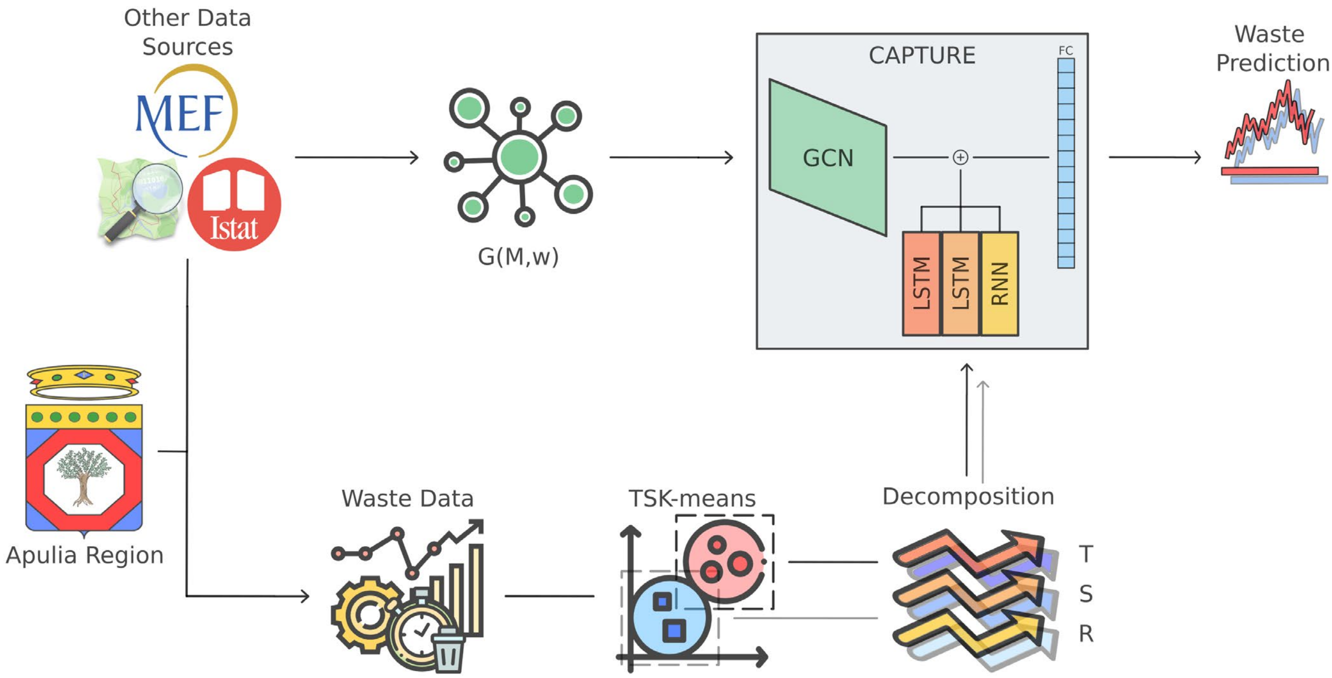

FIGURE 1 | The CAPTURE framework operates by aggregating waste volume data from municipalities across Apulia, pinpointed using geographical coordinates (latitude, longitude). These waste time series are then subjected to clustering, resulting in four distinct clusters. Within each cluster, the time series undergo decomposition into three components: trend, seasonal, and residual. Simultaneously, utilising additional data sources (ISTAT, MEF, and Technical Region Maps), a graph is constructed to capture similarities between municipalities. Subsequently, the CAPTURE model ingests both types of data. The model meticulously dissects the provided information, encompassing a graphical component based on GCN, and a temporal component relying on LSTM and RNN. In addition, it is important to emphasise that the model receives input data for each cluster, undergoing separate training and testing for each individual cluster. Finally, the model generates predictions for waste production by averaging the results obtained for each cluster.

et al. 2016), utilising Dynamic Time Warping (DTW) as the distance metric, to partition the municipalities into distinct clusters.

The TSK-means algorithm is a variant of the traditional `K-means` clustering algorithm adapted for time series data. It operates by iteratively assigning time series data points (representing waste production profiles for municipalities) to cluster centroids in a manner that minimises the DTW distance between the data points and the centroids. The DTW formulation (Sakoe 1971; Sakoe and Chiba 1978; Cuturi and Blondel 2017) can be written as follow:

$$\mathrm{DTW}(x, y) \coloneqq \min_{A \in A_{n,m}} \langle A, \Delta(x, y) \rangle \tag{1}$$

where, $x = (x_1, \ldots, x_n) \in \mathbb{R}^{p \times n}$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^{p \times m}$ are two time series of varying length $[n, m]$ taking values in $\Omega \subset \mathbb{R}^p$, $A_{n,m} \subset \{0,1\}^{n \times m}$ is the set of alignment matrices and $\Delta(x, y) \coloneqq \left[\delta(x_i, y_i)\right]_{ij} \in \mathbb{R}^{n \times m}$ is the cost matrix defined by a differentiable substitution-cost function $\delta: \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^+$.

The clustering process aims to group municipalities with similar waste production patterns into the same cluster while maximising dissimilarities between clusters. Before applying the algorithm, it is crucial to determine the optimal number of clusters for partitioning the municipalities. To achieve this, we utilise the elbow method (Kodinariya and Makwana 2013), a common technique in cluster analysis. The elbow method involves plotting the within-cluster sum of squared distances (WCSS) against different values of clusters number and selecting the point where the rate of decrease in WCSS begins to diminish, signifying an optimal balance between cluster compactness and separation.

Once the optimal number of clusters is determined, we proceed to apply the TSK-means algorithm to partition the municipalities into clusters based on their waste production time series data. The Principal Component Analysis (PCA) and TSK-means results are shown in Figure 2, while Figure 3 shows the geographical representation of the clusters obtained.

This approach allows us to discover hidden patterns and structures within the data, enabling more informed decision-making and resource allocation in MWM.

3.1.2 | Time Series Decomposition

In this subsection, we introduce the process of time series decomposition, a crucial step in our methodology aimed at disentangling the various components underlying waste production dynamics.

We employ the Seasonal and Trend decomposition method (STL) (Cleveland et al. 1990), a powerful technique for decomposing time series data into its seasonal, trend, and remainder components. This technique utilises locally weighted scatterplot smoothing (Loess) to iteratively estimate the seasonal and trend components while simultaneously detrending the series:

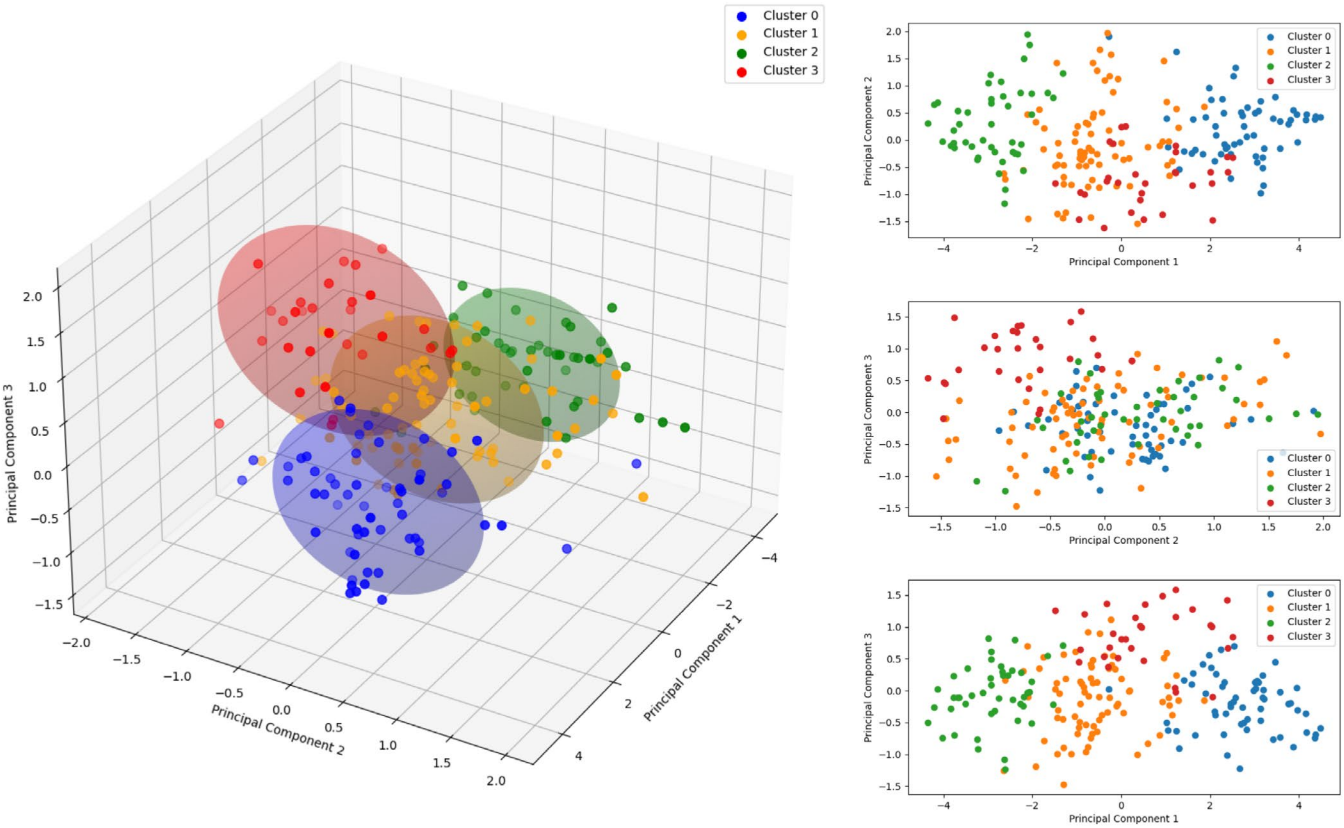

FIGURE 2 | Results from PCA and TSK-means clustering for four distinct clusters. Each data point's colour corresponds to its cluster membership, while ellipsoids delineate the 75th quantile of the density associated with each cluster.

- the seasonal component represents the recurring patterns or seasonal fluctuations in waste production, capturing periodic variations such as monthly or quarterly cycles. In waste generation data, we often observe recurring patterns that correspond to specific seasons or times of the year (i.e., there tends to be an uptick in waste generation during summer months, driven by heightened tourist activity, whereas winter months typically witness a decline in waste generation owing to reduced outdoor engagements). The seasonal component effectively captures these cyclic fluctuations, offering insights into the periodic variations in waste generation;
- the trend component encapsulates the long-term upward or downward trajectory of waste production, providing insights into overall growth or decline trends over time. This component helps in identifying overarching patterns such as steady growth or decline in waste generation. For instance, with increasing urbanisation and population growth, a municipality might experience a consistent upward trend in waste production over several years. Understanding these trends aids in predicting future waste production levels and planning appropriate waste management strategies;
- the remainder component represents the residual variation after removing the seasonal and trend components, encompassing random fluctuations and irregularities in the data. By isolating the remainder component, we can identify anomalous patterns or unexpected deviations from expected waste generation trends, which may be indicative of external factors or exceptional events impacting waste production.

This decomposition is applied to the historical waste production time series for each municipality in the region. By decomposing the time series we gain valuable insights into the temporal patterns and trends shaping waste generation dynamics at the local level. Furthermore, the decomposition facilitates the identification of seasonality and trend variations across different municipalities, enabling us to account for spatial heterogeneity in waste production patterns. This spatial-temporal analysis forms the foundation for subsequent forecasting models, allowing us to model each component separately and aggregate them to obtain accurate predictions of future waste production levels.

3.1.3 | Graph Construction

In this subsection, we delve into the process of constructing a comprehensive graph representation of municipalities. This phase is pivotal in encapsulating the spatial dependencies and temporal dynamics inherent in waste generation processes.

A graph $G = (V, E)$ is defined, where V represents the set of nodes corresponding to municipalities, and E denotes the set of edges representing the connections between municipalities. Each node in V represents a municipality, endowed with a rich set of features derived from heterogeneous data sources. The edges in E encode the relationships between municipalities, to

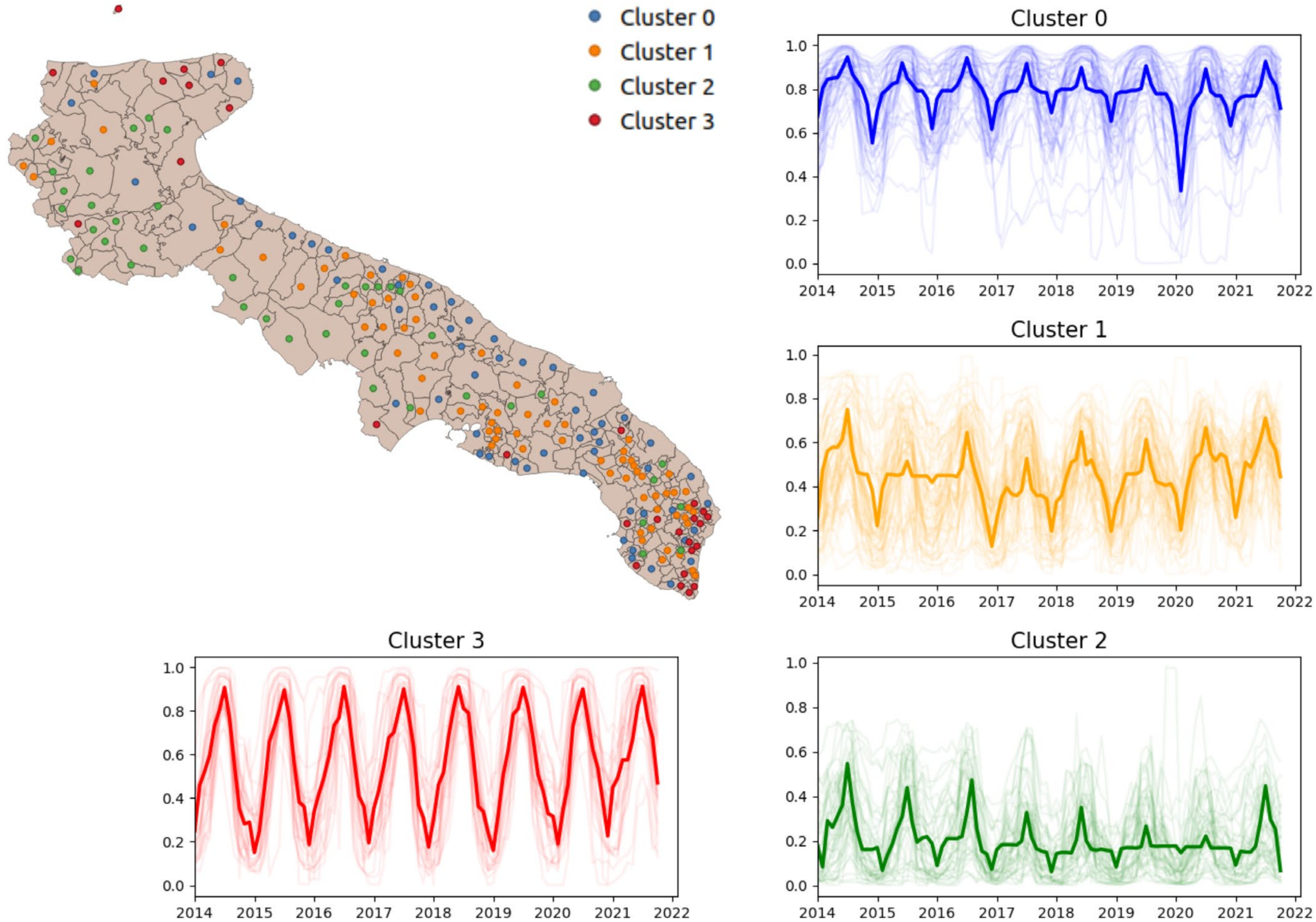

FIGURE 3 | On upper left the map of the Apulia region displaying the various municipalities with available time series data. Each point on the map represents a unique collected time series, distinguished by cluster membership through colour coding. In the subplots, normalised time series are depicted. The pale lines represent individual time series within the cluster, while the brighter line represents the Dynamic Time Warping Barycenter Averaging (DBA) of all-time series within that cluster.

TABLE 1 | Data source, data type and feature extracted for the built graph nodes.

Source	Data	Feature
ISTAT	Census residents on 1st January	• Population
	Housing expenditure: regions and type of municipality	• Spending habits
MEF	Municipal level income variables	• Income per capita
REGION	Regional technical map	• (Lat, lon)
		• Shape

quantify these relationships edge weights are computed based on similarity metrics derived from data.

The incorporation of diverse features enables the representation of municipalities in a multidimensional feature space, capturing the underlying socio-economic, environmental, and geographical factors influencing waste generation dynamics. These features serve as the basis for modelling the complex interactions and dependencies between municipalities within the graph structure. As previously noted, the primary data sources encompass three main categories: time series, demographic, and geographical data.

Specifically, demographic data, comprising population statistics alongside monthly spending habits and per capita income, as well as spatial data, were utilised in constructing the graph. Each node $v_i \in V$, representing a municipality, is thus endowed with the features in Table 1.

Node positions and feature attributes facilitate the generation of spatial adjacency and similarity matrices. In this paper, we concentrate on three factors that could affect waste generation:

1. Taking into account the feature **spending habits**, we investigate how the similarity in monthly spending behaviour

between two municipalities influences their waste production patterns. We hypothesise that municipalities with similar spending habits might contribute similarly to rubbish generation. The spending coefficient between distinct nodes k and j is computed as:

$$\text{coeff}_{\bar{s}}\ (k,j) = \exp\left(-\left\|\bar{s}_k - \bar{s}_j\right\|_2^2\right) \tag{2}$$

where, $\bar{s}_k$ and $\bar{s}_j$ denote the feature vectors of the two municipalities, normalised by their respective populations.

2. Moving on to account the **shape** feature, we delve into examining the similarity between municipalities based on their geographical characteristics. It is reasonable to posit that the geographical extension of a municipality correlates with the amount of various infrastructures within it, implying that municipalities with similar geographical extensions might yield comparable waste volumes. The area coefficient quantifies the extent of resemblance in terms of area between node pairs (k,j) as follows:

$$\text{coeff}_A\ (k,j) = \exp\left(-\left\|A_k - A_j\right\|_2^2\right) \tag{3}$$

where, A_k and A_j represent the land areas of the two different municipalities considered.

3. Finally, considering the (**lat**, **lon**) information, we investigate the impact of spatial proximity between municipalities on waste production. We assume that nearby municipalities might experience mobility among residents due to various reasons such as work or education, potentially leading to contamination concerning waste generation. By examining this coefficient, we aim to verify, and thus consider in our study, the presence of such patterns. The distance coefficient measures the shortest distance between two points on the Earth's surface employing the geodesic distance metric. Let P_1 and P_2 represent the geographical coordinates (latitude and longitude) of two nodes k and j, respectively. The geodesic distance, denoted as $d_{k,j}$, is computed using the Vincenty formula (Vincenty 1975), then the distance coefficient represents the spatial proximity as

$$\text{coeff}_d\ (k,j) = \exp\left(-d_{k,j}\right) \tag{4}$$

it ranges between 0 and 1, with higher values indicating closer spatial proximity between municipalities.

The final weight of the edge connecting municipalities k and j is computed as a weighted sum of the above coefficients:

$$w_{k,j} = \alpha \cdot \text{coeff}_{\bar{s}} + \beta \cdot \text{coeff}_A + \gamma \cdot \text{coeff}_d \tag{5}$$

where, α, β, and γ are fixed scalars representing the relative importance assigned to each factor.

This process is iterated for all pairs of municipalities within the graph, excluding self-referential comparisons. By aggregating these edge weights, we construct the final adjacency matrix:

$$W = \begin{pmatrix} 0 & w_{12} & \cdots & w_{1n} \\ w_{21} & 0 & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & 0 \end{pmatrix}$$

This matrix captures the spatial relationships and similarities between municipalities, providing a comprehensive representation of the graph structure and guiding subsequent analytical and modelling endeavours.

In Figure 4, we present a representation of the graph $G(M, w)$, where M symbolises the set of nodes corresponding to the municipalities. This visual depiction provides insight into the graph structure, highlighting the interconnection of the nodes.

3.2 | Model Architecture

As previously mentioned, our proposed GNN-based model, is structured around two fundamental components: the graphic component and the temporal component. The graphic component is tasked with analysing the data synthesised by the G graph, while the temporal component manages the processing of time series data specifically related to waste generation. In the subsequent section, we will delve deeper into the architectural intricacies of each of these components.

3.2.1 | Graphical Component

The graphical component of our model is responsible for analysing the structured data represented by the graph G. Leveraging the principles of GNN, this component extracts and integrates information from the interconnected nodes and edges of the graph. By considering the spatial relationships and similarities encoded within the graph, the graphical component captures the contextual information essential for understanding waste generation dynamics at the municipal level.

Specifically, the graphical component employs graph convolutional layers to aggregate information from neighbouring nodes within the graph. These layers facilitate the propagation of information across the graph, allowing each node to incorporate insights from its neighbouring nodes.

It can be formalised as follows:

$$X_s = \tanh(W \cdot \sigma(A \cdot X)) \tag{6}$$

where, X denotes the input feature matrix of the nodes, A represents the adjacency matrix of the graph, W is the weight matrix of the `GCN` layer, and σ denotes the `ReLU` activation function.

3.2.2 | Temporal Component

Complementing the graphical component, the temporal component of our model manages the processing of time series data

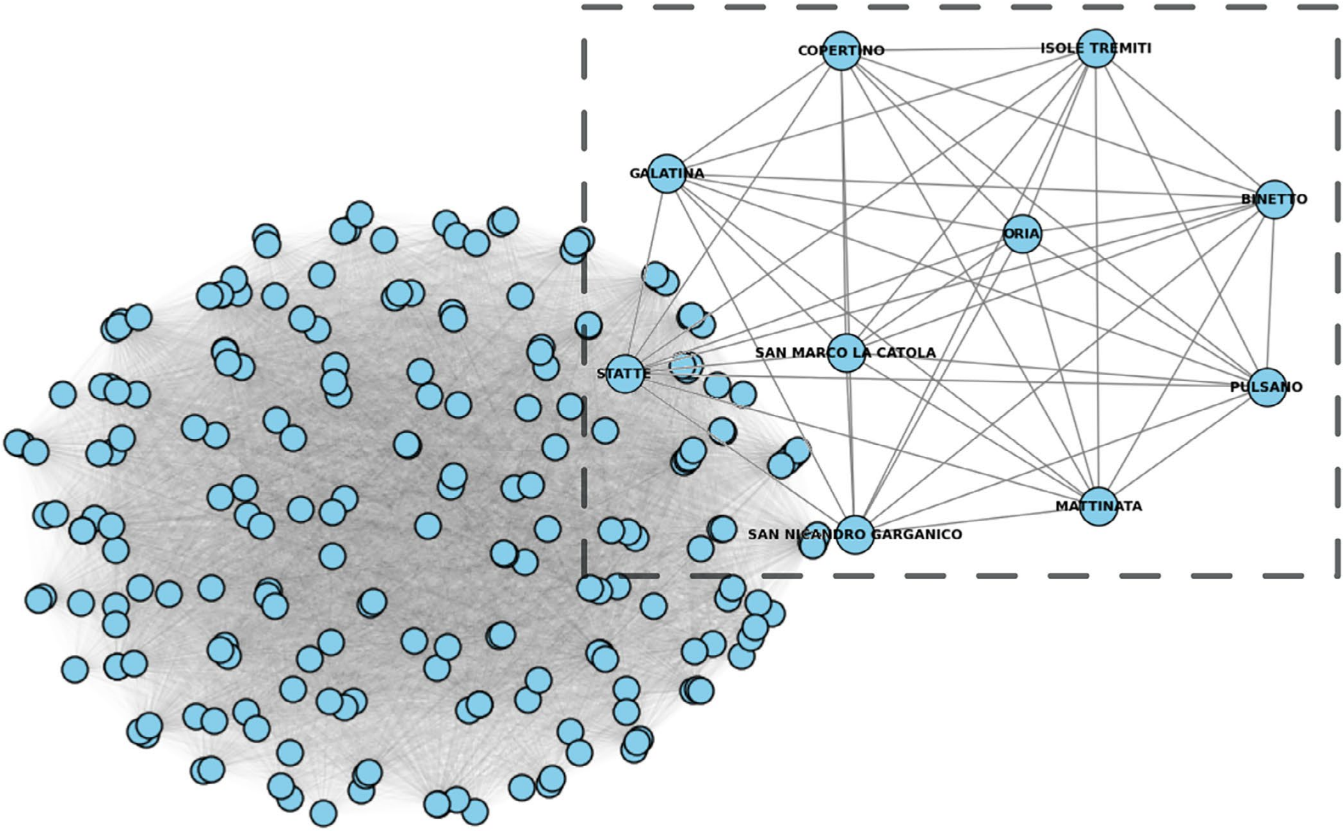

FIGURE 4 | The main image depicts the complete graph $G(M, w)$ containing all 210 municipalities of the Apulia region. In the enlarged image alongside, a zoom-in shows a subgraph containing only 10 randomly selected municipalities from the complete graph. This subgraph provides a detailed and focused view of a small subset of entities, allowing examination of relationships among a selected subset.

pertaining to waste generation. Time series data, comprising historical records of waste production over time, presents unique challenges and opportunities for modelling. The temporal component is designed to address these challenges and effectively capture the temporal dynamics inherent in waste generation patterns.

To achieve this, the temporal component consists of one Recurrent Neural Networks (RNNs) layer and two LSTM layers. These layers are adept at modelling sequential data and capturing long-term dependencies, making them well-suited for analysing time series data. By processing the STL decomposition of waste production time sequences, the temporal component captures the inherent temporal patterns and trends, providing valuable insights for forecasting and analysis.

It can be formalised as

$$X_t = \mathrm{MLP}\left(X^{(1)} \oplus X^{(2)} \oplus X^{(3)}\right) \tag{7}$$

where, $X^{(1)}$, $X^{(2)}$, and $X^{(3)}$ represent the output tensors from the `LSTM` and `RNN` layers, respectively. In this equation, these layers are applied to different input dimensions of X, each corresponding to a component of the STL decomposition. Specifically, the `LSTM` layers are applied to capture seasonality and trend, while the `RNN` layer is applied to the remainder component. The outputs of these layers are then concatenated along the feature dimension (the concatenation is denoted by $\oplus$). Finally, the concatenated tensor is passed through the `MLP` model to obtain the output X_t.

3.2.3 | Integrated Model and Loss Function

In this section, we present the integrated model architecture resulting from the fusion of the spatial and temporal components discussed earlier. Additionally, we introduce the loss function utilised for training the model.

- The integrated model combines the spatial and temporal components to form a comprehensive framework for waste generation prediction. Leveraging the insights captured by the graphical component and the temporal dynamics learned by the temporal component, the integrated model provides a holistic approach to forecasting waste production at the municipal level. The mathematical formulation of the integrated model can be expressed as follows:

$$\mathrm{out} = \mathrm{MLP}\left(X_t \oplus X_s\right) \tag{8}$$

Here, X_t represents the output tensor obtained from the temporal component and X_s denotes the output tensor derived from the spatial component, extracted through the `GCN` layer. The outputs from both components are concatenated along the feature dimension and passed through a multi-layer perceptron for final prediction. This integrated approach ensures robust modelling of waste generation patterns, leveraging both spatial and temporal information for enhanced predictive performance. Notably, the integrated model is trained and tested on each cluster separately, allowing for cluster-specific insights and predictions.

- In our model, we utilise a customised loss function tailored to the unique requirements of waste production prediction. Termed as the Comprehensive Waste Prediction Loss (CWPL), this loss function incorporates multiple components to ensure holistic evaluation and optimisation of model performance. The loss comprises three key components: (1) Mean Squared Error (MSE) loss measures the average squared difference between the predicted and actual waste production values, it quantifies the overall discrepancy between the predicted and ground truth values; (2) R^2 loss evaluates the goodness of fit of the model by assessing the proportion of variance in the waste production data that is explained by the model, it complements the MSE loss by providing insights into the predictive accuracy relative to the variability in the data; (3) annual error (AE) component captures the discrepancy in waste production predictions at an annual granularity, it measures the mean absolute difference between the predicted and actual annual waste production values, offering insights into the model's performance over longer temporal horizons. The loss function is expressed as follows:

$$\mathrm{CWPL} = \underbrace{\frac{1}{N}\sum_{i}^{N}(y_i - \bar{y}_i)^2}_{\mathrm{MSE}} + \underbrace{\frac{\sum_i (y_i - \mathbb{E}[y])^2}{\sum_i (y_i - \bar{y}_i)^2}}_{R^2} + \underbrace{\left|\sum_i y_i - \sum_i \bar{y}_i\right|}_{\mathrm{AE}} \tag{9}$$

where, y and $\bar{y}$ represent the waste production and its prediction, and $\mathbb{E}[\cdot]$ is the mean operator. By incorporating multiple loss components, the CWPL offers a robust framework for training and evaluating waste production prediction models, ensuring accurate and reliable forecasts in real-world scenarios.

3.2.4 | Model Interpretability

To enhance the interpretability of our model and provide greater insights into the logic behind its predictions, we conducted the feature importance analysis using a Random Forest Regressor for the integrated model. Moreover, we also provide a detailed explanation of how the model processes input data, from raw features to final predictions.

3.2.4.1 | Logic Behind the Predictions. Our model operates by processing both temporal and spatial data, capturing patterns that influence waste generation. For each instance, the model takes two inputs: temporal data (i.e., waste production trends over time for a specific municipality) and Spatial data (i.e., a graph where nodes represent municipalities and edges capture relationships such as geographic proximity or shared waste management policies). The LSTM layers capture long-term dependencies and seasonal variations (e.g., recurring peaks in waste production during tourist seasons or holidays), while the RNN focuses on short-term fluctuations (e.g., week-to-week changes). The temporal component outputs a tensor X_t that summarises these patterns. Simultaneously, the GCN processes the graph structure. It aggregates information from neighbouring nodes (i.e., municipalities) and captures spatial relationships, such as waste production similarities between geographically close areas or areas with similar policies. For example, if a municipality has neighbouring cities that recently experienced spikes in waste production due to tourism or local events, this information is propagated through the GCN layers. The result is an output tensor X_s that embeds spatial dependencies. The outputs from the temporal (X_t) and graphical components (X_s) are concatenated. This combination allows the model to consider both historical waste trends and the influence of nearby municipalities simultaneously. For instance, if the time series suggests an upward trend in waste generation due to an ongoing festival, and the spatial data indicates that neighbouring cities are also experiencing increased waste production, the model will weight these factors accordingly in its prediction. The concatenated tensor is passed through a multi-layer perceptron (MLP) that finalises the prediction. This MLP acts as a synthesiser, integrating both temporal and spatial insights to generate the waste forecast for the municipality.

3.2.4.2 | Random Forest Feature Importance. To provide insight into which features are driving the model's predictions, we performed a feature importance analysis using a Random Forest Regressor. This method helps in understanding the contribution of each input feature to the final predictions of waste generation. The Random Forest algorithm was applied as a post-hoc analysis tool, leveraging its ability to rank the importance of input features based on the reduction of impurity in its decision trees. In Figure 5, the feature importance is broken down by clusters, showing the relative importance of the temporal components (i.e., trend, seasonal, and residual) and the spatial component (i.e., KG)

FIGURE 5 | Feature importance for the integrated model using Random Forest Regressor.

in each group of municipalities. The trends observed vary significantly across the clusters. For instance, in Cluster 0, the trend component plays a dominant role, indicating that long-term trends in waste generation are crucial in this cluster. On the other hand, Cluster 1 places more emphasis on the KG (spatial component), highlighting the strong influence of neighbouring municipalities and regional policies on waste generation in this cluster. These differences in feature importance can be attributed to the distinct temporal and spatial dynamics of waste generation within each cluster, as demonstrated in Figure 3. The clustering of time series, as depicted on the map and corresponding plots, reveals varying patterns across municipalities. For example, municipalities in Cluster 2 (shown in green) exhibit more pronounced seasonal fluctuations, with peaks corresponding to tourist seasons. This is reflected in the higher importance assigned to the seasonal component in Cluster 2. Conversely, municipalities in Cluster 3 (shown in red) display strong periodic behaviour, leading to more balanced contributions from both trend and seasonal components.

By incorporating these feature importance analyses, we aim to strengthen the trust decision-makers can place in our model. Understanding the factors driving waste generation predictions allows stakeholders to interpret the results in a more transparent manner. In particular, insights into the temporal dynamics and spatial dependencies captured by the model can inform policy decisions, enhance planning strategies, and foster more effective interventions in waste management.

4 | Experimental Scenario and Case Study

In this section, we provide a detailed overview of the implementation process and the experimental setup employed to evaluate the performance of our proposed methodology for waste generation forecasting.

4.1 | Data

To evaluate the performance of the proposed method in waste generation forecasting, we conduct experiments on real-world datasets.

- *MSW-Apulia*: The Municipal Solid Waste (MSW) Apulia Dataset (Puglia 2016) forms the cornerstone of our analysis, providing valuable time series data on waste production. This dataset encompasses monthly records for each of the 257 municipalities in the region spanning multiple years, allowing for a comprehensive exploration of temporal patterns and trends. For our experiments, we included samples from the years 2014 to 2021, as these were the only ones with continuous data without temporal gaps. This ensures consistency and integrity in the time series data used for analysis. In these measurements, MSW refers to various types of waste, including organic, recyclable and non-recyclable waste, providing a holistic view of the dynamics of waste generation.
- *Demographic Characteristics and Citizenship*: The demographic characteristics and citizenship (DCC) dataset by ISTAT (Istituto Nazionale di Statistica 2016) collects information on persons usually resident in each municipality, thus providing an overview of the demographic and social structure of Italy and its territories. The census is based on the integration between data from administrative sources and those acquired through surveys involving a representative sample of municipalities and households each year.
- *Household Expenditure Dataset*: The household expenditure dataset (HED) from ISTAT provides comprehensive information on average monthly expenditure indicators per household across different population bands, which are classified according to the number of inhabitants in the municipality. The HED is particularly relevant in understanding household expenditure dynamics and their implications for waste generation.
- *Municipal Income Dataset*: The municipal income dataset (MID) by MEF (Ministero dell'Economia e delle Finanze 2016) provides information on income and main personal income tax (IRPEF) variables on a municipal basis. It encompasses various taxpayer types and includes data on the number of taxpayers and total income across municipalities. This dataset is crucial for understanding the economic landscape of different municipalities and it is useful in analysing socioeconomic factors influencing waste generation patterns.

4.2 | Implementation Setting

In the temporal component, both the two `LSTM` and one `RNN` layer consist of 64 units each, with the `Tanh` activation function utilised for the former and `ReLu` for the latter. Concerning the spatial component, the coefficients employed to compute the edge weights in the graph are set as $\alpha = 0.25$, $\beta = 0.25$, and $\gamma = 0.5$. The `GCN` layer precedes a fully connected layer, activated by the `Tanh` function. During experimentation, we employed the Adam optimiser with a fixed learning rate of 0.001 and weight decay ranging from $1e^{-3}$ to $1e^{-5}$, contingent on the time series cluster under consideration. The model was trained for a maximum of 1500 epochs, employing early stopping criteria with a patience of 100. The batch size was set to 256.

As mentioned earlier, the dataset covers the period from 2014 to 2021. For our experimentation, we selected various forecasting horizons that mirror realistic timeframes necessary to evaluate the efficacy of waste management policies. This deliberate choice ensures that the assessment aligns with real-world scenarios, allowing sufficient time to observe the impacts and effectiveness of implemented strategies. Accordingly, we partitioned each dataset into separate training and test sets based on the forecasting horizon. For a 6-month horizon, the test set comprises the last 6 months of 2021, with all preceding years forming the training set. Similarly, for a 12-month horizon, the test set consists of the entire year 2021, while the training set encompasses data from earlier years. Finally, for an 18-month horizon, the test set includes data from 2021 and the last 6 months of 2020, with previous years serving as the training set.

To assess the model's performance, four evaluation metrics were utilised: mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination R^2. Optimal performance is indicated by lower values for MSE, MAE, and RMSE, suggesting improved accuracy.

TABLE 2 | Experimental results of the proposed framework obtained across different clusters and time horizons.

Dataset	Horizon (m)	MSE↓	MAE↓	RMSE↓	R^2↑
Cluster0	6	7.00e−05	6.10e−03	7.17e−03	5.71e−01
	12	1.42e−04	6.78e−03	8.29e−03	7.20e−01
	18	4.31e−04	1.36e−02	1.62e−02	4.31e−01
Cluster1	6	1.96e−04	9.10e−03	1.08e−02	4.39e−01
	12	1.44e−04	7.35e−03	9.38e−03	6.72e−01
	18	2.63e−04	1.06e−02	1.34e−02	4.81e−01
Cluster2	6	2.79e−04	1.05e−02	1.23e−02	4.13e−01
	12	3.56e−04	1.06e−02	1.30e−02	6.10e−01
	18	2.31e−03	2.45e−02	3.04e−02	1.26e−01
Cluster3	6	5.16e−04	1.14e−02	1.32e−02	8.59e−01
	12	2.11e−03	1.74e−02	2.24e−02	8.08e−01
	18	2.47e−03	1.95e−02	2.55e−02	7.37e−01
Mean results	6	2.25e−04	8.90e−03	1.04e−02	5.30e−01
	12	4.69e−04	9.32e−03	1.17e−02	6.92e−01
	18	1.06e−03	1.57e−02	1.95e−02	4.28e−01

Note: For each cluster (Cluster0, Cluster1, Cluster2, and Cluster3), the mean squared error, mean absolute error, root mean square error, and coefficient of determination metrics are reported for three distinct prediction horizons: 6 months (6 m), 12 months (12 m), and 18 months (18 m).

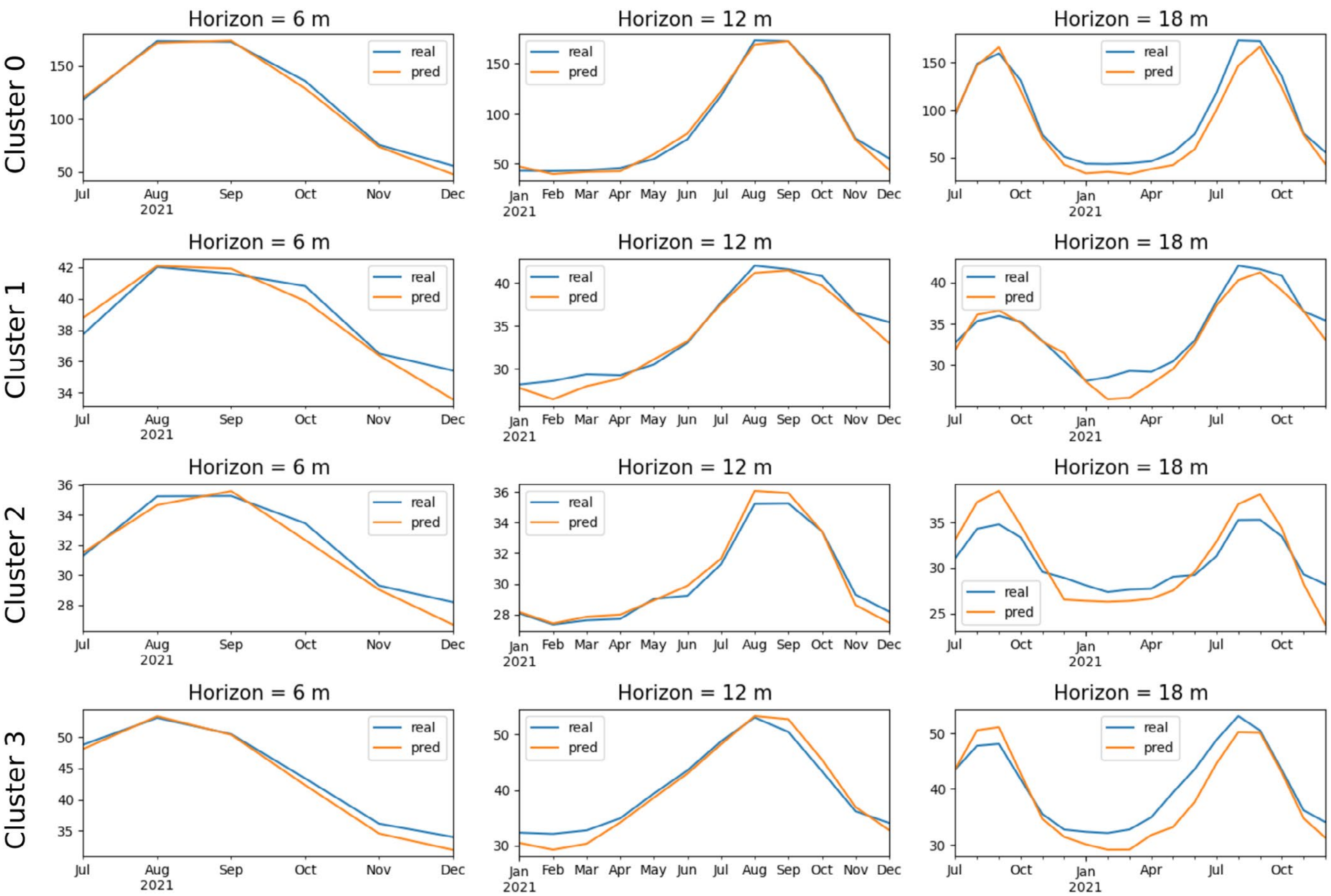

FIGURE 6 | Predicted and actual waste production across varying prediction horizons for the four distinct clusters.

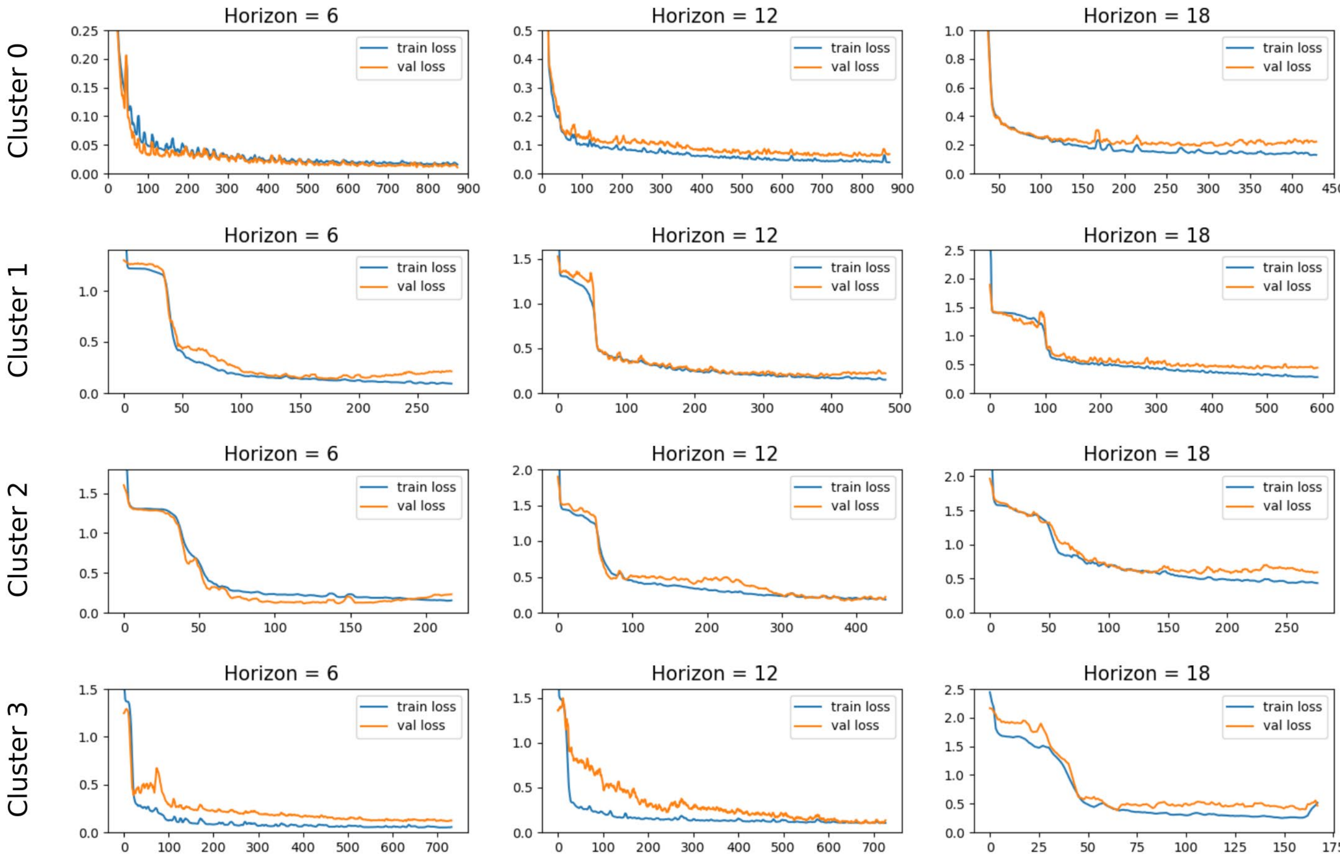

FIGURE 7 | Loss functions depicting the performance of our predictive model across clusters and prediction horizons.

TABLE 3 | The average results of Waste Production Discrepancy and Absolute Waste Discrepancy across different prediction horizons are presented.

Dataset	Horizon (m)	WPD ↓	AWD ↓
Mean results	6	1.35e+04	1.63e−02
	12	**−2.65e+03**	**1.17e−02**
	18	5.75e+04	3.72e−02

Note: It is important to note that for WPD, a negative value suggests an underestimation of waste production, whereas a positive value indicates an overestimation. These metrics provide valuable insights into the accuracy and consistency of the model's predictions across various time horizons. Bold indicates the superior outcomes.

Conversely, higher values of R^2 closer to 1 signify a superior fit of the model to the data. However, it is worth noting that in certain cases (outliers), R^2 may fall outside the typical range of 0 to 1. In instances where R^2 is negative, it suggests poor model fit to the data. To standardise the evaluation process, we set negative R^2 values to zero. This convention ensures a consistent assessment, treating both negative and zero values as indicators of inadequate model adaptation to the data. Hence, in general, we apply the transformation: $R^2 = \max(0, R^2)$.

4.3 | Computational Resources

The computational experiments for our proposed model were conducted on an NVIDIA GeForce RTX 3090 GPU, which features 24 GB of GDDR6X memory. The entire model was implemented using the PyTorch DL framework, leveraging GPU to reduce training and inference times. Despite the use of powerful hardware, it is important to note that the model's architecture is designed to be scalable and adaptable to less resource-intensive environments. For smaller datasets or regions with fewer nodes, the computational burden decreases significantly. Additionally, graph partitioning techniques can be employed to further optimise the processing of larger regions. From a practical perspective, while the model performs optimally on a GPU, it can still be run on standard multi-core CPUs with moderate hardware specifications. Moreover municipalities with limited computational infrastructure could still use the model for inference after it has been pre-trained. This makes the model flexible for deployment in regions where access to advanced computational resources may be constrained. In terms of scalability, future work will focus on optimising the computational efficiency of the model. We aim to strike a balance between performance and resource consumption, ensuring that the model remains accessible and effective in various socio-economic and geographic contexts.

4.4 | Experimental Results

In this subsection, we delineate the fundamental aspects of the experiment, initially concentrating on the primary experiments and main results, which entail comparing the performance of models. Subsequently, we delve into the ablation study. Our aim is to showcase the efficacy of our proposed framework in

TABLE 4 | Comparison of the performance between the proposed model and various other models in predicting waste production.

Horizon (m)	Model	MSE↓	MAE↓	RMSE↓	R^2↑
6	SARIMA	7.90e−04	1.16e−02	2.23e−02	3.06e−01
	PROPHET	7.34e−03	4.01e−02	6.99e−02	0.00
	N-BEATS	2.93e−03	3.56e−02	3.93e−02	2.36e−01
	GCN	**2.20e−04**	**8.90e−03**	**1.04e−02**	**5.30e−01**
12	SARIMA	9.00e−04	1.33e−02	2.80e−02	3.38e−01
	PROPHET	1.83e−03	2.05e−02	3.68e−02	1.23e−01
	N-BEATS	8.16e−03	6.42e−02	7.08e−02	1.85e−01
	GCN	**4.70e−04**	**9.32e−03**	**1.17e−02**	**6.92e−01**
18	SARIMA	1.56e−03	1.86e−02	3.45e−02	1.73e−01
	PROPHET	5.69e−03	3.72e−02	6.25e−02	0.00
	N-BEATS	6.17e−03	6.04e−02	6.49e−02	1.55e−01
	GCN	**1.06e−03**	**1.57e−02**	**1.95e−02**	**4.28e−01**

Note: Bold indicates the superior outcomes.

forecasting future waste production across various time horizons and to establish the superiority of our proposed model.

4.4.1 | Main Results

In this subsection, we detail the experiments conducted and the principal findings obtained by applying the proposed model to the time series data.

As discussed in Section 3.1.1, the TSK-means algorithm is employed to cluster the data. By utilising the Elbow method, we identified the optimal number of clusters to be 4. Specifically, in `Cluster0`, there are 58 municipalities, while in `Cluster1`, there are 77 municipalities. Similarly, `Cluster2` comprises 45 municipalities, and `Cluster3` includes 30 municipalities. It is worth noting that the total number of municipalities across all clusters is less than 257, the total number for the Apulia region. This discrepancy arises because municipalities lacking MSW data for the entire observation period are excluded.

The datasets for all clusters encompass the years 2014 to 2021, with monthly granularity, resulting in 96 samples for each municipality. To ensure consistency, we maintain uniform configurations for the step size settings of the historical input, set to 1, and the prediction horizon for each cluster. This standardised approach facilitates an effective assessment of the model's performance across varying forecast horizons.

The Table 2 presents the experimental outcomes of our proposed framework across each cluster, delineating results for three distinct time horizons. Additionally, it provides aggregated results for the entire dataset, derived by averaging the aforementioned metrics. This comprehensive analysis offers insights into the performance of our framework across various temporal scopes, facilitating a holistic understanding of its predictive capabilities.

In order to demonstrate the efficacy and precision of the proposed approach, Figure 6 illustrates the flow curves for three distinct prediction intervals within each cluster. Through a comparative analysis of the actual flow curves and the predicted ones, it becomes apparent that the proposed method adeptly captures variations in the flow patterns, leading to precise predictions. Notably, it accurately depicts fluctuations in waste generation during peak seasons, showcasing a robust alignment between the model's predictions and observed waste patterns. This visual representation underscores the reliability and effectiveness of our proposed method in forecasting waste production dynamics.

Additionally, Figure 7 presents the loss functions for each cluster across different prediction horizons. These loss functions provide a quantitative measure of the performance of our predictive model. By examining the loss curves, it can be observed how the model's performance varies across different clusters and prediction horizons. This analysis further supplements our understanding of the predictive capabilities of our framework, highlighting areas where improvements may be needed and reaffirming its effectiveness in capturing temporal dynamics in waste production.

Moreover, for a comprehensive assessment of our model's performance aligned with the real-world problem, we introduce two bespoke metrics crafted specifically for this study: the Waste Production Discrepancy (WPD) and the Absolute Waste Discrepancy (AWD). These metrics are tailored to capture waste generation dynamics; they are thought to align with the intricacies of the problem at hand as:

$$\text{WPD} = \frac{1}{1000}\sum_{m=1}^{N_m}\left(T_m - \overline{T}_m\right) \quad \text{and} \quad \text{AWD} = \frac{1}{N_m}\sum_{m=1}^{N_m}\frac{\left|T_m - \overline{T}_m\right|}{T_m} \tag{10}$$

Here, WPD represents the difference in waste production between the sum of the real values (T_m) and the sum of the model predictions ($\overline{T}_m$), converted from kilograms (kg) to tonnes (ton).

TABLE 5 | Comparison of Waste Production Discrepancy (WPD) and Absolute Waste Discrepancy (AWD) across different models.

Horizon (m)	Model	WPD ↓	AWD ↓
6	SARIMA	**−1.18e+03**	4.01e−02
	PROPHET	−1.13e+07	3.63e+03
	N-BEATS	4.70e+04	8.43e−02
	GCN	1.35e+04	**1.63e−02**
12	SARIMA	2.34e+04	5.64e−02
	PROPHET	−1.04e+07	1.75e+03
	N-BEATS	2.54e+05	1.55e−01
	GCN	**−2.65e+03**	**1.17e−02**
18	SARIMA	1.35e+05	8.50e−02
	PROPHET	−9.52e+06	1.05e+03
	N-BEATS	3.78e+05	1.55e−01
	GCN	**5.75e+04**	**3.72e−02**

Note: Bold indicates the superior outcomes.

AWD indicates the mean absolute difference between the actual and predicted waste production, normalised by the total real waste production across all municipalities. These metrics offer valuable insights into the disparity between observed and predicted waste generation, thereby providing a nuanced evaluation of the model's performance within the real-world context. The Table 3 shows the average results for various prediction horizons.

4.4.2 | Comparative Analysis

Moreover, to validate the superior efficacy of our proposed model architecture, we conducted comparisons with three other commonly used models in time series forecasting: SARIMA, PROPHET, and N-BEATS. These models serve as benchmarks in the field, each employing distinct methodologies for processing time series data.

- SARIMA (Shumway et al. 2017), which stands for Seasonal Autoregressive Integrated Moving Average, is a statistical method that extends the traditional ARIMA model by incorporating seasonal components (Box et al. 2015). It is particularly adept at capturing both long-term trends and recurring seasonal patterns in time series data, making it useful in contexts where waste generation exhibits seasonal fluctuations (e.g., increased waste during holidays or tourist seasons). However, SARIMA models rely heavily on assumptions of linearity and stationarity, which may limit their effectiveness when dealing with more complex, non-linear patterns of waste generation.
- PROPHET (Facebook 2017), developed by Facebook, is a robust forecasting algorithm that models annual, weekly, and daily time series trends using additive components. It is designed to handle missing data and outliers, and is particularly well-suited for datasets with strong seasonal trends. In the context of urban waste management, PROPHET's ability to model periodic trends could prove advantageous in predicting waste peaks and troughs related to recurring events (e.g., weekends, public holidays). However, while PROPHET is easy to use and interpretable, it may struggle with capturing more granular, short-term variations in waste production.
- N-BEATS (Neural Basis Expansion Analysis Time Series) (Oreshkin et al. 2019), on the other hand, is a DL-based model specifically designed for time series forecasting. It uses a hierarchical structure with multi-scale building blocks, allowing it to learn from different time frequencies simultaneously. This capability makes it particularly powerful in capturing both short-term variations and long-term trends in time series data. In the field of urban waste management, N-BEATS could be especially useful for cities with highly variable waste production patterns, as it can model complex, non-linear dynamics that are often present in urban environments. However, its main limitation lies in its need for large amounts of data to train the model effectively, which may not always be available in all regions.

It is important to note that our framework introduces a novel approach by integrating GNNs to encode spatial information, marking a departure from conventional forecasting methods. As such, direct comparisons with existing models for waste generation forecasting are not straightforward. Instead, the focus is on comparing these models within the context of time series forecasting based on seasonal patterns. This distinction underscores the innovative nature of our approach and the unique insights it offers into waste generation forecasting. In our experiments, we chose to compare the performance of various models across the three distinct prediction horizons. The comparative outcomes are depicted in the Table 4, showcasing the assessment of model performance relative to these alternative approaches.

Moreover, as mentioned earlier, to provide a more comprehensive evaluation of our model's performance in addressing real-world challenges, we introduced WPD and AWD. Below, we present a comparison Table 5 alongside other considered models to highlight the efficacy of our approach. Our model consistently demonstrates superior performance compared to SARIMA, PROPHET, and N-BEATS models, exhibiting lower MSE, MAE, and RMSE values, along with higher values of R^2. Notably, the PROPHET model exhibits comparatively higher error rates and lower R^2 values, indicating his limitations in achieving precise predictions and optimal alignment with the datasets.

Finally, our proposed model consistently outperforms across the three time horizons in terms of predictive accuracy. Its sustained superiority in accuracy and alignment establishes it as the preferred choice among the considered models, offering enhanced predictive capabilities across diverse time horizons.

4.4.3 | Ablation Study

This paper underscores the adoption of a tailored loss function designed specifically for the task at hand, which involves

TABLE 6 | Experimental results of the ablation study on the proposed methodology, examining the impact of integrating individual components of the loss function.

Horizon (m)	Loss	MSE↓	MAE↓	RMSE↓	R^2↑
6	MSE	3.82e−03	2.99e−02	3.35e−02	1.77e−01
	R^2	2.24e+01	3.67e+00	4.34e+00	0.00
	AE	1.94e−02	1.12e−01	1.34e−01	4.91e−03
	MSE + R^2	3.00e−04	**8.45e−03**	1.01e−02	**5.57e−01**
	MSE + AE	6.30e−04	1.37e−02	1.58e−02	4.10e−01
	R^2 + AE	**2.00e−04**	8.49e−03	**9.97e−03**	5.40e−01
	CWPL	2.20e−04	8.90e−03	1.04e−02	5.30e−01
12	MSE	2.68e−3	2.85e−2	3.36e−2	1.48e−1
	R^2	3.16e+0	1.42e+0	1.75e+0	0.00
	AE	1.66e−2	1.02e−1	1.24e−1	0.00
	MSE + R^2	1.30e−3	1.20e−2	1.51e−2	6.09e−1
	MSE + AE	1.90e−3	2.44e−2	2.90e−2	1.01e−1
	R^2 + AE	7.75e−2	1.04e−1	1.23e−1	3.26e−1
	CWPL	**4.70e−4**	**9.32e−3**	**1.17e−2**	**6.92e−1**
18	MSE	2.34e−03	2.73e−02	3.26e−02	1.52e−01
	R^2	6.14e+00	1.80e+00	2.22e+00	0.00
	AE	3.51e−02	1.35e−01	1.74e−01	0.00
	MSE + R^2	1.07e−03	1.80e−02	2.17e−02	3.67e−01
	MSE + AE	1.92e−03	2.50e−02	3.10e−02	1.50e−01
	R^2 + AE	1.12e−03	1.84e−02	2.26e−02	3.42e−01
	CWPL	**1.06e−03**	**1.57e−02**	**1.95e−02**	**4.28e−01**

Note: Bold indicates the superior outcomes.

predicting the waste generated by individual municipalities. As a result, assessing the significance of each component of this loss function becomes paramount. In Table 6, we present the experimental findings from the ablation study of our proposed methodology, focusing on investigating the influence of integrating solely the MSE component, solely the R^2 component, and solely the AE component.

The experimental findings from the ablation study underscore the superiority of the proposed methodology, particularly when integrating all three components of the loss function. Among these components, the MSE exhibits the most substantial influence on overall performance, significantly enhancing the accuracy and alignment of the model with the observed data.

In essence, the results consistently highlight the efficacy of the proposed approach, with notable performance improvements observed, especially for the 12-month prediction horizon. This underscores the importance of a comprehensive understanding of the intricacies inherent in the real-world problem domain, especially when employing predictive models that integrate diverse sources of information.

5 | Conclusions

This paper presented a methodology for forecasting municipal waste production using DL and GNNs, specifically focusing on the integration of diverse data sources to model waste generation dynamics. Our approach, tested in the Apulia region, demonstrates the utility of GNNs in capturing the complex interplay of factors affecting waste production at the municipal level. The methodology's accuracy in forecasting waste production suggests its potential as a tool for improving waste management strategies and planning.

The adaptability and scalability of the proposed framework are among its key strengths, indicating its applicability to various geographic and demographic contexts. This flexibility allows for the possibility of its use in different municipalities, not limited to the case study region, thereby broadening the scope of its utility in MWM.

Future research directions could include exploring the incorporation of additional types of data, such as economic indicators or environmental factors, to further enhance the model's predictive accuracy. Additionally, the development of real-time data processing capabilities could transform the framework into a dynamic tool for waste management, capable of adapting forecasts based on new data.

In summary, our study contributes to the field of MWM by offering a novel forecasting methodology that leverages DL and GNNs. While the results are promising, continued exploration and refinement are necessary to fully realise the potential of this approach in contributing to sustainable waste management practices.

Acknowledgements

This work was supported by PNRR Centro Nazionale HPC, Big Data e Quantum Computing, (CN_00000013) (CUP: E63C22000980007), under the NRRP MUR program funded by the NextGenerationEU.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available in CAPTURE at https://github.com/MODAL-UNINA/CAPTURE. These data were derived from the following resources available in the public domain: Municipal Solid Waste Apulia Dataset, https://pugliacon.regione.puglia.it/orp/public/servizi/rsu-per-comune; Demographic Characteristics and Citizenship, http://dati-censimentipermanenti.istat.it; Municipal Income Dataset, https://www1.finanze.gov.it/finanze/analisi_stat/public/index.php?search_class[0]=cCOMUNE&opendata=yes.

References

Ahmed, S., S. Mubarak, J. T. Du, and S. Wibowo. 2022. "Forecasting the Status of Municipal Waste in Smart Bins Using Deep Learning." *International Journal of Environmental Research and Public Health* 19, no. 24. https://www.mdpi.com/1660-4601/19/24/16798.

Box, G. E., G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

Christensen, T. H. 2010. *Solid Waste Technology & Management*. John Wiley & Sons.

Cleveland, R. B., W. S. Cleveland, J. E. McRae, and I. Terpenning. 1990. "STL: A Seasonal-Trend Decomposition." *Journal of Official Statistics* 6, no. 1: 3–73.

Cuturi, M., and M. Blondel. 2017. "Soft-dtw: A Differentiable Loss Function for Time-Series." In *International Conference on Machine Learning*, 894–903. PMLR.

Facebook. 2017. "Prophet—Forecasting at Scale." https://facebook.github.io/prophet/.

Fang, B., J. Yu, Z. Chen, et al. 2023. "Artificial Intelligence for Waste Management in Smart Cities: A Review." *Environmental Chemistry Letters* 21, no. 4: 1959–1989.

Ferronato, N. T. V. 2019. "Waste Mismanagement in Developing Countries: A Review of Global Issues." *International Journal of Environmental Research and Public Health* 16, no. 6: 1060.

Gatti, A., E. Barbierato, and A. Pozzi. 2024. "Toward Greener Smart Cities: A Critical Review of Classic and Machine-Learning-Based Algorithms for Smart Bin Collection." *Electronics* 13, no. 5: 836. https://www.mdpi.com/2079-9292/13/5/836.

Hembert, P., C. Ghnatios, J. Cotton, and F. Chinesta. 2024. "Assessing Sensor Integrity for Nuclear Waste Monitoring Using Graph Neural Networks." *Sensors* 24, no. 5: 1580. https://www.mdpi.com/1424-8220/24/5/1580.

Huang, X., Y. Ye, L. Xiong, R. Y. Lau, N. Jiang, and S. Wang. 2016. "Time Series k-Means: A New k-Means Type Smooth Subspace Clustering for Time Series Data." *Information Sciences* 367: 1–13.

Ihsanullah, I., G. Alam, A. Jamal, and F. Shaik. 2022. "Recent Advances in Applications of Artificial Intelligence in Solid Waste Management: A Review." *Chemosphere* 309: 136631.

Istituto Nazionale di Statistica. 2016. "Demographic Characteristics and Citizenship." http://dati-censimentipermanenti.istat.it.

Johnson, N. E., O. Ianiuk, D. Cazap, et al. 2017. "Patterns of Waste Generation: A Gradient Boosting Model for Short-Term Waste Prediction in New York City." *Waste Management* 62: 3–11.

Kakad, R. 2022. "Forecasting Municipal Solid Waste Generation Using a Multiple Linear Regression Model." *International Research Journal of Engineering and Technology* 6: 1–14.

Kannangara, M., R. Dua, L. Ahmadi, and F. Bensebaa. 2018. "Modeling and Prediction of Regional Municipal Solid Waste Generation and Diversion in Canada Using Machine Learning Approaches." *Waste Management* 74: 3–15.

Kodinariya, T. M., and P. R. Makwana. 2013. "Review on Determining Number of Cluster in K-Means Clustering." *International Journal* 1, no. 6: 90–95.

Mauliska, N., W. Lestari, E. T. Wisudaningsih, and M. H. Islam. 2023. "Application of Spatial Temporal Graph Neural Networks for Forecasting Data Time Series River Pollution Waste Content in Probolinggo." In *1st International Conference on Neural Networks and Machine Learning 2022 (ICONNSMAL 2022)*, 257–272. Atlantis Press.

Ministero dell'Economia e delle Finanze. 2016. "Income and Main Irpef Variables on a Municipal Basis." https://www1.finanze.gov.it/finanze/analisi_stat/public/index.php?search_class[0]=cCOMUNE&opendata=yes.

Mukherjee, A. G., U. R. Wanjari, R. Chakraborty, et al. 2021. "A Review on Modern and Smart Technologies for Efficient Waste Disposal and Management." *Journal of Environmental Management* 297: 113347.

Oreshkin, B. N., D. Carpov, N. Chapados, and Y. Bengio. 2019. "N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting." arXiv Preprint arXiv:190510437.

Puglia, R. 2016. "Produzione di rifiuti solidi urbani (RSU) per comune." https://pugliacon.regione.puglia.it/orp/public/servizi/rsu-per-comune.

Sakoe, H. 1971. "Dynamic-Programming Approach to Continuous Speech Recognition." In *1971 Proc. the International Congress of Acoustics, Budapest*, 43–49. IEEE.

Sakoe, H., and S. Chiba. 1978. "Dynamic Programming Algorithm Optimization for Spoken Word Recognition." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26, no. 1: 43–49.

Shumway, R. H., D. S. Stoffer, R. H. Shumway, and D. S. Stoffer. 2017. "ARIMA Models." In *Time Series Analysis and Its Applications: With R Examples*, 75–163. Wordpress.

Srivastava, A., and P. K. Jha. 2023. "A Multi-Model Forecasting Approach for Solid Waste Generation by Integrating Demographic and Socioeconomic Factors: A Case Study of Prayagraj, India." *Environmental Monitoring and Assessment* 195, no. 6: 768.

Vincenty, T. 1975. "Direct and Inverse Solutions of Geodesics on the Ellipsoid With Application of Nested Equations." *Survey Review* 23, no. 176: 88–93. https://doi.org/10.1179/sre.1975.23.176.88.

Wu, F., D. Niu, S. Dai, and B. Wu. 2020. "New Insights Into Regional Differences of the Predictions of Municipal Solid Waste Generation Rates Using Artificial Neural Networks." *Waste Management* 107: 182–190.