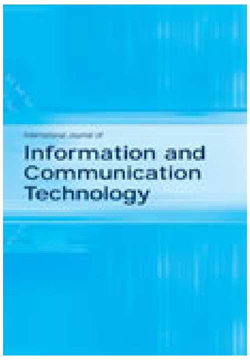


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Modelling and predicting energy consumption patterns using generative adversarial networks for effective carbon management

Yongting Liu, Baotong Li

DOI: 10.1504/IJICT.2026.10080316

Article History:

Received:	01 March 2026
Last revised:	02 May 2026
Accepted:	18 May 2026
Published online:	26 August 2026

Modelling and predicting energy consumption patterns using generative adversarial networks for effective carbon management

Yongting Liu* and Baotong Li

State Grid Xinjiang Electric Power Co., Ltd.,
Urumqi, 830000, No. 68, Nanhu East Road,
Shuimogou District, Urumqi, Xinjiang, China
Email: yonghao_leo@126.com
Email: jeydragon@126.com
*Corresponding author

Abstract: Effective carbon management requires accurate prediction of energy consumption patterns, yet conventional models struggle with complex temporal dependencies. This study proposes a novel integrated framework combining transformer-based generative adversarial networks with Bayesian optimisation (BO-TransGAN) for energy forecasting and carbon optimisation. Using 2,000 hourly records of energy use and emissions, data were pre-processed via imputation, outlier removal, and normalisation. The TransGAN captures nonlinear temporal dependencies through adversarial learning, while Bobcat optimisation tunes hyperparameters for enhanced convergence and stability. BO-TransGAN achieves 0.987 accuracy, 0.995 R^2, minimal losses, and low training (2.12 s) and run times (7.35 s). It generates realistic synthetic sequences and provides actionable insights for reducing carbon emissions, offering a scalable tool for sustainable energy planning and real-time carbon management.

Keywords: carbon management; generative adversarial networks; GANs; energy consumption prediction; sustainable energy planning; energy-carbon pattern modelling.

Reference to this paper should be made as follows: Liu, Y. and Li, B. (2026) ‘Modelling and predicting energy consumption patterns using generative adversarial networks for effective carbon management’, *Int. J. Information and Communication Technology*, Vol. 27, No. 92, pp.79–110.

Biographical notes: Yongting Liu received his Master’s, is an Associate Senior Engineer, and is currently with State Grid Xinjiang Electric Power Co., Ltd. His research interests include small-scale power grid infrastructure project management, and energy-saving and carbon-reduction retrofitting of existing buildings.

Baotong Li received his Master’s, is an Associate Senior Engineer, and is currently with State Grid Xinjiang Electric Power Co., Ltd. His research interests cover intelligent operation and maintenance of power infrastructure and low-carbon planning.

1 Introduction

Carbon management strategies are dependent upon the accurate forecasting of energy use due to the fact that energy use is the largest contributor of global carbon-dioxide (CO_2) emissions, and the consideration of demand forecasting energy use to emissions is critical to climate value (Sharma et al., 2024). Reliable load and consumption forecasting allows utility operators, building managers, and policy makers to defer generation, shift demand, integrate renewables, and reduce peak load, each of which improves carbon footprints directly (Wan and Liu, 2023). Furthermore, reliable forecasting allows organisations to identify inefficiencies, target high-impact interventions, and avoid exacerbating carbon emissions before they are ever produced (Keerthana et al., 2023). Recent data-driven forecasting methods have demonstrated the efficacy of pairing energy predictions and carbon-based metrics, which improves strategic decision-making for environmentally sustainable initiatives while aligning with carbon-neutrality objectives (Nassef et al., 2023). Not having reliable forecasting models could result in carbon management strategies that are reactive and not proactive, which exposes mid-to long-term decarbonisation goals and increases climate-related hazards (Maheen et al., 2023).

Conventional AutoRegressive integrated moving average (ARIMA)-derived forecasting approaches fall short in handling the dynamic and nonlinear patterns seen in modern energy usage data: assumptions of linearity and stationarity, limitations with high-dimensional exogenous inputs, and brittleness in the face of non-stationary trends and sudden regime shifts typical of smart grids (SGs) (Sun et al., 2024). They necessitate some form of manual or coded differencing, requiring advanced feature engineering, which is not scalable across a large heterogeneous building or site context (Pierre et al., 2023). Moreover, ARIMA's point estimates cannot capture complex nonlinear temporal dependencies nor account for multimodal distributions of real-world load time series, making them poor representations of realistic synthetic sequences that are needed for scenario testing and carbon-impact simulations (Al-Duais and Al-Sharpi, 2023). Empirical evaluations have found that hybrid and deep learning (DL) approaches have provided superior results than ARIMA, particularly regarding short-and medium-term horizons, and particularly with high volatility in the load and when working with meteorological or events-based datasets (Tarmanini et al., 2023). ARIMA and its variants can be sensitive to noise and do not provide effective performance in transfer learning contexts, which would limit application in data-limited areas, while data-driven nonlinear models provide increased robustness, adaptability, and greater representation of uncertainty, which is needed for proactive carbon management and demand-side interventions (Son and Shin, 2023).

Energy consumption data would always display nonlinear characteristics, temporal dependencies, and contextual variability from environmental and behavioural contexts, which could all be captured by models (Ji et al., 2023). Even statistical models and their hybrid approaches are unlikely to be able to model long-range dependencies and contextual interactions among multiple variables driving consumption. DL architectures, including recurrent neural networks (RNNs), long short-term memory (LSTM), and transformer-based architectures, are superior at learning sequential correlations and establishing temporal hierarchies without extensive manual feature engineering (Cavus and Allahham, 2025; Balakrishnan et al., 2025). These architectures can recall previous states of the system and update for any non-stationary variations, enabling more accurate and contextually informed predictions that are vital for energy optimisation and carbon management (Rani et al., 2025). Furthermore, recent advances combining attention mechanisms and generative modelling frameworks enable these architectures to forecast energy consumption but also to generate realistic synthetic sequences for simulation and policy evaluation, enabling data-driven technologies and enhanced sustainability and demand side management (Rastgoo et al., 2024). A smart framework is created that can accurately model and predict energy use to enable the management of emissions. The BO-TransGAN is designed to learn complex temporal dependencies and variations in the environment based on a dataset of energy use. The integration of adversarial learning with optimisation for carbon-aware systems would result in improved forecasts, stability, and sustainability outcomes.

While prior works have explored GANs for time-series generation and metaheuristics for hyperparameter tuning, the present study contributes a unified framework that co-optimises forecasting accuracy and carbon emission deviation through a specifically designed adversarial objective. Unlike generic hyperparameter tuning, the Bobcat Optimiser is adapted to search for configurations that simultaneously minimise prediction error and carbon inconsistency, a requirement unique to carbon-conscious energy management. This joint optimisation, coupled with the transformer-based adversarial architecture, forms the novel methodological core of BO-TransGAN.

2 Related works

More recently, machine learning (ML), DL, and hybrid models have been used in the literature for forecasting energy use to increase prediction accuracy and the sustainability of carbon management. Most of the literature has focused either on traditional approaches, artificial intelligence (AI) approaches, or transformer approaches for different applications.

The different techniques for load forecasting (LF) were analysed to determine the reliability and efficiencies of SGs (Habbak et al., 2023). The evaluation was of traditional, clustering-based, AI, and time series-derived prediction methods. Findings indicated that AI-derived approaches, ML and neural network (NN) approaches were more reliable, achieving higher accuracy and lower positive error values. However, there was a problem with the excessive reliance on established models without proposing an integrated framework for real-time SG applications. An AI-integrated energy forecasting framework utilised to manage complex power systems and variable renewable energy was examined (Bacanin et al., 2023). The framework was used LSTM and gated recurrent unit (GRU) NNs that were optimised through an improved sine cosine algorithm (SCA) for effective hyperparameter optimisation. The outcomes of the experimental results demonstrated more reliable forecasting accuracy with the lowest mean squared error (MSE) values recorded for solar, wind, and grid LF at 0.0132, 0.00287, and 0.01504. However, the possible drawbacks included both computational complexity and unsuitability to larger real-time datasets. The forecast accuracy and generalisation of renewable energy were enhanced for solar radiation and wind speed (Alghamdi et al., 2023). A hybrid optimisation algorithm was utilised for optimising a stacked ensemble prediction model. After experimenting on two real-world datasets, it was shown that the genetic algorithm-Al-Biruni Earth Radius (GABER) method outperformed in terms of accuracy, effectiveness, and generalisability. However, the downside of the model was that it adds computational complexity.

The accuracy of predictions of sequential energy usage data within manufacturing processes was examined (Tang et al., 2025). An advanced TimeGAN was applied for augmentation, and an integrated convolutional neural network-GRU (CNN-GRU) handled prediction data. It increased an R^2 score, showed an enhanced prediction performance. While a performance of the model was enhanced, its effectiveness was dependent on the ideal ratio of synthetic to real data. The LSTM model was applied to estimate energy consumption in a steel manufacturing plant to enhance energy efficiency in energy-consuming facilities (Uzun et al., 2025). The LSTM approach was compared to a back-propagation neural network (BPNN) models using operational and weather data. The results showed the LSTM method performed with lower error rates, over 3.6% and 33.3% less than BPNN and regression models. However, the LSTM method was limited to historical data from one facility, which limited generalisation to the industrial context. Forecasting mid or long-term electricity use in industrial forging operations enables efficient power systems planning and contributes to reducing environmental impacts (Moon et al., 2024). An extreme gradient boosting (XGBoost) and an LSTM Autoencoder (LSTM-AE) were forecasted electricity consumption and detected anomalies. The model gave high prediction performance, with a mean absolute error (MAE) of 0.020, MSE of 0.021, and a symmetric mean absolute percentage error (SMAPE) of 4.24, suggesting efficacy for energy management systems. However, a restriction of the model was its reliance on high-quality and process-specific data.

A framework for carbon-conscious grid management was used to minimise emissions with grid stability considerations (Otshwe et al., 2025). The modelling framework utilised graph convolutional networks (GCNs) as the main ML component, combined with other hybrid optimisation techniques. Validation for the experimental framework demonstrated efficacy using a 24-bus microgrid and showed optimised operation using particle swarm optimisation (PSO) techniques could produce the lowest emissions (1.59 kg CO_2) in an efficient computing time frame. However, the framework's limitations scale with the complexity of the computing. The deep reinforcement learning (DRL) optimised the models for low-carbon economic dispatch (Zhong et al., 2024). The modelling framework produced a method for a dynamic calculation of carbon intensity and low-carbon economic dispatch models. Simulations demonstrated a significant reduction in operating costs along with reductions in carbon emissions in the modelled framework. These results showed both economic and environmental performance for the modelling framework. However, its applicability was limited by computational complexity and dependency on high-quality real-time energy data.

ML models with SARIMA were compared for predicting heating energy in buildings (El Alaoui et al., 2023). Bagging, boosting, support vector machine (SVM), artificial neural network (ANN) and 14 SARIMA models were used, where ANN performed best with $R = 0.97$ and a normalised root MSE (nRMSE) of 12.60%, although the reliance on simulated building data limited real-world applicability. The predictive models were assessed for industrial energy consumption using historical energy data (Sarswatula et al., 2022). Multi-linear regression (MLR), random forest (RF), decision tree (DT), XGBoost, SVM, KNN were used. RF classifier attained an accuracy = 0.883. The used industrial assessment centres (IAC) dataset was outdated, and performance depends on feature diversity were the limitations. An electricity consumption forecasting was enhanced through a hybrid DL model (Ragupathi et al., 2024). Deep energy predictor model (DEPM) integrating XGBoost and deep neural network (DNN) was used. The approach attained 98% accuracy and F1 = 0.98. However, the approach was computationally intensive, and required graphics processing unit (GPU)-based systems for training. The residential heating energy prediction was investigated using optimised ensemble learning (Dinmohammadi et al., 2023). It used PSO-optimised RF and stacking ensemble. Stacking model achieved 95.4% accuracy, and a causal inference identified key factors. However, the limitations were high model complexity, and limited scalability for large datasets. An energy consumption was forecasted using a novel LSTM model optimised by a hybrid algorithm (Khafaga et al., 2023). Dynamic DTOSFS + LSTM combined dipper-throated optimisation (DTO) and stochastic fractal search (SFS). It achieved lowest RMSE of 0.00013. But the complex parameter tuning would face overfitting with small datasets. The residential electric consumption was predicted using an LSTM for early building design (Sundaram et al., 2024). It used LSTM with DNN and ANN. LSTM achieved R^2 of 0.97, and it required only 2.69 s for 500 test cases. Focused on specific residential settings limited cross-domain applicability. The building energy consumption was predicted using transformer-based DL (Nazir et al., 2023). Temporal fusion transformer (TFT) with LSTM, Interpretable LSTM, and temporal convolution network (TCN) were used. TFT achieved MSE = 4.09, RMSE = 2.02, MAE = 1.50, sMAPE = 26.46%. Model training requires large labelled datasets and high computational resources were the limitations.

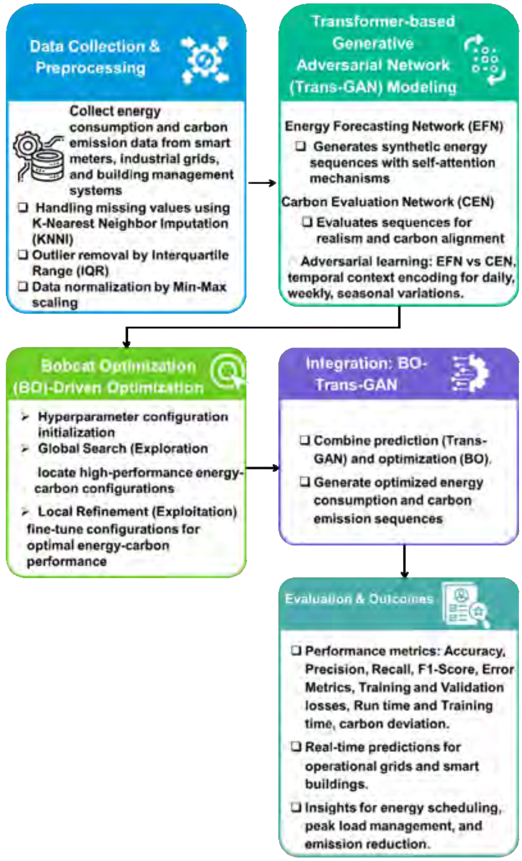
2.1 Research gaps

Though substantial advancements have been made in energy forecasting and carbon-conscious decisions, several instantaneous limitations remain in overall models. AI-driven and time series forecasting approaches (Habbak et al., 2023) continue to lack a systematic integration of real-time smart-grid applications or guidelines, while SCA-optimised LSTM-GRU models (Bacanin et al., 2023) have scalability and computation-related disadvantages. Hybrid ensemble-type methods, such as GABER (Alghamdi et al., 2023), provide increased accuracy but remain labour-intensive. TimeGAN and CNN-GRU approaches (Tang et al., 2025) heavily rely on the appropriate ratio of synthetic to real data input to achieve stness. Models pertaining to LSTM at the industrial level (Uzun et al., 2025) rely on specific information for specific sites and cannot be generalised; and ensemble LSTM-AE methods (Moon et al., 2024) were dependent on the quality of datasets and for certain processes. GCN-based carbon management systems (Otshwe et al., 2025) continue to face large-scale stochastic processing issues, while DRL-based dispatch frameworks (Zhong et al., 2024) were limited by computation overhead and their dependence on vast datasets. The BO-TransGAN fills these gaps by an agreement on optimal hyperparameter optimisation, efficient adversarial learning, and carbon-aware adaptation, providing scalable, accurate, and computationally efficient energy forecasting.

3 Methodology

The methodology describes the procedure used to model and forecast energy usage patterns for carbon management. The research uses a BO-TransGAN to create realistic energy sequences with a carbon optimisation model. This included data pre-processing, feature engineering and hyperparameter tuning to ensure model stability and convergence and high predictive power, and Figure 1 displayed the flow diagram.

Figure 1 Workflow of the BO-TransGAN framework for energy usage modelling and carbon optimisation (see online version for colours)



3.1 *Energy data pre-processing framework*

Data pre-processing describes the organised conversion of raw datasets related to energy and emissions into an analysable standard format appropriate for predictive modelling. The pre-processing techniques used are KNNI of missing values, IQR-based removal of outliers, and min-max normalisation of data.

3.1.1 *KNNI for energy consumption data*

When predicting energy consumption and associated carbon emissions, data being absent in critical aspects of the information can limit the usefulness of prediction models. The KNNI method approximates the misplaced information through similarities of records. The similarity between two energy consumption records *g* and *h* are calculated using a modified Euclidean distance, as defined in equation (1).

$$D(g,h)=\sqrt{\sum_{f=1}^{F}(g_f-h_f)^2} \tag{1}$$

Here, $D(g, h)$ represents the distance between records g and h, g_f and h_f are the normalised values of the f^{th} feature, and F is the total count of the attributes considered. Once the NN are identified, the missing values for a feature in record h is computed by averaging the corresponding values of the k closest neighbours using equation (2).

$$U_h = \frac{1}{K}\sum_{k=1}^{K} F_k \tag{2}$$

Here, U_h is the imputed value for the missing feature, F_k denotes the observed feature value from the k^{th} closest neighbour, and K is the total neighbours used. This preserves correlations among energy, environmental, and occupancy variables.

3.1.2 Outlier removal for reliable energy and carbon prediction

To ensure predictions of energy use and carbon emissions are valid and robust, outliers that would affect model performance are tracked and eliminated using an IQR method. The lesser (LQ) and higher (UQ) energy-carbon thresholds are calculated for each metric to capture the 25th and 75th percentiles of each metric. The IQR is defined as follows in equation (3).

$$IQR = UQ - LQ \tag{3}$$

In summary, data instances are considered outliers if their energy or carbon values are observed to be outside of the computed thresholds found in equations (4)–(5).

$$Lower\ Energy\ Bound = LQ - 1.5 \times IQR \tag{4}$$

$$Upper\ Energy\ Bound = UQ + 1.5 \times IQR \tag{5}$$

Values that are considered lower than the lower bound or greater than the upper bound were flagged and removed, which removed outliers from the data for energy consumption (kWh) and carbon emissions (kg CO_2). This is important to preserve the integrity of the data for the BO-TransGAN to learn from real and valid behaviour, which improves the overall accuracy and robustness of the predictions.

3.1.3 Normalisation for energy parameters

All features were scaled to the [0, 1] interval through min-max normalisation [equation (6)] to eliminate scale-induced bias and to stabilise training.

$$N_{scaled} = \frac{A - A_{\min}}{A_{\max} - A_{\min}} \tag{6}$$

Here, A is an original value, $A_{\max}$ and $A_{\min}$ represent the larger and smaller values of the respective energy or carbon parameter, and N_{scaled} denotes the normalised values used for model training. This normalisation approach assures that all energy and carbon parameters contribute proportionally to the model during training.

3.2 BO-TransGAN framework for optimised energy consumption and carbon forecasting

The suggested framework combines TransGAN with BO to improve accuracy and carbon efficiency in energy forecasts. TransGAN describes nonlinear temporal dependencies and environmental dynamics by generating realistic sequences of energy consumption and emissions through adversarial learning and self-attention mechanisms. BO serves as an adaptable optimisation layer to refine the model's parameters in an efficient manner resulting in less energy-carbon balance. The BO-TransGAN framework achieves better energy prediction accuracy and carbon forecasting sustainability through the combination of adaptable learning and optimisation stability.

To intuitively understand how each component contributes, consider the framework as an interactive feedback loop. The energy forecasting network (*EFN*) acts as a synthetic data generator that learns the complex, nonlinear temporal patterns hidden in real energy consumption and carbon emissions; its transformer architecture, equipped with self-attention and temporal context encoding, automatically captures long-range dependencies such as daily and seasonal cycles without manual feature engineering. The carbon evaluation network (*CEN*) serves as a critic that distinguishes real sequences from generated ones, but it is augmented with a carbon consistency term that penalises sequences deviating from the observed energy-carbon correlation, thereby compelling the EFN to produce not only realistic but also carbon-aware patterns. The adversarial min-max game between EFN and CEN drives the generator to produce sequences that are nearly indistinguishable from real data, sharpening prediction fidelity. Bobcat Optimisation then adaptively searches the hyperparameter space in two phases: a broad energy-pattern search identifies high-performance regions, and a fine-grained adaptive refinement stabilises convergence on configurations that simultaneously minimise forecasting error and carbon deviation. This joint optimisation directly targets a composite objective that balances accuracy with carbon consciousness, explaining the model's high accuracy, R^2, and low training/inference times. Together, adversarial learning enforces statistical realism, self-attention models temporal complexity, and Bayesian optimisation aligns hyperparameters with carbon-aware goals, forming a coherent system that is both accurate and sustainable.

3.2.1 TransGAN for energy consumption and carbon forecasting

Accurately modelling energy consumption behaviour is essential to sustainable carbon management and energy planning. Most conventional models, however, cannot adequately capture nonlinear temporal structures in actual energy consumption and emissions data. To address these limitations in conventional modelling, TransGAN is introduced, comprising self-attention mechanisms and adversarial learning capabilities within a modelling framework, presented in Figure 2, to help capture nonlinear temporal structures and contingent environmental dynamics of carbon intensity.

The EFN produces synthetic energy consumption sequences, and the architecture is shown in Figure 3, while the CEN, demonstrated in Figure 4, identifies whether an input sequence originates from real observations.

Figure 2 Framework of the TransGAN for energy forecasting (see online version for colours)

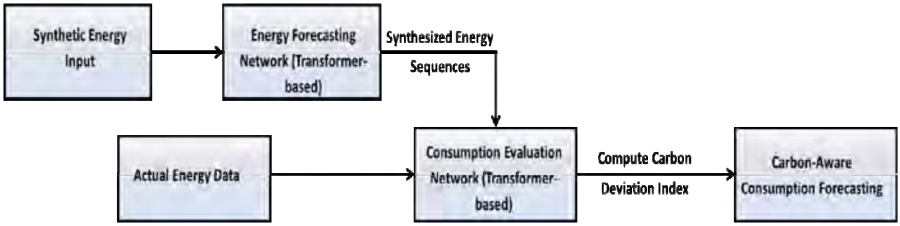
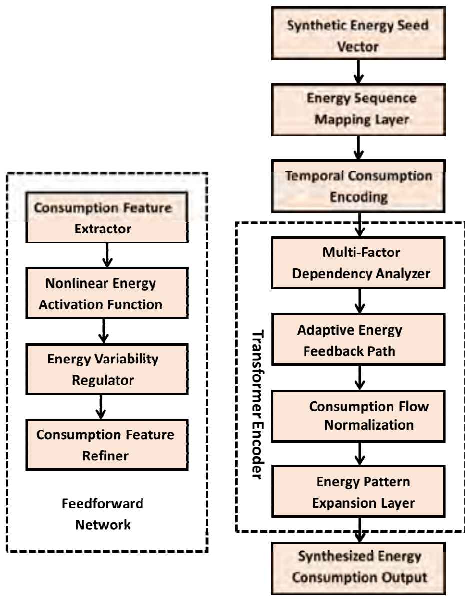


Figure 3 Transformer synthesiser for energy sequence generation (see online version for colours)



The adversarial learning process is governed by the min-max function $\min_\alpha \max_\beta(\text{CEN}_\beta, \text{EFN}_\alpha)$ mentioned in equation (7).

$$\begin{aligned}\min_\alpha \max_\beta \left(\text{CEN}_\beta, \text{EFN}_\alpha\right) = {} & F_{C_r \sim Q_{real}(C)}\left[\log \text{CEN}_\beta\left(C_r\right)\right] \\ & + F_{C_w \sim Q_w(w)}\left[\log\left(1 - \text{CEN}_\beta\left(\text{EFN}_\alpha\left(C_w\right)\right)\right)\right]\end{aligned} \tag{7}$$

Here, $F_{C_r \sim Q_{real}(C)}$ represents the real consumption sequences, C_r, C_w denote real and latest noise inputs, $Q_w(w)$ refers to the prior probability distribution $F_{C_w \sim Q_w(w)}$ of latent variables, α and β are the learnable parameters EFN_α, and CEN_β are EFN, and CEN on the learnable parameters, and $Q_{real}(C)$ denotes the probability distribution of true energy consumption data. To preserve the sequential dependencies between hourly, daily, and seasonal variations, temporal context-encoding Γ is incorporated in the transformer blocks, defined as in equation (8).

$$\begin{cases} \Gamma(\tau, 2h) = \sin\left(\dfrac{\tau}{10,000^{2h/g_e}}\right) \\ \Gamma(\tau, 2h+1) = \cos\left(\dfrac{\tau}{10,000^{2h/g_e}}\right) \end{cases} \tag{8}$$

Here, τ is the timestep, h refers to the embedded dimension, g_e represents the embedding depth which conveys temporal energy features, $10,000^{2h/g_e}$ acts as the modulation factor for energy, and $h + 1$ acts as a temporal index. The dynamic energy focus method to relate energy attributes $Focus(E_c, E_d, E_v)$ is defined in equation (9).

$$Focus\left(E_c, E_d, E_v\right) = softmax\left(\frac{E_c, E_d}{\sqrt{\eta}}\right) E_v \tag{9}$$

Here, E_c is the current energy demand signal, E_d is the historical energy distribution dependencies, E_v denotes the contextual variations in carbon emission and energy load, and η is the dimensional factor. The softmax operation dynamically weights the most relevant temporal features. During the feature alignment phase, the output of the EFN is mapped to the CEN feature space domain to ensure representations are believable. The alignment loss Ralign is as shown in equation (10).

(10)

Here, $C_{true}(C_t)$ and $C_{gen}(\text{EFN}_\alpha(C_w))$ denote the feature embeddings of real and synthetic sequences learned from the CEN. The aim of the EFN is to minimise the adversarial and find alignment loss R_{EFN} as stated in equation (11).

$$R_{EFN} = -F_{C_w \sim Q_w(w)}\left[\text{CEN}_\beta\left(\text{EFN}_\alpha\left(C_w\right)\right)\right] + \varphi_1 \cdot R_{align} \tag{11}$$

In this instance, φ signifies the regularisation coefficient. The objective of the CEN term R_{CEN} is to distinguish real sequences from artificial sequences while maintaining emissions fidelity. It is expressed as in equation (12).

$$\begin{aligned} R_{CEN} &= F_{C_r \sim Q_{real}(C)}\left[\text{CEN}_\beta\left(C_r\right)\right] - F_{C_w \sim Q_w(w)}\left[\text{CEN}_\beta\left(\text{EFN}_\alpha\left(C_w\right)\right)\right] \\ &\quad + \varphi_2 \cdot F_{C_r \sim Q_{real}(C)}\left[\left(C_{pred}\left(C_r\right) - C_{true}\right)^2\right] \end{aligned} \tag{12}$$

Here, $C_{pred}(C_r)$ and C_{true} represent the estimated and observed levels of carbon emissions, and φ_2 adjusts a weight of the carbon emissions, or consistency. The final joint optimisation objective R_T is shown in equation (13).

$$R_T = R_{EFN} + R_{CEN} \tag{13}$$

Table 1 establishes the final configuration and optimal values used to train a TransGAN in forecasting energy use and carbon emissions.

Figure 4 Transformer evaluator for energy validation (see online version for colours)

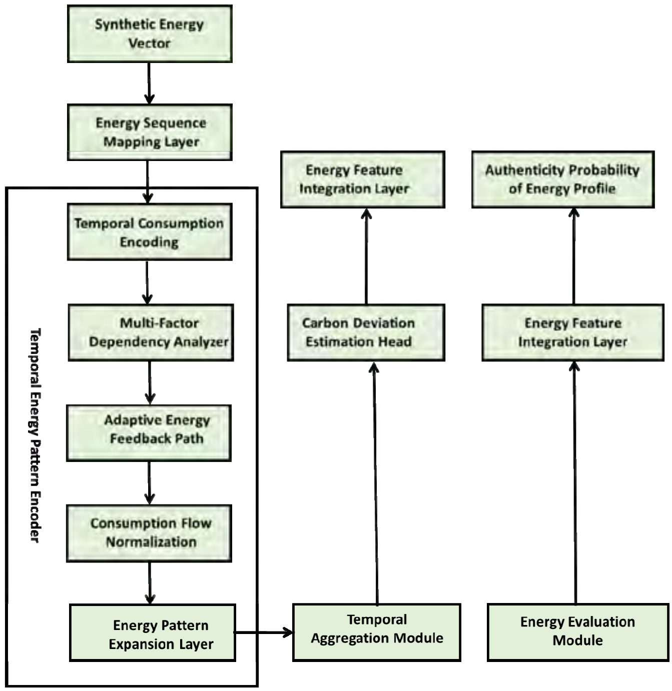
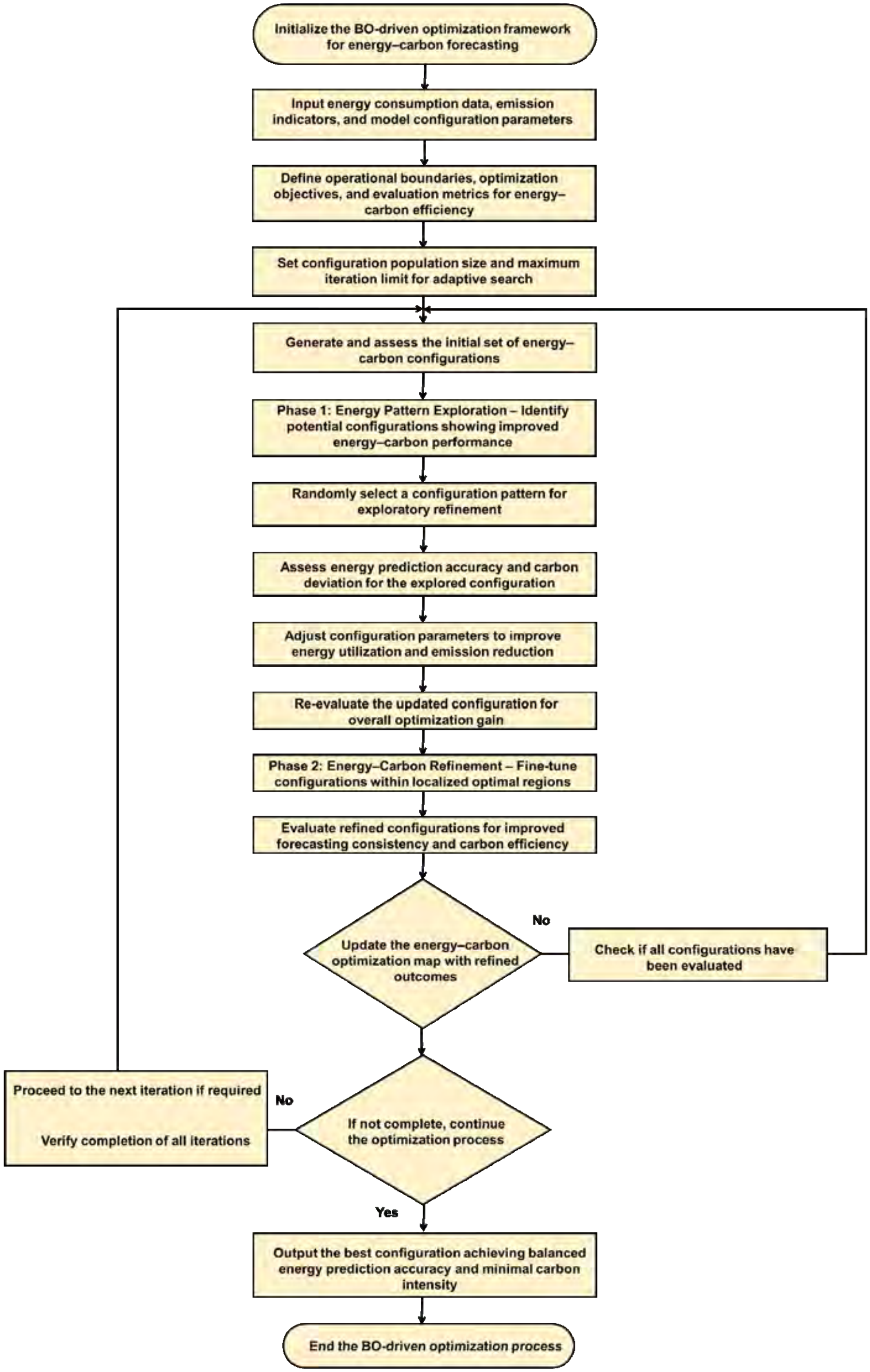


Table 1 Hyperparameter tuning configuration for the TransGAN

Parameter	*Search range/tested values*	*Optimal value*	*Associated network*
Batch size	{16, 32, 64, 128}	64	EFN, CEN
Embedding depth (g_e)	{32, 64, 128, 256}	128	EFN
Number of transformer layers	{2, 4, 6, 8}	6	EFN, CEN
Attention heads	{2, 4, 8, 12}	8	EFN
Dropout rate	{0.1, 0.2, 0.3, 0.4, 0.5}	0.2	EFN, CEN
Regularisation coefficients (φ_1, φ_2)	{0.1, 0.3, 0.5, 0.7, 1.0}	$\varphi_1 = 0.5$, $\varphi_2 = 0.7$	EFN, CEN
Optimiser	{Adam, RMSProp, SGD}	Adam	EFN, CEN
Latent dimension (Q_x)	{32, 64, 100, 128}	100	EFN
Epochs	{50, 100, 200, 300}	200	EFN, CEN
Activation function	{ReLU, LeakyReLU, GELU, ELU}	GELU	EFN
Gradient clipping	{0.1, 0.5, 1.0, 5.0}	1.0	EFN, CEN
Adversarial loss weight	{0.2, 0.5, 1.0, 2.0}	1.0	EFN

Figure 5 Methodological workflow of BO-TransGAN for energy consumption and carbon forecasting (see online version for colours)



3.2.2 BO-driven adaptive configuration strategy for energy-carbon optimisation

The BO operates in two phases:

1 a broad exploration phase that identifies high-performance regions of the hyperparameter space

2 a local exploitation phase that fine-tunes the best configurations. The workflow is shown in Figure 5.

The algorithm starts with the initialisation by representing configuration instances for various apparent policies distributed throughout the multi-dimensional energy-carbon optimisation space. Each configuration unit indicates a possible hyper-parameter configuration for the TransGAN model, which influences the accuracy of energy consumption forecasting and carbon efficiency. Mathematically, the overall position of the configuration unit H is denoted as follows in equation (14).

$$H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_m \end{bmatrix}_{m \times k} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,k} \\ \vdots & \ddots & \vdots \\ h_{m,1} & \cdots & h_{n,k} \end{bmatrix}_{m \times k} \tag{14}$$

Here, H_1, H_2, …. H_m are the configuration performance matrix for BO-driven energy-carbon optimisation, m is the total evaluated configuration, k is the total configuration parameters, $h_{1,1}$ …… $h_{n,k}$ are the values of the k^{th} control parameter in the m^{th} configuration vector. Each position is initialised randomly within defined lower L_k and upper limits U_k using equation (15).

$$h_{n,k} = L_k + \rho.(U_k - L_k) \tag{15}$$

Here, $\rho \in [0, 1]$ is the random uniform scaling. Each position of a configuration unit represents a possible configuration of which the potential capability is determined by the energy-carbon objective function, where forecasting error and emissions rate are utilised. The evaluation Φ is denoted as follows in equation (16).

$$\Phi = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_m \end{bmatrix}_{m \times 1} = \begin{bmatrix} F(H_1) \\ F(H_2) \\ \vdots \\ F(H_m) \end{bmatrix} \tag{16}$$

Here, Φ_m is the objective function for the m^{th} configuration unit, and $F(H_m)$ refers to the function combining prediction loss and carbon deviation metrics.

3.2.2.1 Phase 1: energy flow optimisation

During the energy pattern search phase, each of the configuration units identifies other configurations that yield better energy-carbon balance and begins to move into that optimisation refinement phase on the search for globally optimal regions. The set of configuration units producing better performance is given by equation (17).

$$\Omega_m = \{H_g : \Phi_g < \Phi_m, g \neq m\} \tag{17}$$

Here, Ω_m is the set of effective energy configuration for the m^{th} configuration unit, H_g is the energy position of g^{th} configuration, and Φ_g, Φ_m denote the performance metric of g^{th} and m^{th} configuration. Each configuration unit H_m selects one effective configuration from Ω_m and updates its energy parameters according to equation (18).

$$H_m = \begin{cases} H'_m, & if \ \Phi(H'_m) \leq \Phi(H_m) \\ H_m, & otherwise \end{cases} \tag{18}$$

Here, H'_m is an updated energy configuration, $\Phi(H_m)$ and $\Phi(H'_m)$ is the evaluated energy-carbon efficiency of the current and updated states. This framework creates a dynamically diverse search direction for the algorithm to locate regions associated with high efficiency in energy consumption and low carbon intensity.

3.2.2.2 Phase 2: adaptive refinement for energy-carbon optimisation

Once the high-performance configurations are located, the configuration units fine-tune their positions through incremental adjustments, which analogise stabilisation of energy and optimisation of carbon. The fine-tuned position for localised exploitation would be generated by equations (19)–(20).

$$H''_{m,k} = H_{m,k} + \frac{(1 - 2.\lambda_{m,k})}{1 + \eta.H_{m,k}} \tag{19}$$

$$H_m = \begin{cases} H''_m, & if \ \Phi(H''_m) \leq \Phi(H_m) \\ H_m, & otherwise \end{cases} \tag{20}$$

Here, $H''_{m,k}$ is the refined configuration, $H_{m,k}$ is an updated configuration unit, η is an adjustment factor, $\lambda_{m,k}$ is the random factor, and $\Phi(H''_m)$ s an updated objective. This stage allows the model for localised fine-tuning of the transformer parameters and develops capabilities for optimal convergence on the balance of energy consumption accuracy with respect to carbon minimisation. A pseudocode of the BO-TransGAN approach for energy consumption forecasting is outlined in Algorithm 1.

Algorithm 1 BO-TransGAN framework for energy – carbon optimisation

Step 1: Trans-GAN model
Initialise EFN and CEN networks
for each training epoch:
    Generate synthetic energy sequence G = EFN(noise)
    Evaluate realism D = CEN(G)
    Compute losses:
        L_adv = log(CEN(real)) + log(1 – CEN(G))
        L_align = ||real_features – generated_features||^2
    Update EFN to minimise (L_adv + φ_1 * L_align)
    Update CEN to maximise (L_adv – φ_2 * carbon_error)

```
    if convergence_criteria met:
        Save EFN,CEN models
    else:
        Continue training
Step 2: Bayesian optimisation (BO)
Initialise configuration space for model parameters
for each configuration P:
    Evaluate energy-carbon objective Φ(H)
    if Φ(H) < best_score:
        best_H = H
    else:
        Adjust P using exploration factor ρ
Refine top configurations for local optimisation
Step 3: BO-Trans-GAN integration
for each BO iteration:
    Apply best_H to Trans-GAN hyperparameters
    Train Trans-GAN using updated settings
    if performance_improves:
        Retain configuration (energy accuracy ↑,carbon ↓)
    else:
        Revert to previous configuration
Output optimised Trans-GAN with minimum carbon deviation and high energy forecast
accuracy
```

The suggested algorithm utilises TransGAN to model complex temporal and spatial dependencies in energy consumption. BO then adaptively tailors the primary hyperparameters to improve the energy-carbon efficiency. The continuous iterations of the BO-TransGAN improve prediction accuracy and emissions optimisation for acceptably rapid and stable convergence to sustainable forecasting performance.

4 Results and discussion

An extensive examination of the BO TransGAN framework's performance with existing energy forecasting models was presented. It presents an examination into the forecasting accuracy, calculation efficiency, and loss metrics to demonstrate how the model forecast includes future complex temporal-data dependencies and contributes to carbon-aware energy management. Table 2 presents the experimental configuration of the BO-TransGAN framework for sustainable energy planning.

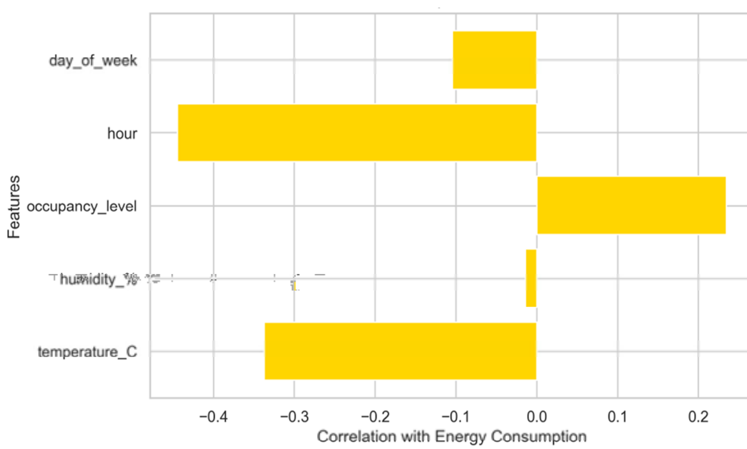
4.1 Dataset description and scope

The Energy Consumption and Carbon Emission Dataset from the Kaggle open site was an hourly record of 2,000 records of carbon emissions and energy consumption for various buildings and industrial units. The dataset contains time-stamps, building identification (ID), temperature, humidity, occupancy, season, energy consumption (kWh), and carbon emissions (kg CO_2). It was observational of temporal variability, environmental variables, and occupancy patterns, and suitable for the analysis of energy consumption patterns and carbon management.

Table 2 Hardware and software for the BO-TransGAN framework

Specification	*Details*
Processor (CPU)	Intel Core i7-12700K, 12 cores
Graphics (GPU)	NVIDIA RTX 3080, 10 GB VRAM
RAM	32 GB DDR4
Storage	1 TB NVMe SSD
Operating system	Windows 11
Programming language	Python 3.10
DL libraries	TensorFlow 2.x, PyTorch 2.x
Optimisation tool	Bobcat Optimisation Library
Visualisation	Matplotlib, Seaborn

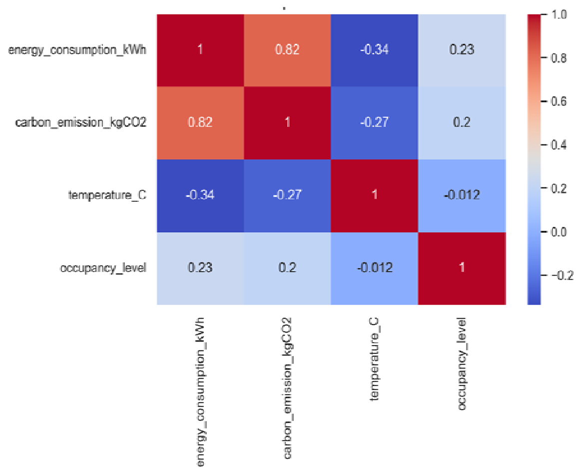
Figure 6 Feature importance for energy consumption drivers (see online version for colours)



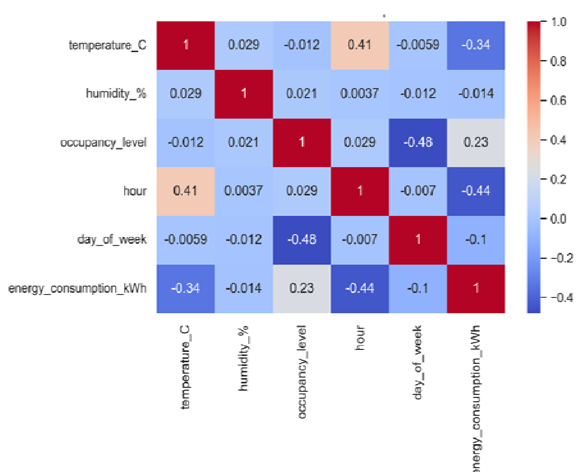
Source: https://www.kaggle.com/datasets/zyan1999/energy-consumption-and-carbon-emissions-dataset/data

Figure 6 illustrates the linear relationships between total energy consumption (kWh) and the time-series and environmental characteristics. Among the five selected variables, occupancy level exhibits the strongest positive linear relationship with energy consumption. This implies that occupancy-driven loads dominate the consumption profile, making occupancy-based demand management a high-leverage carbon reduction strategy. Conversely, the negative correlations with hour and temperature indicate that during early morning hours or in milder thermal conditions, baseline consumption decreases. These insights directly inform the carbon optimisation objective within the BO-TransGAN framework, which prioritises configurations that minimise emissions under varying temporal and environmental contexts.

Figure 7 Inter feature dependencies, (a) energy consumption and carbon emission correlation (b) temporal and environmental drivers of energy consumption (see online version for colours)



(a)

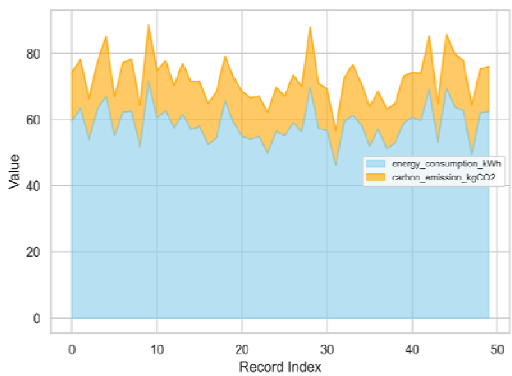


(b

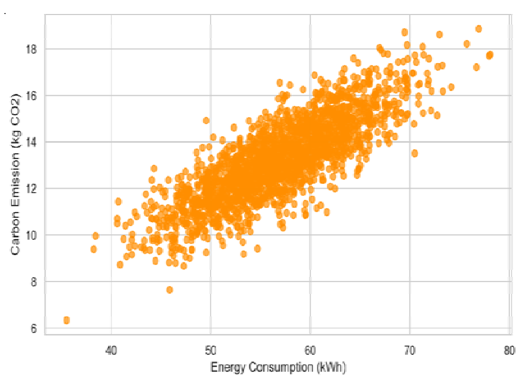
Figure 7 displays the associate relationships between (a) energy emission and (b) important predictor features in the BO-TransGAN model. The strong positive correlation (0.82) between energy consumption and carbon emissions establishes a near-linear proportionality, suggesting that each kWh reduction yields an almost constant marginal emission saving under the prevailing grid mix. This provides a straightforward metric for evaluating the carbon impact of energy efficiency measures. Nevertheless, the moderate negative correlations of temperature (–0.34) and hour (–0.44) with energy use reveal that environmental and temporal factors significantly modulate consumption independently of occupancy. Such conditional effects underscore the necessity of a model that can capture nonlinear interactions and time-varying emission factors. The BO-TransGAN's multi-head attention mechanism is specifically designed to weight these contextual variables dynamically, thereby generating carbon-aware forecasts that cannot be achieved through simple linear extrapolation.

Figure 8 demonstrates a significant linear relationship between energy usage and levels of carbon emissions, which was important for accuracy in carbon management. Figure 8(a) shows carbon emissions rise continuously with energy consumption during the observed period. Figure 8(b) quantifies and confirms a strong positive correlation of 0.82 between the two measures, affirming the idea that energy consumption reduction would lead to emission reduction. Coupled with temperature, hourly and occupancy factors, temperature and occupancy would play another role in shaping energy consumption patterns. The confirmed correlation serves as of one component of the appropriate foundation in developing the BO-TransGAN model to provide improved energy forecasts and improve carbon metrics.

Figure 8 Analysis of energy consumption with (a) carbon emission over time and (b) carbon proportionality (see online version for colours)



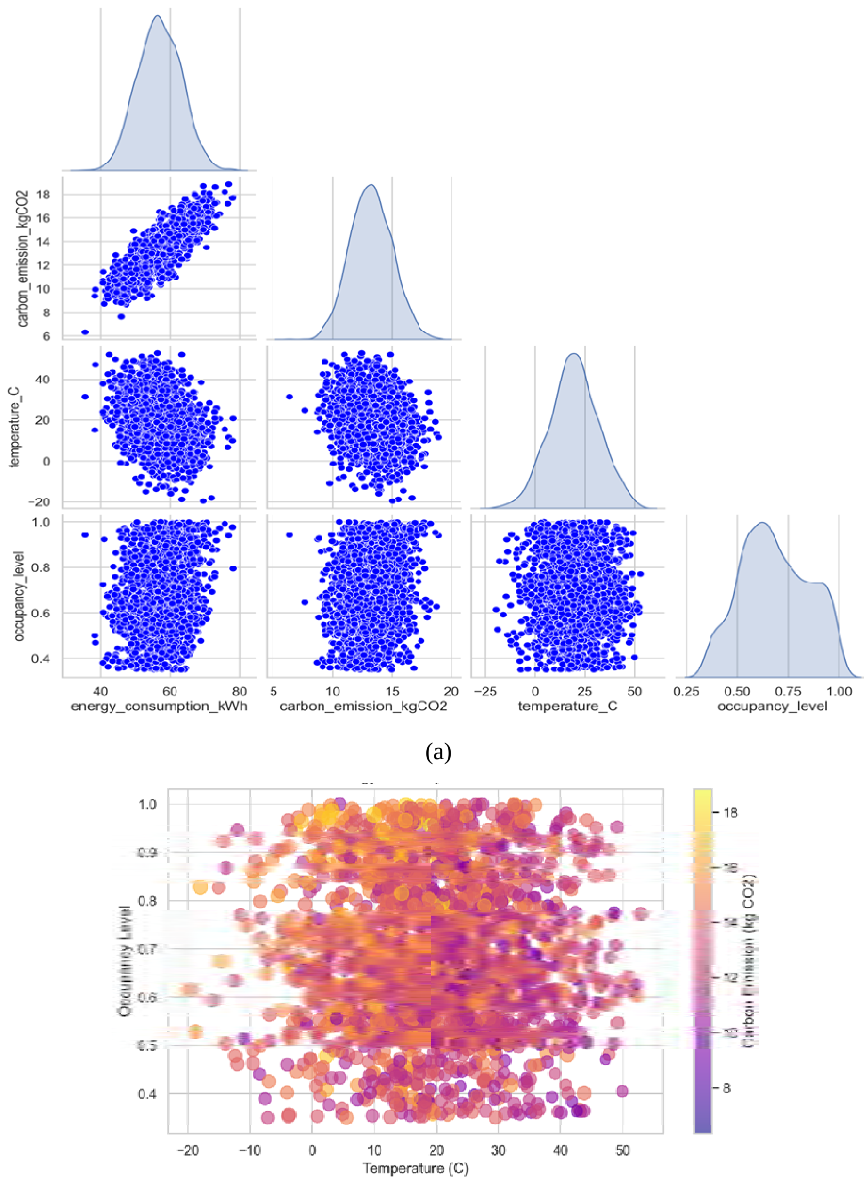
(a)



(b)

Figure 9 demonstrates the interaction of several environmental factors on energy consumption and carbon emission, supporting BO-TransGAN's forecasting function. Figure 9(a) demonstrates a simple linear relationship between energy consumption and carbon emissions but also shows, by occupancy and temperature, possibly independent functions and complex interaction effects. Figure 9(b) shows that high carbon emissions occur in a variety of temperature ranges, and typically occurs when occupancy is high. These patterns of relationship in features heightened the importance of a model that captures interactions between features or factors, strengthening the rationale for carbon aware forecasting and optimisation in the BO-TransGAN approach.

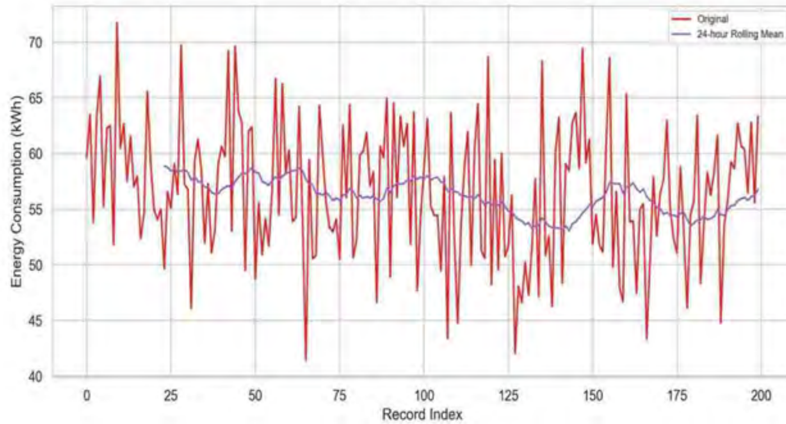
Figure 9 Distribution of (a) features and (b) temperature, occupancy, and carbon interactions (see online version for colours)



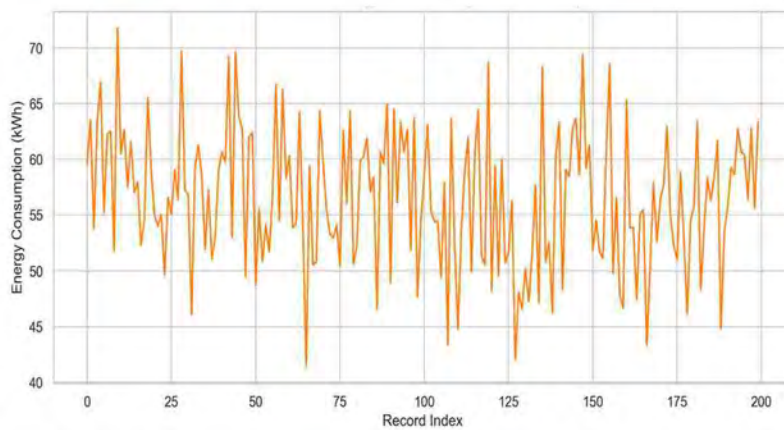
(a)

(b)

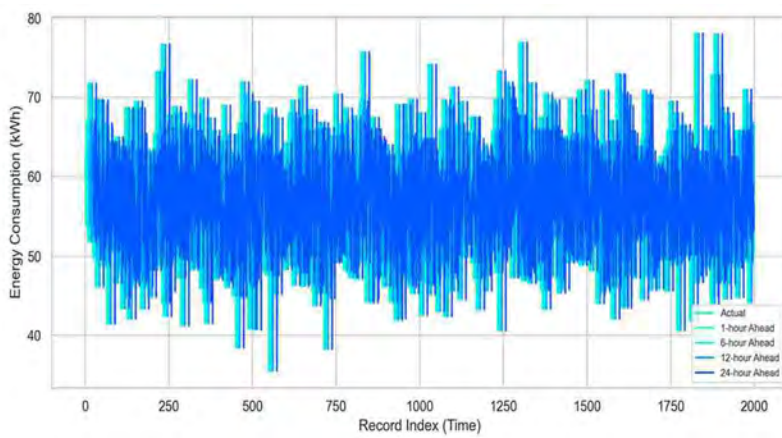
Figure 10 Evaluation of (a) energy volatility, (b) hourly fluctuation and (c) multi-step forecasting (see online version for colours)



(a)

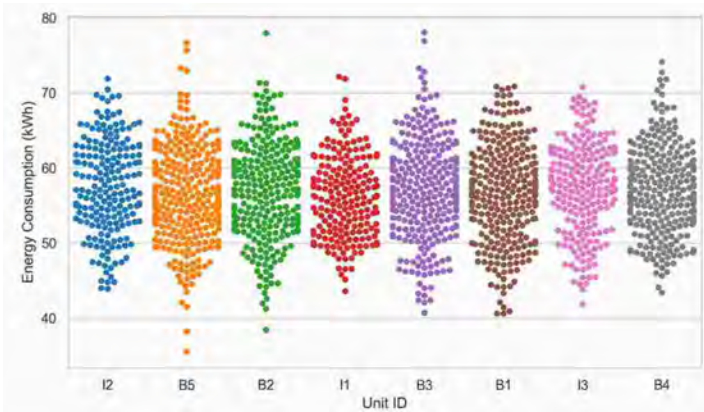


(b)

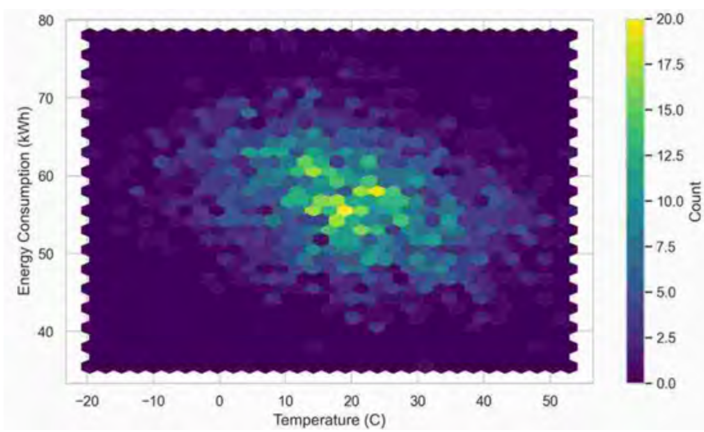


(c)

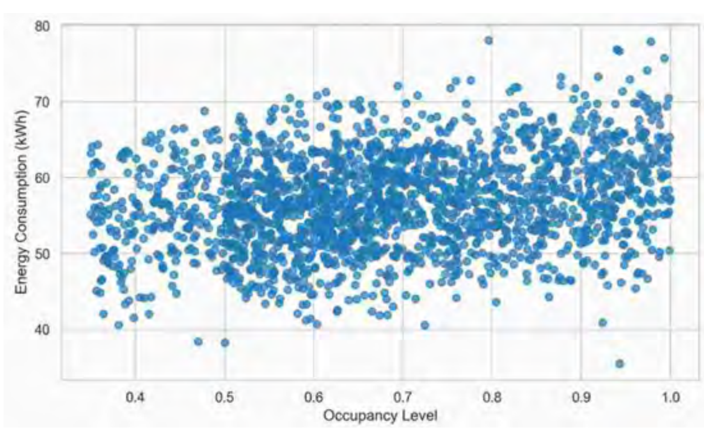
Figure 11 Assessment of (a) unit consumption, (b) temperature with energy density, (c) occupancy with energy trend and (d) autocorrelation check (see online version for colours)



(a)

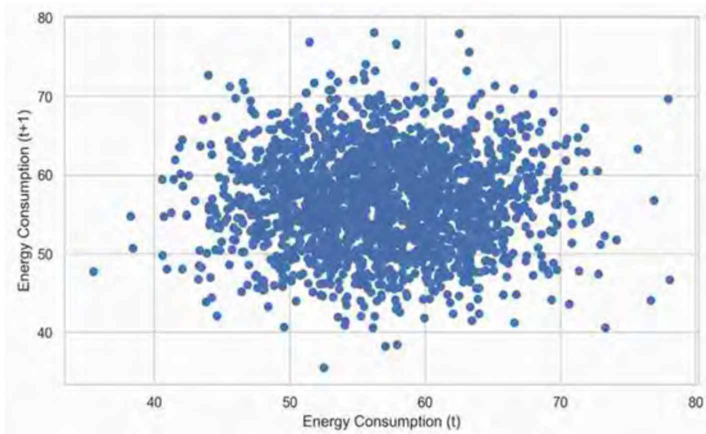


(b)



(c)

Figure 11 Assessment of (a) unit consumption, (b) temperature with energy density, (c) occupancy with energy trend and (d) autocorrelation check (continued) (see online version for colours)



(d)

Figure 10 confirms the time series properties of the energy consumption dataset and highlights the forecasting capability of the BO-TransGAN. Figure 10(a) illustrates the significant volatility in the original data, which indicates the presence of both noise and rich temporal dependencies that permit the use of the TransGAN. The volatility is observed in Figure 10(b), which indicated high inter-hour variability for the first 200 records. Figure 10(c) demonstrates the robustness of the BO-TransGAN model, as the forecasted consumption closely followed the actual data over the entire time series of 2,000 records with high accuracy, including effective longer horizon primitives for carbon managers.

Figure 11 examines unit-level heterogeneity and certain features related to energy use, which were required for precise forecasting and individualised carbon management. Figure 11(a) shows that even though distributions of energy use are largely consistent across units, there were small baselines and volatilities for each unit. Therefore, the TransGAN had to be able to model and capture multiple time series. Figure 11(b) and Figure 11(c) confirm the corresponding nonlinear relationship of temperature and a weak positive trend with occupancy. The confirmation of a need for a DL model to capture these complexities was ideal for the distribution of data and environmental characteristics and prior conditions. Figure 11(d) justifies the modelling of the deeper, nonlinear temporal dependencies through a learner deep in the transformer space to get the BO-TransGAN modelling estimates correct. These multi-scale dependencies validate the inclusion of temporal context encoding and the multi-layer transformer stack, which jointly allow the model to aggregate information across hours, days, and weeks without a priori feature engineering.

4.2 Comparative analysis of energy consumption prediction approaches

To assess the validity of the BO-TransGAN model, a thorough comparative evaluation was performed with previous energy forecasting models using key metrics, which were demonstrated in Table 3 and their mathematical formulations were given in equations (21)–(26).

This evaluation consisted of traditional ML models, such as ANN, bagging (BG), boosting (BT) (El Alaoui et al., 2023), SVM, RF, KNN (Sarswatula et al., 2022), logistic regression (LR), DNN, gradient boosting (GB), naves Bayes (NB), DT, XGBoost + DNN, optimised (Ragupathi et al., 2024), XGBoost, light gradient boosting machine (LightGBM), categorial boosting (CatBoost), stacking (Dinmohammadi et al., 2023), DL frameworks support vector regression (SVR), KNN, multi-layer perceptron (MLP) (Khafaga et al., 2023), LSTM, DNN (Sundaram et al., 2024) and LSTM, blocked LSTM, TCN, TFT (Nazir et al., 2023). Evaluated key performance metrics to demonstrate the gains on predictive accuracy, computational performance, and generalisability of the BO-TransGAN model for improving sustainability in energy management. To guarantee a fully fair comparison, all baseline models were trained and evaluated on the identical pre-processed dataset of 2,000 hourly records, using exactly the same 70 % training and 30 % validation split that was employed for the proposed BO-TransGAN framework. The raw data underwent the same KNN-imputation, IQR-based outlier removal, and min-max normalisation before any model was applied, eliminating variability from differing pre-processing pipelines. Hyperparameter tuning for every baseline was carried out systematically using either grid search or random search with 5-fold cross-validation on the training partition, exploring the parameter ranges recommended in their original publications; the configuration yielding the lowest validation loss was selected for final evaluation. All models were implemented in Python 3.10 with standard libraries (Scikit-learn, TensorFlow 2.x, PyTorch 2.x) and executed on the same hardware platform (Intel Core i7-12700K, NVIDIA RTX 3080, 32 GB RAM) to ensure consistent runtime measurements. The reported accuracy, R^2, MAE, and loss metrics for each baseline represent the mean of five independent runs initialised with different random seeds, and none of the baseline models were exposed to synthetic data generated by TransGAN. This unified protocol guarantees that the observed performance differences stem from model architecture and learning capacity rather than from discrepancies in data preparation, tuning effort, or computational environment. Table 4 compares the efficacy of previous ML and DL models in comparison with the BO-TransGAN model, and Figure 12 provides the visualisation of precision and F1-sore measures.

It was found that the standard methods of SVM, RF, and KNN achieved reasonable levels of accuracy and precision, while more advanced methods, such as XGBoost+ DNN achieved better performance. However, the BO-TransGAN model outperformed all of them, resulting in the highest accuracy of 0.987, the highest precision of 0.978 and 0.981 of recall and an F1 score of 0.989. Results show a BO-TransGAN model's ability to properly assist with real-time energy consumption modelling and improve prediction reliability. Table 5 compares multiple energy consumption forecasting models based on the two metrics – R^2 and MAE, and the graphical plot for R^2 was presented in Figure 13.

It is shown that the conventional and ML models of ANN, SVM and RF achieved a moderate to high accuracy of energy consumption forecasting. The BO-TransGAN model achieved a highest R^2 of 0.995 and lowest MAE of 0.006. The BO-TransGAN model is more accurate in predicting energy consumption and it detects the level of carbon-related consumption awareness. Table 6 summarises the run-time and training performance of several energy forecasting models.

Table 3 Evaluation metrics for energy consumption prediction

Metric	*Definition*	*Equation*		*Variables*
Accuracy	Measures the proportion of correctly predicted energy consumption intervals out of all intervals.	$Accuracy = \frac{TP + TN}{TP + TN + FP + TN}$	(21)	(*TP*): true positives (high consumption correctly predicted), (*TN*): true negatives (low consumption correctly predicted), (*FP*): false positives, (*FN*): false negatives
Precision	Fraction of predicted high-consumption intervals that are actually high-consumption, assessing reliability of positive predictions.	$Precision = \frac{TP}{TP + FP}$	(22)	-
Recall	Fraction of actual high-consumption intervals correctly identified by the model, evaluating detection capability.	$Recall = \frac{TP}{TP + FN}$	(23)	-
F1-score	It indicates an overall classification balance of the energy consumption prediction framework.	$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$	(24)	-
Coefficient of determination (R^2)	Measures proportion of variance in actual energy consumption explained by the model, indicating predictive accuracy.	$R^2 = 1 - \frac{\sum_{l=1}^{k} \left(z_l - \widehat{z_l}\right)^2}{\sum_{l=1}^{k} \left(z_l - \overline{z_l}\right)^2}$	(25)	(z_l): actual energy consumption, $\left(\widehat{z_l}\right)$: predicted consumption, $\left(\overline{z_l}\right)$: average of real values
RMSE	Quantifies prediction error magnitude, penalising larger errors in energy forecasting.	$RMSE = \sqrt{\frac{1}{k} \sum_{l=1}^{k} \left(z_l - \widehat{z_l}\right)^2}$	(26)	(l): total number of intervals, (k) is the individual instances
Run time	Total computational time for generating energy predictions over the dataset, reflecting efficiency.	–		Measured in seconds (s)
Training time	Time required for BO-TransGAN to train on historical energy sequences, including generator-discriminator updates.	–		Measured in seconds (s)
Validation loss	Error metric evaluated on unseen validation energy sequences, guiding model generalisation.	-		-
Training loss	Error metric during training, reflecting model learning progression and convergence.	-		-

Figure 12 Precision and F1-measure performance of energy consumption prediction approaches (see online version for colours)

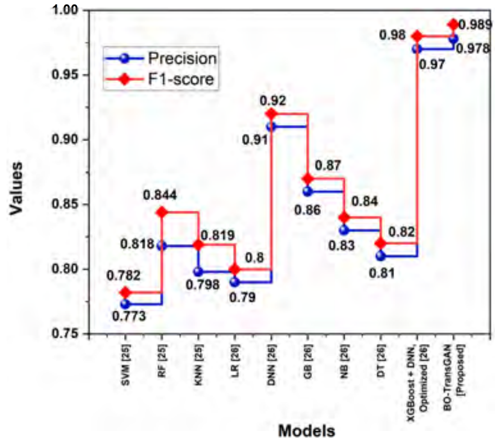


Table 4 Comparative performance analysis for energy consumption prediction models

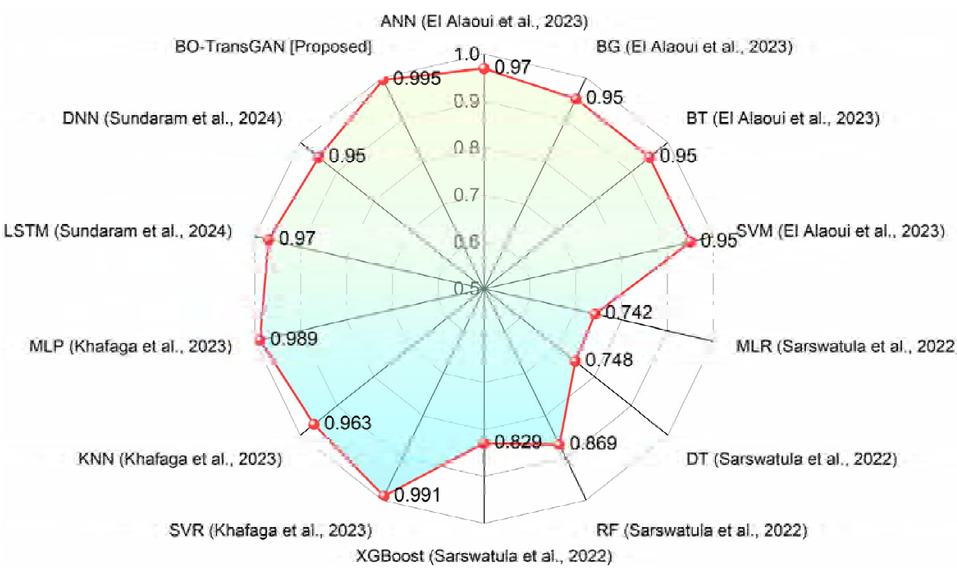
Models	*Accuracy*	*Precision*	*Recall*	*F1-score*
SVM (Sarswatula et al., 2022)	0.87	0.773	0.792	0.782
RF (Sarswatula et al., 2022)	0.883	0.818	0.884	0.844
KNN (Sarswatula et al., 2022)	0.854	0.798	0.854	0.819
LR (Ragupathi et al., 2024)	0.8	0.79	-	0.8
DNN (Ragupathi et al., 2024)	0.92	0.91	-	0.92
GB (Ragupathi et al., 2024)	0.87	0.86	-	0.87
NB (Ragupathi et al., 2024)	0.84	0.83	-	0.84
DT (Ragupathi et al., 2024)	0.82	0.81	-	0.82
XGBoost + DNN, optimised (Ragupathi et al., 2024)	0.98	0.97	-	0.98
LightGBM (Dinmohammadi et al., 2023)	0.828	-	-	-
CatBoost (Dinmohammadi et al., 2023)	0.748	-	-	-
XGBoost (Dinmohammadi et al., 2023)	0.8568	-	-	-
Stacking (Dinmohammadi et al., 2023)	0.904	-	-	-
BO-TransGAN [Proposed]	0.987	0.978	0.981	0.989

Traditional ML approaches such as LightGBM, CatBoost, and XGBoost had longer run-times, while DL was related to longer training times. The BO-TransGAN model achieved the lowest training time (2.12 s) and competitive run-time (7.35 s) in terms of efficiency to make fast and accurate energy consumption forecasts while taking into consideration limited resources. Table 7 and Figure 14 demonstrate the training and validation loss of each energy forecasting model, which represents the model's accuracy of prediction and generalisability.

Table 5 Forecasting error and model performance for accurate energy consumption prediction

Models	R^2	*MAE*
ANN (El Alaoui et al., 2023)	0.97	0.019
BG (El Alaoui et al., 2023)	0.95	0.026
BT (El Alaoui et al., 2023)	0.95	0.032
SVM (El Alaoui et al., 2023)	0.95	0.025
MLR (Sarswatula et al., 2022)	0.742	-
DT (Sarswatula et al., 2022)	0.748	-
RF (Sarswatula et al., 2022)	0.869	-
XGBoost (Sarswatula et al., 2022)	0.829	-
SVR (Khafaga et al., 2023)	0.991	0.048
KNN (Khafaga et al., 2023)	0.963	0.007
MLP (Khafaga et al., 2023)	0.989	0.008
LSTM (Sundaram et al., 2024)	0.97	0.90
DNN (Sundaram et al., 2024)	0.95	0.92
BO-TransGAN [Proposed]	0.995	0.006

Figure 13 R^2 comparison of energy forecasting models (see online version for colours)



The LSTM, Blocked LSTM, TCN, and TFT models achieved moderate loss; however, the BO-TransGAN model achieved a training loss of 0.00068 and a validation loss of 0.00074. This shows that BO-TransGAN exhibited the best learning of complex temporal dependencies and generalisation to previously unseen data. The results demonstrate the effectiveness reported in BO-TransGAN accuracy and reliability to make energy consumption forecasts.

Table 6 Training and run-time comparison of energy forecasting models

Models	*Run-time (s)*	*Training time*
LightGBM (Dinmohammadi et al., 2023)	14.28	-
CatBoost (Dinmohammadi et al., 2023)	7.69	-
XGBoost (Dinmohammadi et al., 2023)	62.56	-
LSTM (Sundaram et al., 2024)	-	2.69
DNN (Sundaram et al., 2024)	-	5.26
ANN (Sundaram et al., 2024)	-	3.88
BO-TransGAN [Proposed]	7.35	2.12

Table 7 Training and loss comparison of energy forecasting models

Models/measures	*Train loss*	*Validation loss*
LSTM (Nazir et al., 2023)	0.00101	0.000948
Blocked LSTM (Nazir et al., 2023)	0.000823	0.00130
TCN (Nazir et al., 2023)	0.00186	0.00149
TFT (Nazir et al., 2023)	0.00171	0.00169
BO-TransGAN [Proposed]	0.00068	0.00074

Figure 14 Comparative training-validation trends for carbon-aware energy forecasting models (see online version for colours)

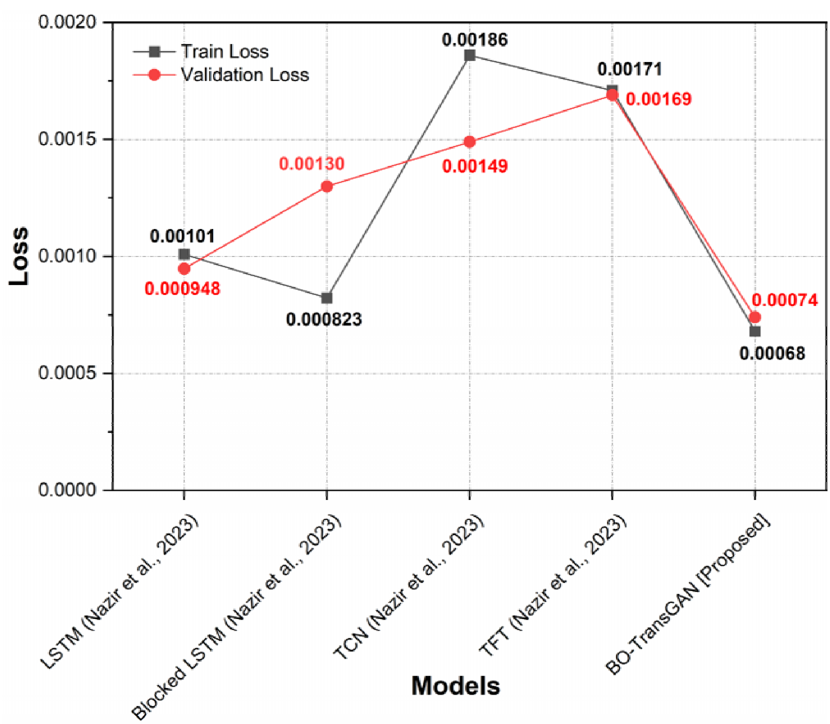


Table 8 Cross-run variability of the BO-TransGAN model

Metric	*Mean*	*Std. Dev.*	*95% confidence interval*
Accuracy	0.987	0.002	[0.985, 0.989]
Precision	0.978	0.003	[0.975, 0.981]
Recall	0.981	0.003	[0.978, 0.984]
F1-score	0.989	0.002	[0.987, 0.991]
R^2	0.995	0.001	[0.994, 0.996]
MAE	0.006	0.0005	[0.0055, 0.0065]
Training loss	0.00068	0.00002	[0.00066, 0.00070]
Validation loss	0.00074	0.00003	[0.00071, 0.00077]
Training time (s)	2.12	0.08	[2.04, 2.20]
Run time (s)	7.35	0.15	[7.20, 7.50]

To ensure that the reported performance metrics are not the product of a particularly fortunate initialisation, all metrics for BO-TransGAN were recomputed as the mean ± standard deviation over five independent runs with distinct random seeds, using the same 70% training/30% validation split. The training protocol, pre-processing pipeline, and hardware environment were identical across runs. Table 8 summarises the aggregated results.

All primary metrics exhibit very small standard deviations, confirming that the BO-TransGAN framework provides highly stable predictions and is insensitive to random initialisation.

4.3 Discussion

The previously adopted energy forecasting models, including ML approaches, such as RF, XGBoost, LightGBM, and CatBoost (Sarswatula et al., 2022; Ragupathi et al., 2024; Dinmohammadi et al., 2023), had relatively effective predictive performance; however, they had longer run-times, additional computational requirements, and limited ability to model complicated temporal dependencies. DL approaches, including LSTM, DNN, ANN, and TFT (Khafaga et al., 2023; Sundaram et al., 2024; Nazir et al., 2023), improved forecasting performance but were not real-time friendly due to the significant time to train, size of labelled data, and overall computational resources. Hybrid models achieved even greater generalisation performance but had to grapple with other disadvantages such as data ratios for synthetic to real data sample size dependency, overfitting with small sample training datasets, and modelling increasing levels of complexity while producing scaled models for high-performance forecasting. The suggested model addresses these issues by combining a TransGAN and BO to speed up convergence along with decreased time to train and time to inference with predictive performance on complex energy consumption sequences. The BO-TransGAN was effective at capturing temporal dependencies for energy forecasting, generating realistic synthetic data samples, and is an example of a scalable, low-carbon solution using advanced techniques for energy forecasting.

Despite the strong predictive performance demonstrated on the evaluated dataset, several important limitations warrant explicit discussion. To begin with, the dataset comprises only 2,000 hourly records, a sample size that is modest relative to the complexity of the Transformer based adversarial architecture; although the model mitigated overfitting through synthetic sequence generation, implicit data augmentation, and close alignment of training and validation losses, the generalisability of the findings to larger, more heterogeneous building stocks and industrial systems remains constrained without additional validation on substantially larger benchmarks. In addition, the computational complexity of the TransGAN framework, which employs self attention mechanisms and adversarial training, may become a practical bottleneck when scaling to datasets with tens of thousands of multivariate streams or longer sequences; while the reported training time of 2.12 s and run time of 7.35 s are low for the current 2,000 record setting, these values are likely to increase nonlinearly with input size, and the model's resource requirements have not been profiled on low power edge hardware. Beyond computational considerations, scalability to multi region, multi modal energy networks – where hundreds or thousands of individual building metres stream data concurrently – remains untested; the present implementation does not incorporate distributed training, parameter sharing, or hierarchical architectures that would be necessary to maintain convergence and efficiency in such settings. From a deployment perspective, real world deployment feasibility is limited by the absence of integration with live sensor data pipelines, a lack of built in uncertainty quantification, and the inability to adapt to concept drift without periodic retraining, all of which are essential for operational carbon management systems that must respond to dynamic consumption patterns and evolving emission factors. Accordingly, future work will therefore prioritise validation on larger public datasets, the development of lightweight model variants through knowledge distillation or model pruning, the exploration of federated learning architectures to enable privacy preserving multi building optimisation, and the incorporation of online learning mechanisms so that the model can continually update its forecasts in deployment environments.

5 Conclusions

An accurate and carbon-aware energy consumption forecasting framework was developed to assist in sustainable energy management. Energy consumption data with 2,000 records was collected through smart meters, building management systems, and industrial grids. The data was pre-processed by missing value imputation using KNNI, outlier removal by IQR, and min-max normalisation. The BO-TransGAN model was developed, which combines a TransGAN with BO to extract complex temporal dependencies to optimise the carbon management practices. The BO-TransGAN was compared with existing models, and the results showed high accuracy, with 0.987, precision of 0.978, recall of 0.981, F1-score of 0.989, R^2 of 0.995, MAE of 0.006, the lowest training loss of 0.00068, and a validation loss of 0.00074, with training and run-times of 2.12 s and 7.35 s. These results demonstrate that the BO-TransGAN generates realistic synthetic energy sequences, and also improves prediction accuracy and increases computational efficiency. The BO-TransGAN forecasting framework offers a practical, scalable solution for real-time energy consumption forecasting with actionable recommendations to inform energy efficiency and sustainable energy system trends to reduce carbon emissions and improve sustainable energy consumption.

Data availability statement

The authors declare that the data supporting the findings of this study are available within the article. The raw/derived data supporting the findings of this study are available from the corresponding author at request.

Declarations

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Al-Duais, F.S. and Al-Sharpi, R.S. (2023) 'A unique Markov chain Monte Carlo method for forecasting wind power utilizing time series model', *Alexandria Engineering Journal*, 1 July, Vol. 74, pp.51–63, https://doi.org/10.1016/j.aej.2023.05.019.

Alghamdi, A.A., Ibrahim, A., El-Kenawy, E.S. and Abdelhamid, A.A. (2023) 'Renewable energy forecasting based on stacking ensemble model and Al-Biruni earth radius optimization algorithm', *Energies*, 28 January, Vol. 16, No. 3, p.1370, https://doi.org/10.3390/en16031370.

Bacanin, N., Jovanovic, L., Zivkovic, M., Kandasamy, V., Antonijevic, M., Deveci, M. and Strumberger, I. (2023) 'Multivariate energy forecasting via metaheuristic tuned long-short term memory and gated recurrent unit neural networks', *Information Sciences*, 1 September, Vol. 642, p.119122, https://doi.org/10.1016/j.ins.2023.119122.

Balakrishnan, A., Sanisetty, B. and Bandaru, R.B. (2025) 'Stacked hybrid model for load forecasting: integrating transformers, ANN, and fuzzy logic', *Scientific Reports*, 4 June, Vol. 15, No. 1, pp.1–24, https://doi.org/10.1038/s41598-025-04210-1.

Cavus, M. and Allahham, A. (2025) 'Spatio-temporal attention-based deep learning for smart grid demand prediction', *Electronics*, 20 June, Vol. 14, No. 13, p.2514, https://doi.org/10.3390/electronics14132514.

Dinmohammadi, F., Han, Y. and Shafiee, M. (2023) 'Predicting energy consumption in residential buildings using advanced machine learning algorithms', *Energies*, 27 April, Vol. 16, No. 9, p.3748, https://doi.org/10.3390/en16093748.

El Alaoui, M., Chahidi, L.O., Rougui, M., Lamrani, A. and Mechaqrane, A. (2023) 'Prediction of energy consumption of an administrative building using machine learning and statistical methods', *Civ. Eng. J.*, May, Vol. 9, No. 5, pp.1007–1022, http://dx.doi.org/10.28991/CEJ-2023-09-05-01.

Habbak, H., Mahmoud, M., Metwally, K., Fouda, M.M. and Ibrahem, M.I. (2023) 'Load forecasting techniques and their applications in smart grids', *Energies*, 2 February, Vol. 16, No. 3, p.1480, https://doi.org/10.3390/en16031480.

Ji, X., Huang, H., Chen, D., Yin, K., Zuo, Y., Chen, Z. and Bai, R. (2023) 'A hybrid residential short-term load forecasting method using attention mechanism and deep learning', *Buildings*, January, Vol. 13, No. 1, p.72, https://doi.org/10.3390/buildings13010072.

Keerthana, K.B., Wu, S.W., Wu, M.E. and Kokulnathan, T. (2023) 'The United States energy consumption and carbon dioxide emissions: a comprehensive forecast using a regression model', *Sustainability*, 12 May, Vol. 15, No. 10, p.7932, https://doi.org/10.3390/su15107932.

Khafaga, D.S., El-Kenawy, E.S., Alhussan, A.A. and Eid, M.M. (2023) 'Forecasting energy consumption using a novel hybrid dipper throated optimization and stochastic fractal search algorithm', *Intelligent Automation & Soft Computing*, 1 January, Vol. 37, No. 2, pp.2117–2132, http://dx.doi.org/10.32604/iasc.2023.038811.

Maheen, R., Cai, L., Zhang, Y.S. and Zhao, M. (2023) 'Quantitative analysis of carbon dioxide emission reduction pathways: towards carbon neutrality in China's power sector', *Carbon Capture Science & Technology*, 1 June, Vol. 7, p.100112, https://doi.org/10.1016/j.ccst.2023.100112.

Moon, Y., Lee, Y., Hwang, Y. and Jeong, J. (2024) 'Long short-term memory autoencoder and extreme gradient boosting-based factory energy management framework for power consumption forecasting', *Energies*, 25 July, Vol. 17, No. 15, p.3666, https://doi.org/10.3390/en17153666.

Nassef, A.M., Olabi, A.G., Rezk, H. and Abdelkareem, M.A. (2023) 'Application of artificial intelligence to predict CO2 emissions: critical step towards sustainable environment', *Sustainability*, 6 May, Vol. 15, No. 9, p.7648, https://doi.org/10.3390/su15097648.

Nazir, A., Shaikh, A.K., Shah, A.S. and Khalil, A. (2023) 'Forecasting energy consumption demand of customers in smart grid using temporal fusion transformer (TFT)', *Results in Engineering*, 1 March, Vol. 17, p.100888, https://doi.org/10.1016/j.rineng.2023.100888.

Otshwe, J.N., Li, B., Chabrol, N.J., Qi, B. and Tshish, L.M. (2025) 'Graph convolutional-'optimization framework for carbon-conscious grid management', *Sustainability*, 3 September, Vol. 17, No. 17, p.7940, https://doi.org/10.3390/su17177940.

Pierre, A.A., Akim, S.A., Semenyo, A.K. and Babiga, B. (2023) 'Peak electrical energy consumption prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU approaches', *Energies*, 15 June, Vol. 16, No. 12, p.4739, https://doi.org/10.3390/en16124739.

Ragupathi, C., Dhanasekaran, S., Vijayalakshmi, N. and Salau, A.O. (2024) 'Prediction of electricity consumption using an innovative deep energy predictor model for enhanced accuracy and efficiency', *Energy Reports*, 1 December, Vol. 12, pp.5320–5337, https://doi.org/10.1016/j.egyr.2024.11.018.

Rani, S., Ahmed, U., Mahmood, A., Manzoor, S., Razzaq, S., Jamil, U., Zia, M.F. and Benbouzid, M. (2025) 'Dual attention-empowered bidirectional long short-term memory network for short-term electric load forecasting', *Energy Technology*, https://doi.org/10.1002/ente.202500914.

Rastgoo, R., Amjady, N. and Zareipour, H. (2024) 'A deep generative model for selecting representative periods in renewable energy-integrated power systems', *Applied Soft Computing*, 1 November, Vol. 165, p.112107, https://doi.org/10.1016/j.asoc.2024.112107.

Sarswatula, S.A., Pugh, T. and Prabhu, V. (2022) 'Modeling energy consumption using machine learning', *Frontiers in Manufacturing Technology*, 22 July, Vol. 2, p.855208, https://doi.org/10.3389/fmtec.2022.855208.

Sharma, K., Dwivedi, Y.K. and Metri, B. (2024) 'Incorporating causality in energy consumption forecasting using deep neural networks', *Annals of Operations Research*, August, Vol. 339, No. 1, pp.537–572, https://doi.org/10.1007/s10479-022-04857-3.

Son, N. and Shin, Y. (2023) 'Short-and medium-term electricity consumption forecasting using Prophet and GRU', *Sustainability*, 11 November, Vol. 15, No. 22, p.15860, https://doi.org/10.3390/su152215860.

Sun, H., Yu, Z. and Zhang, B. (2024) 'Research on short-term power load forecasting based on VMD and GRU', *PloS One*, 11 July, Vol. 19, No. 7, p.e0306566, https://doi.org/10.1371/journal.pone.0306566.

Sundaram, K., Preethaa, K.S., Natarajan, Y., Muthuramalingam, A. and Ali, A.A. (2024) 'Advancing building energy efficiency: a deep learning approach to early-stage prediction of residential electric consumption', *Energy Reports*, 1 December, Vol. 12, pp.1281–1292, https://doi.org/10.1016/j.egyr.2024.07.034.

Tang, P., Li, Z., Wang, X., Liu, X. and Mou, P. (2025) 'Time series data augmentation for energy consumption data based on improved TimeGAN', *Sensors*, 16 January, Vol. 25, No. 2, p.493, https://doi.org/10.3390/s25020493.

Tarmanini, C., Sarma, N., Gezegin, C. and Ozgonenel, O. (2023) 'Short term load forecasting based on ARIMA and ANN approaches', *Energy Reports*, 1 May, Vol. 9, pp.550–557, https://doi.org/10.1016/j.egyr.2023.01.060.

Uzun, E., Karabinaoğlu, M.S. and Ayaz, M. (2025) 'LSTM-based estimation of energy consumption in energy-intensive facilities', *Engineering Research Express*, 18 February, Vol. 7, No. 1, p.015420, https://doi.org/10.1088/2631-8695/adb378.

Wan, Q. and Liu, J. (2023) 'Energy efficiency optimization and carbon emission reduction targets of resource-based cities based on BiLSTM-CNN-GAN model', *Frontiers in Ecology and Evolution*, 31 August, Vol. 11, p.1248426, https://doi.org/10.3389/fevo.2023.1248426.

Zhong, L., Ye, T., Yang, Y., Pan, F., Feng, L., Qi, S. and Huang, Y. (2024) 'Deep reinforcement learning-based joint low-carbon optimization for user-side shared energy storage-distribution networks', *Processes*, 23 August, Vol. 12, No. 9, p.1791, https://doi.org/10.3390/pr12091791.