9-1-2023

A new approach to seasonal energy consumption forecasting using temporal convolutional networks

Abdul Khalique Shaikh
Sultan Qaboos University

Amril Nazir
Zayed University

Nadia Khalique
Sultan Qaboos University

Abdul Salam Shah
Universiti Kuala Lumpur

Naresh Adhikari
Slippery Rock University

Results in Engineering 19 (2023) 101296

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Research paper

A new approach to seasonal energy consumption forecasting using temporal convolutional networks

Abdul Khalique Shaikh [a,*], Amril Nazir [b], Nadia Khalique [a], Abdul Salam Shah [c], Naresh Adhikari [d]

[a] *Department of Information Systems, Sultan Qaboos University, Muscat, 123, Oman*
[b] *Department of Information Systems and Technology Management, Zayed University, Abu Dhabi, 144534, United Arab Emirates*
[c] *Department of Computer Engineering, University of Kuala Lumpur (UniKl-MIIT), Kuala Lumpur, 50250, Malaysia*
[d] *Department of Computer Science, Slippery Rock University, 1 Morrow Way, Slippery Rock, PA 16057, USA*

A R T I C L E I N F O

Keywords:
Energy forecasting
Seasonal energy
Smart grids
Temporal convolutional networks

A B S T R A C T

There has been a significant increase in the attention paid to resource management in smart grids, and several energy forecasting models have been published in the literature. It is well known that energy forecasting plays a crucial role in several applications in smart grids, including demand-side management, optimum dispatch, and load shedding. A significant challenge in smart grid models is managing forecasts efficiently while ensuring the slightest feasible prediction error. A type of artificial neural networks such as recurrent neural networks, are frequently used to forecast time series data. However, due to certain limitations like vanishing gradients and lack of memory retention of recurrent neural networks, sequential data should be modeled using convolutional networks. The reason is that they have strong capabilities to solve complex problems better than recurrent neural networks. In this research, a temporal convolutional network is proposed to handle seasonal short-term energy forecasting. The proposed temporal convolutional network computes outputs in parallel, reducing the computation time compared to the recurrent neural networks. Further performance comparison with the traditional long short-term memory in terms of MAD and sMAPE has proved that the proposed model has outperformed the recurrent neural network.

1. Introduction

Different smart city initiatives by government and private organizations have incorporated information and communication technologies (ICTs) to meet cities' growing challenges. International policies and scientific literature have widely embraced the smart homes, smart grids, and overall intelligent city concept. This concept makes cities smarter for citizens by utilizing many ICT innovations hitting us alarmingly. According to scientific evidence, smart cities are based on the following foundational theories: ICTs, urban planning, environmental considerations, living labs, and creative industries [16]. In addition, the associated concepts illustrate how ICT can assist in addressing almost every urban challenge. The latest ICT trends identified in the literature analysis are smart grids, IoT, big data, open data, and e-government [20,32,45]. The topic of consideration for this study is smart grids. The world's urban population makes up about half of the total population of the entire planet [59]. Cities are increasingly crowded, resulting in declining quality and quantity of services for their residents. The increased population has caused challenges for the energy sector to produce energy as per future demand [12]. The prediction of energy consumption is essential for effective demand-side management.

The models of energy consumption prediction mostly use statistical and machine learning models [1,46]. The previously recurrent neural network-based models have been applied for short-term energy prediction, while the seasonal factor has not been considered [34]. Traditional statistical models have performed better with the small amount of data, but the energy consumption data has drastically increased with the increase in the urban population [41]. The statistical models have certain limitations; hence the deep learning models have been widely adopted to handle the time series energy consumption data. Recurrent neural networks are essential algorithms for forecasting energy consumption [40]. The other most prominent algorithms are temporal convolutional neural networks (TCNs). The governments focus on providing effective

* Corresponding author.
E-mail address: shaikh@squ.edu.om (A.K. Shaikh).

https://doi.org/10.1016/j.rineng.2023.101296
Received 31 March 2023; Received in revised form 10 July 2023; Accepted 11 July 2023
Available online 17 July 2023

solutions to the people for a comfortable life. In terms of comfort, the heating, ventilation, and air conditioning (HVAC) system in any house plays significant importance, and people are ready to spend enough for a better standard of HVAC. The maintenance of HVAC standards requires enough amount of energy to be produced by the smart grids [29,47]. Energy production always remains a costly solution; hence researchers have focused on an alternative for energy production. Based on historical energy consumption data, the deep learning prediction models can avoid energy wastage and production of the exact amount in the grid environment [39].

The prediction methods remained typical for future energy consumption forecasting in the smart grids. Deep learning, specifically recurrent neural networks, successfully managed the extensive energy consumption data. The most common problem with recurrent neural networks is the vanishing gradients and lack of memory retention [9]. The temporal convolutional networks (TCN) have a strong capability to handle sequential data. This paper proposes an energy consumption prediction model using a TCN. In traditional models, energy consumption forecasts are based on the overall patterns, while seasonal patterns vary; hence, an approach for seasonal energy consumption forecasting is needed. The performance comparison with the LSTM has proved that the temporal convolutional network can handle the time series data better.

There are several benefits associated with this model.

- TCN solves the vanishing gradient and memory retention problem, reducing computation time for effective energy consumption management through parallel computation.
- Compared to a traditional RNN, TCNs train and evaluate long input sequences together rather than sequentially and maintain the customer's seasonal energy consumption patterns over time.
- Since TCN filters only share a single back-propagation path among layers, they are an excellent alternative to RNNs for sequences of arbitrary length.
- Different performance evaluation indicators are used to determine how well the proposed TCN procedure performs and some conventional methods, such as LSTM.

The remaining paper is organized as: section 2 provides related work, section 3 contains research design and methodology. The results are presented in section 4, the discussion in Section 5, and finally the conclusion is presented in section 6.

2. Related work

The traditional methods of energy consumption forecasting have used deep learning and recurrent neural networks, along with other machine learning algorithms. Deep learning significantly improved the prediction in time series data as suggested in a survey conducted by Torres et al. [43]. The focus includes the smart grid demand side management and energy consumption prediction at smart homes. The most recent model used NBEATS for the energy consumption forecasting of multiple customers in a smart grid environment [40]. The model has performed better, but it has not considered the seasonal factor of the energy consumption prediction. Deep learning remains significant in optimization-based studies focusing on energy optimization while improving the comfort index [14]. The neural network-based models have handled the energy prediction problem to some extent with higher accuracy but have failed to tackle the complex seasonal patterns of the data [13]. Unlike the traditional model, the deep generative short-term load forecasting model by Langevin et al. [21] used the appliances' energy consumption consumed in the past and the future consumption. The focus on appliances' energy consumption benefits more than historical forecasts.

Additionally, Wahid et al. [48] proposes a multi-layer perceptron and random forest technique for classifying buildings as high or low power consumers. The MLP remains computation efficient for smaller datasets; hence many authors prefer it over deep learning models while having smaller datasets. Oldewurtel et al. [35] used indoor weather prediction to use the thermal capacity of buildings efficiently using model predictive control (MPC). The weather prediction aimed to avoid using high-cost actuators and reduce waste. Zeng et al. [57] used a hybrid approach based on an extreme learning machine and switching delayed particle swarm optimization (SDPSO) algorithm for short-term load forecasting. The SDPSO has the enhanced capability of global searching to reach the optimal solution for optimizing hidden node parameters of the extreme learning machine. Li et al. [26] evaluated the impact of nonlinearity and response time on the accuracy of the system identification process of the energy forecasting model for building. The model tried to solve the scaling problems of buildings because the performance of the systems designed for small commercial buildings was not satisfactory with the larger buildings. By adding electric vehicles, the model develops a power-pollution dynamic load dispatch algorithm that solves a multi-objective optimization problem in an attempt to reduce both fuel costs and pollution emissions at the same time [51]. The model by Araghian et al. [4] optimizes economic and discomfort metrics through integrated energy management. Residents' dissatisfaction with indoor temperatures is measured using a novel metric called discomfort degree-day, which takes both magnitude and duration into account. The hybrid short-term energy consumption model by Sekhar and Dahiya [38] used different recurrent networks and the gray wolf optimization algorithm.

The optimization algorithm has been utilized to select the optimal parameters of Bi-LSTM and CNN. An effective time series feature extraction method is based on a one-dimensional CNN. Based on a neural network with a learning process control algorithm, a method is developed for predicting wind speed in an isolated power system. An hourly retrospective meteorological data set of wind speed observations is used to make predictions for four seasons throughout the year [28]. Modeling of renewable energy sources, energy storage devices, electric vehicles, and distributed generation systems was conducted here to optimize the management of a VPP. As well as the forecasting of wind speed, electricity price, load demand, and the behavior of electric vehicles is examined by using a method based on two-way long short-term memory networks [2]. An improved method is proposed in this paper, which uses a two-stage process which can be used to forecast grid load at the system-wide level. The first benefit of this method is that it reduces the original net load by removing the low-frequency components whose energy is quite insignificant [58].

This paper proposes an approach for selecting the optimal set of features for short-term load forecasting (STLF) problems using hybrid feature selection (HFS). The HFS uses a hybrid genetic algorithm (EGA) combined with a random forest method as a component of its online feature selection algorithm, combining the elitist genetic algorithm and random forest methods [42]. The smart grid energy optimization methods using deep learning are presented in this paper using the new type of wild horse optimization algorithm [31]. The results of this study demonstrate that there should be a consideration of the factors that affect the generation of renewable energy, along with a qualitative model for how the generation of renewable energy will affect electricity prices. This model is based on random forests and improved mahalanobis distance [50].

The study addresses the challenge of imprecise energy efficiency services and insufficient household response to energy use through a dynamic load-priority scheduling strategy [54]. Model predictive control-based reinforcement learning and Shapley values are used in this paper to propose an energy management strategy for residential micro-grid systems [7]. The model proposed by Velimirović et al. [44] is an evident example of the entropy-based fuzzy models for the problem of short-term energy forecasting. To estimate the energy production, two different approaches were used during the analysis: the first user data and artificial intelligence techniques to estimate energy production, specifically multilayer perceptrons in conjunction with radial basis functions; and the second used models to estimate energy production [11]. This paper presents a novel encoding method based on neural ordinary differential equations that can be viewed as continuous residual networks (ResNets) for learning time series dynamics of electricity load [18].

This study proposes a combined framework for solving the problem of low prediction accuracy in electricity load forecasting with a modified noise processing strategy, a multi-objective optimization algorithm, and deep neural networks to resolve the problem of low prediction accuracy [52]. Several efficient algorithms have been applied to DSM methodology for a residential community that aims to reduce peak energy consumption, including binary orientation search algorithms (BOSA), cockroach swarm optimization (CSO), and sparrow search algorithms (SSA). A smart grid DSM optimized on algorithms will reduce electricity consumption costs while increasing smart grid efficiency. Using a load-shifting technique, the proposed DSM methodology is able to achieve the above objective. Using renewable energy sources on-site, the proposed system will reduce electricity costs and reduce peaking at power plants [19]. Alrwashdeh [3] conducted study considering local climate, average monthly energy consumption, and electricity rates in Jordan to determine the most profitable way to install a photovoltaic (PV) system on a residential building.

This review addresses several dimensions, but majority of the studies have considered the normal energy consumption forecasting. The divisions base on short, medium and long term predictions. Energy storage has always been viewed as critical in reducing energy costs. As a part of the studies, green energy was also considered from a pollution reduction perspective. In most prediction problems, deep learning and reinforcement learning techniques have been used to build models. The implementation of a smart grid requires developing a model to predict energy according to the season.

3. Research design and methodology

Most of the models are based on synthetic data, which means they do not accurately reflect the real smart grid environment and often don't perform well when implemented. The proposed study is based on different experiments using the energy consumption dataset. The model has three main layers, i.e., pre-processing, prediction, and performance evaluation. The experimental design is based on seasonal energy consumption, where the model predicts the four seasons' energy consumption based on historical data. The pre-processing steps are helpful for the removal of noise and irregularities in the data while the main responsibility relies on the prediction module, where the temporal convolutional network predicts energy consumption.

3.1. Database acquisition, pre-processing and description

The dataset has been acquired having energy consumption data for the customers of London households. The same data has been used in our previous energy forecasting model [34,40]. Over the course of November 2011 to February 2014, data was collected from different household customers. The objective of the proposed method is seasonal energy consumption forecasting; hence after analysis, the unnecessary data has been eliminated [10]. The proposed study considers 150 customers, with only seasonal data of 12 months for each customer and a dynamic time-of-use pricing strategy (dToU). The experiment has been carried out with daily consumption data in kWh units. The training data has 75% while the rest 25% used to evaluate the model in terms of the prediction. By using the moving average, the model has been able to eliminate the outliers that cause lower accuracy. Using the Python package NumPy, outliers have been identified as observations lower or higher Q1 + 1.5 IQR [34,40]. We used (1) and (2) to normalize and denormalize the data into a 0–1 scale.

$$N(a) = \frac{x(a) - \min(a)}{\max(a) - \min(a)} \tag{1}$$

Table 1
Parameter setting of LSTM and TCN.

Parameter	LSTM	Convolutional Network (Temporal)
Activation function	ReLU	ReLU
Hidden Dim/Size/Widths	10	
Number of Layers	1	6
Random State	42	0
Training length/Input chunk length	30	30
Output chunk length	15	15
No of epochs val period	1	
Batch size	32	512
Learning rate	1–3	1–3
Epochs	150–250	200–250
Dilation base		2
Kernel Size		5
Number of Filters		3
Dropout rate	0.1	0.1
Future/Past covariates	Day Series	Day Series

$$D(a) = N(a) * (\max(a) - \min(a)) + \min(a) \tag{2}$$

Where the data after normalization has been denoted by $N(a)$, once the training is completed and the data is denormalized and denoted as $D(a)$. The data under consideration for the normalization is represented as $x(a)$. These conversions use the minimum and maximum values of the dataset denoted by $\min(a)$ and $\max(a)$ respectively.

3.2. Prediction module

The prediction module contains a temporal convolutional network that takes the pre-processed input data of energy consumption and trains the network. The TCN provides various advantages over traditional convolutional neural networks (CNN). In computer vision, convolutional neural networks are commonly used to recognize and classify objects in images or videos. Based on studies of the visual cortex, convolutional neural networks were first proposed in 1979. Recently, it has been used for prediction tasks in various areas, including weather and energy consumption prediction. Convolutional neural networks (CNNs) can handle sequential data with temporal aspects and large receptive fields; however, temporal convolutional networks depend on casual convolutions and dilations. As a result of its inherent limitations, CNNs are unsuitable for time-series prediction due to constraints such as a fixed-size input vector and inconsistent input and output sizes [15,24]. At the same time, the TCN resolves the limitations of CNN models and remains suitable for the proposed energy consumption forecasting model. As a crucial factor in defining neural networks' architectural configuration, the number of hidden layers becomes critical to the final result. The number of layers in deep learning networks has been widely acknowledged as an important factor in capturing intricate features and achieving relatively high accuracy levels. To enhance performance, additional layers are commonly added to neural networks to increase their size [25]. In the proposed model we have used random search as it can be argued that random search is a fundamental improvement over grid search. The hyper-parameters are explored randomized to determine potential parameter values by sampling from specific distributions [30]. As soon as the desired level of accuracy is reached, the search process continues. Random search has consistently produced better results than grid search, despite being similar to it, although it is similar to grid search. The parameters selected for TCN and LSTM algorithms can be seen in Table 1.

PyTorch forecasting is used to implement TCN in the model. Input chunks are used as past covariates, and output chunks are used as future covariates [27]. Future covariates are used as mandatory inputs for multi-head attention queries. Furthermore, encoders were used to automatically generate the day covariates (future covariates) in addition to the past covariates.

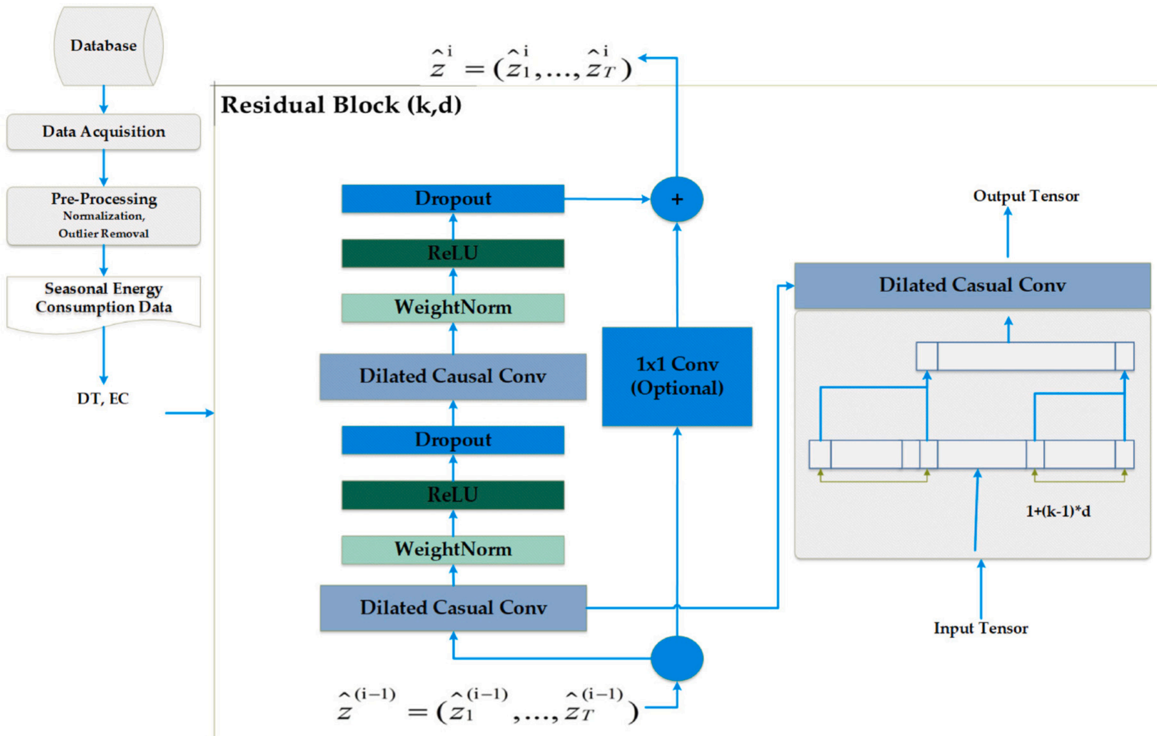

Fig. 1. Proposed TCN model.

3.2.1. *Temporal convolutional network (TCN)*

Like traditional deep learning networks, the convolutional neural networks also contain three prominent layers, i.e., input, hidden, such as the convolutional layer, and an output layer. The input and output layer size depends on the number of input variables and desired output variable. The details of these layers have already been defined in the Table 3. In the proposed TCN model, the hidden layers are responsible for the convolutions. The concept of convolution is commonly used in mathematics, where it refers to the conversion of two functions into a new function that expresses how one function's shape is modified by another [37]. To calculate the integral of the product of two functions, one of which is reversed and shifted, it is necessary to take two functions and calculate the integral of their products. Using the convolution function, we can express how one function changes the shape of another. The convolution is similar to cross-correlation involving two similar series or sequences [36]. It is the hidden layers that perform computations, while it is the pooling layers that consolidate and collect the results from the hidden layers. Constructing a fully connected layer maps the input and output values using a nonlinear function. In addition to RNNs, TCNs have cost or loss functions that are minimized to reduce errors in prediction [17]. In the TCN, the input node's receptive field is also known as a patch in computer vision terminology. Inputs are received from a limited area of the previous layer of the network, the node's receptive field. A feature detection algorithm primarily works by applying filters to the input to make detecting the features easier for the TCN. These filters are made up of neurons or nodes [55]. Similarly, some filters connect to other features. It is essential to isolate the trend or seasonality of a time series. During training, the TCN calibrates the filters' values based on the loss function to minimize prediction error. Using the filters learned, the TCN predicts future energy consumption. Trend and seasonality patterns can be recognized by it when dealing with time series [8]. Feature maps are generated by the nodes, indicating where features are located. A filter scans an input patch to check if the feature is present. A feature map is created based on the results. A kernel is an array of weights arranged in two dimensions. As the kernel moves through a receptive field, it detects the features [33]. The stride of the filter determines how far it moves across the input. An activation function is applied to the extracted features by fully connected layers. Nodes in this layer receive information from nodes in previous layers, so its name arises from connecting them all: in a fully connected layer, a node receives information from all the nodes in previous layers [56]. Through the fully connected layer, features of earlier layers are tied together in a typically nonlinear manner. TCNs learn to sort out unimportant features from essential ones with successive training epochs – feed-forward followed by back-propagation. The TCN ensures causal convolution. Values at the beginning of the input sequence must determine an output value. A TCN dilation expands a node's receptive field to encompass more periods of history [23,53]. A typical residual block of TCN can be seen in the following Fig. 1. The computation and mathematical expressions of the layers and residual block has been adopted from Bai et al. [5], Lässig [22].

If we have one-dimensional sequence input $X \in R^n$ having a filter $f : \{0, \ldots, k-1\} \to R$, we can define the F_s dilated convolution operation as in (3).

$$F(s) = (X * d f)(s) = \Sigma_{i=0}^{k-1} f(i) \cdot X s - d \cdot i \tag{3}$$

The dilation factor can be defined by d, the filter size is represented by d, and the direction of the past is denoted by $s - d \cdot i$ as defined in (4).

$$d_i = 2^i, 1 \leq i \leq n \tag{4}$$

Generalizing, a 1D convolutional network with n layers and k kernels has a receptive field r as in (5).

$$r = 1 + n * (k-1) \tag{5}$$

We can calculate the number of layers n by setting the receptive field size to input length l as in (6).

$$n = \left\lceil \frac{(l-1)}{k-1} \right\rceil \tag{6}$$

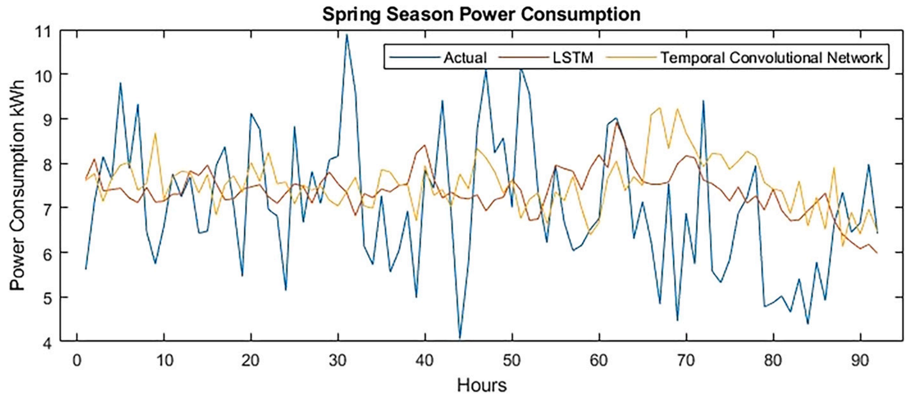

Fig. 2. Comparison of spring season energy consumption.

By computing d as a function of i, we can compute the dilation of a particular layer based on its dilation base integer b as in (7).

$$d = b ** i \tag{7}$$

Where in i represents the number of layers. As a result of these observations, we can calculate how many layers our network needs to provide full historical coverage. If we take a kernel with size k, a dilation base b, where k for each base equals b, and an input length l, then we can state the following inequality using (8).

$$1 + (k-1).\frac{b^n - 1}{b-1} \geq l \tag{8}$$

The minimum number of layers required can be calculated by solving for n using (9).

$$n = \left\lceil log_b \left(\frac{(l-1).(b-1)}{(k-1)} + 1 \right) \right\rceil \tag{9}$$

We can compute the number of zero-padding entries p needed for our current layer given the dilation base b, kernel size k, and the number of layers below it using (10).

$$p = b^i.(k-1) \tag{10}$$

Based on dilation base b and kernel size kxb, we can compute the total size r of the receptive field r for a TCN using (11).

$$r = 1 + \sum_{i=0}^{n-1} 2.(k-1).b^i = 1 + 2.(k-1).\frac{b^{n-1}}{b-1} \tag{11}$$

Thus, there will be a minimum number of residual blocks n for a full history coverage of input length l as in (12).

$$n = \left\lceil log_b \left(\frac{(l-1).(b-1)}{(k-1).2} + 1 \right) \right\rceil \tag{12}$$

The th layer of the network has been denoted by i total dilated convolutional layers are represented by the n as in (13).

$$k_{i+1} = d_i * (k_i - 1) + 1, 1 \leq i \leq n, k1 \epsilon N^* \tag{13}$$

The th layer of the network has been denoted by i total dilated convolutional layers are represented by the n. Data situations usually determine the initial size of the filter, the default is $k_i = 2$. However, it can be adjusted to other values if necessary.

3.3. Performance evaluation metrics

For time series data, the smaller the error value, the better the accuracy with the mean absolute deviation (MAD) [6,49]. Using the MAD formula, a prediction is made, and the actual value is calculated at least one period ahead. In the following equation, you can see the MAD's mathematical expression (14). The other performance metric is the sMAPE, as expressed in (15).

$$MAD = \frac{1}{N} \sum_{i=1}^{n} |A_i - m(X)_i| \tag{14}$$

$$sMAPE = \frac{100\%}{N} \sum_{i=1}^{n} \frac{\frac{|P_i - A_i|}{|A_i| + |P_i|}}{2} \tag{15}$$

Where $m(X)$ denote average value of the dataset, xi is the data values in the set. The N stands for total observations, A for actual values, and P for predicted values.

4. Results

4.1. Seasonal power consumption

The proposed model aims to forecast the smart grids' seasonal energy consumption to produce energy consumption per the customers' seasonal demand. The dataset for evaluating the proposed model belongs to the customers of England, where typical weather has four seasons, i.e., spring, summer, autumn, and winter. The spring season starts from March to May. If we observe the graph of the spring season, the highest energy consumption of randomly selected customers is 11 kWh. Hence, the LSTM and proposed TCN model has performed better, although the actual energy consumption has higher fluctuations, and the model has been trained on energy consumption data of multiple customers having different energy consumption behaviors; hence there is the provision of the improvement in the graphs. Overall, the LSTM graph shows that it has struggled with the frequent fluctuations of energy consumption, while on the other hand, the TCN has a better energy consumption graph.

In the Fig. 2 during the spring, the temperature varies between 1 to 30 centigrade; hence both the heating and cooling systems operate. While from March to May, the highest temperature keeps increasing. The increased temperature usually causes higher energy consumption as most residents start operating cooling systems which is evident from the above graph. The summer season starts from June to August and is considered the time of outings and holidays by most residents. While the weather mostly remains hot, the summer's energy consumption is higher than in spring, as seen in the graph. The highest temperature during August month even reaches 38 centigrade. The highest energy consumption during this period is due to the cooling systems operating in the houses.

The Fig. 3 for the summer show that during June, the customer consumed higher energy, so the LSTM and TCN have predicted higher, wherein both algorithms have shown a similar pattern, which means the fluctuations are challenging for the algorithms. At the same time, the highest energy consumption can be noticed during July. The autumn season has normal weather where from September to November,

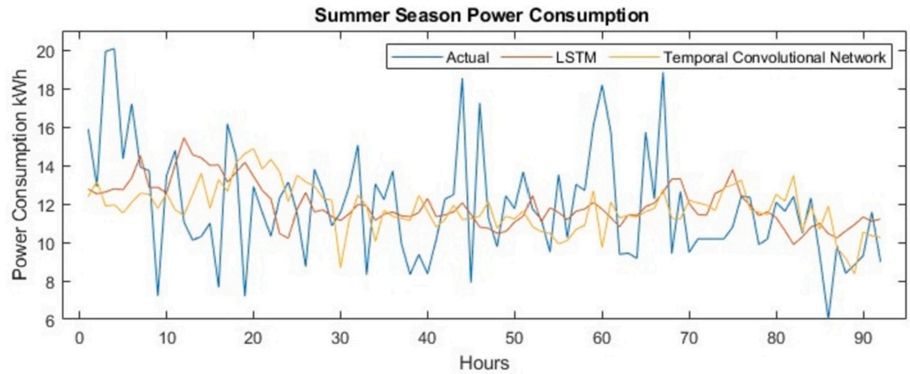

Fig. 3. Comparison of summer season energy consumption.

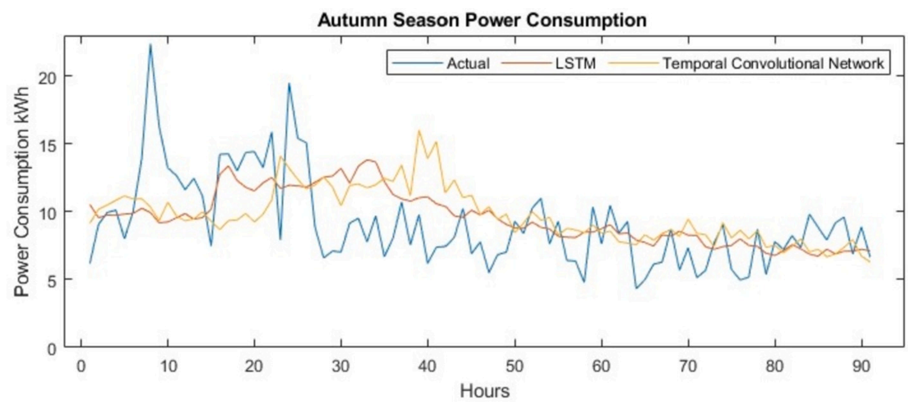

Fig. 4. Comparison of autumn season energy consumption.

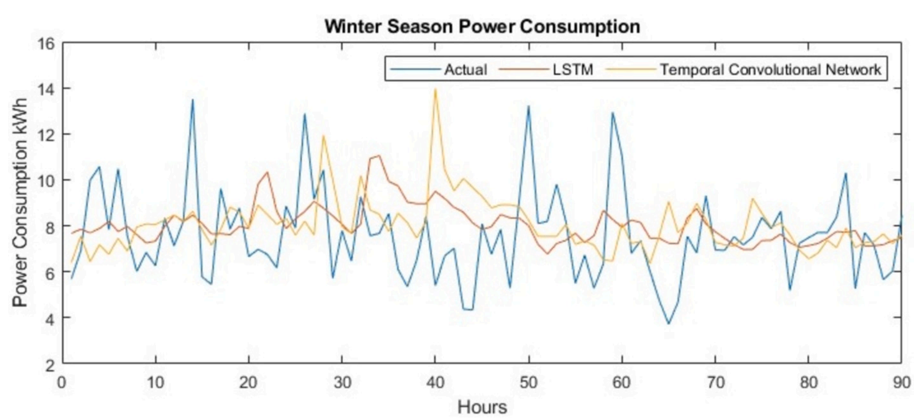

Fig. 5. Comparison of winter season energy consumption.

the temperature gradually reduces; hence Fig. 4 shows the same trend as during September the energy consumption is high while with the passage of days, it reduces to the range of 5 to 10 kWh. The reason is typical weather and a reduction in the usage of cooling systems. Here the LSTM and TCN algorithms have shown better performance with linear graphs of the same pattern without too many fluctuations. Overall, the TCN has performed better than LSTM.

The winter lasts from December to February, and most weather remains cold, with snow and frost everywhere. The Fig. 5 shows a constant pattern of energy consumption during these three months because most of the users operate heating systems during the winter, and the systems operate for the day and night; hence the same amount of energy is consumed daily. The graph of predicted energy consumption by LSTM and TCN shows that both algorithms have performed better with the winter season prediction than the other three seasons. The reason is a similar pattern during the three months. So it is evident that deep learning might require more data for the same customers to perform with better accuracy.

5. Discussion

A comparative analysis has been carried out with the LSTM model compared to the proposed TCN model based on customers' energy consumption data in England. Combining these data has enabled the proposed TCN model to be tested and evaluated. There is no doubt that both deep learning algorithms can handle energy consumption forecasting problems in a smart grid environment when looking at the MAD and sMAPE errors of both algorithms. The performance of both algorithms seems to differ slightly from one another. Due to its strong capabilities and parallel computations, the TCN model has performed better in MAD

Table 2
Comparison of seasonal MAD and sMAPE error.

Model	Season							
	Spring		Summer		Autumn		Winter	
	MAD	sMAPE	MAD	sMAPE	MAD	sMAPE	MAD	sMAPE
LSTM	1.346	0.189	2.307	0.192	2.83	0.292	1.702	0.219
TCN	1.291	0.184	2.295	0.190	2.43	0.256	1.69	0.215

and sMAPE. There is also a significant decrease in training time for the TCN model compared to the LSTM model. An evaluation of the TCN based on modified parameters has been conducted to evaluate its effectiveness. Modifications have been made to epochs, learning rates, batch sizes, and random states. A close examination of the results revealed that the batch size dramatically impacts the computation time, and the accuracy has been reduced with larger batches. The moderate approach was adopted for the experimentation to avoid the extra computation time, even though the slower learning rate improved it significantly. There was an improvement in accuracy due to the increase in epoch, but after a certain level of advancement, there was no further improvement; as a result, the maximum size was selected once there was no more improvement. A significant problem with the data was that it frequently fluctuated, ultimately creating challenges for the deep learning algorithms as they attempted to analyze it. Although normalization was applied, it was limited to a certain extent to keep the original pattern of energy consumption since the model was designed for a smart grid environment where each user consumes energy differently. Since the model uses batch sizes, it has certain limitations; therefore, if new data is added to it in future work, it has to be retrained. As new data is added to a model in real time, the weights and biases will be updated, and the model will be retrained to predict data in line with the customer's requirements.

5.1. Comparison of MAD and sMAPE errors

The performance comparison of the proposed seasonal TCN model with the LSTM has been carried out, where the data has been distributed according to the four seasons. The MAD by LSTM during spring month is 1.346 and sMAPE 0.189, while the TCN model has performed better with slightly less error. The reason is that the TCN model can observe the trends in large datasets. There is a very slight improvement in the sMAPE error. The LSTM and TCN both have strong capability to tackle the prediction tasks. The summer season has slightly higher MAD and sMAPE errors by the LSTM compared with the spring season. The reason is that there are fluctuations in the temperature during summer; hence the energy consumption behavior of the customers changes accordingly. Most users operate cooling systems during the day while night has standard energy consumption patterns; therefore, the fluctuations in the energy consumption have caused challenges for the LSTM and TCN models. Autumn has a similar range of error differences, although it has increased compared to the summer due to the same reason of differentiation in the energy consumption behavior by the residents of homes. Tables 2 and 3 show the TCN's smallest error scale compared to the four comparative methods.

The winter season has again normalized the error difference because the energy consumption pattern remains the same throughout the day while it may reduce further at night. So the MAD and sMAPE errors have reduced compared to the summer and autumn seasons. The TCN model has performed better with the data of all four seasons; hence the model can handle the seasonal data along with the normal overall or customer-based energy consumption in the smart grid environment.

5.2. Comparison of energy consumption

Table 3 provides an overview of the performance of the proposed TCN model in terms of the actual energy consumption vs. forecast by the LSTM and TCN. It can be seen that the customer has consumed higher energy during the summer due to hot weather and the operation of the cooling system inside the homes. The actual energy consumption during summer was 1095.00, while the TCN has predicted it as 1099.03. Hence it can be concluded that for seasonal energy consumption, the LSTM and TCN models are reliable to be implemented in the smart grid environment for future demand prediction. The advantage of the TCN model over LSTM is that it can handle larger datasets, and due to parallel computation, it can provide higher accuracy. The autumn season also has higher consumption; hence, the TCN and LSTM have slightly predicted higher consumption than the original. The autumn season also has hot weather; therefore, people still operate the cooling systems, consuming considerable energy. The spring and winter seasons have a similar range of energy consumption due to cold weather compared to the summer and autumn; hence mostly during the winter season, people operate heating systems to maintain the inside weather conditions under an acceptable range.

Table 3
Comparison of seasonal energy consumption.

Model	Energy Consumption in kWh			
	Spring	Summer	Autumn	Winter
Actual	647.89	1095.00	833.17	680.41
LSTM	694.03	1082.78	886.04	724.41
TCN	679.59	1099.03	869.92	723.77

The performance analysis shows that the TCN and LSTM have performed better with the summer data due to a similar pattern of operating electrical equipment. At the same time, there is enough provision to improve the results for the spring, autumn and winter. Still, if we consider the three months duration for each season, the difference is acceptable and smart grids can avoid enough energy overproduction using the proposed TCN and LSTM models.

6. Conclusion

Various energy consumption models have been proposed in the literature focused on short, medium and long term energy consumption prediction. Due to limitations of the data, most models do not consider the seasonal factor for energy predictions. In traditional recurrent neural network based models, energy consumption has been predicted with certain accuracy, but the complexity of the data remains a challenge. It is difficult for the algorithms to properly train due to the noise and the different patterns of energy consumption among the customers. Using temporal convolutional networks, we have developed a seasonal energy forecasting model that avoids these issues. This model is new in that it uses a temporal convolution network as well as a seasonal component, as traditional models only consider short, medium, and long-term forecasts of energy consumption. For smart grid environments, seasonal prediction can be used instead of predicting energy consumption annually to produce energy based on the seasons. A comparison of TCN's performance with that of LSTM has shown that TCN generally performs better than LSTM. With smoothing of the data, the model's performance can further be improved, but the actual pattern of energy consumption will be affected. Our future work will examine how smoothing affects energy consumption data by comparing actual with smoothed data.

Real-time forecasts can be performed using TCN models due to their ability to process data in parallel. The energy sector is particularly relevant, where real-time prediction ensures stable power supply, optimization of electricity grids, and management of electricity grids. Grid stability, load balancing, and demand response can be improved using TCN models. Solar and wind generation from intermittent sources such as solar and wind can be integrated with the power grid by using TCN models. Grid operators can take proactive measures to maintain grid stability by forecasting renewable energy production accurately and balancing supply and demand. It can be concluded that the primary practical implications of TCN energy forecasting models lie in their ability to deliver accurate, real-time, and scalable predictions regarding the consumption and generation of energy as well as the dynamics of the energy market. To put forward a more sustainable and efficient energy future, these models empower actors and stakeholders involved in the energy industry to make informed choices, optimize energy systems, and reduce carbon emissions.

CRediT authorship contribution statement

The authors hereby confirm that we all have made a substantial contribution to complete this article. **Amril Nazir:** Idea generation, methodology and, engaged in writing paper. **Abdul Khalique Shaikh:** Methodology, writing, reviewing and editing. **Nadia Khalique:** proofreading and structuring. **Abdul Salam Shah:** Experiment and results. **Naresh Adhikari:** Guidance and supervision. We all read and approved the final manuscript.

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.

Data availability

Data will be made available on request.

References

[1] S. Abba, B.G. Najashi, A. Rotimi, B. Musa, N. Yimen, S. Kawu, S. Lawan, M. Dagbasi, Emerging Harris Hawks Optimization based load demand forecasting and optimal sizing of stand-alone hybrid renewable energy systems – a case study of Kano and Abuja, Nigeria, Results Eng. 12 (2021) 100260, https://doi.org/10.1016/j.rineng.2021.100260.
[2] A. Ahmadian, K. Ponnambalam, A. Almansoori, A. Elkamel, Optimal management of a virtual power plant consisting of renewable energy resources and electric vehicles using mixed-integer linear programming and deep learning, Energies 16 (2023) 1000.
[3] S.S. Alrwashdeh, Energy profit evaluation of a photovoltaic system from a selected building in Jordan, Results Eng. 18 (2023) 101177, https://doi.org/10.1016/j.rineng.2023.101177.
[4] M.H. Araghian, M. Rahimiyan, M. Zamen, Robust integrated energy management of a smart home considering discomfort degree-day, IEEE Trans. Ind. Inform. (2023).
[5] S. Bai, J.Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint, arXiv:1803.01271, 2018.
[6] I. Barić, R. Grbić, E.K. Nyarko, Short-term forecasting of electricity consumption using artificial neural networks – an overview, in: 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO, IEEE, 2019, pp. 1076–1081.
[7] W. Cai, A.B. Kordabad, S. Gros, Energy management in residential microgrid using model predictive control-based reinforcement learning and Shapley value, Eng. Appl. Artif. Intell. 119 (2023) 105793.
[8] L. Chen, P. Thakuriah, K. Ampountolas, Short-term prediction of demand for ride-hailing services: a deep learning approach, J. Big Data Anal. Transp. 3 (2021) 175–195.
[9] M.K. Chowdary, J. Anitha, D.J. Hemanth, Emotion recognition from EEG signals using recurrent neural networks, Electronics 11 (2022) 2387.
[10] L. DataStore, SmartMeter energy consumption data in London households, 2018.
[11] O.E. Dragomir, F. Dragomir, Application of scheduling techniques for load-shifting in smart homes with renewable-energy-sources integration, Buildings 13 (2023) 134.
[12] M. Fayaz, D. Kim, A prediction methodology of energy consumption based on deep extreme learning machine and comparative analysis in residential buildings, Electronics 7 (2018) 222.
[13] M. Fayaz, H. Shah, A.M. Aseere, W.K. Mashwani, A.S. Shah, A framework for prediction of household energy consumption using feed forward back propagation neural network, Technologies 7 (2019) 30.
[14] M. Fayaz, I. Ullah, A.S. Shah, D. Kim, An efficient energy consumption and user comfort maximization methodology based on learning to optimization and learning to control algorithms, J. Intell. Fuzzy Syst. 37 (2019) 6683–6706.
[15] R. Gong, X. Li, A short-term load forecasting model based on crisscross grey wolf optimizer and dual-stage attention mechanism, Energies 16 (2023) 2878.
[16] Q. Hassan, A.Z. Sameen, H.M. Salman, M. Jaszczur, M. Al-Hitmi, M. Alghoul, Energy futures and green hydrogen production: is Saudi Arabia trend?, Results Eng. 18 (2023) 101165, https://doi.org/10.1016/j.rineng.2023.101165.
[17] S. He, C. Li, X. Liu, X. Chen, M. Shahidehpour, T. Chen, B. Zhou, Q. Wu, A per-unit curve rotated decoupling method for CNN-TCN based day-ahead load forecasting, IET Gener. Transm. Distrib. 15 (2021) 2773–2786.
[18] S. Huang, J. Shen, Q. Lv, Q. Zhou, B. Yong, A novel node approach combined with LSTM for short-term electricity load forecasting, Future Internet 15 (2022) 22.
[19] A.M. Jasim, B.H. Jasim, B.C. Neagu, B.N. Alhasnawi, Efficient optimization algorithm-based demand-side management program for smart grid residential load, Axioms 12 (2022) 33.
[20] N. Khan, Z. Shahid, M.M. Alam, A.A. Bakar Sajak, M. Mazliham, T.A. Khan, A. Rizvi, S. Safdar, et al., Energy management systems using smart grids: an exhaustive parametric comprehensive analysis of existing trends, significance, opportunities, and challenges, Int. Trans. Electr. Energy Syst. 2022 (2022).
[21] A. Langevin, M. Cheriet, G. Gagnon, Efficient deep generative model for short-term household load forecasting using non-intrusive load monitoring, Sustain. Energy Grids Netw. 34 (2023) 101006.
[22] F. Lässig, Temporal convolutional networks and forecasting – Unit8, 2020.
[23] C. Lea, M.D. Flynn, R. Vidal, A. Reiter, G.D. Hager, Temporal convolutional networks for action segmentation and detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 156–165.
[24] G. Li, V.L. Knoop, H. Van Lint, Multistep traffic forecasting by dynamic graph convolution: interpretations of real-time spatial correlations, Transp. Res., Part C, Emerg. Technol. 128 (2021) 103185.
[25] L. Li, A. Talwalkar, Random search and reproducibility for neural architecture search, in: Uncertainty in Artificial Intelligence, PMLR, 2020, pp. 367–377.
[26] X. Li, J. Wen, E.W. Bai, Developing a whole building cooling energy forecasting model for on-line operation optimization using proactive system identification, Appl. Energy 164 (2016) 69–88.
[27] B. Lim, S.Ö. Arık, N. Loeff, T. Pfister, Temporal fusion transformers for interpretable multi-horizon time series forecasting, Int. J. Forecast. 37 (2021) 1748–1764.
[28] V. Manusov, P. Matrenin, M. Nazarov, S. Beryozkina, M. Safaraliev, I. Zicmane, A. Ghulomzoda, Short-term prediction of the wind speed based on a learning process control algorithm in isolated power systems, Sustainability 15 (2023) 1730.
[29] M. Marwan, J. Jamal, A. Hamid, N. Nasir, N.A. La Nafie, A. Gunawan, S. Syamsuddin, B.A. Razak, M. Musaruddin, Optimize electrical energy cost of air conditioning considering to different wall characteristics, Results Eng. 100990 (2023).
[30] A. Meola, M. Winkler, S. Weinrich, Metaheuristic optimization of data preparation and machine learning hyperparameters for prediction of dynamic methane production, Bioresour. Technol. 128604 (2023).
[31] A. Motwakel, E. Alabdulkreem, A. Gaddah, R. Marzouk, N.M. Salem, A.S. Zamani, A.A. Abdelmageed, M.I. Eldesouki, Wild horse optimization with deep learning-driven short-term load forecasting scheme for smart grids, Sustainability 15 (2023) 1524.
[32] S. Myeong, J. Park, M. Lee, Research models and methodologies on the smart city: a systematic literature review, Sustainability 14 (2022) 1687.
[33] I. Namatēvs, Deep convolutional neural networks: structure, feature extraction and training, Inf. Technol. Manag. Sci. 20 (2017) 40–47.
[34] A. Nazir, A.K. Shaikh, A.S. Shah, A. Khalil, Forecasting energy consumption demand of customers in smart grid using temporal fusion transformer (TFT), Results Eng. 100888 (2023).
[35] F. Oldewurtel, A. Parisio, C.N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, M. Morari, Use of model predictive control and weather forecasts for energy efficient building climate control, Energy Build. 45 (2012) 15–27.
[36] A. Poulenard, M. Ovsjanikov, Multi-directional geodesic neural networks via equivariant convolution, ACM Trans. Graph. 37 (2018) 1–14.
[37] S. Sana, G. Sruthi, D. Suresh, G. Rajesh, G.S. Reddy, Facial emotion recognition based music system using convolutional neural networks, Mater. Today Proc. 62 (2022) 4699–4706.
[38] C. Sekhar, R. Dahiya, Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand, Energy 126660 (2023).
[39] A.S. Shah, H. Nasir, M. Fayaz, A. Lajis, A. Shah, A review on energy consumption optimization techniques in IoT based smart building environments, Information 10 (2019) 108.
[40] A.K. Shaikh, A. Nazir, I. Khan, A.S. Shah, Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series, Sci. Rep. 12 (2022) 22562.
[41] S.Y. Shin, H.G. Woo, Energy consumption forecasting in Korea using machine learning algorithms, Energies 15 (2022) 4880.

[42] A.K. Srivastava, A.S. Pandey, M.A. Houran, V. Kumar, D. Kumar, S.M. Tripathi, S. Gangatharan, R.M. Elavarasan, A day-ahead short-term load forecasting using M5P machine learning algorithm along with elitist genetic algorithm (EGA) and random forest-based hybrid feature selection, Energies 16 (2023) 867.

[43] J.F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, A. Troncoso, Deep learning for time series forecasting: a survey, Big Data 9 (2021) 3–21.

[44] L.Z. Velimirović, A. Janjić, J.D. Velimirović, Entropy-based fuzzy model for short-term load forecasting in smart grid, in: Multi-criteria Decision Making for Smart Grid Design and Operation: A Society 5.0 Perspective, Springer, 2023, pp. 81–108.

[45] P. Vujković, D. Ravšelj, L. Umek, A. Aristovnik, Bibliometric analysis of smart public governance research: smart city and smart government in comparative perspective, Soc. Sci. 11 (2022) 293.

[46] F. Wahid, R. Ghazali, M. Fayaz, A.S. Shah, A simple and easy approach for home appliances energy consumption prediction in residential buildings using machine learning techniques, J. Appl. Environ. Biol. Sci. 7 (2017) 108–119.

[47] F. Wahid, R. Ghazali, M. Fayaz, A.S. Shah, Statistical features based approach (SFBA) for hourly energy consumption prediction using neural network, Networks 8 (2017).

[48] F. Wahid, R. Ghazali, A.S. Shah, M. Fayaz, Prediction of energy consumption in the buildings using multi-layer perceptron and random forest, Int. J. Adv. Sci. Technol. 101 (2017) 13–22.

[49] F. Wahid, D.H. Kim, Short-term energy consumption prediction in Korean residential buildings using optimized multi-layer perceptron, Kuwait J. Sci. 44 (2017).

[50] K. Wang, M. Yu, D. Niu, Y. Liang, S. Peng, X. Xu, Short-term electricity price forecasting based on similarity day screening, two-layer decomposition technique and Bi-LSTM neural network, Appl. Soft Comput. 110018 (2023).

[51] Y. Wang, J. Wang, M. Cao, X. Kong, B. Abderrahim, L. Yuan, A. Vartosh, Dynamic emission dispatch considering the probabilistic model with multiple smart energy system players based on a developed fuzzy theory-based harmony search algorithm, Energy 269 (2023) 126417.

[52] Y. Xia, J. Wang, D. Wei, Z. Zhang, Combined framework based on data preprocessing and multi-objective optimizer for electricity load forecasting, Eng. Appl. Artif. Intell. 119 (2023) 105776.

[53] J. Yan, L. Mu, L. Wang, R. Ranjan, A.Y. Zomaya, Temporal convolutional networks for the advance prediction of ENSO, Sci. Rep. 10 (2020) 1–15.

[54] J. Yang, Q. Sun, L. Yao, Y. Liu, T. Yang, C. Chu, L. Fang, L. Zhu, A novel dynamic load-priority-based scheduling strategy for home energy management system, J. Clean. Prod. 135978 (2023).

[55] C. Yue, L. Wang, D. Wang, R. Duo, H. Yan, Detecting temporal attacks: an intrusion detection system for train communication ethernet based on dynamic temporal convolutional network, Secur. Commun. Netw. 2021 (2021) 1–21.

[56] M. Zeng, L.T. Nguyen, B. Yu, O.J. Mengshoel, J. Zhu, P. Wu, J. Zhang, Convolutional neural networks for human activity recognition using mobile sensors, in: 6th International Conference on Mobile Computing, Applications and Services, IEEE, 2014, pp. 197–205.

[57] N. Zeng, H. Zhang, W. Liu, J. Liang, F.E. Alsaadi, A switching delayed PSO optimized extreme learning machine for short-term load forecasting, Neurocomputing 240 (2017) 175–182.

[58] T. Zhang, X. Zhang, T.K. Chau, Y. Chow, T. Fernando, H.H.C. Iu, et al., Highly accurate peak and valley prediction short-term net load forecasting approach based on decomposition for power systems with high PV penetration, Appl. Energy 333 (2023) 120641.

[59] X.X. Zhu, C. Qiu, J. Hu, Y. Shi, Y. Wang, M. Schmitt, H. Taubenböck, The urban morphology on our planet – global perspectives from space, Remote Sens. Environ. 269 (2022) 112794.