



Signal decomposition–driven deep learning framework for passenger demand forecasting: Evidence from Dublin Luas

Afaq Khattak^{*}, Brian Caulfield

Centre for Transport Research, Department of Civil, Structural and Environmental Engineering, Trinity College Dublin, College Green, Dublin 2, Ireland

HIGHLIGHTS

- Signal decomposition–driven deep learning for multi-horizon Dublin Luas passenger demand forecasting.
- VMD+BiGRU outperforms other deep learning models.
- Statistical tests confirm proposed framework is statistically significant.

ARTICLE INFO

Keywords:

Public transport
Time series analysis
Signal processing
Deep learning

ABSTRACT

The Dublin Luas is an urban light rail system that exhibits complex and dynamic passenger demand, which poses significant challenges for planning and scheduling decisions. Effective management of the system requires reliable passenger demand forecasts across short-, medium-, and long-term horizons to support operational control, scheduling, and capacity investment decisions. This study proposes a hybrid Artificial Intelligence (AI) framework that integrates signal processing and deep learning for multi-horizon passenger demand forecasting. The framework first applies Variational Mode Decomposition (VMD) to the weekly passenger demand series from the Dublin Luas Red and Green lines. This decomposes the data into distinct temporal components. These components are then modelled with Bidirectional Gated Recurrent Units (BiGRU) to capture their frequency-specific temporal dependencies. The BiGRU hyperparameters are optimized using the Tree-structured Parzen Estimator (TPE) algorithm. Results demonstrate that the proposed hybrid VMD+BiGRU framework outperforms the baseline models. On the Red Line, it achieved an RMSE of 7.88, an MAE of 5.85, and an R^2 of 0.976. On the Green Line, it achieved an RMSE of 9.65, an MAE of 8.47, and an R^2 of 0.958. Wilcoxon signed-rank tests further confirmed the superiority of the proposed framework. To further evaluate model robustness and temporal transferability, a multi-horizon forecasting analysis was conducted for short-term (1-week), medium-term (4-week), and long-term (12-week) intervals. The VMD+BiGRU model maintains low error growth and high predictive accuracy across all horizons compared with competing models. The proposed framework provides reliable forecasts that support resource allocation, timetable adjustment, and operational management of the Dublin Luas system.

1. Introduction

The Luas, meaning "speed" in Irish, serves as the light rail system for the Dublin area and plays a central role in its public transportation framework [1]. It provides high-capacity, reliable mobility along key urban corridors, supports daily commuting for work and education, and reduces dependence on private cars [2,3]. In addition, it contributes to congestion reduction, lower emissions, and improved accessibility, which makes it a critical component of sustainable urban mobility in the

city. Among the various public transportation systems in Dublin, Transport Infrastructure Ireland (TII) has reported the increasing significance and operational importance of the Dublin Luas within the city's public transport network [4], as shown in Fig. 1. The report indicates that the Dublin Luas carries approximately one million passenger trips per week. Around 75% of passengers use the Luas at least twice per week, with nearly two-thirds traveling for work, school, or college. In addition, it accounts for one in every five public transport trips within the Dublin area.

^{*} Corresponding author.

E-mail addresses: akhatak@tcd.ie (A. Khattak), brian.caulfield@tcd.ie (B. Caulfield).

<https://doi.org/10.1016/j.asoc.2026.115245>

Received 17 June 2025; Received in revised form 12 April 2026; Accepted 13 April 2026

Available online 15 April 2026

1568-4946/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Since it began operations in 2004 under the management of Transdev, the network has expanded with the Red Line and the Green Line, together serving a total of 67 stations. The Red Line operates between Tallaght or Saggart and Connolly Station or The Point, with key interchange and Park and Ride facilities at Cheeverstown, Red Cow, and Tallaght. The Green Line runs between Brides Glen or Sandyford and Parnell or Broombridge, with Park and Ride access at Carrickmines, Sandyford, Stillorgan, and Balally, as shown in Fig. 2.

Over the past decade, passenger demand on the Dublin Luas has increased by about 58%. To meet this demand, the network deploys 58 tram units during the morning peak, carrying up to 18,000 passengers per hour [4]. Central Statistics Office (CSO) of Ireland provides detailed statistics on Dublin Luas passenger demand [5]. It has been reported that prior to the COVID-19 pandemic, combined passenger demand across both the Red and Green lines reached a peak in 2019, with a total of 48.38 million passenger journeys. Both lines had severe declines in 2020, almost 60% below the base level, followed by only slight changes in 2021, as shown in Fig. 3. A strong rebound appeared in 2022, when the Red Line rose by about 109% and the Green Line by about 95%. The pace of growth slowed afterward and reached modest positive rates in 2023 and 2024. These statistics show a complete post-pandemic recovery phase and a steady rise in passenger demand across the Dublin Luas network. Similarly, Fig. 4 shows the year-on-year growth rate in Dublin Luas passenger journeys for the same period, with 2019 as the base year.

Moreover, passenger demand is expected to increase, which is likely to place additional operational pressure on the core sections of the existing network. TII has already highlighted limitations within the current network, noting that one line is already operating at or over capacity. Passenger demand is expected to reach approximately 11 million annual journeys by 2035 and around 15 million by 2050 [6]. To accommodate the growing passenger demand, the proposed Luas network includes the development of new lines that extend the network into suburban areas. Future expansion plans include the proposed Luas Lucan line, as illustrated in Fig. 5.

The rise in passenger demand signals a significant shift in urban mobility and a greater reliance on high-capacity public transportation across the city. This growth places additional pressure on Dublin Luas service planning and operations. Therefore, modelling of passenger demand plays a key role in effective service planning, optimal capacity allocation across different demand periods, and improved operational management and network resilience. Reliable demand forecasts also serve as a key input for future network expansion, which supports long-term infrastructure investment, network development strategies, and the design of new routes and services.

1.1. Motivation and significance of the study

Passenger demand on the Dublin Luas exhibits complex non-linearity, non-stationarity, and temporal variability, which limits the effectiveness of traditional forecasting strategies and requires advanced time series models. The demand signal contains substantial noise, driven by irregular patterns and abrupt fluctuations associated with public holidays, events, weather conditions, and operational disruptions. Such characteristics lead to error propagation and reduced prediction reliability, with significant impact on both short- and long-term forecast horizons.

While existing studies have investigated passenger demand forecasting for high-speed rail [7–9], and bus systems [10–12], relatively limited attention has been directed toward short- and long-term demand analysis within urban light rail systems, such as tram and Luas-type networks. This gap is non-trivial, as these systems operate under distinct demand regimes shaped by dense urban interactions, short trip lengths, clear weekday–weekend contrasts, and high sensitivity to external disruptions. Moreover, most forecasting studies focus on single-step or short-term prediction, which restricts their practical value for medium- and long-term transit planning. To address the complex characteristics of passenger demand time series, prior research has incorporated signal decomposition techniques such as Empirical Mode Decomposition and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise in combination with deep learning models such as Long Short-Term Memory (LSTM), and Bidirectional LSTM (BiLSTM), and Gated Recurrent Unit (GRU) neural network to improve forecasting performance [13–16]. Although these hybrid strategies facilitate noise attenuation and multi-scale feature extraction, they remain susceptible to mode mixing, endpoint distortions, and sensitivity to noise, which undermine forecast reliability and temporal transferability. These limitations intensify when forecasts extend beyond short horizons.

To address these shortcomings, Variational Mode Decomposition (VMD) acts as an alternative signal processing method that applies a mathematical decomposition framework, which leads to improved mode separation and reduced sensitivity to noise [17,18]. This aspect makes VMD suitable for modelling complex and non-stationary passenger demand patterns. When this approach pairs with bidirectional recurrent architectures such as Bidirectional Gated Recurrent Units (BiGRU) [19], it facilitates frequency-specific modelling of demand dynamics while it preserves temporal dependencies in both forward and backward directions. The significance of this study lies in the advancement of a signal processing–driven deep learning framework. This approach directly addresses the structural complexity of passenger demand on the Dublin Luas and provides accurate and stable multi-horizon

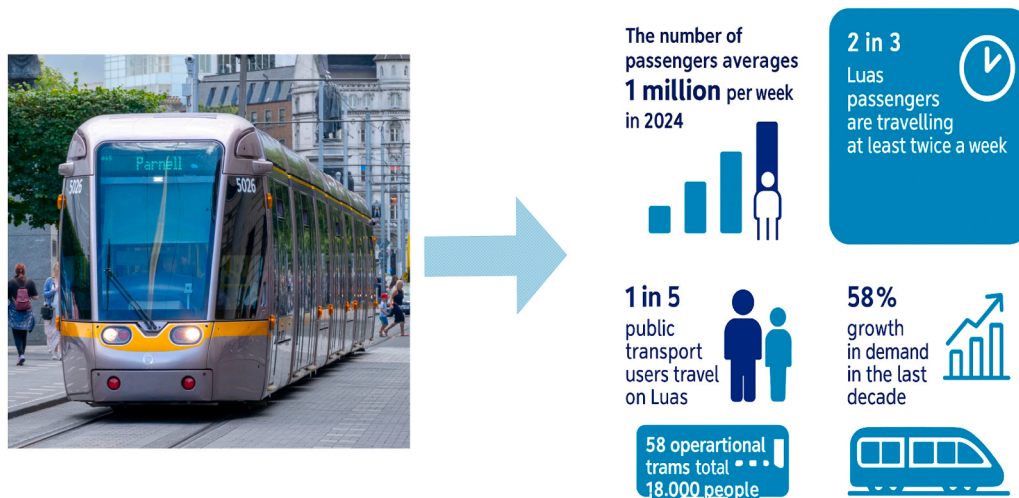


Fig. 1. Dublin Luas operational and ridership statistics in 2024 (Source: [4]).

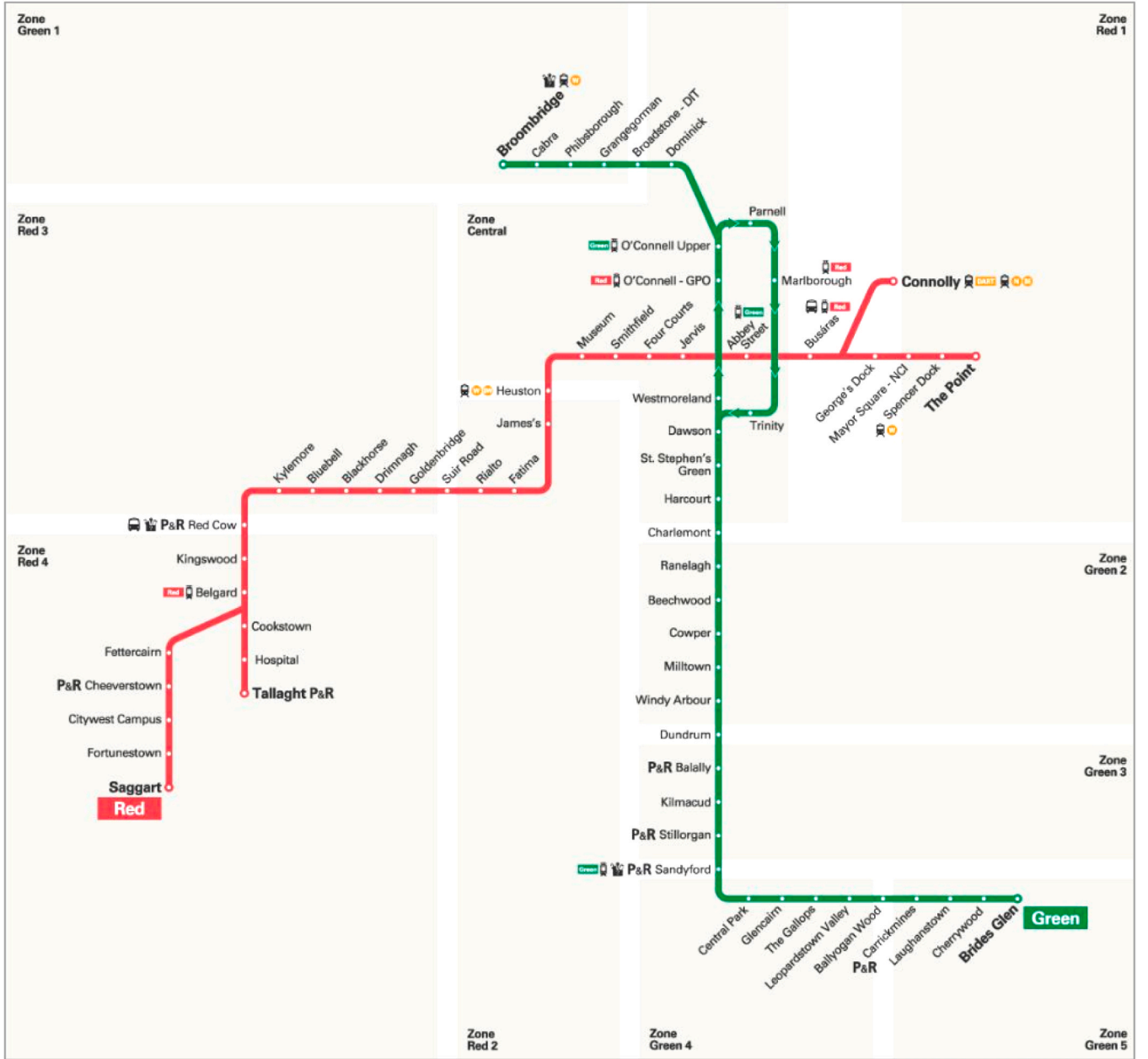


Fig. 2. Dublin Luas Red and Green line network with stations.

forecasts. The work contributes an operational solution for improving demand prediction in urban rail systems.

1.2. Study contributions and scope

This study applies an optimized VMD-based decomposition strategy integrated with BiGRU to improve passenger demand forecasting. By addressing noise, abrupt fluctuations, and multi-scale temporal variability, the proposed framework produces more reliable passenger demand forecasts across different time horizons. The main contributions are as follows:

1. An optimized VMD-based signal processing strategy is proposed through the integration of Fast Fourier Transform and Mutual Information (FFT-MI) to decompose Dublin Luas passenger demand into multiple Intrinsic Mode Functions (IMFs) [20,21], with each IMF representing a distinct frequency component. Compared with
- conventional EMD and CEEMDAN approaches, the proposed strategy reduces mode mixing, yields smoother and more orthogonal IMFs, and mitigates endpoint effects and noise sensitivity that commonly affect EMD-based methods.
2. Frequency-specific modelling of passenger demand is achieved through the application of BiGRU to each IMF [22]. By processing temporal information in both forward and backward directions, BiGRU captures asymmetric and long-range temporal dependencies more effectively than unidirectional GRU and conventional LSTM architectures, while requiring fewer parameters [23]. This structure improves representation of uneven demand peaks, abrupt variations, and recurring cycles within passenger demand time series [24].
3. Efficient hyperparameter optimization is introduced through the integration of the Tree-structured Parzen Estimator (TPE) algorithm during BiGRU training [25]. TPE adopts a Bayesian optimization scheme that models conditional hyperparameter distributions by separating high-loss and low-loss regions, then sampling new

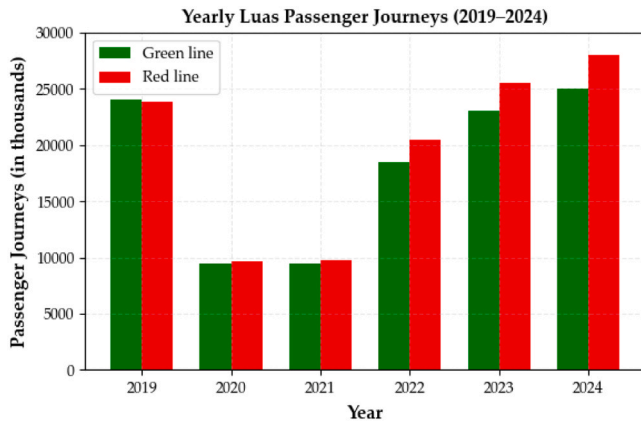


Fig. 3. Annual passenger journeys on the Dublin Luas Red and Green lines from 2019 to 2024 (Source: [5]).

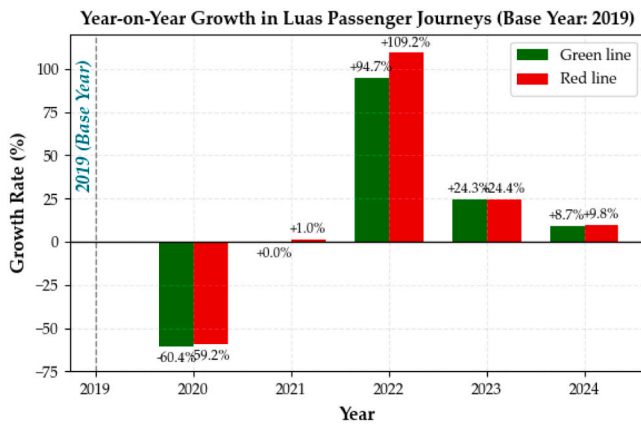


Fig. 4. Year-on-year growth rate in Dublin Luas passenger journeys (Source: [5]).

configurations from regions associated with lower loss [26]. This strategy facilitates efficient exploration of complex and high-dimensional hyperparameter spaces and has demonstrated greater efficiency than grid search and random search across a wide range of applied deep learning research [27,28]. Prior studies have also reported the effectiveness of TPE for time series forecasting tasks [29,30], as well as public transport and passenger demand modelling [31,32], and traffic flow prediction [27,33].

4. A systematic ablation study is conducted to quantify the contribution of the FFT-MI-based VMD component by comparing BiGRU performance with and without decomposition, as well as against baseline forecasting models. This analysis directly demonstrates the added value of the proposed decomposition-driven framework.

The rest of the paper is organized as follows. Section 2 reviews previous studies related to the application of AI models such as machine learning and deep learning within public transport. Section 3 provides a theoretical overview of the proposed integrated VMD+BiGRU model for forecasting passenger demand on Dublin Luas Red and Green lines. Section 4 presents and compares the performance of the proposed model results to that of baseline models. Finally, Section 5 presents the conclusions of the study, discusses its limitations, and outlines directions for future research.

2. Related work

Reliable short- and long-term passenger demand forecasting forms a

strong basis for effective network planning, service scheduling, and operational management in the public transport sector [12,34]. In the literature, passenger demand forecasting strategies generally fall into two main categories, which comprise statistical approaches and artificial intelligence-based models. In the case of statistical approaches, the Auto-Regressive Integrated Moving Average (ARIMA) and its seasonal variant, the Seasonal Auto-Regressive Integrated Moving Average (SARIMA), have been widely applied to model passenger flow patterns using historical data [35–37]. These approaches rely on structured modelling frameworks that assume data linearity and stationarity. However, passenger flows in public transport systems frequently exhibit non-linear and dynamic behaviors, which are difficult for traditional statistical models to capture accurately [38].

To address the limitations of conventional statistical models, researchers have begun to shift to AI-based models, including machine learning and deep learning, due to their inherent flexibility and adaptability in modelling complex passenger flow patterns [39,40]. Unlike statistical methods, these models do not rely on strict assumptions about data distribution or stationarity, which allow them to capture irregular fluctuations and reveal hidden patterns in passenger flow [41,42]. Recent research has explored various AI models to address key challenges in public transport systems, with emphasis on passenger demand forecasting, anomaly detection, and schedule and delay analysis. Table 1 provides a summary of recent applications of AI models in the public transport system. These developments reflect the growing application of AI models in improving the management and reliability of public transport operations.

Although many AI models perform well in time series forecasting tasks, their effectiveness can be limited for passenger demand data that exhibit strong non-linearity, non-stationarity, and multi-scale temporal patterns [37,38]. Passenger demand shows daily, weekly, and seasonal periodicity, differs between weekdays and weekends, and is sensitive to events, weather, and disruptions. These characteristics produce highly fluctuating time series that standalone AI models often struggle to capture. This illustrates the need for signal decomposition techniques that separate passenger demand series into distinct frequency components for more accurate forecasting.

To take advantage of both signal processing and deep learning, several studies have adopted hybrid forecasting frameworks that integrate decomposition techniques with neural networks across domains such as power and energy systems [57,58], traffic and transportation modelling [30,59,60], and weather and natural hazard prediction [61–63]. However, their application to passenger demand forecasting in light rail systems remains limited. Current research is still relatively new, with only a small number of studies that focus on exhaustive multi-horizon analysis in urban light rail networks.

3. Research method

This section details the methodological framework formulated to estimate and forecast passenger demand patterns within the Dublin Luas. It begins with the development of the proposed hybrid forecasting framework that integrates VMD, BiGRU, and TPE optimization. The subsequent subsections describe each component of the framework and explain the model training and reconstruction processes. They also introduce benchmark models and evaluation techniques used for comparative analysis and statistical validation.

3.1. Theoretical overview of integrated VMD+BiGRU framework

This study presents an integrated forecasting framework that combines VMD with the BiGRU model, with optimization through TPE, for estimation and multi-horizon forecasting of passenger demand on the Dublin Luas network. The framework consists of three main steps. First, optimized VMD decomposes the original Dublin Luas passenger demand time series into a set of IMFs, where each IMF represents a distinct

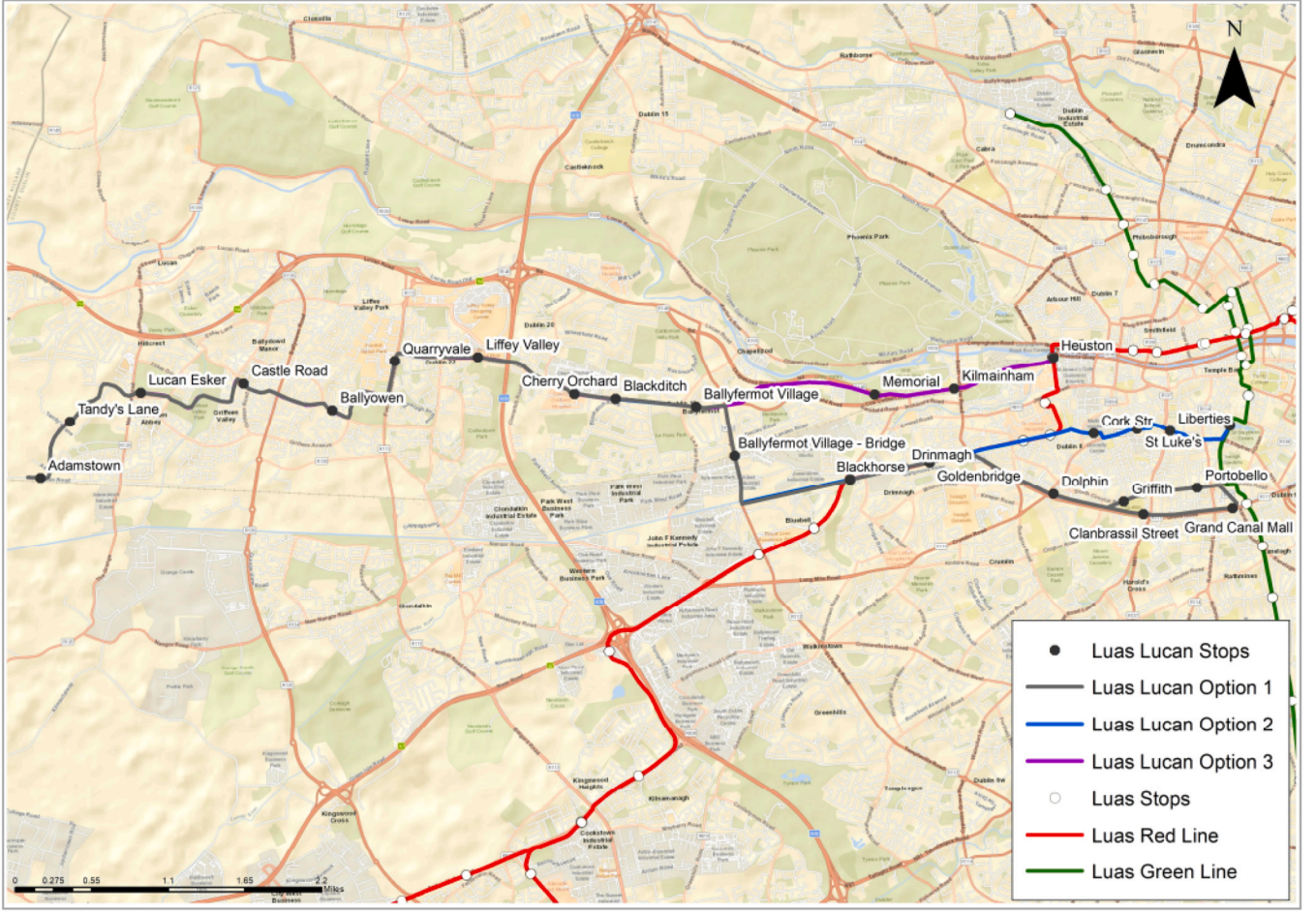


Fig. 5. Proposed route corridors for the Dublin Luas extension (Source: [6]).

frequency characteristic of passenger flow. This decomposition improves signal clarity by isolating localized temporal patterns and reducing noise. Second, each IMF is modelled separately through a BiGRU network that captures temporal dependence in both forward and backward directions. TPE optimization then calibrates the hyperparameters of each BiGRU model to improve predictive accuracy across forecast horizons. Third, the signal reconstruction step combines the predicted IMFs to generate the final passenger demand forecasts. Fig. 6 presents the structure of the proposed framework. The key stages and the underlying working mechanism of the hybrid framework are described below in a stepwise manner.

structures. VMD seeks to decompose a real-valued signal $x(t)$ into K band-limited modes $\{u_k(t)\}_{k=1}^K$, each centered around an estimated frequency ω_k . The constrained variational problem is expressed as Eq. 1.

$$\min_{\{\mu_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \text{ s.t. } \sum_{k=1}^K \mu_k(t) = x(t) \quad (1)$$

To solve this constrained variational problem, an augmented Lagrangian formulation is introduced as shown in Eq. 2. This reformulation facilitates the decomposition of the original optimization task into sub-problems that are more tractable for numerical computation.

$$\begin{aligned} \varpi(\{\mu_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| x(t) - \sum_{k=1}^K \mu_k(t) \right\|_2^2 \\ & + \left\langle \lambda(t), x(t) - \sum_{k=1}^K \mu_k(t) \right\rangle \end{aligned} \quad (2)$$

3.1.1. Signal decomposition via VMD strategy

The first stage of the proposed hybrid framework serves as a signal preprocessing tool that decomposes the original passenger time series into a finite set of IMFs, each associated with a distinct frequency band. This process improves feature separability, reduces noise interference, and enhances the capacity of model to learn underlying temporal

Where; α is the quadratic penalty factor, $\lambda(t)$ is the Lagrange multiplier, $*$ denotes convolution, $\delta(t)$ is the Dirac delta function, and j is the imaginary unit.

Although Eqs. 1–2 describe how VMD decomposes a signal for a given number of modes K and penalty factor α , the quality of the decomposition depends strongly on how these two parameters are

Table 1
Summary of recent AI models application in public transport.

Aspects	Description	Model	Ref.
Passengers flow pattern	Short term forecasting of metro passenger demand	STL-LSTM	[43]
	Metro passenger flow analysis to capture daily and peak-hour inflow and outflow patterns in the station	DeepPF	[44]
	Bus passenger demand forecasting captures boarding patterns and temporal demand variation across routes.	BiLSTM	[12]
	Forecasting hourly passenger demand along Bus Rapid Transit (BRT) corridors to improve traffic management and resource allocation	SAE-DNN	[41]
	Multi-horizon passenger demand forecasting for urban rail transit system	BiLSTM-CTPA (BO-optimized)	[45]
	Short-term passenger demand forecasting for urban rail transit system to improve service scheduling and operations	TCN-LSTM	[46]
	Short-term passenger flow forecasting at metro stations to manage congestion and train dispatch.	Spatio-temporal LSTM	[47]
	Time series analysis of rail transit devices detects irregular patterns and early signs of equipment malfunction.	LSTM	[48]
	Network-level passenger flow anomaly assessment that identifies unusual passenger volume variation across stations or lines.	GCN-informer	[49]
	Identification of fraudulent activities in public transport analyses ticketing data to detect fare evasion and abnormal travel behavior.	IF	[50]
Anomaly Detection	Identification of passenger flow anomalies and assessing their alarm level focuses.	Hybrid K-Mean with HC	[51]
	Analysis of recurrent delay patterns across public transport routes to identify operational inefficiencies.	K-Mean clustering with ResNet50	[52]
Schedule and Delay Pattern	Network-level bus delay forecasting across multiple routes to improve scheduling and fleet management.	TabNet	[53]
	Network-wide rail transit delay analysis to identify system-wide congestion sources.	BN	[54]
	Tram delay forecasting to assess short-term service disruption and timing deviation	RF	[55]
	Rail transit delay assessment to evaluate causes of delays and improve service reliability	ISVM	[56]

Abbreviations; STL: Seasonal-Trend Decomposition using Loess; SAE: Stacked Auto encoder; DNN: Deep Neural Network; TCN: Temporal Convolutional Network; GCN-Informer: Graph Convolutional Network combined with Informer; IF: Isolated Forest; HC: hierarchical clustering; TabNet: Attentive Interpretable Tabular Network; BN: Bayesian Network; RF: Random Forest; ISVM: Improved Support Vector Machine.

selected. Inappropriate values of α may lead to mode mixing, information loss, or over-decomposition [64,65]. Therefore, an outer optimization layer is required to determine K and α in a data-driven manner. In this study, this is achieved through the integration of FFT-MI strategies [66]. The steps of the FFT-MI-based optimization of VMD are described below.

3.1.1.1. Step 1: Determination of the number of modes K by FFT. The original passenger demand time series $x(t)$, $t = 1, 2, \dots, T$, is transformed into the frequency domain by FFT to identify its dominant oscillatory components. Each significant peak in the frequency spectrum corresponds to a distinct frequency band in the signal, and the number of such peaks determines the optimal number of IMFs K , which controls under-decomposition and prevents redundant mode splitting. The discrete Fourier transform of $x(t)$ is computed as Eq. 3.

$$X(f_m) = \sum_{t=1}^T x(t) e^{-j2\pi f_m t / T}, m = 1, 2, \dots, T-1 \quad (3)$$

Where, $f_m = m/T$ denotes the discrete frequency and $j = \sqrt{-1}$.

The corresponding amplitude spectrum is Eq. 4, where the magnitude of the Fourier transform represents the contribution strength of each frequency component within the signal and provides a clear description of its frequency-domain characteristics.

$$A(f_m) = |X(f_m)| \quad (4)$$

Based on the amplitude spectrum, the dominant frequency components of the signal are identified as a set of feature frequencies, as defined in Eq. (5). These frequencies correspond to local maxima whose amplitudes exceed a predefined threshold, indicating their significance in the overall signal structure.

$$\Omega = \{f_i | A(f_i) > \tau, f_i \text{ is a local maximum of } A(f_i)\} \quad (5)$$

Where, τ is the threshold used to remove insignificant spectral fluctuations. Each $f_i \in \Omega$ corresponds to a principal oscillatory component of the passenger demand time series. The optimal number of decomposition modes is therefore determined as shown in Eq. 6.

$$K = |\Omega| \quad (6)$$

This FFT-based feature-frequency analysis allows the number of IMFs to align with the inherent spectral structure of the signal, thereby avoiding under-decomposition and unnecessary mode splitting.

3.1.1.2. Step 2: VMD decomposition for a trial penalty factor α . With the number of modes K fixed, VMD is applied using a trial value of α to decompose the original signal into K components $\{IMF_k(t)\}_{k=1}^K$. Each $IMF_k(t)$ represents a band-limited component with distinct frequency behavior. The penalty factor α controls the bandwidth of each mode and regulates the trade-off between spectral compactness and signal reconstruction accuracy.

3.1.1.3. Step 3: MI computation for individual IMFs. For each decomposed component, the MI between the original signal $x(t)$ and each $IMF_k(t)$ is computed as Eq. 7.

$$MI_k = MI(x(t), IMF_k(t)), k = 1, 2, \dots, K \quad (7)$$

This measure quantifies the amount of information from the original passenger demand series that is retained within each mode. The IMF that carries the highest information content is then identified by selecting the maximum MI value across all components, as expressed in Eq. 8.

$$MI_{\max} = \max(MI_1, MI_2, \dots, MI_K) \quad (8)$$

Eq. 8 identifies the IMF that retains the highest information content relative to the original signal and provides a local indicator of decomposition quality.

3.1.1.4. Step 4: Signal reconstruction and fidelity evaluation. The reconstructed signal is obtained by summing all extracted IMFs, which restores the complete passenger demand signal from its decomposed components, as expressed Eq. 9.

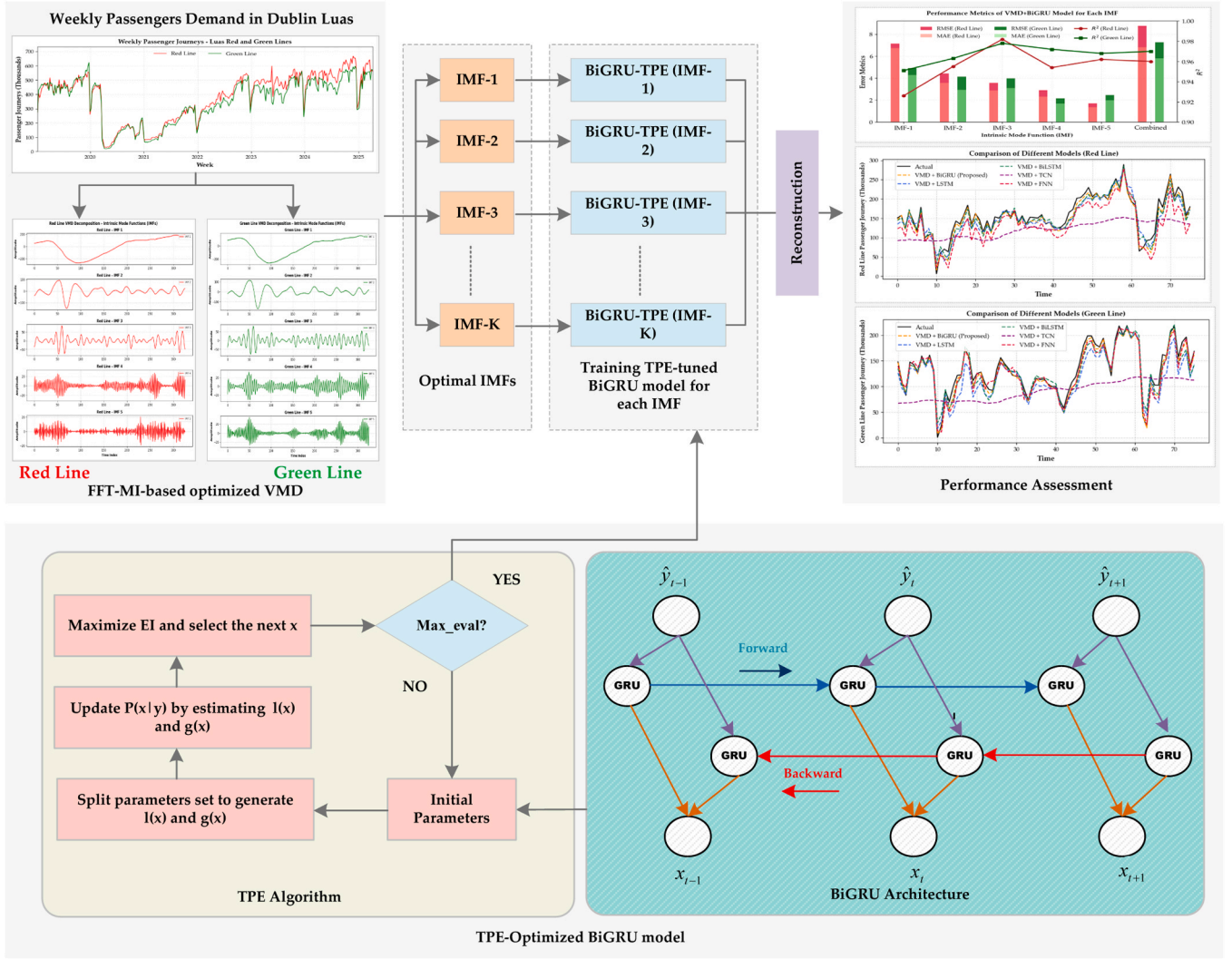


Fig. 6. Overview of the proposed hybrid VMD-BiGRU framework for Dublin Luas passenger demand forecast.

$$x'(t) = \sum_{i=1}^K IMF_i(t) \quad (9)$$

This reconstruction represents the full signal recovered from the VMD components. To evaluate how well the decomposition preserves the original passenger demand series, the reconstruction fidelity is quantified through the MI between the reconstructed signal and the original signal, as illustrated in Eq. 10.

$$\beta_\alpha = MI(x(t), x'(t)) \quad (10)$$

A higher β_α value from Eq. 10 indicates improved reconstruction fidelity and reduced information loss throughout the decomposition process. The quantity β_α reflects the global information similarity between the reconstructed and original series. A larger value of β_α indicates that the selected penalty factor α yields a decomposition that retains more of the original signal structure with lower information loss, which provides the global fidelity measure required for the composite evaluation in Step 5.

3.1.1.5. Step 5: Composite evaluation index. The IMF-level MI from Step 3 and the reconstruction-level MI from Step 4 describe two complementary aspects of the decomposition quality. The former reflects how well the dominant IMF captures the key oscillatory content of the signal, while the latter evaluates how faithfully the full set of IMFs reconstructs the original passenger demand series. To jointly account for both local

and global information preservation, a composite evaluation index γ_α is defined as Eq. 11.

$$\gamma_\alpha = \beta_\alpha \times MI_\alpha^{\max} \quad (11)$$

Where, $MI_\alpha^{\max}$ denotes the maximum MI between the original signal and the individual IMFs under penalty factor α , and β_α represents the MI between the reconstructed signal and the original series.

3.1.1.6. Step 6: Determination of the optimal penalty factor. The composite index γ_α is evaluated for all candidate values of α . Each value of α corresponds to a different trade-off between mode separation and reconstruction fidelity. The optimal penalty factor is selected as the value that maximizes the composite index, as given by Eq. 12.

$$\alpha^* = \underset{\alpha}{\operatorname{argmax}} \gamma_\alpha \quad (12)$$

This choice identifies the decomposition that jointly preserves the dominant information content within the IMFs and maintains the overall structure of the passenger demand series after reconstruction.

3.1.1.7. Step 7: Final VMD decomposition. With the optimized values of K and α , VMD is applied to the original passenger demand time series to obtain the final set of IMFs, as shown in Eq. 13.

Table 2

Pseudocode of the FFT–MI–based VMD parameter selection algorithm.

ALGORITHM: FFT–MI–based VMD Parameter Optimization**Input:** Passenger demand series $x(t)$; candidate penalty factors $\alpha \in A$; , number of modes K **Output:** Optimal penalty factor α^* ; final IMFs $\{IMF_k(t)\}_{k=1}^K$

```

1:   Compute FFT of  $x(t)$ ;
2:   Compute amplitude spectrum  $A(f) = |FFT(x(t))|$ 
3:   Identify dominant frequency peaks
4:   Set  $K \leftarrow$  number of dominant peaks
5:   Initialize:
6:        $\gamma_{best} \leftarrow -\infty$ ,  $\alpha^* \leftarrow \emptyset$ 
7:   for each  $\alpha \in A$  do
8:       Perform VMD on  $x(t)$ ; with parameters  $(K, \alpha)$ 
9:       Obtain IMFs:  $\{IMF_1(t), IMF_2(t), \dots, IMF_K(t)\}$ 
10:      for  $K = 1$  to  $K$ , do
11:          Compute  $MI_k \leftarrow MI(x(t), IMF_k(t))$ 
12:      end for
13:       $MI_{max} \leftarrow \max(MI_1, MI_2, \dots, MI_K)$ 
14:      Reconstruct signal:  $\hat{x}(t) \leftarrow \sum_{k=1}^K IMF_k(t)$ 
15:      Compute reconstruction similarity:  $\rho = MI(x(t), \hat{x}(t)) / MI(x(t), x(t))$ 
16:      Compute composite evaluation index:  $\gamma = MI_{max} \times \rho$ 
17:      If  $\gamma > \gamma_{best}$  then
18:           $\gamma_{best} \leftarrow \gamma$  as  $\alpha^* \leftarrow \alpha$ 
19:      end if
20:  end for
21:  Apply VMD with parameters  $(K, \alpha^*)$  to obtain the final IMFs
22:  Return  $\alpha^*$  and  $\{IMF_k(t)\}_{k=1}^K$ 

```


$$x(t) = \sum_{i=1}^{K^*} IMF_i(t) = \sum_{i=1}^{K^*} A_i(t) \cos(\phi_i(t)) \quad (13)$$

Where; $A_i(t)$ is the amplitude envelope of the i^{th} IMF, and $\phi_i(t)$ is the instantaneous phase.

These IMFs are subsequently used for forecasting. The pseudocode of the algorithm is provided in Table 2.

3.1.2. Individual IMF modelling using BiGRU optimized via TPE

Once the passenger demand time series $x(t)$ is decomposed into K IMFs through VMD, each $IMF_i(t)$ captures a distinct frequency band of the original passenger demand series. Each $IMF_i(t)$ is individually modelled using a BiGRU network to predict its future values. The BiGRU processes time series in forward and backward directions and captures long-term dependencies better than the unidirectional GRU. Let the input sequence for the i^{th} IMF be $X^{(i)} = \{x_1^{(i)}, x_2^{(i)}, \dots, x_T^{(i)}\}$ where, $x_t^{(i)} = IMF_i(t)$ denotes the value of the i^{th} IMF at time t . The goal is to generate the predicted sequence $\hat{Y}^{(i)} = \{\hat{y}_1^{(i)}, \hat{y}_2^{(i)}, \dots, \hat{y}_T^{(i)}\}$, approximating the ground truth values $Y^{(i)}$. This formulation supports direct multi-horizon forecasting and allows assessment of predictive performance at different lead times. At each time step t , the BiGRU processes the input from both temporal directions as shown in Fig. 7. The forward hidden state for the i^{th} IMF is computed as Eq. 14 and simultaneously, the backward hidden state is derived by Eq. 15.

$$\vec{h}_t^{(i)} = GRU_{fwd}(x_t^{(i)}, \vec{h}_{t-1}^{(i)}) \quad (14)$$

$$\overleftarrow{h}_t^{(i)} = GRU_{bwd}(x_t^{(i)}, \overleftarrow{h}_{t+1}^{(i)}) \quad (15)$$

Then, these two hidden states are concatenated to form a complete bidirectional representation that jointly captures past and future temporal information within the IMF, as shown in Eq. 16.

$$h_t^{(i)} = [\vec{h}_t^{(i)}; \overleftarrow{h}_t^{(i)}] \quad (16)$$

The output prediction at each forecast horizon is obtained by passing the hidden representation through a fully connected layer, as illustrated in Eq. 17.

$$\hat{y}_t^{(i)} = W_0 h_t^{(i)} + b_0 \quad (17)$$

Where, $W_0 \in \mathbb{R}^{1 \times 2h}$ and $b_0 \in \mathbb{R}^{1 \times 2h}$ represent the learnable parameters of the output layer.

The model is trained to minimize the Mean Squared Error (MSE) between the actual and predicted values of each IMF, defined as Eq. 18. This formulation allows each decomposed sub-series to contribute independently to the learning process and permits individual adjustment during the optimization stage.

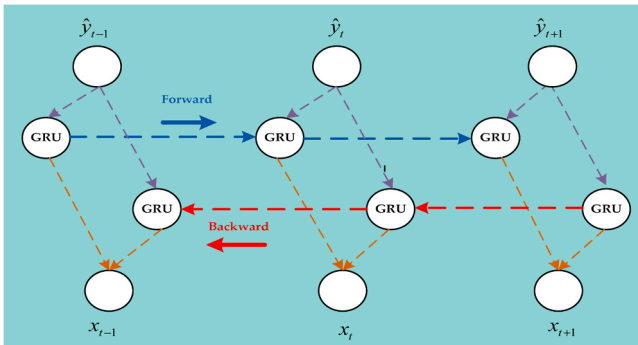


Fig. 7. Structural Architecture of the BiGRU Model.

$$\tau^{(i)} = \frac{1}{T} \sum_{t=1}^T (y_t^{(i)} - \hat{y}_t^{(i)})^2 \quad (18)$$

The performance of the BiGRU model is highly influenced by its hyperparameters. Therefore, for each IMF component, the corresponding BiGRU model is individually fine-tuned using the TPE algorithm, which systematically explores the hyperparameter space to obtain the optimal predictive configuration [67]. For instance, for the BiGRU model trained on a given i^{th} IMF, the set of hyperparameters subject to optimization is represented in Eq.19.

$$\theta^{(i)} = \{u, \eta, b\} \quad (19)$$

Where; $\theta^{(i)} \in \Theta$ represents a candidate configuration from the hyperparameter search space Θ . u denotes the number of GRU units, η is the learning rate, and b is the batch size.

The full hyperparameter configuration for the BiGRU models is provided in Table 3. The objective is to identify the optimal set $\theta^{(i)*}$ that minimizes the validation loss $\tau^{(i)}$, defined earlier in Eq. 18, for the corresponding IMF. This optimization problem can be expressed as Eq. 20.

$$\theta^{(i)*} = \underset{\theta^{(i)} \in \Theta}{\operatorname{argmin}} \tau^{(i)}(\theta^{(i)}) \quad (20)$$

The TPE algorithm models the objective function $\tau^{(i)}$ as a probabilistic process by constructing two density estimators i.e., $p(\theta|\tau)$ is the conditional probability of hyperparameter configurations given observed losses, and $p(\tau)$ is the marginal probability over observed loss values. TPE divides observed hyperparameter and loss pairs into two groups as shown by Eq. 21 and Eq. 22.

$$D_{\text{good}} = \{\theta|\tau(\theta) \leq \tau^*\} \quad (21)$$

$$D_{\text{bad}} = \{\theta|\tau(\theta) > \tau^*\} \quad (22)$$

Where; τ^* is a quantile threshold (typically the 10–25th percentile of losses). The algorithm then models the distributions $l(\theta) = p(\theta|\tau \leq \tau^*)$ and $g(\theta) = p(\theta|\tau > \tau^*)$, and samples new candidates that maximize the expected improvement criterion as shown by Eq. 23.

$$EI(\theta) \propto \frac{l(\theta)}{g(\theta)} \quad (23)$$

This strategy raises the probability that subsequent trials arise from high-potential regions of the search space, which accelerates convergence and limits the chance of entrapment in local minima. Application of TPE for each IMF-specific BiGRU allows the framework to adapt to temporal variation within individual components and produces higher predictive precision.

3.1.3. Reconstruction and Final Prediction

After each IMF_i is modelled and forecast through its corresponding BiGRU sub-model, the final stage reconstructs the overall passenger demand series by combining the predicted sub-signals. Let $\hat{y}_t^{(i)}$ represent

Table 3
BiGRU hyperparameter configuration.

Hyperparameter	Description
No. of GRU Units	Defines the BiGRU capacity and ability to capture temporal dependencies.
Batch Size	Influences the stability of gradient updates and convergence speed.
Learning Rate	Controls the step size for weight adjustments during optimization.
Dropout Rate	Acts as a regularization technique to prevent overfitting.
L2 Regularization	Adds a penalty term to the loss function to constrain large weight values and improve generalization.
Training Epochs	Indicates the number of complete passes through the training dataset.

the predicted value at time t for IMF _{i} component. The reconstructed series $\hat{y}_t$, representing the estimated Dublin Luas passenger demand at time t , is obtained by summing all predicted IMF components, as expressed in Eq. 24.

$$\hat{y}_t = \sum_{i=1}^K \hat{y}_t^{(i)} \quad (24)$$

3.2. Competitive models

To evaluate the performance of the proposed framework, this study benchmarks it against both standalone and VMD-integrated models. The comparison includes Bidirectional Long Short-Term Memory (BiLSTM) [68], Temporal Convolutional Network (TCN) [69] and Feedforward Neural Network (FNN) [70] architectures. A brief overview of each model is provided below.

3.2.1. BiLSTM

BiLSTM model extends the standard LSTM architecture by processing time series data in both forward and backward directions, which allows the model to capture complex temporal dependencies from past and future observations at the same time [68]. This bidirectional structure helps the model represent dynamic passenger demand patterns more effectively, especially in cases where future demand affects present travel decisions, such as during peak hours or in response to scheduled events. The memory cell structure of BiLSTM captures long-term dependencies effectively and represents recurrent and time-varying passenger demand in public transport with higher accuracy [71]. A comprehensive review of BiLSTM architectures and their applications appears in [72].

3.2.2. TCN

The TCN model serves as an effective approach for sequence-based tasks, with strong performance in time series forecasting of passenger demand in public transport [73]. Using dilated causal convolutions, the TCN model captures both local and long-range temporal dependencies without the need for recurrent connections. This convolutional structure allows efficient parallel computation and stable gradient flow, which makes the TCN model suitable for capturing complex temporal patterns in passenger demand data. It also performs well in representing periodic, nonlinear, and highly dynamic variations in the data. Further details on the TCN architecture and its applications appear in [74].

3.2.3. FNN

The FNN model is the simplest type of neural network, which consists of an input layer, one or more hidden layers, and an output layer. In this network, information moves in a single direction from the input to the output layer without any cycles or feedback connections. Each neuron in a layer is fully connected to neurons in the subsequent layer, and nonlinear activation functions are used to capture complex relationships within the data. This architecture allows FNNs to approximate a wide range of functions, which make them effective in modelling diverse input-output mappings. Due to their straightforward architecture and ease of application, FNNs serve as standard reference models in many predictive and classification frameworks. Within passenger demand analysis, FNNs have been utilized for metro systems to estimate short-term demand patterns and to evaluate or refine model performance [75,76]. Further details on the FNN architecture and its applications can be found in [77].

3.3. Performance evaluation of the proposed framework

This section presents a quantitative assessment of the proposed forecasting framework by employing standard statistical error metrics and a post-hoc inferential test for comparative model analysis. The

evaluation process incorporates both deterministic accuracy indicators and hypothesis-based statistical inference to verify performance differences across competing models.

3.3.1. Prediction accuracy metrics for passenger demand forecasting

To assess the performance of the reconstructed signal $\hat{y}_t$, the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2) are used as evaluation metrics. These are defined as follows in Eqs. 25–27.

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (25)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (26)$$

$$R^2 = 1 - \frac{\sum_{t=1}^N (y_t - \hat{y}_t)^2}{\sum_{t=1}^N (y_t - \bar{y})^2} \quad (27)$$

Where; y_t denotes the observed passenger demand at time t , $\hat{y}_t$ represents the reconstructed predicted passenger demand at time t , $\bar{y}_t$ is the mean of observed passenger demand values, and N indicates the total number of time intervals within the evaluation period.

3.3.2. Post-Hoc statistical test for model performance comparison

Although standard performance metrics provide useful information, they may not effectively distinguish models when their prediction performance levels are closely matched. To statistically verify the advantage of the proposed framework over alternative models, the Wilcoxon signed-rank test [78] is applied. This non-parametric method evaluates whether the median of paired differences between two models significantly differs from zero. Let $d_i = x_i - y_i$ denote the difference in prediction errors across n test samples. The test ranks the absolute values $|d_i|$, omits zero differences, and assigns the corresponding sign to each rank. The resulting test statistic W is calculated as expressed in Eq. 28.

$$W = \sum_{i=1}^n \text{sgn}(d_i) \cdot R_i \quad (28)$$

Where; $\text{sgn}(d_i) \in \{-1, 0, +1\}$ is the sign function, and R_i is the rank of $|d_i|$.

The hypothesis formulation is as follows:

- **H₀:** No significant difference exists between the proposed VMD+BiGRU model and other compared models.
- **H₁:** A significant difference exists between the proposed model and the compared model.

At a significance level of $\alpha = 0.05$, the test assesses whether to reject H_0 by comparing the obtained p -value with the corresponding critical threshold.

4. Results and discussion

The Dublin Luas passenger demand dataset is published by TII and is available as open-access data through DATA.GOV.IE [79] and the Central Statistics Office (CSO) [5]. DATA.GOV.IE is Ireland's national open data portal, managed by the Open Data Unit under the Department of Public Expenditure, Infrastructure, Public Service Reform and Digitalization. The portal hosts datasets published and regularly updated by government departments and public bodies, including transport-related statistics. The Dublin Luas passenger demand dataset available on this portal includes monthly and annual aggregated passenger journey data

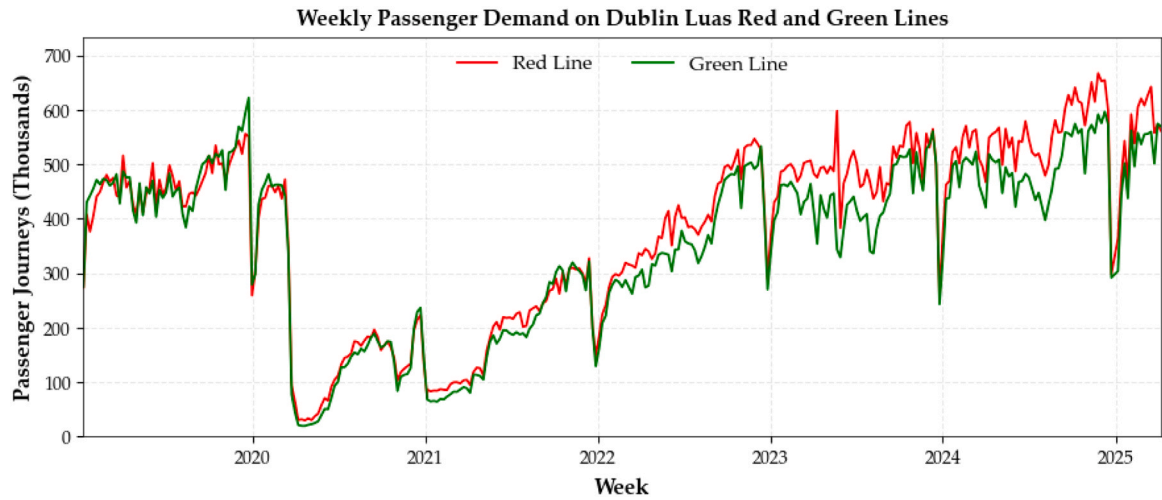


Fig. 8. Weekly passenger demand on the Dublin Luas Red and Green Lines from 2019 to 2025.

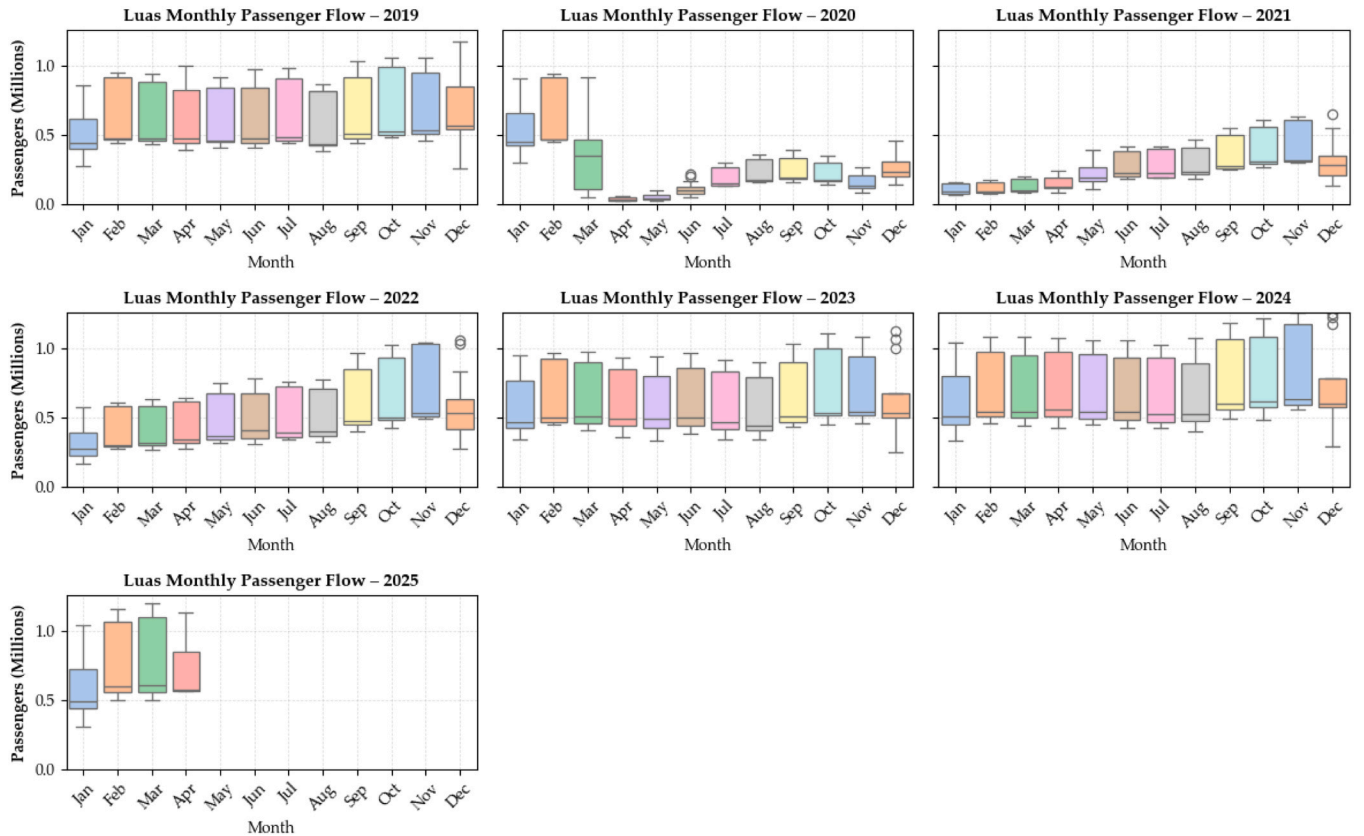


Fig. 9. Monthly distribution of Dublin Luas passenger demand from 2019 to 2025.

for the Dublin Luas network, along with weekly records for the Red and Green Lines that provide finer temporal resolution for analyzing passenger movement trends. It is pertinent to mention that the experiments were implemented in the Python programming language with the TensorFlow/Keras, Scikit-learn, Hyperopt, Matplotlib, and Seaborn packages. All computations were performed on a PC equipped with an Intel® Core™ Ultra 5 125 U processor, 16 GB RAM, and running the Microsoft Windows 11 operating system.

The weekly passenger demand time series exhibits pronounced non-stationary behavior, characterized by irregular fluctuations, abrupt level shifts, and time-varying amplitude, as illustrated in Fig. 8. The series alternates between periods of high and low variability and lacks a stable

or consistent periodic structure, which reflects the influence of external disruptions and evolving travel patterns. Fig. 9 presents box plots from 2019 to 2025 that describe the monthly distribution of Dublin Luas passenger demand and capture the effect of external factors. The plots show strong seasonal variation in 2019, a marked decline after March 2020 due to the COVID-19 outbreak and gradual recovery through 2021 [80–82]. Passenger demand exceeds pre-pandemic levels in 2022 with distinct seasonal peaks, while 2023 and 2024 show high volumes across all months. Moreover, Fig. 10 presents weekly heatmaps for the Red and Green Lines that reflect temporal variation in passenger demand intensity. Light color tones in 2020–2021 indicate reduced passenger activity, while darker tones from 2022 onward indicate higher demand

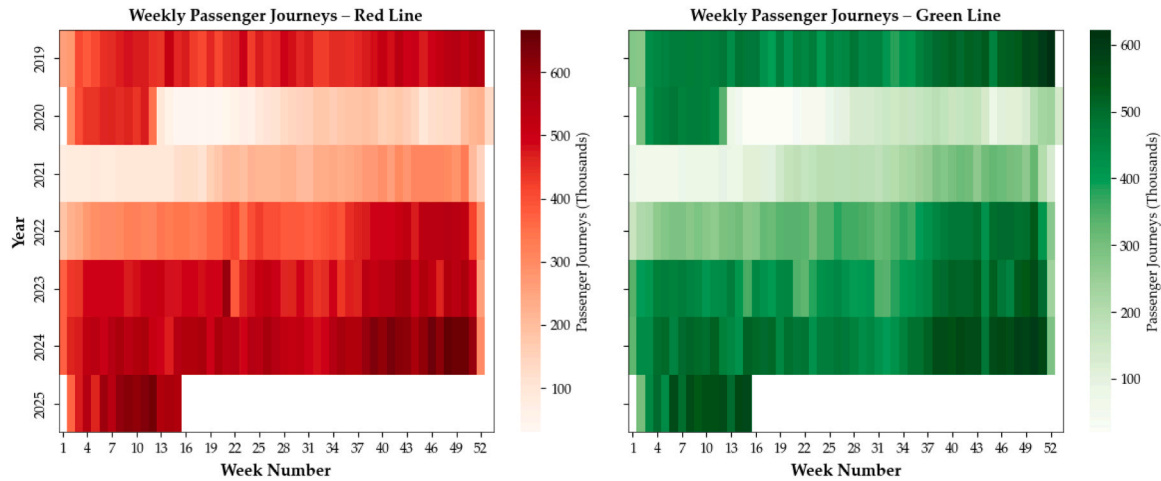


Fig. 10. Weekly passenger demand heatmap for the Dublin Luas Red and Green lines from 2019 to 2025.

Table 4

ADF results for assessing stationarity of weekly Dublin Luas passenger demand.

Line	ADF Statistic	p-value	Critical 1%	Critical 5%	Critical 10%
Red Line	-2.036	0.271	-3.451	-2.871	-2.572
Green Line	-1.652	0.456	-3.451	-2.871	-2.572

intensity and recovery of passenger flow across both routes.

4.1. Stationarity analysis of weekly passenger demand data

To identify irregular patterns in weekly Luas passenger demand, the Interquartile Range (IQR) and rolling z-score methods were applied. No statistical outliers were detected for either the Red or Green Line, which indicated that the major variations, such as those during the COVID-19 period, reflect broader structural shifts rather than isolated anomalies. To assess the temporal characteristics of the Dublin Luas weekly passenger demand data for both the Red and Green Lines, the Augmented Dickey-Fuller (ADF) test was applied [83]. The ADF test results confirm that the weekly passenger journey time series for both the Red and Green Lines are non-stationary as shown in Table 4. The test statistics (−2.036 for Red Line and −1.652 for Green Line) are higher than the

critical values at all conventional significance levels, and the p -values (0.271 and 0.456, respectively) exceed 0.05. The null hypothesis of a unit root is accepted, indicating that both time series exhibit non-stationary behavior.

This outcome reflects the presence of long-term trends and external factors that shape public transport demand over time. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for the Red Line (Fig. 11a) and Green Line (Fig. 11b) also show slow, gradual decay without sharp cutoff and aligns with the non-stationary nature of both series.

4.2. Decomposition of passenger demand signal

The FFT–MI–based VMD strategy was applied to decompose the weekly Dublin Luas passenger demand time series into an optimal set of IMFs both the Red and Green Lines. The optimal number of IMFs was identified as $K = 5$ for both lines. However, the optimal bandwidth constraint and convergence tolerance differed slightly between the two lines. For the Red Line, the optimal bandwidth parameter α was 2400 and the convergence tolerance was 1×10^{-6} . For the Green Line, the optimal α was 2200 and the convergence tolerance was 1×10^{-7} . Fig. 12 (a–b) illustrates the optimal IMFs for the Red Line and Green Line. In both cases, IMF 1 represents the smooth, low-frequency component associated with the long-term demand trend, IMF 2 and

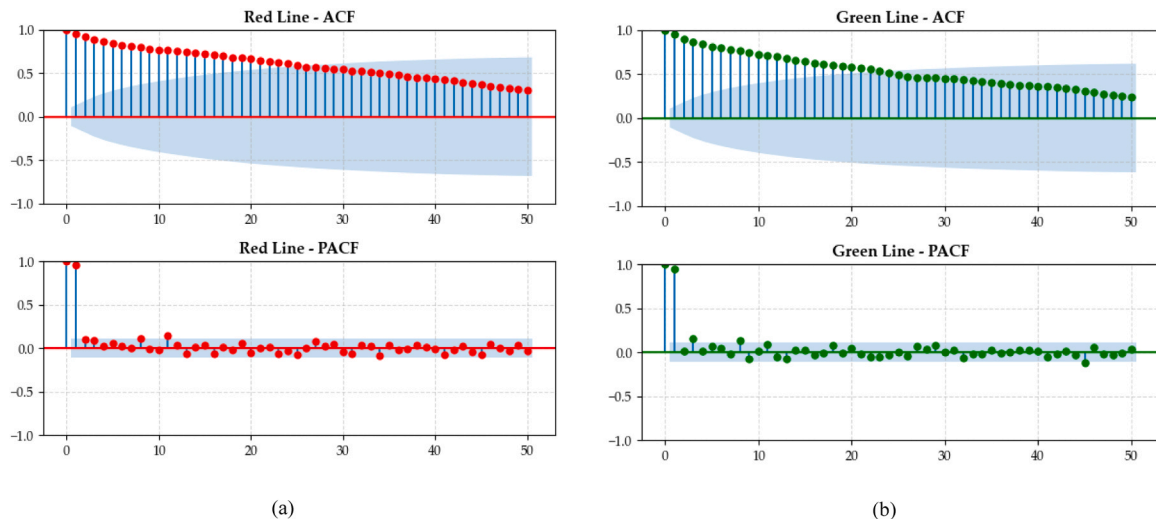


Fig. 11. ACF and PACF plots for the Dublin Luas Passenger demand; (a) Red Luas Line; (b) Green Luas Line.

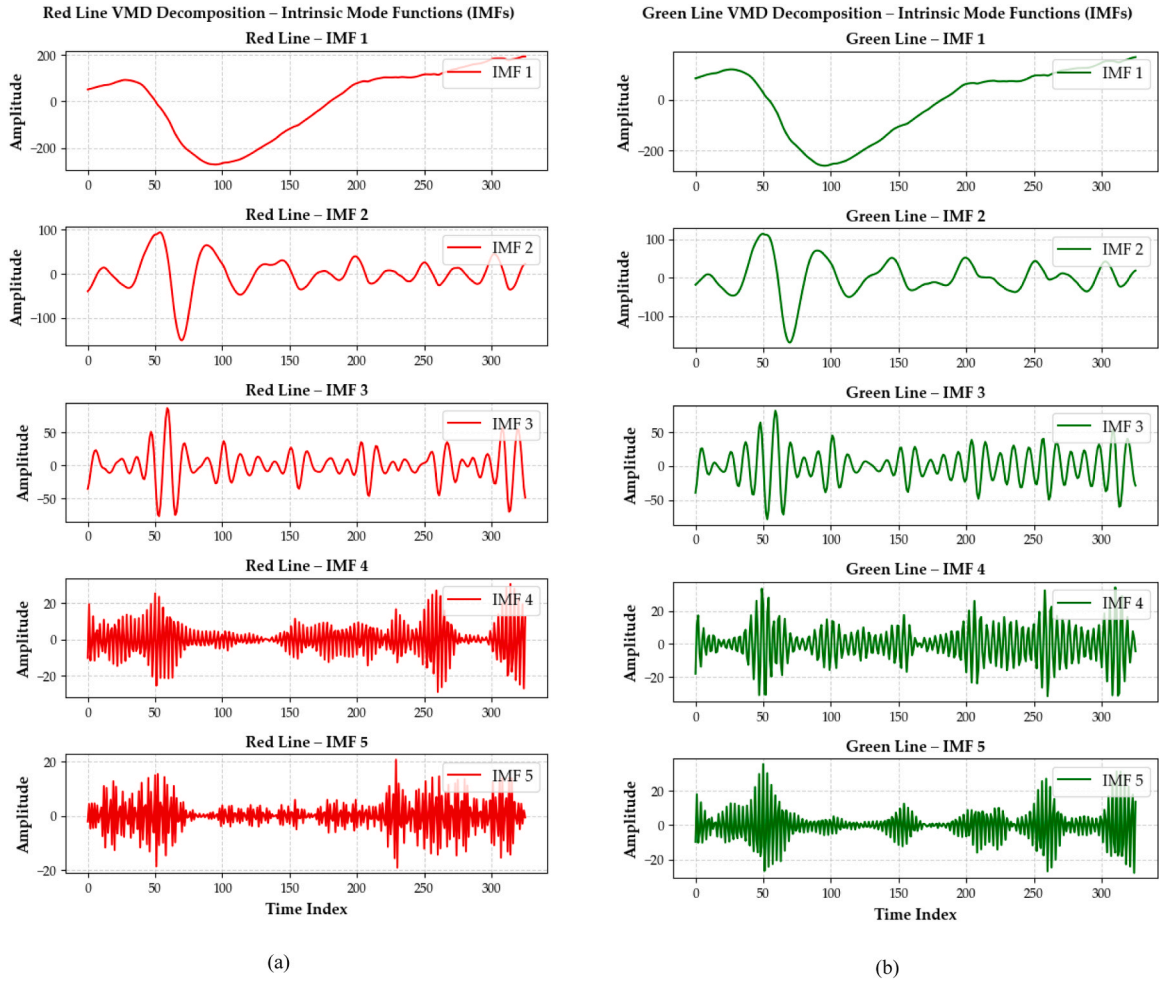


Fig. 12. VMD-based decomposition of the Dublin Luas passenger demand time series into 5 optimal IMFs; (a) Red Luas Line; (b) Green Luas Line.

IMF 3 capture medium-frequency oscillations that align with seasonal cycles, and IMF 4 and IMF 5 depict high-frequency fluctuations linked to short-term operational variations or external calendar effects.

4.3. Hyperparameter tuning of VMD+BiGRU via TPE

Each IMF generated by the FFT-MI-based VMD strategy was modelled with an independent BiGRU network. The hyperparameters of each BiGRU model were optimized separately via the Tree-structured Parzen Estimator (TPE) algorithm. Following standard time-series practice, the dataset was partitioned chronologically into 75% training data and 25% testing data. This proportion allocates most observations to temporal pattern learning while reserving a substantial segment for out-of-sample performance evaluation [84–89]. Although alternative ratios may be adopted, this split provides a practical balance between learning capacity and predictive reliability.

In this study, the hyperparameter search space was defined to cover a comprehensive set of architectural and training hyperparameters of the BiGRU model. The number of GRU units was explored between 32 and 128 with a step size of 16. The dropout rate was treated as discrete within the range of 0.1–0.5 with a step size of 0.1. The batch size was selected from discrete values between 16 and 96 with a step size of 16. The number of training epochs was chosen from {150, 200, 250, 300}. In addition, the L2 regularization coefficient was sampled from a logarithmic distribution within the interval 10^{-7} to 10^{-4} , and the learning rate was drawn from a logarithmic range between 5×10^{-5} and 2×10^{-2} . The same search space configuration was maintained for all IMFs to preserve methodological uniformity and guarantee fair

comparison across IMF-specific models. Thus, differences in optimal hyperparameter values arise from intrinsic IMF characteristics rather than variations in search boundaries.

Each IMF was optimized independently to account for its distinct statistical properties and frequency behavior. The resulting optimal configurations differed across IMFs and between red and green line datasets, as illustrated in Fig. 13. For IMF-1, both lines converged to 64 GRU units, a batch size of 64, 250 training epochs, and a dropout rate of 0.30, with learning rates around 1.7×10^{-3} . For IMF-2 and IMF-3, 64 GRU units with a batch size of 32 were obtained, with dropout values between 0.20 and 0.30 and learning rates ranging from approximately 7.8×10^{-4} to 1.8×10^{-3} . For IMF-4 and IMF-5, TPE optimization selected 96 GRU units, which indicates higher representational capacity requirements for these components. Dropout values remained within 0.20–0.30, while learning rates ranged approximately from 7.76×10^{-4} to 1.96×10^{-3} . The L2 regularization coefficient remained on the order of 10^{-6} , which reflects mild weight penalization across the models.

Furthermore, the total computational burden associated with TPE-based hyperparameter optimization varied across IMFs and between red and green line configurations, as shown in Fig. 14. Lower-index IMFs reached convergence more rapidly during the optimization process, whereas higher-index IMFs showed higher computational demand. For IMF-1, the total computational time remained close to 1000 s for both lines. In contrast, IMF-4 and IMF-5 exhibited the highest computational cost, with values reaching approximately 2300–2750 s across the configurations.

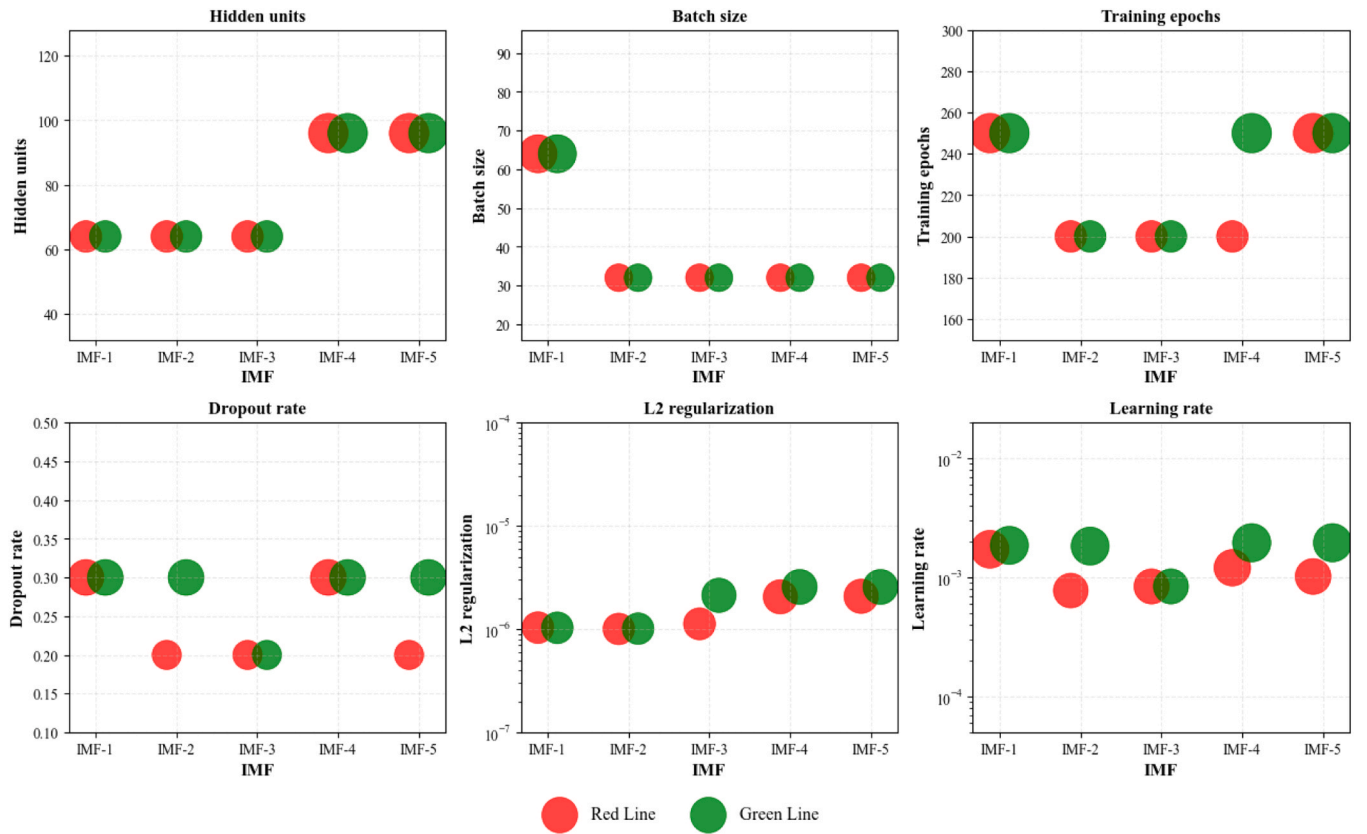


Fig. 13. Hyperparameter optimization results for the BiGRU models of the Dublin Luas Red and Green Lines across five IMF components.

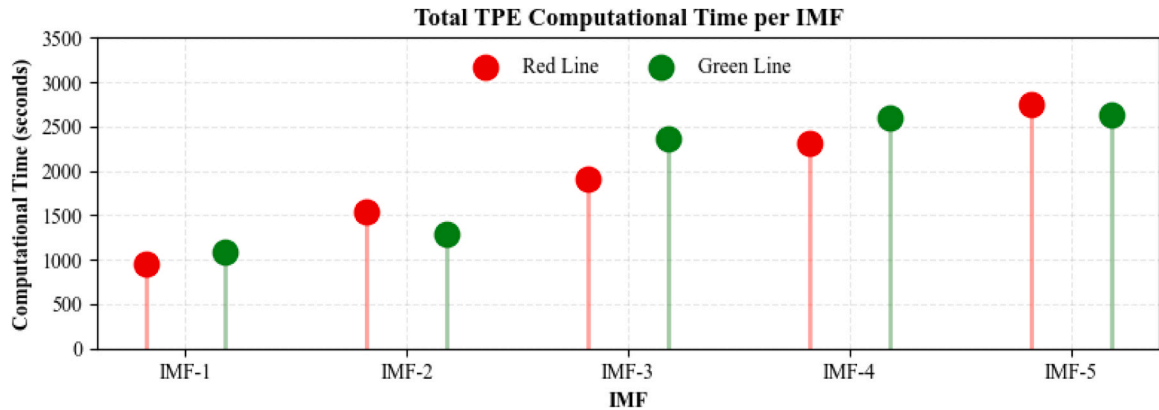


Fig. 14. Total computational cost of TPE-based hyperparameter optimization for each IMF.

4.4. Performance analysis of hybrid VMD+BiGRU

The VMD+BiGRU framework was evaluated by modelling each IMF separately and then reconstructing the overall forecast signal from the individual predictions. The performance of the TPE-optimized model across all IMFs of the Red and Green Line series, as measured by RMSE, MAE, and R^2 , is presented in Fig. 15. The results show high predictive accuracy across all IMF components for both series. For the Red Line, R^2 values range from 0.94 to 0.98 across individual IMFs, while the Green Line exhibits R^2 values between 0.87 and 0.98. A progressive decrease in RMSE and MAE from IMF-1 toward intermediate IMFs reflects improved modelling accuracy at higher-frequency scales. This behavior indicates that the VMD process isolates temporal components with lower variance and more distinct dynamic structure, which allows the BiGRU model to capture their patterns more effectively. The reconstructed forecasts

align closely with the observed series and yield overall R^2 values of 0.976 for the Red Line and 0.958 for the Green Line.

Figs. 16(a–f) and 17(a–f) present the actual and predicted plots for each IMF in the Red and Green Line time series, along with the corresponding reconstructed series. Both cases reveal a close match between the predicted and observed signals. Each subplot reveals that the model follows the original signal with minimal deviation, especially for the mid- and high-frequency components. The predicted curves reflect the overall shape and amplitude of the corresponding IMF trajectories. IMF-3 and IMF-5 show very narrow gaps between actual and forecasted values. The final reconstructed forecasts for both the Red and Green lines closely match the actual data.

Furthermore, the prediction error plots for both the Red and Green lines illustrate the reliability and stability of the proposed VMD+BiGRU model as shown in Fig. 18. In addition to the standard performance

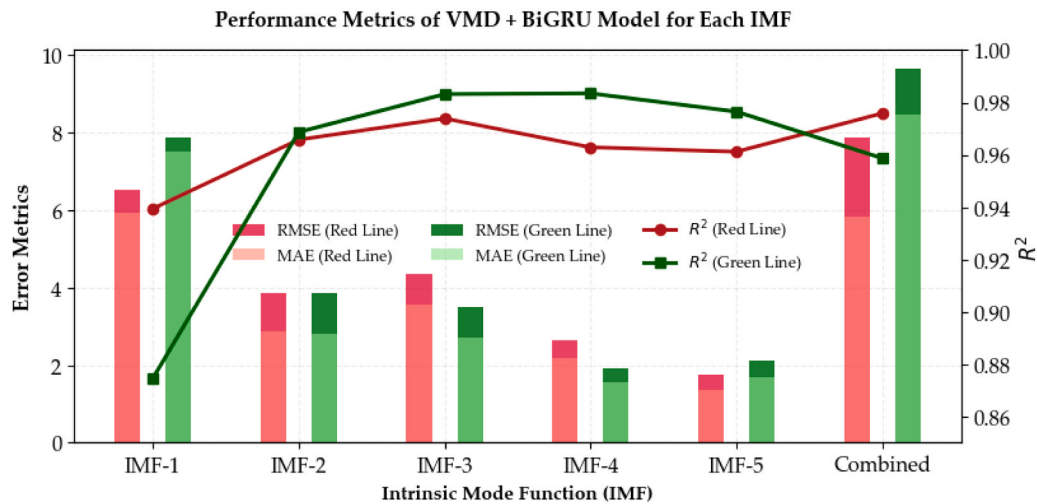


Fig. 15. Performance metrics of the VMD+BiGRU model for each IMF and the combined forecast on the Red and Green Line passenger demand time series.

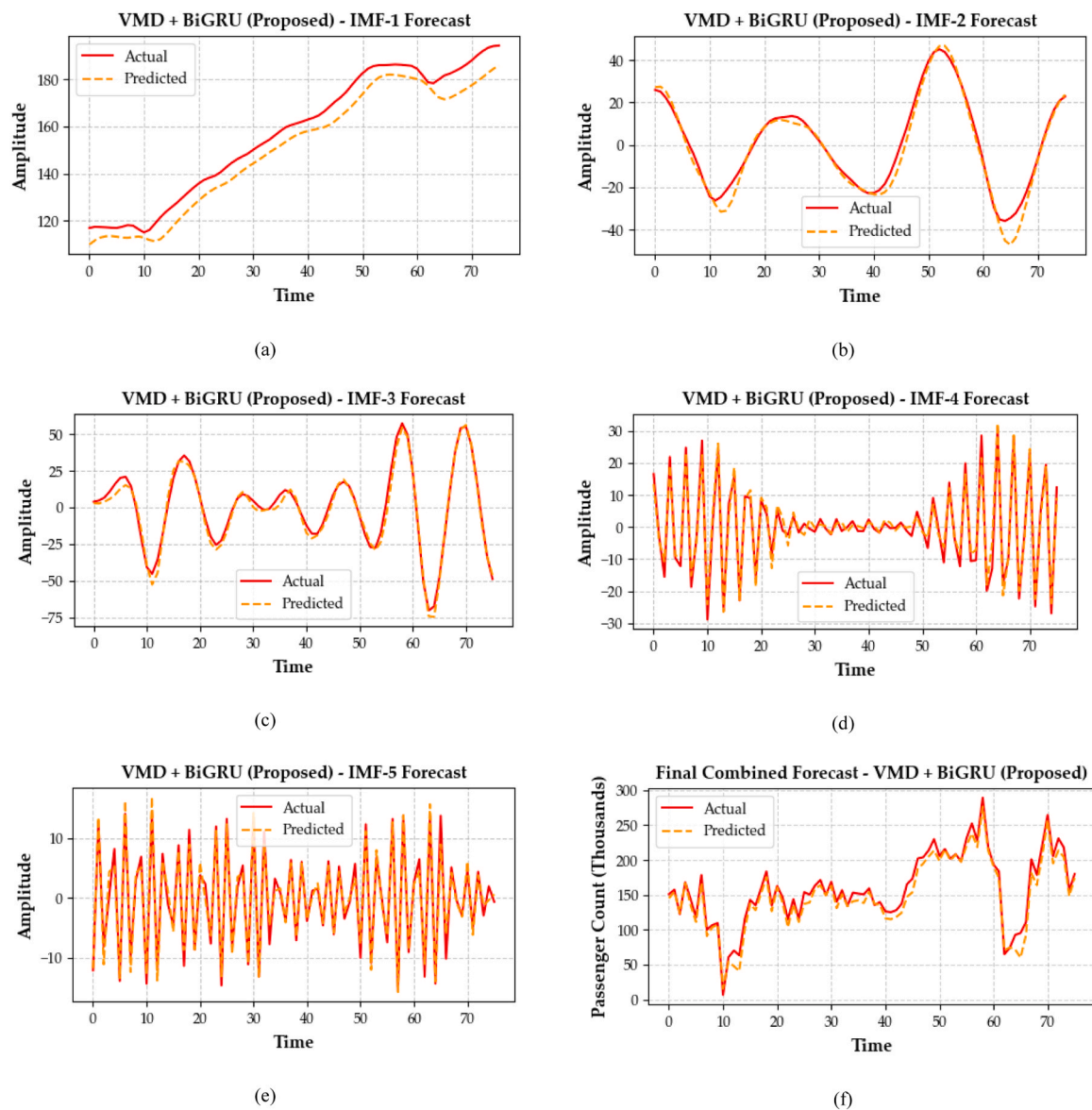


Fig. 16. Actual and predicted series using the VMD+BiGRU model for Dublin Red Luas Passenger demand; (a–e) Comparison between actual and predicted values for individual IMFs, (f) Reconstructed overall forecast generated by the VMD+BiGRU framework.

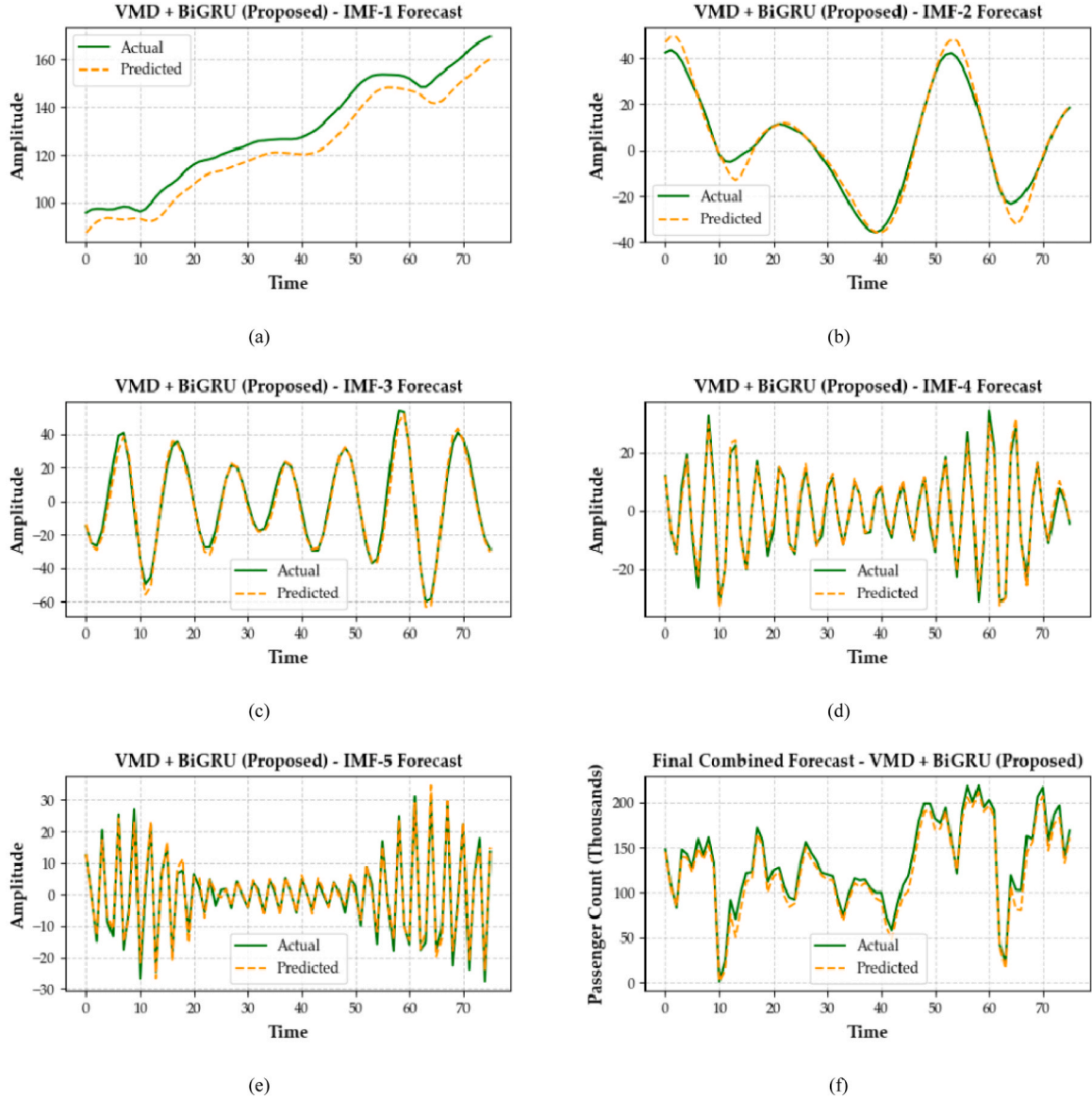


Fig. 17. Actual and predicted series using the VMD+BiGRU model for Dublin Green Luas Passenger demand; (a–e) Comparison between actual and predicted values for individual IMFs, (f) Reconstructed overall forecast generated by the VMD+BiGRU framework.

measures, these plots provide a visual understanding of the predictive accuracy of proposed framework across both training and testing datasets. The close alignment of data points along the ideal diagonal line (black dashed) indicates strong agreement between actual and predicted values, while the red dashed fit line for the test set further reflects the generalization capability of the proposed framework.

4.5. Performance comparison with other deep learning models

Figs. 19 and 20 show the reconstructed passenger time series and forecast results of the proposed VMD+BiGRU model compared with other hybrid models (VMD+LSTM, VMD+BiLSTM, VMD+TCN, and VMD+FNN) for the Red and Green Lines, respectively. The actual passenger demand data is plotted in black, while the outputs from each model are overlaid for comparison. Among all tested models, the VMD+BiGRU framework achieves the best predictive performance, accurately capturing fluctuations in passenger demand for both the Red and Green Line time series. In contrast, VMD+TCN performs the poorest, failing to capture key temporal variations and producing an overly smoothed response that diverges from the observed data. Although

VMD+FNN shows a modest improvement over VMD+TCN, it still fails to reflect important changes in passenger demand.

4.6. Ablation experiments

To assess the contribution of the VMD decomposition step and to compare different deep learning architectures, an ablation study was conducted. Fig. 21(a–b) summarize the model performance in terms of MAE, RMSE, and R^2 , along with the corresponding performance improvement heatmaps for the Red Line. In case of Red Luas Line, models that incorporate VMD show clear performance gains over their standalone counterparts. The BiGRU model reduces MAE from 15.40 to 5.85 and RMSE from 20.30 to 7.88 after VMD integration, while R^2 increases from 0.890 to 0.976. Similar improvements appear for other architectures. BiLSTM decreases MAE from 18.90 to 16.32 and RMSE from 24.80 to 19.23 with VMD, while R^2 changes from 0.860 to 0.856. LSTM reduces MAE from 21.20 to 15.12 and RMSE from 27.60 to 19.67, with R^2 rising from 0.830 to 0.850. These results confirm that decomposition improves predictive accuracy across most architectures, with the strongest absolute performance achieved by VMD+BiGRU.

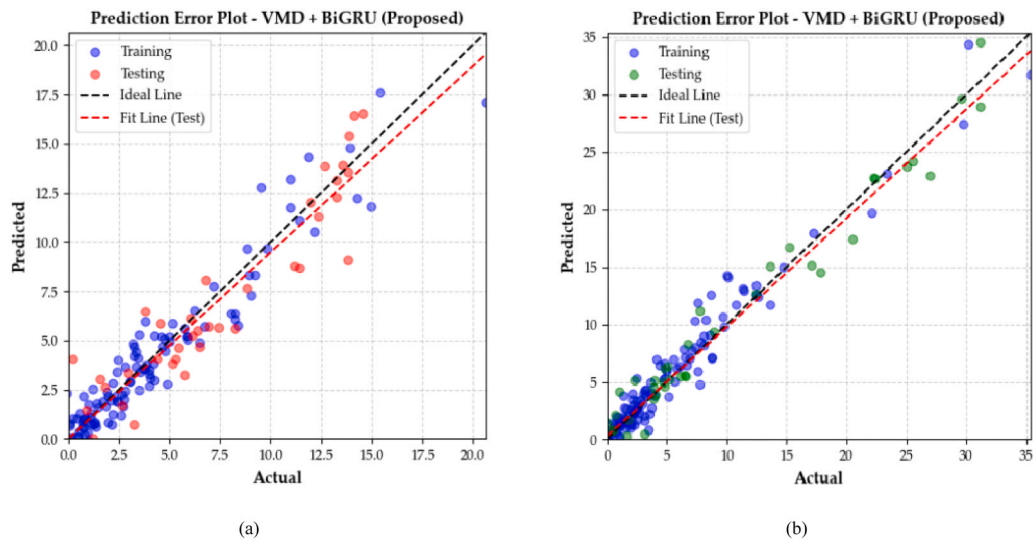


Fig. 18. Prediction error plots of passenger demand for the Dublin Luas using the proposed VMD+BiGRU model; (a) Red Line, (b) Green Line.

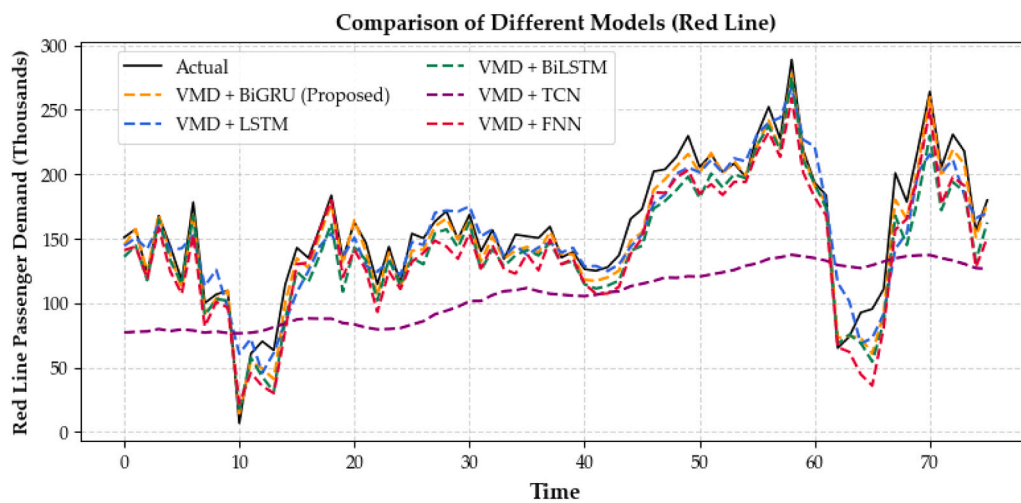


Fig. 19. Forecast results of different deep learning models compared with actual values for the Red Luas Line time series.

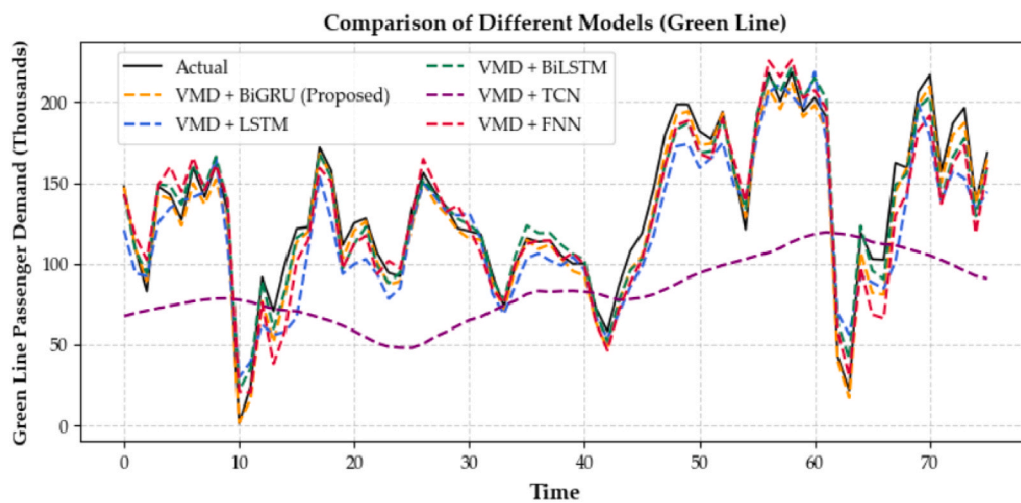


Fig. 20. Forecast results of different deep learning models compared with actual values for the Green Luas Line time series.

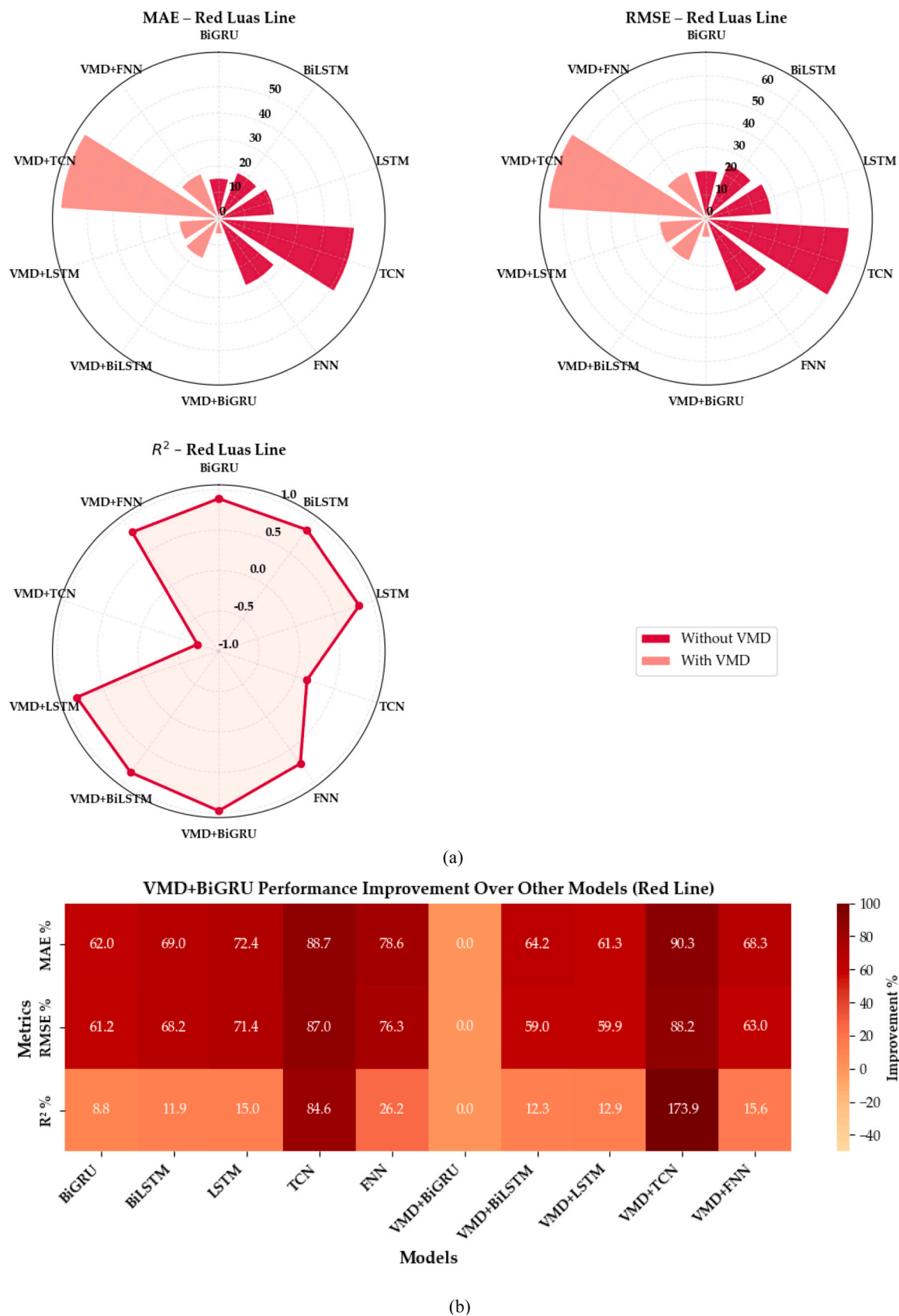


Fig. 21. Performance evaluation on the Red Luas Line; (a) Performance comparison of models with and without VMD; (b) Percentage performance improvement of VMD+BiGRU relative to other models.

When VMD+BiGRU is taken as the reference model, its relative advantage over other architectures becomes clear, both for models without VMD and those with VMD preprocessing. Compared with models without VMD, VMD+BiGRU achieves 62.0% lower MAE and 61.2% lower RMSE than BiGRU, together with an 8.8% increase in R^2 . Against BiLSTM, MAE and RMSE are reduced by 69.0% and 68.2%, while R^2 improves by 11.9%. In comparison with LSTM, the reductions reach 72.4% for MAE and 71.4% for RMSE, with R^2 higher by 15.0%. The performance gap is more pronounced for TCN and FNN, where MAE decreases by 88.7% and 78.6%, RMSE decreases by 87.0% and 76.3%, and R^2 increases by 84.6% and 26.2%, respectively. Focusing on models that also incorporate VMD, VMD+BiGRU continues to show a clear margin. Relative to VMD+BiLSTM, MAE and RMSE are lower by 64.2% and 59.0%, with R^2 higher by 12.3%. Compared with VMD+LSTM, MAE and RMSE remain lower by 61.3% and 59.9%, while R^2 increases by 12.9%. The largest relative difference occurs against VMD+TCN, where MAE and RMSE are reduced by 90.3% and 88.2%, and R^2 increases by 173.9%. Against VMD+FNN, VMD+BiGRU achieves 68.3% lower MAE and 63.0% lower RMSE, together with a 15.6% improvement in R^2 .

For the Green Luas Line, as illustrated in Fig. 22(a–b), models that incorporate VMD preprocessing show clear performance improvements relative to their standalone counterparts. The BiGRU model reduces MAE from 22.00 to 8.47 and RMSE from 28.00 to 9.65 after VMD integration, while R^2 increases from 0.720 to 0.958. BiLSTM also benefits from decomposition, with MAE decreasing from 26.00 to 8.95 and RMSE from 32.00 to 10.10, alongside an increase in R^2 from 0.690 to 0.945. For LSTM, MAE declines from 28.00 to 17.33 and RMSE from 35.00 to 20.95, with R^2 rising from 0.660 to 0.805. These results indicate that VMD enhances predictive performance across all recurrent architectures, with VMD+BiGRU and VMD+BiLSTM showing the strongest absolute accuracy. Taking VMD+BiGRU as the reference model, its relative advantage becomes more evident. Compared with BiGRU, MAE and RMSE are lower by 61.5% and 65.5%, while R^2 increases by 33.1%. Relative to BiLSTM, MAE and RMSE decrease by 67.4% and 69.7%, with R^2 higher by 38.4%. Against LSTM, MAE and RMSE reductions reach 69.7% and 72.4%, while R^2 improves by 45.2%. The contrast is more pronounced for TCN and FNN, where MAE reductions reach 85.9% and 61.5%, RMSE reductions reach 86.2% and 61.9%, and R^2 improvements reach 85.8% and 74.2%, respectively. Within the VMD-based group, VMD+BiGRU maintains a clear margin. Relative to VMD+BiLSTM, MAE and RMSE remain lower by 5.4% and 4.5%, with R^2 higher by 1.4%. Compared with VMD+LSTM, MAE and RMSE are lower by 51.1% and 53.9%, while R^2 increases by 19.0%. The largest relative difference appears against VMD+TCN, where MAE and RMSE decrease by 85.6% and 85.6%, alongside a substantial increase in R^2 . Against VMD+FNN, MAE and RMSE are lower by 30.8% and 36.7%, with R^2 higher by 6.8%.

4.7. Robustness of proposed model (Multi-horizon forecasting)

Urban transit planners often require demand forecasts not only for short-term scheduling but also for medium- and long-term planning related to capacity allocation, resource management, and infrastructure investment [90,91]. To evaluate the temporal stability and predictive robustness of the proposed VMD+BiGRU framework, a multi-horizon forecasting analysis was performed. The evaluation covered both the Red and Green Luas Lines across three forecast horizons: 1 week (short term), 4 weeks (medium term), and 12 weeks (long term). The results were compared with other VMD-based neural architectures, including VMD+BiLSTM, VMD+LSTM, and VMD+FNN, as illustrated in Fig. 23. It is pertinent to mention that the VMD+TCN model, which showed poor single-step performance with high errors and negative R^2 values, was excluded from the multi-step analysis due to its limited predictive reliability.

For both lines, the VMD+BiGRU model attains the lowest MAE and RMSE values across all forecast horizons, confirming its strong ability to

preserve predictive accuracy over time. On the Red Line, RMSE increases gradually from about 10.4 at the 1-week horizon to 16.8 at 12 weeks, accompanied by a slight decrease in R^2 from 0.97 to 0.94, which indicates minimal performance degradation as the prediction window extends. In contrast, the VMD+BiLSTM model yields higher error levels, with RMSE rising from approximately 20.6–28.4, whereas VMD+LSTM and VMD+FNN exhibit larger fluctuations and lower R^2 values and shows a reduced temporal generalization capability. A similar pattern is observed on the Green Line, where the VMD+BiGRU model again shows better performance by maintaining low error levels, with RMSE increasing only from about 8.1 at the 1-week horizon to 11.5 at 12 weeks, and an R^2 value above 0.95 across all horizons. The VMD+BiLSTM model produces moderate errors, with RMSE ranging between 18.2 and 30.1 and a gradual decrease in R^2 , while VMD+LSTM and VMD+FNN exhibit higher errors and a clearer drop in correlation strength as the forecast horizon extends.

These multi-horizon forecasting results verify that the VMD+BiGRU architecture provides the most accurate and reliable predictions for both Luas lines. Its ability to maintain low errors and high R^2 values across extended intervals indicates strong temporal adaptability, which makes it appropriate for short-term operations, medium-term scheduling, and long-term passenger demand forecasting in urban light rail transit systems.

4.8. Post-Hoc statistical comparison using Wilcoxon Signed-Rank test

Although the proposed VMD+BiGRU model shows better accuracy than other deep learning models across all performance metrics, these metrics alone may not provide sufficient evidence when results appear numerically close. To verify whether the performance differences are statistically significant, the Wilcoxon signed-rank test is applied as a post-hoc method. This test compares prediction errors between the proposed model and each alternative across all forecasted time steps. It is applied at a significance level of $\alpha = 0.05$. Table 5 reports the p-values and corresponding decisions for comparisons between VMD+BiGRU and VMD+BiLSTM, VMD+LSTM, VMD+TCN, and VMD+FNN. For the Red Line, the results show that VMD+BiGRU achieves statistically better performance relative to all alternative models, as all comparisons indicate significant differences. For the Green Line, VMD+BiGRU shows better performance than VMD+LSTM, VMD+TCN, and VMD+FNN. However, when compared with VMD+BiLSTM, the p-value of 0.092 exceeds the 0.05 significance threshold, which indicates that the current data sample does not provide sufficient evidence of a statistically significant performance difference between these two models. These results support the effectiveness of VMD+BiGRU on the Red Line, while indicating statistically comparable performance between VMD+BiGRU and VMD+BiLSTM on the Green Line.

4.9. Seasonal demand clustering of Dublin Luas ridership

A Gaussian Mixture modelling (GMM) clustering strategy is applied to the forecasted outputs generated by the VMD+BiGRU model, which was designed to capture temporal ridership dynamics [92]. GMM is adopted because its probabilistic formulation provides superior capability for identifying overlapping and heterogeneous travel demand regimes compared with distance-based methods such as K-means or hierarchical clustering, and it has been widely applied in traffic and transportation studies [93,94]. By enabling soft cluster assignment, this approach provides a more realistic representation of uncertainty and gradual transitions in passenger demand patterns.

The resulting monthly passengers demand forecasts were grouped into three distinct clusters that reflect low, medium, and high demand periods across the Luas network. The clustering outcome shows clear seasonal travel patterns on both the Dublin Red and Green lines as shown in Fig. 24. The low-demand cluster (Cluster 0), which includes the months from January to June, reflects a period of reduced public

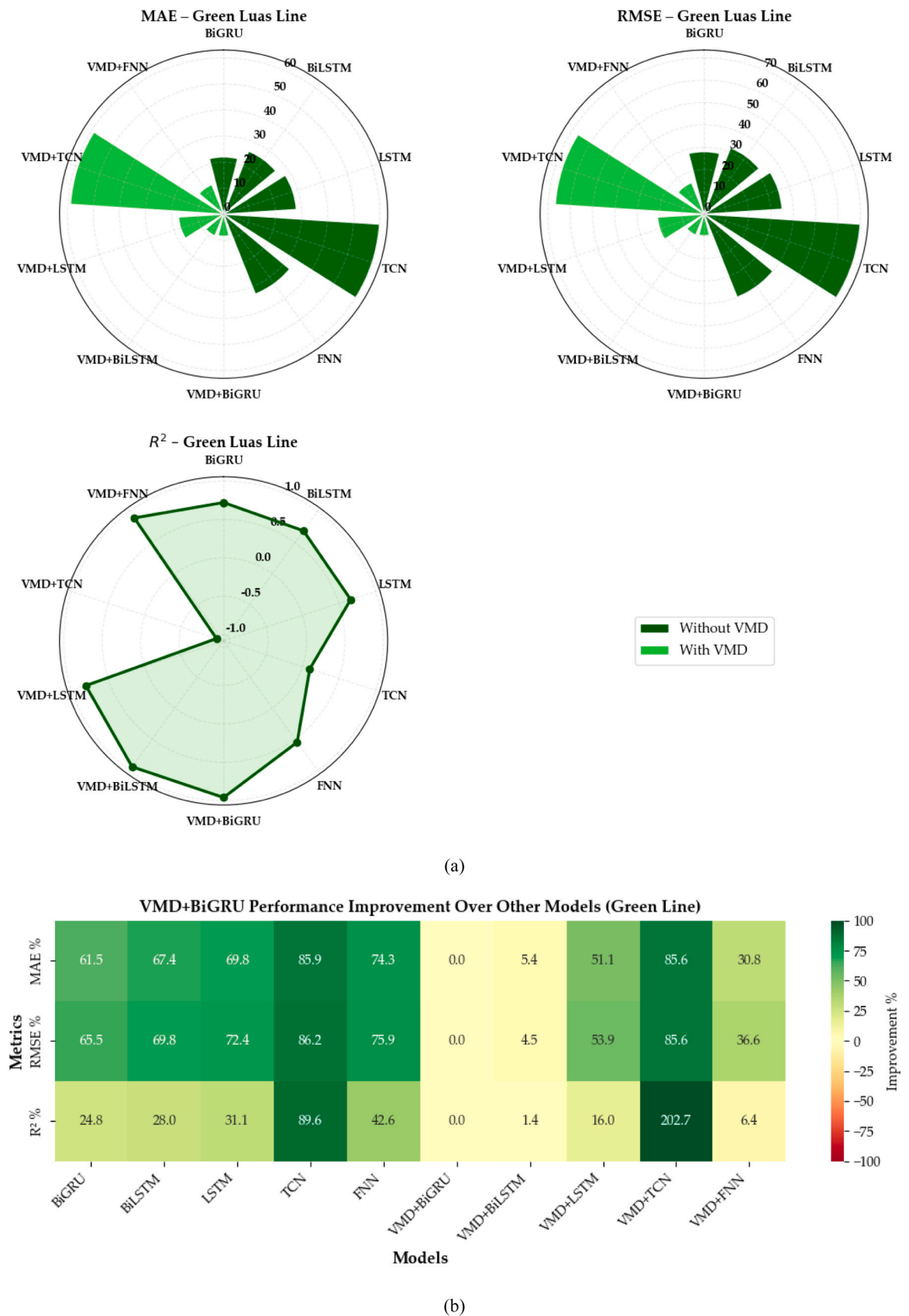


Fig. 22. Performance evaluation on the Green Luas Line; (a) Performance comparison of models with and without VMD; (b) Percentage performance improvement of VMD+BiGRU relative to other models.

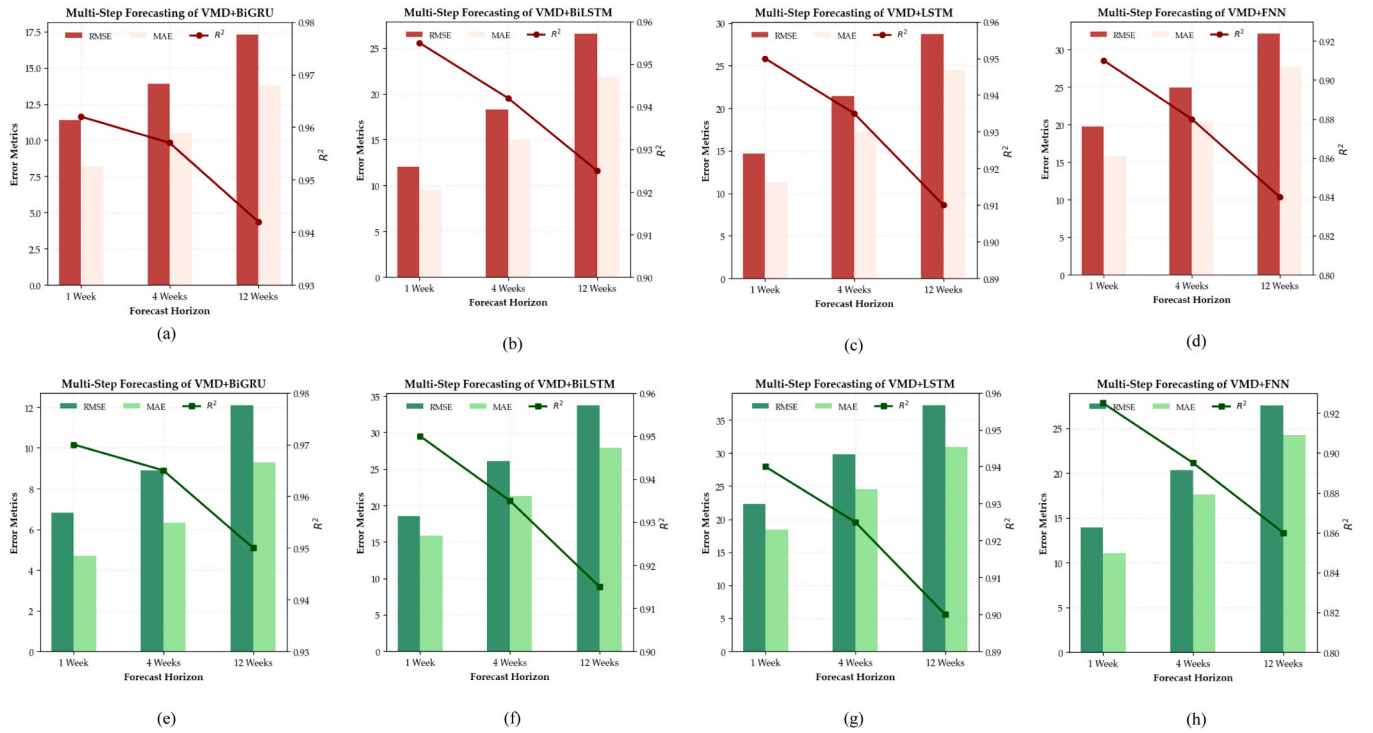


Fig. 23. Multi-horizon forecasting performance of VMD-based models across short-term (1 week), medium-term (4 weeks), and long-term (12 weeks) horizons for both Red and Green Luas Lines; (a) VMD+BiGRU (Red Line); (b) VMD+BiLSTM (Red Line); (c) VMD+LSTM (Red Line); (d) VMD+FNN (Red Line); (e) VMD+BiGRU (Green Line); (f) VMD+BiLSTM (Green Line); (g) VMD+LSTM (Green Line); (h) VMD+FNN (Green Line).

Table 5

Wilcoxon Signed-Rank test results for red and green lines.

Model Compared	<i>p</i> -value (Red Line)	Reject H_0	<i>p</i> -value (Green Line)	Reject H_0
VMD+BiLSTM	0.008	Yes	0.092	No
VMD+LSTM	0.004	Yes	0.006	Yes
VMD+TCN	0.001	Yes	0.003	Yes
VMD+FNN	0.003	Yes	0.03	Yes

transport usage. This decline is likely shaped by several factors, including the post-holiday travel slowdown in January and February, cold and wet winter weather that discourages walk and wait at Luas stops [95], and the academic calendar that results in lower student ridership due to semester breaks and examinations during spring months. In addition, tourism activity during this period tends to be

relatively low compared to the summer and autumn months.

The medium-demand cluster (Cluster 1), comprising July, August, and December, captures transitional travel behavior. July and August mark the peak of Ireland's tourist season, which boosts non-commuter ridership [96]. However, this is partially offset by reduced commuter traffic as many residents take annual leave during these months. December presents a dual effect, with ridership rising in the first half of the month due to Christmas shopping and city-center events, followed by a sharp fall in travel during the holiday break, which leads to a moderate monthly average. The high-demand cluster (Cluster 2) includes September, October, and November, which show the highest passenger volumes. This period reflects the full resumption of schools, universities, and workplaces after summer, which results in strong and stable commute activity. Weather conditions in early autumn are typically milder, which encourages regular use of public transport. Moreover, this period remains largely unaffected by public holidays, which

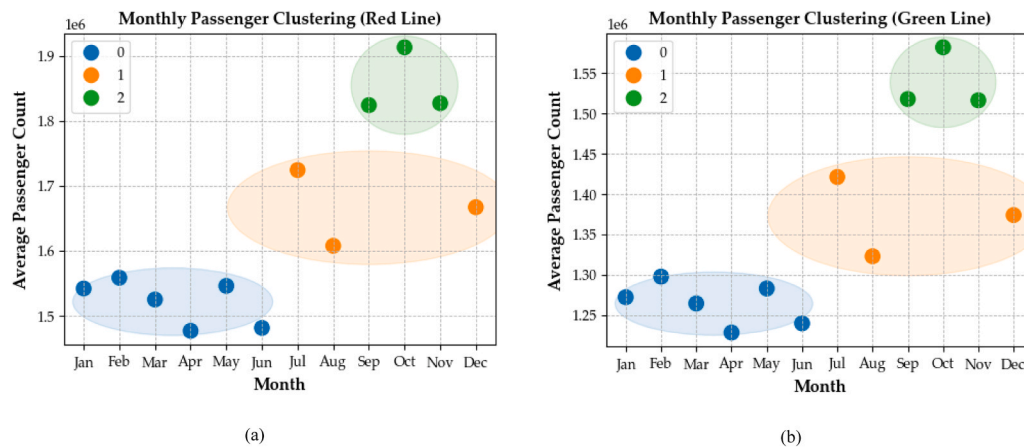


Fig. 24. Monthly Passenger Demand Clustering in Dublin Luas; (a) Red Luas Line; (b) Green Luas Line.

contributes to steady travel patterns [97].

5. Conclusions and future work

This study developed an integrated VMD+BiGRU framework to model the highly non-linear and non-stationary characteristics of weekly passenger demand on the Dublin Luas Red and Green lines. The proposed approach combines an FFT-MI-based VMD strategy for adaptive signal decomposition, frequency-specific modelling through the BiGRU architecture, and hyperparameter optimization using the TPE method. This integrated framework improves decomposition quality, captures bidirectional temporal dependencies within each IMF, and facilitates effective tuning of BiGRU hyperparameters at the component level, which results in improved forecasting performance. The proposed hybrid framework also provides greater adaptability and higher prediction accuracy than other deep learning models.

For the Red Line, the TPE-optimized VMD+BiGRU model achieves the highest prediction accuracy, with an RMSE of 7.88, MAE of 5.85, and R^2 of 0.976. The VMD+BiLSTM model yields the next best results, with RMSE of 19.23 and MAE of 16.32, alongside an R^2 of 0.856. In contrast, VMD+TCN records the lowest predictive accuracy, with an RMSE of 66.59, MAE of 60.10, and a negative R^2 of -0.722 , which indicates limited explanatory capability. A similar pattern appears for the Green Line. The VMD+BiGRU model again achieves the strongest performance, with an RMSE of 9.65, MAE of 8.47, and R^2 of 0.958. The VMD+BiLSTM model follows closely, with an RMSE of 10.10, an MAE of 8.95, and an R^2 of 0.945. The weakest performance is observed for VMD+TCN, which reports an RMSE of 66.82, MAE of 58.91, and a negative R^2 of -0.983 , which reflects a worst predictive reliability on the Green Line.

Beyond single-step evaluation, a multi-horizon forecasting analysis was also performed to assess temporal robustness across short-, medium-, and long-term horizons (1-week, 4-weeks, and 12-weeks). The VMD+BiGRU model maintained the lowest RMSE and MAE across all forecast horizons for both Red and Green lines and shows a stable performance over time. The Wilcoxon signed-rank test verified significant performance differences between the proposed model and all other models on both lines. The exception occurred on the Green Line, where the difference with VMD+BiLSTM was statistically insignificant.

5.1. Limitations of the study

The current study has some limitations that may impact the scope and applicability of the results.

1. The analysis relied solely on univariate time series data derived from weekly passenger journeys and did not incorporate external exogenous factors that may influence passenger demand.
2. Only aggregate ridership at the line level was considered due to availability of limited data, which may overlook variations across individual stations or specific sections of the network.
3. The proposed framework relies solely on the VMD strategy for signal decomposition and does not include a direct comparison with alternative decomposition techniques within the same forecasting framework.
4. The hyperparameter tuning process relies solely on the TPE optimization strategy. Although TPE provides efficient exploration of the hyperparameter space, other Bayesian optimization variants, evolutionary algorithms, or population-based approaches were not examined in this study.
5. The study employed a forward-chaining (fixed-origin) evaluation strategy, which, while providing a controlled and interpretable setup, may introduce uncertainty in the assessment of long-term multi-horizon forecasting performance.

5.2. Future direction

Future research can extend this work in several directions. First, incorporating exogenous variables, such as weather conditions, scheduled events, or traffic disruptions, could improve forecasting performance and model adaptability. Second, a station-level analysis may provide more granular insights into localized travel patterns and passenger demand dynamics. Furthermore, applying alternative decomposition techniques like CEEMDAN or Empirical Wavelet Transform (EWT) within an identical modelling setup would provide a robust benchmark for evaluating the relative efficacy of the VMD strategy. Similarly, employing rolling-origin validation would provide a more robust evaluation of long-term forecasting performance. Future studies could explore the integration of advanced transformer-based models with signal processing strategies to improve performance in capturing complex, long-range temporal patterns in passenger demand. Future work may also assess and compare other hyperparameter tuning strategies, including Gaussian process-based Bayesian optimization, Sequential Model-Based Optimization, Covariance Matrix Adaptation Evolution Strategy, Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Differential Evolution (DE), to further evaluate the robustness and effectiveness of the framework.

CRedit authorship contribution statement

Brian Caulfield: Writing – review & editing, Validation, Supervision, Funding acquisition, Formal analysis, Conceptualization. **Afaq Khattak:** Writing – original draft, Resources, Project administration, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to acknowledge the support of the SEAI National Energy Research Funding Programme for the RATE project (Project code: RDD932). The authors also acknowledge the use of the Grammarly AI tool solely for grammar and language correction. No content generation or rewriting was performed using AI tools.

Data availability

The data used in this study are publicly available through the Irish government's open data portal at data.gov.ie.

References

- [1] P. Mangan, J. Steer, N. Chadwick, The importance of process in planning: the Dublin transportation initiative, *Built Environ.* 48 (2022) 528–547, <https://doi.org/10.2148/benv.48.4.528>.
- [2] M.D. Culleton, Urban Quality Related Effects on Residential Property Prices in Dublin, Ireland, School of Architecture, Urban Planning and Construction Engineering, Master of Science in Urban Planning and Policy Design, 2022.
- [3] E. Sweeney, J. Brunton, Modification of Luas heating and ventilation systems to reduce energy consumption. Irish Transport Research Network Conference, 2013, 10.21427/16za-cg14.
- [4] TII, Passenger Numbers in LUAS, (2024) (<https://www.tii.ie/en/public-transport/operations-and-maintenance/passenger-numbers/>).
- [5] CSO, Luas Passenger Numbers by Month, Year, and Line, (2024) (<https://data.cso.ie/table/TOA10>).
- [6] NTA, Luas Lucan Feasibility Study, (2021) (<https://www.nationaltransport.ie/wp-content/uploads/2021/11/Luas-Lucan-Feasibility-Study.pdf>).
- [7] S. Zhao, X. Mi, A novel hybrid model for short-term high-speed railway passenger demand forecasting, *IEEE Access* 7 (2019) 175681–175692, <https://doi.org/10.1109/ACCESS.2019.2957612>.

- [8] G. Jin, S. He, J. Li, Y. Li, X. Guo, H. Xu, An integrated model for demand forecasting and train stop planning for high-speed rail, *Symmetry* 11 (2019) 720, <https://doi.org/10.3390/sym11050720>.
- [9] H. Su, S. Mo, S. Peng, Short-term prediction of time-varying passenger flow for intercity high-speed railways: A neural network model based on multi-source data, *Mathematics* 11 (2023) 3446, <https://doi.org/10.3390/math11163446>.
- [10] R. Xue, D. Sun, S. Chen, Short-term bus passenger demand prediction based on time series model and interactive multiple model approach, *Discret. Dyn. Nat. Soc.* 2015 (2015) 682390, <https://doi.org/10.1155/2015/682390>.
- [11] A. Cyril, R.H. Mulangi, V. George, Modelling and forecasting bus passenger demand using time series method. *Proceedings of the 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2018, pp. 460–466, [10.1109/ICRITO.2018.8748443](https://doi.org/10.1109/ICRITO.2018.8748443).
- [12] S. Liyanage, R. Abduljabbar, H. Dia, P.-W. Tsai, AI-based neural network models for bus passenger demand forecasting using smart card data, *J. Urban Manag.* 11 (2022) 365–380, <https://doi.org/10.1016/j.jum.2022.05.002>.
- [13] Q. Chen, D. Wen, X. Li, D. Chen, H. Lv, J. Zhang, P. Gao, Empirical mode decomposition based long short-term memory neural network forecasting model for the short-term metro passenger flow, *PloS One* 14 (2019) e0222365, <https://doi.org/10.1371/journal.pone.0222365>.
- [14] L. Zeng, Z. Li, J. Yang, X. Xu, CEEMDAN-IPSO-LSTM: a novel model for short-term passenger flow prediction in urban rail transit systems, *Int. J. Environ. Res. Public Health* 19 (2022) 16433, <https://doi.org/10.3390/ijerph192416433>.
- [15] Q. Qin, Za Wang, B. Li, A. Huang, The HWAM-EMD-GRU forecasting model for short-term passenger flow in an airport light rail transit line, *Urban Rail Transit* 10 (2024) 178–187, <https://doi.org/10.1007/s40864-024-00217-5>.
- [16] R. Deng, J. Zhang, Short-term metro passenger flow prediction based on CEEMDAN-IGWO-CNN-BiLSTM. *Proceedings of the 2025 8th International Conference Computer Information Science Artificial Intelligence*, 2025, pp. 275–280, [10.1145/3773365.3773409](https://doi.org/10.1145/3773365.3773409).
- [17] Y. Lu, J. Wang, Short-term passenger flow forecasting for rail transit integrating multi-scale decomposition and deep attention mechanism, *Sustainability* 17 (2025) 8880, <https://doi.org/10.3390/su17198880>.
- [18] Q. Bing, F. Shen, X. Chen, W. Zhang, Y. Hu, D. Qu, A hybrid short-term traffic flow multistep prediction method based on variational mode decomposition and long short-term memory model, *Discret. Dyn. Nat. Soc.* 2021 (2021) 4097149, <https://doi.org/10.1155/2021/4097149>.
- [19] M. Wand, K. Malinovsky, J. Schmidhuber, I.V. Tetko, Artificial neural networks and machine learning-ICANN 2024. 33rd International Conference on Artificial Neural Networks, Lugano, Switzerland, September 17–20, 2024, *Proceedings, Part I*, Springer Nature, 2024, [10.1007/978-3-031-72341-4](https://doi.org/10.1007/978-3-031-72341-4).
- [20] N. Shahsavari-Pour, A. Heydari, F. Keynia, A. Fekih, A. Shahsavari-Pour, Building electrical consumption patterns forecasting based on a novel hybrid deep learning model, *Results Eng.* 26 (2025) 104522, <https://doi.org/10.1016/j.rineng.2025.104522>.
- [21] H. Li, T. Liu, X. Wu, Q. Chen, An optimized VMD method and its applications in bearing fault diagnosis, *Measurement* 166 (2020) 108185, <https://doi.org/10.1016/j.measurement.2020.108185>.
- [22] H. Balti, A. Ben Abbes, I.R. Farah, A Bi-GRU-based encoder-decoder framework for multivariate time series forecasting, *Soft Comput.* 28 (2024) 6775–6786, <https://doi.org/10.1007/s00500-023-09531-9>.
- [23] X. Zhang, C. He, Y. Lu, B. Chen, L. Zhu, L. Zhang, Fault diagnosis for small samples based on attention mechanism, *Measurement* 187 (2022) 110242, <https://doi.org/10.1016/j.measurement.2021.110242>.
- [24] S. Wang, C. Shao, J. Zhang, Y. Zheng, M. Meng, Traffic flow prediction using bi-directional gated recurrent unit method, *Urban Inform.* 1 (2022) 16, <https://doi.org/10.1007/s44212-022-00015-z>.
- [25] J. Krauß, *Optimizing Hyperparameters for Machine Learning Algorithms in Production*, Apurimus Wissenschaftsverlag, 2022.
- [26] D.S. Huang, Q. Zhang, C. Zhang, W. Chen, *Advanced Intelligent Computing Technology and Applications. 21st International Conference, ICIC 2025, Ningbo, China, July 26–29, 2025, Proceedings, Part VI*, Springer Nature, Singapore, 2025, [10.1007/978-981-96-9921-6](https://doi.org/10.1007/978-981-96-9921-6).
- [27] H. Khan, A. Khan, S. Villar, L. Alonso, A. Almaleh, A. Al-Qahtani, A comparative study of optimized-LSTM models using tree-structured Parzen estimator for traffic flow forecasting in intelligent transportation, *Comput. Mater. & Contin.* 83 (2025) 3369, <https://doi.org/10.32604/cmc.2025.060474>.
- [28] M. Liu, S. Long, T. Li, W. Wang, J. Li, Spatiotemporal prediction of ground surface deformation using TPE-optimized deep learning, *Remote Sens.* 18 (2026) 234, <https://doi.org/10.3390/rs18020234>.
- [29] S. Hanifi, A. Cammarono, H. Zare-Behtash, Advanced hyperparameter optimization of deep learning models for wind power prediction, *Renew. Energy* 221 (2024) 119700, <https://doi.org/10.1016/j.renene.2023.119700>.
- [30] A. Khattak, P.-w. Chan, F. Chen, A. H. Almaliki, A hybrid VMD-TPE-TCN framework for wind shear prediction near an airport runway, *Atmosphere* 16 (2025) 381, <https://doi.org/10.3390/atmos16040381>.
- [31] J. Ma, J. Weng, C. Tang, Z. Liu, J. Yuan, Short-term bus passenger flow forecasting method based on gradient boosting decision tree (GBDT) model, *CICTP 2023*, 2023, pp. 870–880, [10.1061/9780784484869.084](https://doi.org/10.1061/9780784484869.084).
- [32] K. Guo, M. Xiao, Y. Liu, H. Zhang, Outbound time prediction of subway passengers based on deep forest algorithm, *CICTP 2023*, 2023, pp. 728–738, [10.1061/9780784484869.071](https://doi.org/10.1061/9780784484869.071).
- [33] S. Fukuda, H. Uchida, H. Fujii, T. Yamada, Short-term prediction of traffic flow under incident conditions using graph convolutional recurrent neural network and traffic simulation, *IET Intell. Transp. Syst.* 14 (2020) 936–946, <https://doi.org/10.1049/iet-its.2019.0778>.
- [34] J. Hoppe, F. Schwinger, H. Haeger, J. Wernz, M. Jarke, Improving the prediction of passenger numbers in public transit networks by combining short-term forecasts with real-time occupancy data, *IEEE Open J. Intell. Transp. Syst.* 4 (2023) 153–174, <https://doi.org/10.1109/OJITS.2023.3251564>.
- [35] D.D. Chuwang, W. Chen, Forecasting daily and weekly passenger demand for urban rail transit stations based on a time series model approach, *Forecasting* 4 (2022) 904–924, <https://doi.org/10.3390/forecast4040049>.
- [36] S. Feng, G. Cai, Passenger flow forecast of metro station based on the ARIMA model. *Proceedings of the 2015 International Conference on Electrical and Information Technologies for Rail Transportation*, Springer, 2016, pp. 463–470 (Transportation). [10.1007/978-3-662-49370-0_49](https://doi.org/10.1007/978-3-662-49370-0_49).
- [37] B. Moghimi, C. Kamga, A. Safikhani, S. Mudigonda, P. Vicuna, Non-stationary time series model for station-based subway ridership during covid-19 pandemic: case study of New York City, *Transp. Res. Rec.* 2677 (2023) 463–477, <https://doi.org/10.1177/03611981221084698>.
- [38] S. Halyal, R.H. Mulangi, M. Harsha, Forecasting public transit passenger demand: with neural networks using APC data, *Case Stud. Transp. Policy* 10 (2022) 965–975, <https://doi.org/10.1016/j.cstp.2022.03.011>.
- [39] A. Khattak, J. Zhang, P.-w. Chan, F. Chen, A. H. Almaliki, A new frontier in wind shear intensity forecasting: stacked temporal convolutional networks and tree-based models framework, *Atmosphere* 15 (2024) 1369, <https://doi.org/10.3390/atmos15111369>.
- [40] B. Lim, S. Zohren, Time-series forecasting with deep learning: a survey, *Philos. Trans. R. Soc. A* 379 (2021), <https://doi.org/10.1098/rsta.2020.0209>.
- [41] L. Liu, R.-C. Chen, A novel passenger flow prediction model using deep learning methods, *Transp. Res. Part C Emerg. Technol.* 84 (2017) 74–91, <https://doi.org/10.1016/j.trc.2017.08.001>.
- [42] J. Zhang, F. Chen, Z. Cui, Y. Guo, Y. Zhu, Deep learning architecture for short-term passenger flow forecasting in urban rail transit, *IEEE Trans. Intell. Transp. Syst.* 22 (2020) 7004–7014, <https://doi.org/10.1109/TITS.2020.3000761>.
- [43] D. Chen, J. Zhang, S. Jiang, Forecasting the short-term metro ridership with seasonal and trend decomposition using loess and LSTM neural networks, *IEEE Access* 8 (2020) 91181–91187, <https://doi.org/10.1109/ACCESS.2020.2995044>.
- [44] Y. Liu, Z. Liu, R. Jia, DeepPF: A deep learning based architecture for metro passenger flow prediction, *Transp. Res. Part C Emerg. Technol.* 101 (2019) 18–34, <https://doi.org/10.1016/j.trc.2019.01.027>.
- [45] J. Zuo, M. He, Z. Yu, J. Wang, Short-term passenger flow prediction of urban rail transit on Bayesian optimization-bidirectional long short-term memory with causal temporal pattern attention, *Int. J. Transp. Sci. Technol.* (2025), <https://doi.org/10.1016/j.ijst.2025.01.008>.
- [46] Z. Hou, Z. Du, G. Yang, Z. Yang, Short-term passenger flow prediction of urban rail transit based on a combined deep learning model, *Appl. Sci.* 12 (2022) 7597, <https://doi.org/10.3390/app12157597>.
- [47] X. Yang, Q. Xue, M. Ding, J. Wu, Z. Gao, Short-term prediction of passenger volume for urban rail systems: A deep learning approach based on smart-card data, *Int. J. Prod. Econ.* 231 (2021) 107920, <https://doi.org/10.1016/j.jipe.2020.107920>.
- [48] Y. Wang, X. Du, Z. Lu, Q. Duan, J. Wu, Improved LSTM-based time-series anomaly detection in rail transit operation environments, *IEEE Trans. Ind. Inform.* 18 (2022) 9027–9036, <https://doi.org/10.1109/TII.2022.3164087>.
- [49] B. Liu, X. Ma, E. Tan, Z. Ma, Passenger flow anomaly detection in urban rail transit networks with graph convolution network-informer and Gaussian Bayes models, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 381 (2023), <https://doi.org/10.1098/rsta.2022.0253>.
- [50] J. Nilsson, Isolation forest outlier detection for analyzing fraud in public transport, TRITA-SCI-GRU, Independent thesis Advanced level (degree of Master (Two Years)) (2024) 62 (<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-361498>).
- [51] J. Gu, Z. Jiang, W.D. Fan, J. Wu, J. Chen, Real-time passenger flow anomaly detection considering typical time series clustered characteristics at metro stations, *J. Transp. Eng. Part A Syst.* 146 (2020) 04020015, <https://doi.org/10.1061/JTEPBS.0000333>.
- [52] Y. Cheng, P. Krishnakumari, Unsupervised learning for public transport delay pattern analysis, *Transp. Res. Rec.* 2678 (2024) 358–373, <https://doi.org/10.1177/03611981231215333>.
- [53] M. Shoman, A. Aboah, Y. Adu-Gyamfi, Deep learning framework for predicting bus delays on multiple routes using heterogeneous datasets, *J. Big Data Anal. Transp.* 2 (2020) 275–290, <https://doi.org/10.1007/s42421-020-00031-y>.
- [54] M.B. Ulak, A. Yazici, Y. Zhang, Analyzing network-wide patterns of rail transit delays using Bayesian network learning, *Transp. Res. Part C Emerg. Technol.* 119 (2020) 102749, <https://doi.org/10.1016/j.trc.2020.102749>.
- [55] A. Piaskowska, M. Hernes, E. Walaszczyk, A. Kozina, K. Czerniachowska, Prediction of the delay time of public transportation using machine learning. *International Conference on Computational Collective Intelligence*, Springer, 2024, pp. 283–294, [10.1007/978-3-031-70248-8_22](https://doi.org/10.1007/978-3-031-70248-8_22).
- [56] M. Sarhani, S. Voß, Prediction of rail transit delays with machine learning: How to exploit open data sources, *Multimodal Transp.* 3 (2024) 100120, <https://doi.org/10.1016/j.multra.2024.100120>.
- [57] S. Wang, S. Liu, X. Guan, Ultra-short-term power prediction of a photovoltaic power station based on the VMD-CEEMDAN-LSTM model, *Front. Energy Res.* 10 (2022) 2022, <https://doi.org/10.3389/fenrg.2022.945327>.
- [58] Z. Ning, G. Chen, J. Wang, W. Hu, Multi-scale decomposition and hybrid deep learning CEEMDAN-VMD-CNN-BiLSTM approach for wind power forecasting, *Processes* (2025), <https://doi.org/10.3390/pr13072046>.
- [59] C. Ma, Y. Hu, X. Xu, Hybrid deep learning model with VMD-BiLSTM-GRU networks for short-term traffic flow prediction, *Data Sci. Manag.* 8 (2025) 257–269, <https://doi.org/10.1016/j.dsm.2024.10.004>.

- [60] S. Yuan, P. Jia, S. Wang, A secondary decomposition-ensemble approach to interval predicting China's railway container volume, *Appl. Soft Comput.* 143 (2023) 110374, <https://doi.org/10.1016/j.asoc.2023.110374>.
- [61] O.M. Katipoglu, Integration of extreme learning machines with CEEMDAN and VMD techniques in the prediction of the multiscalar standardized runoff index and standardized precipitation evapotranspiration index, *Nat. Hazards* 120 (2024) 825–849, <https://doi.org/10.1007/s11069-023-06238-w>.
- [62] S.A. Gyamerah, V. Owusu, Short- and long-term weather prediction based on a hybrid of CEEMDAN, LMD, and ANN, *PLOS ONE* 19 (2024), <https://doi.org/10.1371/journal.pone.0304754>.
- [63] S. Hou, Q. Geng, Y. Huang, Z. Bian, Rainfall prediction model based on CEEMDAN-VMD-BiLSTM network, *Water Air Soil Pollut.* 235 (2024) 482, <https://doi.org/10.1007/s11270-024-07299-8>.
- [64] T. Liang, H. Lu, H. Sun, Application of parameter optimized variational mode decomposition method in fault feature extraction of rolling bearing, *Entropy* 23 (2021) 520, <https://doi.org/10.3390/e23050520>.
- [65] S. Wu, F. Feng, J. Zhu, C. Wu, G. Zhang, A method for determining intrinsic mode function number in variational mode decomposition and its application to bearing vibration signal processing, *Shock Vib.* 2020 (2020) 8304903, <https://doi.org/10.1155/2020/8304903>.
- [66] L. Wang, Y. Liu, T. Li, X. Xie, C. Chang, Short-term PV power prediction based on optimized VMD and LSTM, *IEEE Access* 8 (2020) 165849–165862, <https://doi.org/10.1109/ACCESS.2020.3022246>.
- [67] D.N.T. How, M.A. Hannan, M.S.H. Lipu, P.J. Ker, M. Mansor, K.S.M. Sahari, K. M. Muttaqi, SOC estimation using deep bidirectional gated recurrent units with tree parzen estimator hyperparameter optimization, *IEEE Trans. Ind. Appl.* 58 (2022) 6629–6638, <https://doi.org/10.1109/TIA.2022.3180282>.
- [68] M. Woźniak, M. Wiecek, J. Silka, BiLSTM deep neural network model for imbalanced medical data of IoT systems, *Future Gener. Comput. Syst.* 141 (2023) 489–499, <https://doi.org/10.1016/j.future.2022.12.004>.
- [69] A.H. Almaliki, A. Khattak, Short-and long-term tidal level forecasting: a novel hybrid TCN+ LSTM framework, *J. Sea Res.* 204 (2025) 102577, <https://doi.org/10.1016/j.seares.2025.102577>.
- [70] B. Abduraxman, P. Hubbard, T. Harrison, C. Ward, D. Fletcher, R. Lewis, B. White, Acceleration-based friction coefficient estimation of a rail vehicle using feedforward NN: validation with track measurements, *Veh. Syst. Dyn.* 62 (2024) 3235–3254, <https://doi.org/10.1080/00423114.2024.2323600>.
- [71] Q. Qi, R. Cheng, H. Ge, Short-term inbound rail transit passenger flow prediction based on BiLSTM model and influence factor analysis, *Digit. Transp. Saf.* 2 (2023) 12–22, <https://doi.org/10.48130/DTS-2023-0002>.
- [72] M. Adil, J.-Z. Wu, R.K. Chakraborty, A. Alahmadi, M.F. Ansari, M.J. Ryan, Attention-based STL-BiLSTM network to forecast tourist arrival, *Processes* 9 (2021) 1759, <https://doi.org/10.3390/pr9101759>.
- [73] J. Cheng, G. Liu, S. Gao, A. Raza, J. Li, W. Juan, Short-term passenger flow prediction in urban rail transit based on points of interest, *IEEE Access* (2024), <https://doi.org/10.1109/ACCESS.2024.3425634>.
- [74] P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, Y. Liu, Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station, *Soft Comput.* 24 (2020) 16453–16482, <https://doi.org/10.1007/s00500-020-04954-0>.
- [75] M. Gallo, G. De Luca, L. D'Acerno, M. Botte, Artificial neural networks for forecasting passenger flows on metro lines, *Sensors* 19 (2019) 3424, <https://doi.org/10.3390/s19153424>.
- [76] J. Mulerikkal, S. Thandassery, V. Rejathalal, D.M.D. Kunnamkody, Performance improvement for metro passenger flow forecast using spatio-temporal deep neural network, *Neural Comput. Appl.* 34 (2022) 983–994, <https://doi.org/10.1007/s00521-021-06522-5>.
- [77] R.S. Pontoh, S. Zahroh, H. Nurahman, R. Aprillion, A. Ramdani, D. Akmal, Applied of feed-forward neural network and facebook prophet model for train passengers forecasting. *Journal of physics: Conference series*, IOP Publishing, 2021 012057, <https://doi.org/10.1088/1742-6596/1776/1/012057>.
- [78] A. Khattak, J. Zhang, P.-w. Chan, F. Chen, SPE-SHAP: Self-paced ensemble with Shapley additive explanation for the analysis of aviation turbulence triggered by wind shear events, *Expert Syst. Appl.* 254 (2024) 124399, <https://doi.org/10.1016/j.eswa.2024.124399>.
- [79] TII, TOA11 - Luas Passenger Numbers, (2024) (<https://data.gov.ie/dataset/toa11-luas-passenger-numbers>).
- [80] M. Hyne, P. Malone, The utility of public transport in Ireland: Post COVID-19, (2020).
- [81] K.I. Quintyne, C. Kelly, A. Sheridan, P. Kenny, COVID-19 transport restrictions in Ireland: impact on air quality and respiratory hospital admissions, *Public Health* 198 (2021) 156–160, <https://doi.org/10.1016/j.puhe.2021.07.008>.
- [82] C.K. Manzira, A. Charly, B. Caulfield, Assessing the impact of mobility on the incidence of COVID-19 in Dublin City, *Sustain. Cities Soc.* 80 (2022) 103770, <https://doi.org/10.1016/j.scs.2022.103770>.
- [83] V.P. Singh, R. Singh, P.K. Paul, D.S. Bisht, S. Gaur, Hydrological processes modelling and data analysis, *Water Sci. Technol. Libr.* 127 (2024), <https://doi.org/10.1007/978-981-97-1316-5>.
- [84] W. Bukaita, G. Garcia de Celis, M. Gurram, Training-testing data ratio selection for accurate time series forecasting: a COVID-19 case study. *Proceedings of the Future Technologies Conference*, Springer, 2024, pp. 227–246, https://doi.org/10.1007/978-3-031-73125-9_14.
- [85] Z. Mohammed, C. Anas, M. El Hammoume, A hybrid learning framework for forecasting uncertainty and adaptive inventory planning in retail supply chains, *Supply Chain Anal.* (2025) 100180, <https://doi.org/10.1016/j.sca.2025.100180>.
- [86] O. Baker, Z. Ziran, M. Mecella, K. Subaramaniam, S. Palaniappan, Predictive modelling for pandemic forecasting: a COVID-19 study in New Zealand and partner countries, *Int. J. Environ. Res. Public Health* 22 (2025) 562, <https://doi.org/10.3390/ijerph22040562>.
- [87] G.M. Weiss, F. Provost, Learning when training data are costly: the effect of class distribution on tree induction, *J. Artif. Intell. Res.* 19 (2003) 315–354, <https://doi.org/10.1613/jair.1199>.
- [88] D. Sarwinda, R.H. Paradisa, A. Bustamam, P. Anggia, Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer, *Procedia Comput. Sci.* 179 (2021) 423–431, <https://doi.org/10.1016/j.procs.2021.01.025>.
- [89] Y. Ensafi, S.H. Amin, G. Zhang, B. Shah, Time-series forecasting of seasonal items sales using machine learning – a comparative analysis, *Int. J. Inf. Manag. Data Insights* 2 (2022) 100058, <https://doi.org/10.1016/j.jjimei.2022.100058>.
- [90] Q. Xue, W. Zhang, M. Ding, X. Yang, J. Wu, Z. Gao, Passenger flow forecasting approaches for urban rail transit: a survey, *Int. J. Gen. Syst.* 52 (2023) 919–947, <https://doi.org/10.1080/03081079.2023.2231133>.
- [91] L. Wang, J.G. Jin, L. Sun, D.-H. Lee, Urban rail transit disruption management: Research progress and future directions, *Front. Eng. Manag.* 11 (2024) 79–91, <https://doi.org/10.1007/s42524-023-0291-z>.
- [92] F. Ros, R. Riad, Feature and dimensionality reduction for clustering with deep learning, *Springer*, 2024, <https://doi.org/10.1007/978-3-031-48743-9>.
- [93] L. Li, R.J. Hansman, R. Palacios, R. Welsch, Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring, *Transp. Res. Part C Emerg. Technol.* 64 (2016) 45–57, <https://doi.org/10.1016/j.trc.2016.01.007>.
- [94] Z. Ma, B.N. Jørgensen, Z. Ma, Fuel efficiency analysis of the public transportation system based on the Gaussian Mixture Model clustering. *EPIA Conference on Artificial Intelligence*, Springer, 2024, pp. 209–219, https://doi.org/10.1007/978-3-031-73500-4_18.
- [95] Met Éireann, Climate of Ireland, (2023) (<https://www.met.ie/climate/what-we-measure/>).
- [96] Fáilte Ireland, Tourism facts and figures, (2023) (<https://www.failteireland.ie/>).
- [97] CSO, Transport Omnibus 2022, (2023) (<https://www.cso.ie/en/releasesandpublications/>).