THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Data-Driven Short-Term Load Forecast Method and Demand-Side Management for Distribution Network

Yuanda Gao

Doctor of Philosophy

THE UNIVERSITY OF EDINBURGH

2022

Abstract

With the development of the power grid, the smart grid makes the system more intelligent, efficient, sustainable and reliable with integrated Information and Communication Systems. Moreover, the data from the advanced system provides chances to utilise machine learning algorithms to improve the system operation further. In addition, the load profile is undergoing altering and becoming more unpredictable because of the increase in smart home appliances, EVs, e-heating systems, energy storage devices, etc. These factors bring more challenges and opportunities for the future power system to improve operational efficiency and demand response quality. In this regard, considering load forecasting is crucial in smart distribution networks for utility companies, especially those employing the demand-side management alternatives, the short-term load forecast could be more accurate and robust, evolve for future load forecasting purposes, and reveal its value in improving demand-side management qualities.

This research proposes a novel Dynamic Adaptive Compensation-Long Short-Term Memory (DAC-LSTM) forecast method. This method uses high time-resolution datasets and LSTM networks as fundamental to give short-term time-series load forecast results. The proposed method dynamically distinguishes the peak and off-peak hours and improves the forecast accuracy separately. For DNOs, the forecast errors, especially during peak hours, lead to penalties to start/stop backup generations or adjust the distribution schedules, and this will result in more operational costs. Further, the proposed method introduces a novel DAC block to compensate for forecast errors

according to the error trend calculated by historical forecast and actual load, then applying dynamic adaptive parameters. The greater the current-to-average forecast error ratio or the closer the forecast step to the present time stamp, the larger the compensation factors. Besides, the factor caps are set to prevent the model from over-compensation conditions. The sensitivity of introduced parameters is analysed, providing the performance of the developed method under different parameter values.

Afterwards, the proposed method is evaluated with six case studies, including varying the forecast steps (compared with LSTM and ARIMA), limiting the size and length of the training datasets (compared with ARIMA and Persistence), comparing with other state-of-art methods qualitatively, and comparing with ELEXON UK domestic load forecast results. Finally, the advantages of the DAC-LSTM method are validated, including providing accurate short-term load forecast results during peak and off-peak simultaneously, with a shorter length of or fewer households' historical datasets, and compared with existing transmission network forecast methods. The system operators, like DNOs, can reduce the operational cost with more accurate forecasts during peak hours as well as own more load curtailment potentials during the off-peak hours.

Additionally, more contributions, including the future bottom-up load scenarios establishment and the improved Stackelberg Game demand response for end-user utility bill reductions, will help system operators develop suitable DSM alternatives and tariffs based on more realistic and accurate analysis. To be more specific, first, based on the Ten Year Network Development Plan (TYNDP) 2018 and the UK government reports, bottom-up load profiles are designed and generated for the UK distribution network for the scenario years 2020, 2030 and 2040. The DAC-LSTM method is evaluated with these scenario profiles, yielding up to 0.989 and 3.79% (measured in R2 and MAPE) forecast accuracy, for various levels of electric vehicle and e-heating penetration when compared with the ARIMA and Persistence methods. Second, a

DSM alternative is built based on a Stackelberg Game to reduce the consumers' utility bill, which considered the forecast error as a constraint. In this game, given that the consumer offers maximum controllable power to the operator, the game achieves the Stackelberg Equilibrium while maximising the operator's revenue and supplying the necessary power to consumers. The case study demonstrates that when considering forecast errors in demand response strategies, higher forecast accuracies reduce the electricity bill up to 10.4% in an ideal circumstance. The improved Stackelberg Game makes the forecast error one primary constraint that most existing DSM alternatives lack. This proves the value of utilising state-of-art forecast methods in the deployment of DSM alternatives.

Lay Summary

With the anticipated altered demand profile and requirement of improving demand-side management efficiency and quality, the short-term load forecast method, which is crucial in smart distribution networks, embraces the opportunity of implementing machine learning techniques to improve forecast qualities. Furthermore, as demand-side management alternatives will be applied to the forecasted load rather than the actual load, state-of-art short-term load forecast methods should be able to provide convincing results under rigorous situations. For example, the historical dataset is limited or incoherent, or the load profile is highly non-linear due to electric vehicle charging, smart home appliances, etc.

In this regard, this thesis provides detailed development and evaluation procedures of a novel machine learning method for short-term load forecast. The proposed method is evaluated with different case studies that the future distribution network might encounter, such as the length and size of the historical dataset being limited and the forecast step requirement changes. In addition, the proposed method is evaluated by quantitatively comparing with other extensively used methods, such as ARIMA and Persistence methods, and qualitatively comparing with other state-of-art hybrid machine learning approaches.

Moreover, bottom-up load profiles are generated for the UK distribution network in the scenario year 2020, 2030 and 2040. These profiles contain an aggregated base, electric vehicle, and e-heating demand profiles from several individual households. These individual households' datasets are collected from Thames Valley Vision Project and the UK Energy Research Centre, or generated from the UK National Travel Survey. Finally, the proposed method is evaluated with these scenarios.

Notably, the impact of utilising an accurate forecast method in demand-side management strategies is assessed by introducing the Stackelberg Game, which maximises the operator's revenue and supplies the required power to consumers, given that the consumer offers maximum controllable power to the operator. This work firstly involves forecast accuracy as a primary constraint to establish the Stackelberg Game. In conclusion, the electricity price is inversely proportional to the controllable power consumers offer. The case study illustrated that the bill reduction is inversely proportional to the forecast error in management.

Acknowledgements

I would like to express my deepest appreciation to my principal supervisor, Prof. Aristides Kiprakis, for his continuous patience and guidance since my first year in the UK. His support encouraged me to start the PhD and allowed me to complete this work. Also, I am deeply grateful for his help in steering this research in the proper direction through many hours of insightful discussions and feedback. Thanks are also extended to the other supervisors, Dr Jonathan Shek and Dr James Robertson, who shared knowledge selflessly and guided me patiently. Many thanks to Prof. Harrison Gareth for his suggestions during my study. Meanwhile, I would like to thank the thesis examiners for their time and efforts in reviewing the thesis and giving comments.

Thanks to my colleagues and friends from the Agile Research Group and the Institute for Energy Systems, particularly all those who have passed through Room A110. Though the pandemic hindered us from seeing each other physically, the online meeting still 'works'. I wish everything back to normal and at peace.

To Dong – Whether happy or sad, poor or rich, I am glad to spend a walk through every four seasons with you.

To my family and friends – Life would have been so mundane without you all.

To my parents – You are my strongest backing, giving me the most selfless support. I wish you a healthy and trouble-free life.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Yuanda Gao

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Research Background

Power systems are defined as a network of electrical components implemented to generate, transfer and distribute electric power. In the past years, most studies focused on the generation and transmission levels, such as increasing the power generation capacity, reducing power generation and transmission cost, stability and security of power network, maintenance of network equipment, etc. However, with the development of technology, the traditional power system issues are not dominated in the future power system. Some new challenges such as better customer services, environmental protection and more social responsibility [1].

With the development of information and communication technologies (ICT) and improved metering infrastructure (AMI), the smart distribution network (SDN), also called active distribution networks (ADN), is substituting the traditional distribution networks [2]. Meanwhile, the smart grid (SG) encouraged the increasing penetration of distributed generations (DGs) and renewable energy sources (RES), and focuses on the reliable operation transition [3, 4]. Due to these facts, the power network is becoming more interactive and sophisticated, which requires improvement in planning and op-

eration management. Notably, the planning and operation of SDN need precise load forecast results from various time steps. Therefore, except for the long-term network planning consideration, the accuracy of short-term load forecast is vital to the modern SDN [5, 6].

Consequently, with state-of-art ICT and AMI implementation, massive system operation data is recorded and awaiting for later analysis. Using big data for load forecast in the power system can improve existing services and benefit both operators and consumers. It can also assist the construction of a data-driven distribution network, the next generation of power systems. In this progress, the development of state-of-art load forecast algorithms requires the consideration of prospects of future distribution networks, such as network decentralisation, distributed renewable generations, demand profile altering, etc.

1.2 Research objectives and scope

To utilise massive system operation data acquired from the network, machine learning techniques are involved in improving existing services. In practice, DSM alternatives are deployed to improve the network operation quality, provide optimised services to consumers, reduce costs, etc. Meanwhile, the DSM alternatives are managed especially based on short-term load forecast results. As research demonstrated, a 1% improvement in forecast accuracy can save 10 million pounds for one electric utility in the UK [7]. Thus, the value of combining machine learning methods with short-term load forecasts to benefit DSM is obvious.

Besides, with many countries having released the carbon neutral target by different target years, the RES and EVs penetration level, home appliances usage pattern, heating system, etc., will significantly change and switch to an environmentally friendly path. The residential and commercial electricity load patterns will change correspondingly. Therefore, when designing the novel forecast method, the method should be suitable for future load scenario applications.

According to the facts mentioned above, the research objectives can be summarised as follow:

1. Development of adaptive short-term load forecast methods based on machine learning approaches. The method is developed for distribution network applications, providing accurate peak and off-peak load forecast results. Besides, the method should be optimised when the input training dataset is insufficient or incoherent. Moreover, the method should be evaluated with other existing methods.
2. Based on existing research and government policies, bottom-up load scenario profiles should be generated for the United Kingdom (UK) distribution network. These load profiles should take the penetration level of different load compositions into account, emulating the actual circumstances. In addition, the developed forecast method should be evaluated with future load scenarios.
3. The value of deploying an accurate forecast method should be evaluated with demand-side management alternatives. To achieve this objective, the game theory is introduced and modified to fit the evaluation purpose.

The scope and boundaries of this research are defined as follows:

1. The dataset used for forecast method development is expanded with linear interpolation and adding random errors to emulate a high-resolution dataset. In addition, the EV charging load profiles are collected or generated from American Adaptive Charging Network and the UK National Travel Survey, which are different from the basic load profile from Thames Valley Vision Project.
2. The forecast method, its evaluations, and demand-side management are implemented and optimised in Python. However, this research does not discuss the required computational power and time consumption for model training.
3. The demand-side management method introduced is based on the existing Stackelberg Game that fits the research purpose with modifications. This is a supplement and perfection of previous work.
4. The thesis focuses on the short-term load forecast for the distribution network only. Because the load profiles for the mid and long-term periods are highly sinusoidal (which can be forecasted using linear approaches) and utilised for investment analysis and power network planning (instead of real-time control), these two periods are not discussed in detail in this work.

1.3 Thesis statement

Machine learning techniques can be utilised to develop adaptive forecast methods which provide accurate forecast results for peak and off-peak load, and adapt to smaller network applications.

1.4 Acknowledgement of the thesis contributions

The main contribution of this research can be summarised as:

1. A novel dynamic error compensation method is developed for short-term load forecast based on long short-term network structures. The proposed method presents robust forecast accuracy when the load is highly non-linear or the training dataset is limited, and this makes the forecast method appropriate for load patterns that include high Electric Vehicle penetration levels and smaller networks.
2. Bottom-up future load scenarios for the UK distribution network are built based on existing research and policies for 2020, 2030, and 2040. The future load scenario aims to evaluate the performance of the proposed forecast method.
3. The proposed forecast method is evaluated with demand-side management strategies, which illustrates the advantage of utilising the forecast method with higher accuracy. The demand-side management method utilises Stackelberg Games which finds an equilibrium point for companies and consumers. This method introduces forecast error factors to emulate real-world conditions and proves the value of utilising forecast methods with higher accuracy.

1.5 Thesis structure

This thesis is divided into seven chapters. Chapter 1 includes an overview of the whole research area, highlighting the contributions of this project and forming the introduction of this doctoral thesis.

Chapter 2 reviews the literature and achievements of related topics in three main subjects. The first is the development of the distribution network, which introduces that the current distribution network is implemented with more distributed generations, undergoing load profile altering, equipped with advanced monitoring and management systems, and becoming much 'smarter'. These changes generate a massive amount of system operation data which can be utilised to improve existing service quality. Load forecasting is crucial to delivering high-quality services such as long-term planning, off-line planning, scheduling, and real-time operational planning. Secondly, short-term load forecast methods widely utilised in the power network are reviewed because the methodologies of these approaches are frequently referred to for novel method development. The typical statistical methods can be concluded as follow: multiple linear regression, exponential smoothing, ARIMA and its variants, and other methods. The typical machine learning methods can be concluded as follow: support vector regression, fuzzy logic, artificial neural network, random forest and hybrid learning methods. Thirdly, applications of machine learning techniques in load forecast are reviewed, including short-term, mid-term and long-term load forecast methods. Referenced forecast approaches can be classified into statistical and machine learning methods, and the positive and negative aspects of these methods are discussed. Finally, the importance of load forecast in demand-side management is reviewed. The main objectives of the smart grid and typical demand-side management techniques are referred to, followed by the discussion of the importance of load forecast in electrical power businesses. In addition, the research gap is noted at the end of this chapter.

Chapter 3 presents the detailed development procedures of a novel short-term load forecast method. First, the Persistence and ARIMA methods are explained and the forecast results will be compared with the proposed method. Next, the development of the proposed forecast method can be concluded as follow: description of long

short-term memory network and development of dynamic adaptive compensation method. The followed chapter 4 carries out six case studies to evaluate the forecast method developed in chapter 3. These case studies emulate real-world application circumstances to test the robustness of the proposed method.

Chapter 5 discusses the build-up of future load scenarios for UK distribution networks. This chapter introduces the standards, research, and procedures required to generate bottom-up future load profiles. Finally, the proposed load forecast method is evaluated with future load profiles.

Chapter 6 introduces a Stackelberg Game as a demand-side management strategy for utility bill reductions. First, the modifications and methodologies of the Stackelberg game are discussed, followed by a case study that illustrates the value of the proposed load forecast method in network management.

Chapter 7 is a summary and overview of all contributions to the research made in the previous chapters. Furthermore, some limitations of the research are discussed, and a future trajectory for improving this research is stated.

Chapter 2

Literature Review

2.1 Introduction

The causal relationship underlying socio-economic growth and the availability of reliable, pervasive, and high-quality electricity has long been acknowledged as one of the significant impediments hindering development in numerous regions of the globe. First, renewable energy can be regarded as a sustainable and reliable source of electricity in these economies [8, 9, 10]. Incorporating large populations and/or distributed energy resources (DER) capacities and intelligent devices, such as small-scale renewables, controllable loads, and storage, into a network may introduce planning, management, and real-time control challenges. Second, the advancement of Information and Communication Technologies (ICTs) in conjunction with smart monitoring devices, including smart meters, Phasor Measurement Units (PMUs), aggregators, etc., generates massive real-time data in terms of volume, velocity, and variety [11]. With the integration of DER and intelligent devices, big data could provide new opportunities for demand-side energy management, such as energy planning, efficient energy generation, and distribution [12].

On the one hand, the load pattern has altered significantly in the distribution network, so user consumption behaviour is considerably more difficult to forecast, and this will result in a greater generation mismatched between supply and demand. In addition, the growth in the energy generated by renewable sources increases the load on power lines due to the significant fluctuation of their performance, particularly in distributed networks. On the other hand, by implementing advanced measuring devices such as PMUs, the system status with high precision and a short time interval could be monitored for analysis, prediction, and management. Furthermore, the acquired data might be utilised to train the machine learning model, such as load pattern recognition, short-time load prediction, etc., even at the substation or bus level via high computation capabilities [13, 14, 15]. This will provide the system operator with an additional opportunity to detect abnormal system operating status and deploy appropriate management techniques with a shorter response time and at a reduced cost.

Furthermore, the future load pattern is influenced by the policies of governments and organisations. One hundred thirty-seven counties have committed to carbon neutrality (CN), with the commitments centred on the year 2050. On the residential side, the CN is achieved by converting to low-carbon emission transportation, low-carbon heating system, and boosting energy efficiency, among other measures [16]. In addition, the CN is realised in the industrial and commercial sectors by increasing the proportion of renewable energy generation and implementing industrial CO_2 capture, removal, storage, and utilisation devices, etc. [17]. Based on these changes, the future load pattern at the distribution or lower network level will become even more non-linear, rendering the current load forecasting approach inaccurate. Consequently, the system operation can be benefited from utilising big data and machine learning techniques.

2.2 Distribution Network

For hierarchical-based traditional grids, the merging problems including increasing demands, ageing infrastructure, and limited expansion options bring more challenges to keep up with current needs [18]. Furthermore, based on the existing technology, the competitive energy market and service requirements have reached limits, which leads to overstressed grid operations, especially at the distribution network (DN) [18]. Before, the DN has been designed to maintain unidirectional power flow with radial architecture to deliver power to the end customer efficiently. However, escalating load requirements over vast geographical distances have resulted in significant technical challenges in DNs, including increased system losses, reduced voltage regulation, impaired power quality, reliability concerns, and costly planning options [19, 20].

Because of the development of information and communication technologies (ICT) and improved metering infrastructure (AMI), the smart distribution network (SDN), also called active distribution networks (ADN), is substituting the traditional distribution networks [2]. Meanwhile, the smart grid (SG) encouraged the increasing penetration of distributed generations (DGs) and renewable energy sources (RES) and focuses on the reliable operation transition [3, 4].

Due to these facts, the power network is becoming more interactive and sophisticated, which requires improvement in planning and operation management. Notably, the planning and operation of SDN need precise load forecast results from various time steps. Therefore, except for the long-term network planning consideration, the accuracy of short-term load forecast is vital to the modern SDN [5, 6]. Furthermore, the modern grid utilises supervisory control and data acquisition (SCADA) systems to monitor the SDN. Meanwhile, distribution management systems (DMS) and energy management systems (EMS) serve as decision-support information systems for the

coordination of remote SDN equipment. In addition, the extensive use of devices such as distribution transformer terminal unit (TTU), feeder terminal unit (FTU), remote terminal unit (RTU), and distribution automation terminal (DTU) contribute to the maturity of SDN [5, 6]. The development of intelligent and interactive distribution grids from the past to the smart future is depicted in figure 2.1(redrew from [18]).

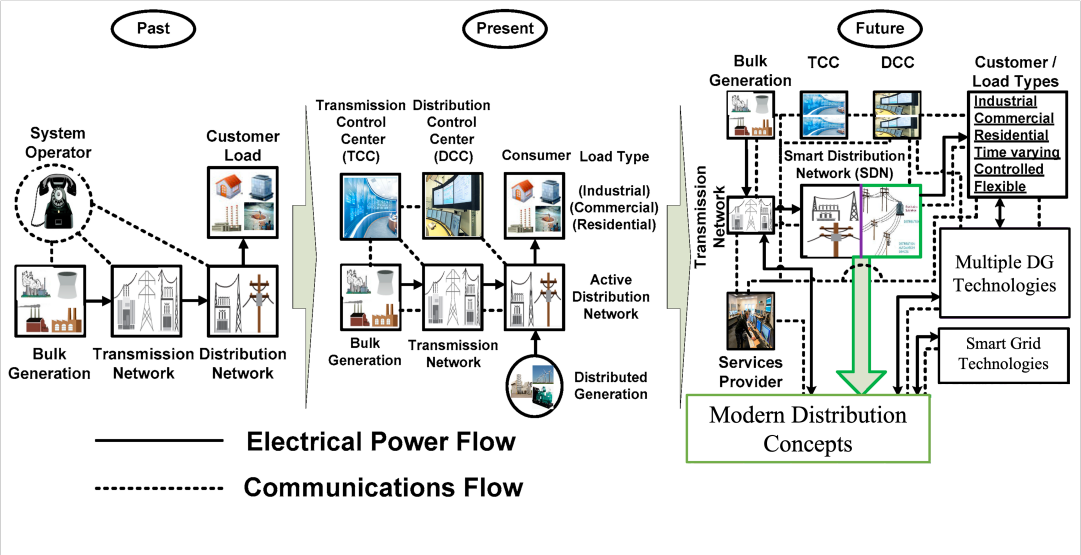


Figure 2.1: Development process of distribution grids from the past to a smart future [18]

The SG paradigm necessitates broad stakeholder input from multiple SDN perspectives, particularly planning. Consequently, a deeper, broader, and aggregated approach to planning in the actual world is contagious. The typical planning issues centre on locating an economically viable (cost-effective) solution. In addition, objectives such as reliability, power quality, minimal (negative) environmental effect, system stability, energy efficiency, and customer satisfaction can be included in the formulation of SDN problems [18]. Furthermore, long-term SDN planning, off-line planning problem, scheduling, and real-time operational planning can be used to classify the SDN planning difficulties planning, off-line planning problem, scheduling, and real-time operational planning [21, 22, 23]. Finally, the load forecast is crucial to delivering high-quality services in the aforementioned planning methodologies.

2.3 Load Forecast Methods

The electric utility makes important short and long-term decisions based on load forecasting, including power system planning, operation and control. Load forecast helps planning and operation necessitate a specific amount of lead time, which is often known as forecasting intervals or forecast stages. Planning for the growth of a power system begins with a projection of anticipated future load requirements. Demand and energy requirements estimation is essential for effective system planning. Demand predictions are used for determining the generation capacity, transmission, distribution system additions, etc. Load forecasts are also used to define procurement plans for construction capital energy estimates, which are necessary for determining future fuel requirements. Therefore, an accurate prediction that reflects current and future trends is essential for all planning. However, as a result of the liberalisation of the energy industries, load forecasting is of even greater importance.

Moreover, the load forecast can also be classified into two categories [24]: demand forecasting and energy forecasting. Demand forecasting is used to determine the capacity of the generation, transmission, and distribution system additions. History and government policy can be used to anticipate future demand based on the rapid growth rate of demand in the past. The energy prediction is used to determine the future facility types required.

The load forecast techniques are classified into statistical methods and machine learning methods. This chapter refers to the most classical and widely utilised methods as the merits and demerits of these methods are considered in structuring this thesis. Moreover, the principles of building these existing methods are referred to in developing the proposed novel forecast method, which is discussed in chapter 4.

In the past, the statistical methods are developed by traditional (conventional) mathematical techniques where load forecasting is considered a linear problem. From the transmission network to the distribution network, however, the linearity of the combined demand curve declines. At the level of transmission networks and for long-term forecasts, statistical methods yield accurate forecast results. But when it comes to the distribution network or even microgrid level and the short-term forecast, the previously used forecast methods are deficient. In recent years, as computing power has increased, people have realised that machine learning methods are better suitable for handling nonlinear problems. Therefore, research and implementations of cutting-edge machine learning techniques are widely utilized today.

Several essential short-term load forecasting methods will be explained in this part, along with references about improved methods.

2.3.1 Statistical Methods

Multiple Linear Regression

Multiple linear regression (MLR) models the relationship between multiple independent variables and one dependent variable. The method can be represented below [25]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon \tag{2.1}$$

where, y is the dependent variable, x_k is k independent variables, β is the regression coefficients and ε is represented as error term. Besides, for multiple observations, we use multiple equations of y and these equations can be expressed as follow:

$$y = X\beta + \varepsilon \tag{2.2}$$

Where,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 1 & x_1 2 & \dots & x_1 k \\ 1 & x_2 1 & x_2 2 & \dots & x_2 k \\ \vdots & & & & \\ 1 & x_n 1 & x_n 2 & \dots & x_n k \end{bmatrix} \quad \beta = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \varepsilon = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{2.3}$$

Based on equation 2.3, the forecast load can be represented as follow:

$$y = X\beta + \varepsilon \tag{2.4}$$

$$\hat{y} = X\beta \tag{2.5}$$

where $\hat{y}$ is the prediction of y.

The MLR method fit the electrical load curve well when the load trend shows obvious periodicity. But as a result of the modification of modern load characteristics and the application of the forecasting method to a smaller network, the load periodicity decreases, and MLR can no longer effectively manage the STLF.

Exponential Smoothing

Exponential smoothing (ES) is a popular and dependable STLF technique over alternative linear methods due to its computational simplicity, intuitiveness, and ability to spot linear trends in a time series. ES bases each new forecast on the preceding forecast in addition to a percentage of the forecast error at that time [26]. The difference between the actual value and the anticipated value is the forecast error. Consequently, ES is suitable for non-stationary data (i.e. data with a trend and seasonal data). The forecast error is the difference between the actual value and the forecast value:

$$ForecastError = ActualValue - ForecastValue \tag{2.6}$$

The Exponential Smoothing forecast is computed using the following equation [27]:

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1}) \tag{2.7}$$

Where, F_t is the forecast for the period at time t, α is the smoothing constant (represents a percentage of the forecast error, A_{t-1} is the actual demand for the previous period. α determines the responsiveness of the adjustment of the forecast to error. When it tends to zero, the forecast is less responsive to the error. Conversely, as it tends to unity, the forecast error influences increase, and the smoothing of the forecast decreases. In real-world applications, the smoothing constant is the value which balances the responsiveness of the model and the smoothing random variations [28]. The value is determined by trial and error, with the forecaster making an informed selection.

Similar to the MLR method, ES shows weakness in forecasting load with lower periodicity. Consequently, the accuracy of forecast results diminishes in the distribution or smaller network.

ARIMA and Other Methods

Besides MLR and ES, Kalman Filter (KF) method is used for bottom-up load forecast in [29]. The method focuses on analysing the behaviour of home appliances and measuring the error rate. But the bottom-up method works only when most appliances are well-modelled and the consumer behaviour follows a fixed pattern. Also, by aggregating the households, the uncertainty of load forecast at a specific period will be

amplified. In [30], a Blind Kalman Filter (BKF) is proposed based on the linear state-space method with unknown state and observation matrices that are sequentially estimated from the data. The methods assume a small segment of the entire time series to be linear and estimate the forecast result.

Another commonly used method is ARIMA. ARIMA is a generalised method of Autoregressive Moving Average (ARMA) that combines the Autoregressive (AR) process and the Moving Average (MA) processes and builds a composite model of the time series. But ARIMA methods should be used on stationary data, therefore, parameters should be predefined to remove the trend from the data. The detailed method description can be found in Section 4.2.2. Beyond the ARIMA method, which is univariate, Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) is developed for multivariate applications [31].

In conclusion, the statistical methods are classical when the computational power is lacking, the machine learning techniques are not well developed, and most load forecasting requirements exist at transmission network levels. However, with the development of power systems, the 'smart' network requires load forecasts for smaller networks and better accuracy. Therefore, the statistical methods become non-competitive.

Moreover, the statistical methods cannot accurately forecast peak and off-peak load simultaneously. ES and ARIMA methods consider compensating the errors during forecasting, but the compensation factors are fixed. Therefore, the robustness of the forecast method is limited. These leave room for future improvements.

2.3.2 Machine Learning Methods

Support Vector Regression

Support Vector Machine (SVM) is a supervised machine learning algorithm, which is designed for classification and regression problems. The Structural Risk Minimisation (SRM) principle, which is superior to the traditional Empirical Risk Minimisation (ERM) principle, is the basis of the SVM. Its formulation is also used by conventional neural networks [32, 33]. SRM minimises an upper bound on the estimated risk, while ERM minimises the training data inaccuracy. This distinction gives SVM a stronger capacity for generalisation [34]. Numerous MTLF and LTLF forecasting methods employ SVM methods alone or in combination with other algorithms. SVM is effective in solving high-dimensional spaces, and memory efficient as it uses subsets of training points in decision function making. However, the algorithm provides the probability estimates through a costly five-fold cross-validation procedure, which is not directly derived. However, the algorithm does not directly provide probability estimates, as they are derived through a costly five-fold cross-validation procedure. Numerous MTLF and LTLF forecasting methods employ SVM methods alone or in combination with other algorithms. In [35], the SVM was compared against other algorithms, such as the AR method. In this case, the SVM allowed the training dataset to be above the limit that the AR method or other neural networks allow. Based on the root-mean-square error (RMSE), the performance of SVM was greatly improved by increasing the training data to around two years.

As SVM was initially designed to tackle binary classification problems and was later adapted to regression problems, the Supported Vector Regression (SVR) approach is developed as an enhanced method for handling nonlinear difficulties. Figure 2.2 (redrew from [36]) displays the primary structure of SVR. The goal is to locate a smooth function in feature space as opposed to input space.

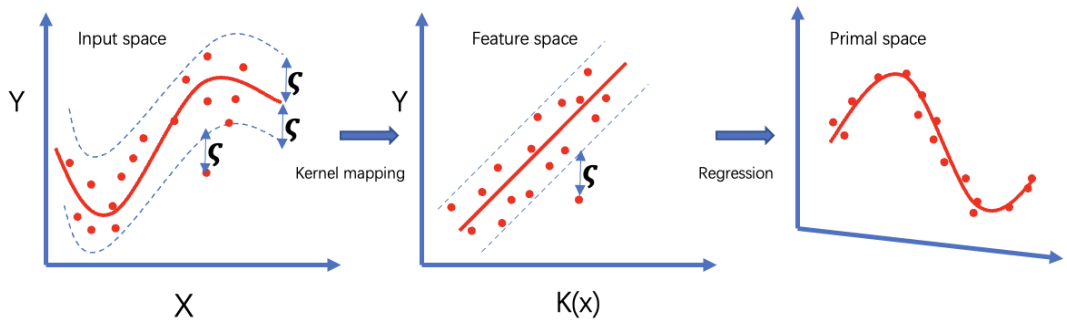


Figure 2.2: The main structure of SVR [36]

Fuzzy Logic

Fuzzy logic (FL) is a multi-valued logic that handles reasoning approximately, unlike precise reasoning. The theory was to enable computers to emulate human thought processes. Unlike the binary sets, which are based on Boolean logic (either 0 or 1), FL has values ranging between 0 and 1. Therefore, particular functions are required to regulate the degree to which a variable belongs to a fuzzy set as FL utilises linguistic variables. Fuzzy sets are useful because they can model unclear data and deliver appropriate decisions.

The relationship between load demand and factors like temperature and humidity is nonlinear. In [37], the FL is used to scale the highly non-linear relationship between the weather elements and the electric load profile. Moreover, FL is utilised to pre-process input data prior to sending it to the neural network structures. In contrast

to the internal neural networks, which are like black boxes and opaque to the user, the fuzzy inference method is defined by the users. Therefore, as the learning and optimising capabilities of neural networks are superior to FL, both methods can be combined while ignoring their weaknesses and providing effective results.

Artificial Neural Network

The Artificial Neural Network (ANN) consists of numerous frameworks, such as RNN, LSTM, CNN, Transformer, etc., each of which has been utilised in a variety of situations. The operation of the ANN is following four principles, which makes the ANN successfully applied in scientific fields [33, 38, 39]:

- To train the network, a big database is necessary. The training process compares the known inputs with the corresponding outputs.
- The output values are produced based on the actual output values and the amendment of the weights in accordance with the network.
- With a proper network, the error reduces as the number of repetitions increases. The trained network is achieved when the error is below the threshold determined at the beginning of the training process.
- The network is regarded as trained when it responds correctly to the new incoming inputs.

The traditional ANN maps the input historical data and output forecast value but lacks consideration of time correlation in the data sequence. This causes the problem that the ANN cannot find the relationship between data and time. An RNN is an improved network of ANN, which uses the temporal information of the input data, where connections between units form a directed cycle within the same layer. Therefore, the output of each time step is affected by the input data from previous steps. But the vanishing gradient problems affect the RNN in processing long time series datasets. As more

layers using certain activation functions are added to neural networks, gradients of the loss function approach zero, making the network hard to train and effectively preventing the weight from changing its value. A fundamental method that can alleviate vanishing gradient problems is to replace the activation function used in the neuron network. But this method affects the performance of the model in specific cases. The LSTM network is a variation of RNN that alleviates difficulties with vanishing gradients in RNN by allowing the LSTM cell to recall or forget relevant or irrelevant input data. The details of the RNN and LSTM network can be found in Chapter 3.

CNN is a class of ANN which can process high-dimensional data and is frequently used for visual image, video, and text classification. Power system sensors and gadgets are highly intelligent. Therefore, the spatial information data recorded by the smart sensors and devices in the SDN is preserved and the position and time sequence of these sensors and devices are stored as well. For the CNN, the spatiotemporal power system data is regarded as an image or, video and text sequences to make the load forecast. The CNN structure is shown in Figure 2.3 (redrew from [40]).

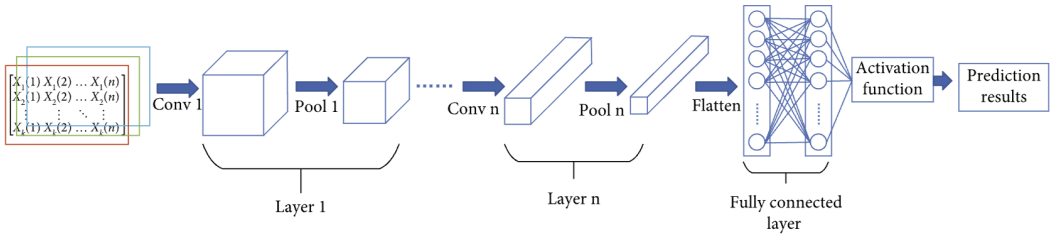


Figure 2.3: Structure of CNN [40]

As shown in figure 2.3, the two-dimensional spatiotemporal matrices are firstly stacked into three-dimensional blocks. Afterwards, these blocks were passed to the convolution operation blocks, where the highly abstract feature of the input dataset will be extracted. After the convolution operation, the outputs are applied to the pooling operation. The pooling operation reduces the size of the matrix and the nodes numbers, therefore, reducing the parameters in the CNN. During the pooling operation,

the depth of the input matrix is maintained. Next, after repeated convolution and pooling operations, the highly abstract feature was obtained and flattened to a one-dimensional vector, which is connected with the full connection layer. In this step, the iterative calculation is applied to the weights and bias parameters obtained from the full connection layer. Finally, outputs from the full connection layer are passed to the activation function and the prediction results are obtained.

Random Forest

The Random Forest (RF) approach is an integrated, decision-tree-based algorithm that is always combined with other methods. In [41], the RF is employed to pick features, which is then followed by the CNN method. In [42], the RF is used as part of the decision tree method for the load forecast. The RF approach is more robust to noise and missing data, has a quicker learning rate, and its primary steps are as follows [42]:

- For each decision tree, the training set P is constructed by N training sample data with N times random sampling and put-back bootstrap. These N training samples have M attributes. For those data which has not been sampled N times, these data are regarded as out-of-bag data and not counted in the training set sample.
- The CART (Classification And Regression Tree) decision tree T is generated based on the training set samples from the previous step. The split attribute set of the current tree node is randomly selected from M attribute sets, with each node containing $m(m < M)$ attributes. Meanwhile, the optimal path to split the node among the m attributes is found. In this process, the stopping point is reached when each tree grows as much as possible or reaches the conditions set in advance.

- By repeating the steps mentioned above, the training sets $P_1, P_2, \ldots, P_n$ and decision trees $T_1, T_2, \ldots, T_n$ are obtained to construct the RF model.
- The prediction is based on the model built. The forecast value is the average of the prediction from all trees.

In the regression tree, the minimum mean square deviation is used as the evaluation standard:

$$\min_{a,t}\left[\min_{c_1}\sum_{x_i\in R_1(a,t)}(y_j-c_1)^2+\min_{c_2}\sum_{x_i\in R_2(a,t)}(y_j-c_2)^2\right] \tag{2.8}$$

In addition, academics have proposed other machine learning approaches, such as backpropagation, optimisation algorithms, etc. However, these procedures are typically included in the aforementioned methods.

In conclusion, the machine learning methods utilise tremendous computational power to solve non-linear time-series problems which demand load curve represents, especially RNN. In general, the performance of RNN based network achieves similar forecast results compared to a random combination of SVR, FL, RF, etc., methods. However, the ability to provide accurate forecast results for both peak and off-peak periods of these approaches is still lacking. Therefore, published approaches focus on improving forecast accuracy for only peak loads. These demerits leave room for future improvement to develop a method that provides an accurate forecast method for peak and off-peak periods.

2.3.3 Hybrid Learning Methods

In recent years, hybrid methods have been developed to handle forecasting difficulties because they combine the benefits of statistical and machine learning approaches. In [37], an RNN-FL method is proposed and the computed results conclude that the synergistic use of FL with the RNN method yields higher forecast results by analysing

the relationship between the weather parameters and the change in load profile. In [43], a hybrid approach, Support Vector Regression-Long Short-Term Memory (SVR-LSTM) is presented for the microgrid (MG) load forecasting and the values of the new methods compared with SVR and LSTM are evaluated in this paper. In [44], a hybrid method for STLF for higher educational institutions, such as universities, using RF and MLP methods for day-ahead load forecast. Also, this method utilises the decision tree as the classifier to pre-process the input dataset. Moreover, dozens of hybrid methods are proposed every year and the forecast accuracy outperform the original methods in each of them in different application range.

2.4 Machine Learning in Electrical Load Forecast

The earliest record of load forecasting dates back to the late 1960s when the first publication on load forecasting techniques was published [45]. Nowadays, the load forecast has become a crucial part of the power system operation. The system operators, suppliers, financial institutions, and power production, transmission, and distribution section players require an accurate load forecast result. The load forecast can be split into three primary groups based on the time scale [46, 47, 48]:

- Short-term load forecast (STLF): STLF periods range from minutes to days or weeks in advance. STLF aims for economic dispatch, optimal generating unit commitment, and real-time control and security evaluation.
- Mid-term load forecast (MTLF): The duration of MTLF is between one month and two years. MTLF tries to balance demand and generation via maintenance scheduling, coordination of load dispatch, and price settlement.

- Long-term load forecast (LTLF): The duration of LTLF is several years or longer. LTLF seeks to plan system growth, including generation, transmission, and distribution. In some instances, it affects the acquisition of new generating units.

Each length of time-period load forecast is crucial to the operation of the power system. The unpredictability of the forecast outcome contributes to the loss and instability of the economy and network control. In STLF, the accuracy of the forecast influences the control efficiency, and In MTLF and LTLF, load forecasting is intimately connected to system development, particularly system design and economics. In addition, the load forecast is essential for dependability analysis [48, 49, 50].

In contrast, the load forecast can be classified as either spatial load forecast (SLF) or hierarchical load forecast (HLF). Traditionally, load forecasting at the small area or equipment level, such as distribution transformer, is often referred to as spatial load forecasting [51, 52]. With the advent of the smart grid, large numbers of smart meters have been deployed during the past decade, providing the system operator with vast quantities of data. This information is both temporally and spatially granular [32]. By combining the data collected from WAMS with advanced computational technologies and forecast methods, spatial load forecast has been transformed into a hierarchical load forecast. The HLF provides forecasts from the home level to the corporate level, from a few minutes to several years in advance. The most significant development of HLF refers to the Global Energy Forecasting Competition 2012 (GEFCom2012) [53]. In addition, probabilistic electric load forecasting (PLF) is proposed as a load prediction category, but its use is limited. PLFs may base on scenarios. However, scenario-based forecasts are not probabilistic forecasts until probabilities are assigned to the scenarios. PLFs could be quantiles, intervals, or density functions [32]. In [54], [55] and [56], a number of PLF techniques are offered for wind power forecasting.

In general, the load forecast follows the steps shown in 2.4, with data collection and model training being the most critical procedures. Several strategies can be employed to process the data and increase the accuracy and resilience of the forecast method.

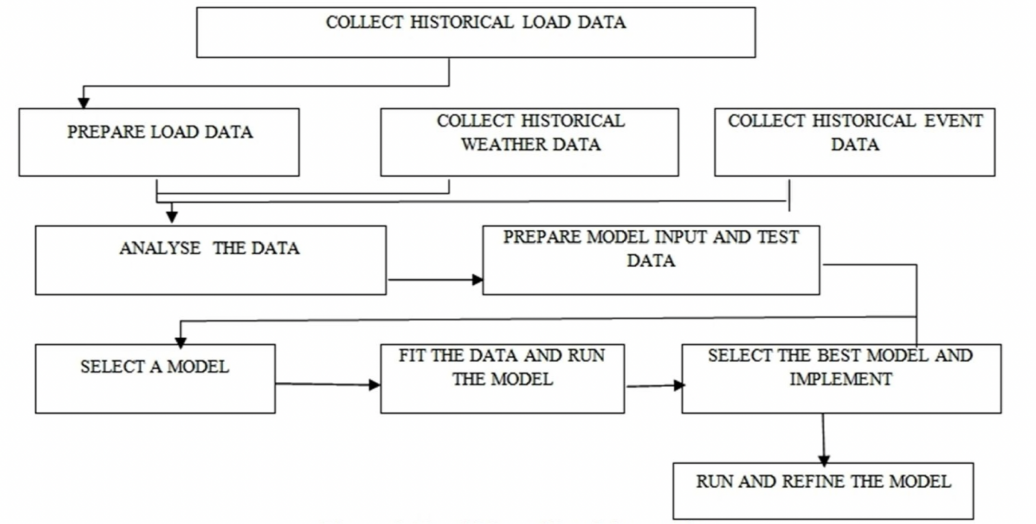


Figure 2.4: Roadmap of load forecasting [57]

2.4.1 Short-Term Load Forecast

Short-term load forecasts are essential for balancing electricity supply and demand, managing and planning power transmission and distribution, providing ancillary services, and supporting operations and maintenance activities [58]. For the electrical utilities, research carried out in [59] shows that a 1% decrease in STLF error can result in a £10 million reduction in the annual operating cost. To be more specific, the improved forecast accuracy minimised the forecast error between forecast and actual load, especially during peak time. The reduced forecast mismatch reduces the backup generators that the system operator should start during peak hours, therefore reducing the cost. While during the off-peak hours, the system operator has plenty of time to balance the power generation and consumption as the load pattern remains flat. In the preceding decades, multiple STLF approaches have been presented, and the STLF has become an active study area.

The STLF approach can be broadly categorised as statistical and artificial intelligence-based methods [60]. First, the statistical methods are classified into two categories: time-series [61] and multiple linear regression (MLR) methods [62]. The time-series methods, typically persistence and Autoregressive Moving Average (ARMA) methods and the variants, are commonly utilised in peak load estimation [63]. The Autoregressive Moving Average methods are improved by combining the Autoregressive and Moving Average methods. In recent years, it has been suggested that the Autoregressive Integrated Moving Average method (ARIMA) and its derivatives should replace the older techniques [64][65][66].

However, the ARIMA approaches only examine the electrical load [67] and do not account for additional effect factors such as weather conditions, special events, etc. Consequently, the forecast inaccuracy may increase in numerous circumstances [60]. In contrast, MLR approaches consider additional variables, such as meteorological conditions, and have been widely used in recent decades. In [68], an MLR-based STLF method is suggested for the day-ahead load forecast. This approach accounts for meteorological conditions and yields a Mean Absolute Percentage Error (MAPE) of 3.5% in the dry season and 4.3% in the wet season. Also, [69] achieves a MAPE of 4.5% in the hourly load forecast.

According to [70] and [71], the stationary method demonstrates good performance only over the stationary data, although the traditional electric load is simple, lacking aspects such as renewable generation, smart home appliances, storage devices, etc. The ARIMA method analyses time-series data based on the assumption that the collected and forecast data are linearly related at the forecast point, while the actual load pattern is highly non-smooth and non-linear [72]. Currently, the load non-linearity at the transmission network (TN) and distribution network level are more significant than in the past. For example, the increasing number of electric vehicles (EV) and

distributed generation (DG) penetrations contribute to a massive portion of usage and uncertain electricity generation across the day [73]. The traditional linear approaches (persistence, ARIMA, MLR, etc.) are inferior to ML methods for load forecasting, particularly in DN. With the massive development of computation ability, ML methods have been wildly applied in the forecast method in the past decades.

Machine learning approaches such as Artificial Neural Network (ANN) [74], Support Vector Machine (SVM) [75], Random Forest (RF) [76], Fuzzy Logic Theory [77], optimisation algorithms [78], feed-forward Multilayer Perceptron (FFMLP) [79], Back Propagation (BP) [80], etc. have been utilised in developing STLF methods in recent years. The ML provides advantages for analysing non-linear data, particularly the ANNs [81, 82]. These approaches are appropriate for modelling the complex relationship among multiple variables, especially the non-linear relationships [82].

In general, ML methods such as BP algorithms, RF, Grey Projection Network, etc., lack the consideration of time correlation of time series data [83]. Furthermore, Fuzzy Logic and Genetic Algorithms concentrate on distinct facets of data processing, such as data classification and solution optimisation [84]. Moreover, the RNN is an improved method based on the conception of ANN, using the temporal information of the input data, where connections between units form a directed cycle within the same layer, although the ANN cannot find the relationship between data and time. Therefore, the output of each time step in the RNN is affected by the input data from previous steps [85, 86]. However, the vanishing gradient problem affects the RNN when the input time-series data becomes deep and complex [87, 88]. In some cases, the gradient information during backpropagation progress will be vanishingly small, preventing the weight stored in the network from changing its value [89].

Thus, in solving STLF problems, the Long Short-Term Memory (LSTM) method, based on Recurrent Neural Networks (RNN), is introduced, which provides the ability to analyse non-linear data, taking into account the time correlation of time series data [90, 91, 83]. To ease the vanishing gradient problem, LSTM blocks are introduced into the cell in RNN to remember values for either long or short duration of time [92]. Hidden units in LSTM blocks trap the coming input data depending on the weight at the input and output gate, as shown in Figure 2.5. Therefore, the gradient can be propagated back across several time steps in the backpropagation progress without exploding and vanishing, which helps the LSTM network learn the long-range dependencies of the time series.

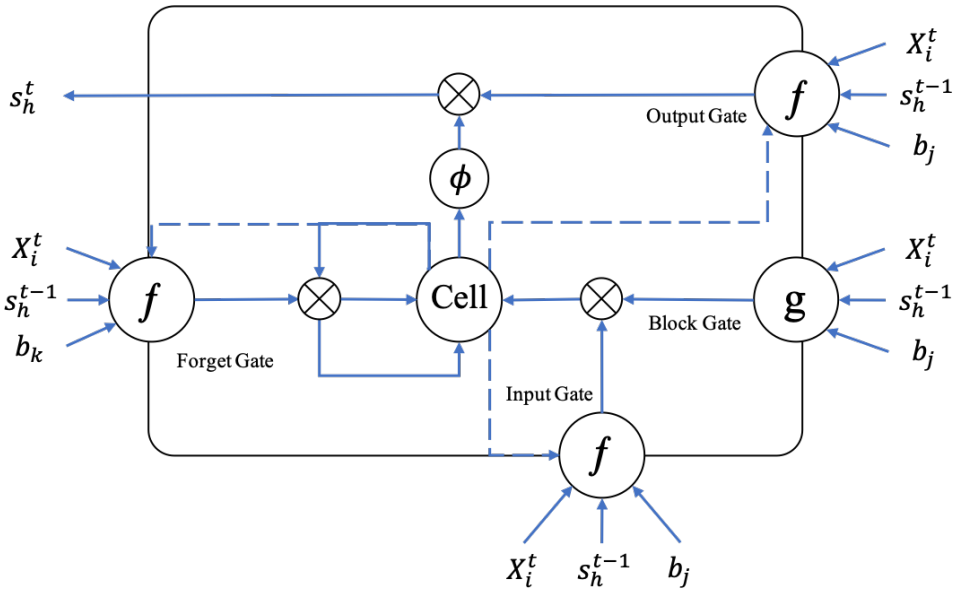


Figure 2.5: The LSTM block

Recent research has demonstrated that a hybrid prediction scheme employing multiple machine learning or statistical approaches outperforms the standard prediction scheme employing a single machine learning algorithm [93]. The hybrid method seeks to deliver the highest possible prediction performance by controlling the strengths and weaknesses of each base method automatically [60, 94, 95, 96, 40, 97, 98]. Kouhi et al. [99] propose a three cascade neural network (CNN) structure to eliminate the

unnecessary and redundant inputs in order to improve forecast error. Fan et al. [100] offer a self-organised map (SOM) and SVM method that can automatically adjust distinct methods simultaneously for ordinary days and anomalous days. Moon et al. [44] suggest an STLF technique employing RF and multilayer perception for small networks.

Variable selection is a crucial aspect in developing STLF methods [101]. Temperature, sunshine duration, relative humidity, cloud cover, and wind speed are common meteorological factors utilised for STLFs [102, 103]. Li et al. [104] reduce the MAPE of the method from 2.4% to 1.4% by incorporating the accumulated influence of high temperatures upon electricity demand. In STLF methods, non-meteorological variables such as time of day and day of the week are also crucial. For instance, according to [105], short-term heat load estimates can be improved by integrating calendar and holiday variables. In [76], the RF method produces a day-ahead forecast with an average MAPE of 2.3%. In [106], the Gradient Boosting method is utilised to produce an hourly load forecast. Finally, in [107], the RF and Gradient Boosting methods are contrasted, while the location and used method influence the forecast outcome. The MAPE ranges between 4% to 7%.

Behind-the-meter installation of solar panels and other DGs are other vital elements that influence electrical demand. In [108], Kaur et al. find that the accuracy of STLFs in California dropped when solar generation capacity was installed. [109] demonstrates that incorporating photovoltaic (PV) variables in STLF methods improve method accuracy. However, prior research has demonstrated that behind-the-meter DGs impact load forecasts and this factor has not been considered in STLF methods until recently. Takeda et al. [101] use estimated solar generation, temperature, solar radiation, hu-

midity and wind speed in STLF methods in Tokyo, Japan. They achieve a MAPE of 1.8% during weekdays. Hasan et al. [110] using an ANN-based method to forecast solar irradiance and then convert it into solar power. They find that this enhances their STLF method forecast accuracy.

2.4.2 Mid and Long Term Load Forecast

MTLF and LTLF play essential roles in maintenance scheduling, fuel reverse planning, unit commitment, energy contracts, load dispatching analysis, revenue from sales, load dispatching coordination, monthly peak load studying, and capacity expansion for electric utilities, Network planning, capital investment, purchase of generating units, purchase of equipment, revenue analysis, and staff hiring, respectively [111, 112]. When making MTLF and LTLF, the method selection and accuracy are influenced by several factors [97, 113, 114, 115, 116]:

- Population: number of electricity consumers, household units, electricity connections, used or provided electric units, etc.
- Weather: temperature, global warming index, rainfall, humidity, etc.
- Economy of the Considered Territory: Gross Domestic Product (GDP), Per Capita Income (PCI), Gross national product (GNP), Gross National Income (GNI), etc.
- Standard of Living: Sales of luxury items including appliances, Technology development, etc.
- Fuel and Electricity Prices: oil price, gas price, petroleum price, electricity price, accessibility to amenities, etc.
- Geographical and Regional Developments
- Government Policies,
- Random Factors

In the time period, MTLF contributes to the possible allocation of available resources and the creation of other infrastructure elements [112]. For example, improving congestion management in transmission grids and total system efficiency, and reducing consumer energy cost [117]. Moreover, the deregulated firms can improve the transmission and distribution systems under the guidance of the LTLF information [48]. Economically, the research about the impact of MTLF is shown in [118, 119]. The precision of MTLF affects energy supply planning, with inadequate supply restraining the economy's growth and overstock resulting in cost overruns that may be passed on to consumers. In addition, articles [120, 121] demonstrate the effects of MTLF on hydro-thermal coordination and cost-efficient fuel purchasing. Nevertheless, the MTLF has not gained much traction because academics have focused on the other two load forecasting horizons, and the MTLF prediction period is wedged between the STLF and LTLF.

When it comes to the LTLF, an electric utility must conduct accurate forecasting for LTLF because the installation of power generation and transmission facilities usually takes years and require substantial investments. The accuracy of LTLF directly has a major impact on future generation and transmission network development. Therefore, the importance of the electricity network planner to predict future circumstances is crucial. In accordance with the anticipated demand, the electric utilities coordinate the resources with a cost-effective plan. In general, the LTLF necessitates the evaluation of a significant number of uncertainties, and in [122], several reasons for the inaccurate forecast are stated:

- Peak demand is quite much dependent on temperature.
- The required meteorological data and economic statistics for LTLF are unavailable.
- Difficulties in electric power storage.

- The period to construct new power plants and transmission networks is years and requires a substantial investment.

Compared with STLF, less work has been completed on LTLF yet. The cause is the same as for MTLF, for example, uncertainty, complexity and difficulty of data collection and processing. In addition, weather and economic and societal variables exacerbate the difficulties of LTLF. In [123, 124], an LSTM method with a Genetic Algorithm (GA) is proposed for LTLF, and it provides more accurate forecast results than statistical methods. Moreover, [125] proposes a method that combines ANN with Fuzzy Logic in which the neural network consists of multilayer sensing and the adaptation of parameters to connect fluctuations in electricity usage with changes in electricity pricing.

In this work, as the values of MTLF and LTLF are revealed mostly in power network expansion planning, infrastructure investment, and other long-term objectives, only STLF is discussed. The load pattern analysis in the mid and long-term window is focused on the aggregated load profile, which is highly sinusoidal. Therefore, the improvement of applying machine learning methods compared with STLF is less.

2.5 Demand-Side Management and Load Forecast in Distribution Network

The main objectives of the smart grid are to improve power system efficiency, reduce the peak-to-average ratio, minimise production costs, and integrate renewable energy sources (RES). The particular objectives can be categorised into six groups [126]:

- Real-time pricing and billing implementation;
- Renewable energy resources integration;

- Accommodation of plug-in hybrid vehicles (PHEV) and plug-in hybrid vehicles (PEV);
- Bi-directional information flow between the utility company and consumer;
- Production cost minimisation;
- Energy production optimisation and greenhouse gas emission reduction.

The demand-side management (DSM) Program is an approach used to indirectly regulate the load profile in order to meet utility objectives. These objectives are [36]:

- Keeping the load factor close to 1.0.
- Maintaining the peak load within the network capacity.

In the achievement of the objectives mentioned above, the utility company obtains the maximum energy from the installed plants, therefore, maximising the overall profit and minimising the average cost. As depicted in Figure 2.6 (redrew from [36]), common load shaping strategies include peak clipping, valley filling, load shifting, strategic conservation, strategic load increase, and variable load shape.

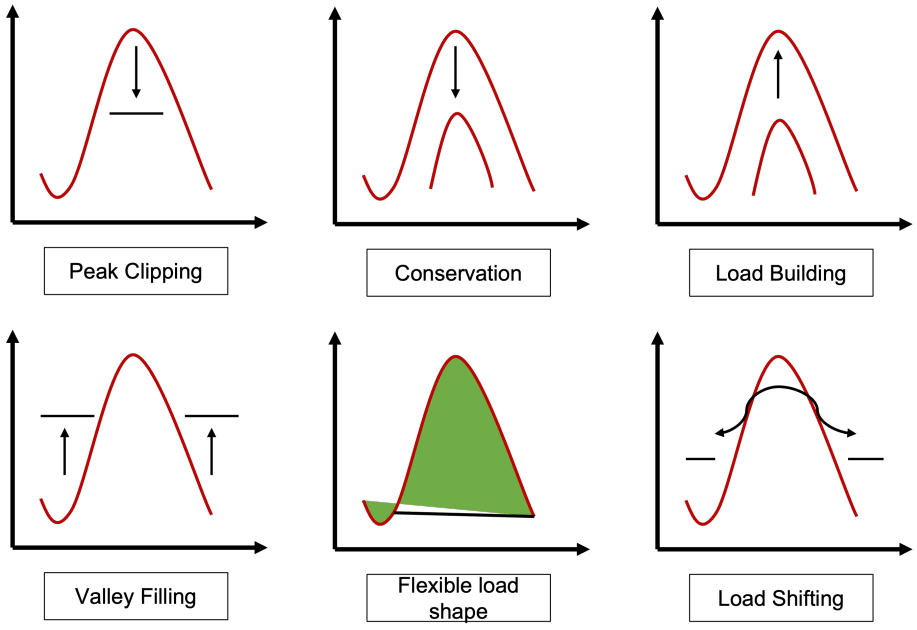


Figure 2.6: Typical demand-side management techniques [36]

Demand response (DR) is the process that the utility company curbs the load or remotely controls appliances with consumers' premises to improve the operation and service quality based on the existing systems and avoid substantial capital investments. DR aims to deal with fuel price spikes, brownouts, blackouts, and other emergency conditions. Consumers are involved in participating in the DR programs via various incentives and penalties [127]. In general, DR programs can be classified into price-based and incentive-based programs. The price-based programs introduce Time-of-Use, Real-Time Pricing, Critical Peak Pricing, etc. The incentive-based programs involve Direct Load Control, Demand Bidding, Interruptable Programs, etc. [128].

Price-Based DR Program

The price-based DR programs and indirect load controls usually encourage consumers to modify their energy consumption patterns at a specific period, such as peak hours, in response to the operator's time-based pricing schemes. The most common schemes include Time-Of-Use pricing (TOU), Critical Peak Pricing (CPP), and Real-Time Pricing (RTP). In [129], the tariff is designed to offer different charges depending on the time in different year seasons or hours of the day. The price during peak hours is generally higher than during off-peak hours. CPP is similar to TOU, which periodically changes prices, especially when the system is overloaded in the summer. In [130], the CPP scheme is notified to consumers a day ahead with proper load forecasting techniques. Further, RTP offers hourly prices, and the participants are notified about the time beforehand. The implementation of RTP requires well-developed smart grid systems to achieve real-time communication among utilities, customers, and control centres [131].

Incentive-Based DR Program

In incentive-based DR programs, consumers receive financial incentives for their load-reduction behaviours. The most common schemes include Direct Load Control (DLC), demand bidding, and interruptible programs. DLC programs control customer appliances remotely by sending signals based on the contract agreed between operators and customers [132]. Demand bidding programs are market-based programs in that customers bid for the load they are willing to offer control availability to the utility companies [133]. Finally, interruptible programs allow customers to shift the unnecessary load to off-peak periods. In some circumstances, the customers are encouraged to shut down during emergencies. Participants receive incentives for their actions or penalties if they fail to respond to the special events [134].

Soft Computing Based DSM

Depending on the machine learning approaches, the widely utilised approaches can be classified as Fuzzy Logic, Artificial Neural Networks, and Evolutionary Computation (EC). The soft computing techniques have advantages in solving complex problems of intelligent building control [135]. First, the Fuzzy Logic is extensively utilised in controlling and monitoring home appliances as this approach outperforms others in dealing with uncertainties and nonlinearities [136]. Second, ANN is one of the machine learning approaches that can consider various factors, therefore making optimisation decisions with pattern and signal predictions[137, 138, 139]. Finally, EC is used to solve complex nonlinear, nonconvex and constrained optimisation problems [131].

Optimisation Based DSM

Game Theory is one of the most commonly utilised and powerful optimisation techniques. In [140], an Autonomous game theory-based DSM is presented. The objectives are to minimise energy costs and peak-to-average ratio. Moreover, for RTP-based DSM, stochastic optimisation techniques are widely used for price minimisation and controlling financial risks [129, 141, 142]. In [143], the DSM program is implemented with Simulated Annealing techniques with white tariffs (an extension of the TOU tariff). Besides, for various DSM program requirements, load scheduling optimisation, multi-objective optimisation, interval number optimisation, etc., are introduced to solve optimisation problems [144, 145, 146, 147].

Load Forecast in DSM

Forecasting electric load is an important component of the electrical power business. It is essential to forecast the future electricity demand as early as feasible. The corporation makes investments and decisions for purchasing energy from generating companies and planning for maintenance and expansion based on the anticipated load. Therefore, it is essential to know future energy usage. Electric power distributors require a tool that allows them to estimate the load in order to facilitate its management and improve the planning formulation. Accurate prediction of electric load is difficult. Predicting the electric load at a future time is a difficult task due to the different properties of the electrical demand and the related uncertainty. A typical daily variation of electric loads is shown in Figure 2.7 (redrew from [36]).

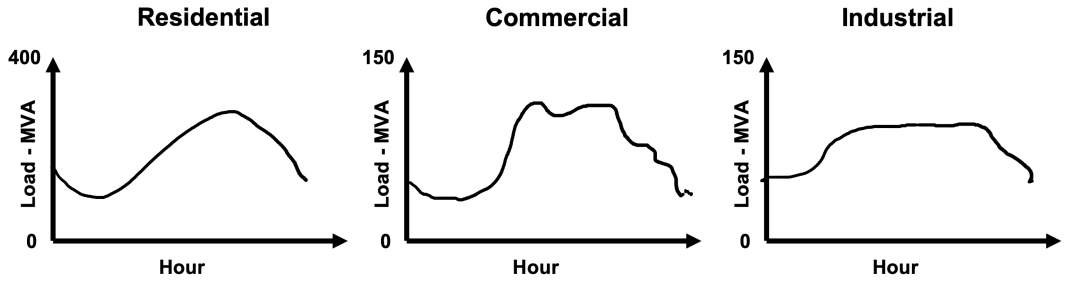


Figure 2.7: Typical daily load curve variations [36]

The features of the electric load are determined by the consumers and the devices such as motors, lighting, heating system, etc. Depending on the end-users and the devices, the electric load can be divided into four categories, including residential, commercial, industrial and agricultural. Load forecast is crucial in SDN for utility companies because these companies offer DSM schemes based on the forecast load rather than the actual load. Moreover, the importance of load forecasting is as follows [36]:

- Purchasing, generation, sales
- Contracts
- DSM
- Area planning
- Infrastructure development/capital expenditure decision making

2.6 Research gaps

According to the comprehensive literature review done in this chapter, the research gaps are noted as follows:

First, based on the review, the demerits of current forecast approaches can be concluded as follow:

- The existing STLF methods utilise sufficient historical data for model training purposes. But, the DSM alternatives, such as Peer-to-Peer (P2P) trading methods, are usually utilised in a small network, where the collectable historical data is insufficient or limited for model training purposes. Moreover, the nonlinearity of the load curve increases further which worse the existing forecast methods and increases the forecast errors. Therefore, the novel developed method should be able to provide accurate results when the training dataset is limited and the nonlinearity of the load increases.
- The trained model only fits the trained historical dataset load patterns, and when the application area changes, the model losses the forecast accuracy. This is caused by the machine learning models being trained based on the specific datasets and these methods lack abilities to learn new features from other datasets without proper training procedures. Therefore, the novel developed method should be adaptive.
- The peak and off-peak load cannot be forecast accurately simultaneously. Although hybrid processes and methods could be combined to ease this problem in existing methods, it increases the model complicity, requires more computational power, and, therefore, narrows the application range. Moreover, the forecast mismatch, especially during peak periods, will cause the extra operational cost for the system operators to increase/decrease backup generations (or extra energy storage devices) to balance the demand. Therefore, the novel forecast method proposed in this work should provide an accurate forecast method for peak and off-peak periods and implement dynamic compensation parameters, making the proposed method adaptive to fit various application purposes.

Second, with many countries releasing the carbon neutral target by different target years, the RES and EVs penetration level, home appliances usage pattern, heating system, etc. will be significantly changed and switch to an environmentally friendly path. The existing forecast methods are validated with the historical dataset, while the load pattern in the future is altered. This will lose the validation convincing for existing forecast methods. Therefore, this research should build up future load scenarios to evaluate the proposed method.

Third, the existing DSM alternatives are developed based on 100% accurate forecast result which deviates from real-world application conditions. Therefore, this research should reveal the value of improving forecast accuracy in deploying the DSM alternatives.

2.7 Conclusion

This chapter presents a general overview of the areas related to the primary research topic of this PhD thesis. First, it summarises the development status of current smart grids and the advantages of applying load forecast in the distribution network. The current power grid is transitioning from passive to active distribution networks due to the accelerated growth of emerging ICT and IMI. Moreover, the increased penetration level of DGs offers more opportunities for operators to improve service quality. With this consideration, an accurate load forecast method will benefit short, mid and long-term power grid management. With state-of-art machine learning techniques, the load accuracy is potentially improved by utilising neuron networks and data collected from the monitoring system. Second, the load forecast methods widely utilised in the power network as reviewed. As methodologies of these approaches are frequently referred to for novel method development, it is worth investigating the value of ex-

isting methods. In this section, both statistical and machine learning methods are reviewed. Typical statistical methods are as follows: multiple linear regression, exponential smoothing, ARIMA and its variants, Kalman filter, etc. Typical machine learning methods can be concluded as follow: support vector regression, fuzzy logic, artificial neural network, random forest, etc. According to recent research, it has been found that hybrid learning methods offer better forecast ability as they compensate for the disadvantages of a single method. Third, applications of machine learning techniques in various load forecast steps are reviewed, including short-term, mid-term and long-term load forecast methods. Referenced forecast approaches can be classified into statistical and machine learning methods, and the positive and negative aspects of these methods are discussed. Finally, the demand-side management strategies are reviewed, including widely utilised optimisation techniques. Also, the importance of load forecast in demand-side management is reviewed.

Chapter 3

Development of Short-Term Load Forecast Method for Distribution Network

3.1 Introduction

Electric load forecast is fundamental in smart grids as it can help suppliers to model and forecast load in advance, balance the demand and supply, adjust demand response plans, implement real-time pricing schemes, etc. In this regard, the accuracy and robustness load forecast method is important. The gap between electricity supply and demand can be minimised with accurate forecast methods.

In a distribution network, the relationship between consumers, suppliers, and distribution system operators (DSO) becomes complex. The consumers become more active in interacting with the electricity market, such as selling electricity to the open market [148]. The situation is also challenging and beneficial for DSO. the DSO provides advanced management methods, for example, the state estimator, the self-healing function, etc., with accurate load forecast methods[149]. On the other hand, the development of the smart grid has created massive real-time and historical data, which is collected from the monitoring devices, such as smart meters, Phasor Meas-

urement Units (PMUs), and other user behaviour data [150]. By taking advantage of Machine Learning (ML), the short-term load forecast methods are expected to increase forecast accuracy and method robustness upon modern load patterns and massive load data.

In the modern smart grid, with the big data generated from the smart grid, the Wide Area Monitoring System provides chances to change the way of energy production and the pattern of energy consumption. We could efficiently analyse and mine the energy big data to support more effective and efficient decision makings, prevent risks, protect privacy, etc. As the electricity load forecast accuracy is crucial in providing better cost-effective risk management plans, the STLF method with high forecasting accuracy is introduced [151].

This chapter will fully illuminate the short-term load forecast method based on the ML algorithm. While the existing top-down load forecast method requires months or years of historical data and cannot simultaneously forecast peak and off-peak load accurately, the key contribution of the proposed method is the accurate forecast result for both peak and off-peak periods at the DN when the historical data and data varieties are insufficient. The novel developed method introduces dynamic error compensation abilities to the existing LSTM network, which helps the forecast method reduce errors in peak and off-peak load forecasting. Moreover, the forecast method is featured in utilising high-time-frequency data with state-of-art machine learning methods to generate the load forecast method for the distribution network. Besides, the developed method uses multi-variables, such as load, temperature, humidity, etc., as the input, coming with the correlation selection function to improve forecast accuracy even further. The surplus of the developed method provides the ability to apply load forecast in small communities with fewer households, where the historical dataset is insufficient.

This chapter demonstrates detailed development procedures of the DAC-LSTM method, including data processing and analysis, the dynamic adaptive compensation module, and sensitivity analysis.

3.2 Long Short-Term Memory Network

Recurrent Neural Network

The traditional ANN maps the input historical data and output forecast value but lacks consideration of time correlation in the data sequence. This causes the problem that the ANN cannot find the relationship between data and time. An RNN is an improved ANN class using the input data's temporal information, where connections between units form a directed cycle within the same layer [85, 152]. Therefore, the output of each time step is affected by the input data from previous steps.

However, the vanishing gradient problem affects the original RNN when the input time series data becomes deep and complex. The 'weight' contributed by the former data 'vanishes' as the input time series data becomes longer. As a result, the perception of later nodes from the previous time step decreases, which prevents the 'weight' stored in cells from being updated in each epoch. The vanishing gradient problem is worse when the activation function is sigmoid, while Gated Neural Networks (GNNs) such as Long-short Term Memory and Gated Recurrent Unit (GRU) provide promising results in time series learning tasks by introducing sophisticated network structures [153].

The RNN structure is shown in Figure 3.1.

The symbols in the figure above are explained as follows. X is the input unit. H is the hidden unit. Y is the output unit. W is the weight matrix.

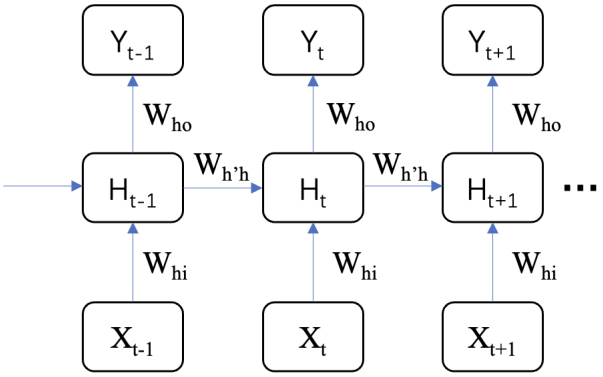


Figure 3.1: The RNN network structure

Besides, the forward propagation can be expressed as follow:

$$a_h^t = \sum_{i=1}^{I} w_{hi} x_i^t + \sum_{h=1}^{H} w_{h'h} s_{h'}^{t-1} \tag{3.1}$$

$$s_h^t = f_h(a_h^t) \tag{3.2}$$

$$a_o^t = \sum_{h=1}^{H} w_{ho} s_h^{t-1} \tag{3.3}$$

The symbols in the equations above are explained as follows. w is the weight. a is the sum weight. f is the activation function. s is the value after passing through the activation function. t is represent the current time. i is the input vector number. h is the hidden vector number at time t. h' is the hidden vector number at time $t-1$. o is the output vector number.

Long Short-Term Memory Network

To solve or ease the vanishing gradient problem, an LSTM block is introduced into the RNN, remembering the values of either long or short duration of time for different cases [92]. To be more specific, the hidden units in the RNN are replaced by the LSTM blocks. Each LSTM block contains three extra gates, which are used to control the information flows into, out of, or trapped in the cell.

The structure of LSTM is shown in Figure 3.2 (redrew from [154, 155]).

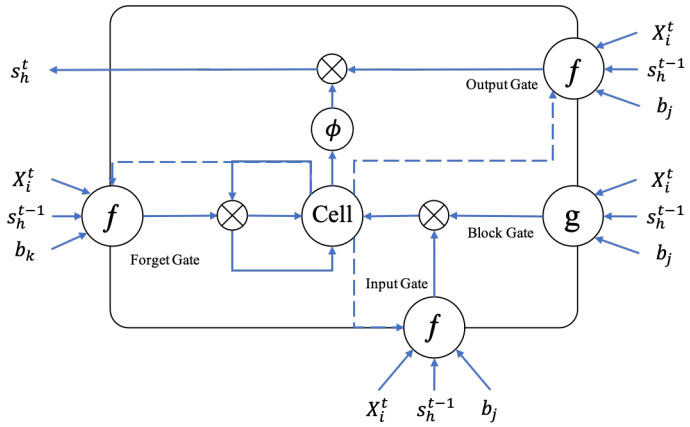


Figure 3.2: The LSTM cell in an RNN network

$$a_j^t = \sum_{i=1}^{I} w_{ij} x_i^t + \sum_{h=1}^{H} w_{hj} s_h^{t-1} + \sum_{c=1}^{C} w_{cj} m_c^{t-1} + b_j \tag{3.4}$$

$$s_j^t = f(a_j^t) \tag{3.5}$$

$$a_k^t = \sum_{i=1}^{I} w_{ik} x_i^t + \sum_{h=1}^{H} w_{hk} s_h^{t-1} + \sum_{c=1}^{C} w_{ck} m_c^{t-1} + b_k \tag{3.6}$$

$$s_k^t = f(a_k^t) \tag{3.7}$$

$$a_c^t = \sum_{i=1}^{I} w_{ic} x_i^t + \sum_{h=1}^{H} w_{hc} s_h^{t-1} + b_c \tag{3.8}$$

$$d_c^t = s_k^t m_c^{t-1} + s_j^t g(a_c^t) \tag{3.9}$$

$$a_l^t = \sum_{i=1}^{I} w_{il} x_i^t + \sum_{h=1}^{H} w_{hl} s_h^{t-1} + \sum_{c=1}^{C} w_{cl} m_c^t + b_l \tag{3.10}$$

$$s_l^t = f(a_l^t) \tag{3.11}$$

$$s_h^t = s_l^t \phi(d_c^t) \tag{3.12}$$

The symbols in the equations above are explained as follows. X is the input unit. w is the weight. a is the sum weight. h is the hidden vector number. s is the value after passing through the activation function. m is the input from the cell to the input gate. i is the input vector number. j, k and l are the input, forget and output gate vector numbers. c is the cell vector number. d is the cell value. f, g, ϕ are activation functions.

Depending on the weight at the input and output gates, the LSTM blocks can let errors into, out of or trapped in the block. The coming input data is trapped in the cell if both input and output gates are closed. As a result, the value remains unchanged, and the impact on the output of the current time step is prohibited. Therefore, when the backpropagation progress occurs, the gradient can be propagated back across several time steps without exploding and vanishing. This memory block allows the LSTM network to learn the long-range dependencies of time series compared with the traditional RNN. Moreover, by taking advantage of using RNN, the input matrix towards the LSTM block can be multi-variables, multi-steps ahead, and the output can be multi-steps after. Besides, the seasonal influences are minimised.

3.3 Development of Dynamic Adaptive Compensation

This section presents a detailed methodology for developing an STLF forecast method. It includes a data processing procedure, peak detection method, and LSTM method developing with an adaptive weight correction procedure. In addition, the novel method introduces dynamic error compensation abilities to the existing LSTM network, which helps the forecast method reduce errors in peak and off-peak load forecasting.

The block diagram is shown in Figure 3.3.

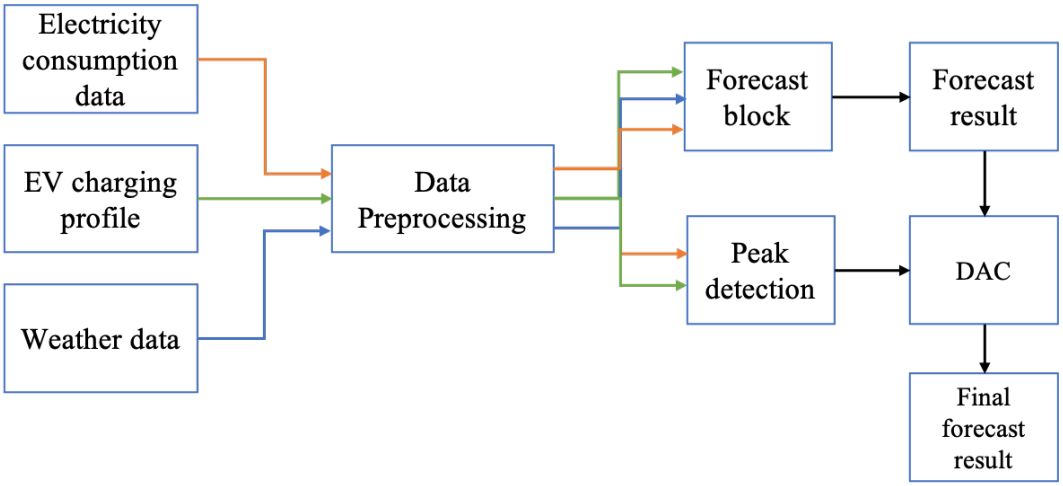


Figure 3.3: Block diagram of the DAC-LSTM scheme

As Figure 3.3 illustrates, the raw data is first fed into the data processing module to generate the required data format. Then by following the path with different colours, consumption, EV, and weather data are passed to the LSTM forecast module. As LSTM blocks cannot accurately forecast both peak and off-peak load, the peak load detection module is added to distinguish the peak and off-peak load dynamically. Different parameter sets will be applied to peak and off-peak loads separately afterwards. A more detailed explanation can be found in 3.3.5. Finally, according to the forecast error from previous steps, the dynamic adaptive compensation (DAC) module is added to improve forecast results for peak and off-peak periods.

Because the DAC module provides different compensation parameters for peak and off-peak hours dynamically, the forecast error is reduced dynamically according to the forecast quality regardless of the forecast methods. The DAC-LSTM method solves problems existing in most other methods, that, first, the forecast accuracy cannot be maintained in peak and off-peak hours simultaneously, and, second, the model performance (forecast accuracy) reduces when the training datasets are insufficient.

Figure 3.4 shows the overall forecast network structure. The forecast network is built based on the LSTM network, consisting of several layers (Figure 3.4 left). Each layer is structured as an RNN network. The hidden units in each layer are replaced by the DAC-LSTM unit (Figure 3.4 middle). The peephole DAC unit is shown on the right of Figure 3.4. The detailed explanation of the DAC-LSTM unit is presented in a later section 3.3.7.

3.3.1 ML Problem Definition

This research aims to develop short-term load forecast methods using ML techniques. As the existing forecast methods cannot forecast peak and off-peak periods simultaneously and the training dataset is assumed to be sufficient, this work intends to develop a forecasting method based on the LSTM network, capable of providing accurate peak and off-peak load forecast results simultaneously and dealing with future smaller distribution network applications. The developed method intends to utilise multiple variants as inputs and high-time resolution datasets to improve forecast accuracy.

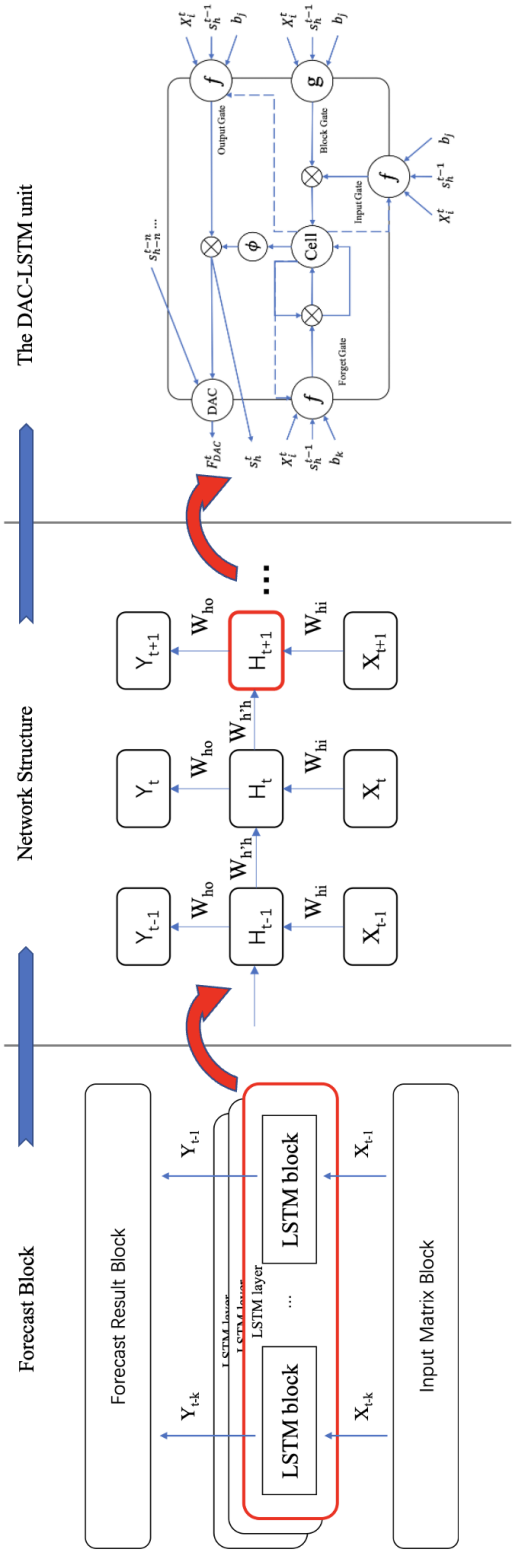


Figure 3.4: DAC-LSTM structure: From the network to cells

3.3.2 Data Description

Three primary datasets are utilised: electrical load, ambient condition, and EV charging profile.

The forecast method is developed by using the following datasets. The real-world historical electrical load data is obtained from the Thames Valley Vision Project (TVVP) [156]. The data is half-hourly recorded, comes from 220 domestic properties, starts in February and March 2013 and ends in November 2014. Also, the half-hourly weather data in the Thames Valley area was collected from the local weather station from January 2013 to December 2014, including 16 features described in the Correlation Analysis Section. The weather data shows the general weather condition in this area instead of the specific property. The EV charging profiles is from Adaptive Charging Network (ACN), including the charging profiles from charging stations[157]. We assume the US EV users are similar to those in the UK. According to the EV charging data source, the dataset is made possible by close collaboration with PowerFlex Systems around the US. These two datasets (datasets from TVVP and ACN) are combined to emulate the real load pattern scenarios with high EV penetrations. The detailed explanation can be found in section 3.3.3.

A one-week snapshot of TVVP and ACN datasets is shown in Figure 3.5. It can be seen that the EV charging loads mainly contribute to the peak hours, constituting about 10-20% of the sum load. Therefore, EV charging increases the Peak-to-Average ratio and brings more challenges for existing forecast methods obtains accurate result at peak and off-peak period simultaneously.

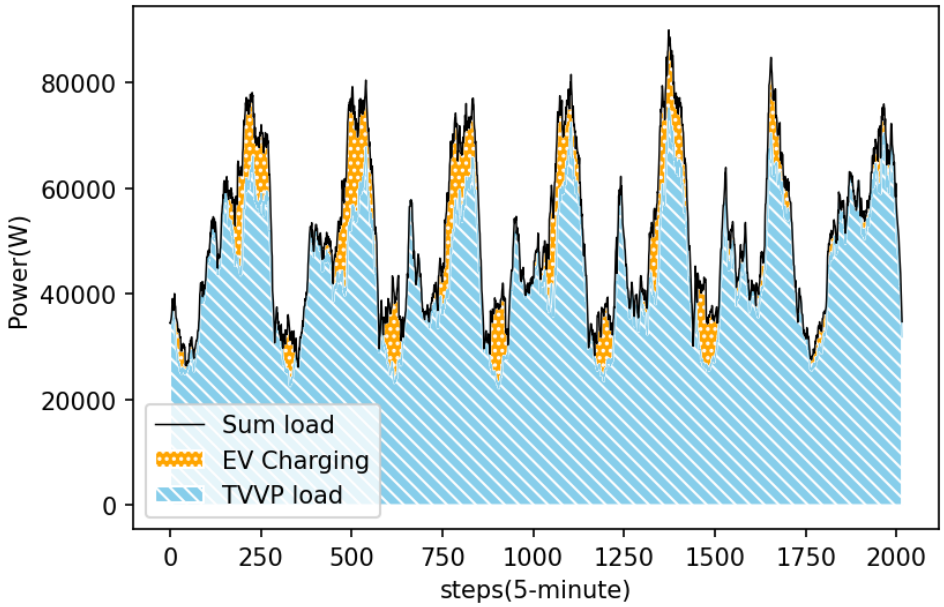


Figure 3.5: Snapshot of TVVP and ACN datasets for 7 days

3.3.3 Data Input

The electrical load data from TVVP is pre-processed to remove the bad point, which may be caused by the delay of recording devices installation, unrecorded days, etc. As shown in Figure 3.6, the spark at the right (1 P.U.) represents the wrongly recorded data or fault data, therefore removed. Moreover, to simulate the high-resolution big data collected from modern smart devices (like smart meters and PMUs), the load data is expanded using linear interpolation and -1.5% to 1.5% random error was added to reduce the time interval from 30 to 5 minutes. The random error is used to simulate the real data recording pattern in the real world. Afterwards, the individual data from 220 households are aggregated to represent the total load. Finally, after the processing, the usable dataset contains 550-day data. Also, the 5-minute interval data is smoothed by passing a moving average block. This helps the methods to learn the load pattern and keep the advantages of using high-time resolution data.

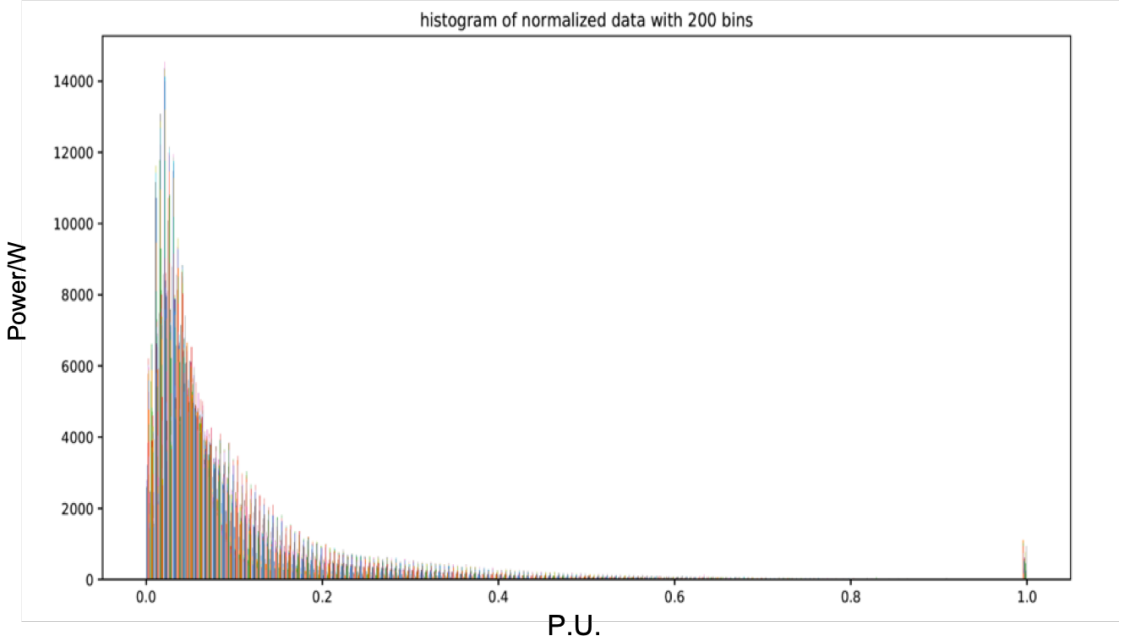


Figure 3.6: Histogram of normalised data with 200 bins (data recorded under 200 is removed)

Then, the EV charging profile from ACN is cleaned by removing the spikes and the wrongly recorded samples. The numerical relationship between ACN and TVV dataset is defined by the following steps below:

1. According to the report [158], the daily charging profile per EV is calculated, peaking at 0.9 kW and 0.37 kW on average (across 24 hours).
2. From TVV, 220 households with 25% EV penetration yield 20.4 kW on average.
3. The dataset from ACN gives an average EV charging power of 53.8 kW. Therefore, the ACN dataset is multiplied by 0.38 and then combined with the TVV dataset.

The input dataset is pre-processed to clean the bad point and be expanded to a 5-min time resolution. Then the following steps are taken as the data processing progress.

3.3.4 Correlation Analysis

This section presents the process of selecting features related to load forecast. The used feature data is recorded from the same area as the TVV Project and within the same period. The collected data includes 16 features, while only 6 of them show a potential relationship with the load selected. Besides, the weather condition feature is simplified into sunny and not sunny. Moreover, 1 extra feature is added as wkornot, which represents weekday or weekend.

Input Features

The input weather data includes 6 features as shown in Table 1.1: load, tempC, weatherCode, humidity, cloudcover, wkorNot.

Table 3.1: Input features and its represents

Term	Represent
Load	Load (W)
tempC	Temperature (°C)
Humidity	Humidity
WeatherCode	Current weather
Cloudcover	Cloud cover rate (%)
wkorNot	weekday or weekend

Correlation Coefficient

The correlation analysis introduces the Pearson Correlation Analysis as it illustrates the correlation between time series data [159]. By applying the equation 3.13, the following Table 3.2 is derived, where n is the number of samples:

$$r_{xy} = \frac{n(\sum x_i y_i - \sum x_i \sum y_i)}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}} \tag{3.13}$$

Table 3.2: Numerical feature correlation coefficient

Index	Load	tempC	Humidity
Coefficient	1	-0.154	0.022
Index	Cloudcover	wkornot	weatherCode
Coefficient	-0.026	-0.02	-0.116

Table 3.2 shows that the load value has a positive correlation with humidity and a negative correlation with temperature. While the weatherCode influences weather, humidity, etc., the weatherCode feature is excluded. In addition, the coefficient of Load represents the self-correlation, which yields 1. In Figure 3.7, 3.8, the plot shows the trend of the load and temperature, which is a repetitively negative correlation.

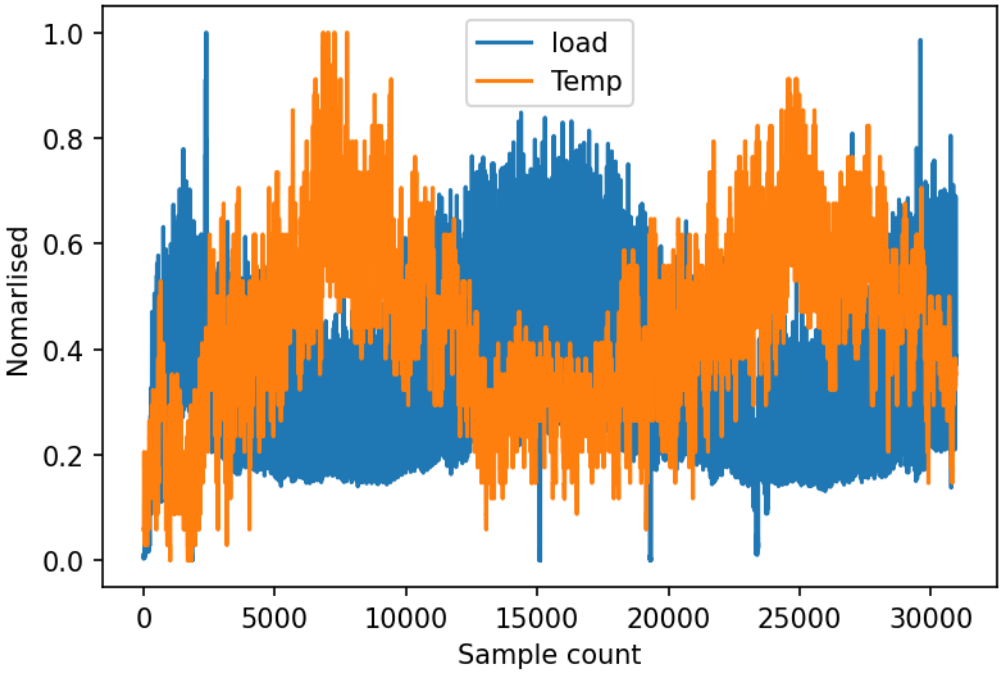


Figure 3.7: Load and temperature profiles in every 5 minutes

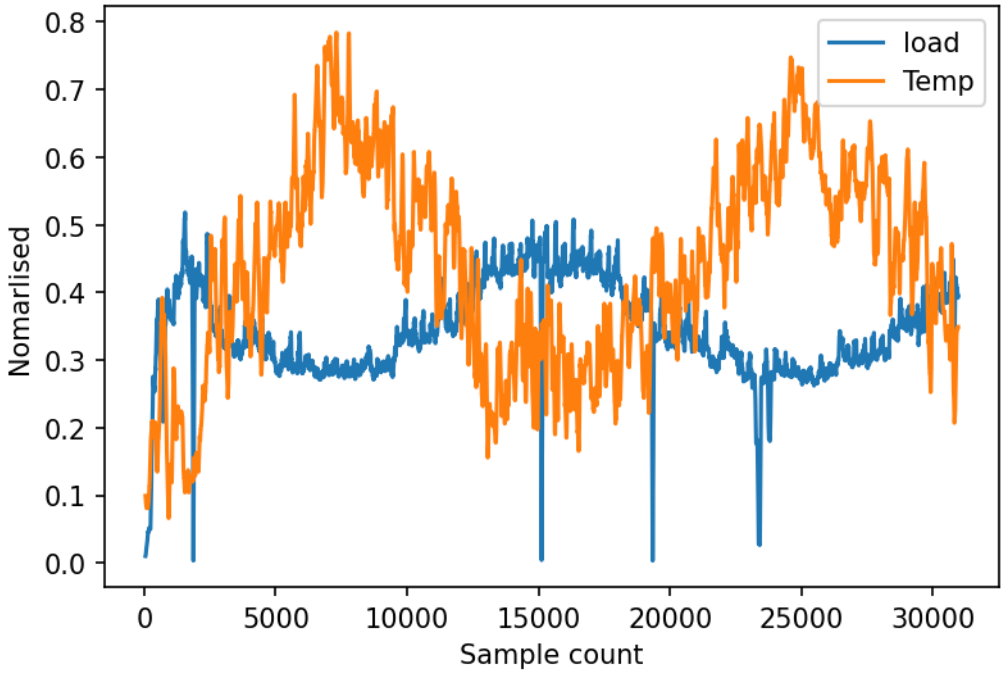


Figure 3.8: Load and temperature profiles every 24 hours

3.3.5 Peak Detection

The LSTM forecast method cannot accurately forecast both peak and off-peak hours simultaneously. Besides, when the peak comes, the forecast result may significantly exceed or exceed the actual peak value due to the forecast hysteresis. Figure 3.9 shows the hysteresis of the forecast method and the large forecast error at the peak (compared with the off-peak). Therefore, the peak detection method is applied to distinguish peak and off-peak load before correcting the error upon the complete forecast result. In the later section, methods are used to improve the average forecast error and minimise the forecast error during the peak.

In figure 3.9, the forecast result is obtained from the LSTM method trained by TVVP and ACN datasets. The forecast step is 30 minutes. The general forecast accuracy is 7.92% (R2) and 0.956 (MAPE). The accuracies during off-peak hours are 7.73% (R2) and 0.944 (MAPE), while during peak-hour are 9.17% (R2) and 0.895 (MAPE).

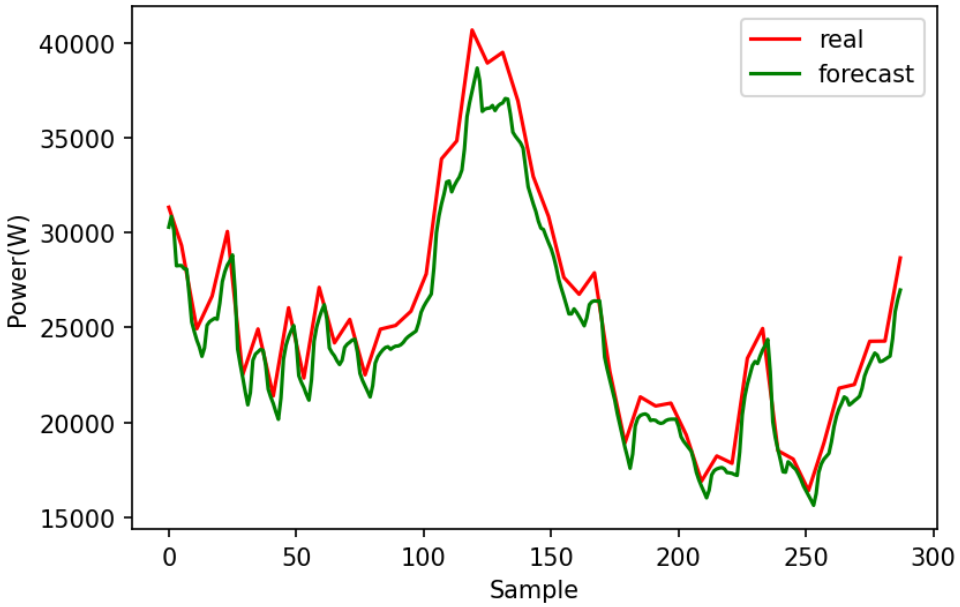


Figure 3.9: Typical load and half-hourly LSTM forecast result for one day

This demonstrates that although the forecast mismatch during peak hour slightly affects the general accuracy (reduced by 0.19%, measured in R2), the forecast error still significantly (increased by 1.25%, measured in R2), which will cause the extra operational cost for the DNOs.

The Z-score (also called the standard score) is based on the dispersion principle. The algorithm is robust as the input time-series does not corrupt the threshold due to the separated moving mean and deviation modules. The coming signals are determined with similar accuracy, which is not affected by the previous signals [160]. According to the Z-score method, the time series will be regarded as peak only when:

$$|y - avg| > threshold * std \tag{3.14}$$

According to equation 3.14, three values will be generated afterwards, including 1, -1 and 0. If the absolute values of the difference are greater than $threshold * std$, a positive (for $y > avg$) or negative (for $y < avg$) signal will be generated, which represents the peak (1) and valley (-1) hours. In this research, only peak (1) and off-peak (0) signals are considered, while the valley hours (-1) are considered as off-peak (0).

Figure 3.10 shows the typical peak load detection result. The positive and negative signal represents the 'peak' and 'bottom' load, while the zeros represent the off-peak load. Depending on the purpose of the utilisation, the sensitivity of the method is adjustable. The Z-score method requires the following parameters:

- lag: the lag of the moving window.
- threshold: the value that the datapoint is away from the moving mean when the signal will be noted.
- influence: the influence (between 0 and 1) of new signals on the mean and standard deviation.

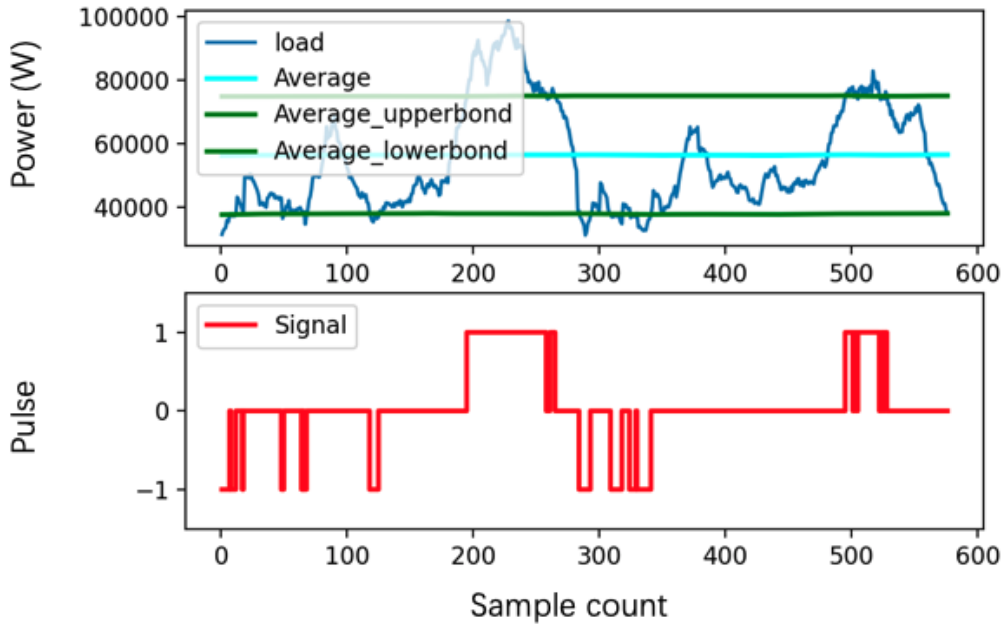


Figure 3.10: Peak load detection using z-score method

For example, when the lag is 10, the last 10 signals will be smoothed. When the threshold is 5, the signal will be noted when a datapoint is 5 standard deviations away from the moving mean. Last, when the influence is 0.5, half of the normal datapoints' influence. Similarly, an influence of 0 gives that all signals will be ignored for recalculating the new threshold, and an influence of 1 represents the threshold remains stationary. Therefore, the most robust mode is achieved when the influence is 0 while the influence of 1 is the least robust. In addition, the influence should remain between 0 and 1 for non-stationary data, which the load profile mostly represents. In this research, the threshold value is defined manually by observing the performance of the Z-score method. The configuration of the derived Figure 3.10 is as follows: lag is 288, the threshold is 1, and influence is 0.8.

3.3.6 LSTM network Development

Python is used for implementing the algorithms along with Keras. The LSTM network is developed using Spyder IDE and Tensorflow as the backend. The overall structure is shown in Figure 3.11. In this Figure, the input matrix contains the preprocessed data, and the three layers in the mid represent the network containing three hidden layers. The following steps are taken in the algorithms:

1. Data normalisation: the data is normalised within [0,1]
2. Frame the data as supervised learning
3. Split the sample into two sets: $train : test = 2 : 1$
4. Reframe the train and test set into 3-dimension: (x, y, z), where x represents the length of datasets, y represents the length of forecast steps, z represents the amount of features.

5. Network design: the network has 3 layers, numbers of cells in each layer are c, $c/2$ and $c/3$. The recurrent activation function is 'sigmoid'. The dropout is set to 0.1. The Dense is d, which equals to the number of forecast steps.
6. The model is compiled with the following settings: the loss is measured by 'MSE', Metrics is using 'MSE', and the optimiser is using RMSprop. The other parameters are default.
7. The model is fitted with the following settings: the maximum number of epochs is 1000, and the batch size is 1152 (selected according to the maximum system memory size). The batch size depends on the computational ability of the computer. Therefore, it is not fixed.
8. When the model is trained, the next step is the prediction
9. Invert the data to the real value from the normalised
10. Calculate the accuracy and generate graphs

In addition, the network parameters are defined based on the following consideration. First, the sample size ratio between the training and test dataset follows the most common ratio utilised in machine learning model training, which is 2:1. Second, the network contains 3 hidden layers as the size of the training sample is 1.4GB, which required a relatively simple network. Therefore, the number of hidden layers is defined as 3. Third, the cell in each hidden layer follows a dropping trend. In this research, the number of cells in each hidden layer is defined as c, $c/2$ and $c/3$.

In the steps above, x is the number of rows of the train or test set, y is the time steps used for prediction, z is the number of features, c is the calculated by equation 3.15, d is the number of forecast steps, MSE represents Mean Squared Error, RMSprop represents Root Mean Squared Propagation.

$$c = |\frac{R_{train}}{2(C_{train} + H_{train})}| \tag{3.15}$$

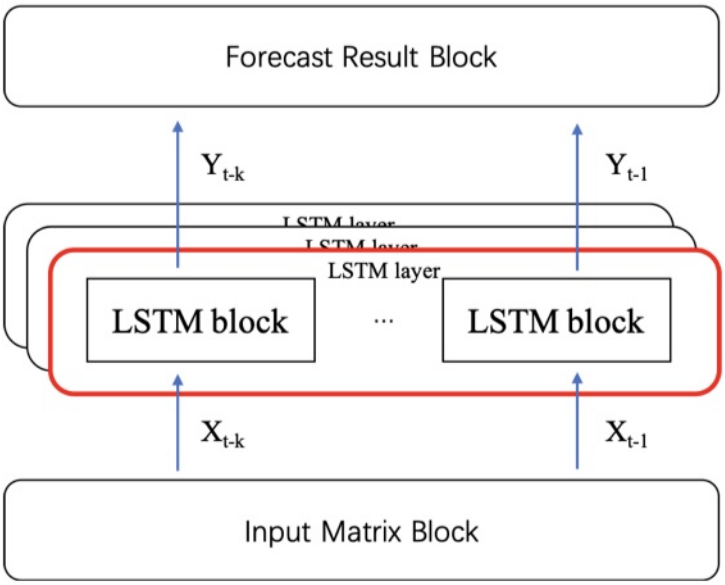


Figure 3.11: An LSTM network with multiple layers

Where R is the number of rows, C is the number of columns, and H is the number of heights.

Equation 3.15 is derived according to LSTM network development guidance provided in [161]. There has no fixed equation to define cell numbers in each hidden layer. In this work, the cell number is calculated based on the shape of the input dataset. The maximum c is the length(R_{train}) of input, the divisor is defined as the sum of width and depth (C_{train} and H_{train}) of the input. Moreover, as the complicity of the network is limited (3 hidden layers), the the divisor is multiplied by two.

3.3.7 Dynamic Adaptive Compensation

This section describes the Dynamic Adaptive Compensation (DAC) methods, which are used to improve the forecast result dynamically. As Figure 3.12 shows, the DAC is an add-on function to the LSTM block. The overall block diagram is shown in Figure 3.13. The pre-forecast value is compared with the real load value, generating an error. According to the forecast step of the method and the rolling forecast procedure, calculated errors are passed through a moving average error block which gives the average error value. Besides, the bias β_{bias} is generated, which gives a dynamic-adjust bias value. Also, the real-time forecast value passes through the Activation function, in this research sigmoid, to generate a parameter $\alpha_{sigmoid}$. The parameter $\alpha_{sigmoid}$ is fixed according to different forecast steps. Finally, Equation 3.16 is introduced to give the final forecast result.

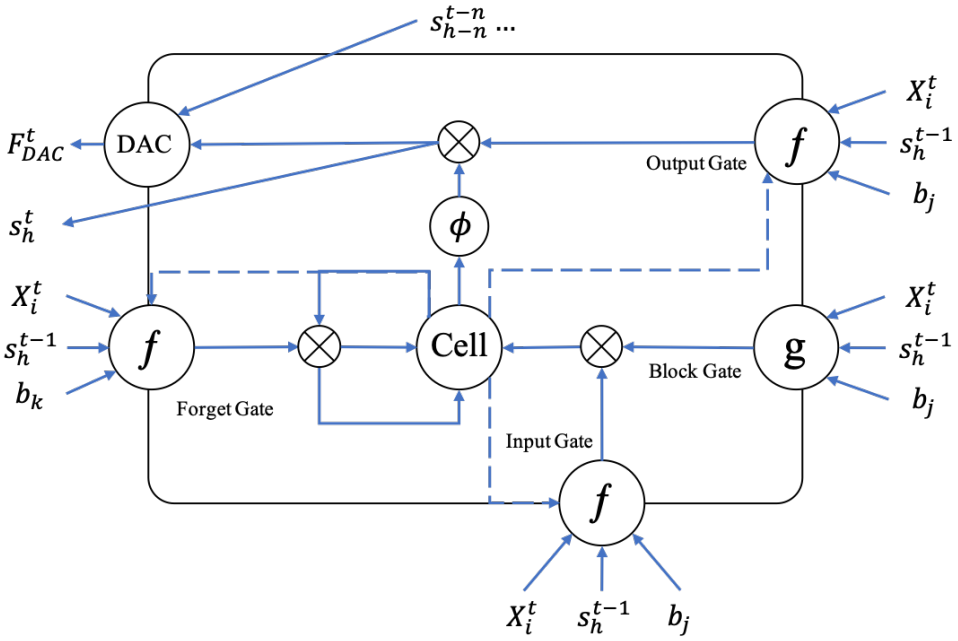


Figure 3.12: The DAC-LSTM unit

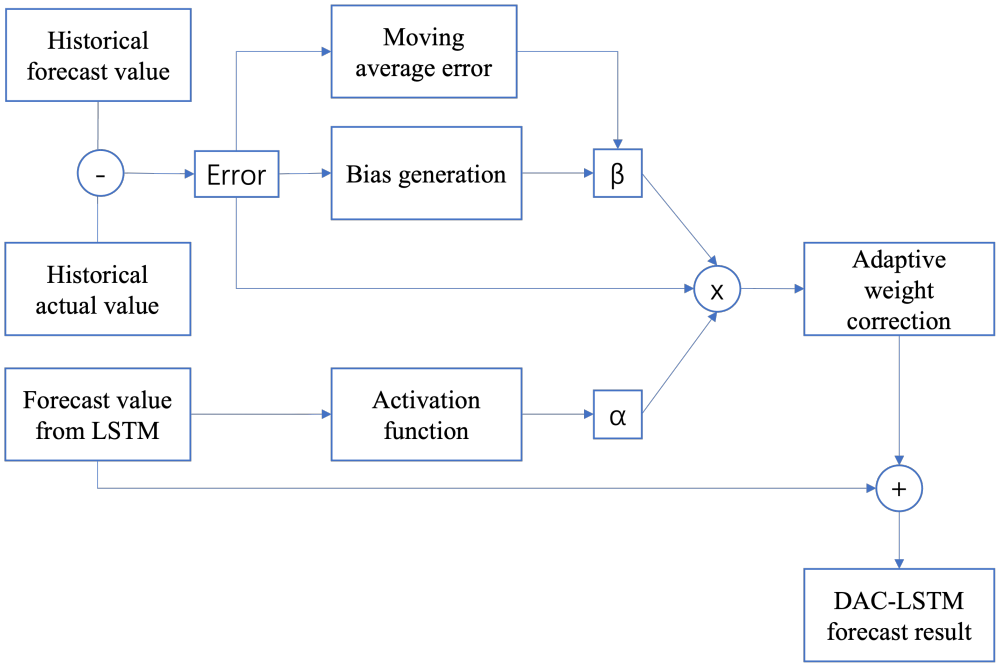


Figure 3.13: Block diagram of adaptive weight forecast method

This procedure can be expressed as the following equations:

$$F_{DAC}^{t+i} = \alpha_{sigmoid}^{i} \beta_{bias}^{i} E^{t} + S_{h}^{t+i} \tag{3.16}$$

In equation 3.16, the present time is t and the forecast step is i. F_{DAC} is the forecast value from DAC-LSTM, S_h^{t+i} is the forecast result from original LSTM block at i^{th} forecast step. The DAC-LSTM forecast result is calculated as the sum of the forecast result from the LSTM block and the adjusted error calculated by 3.17 and coefficients.

$$E^{t} = S_{h}^{t} - R^{t} \tag{3.17}$$

In equation 3.17, E^t is the forecast error at time t, S_h^t is the best historical forecast result at current time t, R^t is the recorded actual load data. Equation 3.17 illustrates that the forecast error used for compensation is the difference between the best forecast result (from LSTM block, uncompensated by DAC) from previous time steps and the actual load at time t.

$$\alpha_{sigmoid}^{i} = f_{\alpha}(i) \tag{3.18}$$

In equation 3.18, $\alpha_{sigmoid}^{i}$ is the fixed forecast step parameter. $f_{\alpha}(i)$ is derived from sigmoid function. For each forecast step i, the value of $\alpha_{sigmoid}^{i}$ is constant.

$$\beta_{bias}^{i} = \frac{E^t}{E_{avg}} f_{\beta}(i) \tag{3.19}$$

In equation 3.19, β_{bias}^{i} is the adaptive error parameter. E_{avg} is the moving average error calculated from several preceding E^t. $f_{\beta}(i)$ is a dynamic changing parameter. This equation illustrates that the compensated error is adjustable according to previous errors. If the forecast error shows an increasing trend, the compensation increases, and vice versa.

The equation of $f_{\alpha}(i)$ is described as follow and plotted in Figure 3.14:

$$f_{\alpha}(i) = 1.5 - \frac{1}{1 + e^{(-i+1)/2}} \tag{3.20}$$

It can be seen that the parameter for the first forecast step is 1, reducing with each later step. $f_{\alpha}(i)$ aims to alleviate the effects of forecast error to the later step because the errors may reduce. This function maintains the parameter for the first step is 1 and the farthest step is above 0.5.

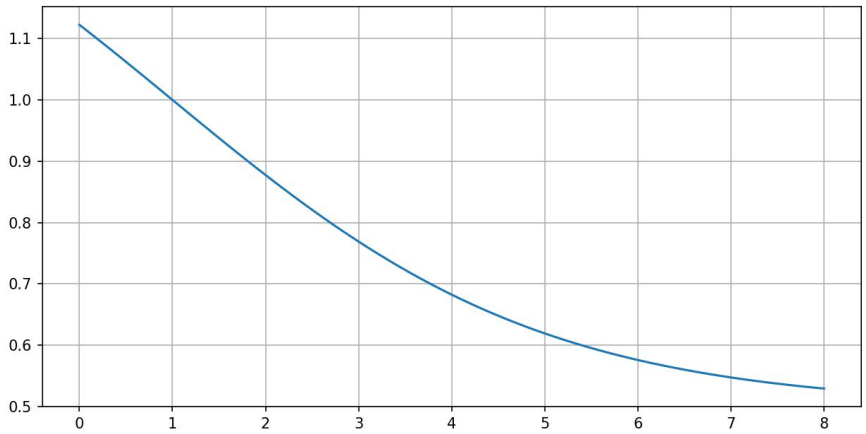


Figure 3.14: The plot of $f_\alpha(i)$

The equation of $f_\beta(i)$ is described as follow:

$$f_\beta(i) = k\delta(\frac{1-i}{m}+1)+1 \tag{3.21}$$

$$\delta(\frac{E_i^t}{E_{avg}}) = \frac{2}{5+5e^{n(1-\frac{E_i^t}{E_{avg}})}} - \frac{1}{5} \tag{3.22}$$

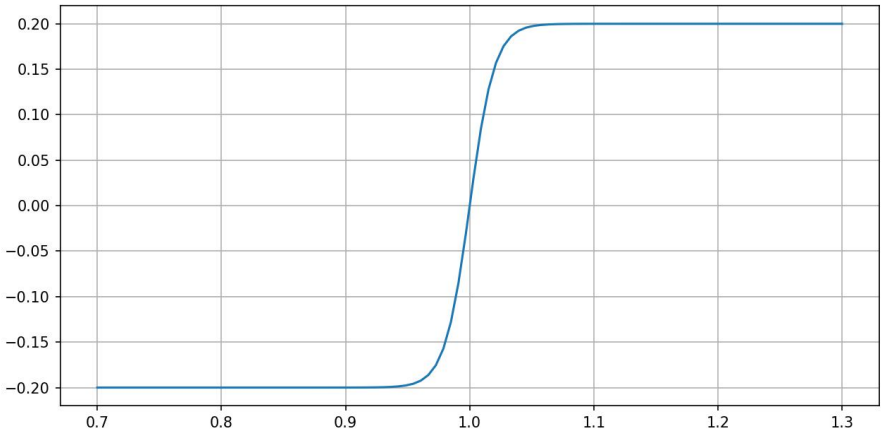


Figure 3.15: The plot of δ, with $n = 100$, x-axis: $\frac{E_t}{E_{avg}}$

In equation 3.21, $k\delta$ is the compensation factor, determining the upper and lower bound of the compensated P.U. value. k is sensitive to the error trend, while δ is sensitive to the forecast step.

First, k is explained as follows. If the E^t keeps increasing compared with previous steps, k increases with a step of Δ and vice-versa. The default value of k is 1 and k is reset to default when $\frac{E_i^t}{E_{avg}}$ changes the positive and negative.

Second, δ is explained as follows. The function of δ is derived from the commonly used activation function in the LSTM network, the sigmoid function. This function intends to limit the range of the compensation, which is $\pm$0.2. If $\frac{E_i^t}{E_{avg}} > 1$, the compensation will be positive and vice versa. If $\frac{E_i^t}{E_{avg}} = 1$, the compensation is cancelled.

Moreover, m and n are constant, which are determined by the test dataset. m controls the sensitivity between forecast steps and compensation factor $k\delta$, and $m \geq (i_{max} - 1)$. This indicate that, for example, if $m = 5$ (given that maximum forecast step is 6), $f_\beta(6) = 1$. Therefore, the compensation is disabled. n controls the sensitivity between error changes ($\frac{E_i^t}{E_{avg}}$) and compensation limit. For example, when ($\frac{E_i^t}{E_{avg}}$) remains unchanged, 3.22 with higher n yields larger $|\delta|$, which means more compensated error.

Thus, when looking back to equation 3.21, the value of $f_{\beta(i)}$ depends on forecast step i and error trend $\frac{E_i^t}{E_{avg}}$. The bound of $f_{\beta(i)}$ is 80-120% with slightly changes caused by Δ and $\frac{1-i}{m}$.

Besides, in this procedure, the compensated forecast result may exceed or is less than the actual value significantly when the error becomes extremely large. Therefore, to improve the forecast accuracy during the peak, the peak detection and compensation cap methods are used to avoid over or under-compensation. The result with or without a cap is shown in Figure 3.16. The max/minimum compensation value is set to $\pm 0.2k\ E^t$.

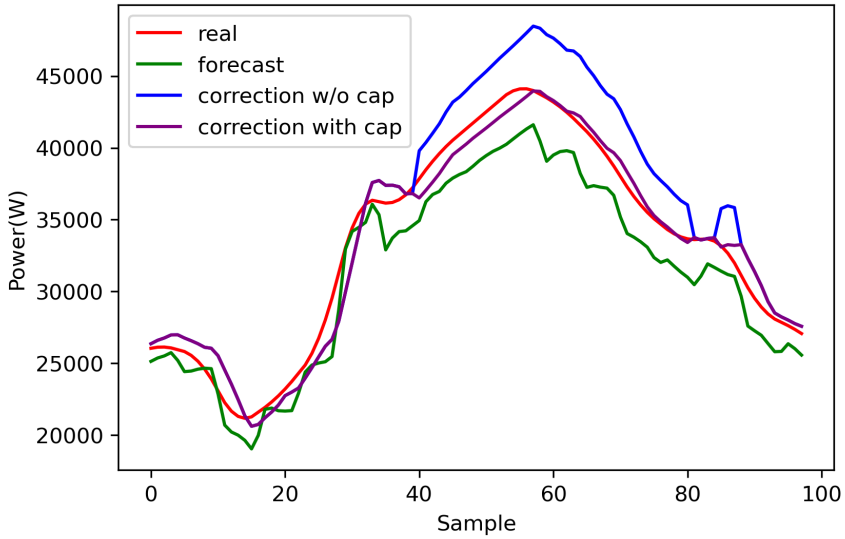


Figure 3.16: The performance of applying peak cap

In addition, Figure 3.17 illustrates the snapshot of the DAC-LSTM forecast method result for 24 hours. The red, green and blue curves represent the actual, LSTM forecast, and DAC-LSTM forecast load profiles. It can be found that the error reduces significantly, while at some points the DAC-LSTM forecast result exceeds the actual load values. Moreover, the parameters for the whole simulation are present in Figure 3.18a, 3.18b, 3.18c and 3.18d. In Figure 3.18a, α remains 0.51 as the simulation forecasts 30minute. Figures 3.18b and 3.18c represent parameters β and k. Figure 3.18d represents $k\delta$ which is within the limit $\pm 0.2k$.

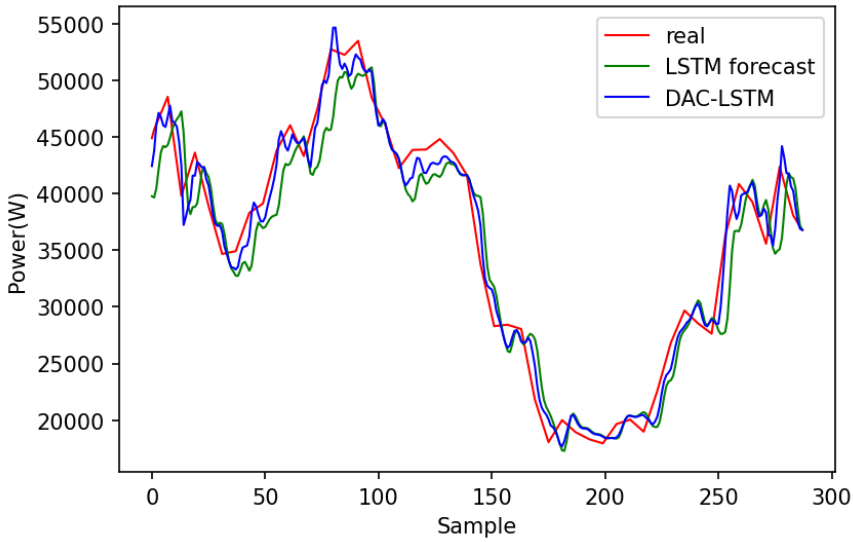
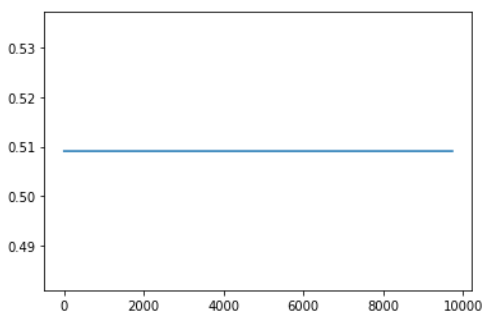
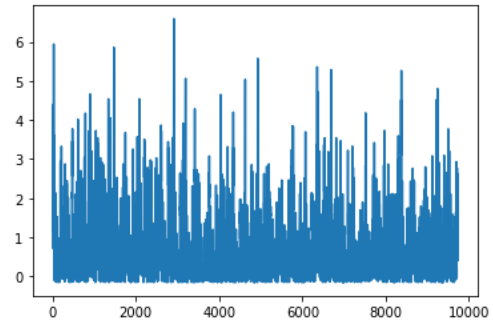


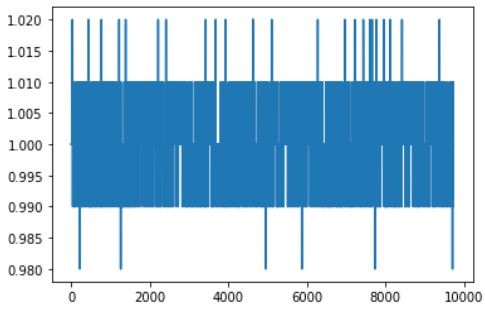
Figure 3.17: Snapshot of the DAC-LSTM method 24-hour forecast result, 30-minute forecast



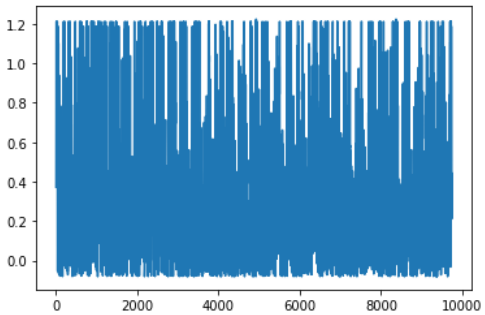
(a) α during the whole simulation



(b) β during the whole simulation



(c) k during the whole simulation



(d) $k\delta$ during the whole simulation

Figure 3.18: The plots for parameters in DAC-LSTM forecasting

Figure 3.19 displays the 30-minute ahead forecast. This Figure gives a visual comparison between the forecast performances of the ML, the persistence, and the ARIMA methods for a period of 7 days taken from the evaluation set. It can be observed that the DAC-LSTM method provides the closest value to the actual load profile.

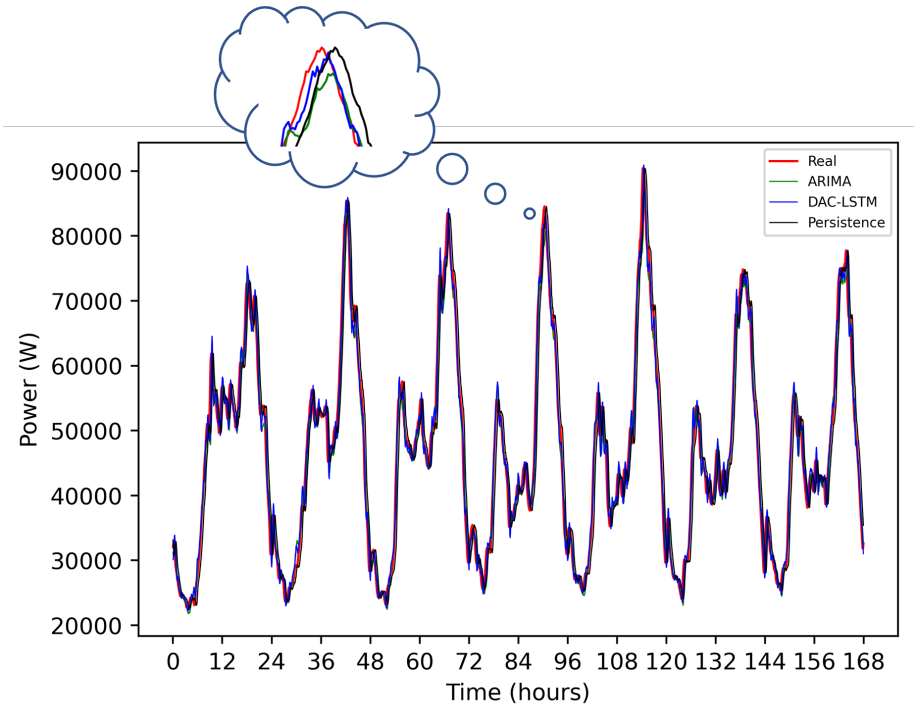


Figure 3.19: Snapshot of 7-day Forecast results from the 30-minute ahead DAC-LSTM method

The following figure 3.20 illustrates the forecast result during peak and off-peak hours in detail. From this figure, it can be noticed that during off-peak hours, the forecast result from ARIMA, Persistence and DAC-LSTM methods remains similar, while during the peak hours, the DAC-LSTM method provides the best results. The detailed numerical result is shown in table 3.3.

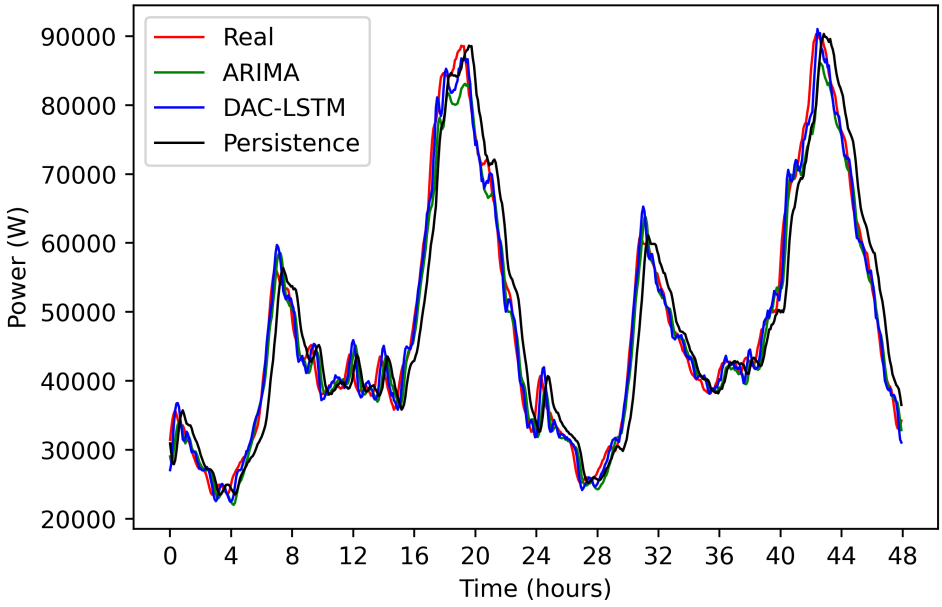


Figure 3.20: Snapshot of 2-day Forecast results from the 30-minute ahead DAC-LSTM method

Table 3.3: Forecast accuracy comparison among Persistence, ARIMA and DAC-LSTM methods, evaluated in R2

	Persistence	ARIMA	DAC-LSTM
General	0.881	0.920	0.991
Peak	0.884	0.892	0.990
off-peak	0.880	0.921	0.991

From table 3.3, it can be observed that the DAC-LSTM method achieves the best forecast accuracy during general, peak and off-peak forecasts, all above 0.99. While forecast accuracy from the ARIMA method during peak hours significantly reduces, from 0.920 to 0.892. Because the Persistence method uses a hysteresis forecast, therefore, the forecast accuracy changes slightly.

3.3.8 Sensitivity Analysis

The parameters used in the DAC-LSTM method are analysed in this section, including E_{avg}, m, n. To analyse the contribution of each parameter to the DAC-LSTM method, the experiment is carried out as follows:

1. The utilised dataset is from TVVP and ACN, containing 220 households and 550 historical days.
2. The range for E_{avg} varies from 1 to 288, which represents 5 minutes to 1 day.
3. The range for m and n varies from 1 to 100.

The maximum accurate values are achieved when m is in the range 74 to 92 and n is in the range 18 to 50, with $R2 = 0.984$ and $MAPE = 4.3\%$. The following Figures 3.21 and 3.22 illustrate the R2 and MAPE values under different m and n when 221 samples are used to calculate E_{avg}.

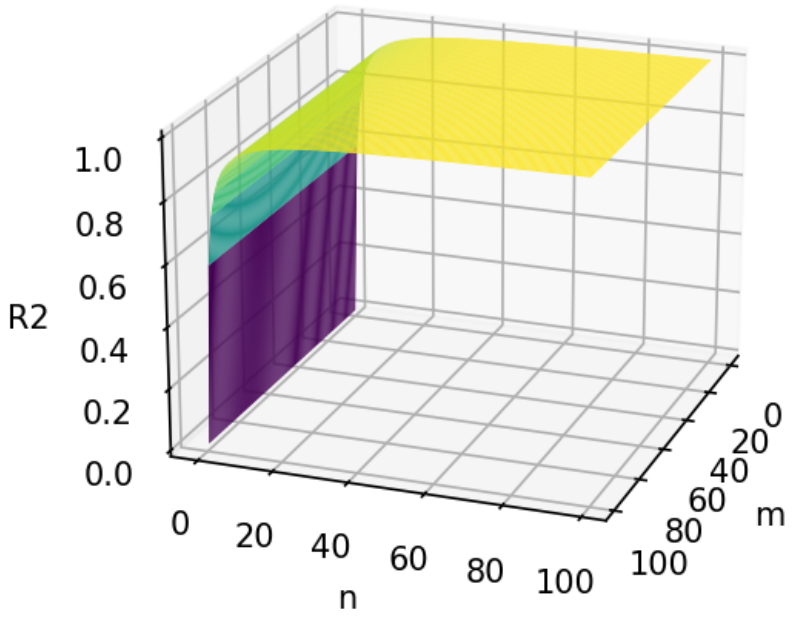


Figure 3.21: R2 under different m and n values, when 221 samples are used to calculate E_{avg}

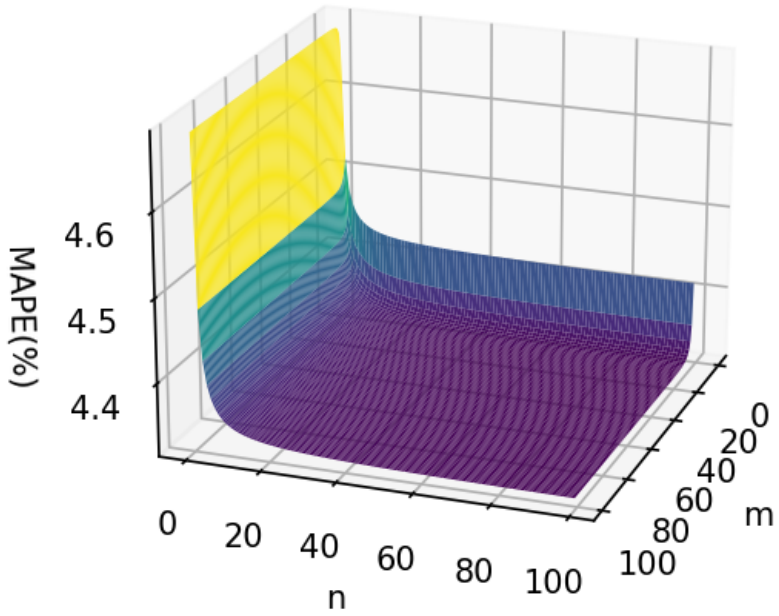


Figure 3.22: MAPE under different m and n values, when 221 samples are used to calculate E_{avg}

Figure 3.23 illustrates the R2 and MAPE values for different lengths of datasets to calculate E_{avg} (given that $m = 85$ and $n = 30$). It can be found that the highest R2 is at the 221st sample, and the smallest MAPE is at the 38th sample. Moreover, as the MAPE at 38th and 221st are 0.430 and 0.433 separately, the E_{avg} is calculated using 221 samples.

In conclusion, for the TVVP and ACN datasets, the optimised and selected parameters are $m = 85$, $n = 30$, and E_{avg} is calculated by 221 samples. For different training datasets, the parameters are required to be re-calculated. In this case, the optimal range for m and n are 18-50 and 74-92.

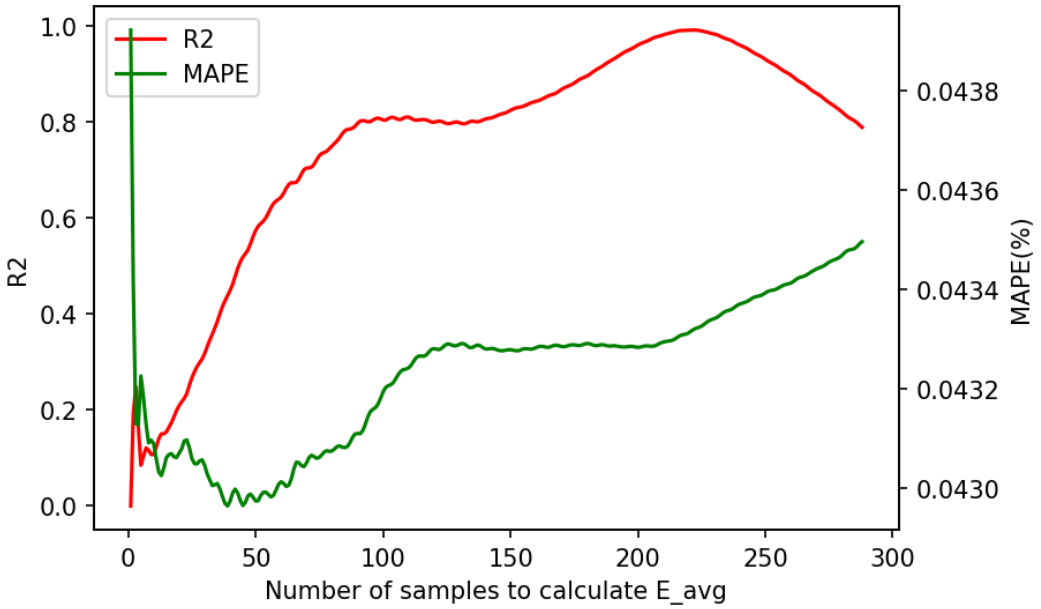


Figure 3.23: The MAPE and R2 for different sample numbers for E_avg calculation

3.4 Conclusion

This work explains Persistence and ARIMA methods with equations, and detailed parameter determination procedures are discussed with the TVVP and ACN datasets. Further, a novel robust DAC-LSTM forecast method is proposed for short-term electricity load forecast based on the existing LSTM network, including a DAC module as the add-on function. As most existing statistical and machine learning methods lack the ability to forecast peak and off-peak hour loads simultaneously. Therefore, hybrid methods or separated methods for peak and off-peak hours are required. Moreover, existing methods assume the historical dataset is sufficient to train the required network. However, with the construction of small-scale smart grids, the historical datasets tend to be high time-resolution, high variety and less diversity (smaller networks with fewer consumers). Meanwhile, methods with constant compensation parameters

lead to more forecast errors. Therefore, the existing methods show weaknesses in distribution network applications. This work proposes a robust forecast method which provides accurate forecast results for peak and off-peak load and adapts to smaller network applications.

The proposed STLF method follows a procedure including data processing, LSTM network forecasting and peak load distinguishing, DAC for dynamic error compensation, and forecast result generation. The key contribution of the proposed method is the accurate forecast result for both peak and off-peak periods at the DN when the historical data and data varieties are insufficient, while the existing top-down load forecast methods require months or years of historical data and cannot forecast peak and off-peak load simultaneously. The novel forecast method first introduces dynamic peak load distinguishing blocks to separate peak and off-peak loads. Second, the dynamic error compensation module is introduced to compensate for the forecast error with dynamic adjusting parameters for peak and off-peak loads. These functions help the forecast method reduce errors in peak and off-peak load forecasting, handle various datasets, and provide accurate forecast results.

Chapter 4

Case Studies

4.1 Introduction

The development procedures of the DAC-LSTM method are present in Chapter 3. The key contributions of the proposed method are the accurate forecast result for both peak and off-peak periods at the DN when the historical data and data varieties are insufficient. Furthermore, to validate the value of the DAC-LSTM method, this chapter intends to carry out several case studies from aspects, including varying forecast steps, varying length of training dataset, varying size of training dataset, expanding application range to the transmission network, and qualitatively comparison with other methods. Moreover, the DAC-LSTM method is benchmarked with the Persistence and ARIMA methods.

4.2 Persistence and ARIMA Methods

This section includes an explanation of Persistence and ARIMA methods, which are two commonly utilised methods in the academic and industrial areas. These two methods are introduced as the benchmark to evaluate the performance of the DAC-LSTM method.

4.2.1 Persistence Method

The traditional short-term load forecast method is based on the fact that consumer behaviours could be considered 'quasi-stationary' on a time scale of a period[162]. However, the power consumption varies slowly over days at the transmission network level. Therefore, the persistence method is commonly utilised. The equation is shown below:

$$P_{t+k} = P_i \tag{4.1}$$

In the equation above, P is the power. t represents the current status. i is the input power time stamp. k is the forecast step. This equation indicates that the forecasted value at step k is assumed to be the same as the previous value at time i.

The persistence method performs well at the transmission network level, but when it comes to the distribution network or community network, the uncertainty of the load forecast increases. Meanwhile, the load is influenced by special events (like the football match will increase the evening peak), weekdays or weekend days, seasons, weather, etc. Therefore, the persistence method shows a weakness in the short-term load forecast.

4.2.2 ARIMA Method

To improve the forecast accuracy for fewer consumers and short-term loads, Autoregressive Integrated Moving Average (ARIMA) method with its variations has been widely proposed in recent years. ARIMA is a generalised model of Autoregressive Moving Average (ARMA) that combines the Autoregressive (AR) process and Moving Average (MA) processes and builds a composite model of the time series. As the acronym indicates, ARIMA captures the key elements of the model:

- AR: Autoregression. The dependencies between the input and several lagged inputs are utilised to build the regression model (p).
- I: Integrated. The input dataset is measured in derivation at different times to reach stationary status (d).
- MA: Moving Average. The dependency between inputs and the residual error terms is considered when a moving average model is used for the lagged observations (q).

Equation expressions of the ARIMA are shown as follows, from [163]. First is the AR (p), which can be written as a linear process:

$$x_t = c + \sum_{i=1}^{P} \phi_i x_{t-i} + \varepsilon_t \tag{4.2}$$

In the equation above, x_t is the stationary variable. c is constant. ϕ_i are autocorrelation coefficients at lags from 1 to P. ε_t is the residuals, the Gaussian white noise series with mean zero and variance σ_ε^2.

Second is the MA, which can be written in the form:

$$x_t = \mu + \sum_{i=0}^{q} \theta_i \varepsilon_{t-i} \tag{4.3}$$

In the equation above, μ is the expectation of x_t (usually assumed equal to zero). θ is the weights applied to the current and prior values of a stochastic term in the time series, $\theta_0 = 1$.

By combining equations 4.2 and 4.3 and form an ARIMA model of order (p,q):

$$x_t = c + \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon_t + \sum_{i=0}^{q} \theta_i \varepsilon_{t-i}$$

Where $\phi_i \neq 0$, $\theta_i \neq 0$, and $\sigma_\varepsilon^2 > 0$. p and q are the AR and MA orders. The 'Integrate' component involves differencing the time series to convert a non-stationary time series into a stationary. The general form of an ARIMA model is denoted as ARIMA (p, d, q).

The critical step in building up the ARIMA model is the determination of (p, d, q) values. By observing the plot of the input dataset, for instance, if the variance grows or reduces with time, derivations should be used to stabilise the dataset and determine d. The parameter d is the order of difference frequency changing from non-stationary time series to stationary time series. Then, the autocorrelation function (ACF) is introduced to measure the amount of linear dependence between inputs, and this determines the lag p. Moreover, the partial autocorrelation function (PACF) is used to determine the autoregressive terms q [163].

Figure 4.1, 4.2 and 4.3 illustrate the original series, first-order differencing, and second-order differencing of TVVP datasets and their corresponding ACF and PACF. The second-order differencing dataset is used for ARIMA parameters determination. The determination process of parameters p, d, and q are explained as follows:

1. By comparing the ACF of original, first-order differencing and second-order differencing, the time series reaches stationary with two orders of differencing. However, looking at the ACF plot for the 2nd differencing, the lag goes into the far negative zone reasonably quickly, indicating that the series might have been over-differenced. Therefore, for this specific dataset, d should be 1.
2. Any autocorrelation in a stationarised series can be rectified by adding enough AR terms. Thus, initially, The order of the AR term is taken to be equal to as many lags that crosses the significance limit (the blue area) in the PACF plot. It can be observed that the PACF lag 4 is quite significant since it is well above the significance line. Thus, p is fixed as 4.

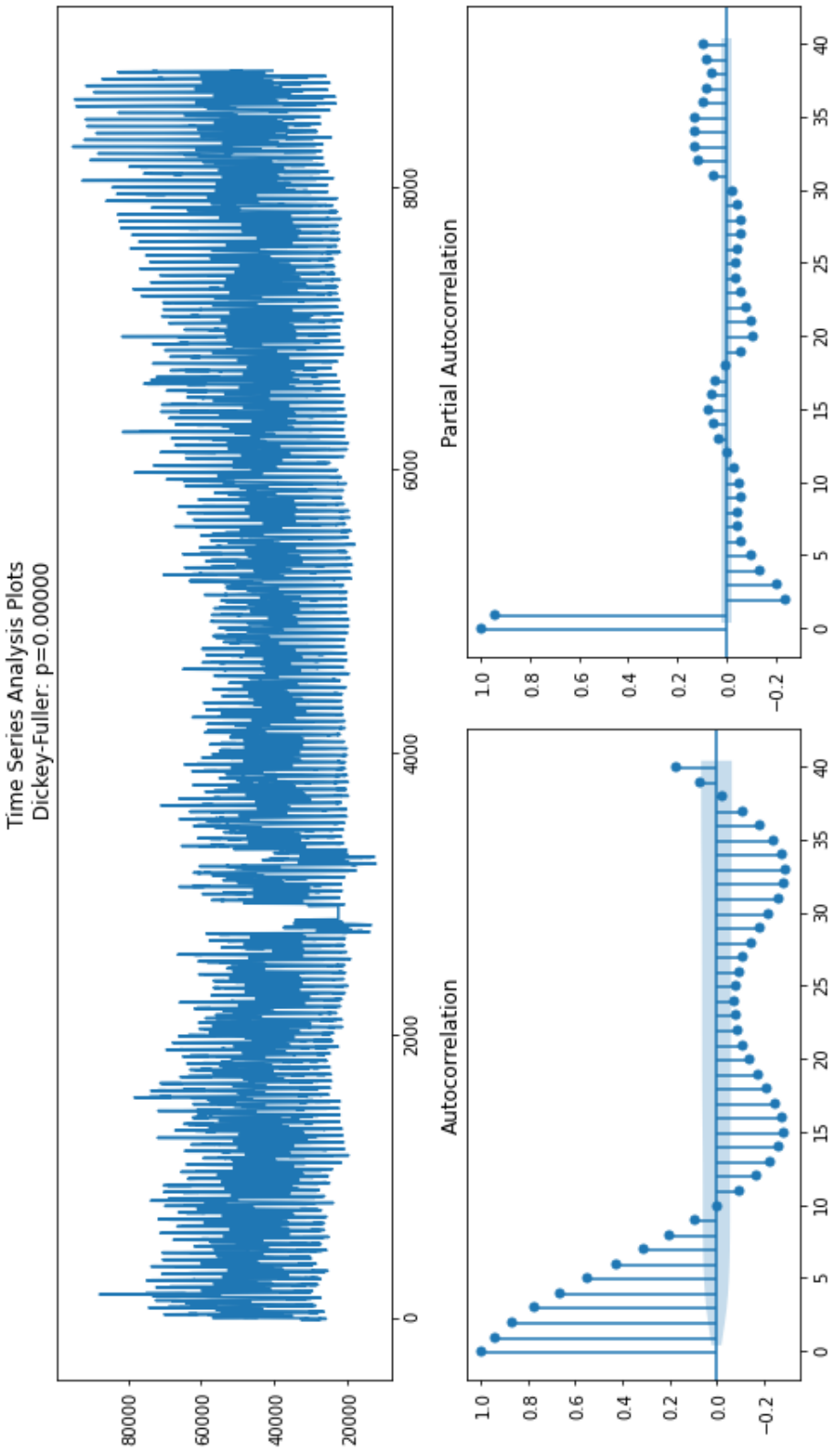


Figure 4.1: Plot of original series, and its ACF and PACF for the TVVP dataset

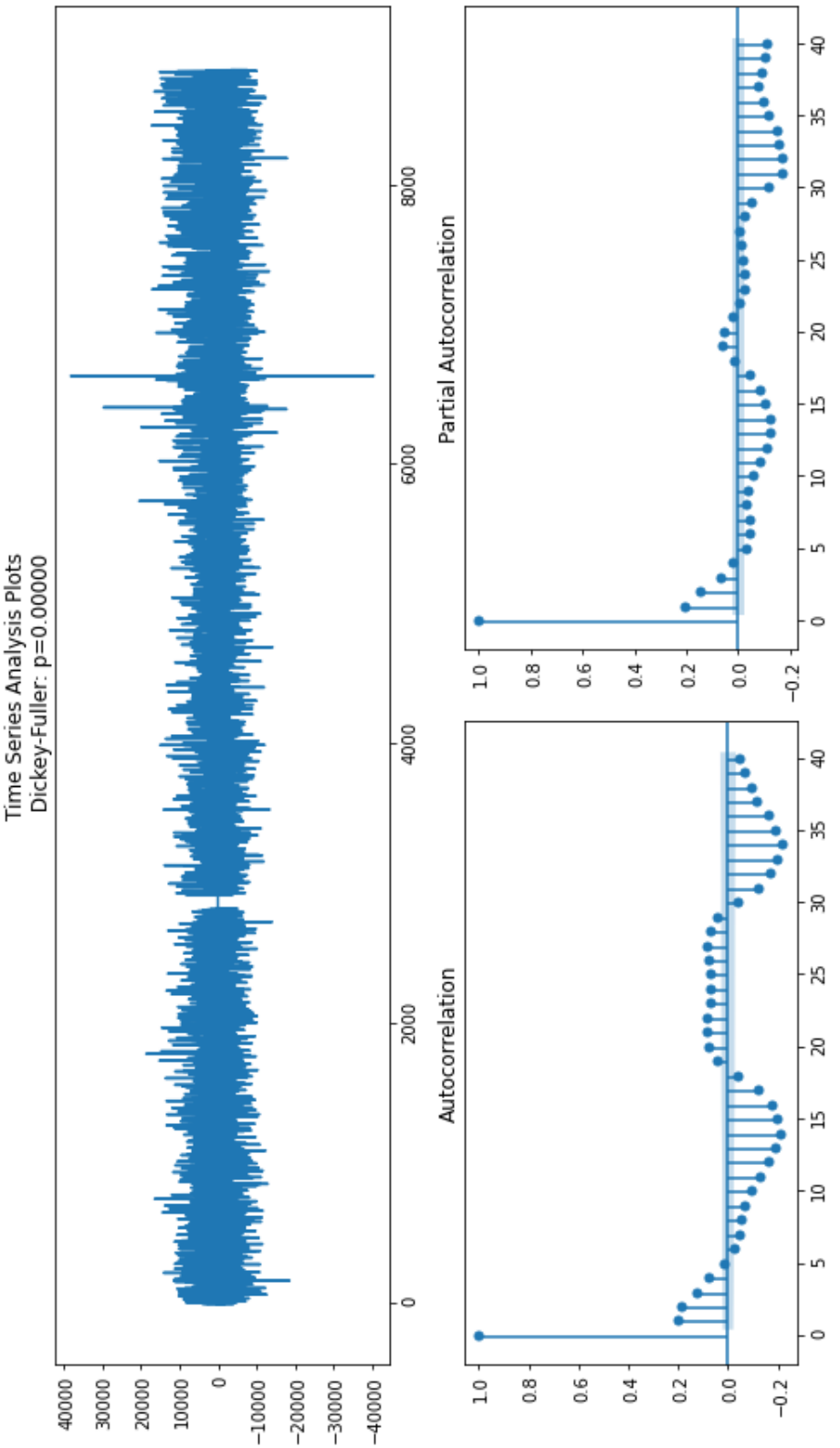


Figure 4.2: Plot of 1^{st} differencing, and its ACF and PACF for the TVVP dataset

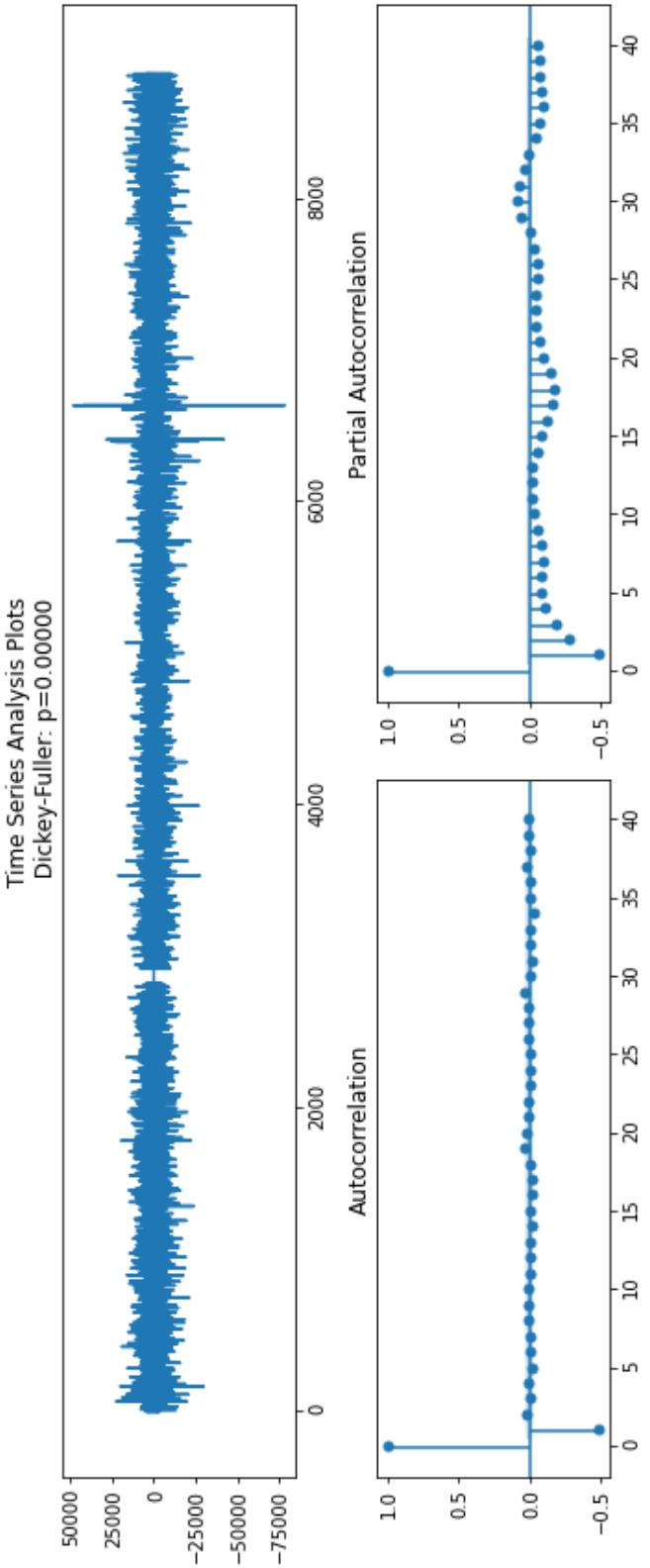


Figure 4.3: Plot of 2^{nd} differencing, and its ACF and PACF for the TVVP dataset

3. The ACF tells how many MA terms are required to remove any autocorrelation in the stationarised series. A couple of lags are well above the significance line. Therefore, q is fixed as 4.

According to the steps above, ARIMA model parameters are defined as follows: $(p,d,q) = (4,1,4)$.

4.3 Performance Metrics

In this chapter, the performance of Dynamic Adaptive Compensation DAC-LSTM is evaluated and carried out with case studies. TVVP and ACN datasets are used for each case study. Moreover, the EXLETRON UK domestic load dataset is used to test the application of DAC-LSTM in a higher-level power network. When evaluating the DAC-LSTM, the following settings are made:

1. Each model is trained ten times and then takes the average error value to show a steady performance of DAC-LSTM.
2. When training the LSTM model, the model parameters are adjusted according to the training data pool size, forecast steps, and forecast steps ahead.
3. Each model is trained for 500 epochs and then returns to the best epoch.
4. A 30-day electric load is forecasted using the trained model and compared with actual values to derive the $R2$, MAPE and correlation.
5. When needed, the input dataset is expanded to 5 min time resolution with linear interpolation and random errors. The random error is typical -1.5% to 1.5%.
6. The configuration of parameters in the peak detection is as follows: lag is 288, the threshold is 1, and influence is 0.8.
7. The configuration of the DAC module parameters is as follows: $m = 85$, $n = 35$.
8. The configuration of parameters in the ARIMA is as follows p, d, q = 4, 1, 4.

The performance is evaluated by using the following values: R-squared (R2), Mean absolute percentage error (MAPE) and Pearson Correlation (Correlation). Regression models are usually evaluated by $R2$ and MAPE, both functions of errors between predicted and actual values. $R2$ is a statistical measure representing the proportion of the variance for a dependent variable explained by an independent variable or variables in a regression model. For example, if $R2$ equals 0.5 in a model, then approximately half of the observed variation can be explained by the model's inputs. MAPE is the sum of the individual absolute errors divided by the demand (each period separately). It considers the error between forecast and actual and the percentage of error and actual.

$$R^2 = 1 - \frac{UnexplainedVariation}{TotalVariation} \tag{4.4}$$

$$MAPE = \frac{1}{n}\sum |\frac{A_i - F_i}{A_i}| \tag{4.5}$$

$$\rho(x,y) = \frac{E[xy]}{\sigma_x \sigma_y} \tag{4.6}$$

where $E[xy]$ is the cross-correlation between x and y, and $\sigma_x^2 = E[x^2]$ and $\sigma_y^2 = E[y^2]$ are variances of signals x and y, respectively. The $R2$ and MAPE reflect the

4.4 Case 1: Comparison Between DAC-LSTM and LSTM Methods with Various Forecast Steps

First, the proposed DAC-LSTM method is compared with the original LSTM method. The forecast step varies from half an hour (1 step) to 24 hours (48 steps). The numeric value is shown in Appendix, Table A.1. The data from Table A.1 is plot in Figure 4.4, 4.5, 4.6:

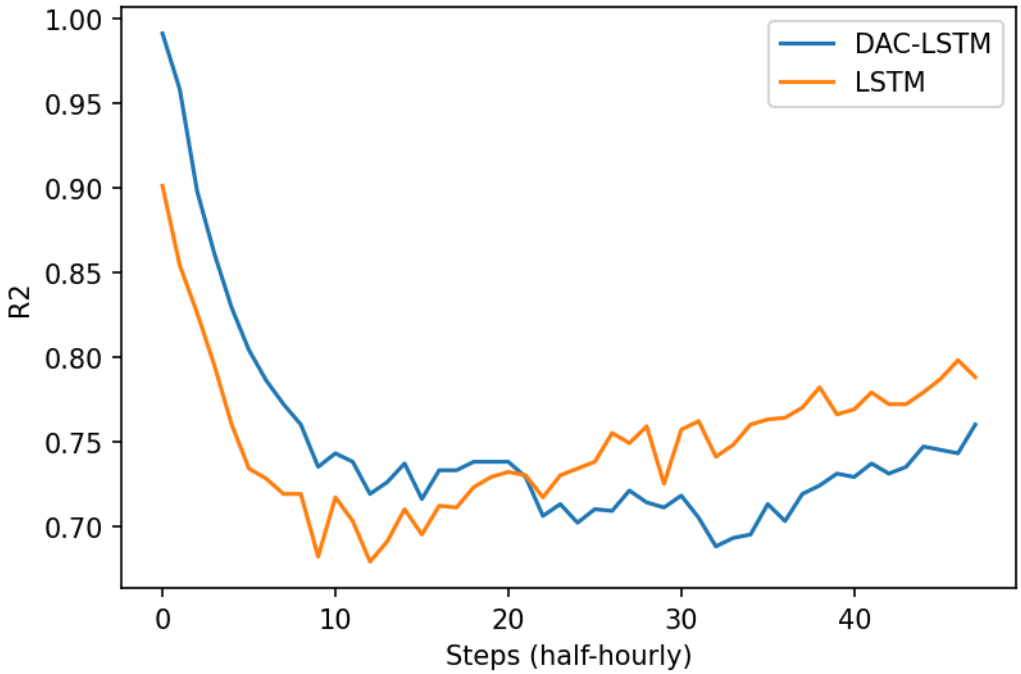


Figure 4.4: The forecast results comparison between LSTM and DAC-LSTM methods, forecast step varies from 0.5 hours to 24 hours, evaluated in R2

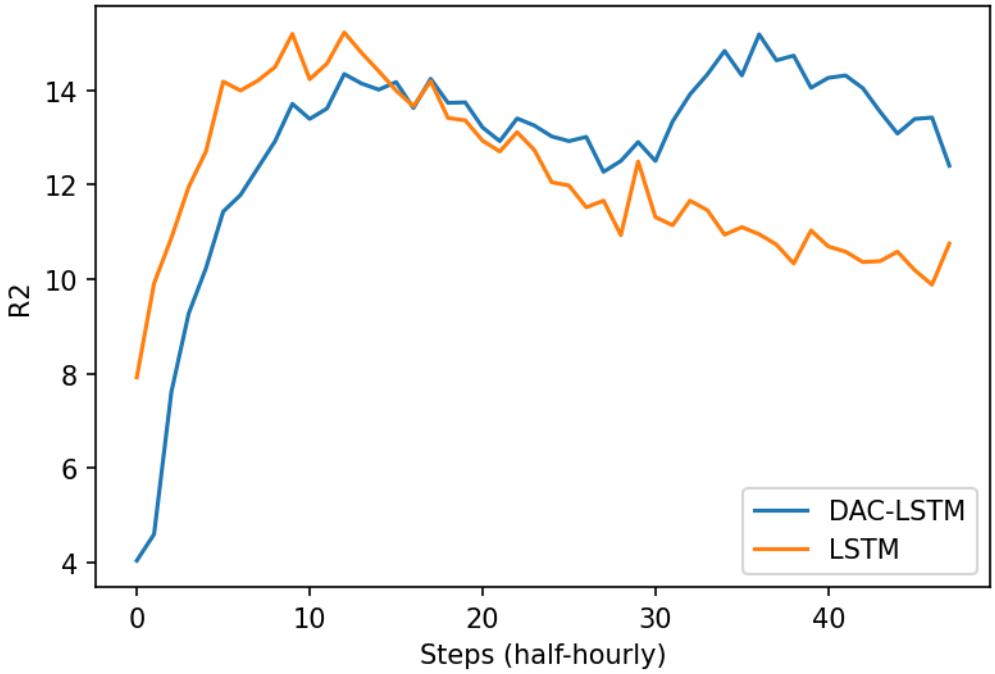


Figure 4.5: The forecast results comparison between LSTM and DAC-LSTM methods, evaluated in MAPE

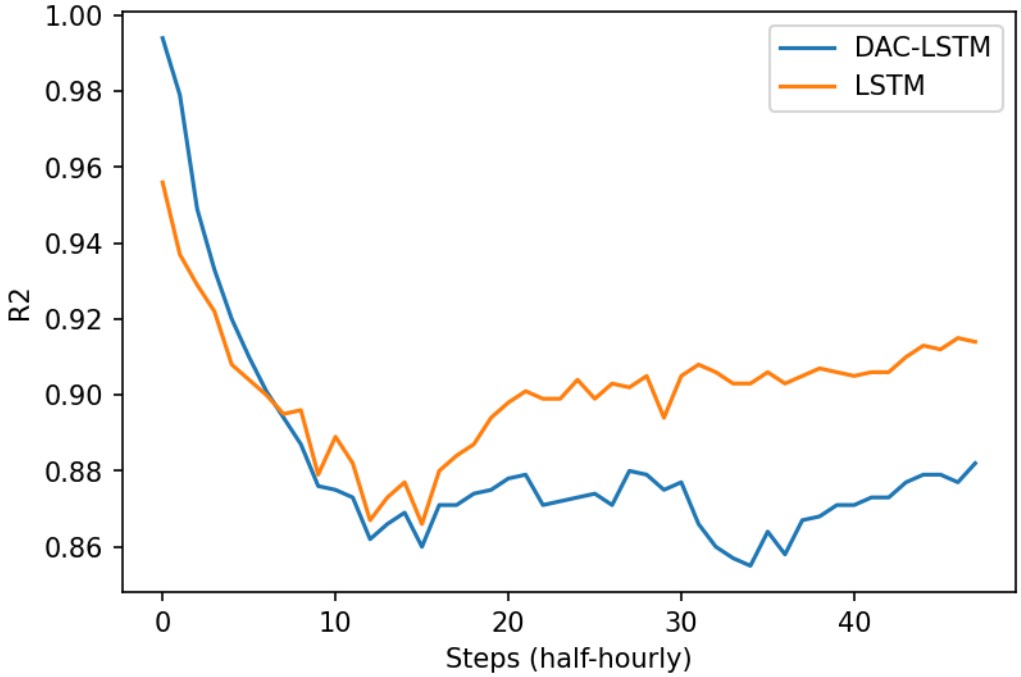


Figure 4.6: The forecast results comparison between LSTM and DAC-LSTM methods, forecast step varies from 0.5 hours to 24 hours, evaluated in Corr

The improvement of the DAC-LSTM method compared with the LSTM method for the 24-hour forecast step is shown in Figure 4.7, 4.8:

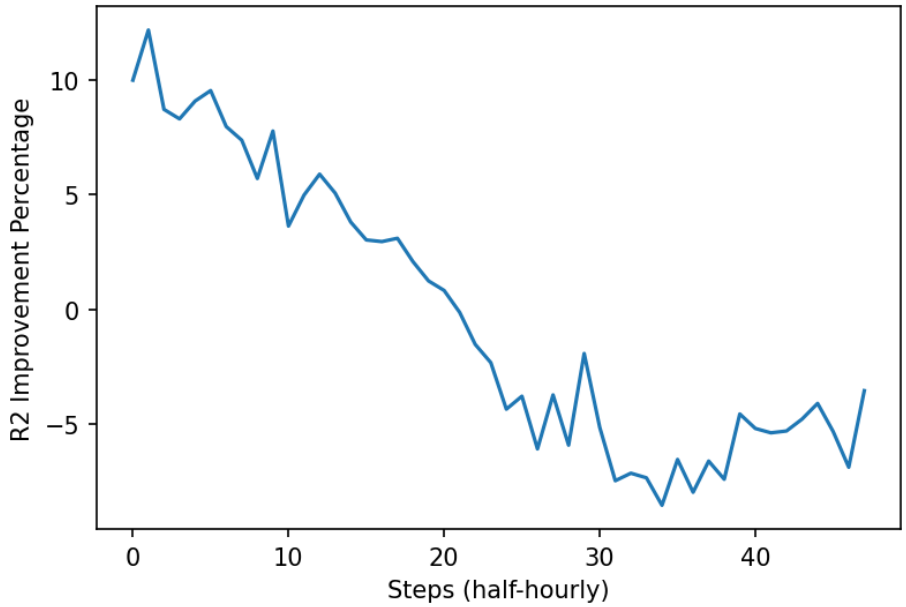


Figure 4.7: Case 1: R2 Improvement percentage of DAC-LSTM compared with LSTM

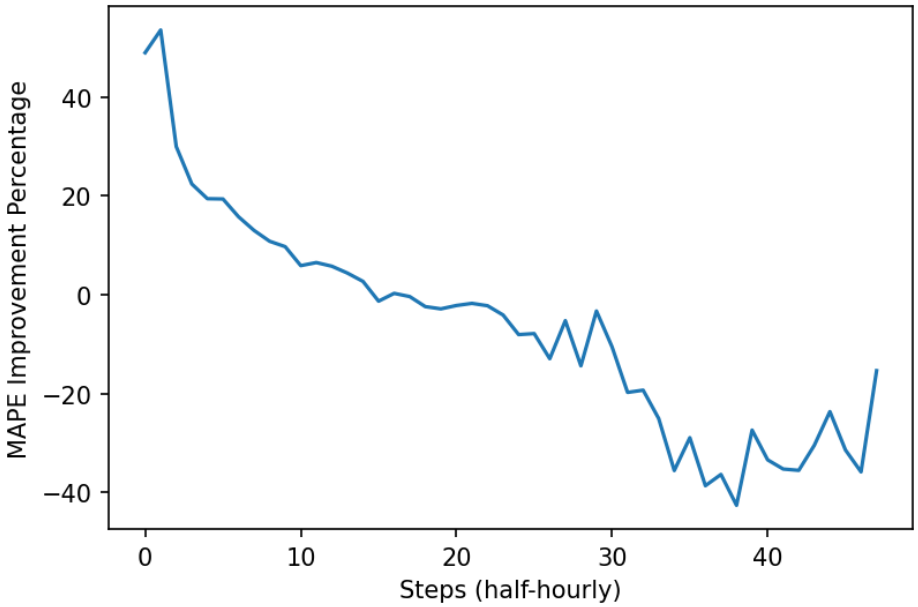


Figure 4.8: Case 1: MAPE Improvement percentage of DAC-LSTM compared with LSTM

In Figure 4.4 - 4.8, the DAC-LSTM method shows significant improvement within the 12-hour forecast step. When forecast steps range from 1 (30 minutes) to 18 (9 hours), the DAC-LSTM method performs better than the LSTM method. The improvement of R_2 and MAPE ranges from 12.2% (R2) and 53% (MAPE) to 3.1% (R2) and 2.7%. In addition, the forecast accuracy of both LSTM and DAC-LSTM minimises at around the 6-hour forecast step, which is followed by a slight increase. Eventually, MAPE fluctuated at 14% (LSTM) and 11% (DAC-LSTM) and R_2 fluctuated at 4.2% (LSTM) 3.8% (DAC-LSTM).

This result illustrates that at shorter-term load forecast, the DAC-LSTM method significantly improves compared with the LSTM method while the forecast step is less than 9 hours. However, when the forecast step approaches 24 hours, the DAC-LSTM method results in 5% more error than the LSTM method on average. According to section 3.3.7, the DAC module first monitors the moving average error from the previous forecast and real load values. Then, the dynamic compensation fixes forecast errors based on pre-defined equations and dynamic parameters. When the forecast step increases, the correlation between the previous forecast error and the compensation coefficient reduces. Therefore, DAC-LSTM outperforms LSTM when the forecast step is less than 18 (9 hours) and slightly falls behind LSTM at the 24-hour forecast.

4.5 Case 2: Comparison Between DAC-LSTM and ARIMA with Various Forecast Steps

The second case study compares the DAC-LSTM method with the ARIMA method. The forecast step varies from 1 (30-minute) to 48 (24 hours). The historical data pool for training is fixed at 550 days and 220 households. The numeric value is shown in Appendix, Table A.2. The data from Table A.2 is plot in Figure 4.9, 4.10, 4.11:

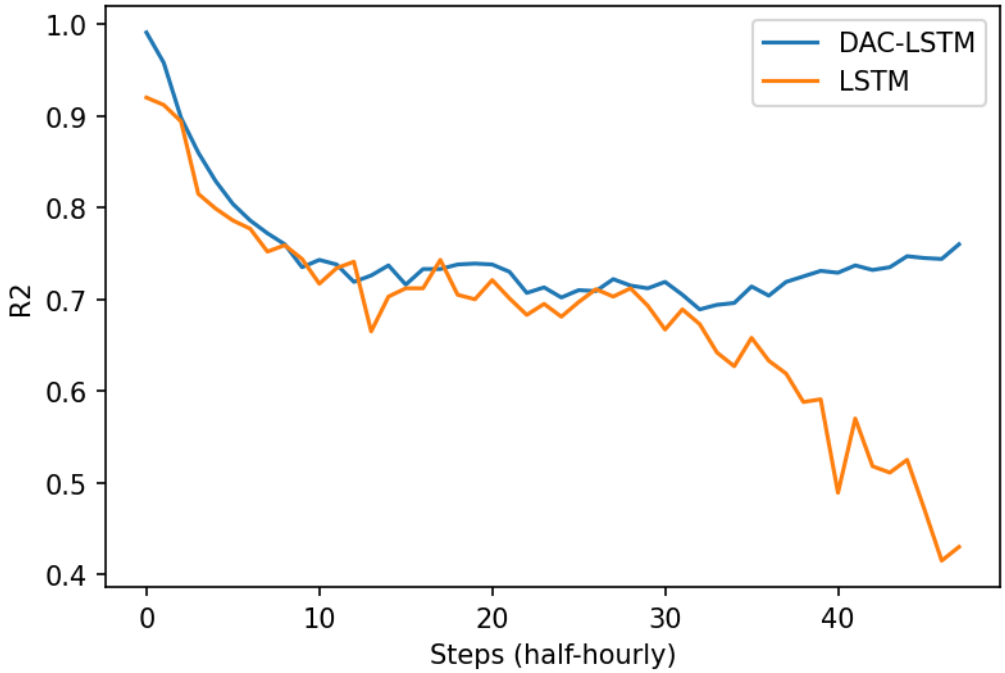


Figure 4.9: The forecast results comparison between ARIMA and DAC-LSTM methods, forecast step varies from 0.5 hours to 24 hours, evaluated in R2

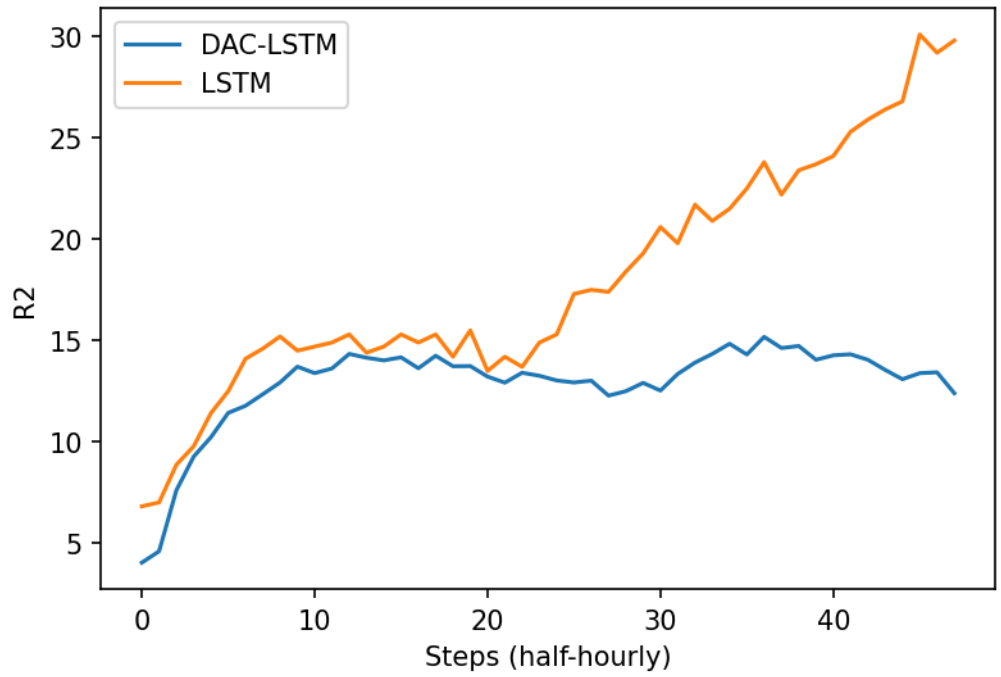


Figure 4.10: The forecast results comparison between ARIMA and DAC-LSTM methods, evaluated in MAPE

The improvement of the DAC-LSTM method compared with the ARIMA method for the 24-hour forecast step is shown in Figure 4.12 and 4.13:

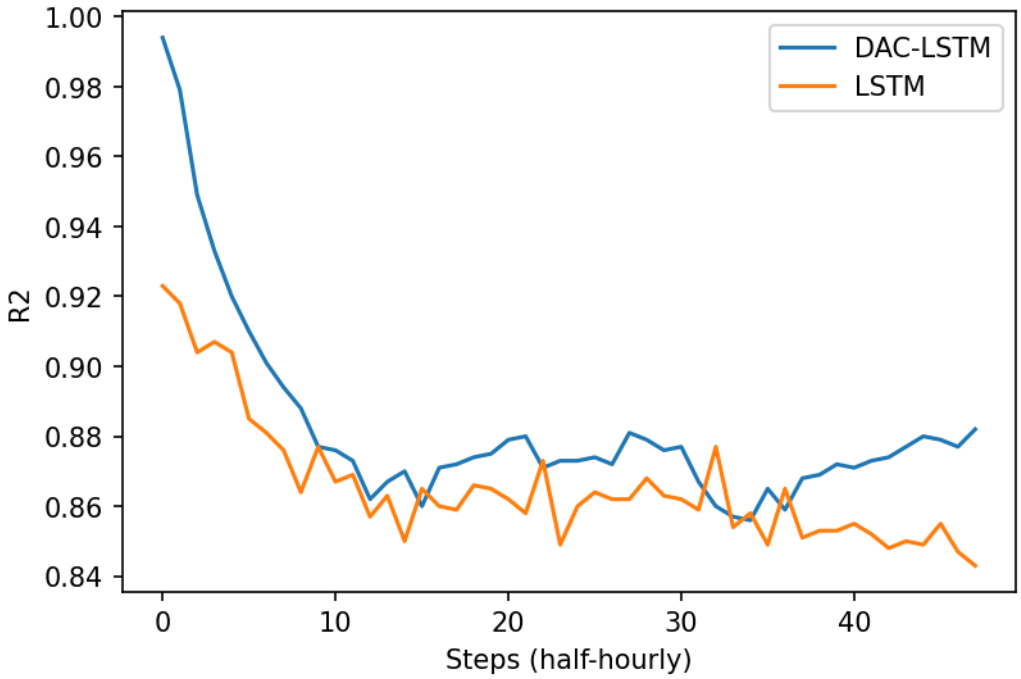


Figure 4.11: The forecast results comparison between ARIMA and DAC-LSTM methods, forecast step varies from 0.5 hours to 24 hours, evaluated in Corr

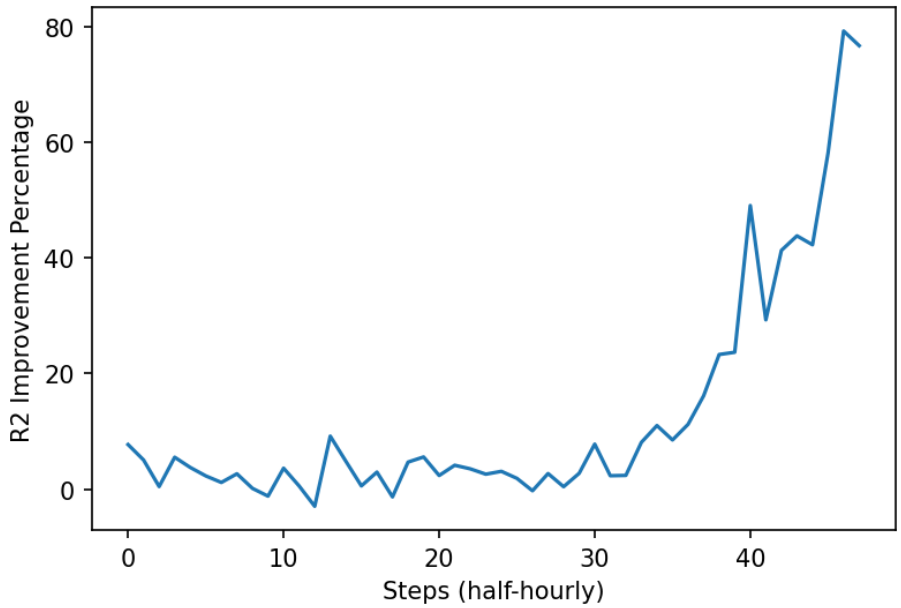


Figure 4.12: Case 2: R2 Improvement percentage of DAC-LSTM compared with ARIMA

In Figure 4.9 - 4.13, the DAC-LSTM method shows significant improvement within the 12-hour forecast step. For this 48 steps comparison, DAC-LSTM performs better than ARIMA except for steps 10 (-1.20%), 13 (-2.97%), 18 (-1.35%), 23 (-3.95%), and 27 (-0.28%) which improvements are negative. Despite these negative values,

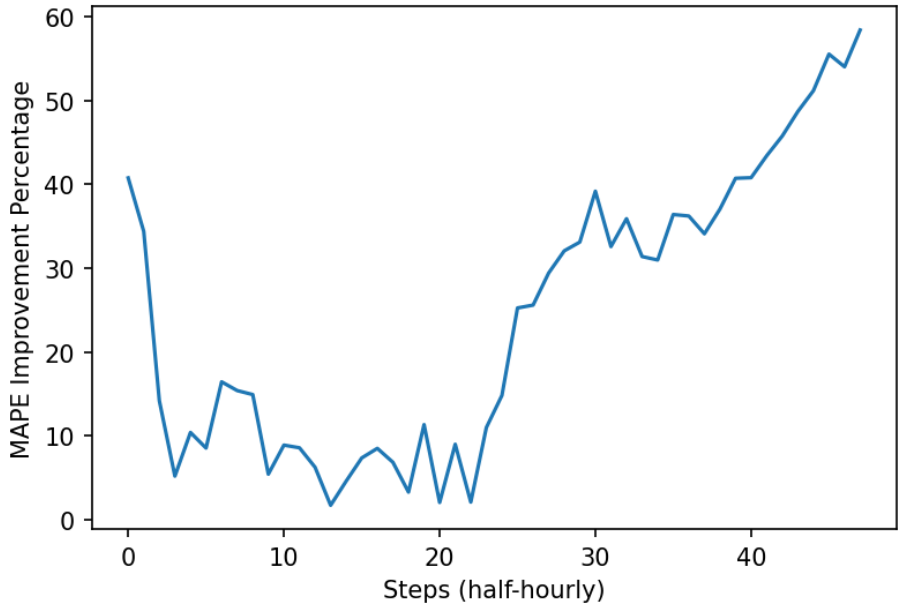


Figure 4.13: Case 2: MAPE Improvement percentage of DAC-LSTM compared with ARIMA

which may be caused by a specific dataset that reduces the model accuracy, the overall improvement is significant, which achieves 79.6% (R2) and 58.4% (MAPE). This result proves that at a shorter-term load forecast (within 24 hours), the DAC-LSTM method provides a significant improvement compared with the ARIMA method. In addition, various forecast step comparisons between DAC-LSTM and Persistence are not included as the Persistence method assumes the forecast value equals the forecast-step ahead value. This limits the application of the Persistence method to an hour or day-ahead forecast.

4.6 Case 3: Various Lengths of Historical Data

The third case study compares forecast results with different amounts of days for training. The historical data used are from TVVP and ACN, as described in Section 3.3.2. First, the model is tested with different training pools, from 550-day to 14-day historical data. The expected forecast step is 6, which represents 30 minutes forecast. This case study tests the model's robustness when historical input data is insufficient.

Tables A.3, A.4, A.5 in Appendix contain numeric values of R2, MAPE, and correlation according to different amounts of days for training by using Persistence, ARIMA, and DAC-LSTM methods. The forecast accuracy values are plotted as bar charts in Figure 4.14, 4.16, and 4.18.

It illustrates that the forecast accuracy reduces with the reduction of training samples. However, the DAC-LSTM method provides good forecast accuracy when the data pool is reduced. The forecast accuracy from DAC-LSTM remains over 0.95 by using more than 100-day historical data.

4.6.1 Comparison of R2

Figures 4.14 and 4.15 indicate $R2$ values and percentage of improvements of DAC-LSTM compared with Persistence and ARIMA methods.

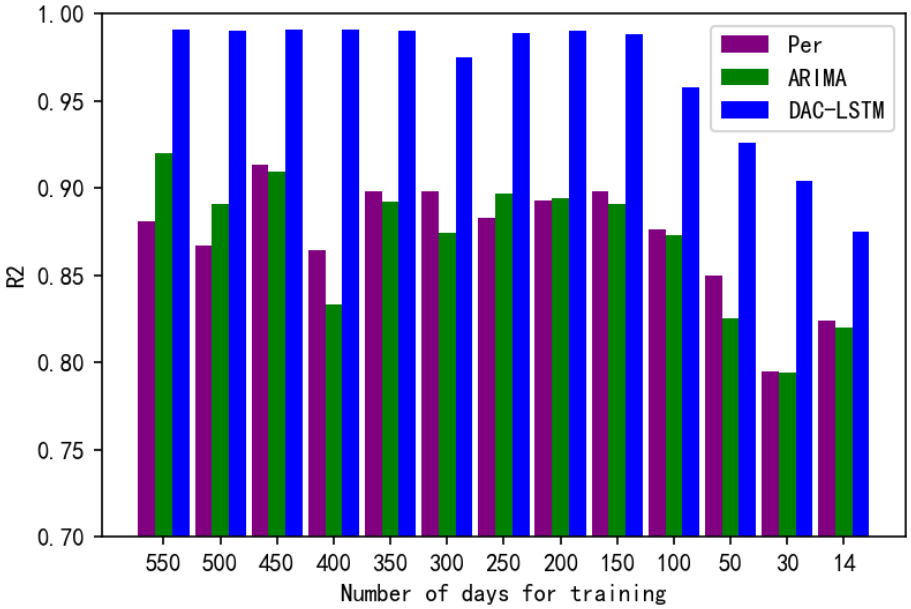


Figure 4.14: Case 3: $R2$ with different amount of households for training

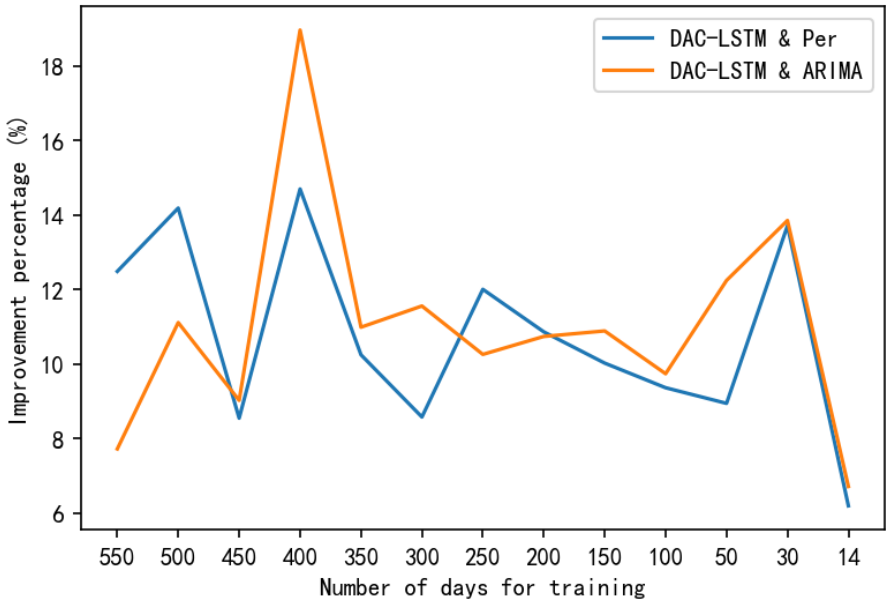


Figure 4.15: Case 3: $R2$ with different amount of households for training

As the figures show, the overall $R2$ increases by applying the DAC-LSTM method. Figure 4.14 illustrates the DAC-LSTM method maintains forecast accuracy over 0.97 when training samples are obtained from more than 150 historical days. While Persistence and ARIMA methods represent a decreasing trend with the reduced number of historical days, none of these methods provides forecast accuracy above 0.93. In addition, Figure 4.15 illustrates the DAC-LSTM method provides 7% - 19% improvement upon the ARIMA method and 6% - 15% improvement upon the Persistence method.

4.6.2 Comparison of MAPE

Figures 4.16 and 4.17 indicate MAPE values and percentage of improvements of DAC-LSTM compared with Persistence and ARIMA methods.

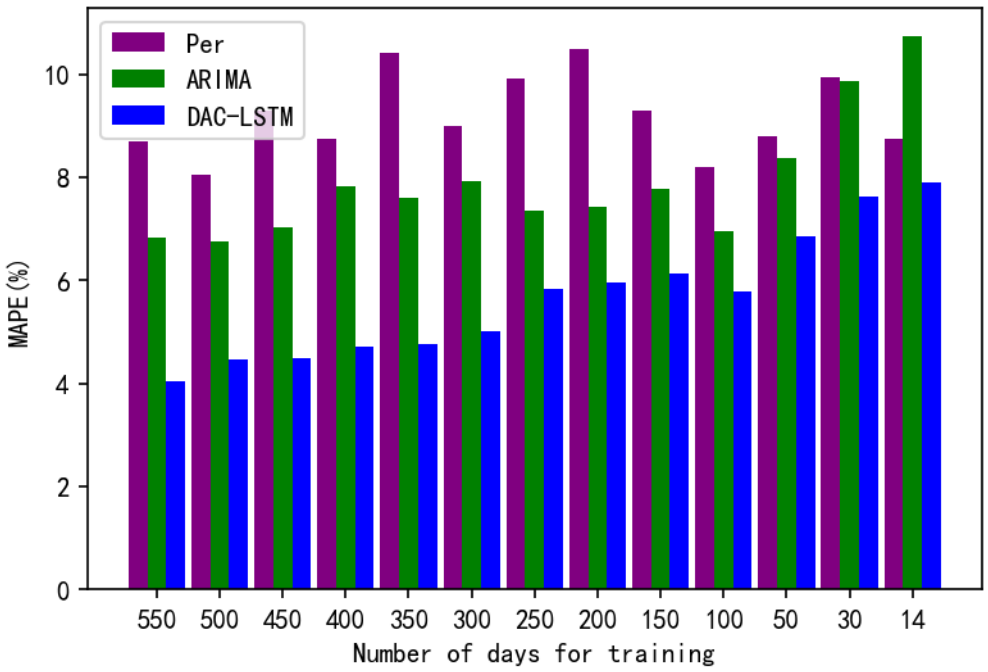


Figure 4.16: Case 3: MAPE with different amount of households for training

As the figures show, the overall MAPE increases by applying the DAC-LSTM method. Figure 4.16 illustrates that the DAC-LSTM method maintains forecast accuracy below 5% when training samples are obtained from more than 300 historical days. While Persistence and ARIMA methods represent an increasing trend with the reduced number of historical days, none of these methods provides 5% or lower forecast error. In addition, Figure 4.17 illustrates the DAC-LSTM method provides 16% - 41% improvement upon the ARIMA method and 10% - 53% improvement upon the Persistence method.

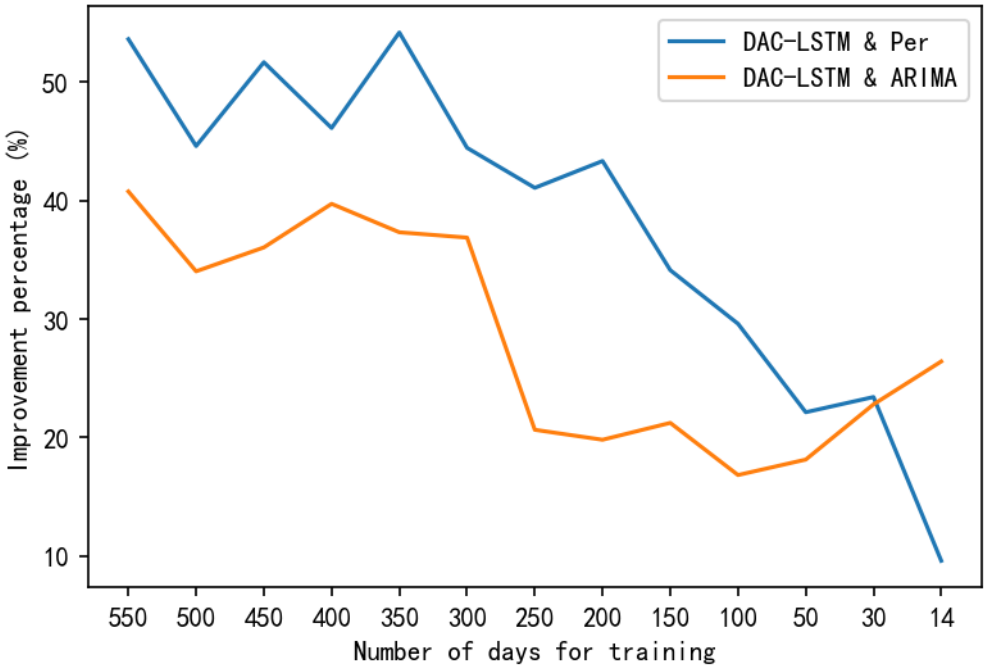


Figure 4.17: Case 3: MAPE improvement with different amount of households for training

4.6.3 Comparison of Correlation

Figures 4.18 and 4.19 indicate Correlation values and percentage of improvements of DAC-LSTM compared with Persistence and ARIMA methods.

As the figures show, the overall correlation between forecasted and real loads increases after applying the DAC-LSTM method. Figure 4.18 illustrates that the DAC-LSTM method maintains a correlation of forecast and real values over 0.988, representing that the forecasted load maintains the most similar load shape compared with the real load. Correlations of ARIMA and Persistence show a decreasing trend with the reduction of the length of the historical dataset. In addition, the green bars in figure 4.18 are higher than the red bars, proving that the ARIMA method provides more convincing results than the Persistence method. In figure 4.19, improvements are 2% - 11% and 4% - 11% for ARIMA and Persistence methods separately.

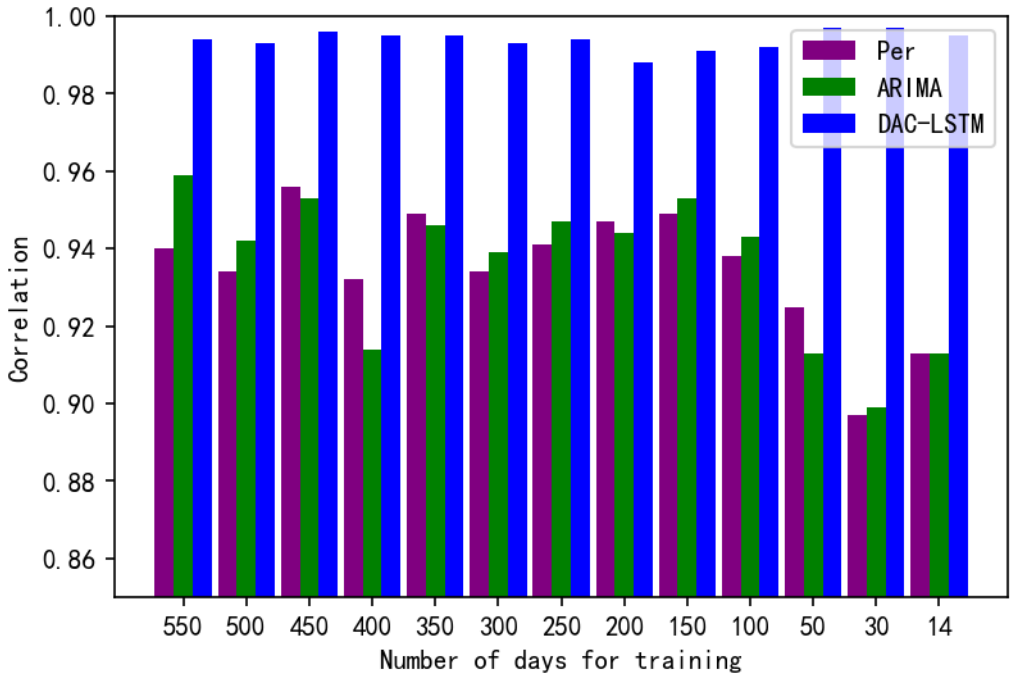


Figure 4.18: Case 3: Correlation with different amount of households for training

In conclusion, with the reduced length of the historical dataset, the DAC-LSTM method provides the most convincing forecast result. $R2$ values remain over 0.95 when the length of historical data is more than 130 days, where the input changes can efficiently affect the predicted load. Moreover, the MAPE remains below 5% when the input data is more than 300 days, where the predicted load gives relatively low errors. Therefore, for this specific training dataset, the DAC-LSTM method provides accurate forecast results ($R2$ over 0.95 and MAPE below 5%) when the length of the historical dataset is longer than 300 days. Moreover, the DAC-LSTM method provides a convincing result ($R2$ over 0.95 and MAPE below 6.1%) if the length of the historical dataset is longer than 100 days. This case study illustrates that the proposed DAC-LSTM method can be applied in a situation where the length of the historical dataset is limited, which may be caused by temporal device failure, lack of infrastructure development, etc.

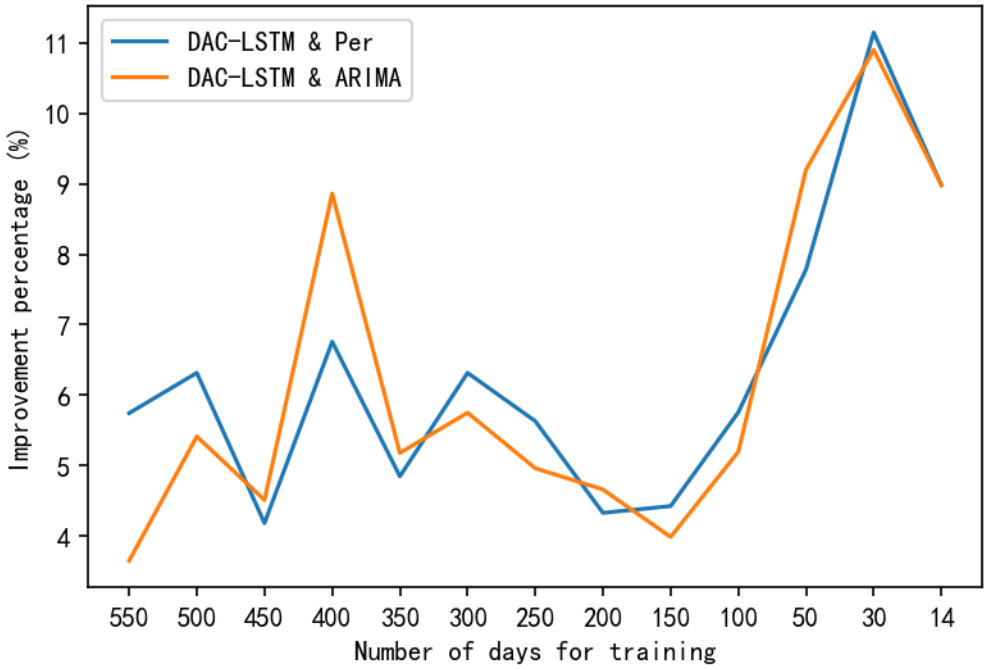


Figure 4.19: Case 3: Correlation improvement with different amount of households for training

In addition, the improvement is significant for correlation when comparing DAC-LSTM with ARIMA and Persistence methods. This shows that DAC-LSTM keeps the load shape and trend the closest to the actual shape.

4.7 Case 4: Various Sizes of Household Numbers

The fourth case study compares forecast results with different amounts of households used for training. The historical data pool for training is fixed at 550 days. The expected forecast step is 6, which represents 30 minutes forecast. The trained models are used to forecast 220 households' electrical load. Table A.6, A.7, A.8 in the Appendix contain numerical values of R2, MAPE, and correlation according to different amounts of households for training by using Persistence, ARIMA, and DAC-LSTM methods. These values are plotted as bar charts in Figure 4.20, 4.22, and 4.25.

It shows that the forecast accuracy reduces with the reduction of training samples. However, the DAC-LSTM method provides good forecast accuracy when the data pool is reduced. The forecast accuracy from DAC-LSTM remains over 0.95 by using more than 30 household historical data.

4.7.1 Comparison of R2

Figure 4.20 and 4.21 indicate $R2$ values and percentage of improvements of DAC-LSTM compared with Persistence and ARIMA.

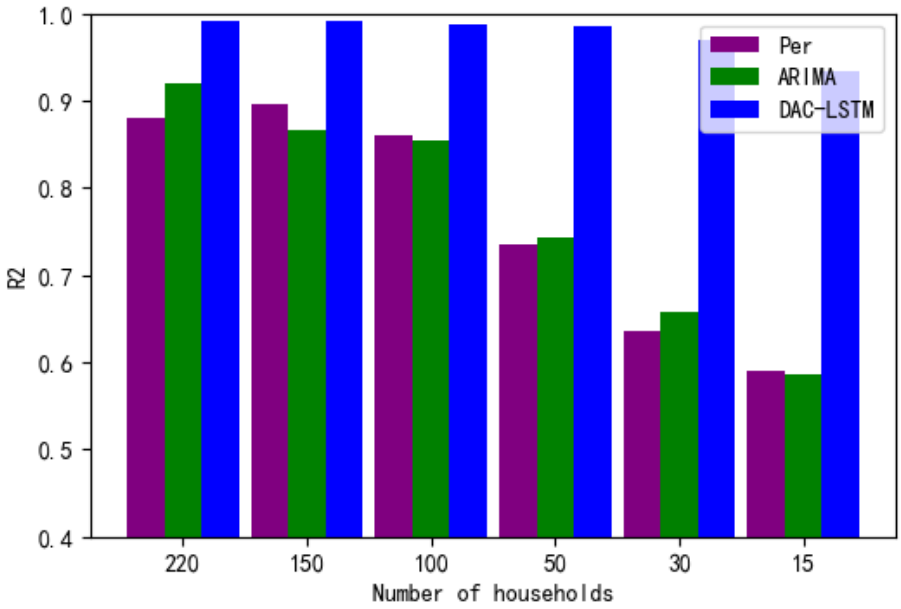


Figure 4.20: Case 4: R2 with different number of households' data for training

As the figures show, the overall R2 increases in using the DAC-LSTM method. Figure 4.20 illustrates that the DAC-LSTM method maintains forecast accuracies above 0.97 when training samples are obtained from more than 30 households. While Persistence and ARIMA methods illustrate a linearly decreasing trend with the reduced number of households and forecast accuracy reduced from 0.881 to 0.591 (Persistence) and 0.92 to 0.587 (ARIMA). The forecast accuracy provided by Persistence

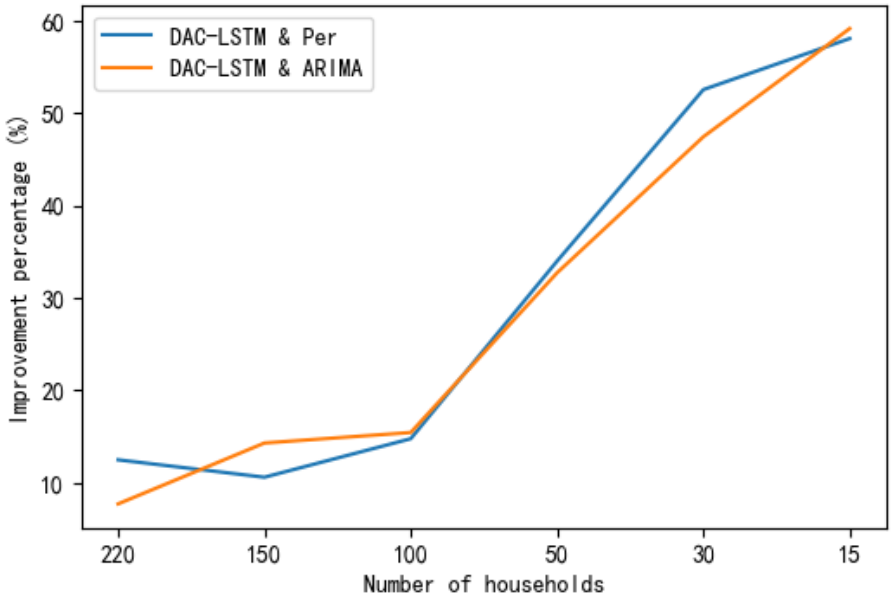


Figure 4.21: Case 4: R2 improvement with different number of households' data for training

and ARIMA methods reduces significantly when the number of households is less than 100. In addition, Figure 4.21 illustrates that the DAC-LSTM method provides 12% - 58% improvement upon the Persistence method and 7% - 59% improvement upon the ARIMA method.

4.7.2 Comparison of MAPE

Figure 4.22 and 4.23 indicate the MAPE value and percentage of improvement of DAC-LSTM compared with Persistence and ARIMA.

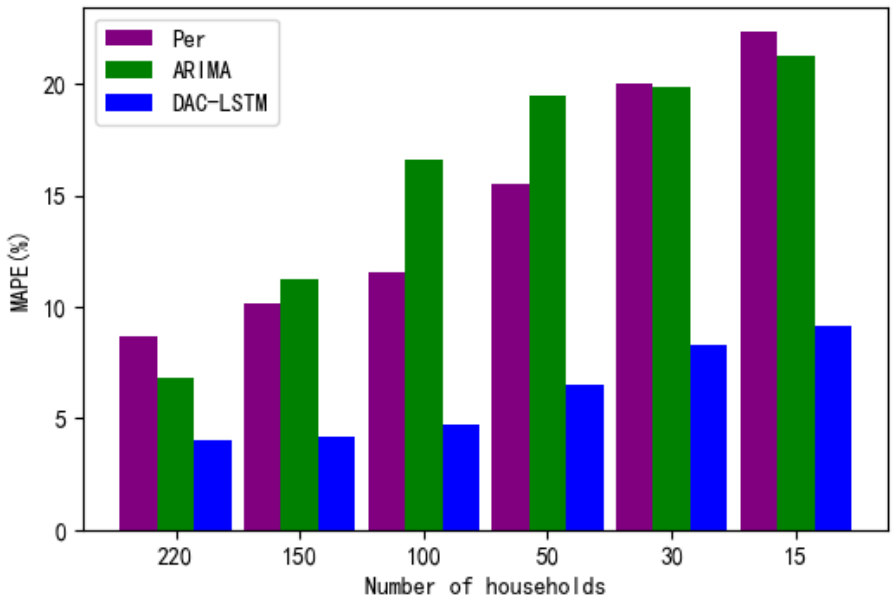


Figure 4.22: Case 4: MAPE with different number of households for training

As the figures show, the overall MAPE increases by applying the DAC-LSTM method. Figure 4.22 illustrates that the DAC-LSTM methods maintain the MAPE below 5% when the training sample contains more than 100 household data. While Persistence and ARIMA methods give an increasing error trend with the reduced number of households, none of these methods provides 5% or lower forecast error. In addition, Figure 4.23 illustrates the DAC-LSTM method provides 53% -59% and 40% - 71% improvement compared with the Persistence and ARIMA methods, respectively.

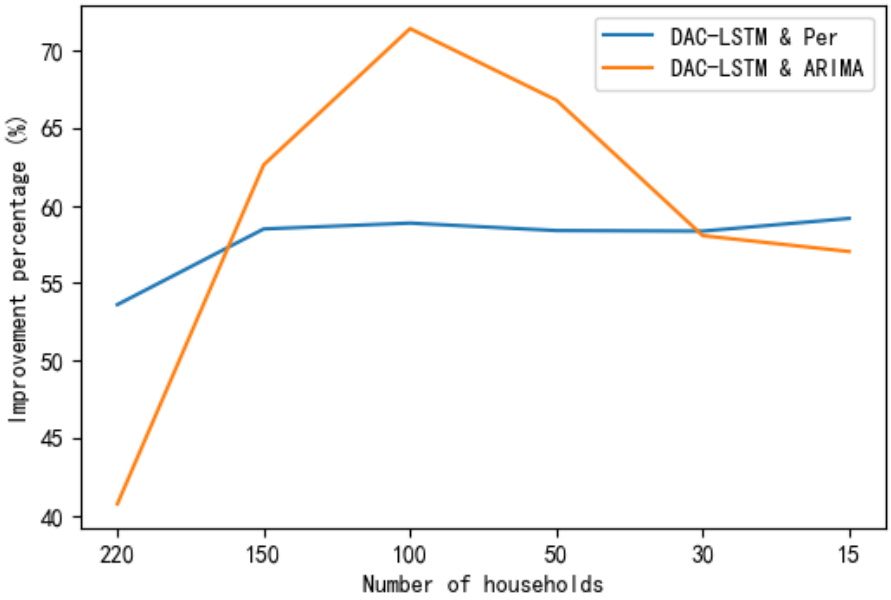


Figure 4.23: Case 4: MAPE improvement with different number of households for training

4.7.3 Comparison of Correlation

Figure 4.24 and 4.25 indicate correlation values and percentage of improvement of DAC-LSTM compared with Persistence and ARIMA methods.

As the figures show, the overall correlation was reduced by applying the DAC-LSTM method. Figure 4.24 illustrates the DAC-LSTM method maintains a correlation of forecast and actual values over 0.98 regardless of the number of households. While the correlation of ARIMA and Persistence methods shows a decreasing trend. In addition, Figure 4.25 illustrates the DAC-LSTM method provides 5% - 24% and 4% - 37% improvement compared with Persistence and ARIMA methods, respectively.

In conclusion, with the reduced number of households available for load monitoring, the DAC-LSTM provides the most convincing forecast result. $R2$ values remain over 0.95 with a minimum number of 30 households. Moreover, the MAPE remains below 5% when at least 100 households are available. Therefore, for this specific training

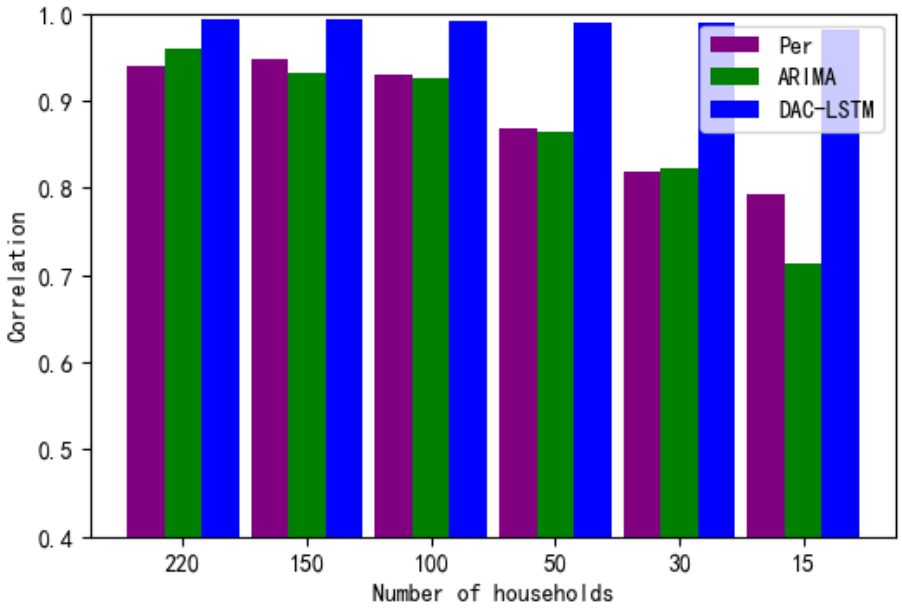


Figure 4.24: Case 4: Correlation with different number of households for training

dataset, the DAC-LSTM method provides accurate forecast results ($R2$ over 0.95 and MAPE below 5%) when the number of households is more than 100. Moreover, the DAC-LSTM method also provides a convincing result ($R2$ over 0.95 and MAPE below 6.5%) if the number of households is more than 30. This case study illustrates that the proposed DAC-LSTM method can be applied in small distribution networks or micro-grids with fewer consumers for the purpose of DSM, P2P, etc.

In addition, the improvement is significant for correlation when comparing DAC-LSTM with ARIMA and Persistence methods. This illustrates that DAC-LSTM keeps the load shape and trend the closest to the actual shape.

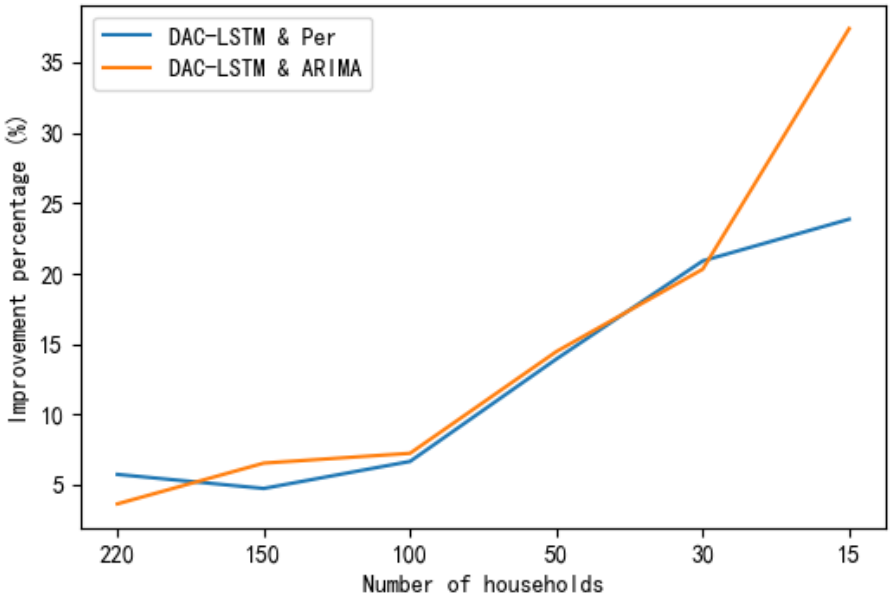


Figure 4.25: Case 4: Correlation improvement with different number of households for training

4.8 Case 5: Comparison with ELEXON UK Domestic Load

In this case study, the DAC-LSTM method is applied to the UK domestic load from ELEXON to expand the application range. As Figure 4.26 shows, the load trend from ELEXON is more smooth and shows less distortion among days compared with distribution network load data. The ELEXON dataset is recorded half-hourly and without linear interpolation before training. Table 4.1 shows the forecast accuracy from ELEXON and DAC-LSTM. The R2 of the ELEXON forecast and DAC-LSTM forecast results are 0.97 and 0.99 separately, while the DAC-LSTM method gives less average error. Figure 4.27 compares the actual and forecast result from ELEXON, LSTM, and DAC-LSTM.

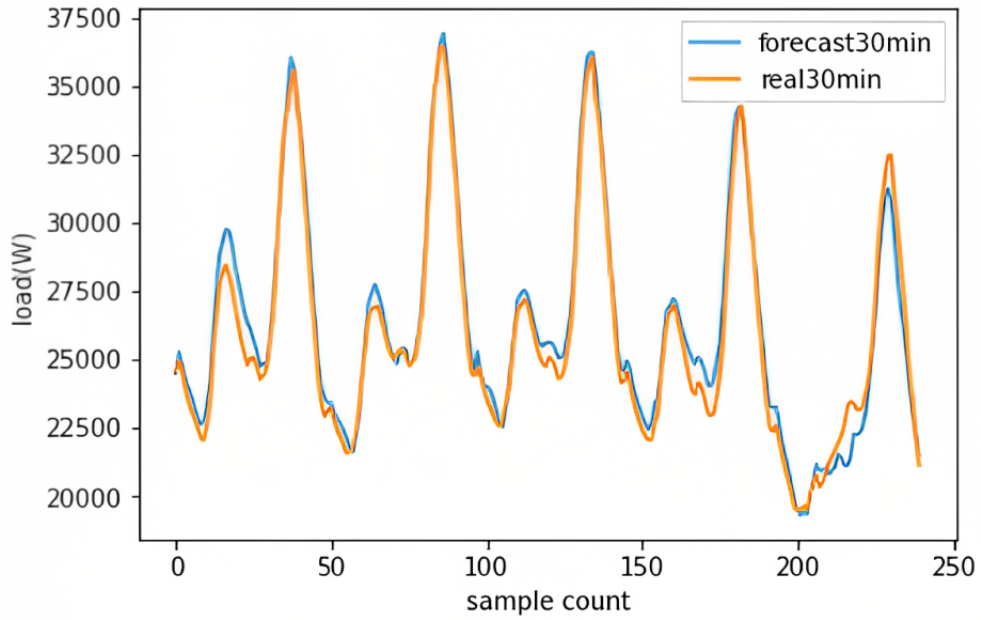


Figure 4.26: Case 5: real and forecast load from ELEXON, at 30-min time-step

Table 4.1: Case 5: UK domestic load dataset forecast result

	R2	**MAPE (%)**	RMSE (%, normalised)
ELEXON	0.97	0.17	0.13
DAC-LSTM	0.99	0.13	0.11

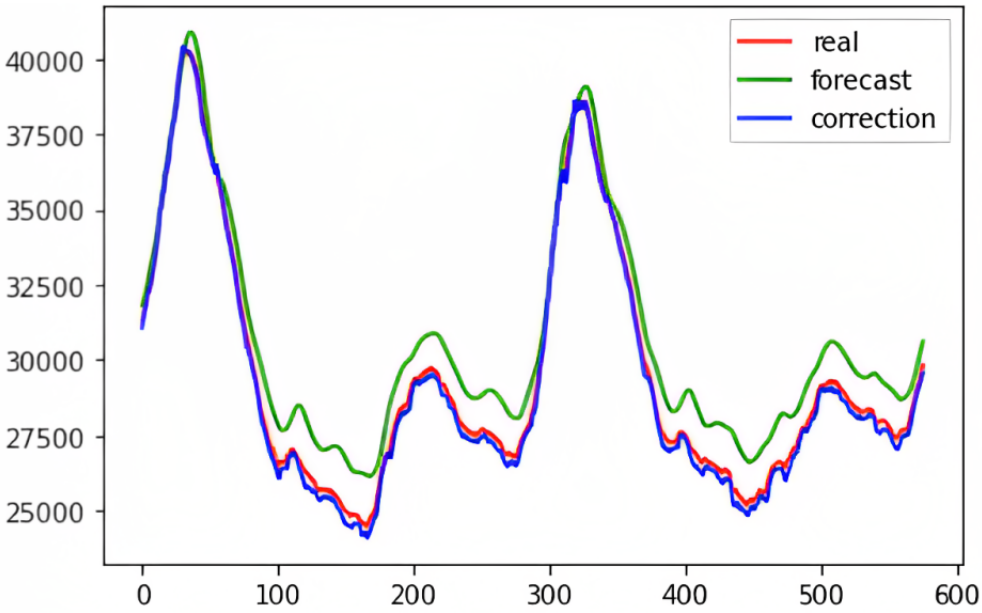


Figure 4.27: Case 5: real ELEXON recorded load, LSTM forecast result, and DAC-LSTM forecast result

4.9 Case 6: Comparison with Other Methods

The sixth case study compares the forecast result from the DAC-LSTM method with other state-of-art machine learning methods. All referred methods are compared instead of using quantitative numbers but measuring the improvements based on the benchmarks. Moreover, as discussed in the literature review chapter, the LSTM network shows advantages in long-time-series data forecast. Therefore, the forecast accuracy of the LSTM method is set as the comparison baseline (benchmark) for each method. As different publishments introduce various performance metrics, Table 4.2 summarises the most commonly used metrics, MAPE (%), as the standard. Moreover, because different methods are developed for various forecast steps, the case study only compares methods developed for a 30-minute to 2-hour forecast.

Table 4.2: Case 6: comparison among DAC-LSTM and other methods with the performance of LSTM set as the baseline

	DAC-LSTM	[164]	[165]	[166]	[167]	[90]
LSTM	5.44	2.56	5.54	25.9	2.86	4.80
Proposed	4.04	1.96	4.37	24.4	2.57	3.96
Improvement	0.257	0.234	0.211	0.058	0.101	0.175

In [164, 165, 166, 167, 90], proposed methods are HEL-LSTM, MS-CNN, ResNet-LSTM, EEMD-LSTM and CNN-LSTM respectively. It can be noticed that the hybrid methods illustrate better performance compared with the original machine learning method. Moreover, compared with other STLF methods, the DAC-LSTM method provides more improvement, 25.7%, while the others vary from 10.1% to 23.4%.

4.10 Conclusion

In this work, the proposed DAC-LSTM method is evaluated with 6 case studies, including:

1. Compared with LSTM and ARIMA for various forecast steps (cases 1 and 2). These two cases study evaluate the performance improvement upon LSTM and ARIMA methods for forecast length within 24 hours. The result shows that DAC-LSTM provides more convincing results than ARIMA within a 24-hour length forecast or LSTM within a 9-hour length forecast.
2. Compared with ARIMA and Persistence for various lengths of historical data-sets as input (case 3). This case study evaluates the performance of DAC-LSTM while the length of the input dataset is deficient. The result proves that the DAC-LSTM can be applied in the area where the load profile is not continuously monitored across the year.
3. Compared with ARIMA and Persistence for various sizes of households' data as input (case 4). This case study evaluates the performance of DAC-LSTM while the load profile variety is limited, which leads to a highly non-linear load profile. The result proves that the DAC-LSTM can be applied in areas with fewer households or utility companies.
4. Compared with ELEXON UK domestic load forecast result (case 5). This case study extends the application range of the proposed method to the transmission level. DAC-LSTM provides slightly more accurate less RMSE results than the official half-hourly load forecast result.
5. Finally, qualitatively compared with other machine learning methods (case 6). This case study qualitatively evaluates the performance of DAC-LSTM with other state-of-art machine learning methods.

The proposed method is evaluated in different scenarios based on the above case studies. These results prove the DAC-LSTM method can be utilised from the distribution level network (highly non-linear) to the domestic load level. These features offer opportunities for the electricity supplier and the grid operator to improve the existing services, DR plan, real-time pricing schemes, etc. For example, in a P2P trading system, more accurate STLF methods could reduce the Ethereum gas trading cost, therefore reducing the operation cost and improving efficiency. Also, in a network with high DG penetration, the share of renewable generation could be increased with more accurate forecast methods to reduce carbon footprint.

Chapter 5

Future Load Scenarios

5.1 Introduction

Carbon neutralisation (CN) represents that carbon dioxide (CO_2) should be compensated through technologies such as carbon capture, carbon storage, and carbon conversion within a period. The ultimate goal of CN is to achieve the net zero emission of greenhouse gases [168]. The concept of carbon neutrality was originated in Samsoe Island, Denmark in 1997 and later accepted by industries around the world. In 2003, the International Air Transport Association announce the goal of Carbon Neutrality in the Aviation Industry by 2020. Afterwards, the 1.5℃ above the temperature of the pre-industrial level increasing goal is announced in the Paris Agreement to limit global warming affections. According to the objective of the Intergovernmental Panel on Climate Change (IPCC), the middle of the 21st century is the deadline for achieving the carbon neutral target agreed upon by countries [169]. To realise the goal proposed by IPCC, countries and unions including China, The Europe Union, The United Kingdom, the United States, etc., announced their commitments to reducing carbon emissions and achieving carbon neutrality before specific years (usually between 2050 and 2060). However, the United Nations Environment Programme

(UNEP) presents that the gap between countries' targets emissions and the 1.5℃ goal is still big, as the "Emissions Gap Report 2019" released [170]. In order to reach carbon neutrality, the IPCC states the importance of a multi-pronged approach based on [171]:

- Phasing out fossil fuels, using renewable energy sources (RES);
- Facilitating behaviour change, improving energy efficiency (EE) for both supply and demand;
- Introducing negative emissions measures.

Climate change has become a major global challenge in recent decades. To achieve the emission reduction targets, the SDN is adequate to handle the increasing RES share and associated management methods with appropriate ancillary services on both supply and demand sides. In this progress, the STLF play a vital role in order to improve the system management quality, such as the renewable energy generation forecast, DSM algorithms deployment, etc. To be more specific, the load curve in the future will shift to different patterns, which will include more controllable devices, energy storage systems, renewable generation plants, etc., and the energy consumption pattern changes due to the energy policies, heating habits, travellers' behaviours, etc. Also, the importance of microgrid construction and management increases while peer-to-peer (P2P) energy trading creates a platform for all microgrid users to exchange their energy surplus and demand without the presence of a centralised authority. Therefore, dozens of studies on the relationship between energy policies and load patterns have been published in recent years and the STLF methods should also be robust in future load scenario applications.

This chapter intends to build up bottom-up future load scenarios based on existing research and policies. The DAC-LSTM is present and evaluated in the previous two chapters with 6 case studies from various aspects. The DAC-LSTM method can provide accurate forecast results, while the training dataset is limited to various forecast steps. This chapter extends this study into the future by using future load scenarios developed for different scenario years. The main differences in each scenario are the EV and e-heating penetration levels, estimated according to research and policies. The proposed method forecasts aggregated load profiles based on each scenario. Eventually, the results will be analysed, and the performance of the DAC-LSTM method will be discussed.

5.2 Policies

The United Kingdom

In 2015, “The Paris Agreement” is signed by 178 countries, including the UK. To control the climate changing problems, the UK shows great responsibility, and in June 2019, clearly stated in the “Climate Change Act” that carbon neutrality should be achieved by 2050. Moreover, the carbon neutrality target is written into the law first by the UK in developed countries. The UK is the first developed country to start carbon neutrality practices. The British Standards Institute issued the world’s first carbon neutral specification (PAS 2060) (BSI) [172]. In the industries, the UK issued a program for The British civil aviation industry to achieve carbon neutrality by 2050. More than announcing the Ten Point Plan for a Green Industrial Revolution, as the capital of the UK, London makes clear low-carbon requirements in power, construction, transportation and other sectors to reduce carbon emissions by 80% before 2050.

The European Union

Since the Paris Agreement, countries including Sweden, France, New Zealand, etc. also passed legislation pledging to achieve the carbon goal by 2050 or earlier [171, 173]. According to the ambitious climate and energy targets set by the European Union (EU), the Resolution on Climate Change of the European Parliament endorsed a net-zero GHG emissions target by 2050 in March 2019 and “urged the Member States to do the same as part of the Future of Europe”. Moreover, a target of 30% carbon reduction is set as the intermediate target for most countries, it includes the reduction in many aspects: GHG emission reduction by at least 40%, consumed energy generation from RES account for at least 32%, improved EE by at least 32.5% (comparing with business as usual). Until now, the goal is written into the draft of the European Climate Law.

China

China is the world’s largest carbon emitter, contributing to 32% of the global carbon emission by the end of 2020. The Chinese government set goal to achieve the peak CO_2 emission by 2030 on September 22, 2020. Afterwards, China sets the carbon-neutral goal, achieved by 2060. Til now, China has made remarkable achievements in controlling climate change during the 13th Five-Year Plan period. By the end of 2020, compared with the years 2005 and 2015, greenhouse emissions have been reduced significantly, by 48.1% and 18.2%, and are effectively controlled. In addition, renewable energy vehicles are growing rapidly in China and energy conservation in some areas is proceeding steadily.

The United States

On January 20, 2021, the United States government announced the return of the United States to the Paris Agreement. During the campaign, a plan called “carbon-free electricity society” is announced by Biden. This plan intends to achieve carbon-free electricity by 2035 by utilising renewable energy and carbon neutrality with a 100% clean energy economy by 2050.

5.3 Future load structure in the UK

To demonstrate the future load scenarios, several publishes, government reports, and projects have proposed future load models considering current policies, load structures, user consumption patterns, weather impacts, etc. Typically, [174] proposed a numerical modelling method for Hungary’s urban area, [175] provides a sustainable energy development roadmap for Scotland, which includes the energy consumption for different sections in 2030 and 2050. Moreover, [176, 177] analyses the future passenger transport transformation and the impact of EVs in future scenarios. Also, the European Network for Transmission System Operators Electricity (ENTSO-E) proposed a comprehensive analysis of future scenarios across the European countries [178].

5.3.1 TYNDP 2018

The 2018 Ten Year Network Development Plan (TYNDP 2018) was adopted by ENTSO-E and publicly released on 19 November 2018 after a public consultation that ended on 21 September 2018 [178]. The Plan comes at a pivotal time in delivering the European policy objective of cost-efficiently achieving a clean future for the power

system by 2050 while maintaining system security. The Plan gives a scenario analysis for most European countries for 2020, 2025, 2030 and 2040. Thus, TYNDP 2018 is referred to in this work to build up future load profiles. In TYNDP 2018, the year 2020 and 2025 has the best estimate scenarios; 2030 and 2040 scenarios have been designed with European 2050 targets as an objective. Therefore, the scenarios for 2020 and 2025 are built as bottom-up topologies, while 2030 and 2040 are top-down topologies. As of today, the year 2020 scenario is no longer required. Thus only the analysis of the best years will be considered.

The project has three scenario storylines build-up with detailed electric load, various types of generations, along with gas demand and supply, within the climate target framework sets by the EU and commodity prices. The TYNDP scenarios include the best estimate scenario before 2025 and three storylines for 2030 to 2050. The best estimate scenario is developed for the short and medium term. In this scenario, it includes a merit order sensitivity between coal and gas in 2025. While the longer storylines consider the increasing uncertainties. In addition, these scenarios are built to meet the EU's carbon neutrality targets. The scenario pathways from 2020 to 2050 are in figure 5.1.

The storylines of the TYNDP 2018 are structured as follows:

- Distributed Generation: prosumers at the centre - small-scale generation, batteries and fuel switching society engaged and empowered.
- Sustainable Transition: national regulation, emission trading schemes and subsidies, maximising the use of existing infrastructure.
- Global Climate Action: global decarbonisation, large-scale renewables development in both electricity and gas sectors.

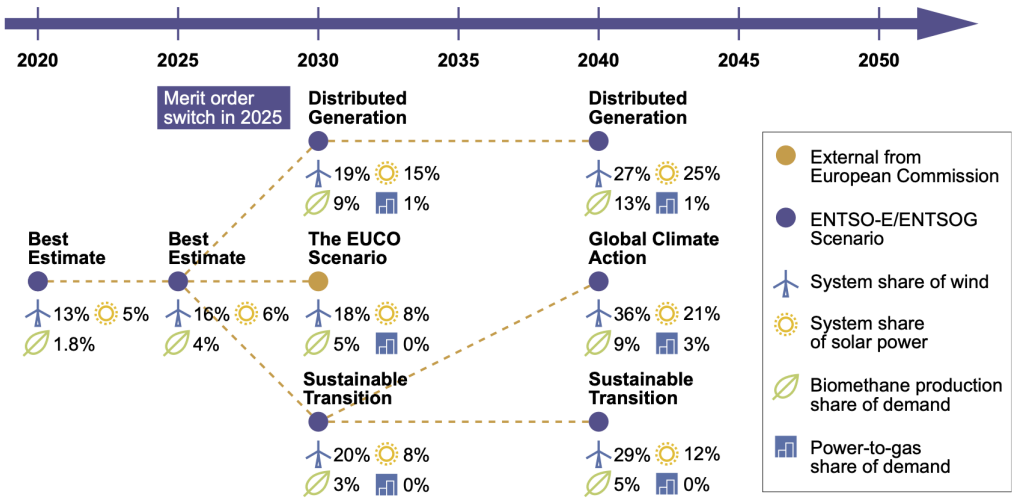


Figure 5.1: The scenario pathways from 2020 to 2050 and the RES shares [178]

- External scenario: based on EUCO 30 (European Commission) - an energy-efficient target of 30% and energy targets agreed by the European Council in 2014.

The Sustainable Transition (ST) scenario is selected as the main storyline in this project. For such a scenario, the following building key parameters and their changes are taken into consideration:

- Power generation: increase in small-scale generations, P2G, electricity demand flexibility, and heating needs in winter.
- Transport: increased electricity and gaseous usage, lower battery cost, and liquefied natural gas (LNG) for long transport.
- Heat: more electric and hybrid heat pumps, higher building efficiencies, and more district heating from combined heat and power (CHP) plants.
- Electricity demand: increased overall, residential growth reduced, peak demand reduced because of proper management.
- Gas demand: transport sector increase, residential sector decrease, required for peak demand situation (winter cold weather, etc.)

The following sections are the scenario-building process, covering the electricity demand, electric vehicle (EV) increase, and heating demand. The figures below illustrate the numerical number and trend for the European countries. These numbers and trends are considered in building the UK DN scenarios.

Electricity Demand

The ST scenario has the lowest electricity demand in 2030 and 2040 as the heating, power generation and transport sectors are predominately supplied by gas and other fuel types. The detailed comparison of different scenarios is shown in Figure 5.2.

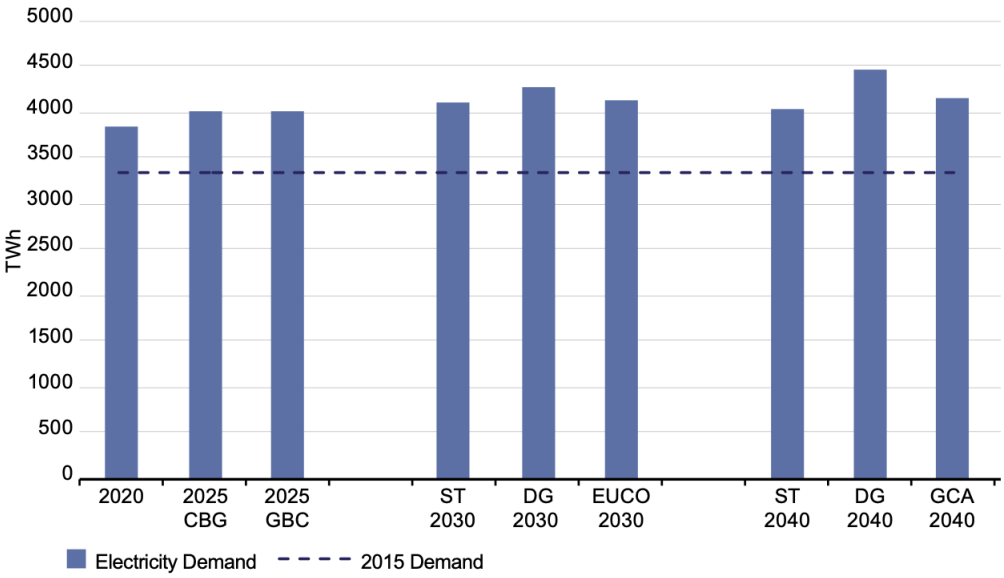


Figure 5.2: Electricity annual demand by scenario [178]

EV, HP, and Hybrid HP

In this section, EV, HP and hybrid HP represent Electric Vehicles, Heat Pumps and Hybrid Heat Pumps. The growth in electric vehicles is exponential throughout the timeline in all scenario paths, as Figure 5.3 shows:

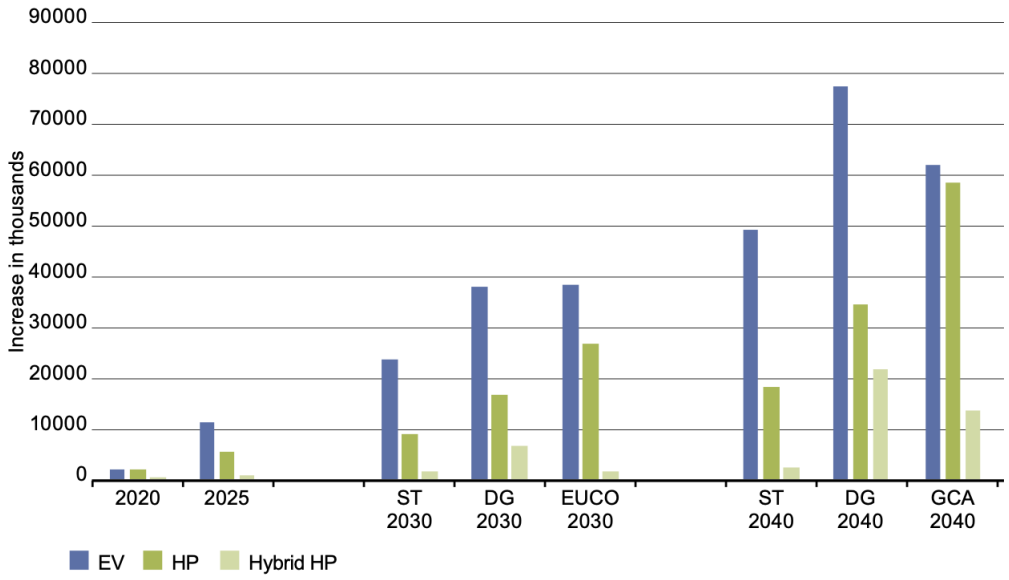


Figure 5.3: Increase in numbers of EV, e-HP and hybrid HP for each scenarios[178]

About the EV, the lowest growth is seen in the ST scenario in both 2030 and 2040 because the gas prices are at the lowest. However, the number of heat pumps significantly increases. The HP increases as the heating sector involve more electrification heating systems to contribute to the decarbonise. The hybrid HP grows slightly due to the moderate economic growth and increase in EVs.

5.3.2 UK Scenario for Distribution Network

Based on the scenario built in TYNDP 2018, the UK DN load profile is constructed considering the base load profile, heating load profile, EV charging load profile and EV penetration levels. The number of households is set to 220, which matches the TVVP dataset. The simulation data for each year starts from 26th February 2013 to 25th February 2014 (which matches the TVVP dataset). The simulated scenario years are 2030 and 2040. In addition, the assumption is made that user behaviour in these years is unchanged.

Base Load

The basic load profile inherits the TVVP dataset, which includes 220 households. According to the UK government report, as Figure 5.4 shows, the non-transport final energy consumption of heat energy used for space heating and water heating in 2013 is 7.7% (the numeric value can be found in Appendix A.1). In addition, industrial reports show that only 0.16% of cars in 2013 were registered as plug-in electric vehicles. Thus, the portion of EV and electrified heating loads in TVVP datasets are assumed to be minimum. Moreover, consumer energy usage patterns are assumed to be unchanged in future load scenarios build-up.

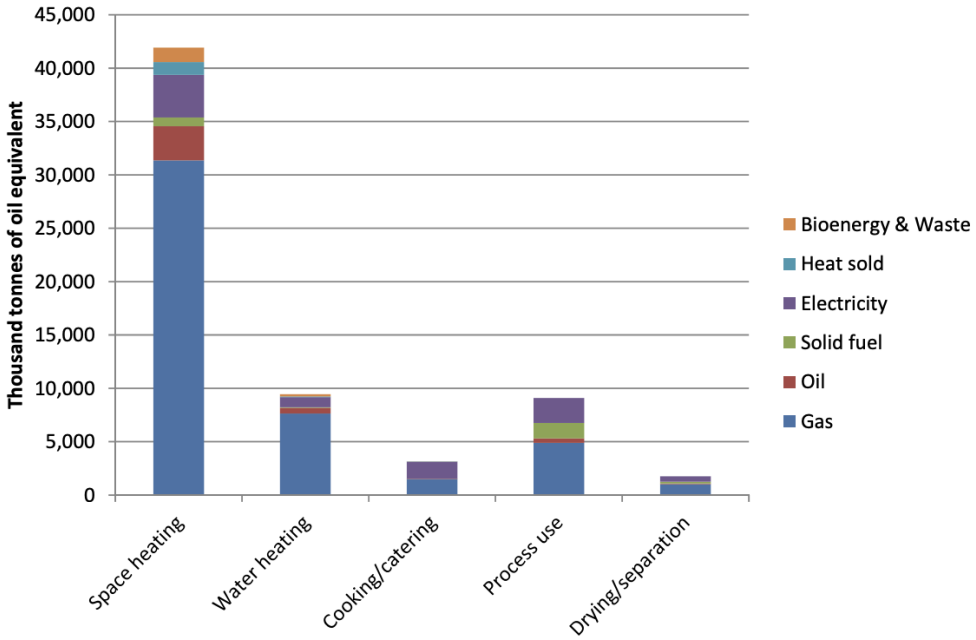


Figure 5.4: Non-transport final energy consumption of heat energy in 2013 [179]

Heating Load Profile

The heating load profile is from Data Catalogue, UKERC Energy Data Centre. The dataset includes annual heat demand data for England and Wales at Lower Layer Super Output Area (LSOA) level, before and after energy efficiency measures. The project can be found in [180]. The heating load is collected from the same area as the TVVP dataset, Bracknell and surrounding areas. The number of households providing a heating load profile is 26 (derived from Table 5.2). The heating profile is half-hourly monitored. The profile includes Air Source Heat Pump (ASHP), Ground Source Heat Pump (GSHP) and Resistance Heater (RH). Only heating types consuming electricity are used in the UK DN scenario development. Figure 5.5 illustrates the 1-year E-heating load profile of 26 households in Bracknell and surrounding areas.

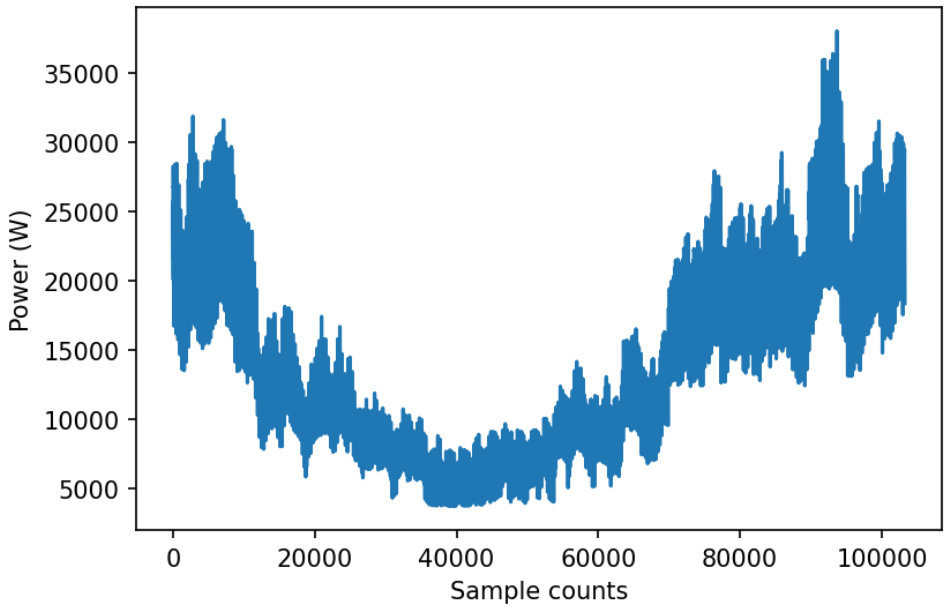


Figure 5.5: Aggregated e-heating load profile of 26 households [179]

EV Charging Profile and Penetration Level

The EV charging profile is generated using the same method in [181]. The mobility data obtained from the U.K. National Travel Survey [182] is utilised to emulate customer behaviours. This survey includes disaggregated data on means, demographics and personal travel behaviour based on diaries dating back to 2002. According to this survey, the average μ and standard deviation σ of daily mileage, arrival and departure times of individual households were determined.

The arrival time refers to the arrival after the last trip of day one and the departure time refers to the departure before the first trip of day two. In each household's emulation, the average and standard deviation of the arrival and departure times are calculated and compiled as two sets of (μ,σ). These tuples represent the EV availability of each household. The sum of all individual trips for each household on the record day is the daily trip mileage. According to the mileage from each household, the required battery charging capacity is derived. The mileage values are described and recorded as a set of (μ,σ). The following parameters in Figure 5.6 are utilised in the modelling:

			Mean μ	Variance σ^2	Skewness γ_1	Kurtosis γ_2
Arrival time	μ	empirical	1005.40	16907	-0.31	3.48
		logistic fit	1009.13	17565	0	1.20
	σ	empirical	156.44	8312	1.007	5.72
		logisitc fit	151.36	8105	0	1.20
Departure time	μ	empirical	624.56	15231	0.723	3.89
		gamma fit	624.55	15945	0.194	0.23
	σ	empirical	129.76	8571	0.994	4.83
		halfnormal fit	127.18	9232	0.995	0.87
Daily mileage	μ	empirical	21.63	192.83	1.602	7.65
		gamma fit	21.63	178.98	0.619	2.30
	σ	empirical	14.92	140.35	1.173	4.07
		exponential fit	14.92	140.35	2	6

Figure 5.6: Distribution parameters for travel patterns, and values of arrival and departure time represents minutes after midnight)

The simulation result gives the EV charging profile for 93 households and 365 days (derived from Table 5.2). Assuming the charging power is limited to 3.3kW, only 50% of cars are charged daily. In Figure 5.7, the EV charging profile includes 93 (derived from 5.2) individual cars is presented:

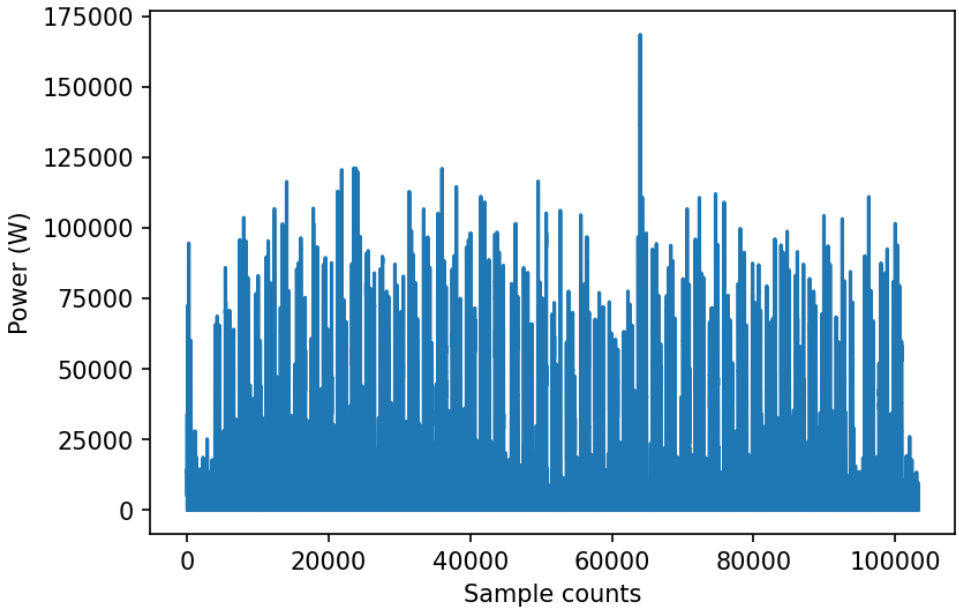


Figure 5.7: Aggregated 93 EVs charging profile for 365 days

Scenario Build-up

According to the number of households in England predicted by the UK government, which is shown in Figure 5.8, the number of households increases steadily:

As Figure 5.8 illustrates, the total number of households in the UK is estimated based on the same increase rate as England. Moreover, as previously discussed, the TYNDP 2018 Plan is referred to generate the required parameters to build up the UK DN scenario. Thus, the following Table 5.1 shows the number of households, EVs, HP, and hybrid HP calculated based on TYNDP 2018 project data source and Figure 5.8:

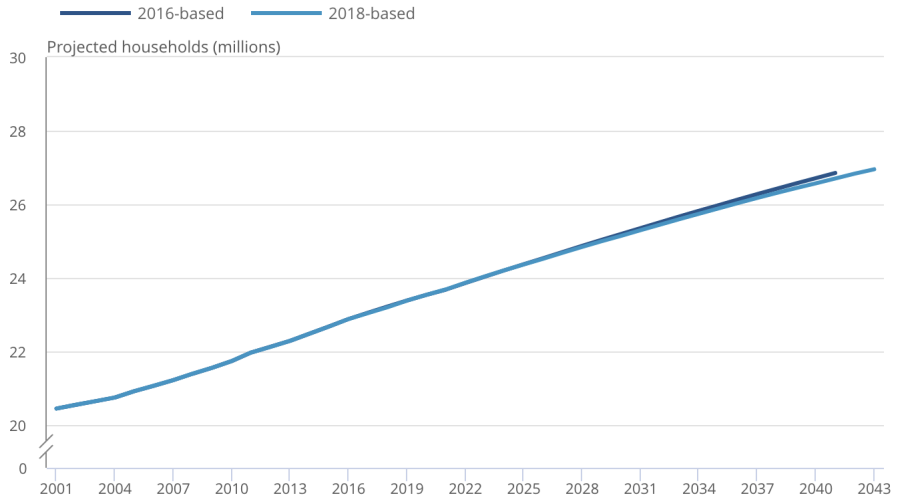


Figure 5.8: Projected number of households, 2016-based and 2018-based household projections, England, 2001 to 2043 [183]

Table 5.1: Scenario build-up parameters

Year	HH (England)	HH (UK)	EV (UK)	HP (UK)	HHP (UK)
2020	23.5	27.8	308399	153627	99016
2030	25.1	29.7	5014535	995567	1255638
2040	26.6	31.5	13324469	2165719	1505120

In Table 5.1, HH represents the number of households (million), and the number of EV, HP and hybrid HP derive from the following sources [184]. Moreover, as percentages of households that own EVs, HP and HHP can be derived from Table 5.1, parameters required for UK DN scenarios are illustrated in the following Table 5.2:

Table 5.2: EV penetration level, number of EV and HP+HPP for UK DN scenario

Year	PL	EV_{UKDN}	$HP_{UKDN} + HPP_{UKDN}$
2020	0.01	2	2
2030	0.17	37	17
2040	0.42	93	26

In Table 5.2, the EV penetration level, number of EV and HP+HPP are calculated according to the parameters given in Table 5.1, following the same share:

$$PL = EV/HH \tag{5.1}$$

$$EV_{UKDN} = PL * 220 \tag{5.2}$$

$$HP_{UKDN} + HPP_{UKDN} = (HP + HHP)/HH * 220 \tag{5.3}$$

A 24-hour load profile generated based on Section 5.3.2 is shown in Figure 5.9. It includes the base load, heat load, EV charging load and their sum, representing the full load. This figure is plotted based on scenario 2040, which represents 220 households in total, 0.42 EV penetration level (93 EVs), and 26 households who own HP or HPP.

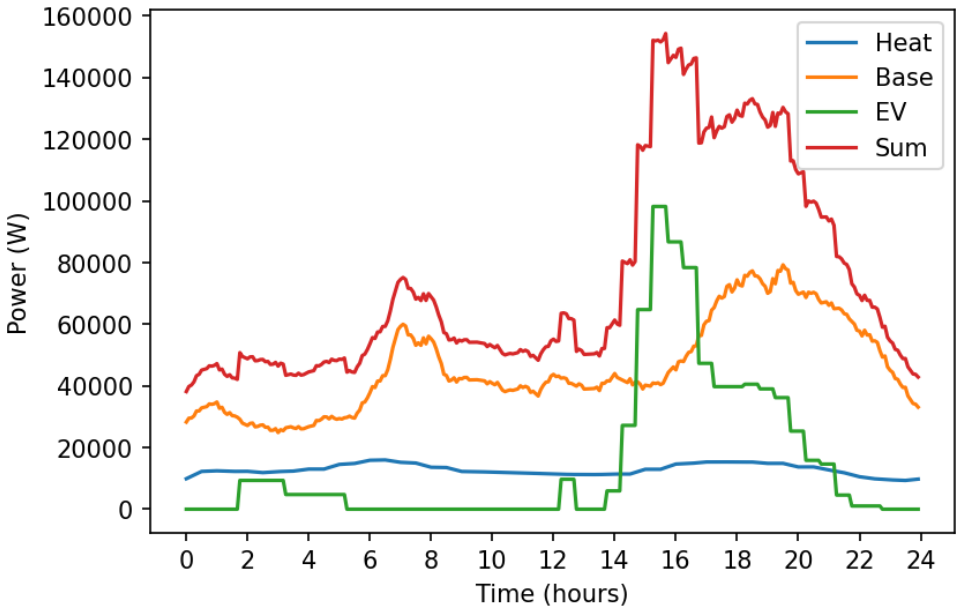


Figure 5.9: Base, Heat, EV charging and Sum load profiles for 24 hours

5.4 DAC Based on Future Scenarios

As the UK DN scenario load profile is defined in the previous section, the DAC-LSTM method is tested with this dataset. The performance metrics are the same as in Chapter 4.3, which includes $R2$, MAPE and correlation. In this section, when evaluating the DAC-LSTM, the following settings are made:

1. Each model is trained 10 times, then takes the average error value to show a steady performance of DAC-LSTM.
2. When training the LSTM model, the model parameters are adjusted according to the training data pool size, forecast steps, and forecast steps ahead.
3. Each model is trained for 500 epochs and then returns to the best epoch.
4. When needed, the input dataset is expanded to 5 min time resolution with linear interpolation and random errors. The random error is typical -1.5% to 1.5%.
5. The configuration of parameters in the peak detection is as follows: lag is 288, the threshold is 1, and influence is 0.8.
6. The configuration of parameters in the ARIMA is as follows: p, d, q = 6, 1, 2.

5.4.1 Case 1: The scenario Year 2020

The first case study compares forecast results among DAC-LSTM, ARIMA and Persistence. The load profile is generated following the UK scenario 2030, shown in Table 5.3. The EV penetration level is 0.01, representing 2 EVs in the simulation area, and the number of households utilising HP and HPP is 2.

Table 5.3: Case 1 parameters: the scenario year 2020

Year	HH	PL	EV_{UKDN}	$HP_{UKDN} + HPP_{UKDN}$
2030	220	0.01	2	2

Forecast accuracies are summarised in Table 5.4 and a 24-hour simulation result is shown in Figure 5.10.

Table 5.4: Case 1 results: the scenario year 2020

	R2	MAPE (%)	Corr
DAC-LSTM	0.989	3.67	0.993
ARIMA	0.942	5.38	0.975
Per	0.911	6.89	0.948

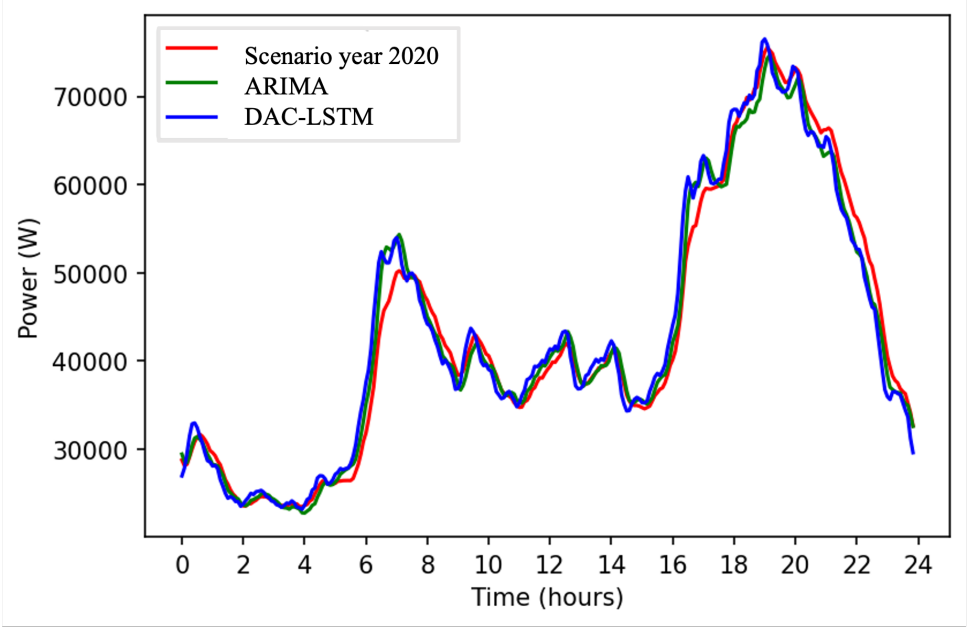


Figure 5.10: Case 1: 24-hour DAC-LSTM forecast result (the scenario year 2020)

5.4.2 Case 2: The scenario Year 2030

The second case study compares forecast results among DAC-LSTM, ARIMA and Persistence. The load profile is generated following the UK scenario 2030, shown in Table 5.5. The EV penetration level is 0.17, representing 37 EVs in the simulation area, and the number of households utilising HP and HPP is 17.

Table 5.5: Case 2 parameters: the scenario year 2030

Year	HH	PL	EV_{UKDN}	$HP_{UKDN} + HPP_{UKDN}$
2030	220	0.17	37	17

Forecast accuracies are summarised in Table 5.6 and a 24-hour simulation result is shown in Figure 5.11.

Table 5.6: Case 2 results: the scenario year 2030

	R2	MAPE (%)	Corr
DAC-LSTM	0.976	3.72	0.992
ARIMA	0.905	6.44	0.941
Per	0.927	7.08	0.939

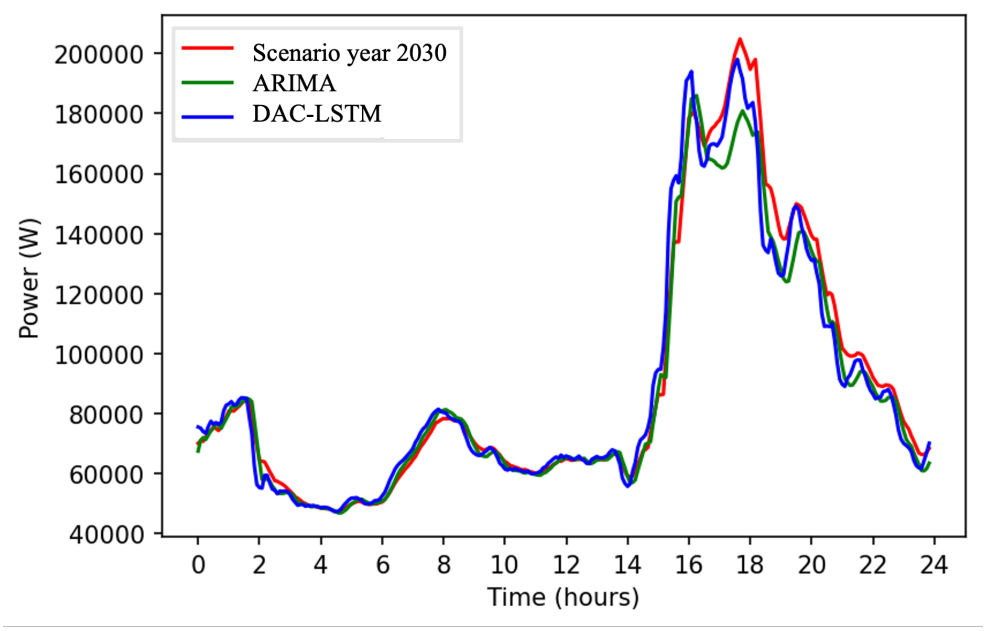


Figure 5.11: Case 2: 24-hour DAC-LSTM forecast result (the scenario year 2030)

5.4.3 Case 3: The scenario Year 2040

The third case study compares forecast results among DAC-LSTM, ARIMA and Persistence. The load profile is generated following the UK scenario 2040, shown in Table 5.7. The EV penetration level is 0.42, representing 93 EVs in the simulation area, and the number of households utilising HP and HPP is 26.

Table 5.7: Case 3 parameters: the scenario year 2040

Year	HH	PL	EV_{UKDN}	$HP_{UKDN} + HPP_{UKDN}$
2040	220	0.42	93	26

Forecast accuracies are summarised in Table 5.8 and a 24-hour simulation result is shown in Figure 5.12.

Table 5.8: Case 3 results: the scenario year 2040

	R2	MAPE (%)	Corr
DAC-LSTM	0.957	3.79	0.988
ARIMA	0.868	7.71	0.916
Per	0.908	7.04	0.954

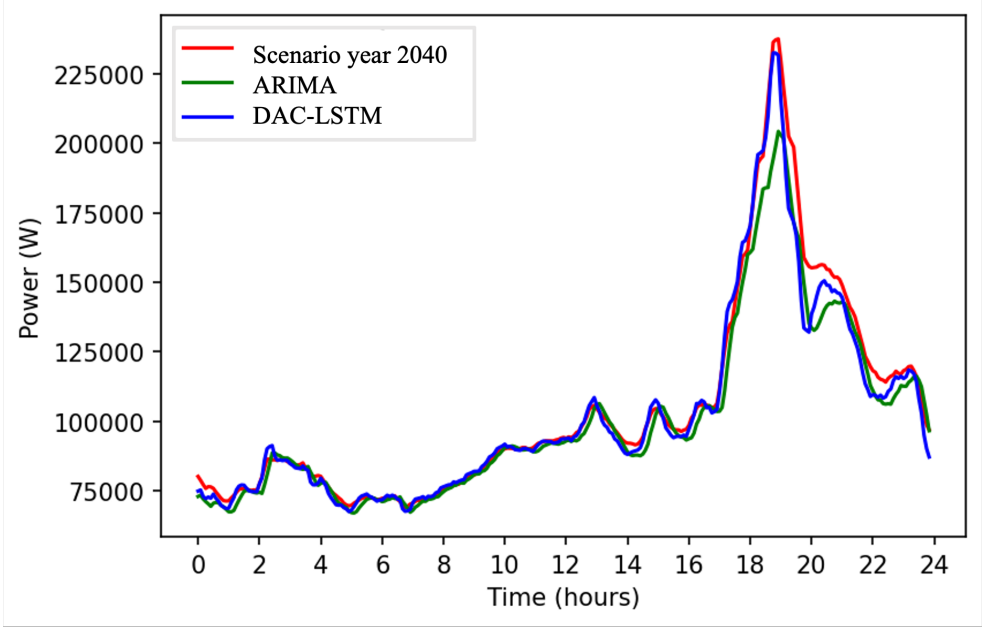


Figure 5.12: Case 3: 24-hour DAC-LSTM forecast result (the scenario year 2040)

5.4.4 Results Snalysis

Average values of base, EV, heat and sum load and their percentages in sum load are shown in Table 5.9.

As Table 5.4, 5.6, 5.8 and 5.9 indicate, with the increase of EV penetration level and e-heating demand portions, forecast accuracy of DAC-LSTM, ARIMA and Persistence methods reduce. $R2$ of DAC-LSTM reduces from 0.989 to 0.957 (3.24% reduction), of ARIMA from 0.942 to 0.868 (7.86%reduction), and of Persistence from 0.911 to

Table 5.9: Average values of Base, EV, Heat, and Sum load and their percentages in years 2020, 2030 and 2040

	Base	EV	Heat	Sum
Average (2020)	49894	304	1029	51227
Average (2030)	49894	5274	9738	64906
Average (2040)	49894	13510	14641	78046
Percentage (2020)	97.4%	0.6%	2.0%	100%
Percentage (2030)	76.9%	8.1%	15.0%	100%
Percentage (2040)	63.9%	17.3%	18.8%	100%

0.908 (0.33%reduction). While MAPE follows an opposite pattern as $R2$, of DAC-LSTM increases from 3.67% to 3.79% (3.27% increase), of ARIMA increases from 5.38% to 7.71% (43.41% increase), of Persistence increases from 6.89% to 7.08%, then reduces to 7.04% (2.18% increase).

It can be noticed that the forecast errors at the peak period, especially from the ARIMA method, increase significantly due to the increasing EV penetration levels because the ARIMA method cannot provide accurate peak and off-peak forecast results simultaneously. Persistence gives the minor reduction ($R2$)/increase (MAPE) from 2020 to 2040 as it assumes the load after 30 minutes equals the present. The DAC-LSTM method maintains the most accurate result compared with ARIMA and Persistence. Benefiting from the DAC module, the compensation parameters adapt to the peak load changes while the parameters for the off-peak load change slightly. As a result, DAC-LSTM minimises the forecast error when the peak load increases further due to the EV charging (without DSM management), as Figure 5.11 and 5.12 shows, while ARIMA lacks dynamic error compensation ability.

5.5 Conclusion

In the future, governments are taking steps to achieve carbon reduction or carbon neutralisation and, therefore, publish policies and restrictions to alter current load patterns. As reported by several organisations, there will be more electrified elements in the load, especially EVs and e-heating systems. According to existing works and government reports, this chapter builds bottom-up scenarios for the UK distribution network load in the years 2020, 2030 and 2040. In these scenarios, the EV penetration level, HP and HPP load increase. Based on existing works and datasets, bottom-up load profiles are generated to simulate different percentages of EVs, HPs and HPPs.

This work simulates a bottom-up distribution network load with 220 households, which matches the inherited baseload dataset from TVVP. The heat load derives from UKERC Energy Data Centre, which is in the same area as TVVP. The EV load is generated from the existing model with mobility data obtained from the UK National Travel Survey. Figure 5.9 shows that the evening peak load rises significantly without proper DSM algorithms. Therefore, the peak detection and DAC module in the DAC-LSTM method efficiently minimise the error at the peak. When comparing the forecast result, DAC-LSTM provides the best result among all case studies. ARIMA method was the second best in 2020, and forecast accuracy drops the most in 2040 due to the increase in EV and heat load. The Persistence method accuracy changes the least because the forecasted value equals the present. These results illustrate that the DAC-LSTM method is more robust in distribution network load forecast than existing extensively used methods.

Chapter 6

DSM with DAC-LSTM Method for End-Users Electricity Cost Reduction

6.1 Introduction

One critical aspect of DSM in the smart grid is demand response, which is the response of consumers' demands to price signals from utility companies. Demand Response (DR) allows companies to control or manage consumers' appliances directly or indirectly, through direct load control or pricing incentives. Proper DR can improve the electricity market efficiency [185]. However, the deployment of DR schemes always comes with difficulties [186]. Using the game theory framework, load-adaptive pricing was introduced decades ago. This chapter introduces a game theory approach to DR management development. In this game, the utility company and consumers reach a Nash equilibrium where the prices and demands and optimally chosen, the company maximise the revenues and the consumers minimise the electricity bills by offering available power to the company.

The novel method, DAC-LSTM, is developed in previous chapters for short-term load forecasting, especially in distribution networks. Compared with existing and commonly utilised forecast methods, the DAC-LSTM method can provide accurate forecast results, while the training dataset is limited for various forecast steps. Here the application of the DAC-LSTM method in the supply system, reducing the cost of electricity for end-users, is present in this chapter.

This chapter proposes a demand response algorithm based on the game theory stated in [187]. The game theory method models the interactions of a Stackelberg game, where companies set their prices and consumers respond by choosing their demands. In this progress, companies deliver power to consumers while consumers offer a minimum energy consumption threshold and flexible power availability. In [187], the consumption threshold is 100% which means flexible power is zero. This research modifies the method by improving the optimisation function with constraints that offer flexible power and keep the base load uninfluenced.

6.2 Methodology

6.2.1 Game Theory Preliminaries

A static N-person noncooperative game is comprised of the player set, action sets, and utility functions. Assume the play set is $\mathscr{N} := 1,\ldots,N$, where N is the number of players. The action set is $\mathscr{A}_i$ for each player The decision of player i is $a_i \in A_i$. The decisions taken by other players is a set of vector $a_{-i} := (a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_N)$.

For each player, his/her objective is to maximise the utility function $u_i(a_i, a_{-i})$. In game theory, the utility function of player i depends on both his/her actions and others. To solve the problem, an equilibrium concept is developed which is the Nash equilibrium (NE) [187]:

The action vector $\mathbf{a}^* \in \mathscr{A}_1 * \ldots * \mathscr{A}_N$ constitutes an NE for the N-person static noncooperative game in pure strategies if:

$$u_i(\mathbf{a}_i^*, \mathbf{a}_{-\mathbf{i}}^*) \geq u_i(\mathbf{a}_i, \mathbf{a}_{-\mathbf{i}}^*) \quad \forall \mathbf{a}_i \in \mathscr{A}_i,\ i \in \mathscr{N}. \tag{6.1}$$

Based on equation 6.1, the Game Theory can be improved and benefited by introducing the hierarchical structure. In this case, the players are leaders and followers. The leaders send dominant decisions and the followers respond to the leaders' decisions. This game is also called the Stackelberg game, the optimised solution concept is called the Stackelberg equilibrium. In such a game, the leaders start trading with the followers and making corresponding actions according to the followers' reactions. Assume there are K leaders and N followers, with $\mathscr{N} := 1, \ldots, N$ and $\mathscr{K} := 1, \ldots, K$, and with the action sets $(\mathscr{F}_i)_{i \in \mathscr{N}}$ and $(\mathscr{L}_j)_{j \in \mathscr{K}}$. The action vector $\mathbf{a}^* \in \mathscr{L}_1 * \ldots * \mathscr{L}_N$ is a Stackelberg equilibrium strategy for all the K leaders in pure strategies if, for each $j \in \mathscr{K}$:

$$u_j(\mathbf{a}_j^*, \mathbf{a}_{-\mathbf{j}}^*, \mathbf{b}^*(\mathbf{a}^*)) \geq u_j(\mathbf{a}_j, \mathbf{a}_{-\mathbf{j}}^*, \mathbf{b}^*(\mathbf{a}_j; a_{-j}^*))\ \forall \mathbf{a}_j \in \mathscr{L}_j \tag{6.2}$$

Where, $\mathbf{b}^*(\mathbf{a}) \in \mathscr{F}$ is the followers' optimal response. Generally, the equilibrium is also the NE. But when the coupling between diffident followers does not exist, which means the decision made by other followers has no impact on follower i, followers become independent. In this case, it is assumed that each follower is independent, therefore, for a Stackelberg game, the pair $\mathbf{a}^*, \mathbf{b}^*(\mathbf{a}^*)$ constitutes the equilibrium strategy.

6.2.2 Optimisation Problem Definition

The optimisation problem definition functions are referred to from work published in [187]. In this work, optimisation functions are modified to fit simulation purposes: the amount of electricity supply must meet the consumer's requirement each day, and the demand (available for power control) considers the forecast error factor. Besides, steps in solving the optimisation functions are from [187]. The original method plays Stackelberg games around multiple companies, while only one company is involved in this work.

Consumer Side Analysis

Because of energy scheduling and storage devices (EV), consumers may have flexibility on when to receive a certain amount of energy and postpone a certain percentage of their energy consumption (switching e-heat to gas heating). On the consumer side, the optimisation problem is defined as follows, which aims to maximise the utility of consumers:

$$u_n(\mathbf{d}_n) = \gamma_n \sum_{t \in \mathscr{T}} \ln(\zeta_n + d_n(t)) \tag{6.3}$$

Where $\gamma_n > 0$ and $\zeta_n \geq 1$are the preference parameters. In this research is set to 1. n represents consumers and companies. t is the time slot. $d_n(t)) \geq 0$ is the demand at the consumer side from the company.

The consumer-side optimisation problem is formulated as follows:

$$\max_{\mathbf{d}_n} U_n(\mathbf{d}_n) \tag{6.4}$$

$$\text{s.t.} \sum_{t \in \mathscr{T}} p(t) d_n(t) \leq B_n \tag{6.5}$$

$$\sum_{t \in \mathscr{T}} d_n(t) \geq E_n^{\min} \tag{6.6}$$

$$\sum_{t \in \mathscr{T}} d_n(t) = \mathbf{G} \tag{6.7}$$

$$d_n(t) \geq E_{min}(t) \geq 0 \ \ \forall t \in \mathscr{T}. \tag{6.8}$$

Where p is the price, B_n is the budget of the consumer, $E_n^{\min}$ and $E_{min}(t)$ denote the minimum energy need for the entire time horizon and at time slot t. Constrains above denote that the total cost for consumers must be less than budget B_n, the total demand $d_n(t)$ must be greater than the minimum requirements and at time slot t, demand $d_n(t)$ must greater than the base load requirement.

Company Side Analysis

The optimisation function for companies is to maximise the revenue:

$$\pi(\mathbf{p}) := \sum_{t \in \mathscr{T}} p(t) \sum_{n \in \mathscr{N}} d_n(\mathbf{p}, t). \tag{7}$$

The company-side optimisation problem is formulated as follows:

$$\max_{\mathbf{p}} \pi(\mathbf{p}) \tag{6.9}$$

$$\text{s.t.} \sum_{n \in \mathcal{N}} d_n(\mathbf{p}, t) \leq G(t)(1 - e(t)) \ \ \forall\, t \in \mathcal{T} \tag{6.10}$$

$$p(t) > 0 \ \ \forall\, t \in \mathcal{T} \tag{6.11}$$

$$p_{max} \geq p(t) \geq p_{min} > 0. \tag{6.12}$$

Where $e(t)$ denotes the forecast error at time slot t. Constrains above denote that the demand at price signal $\mathbf{p}$ at time slot t must be less than the power that could be supplied by the company, and the price must be between the maximum and minimum price value required by the government.

Equilibrium Strategy

Equation 6.6 can be solved by finding the associated Lagrange function 6.13, and this yields the following equation for $d_n^*(t)$ (minimising the power required, $d_n(t)$, by the consumer):

$$\begin{aligned} L_n = & \gamma_n \sum_{t \in \mathcal{T}} \ln(\zeta_n + d_n(t)) \\ & - \lambda_{n,1} \left(\sum_{t \in \mathcal{T}} p(t) d_n(t) - B_n \right) \\ & + \sum_{t \in \mathcal{T}} \lambda_{n,2}(t) d_n(t) \end{aligned} \tag{6.13}$$

$$d_n^*(t) = \frac{B_n + \sum_{h\in\mathscr{T}} p(h)\zeta_n}{KT\,p(t)} - \zeta_n \quad \forall\, t \in \mathscr{T}, \tag{6.14}$$

Where K denotes the number of companies involved in the game. When considering equation 6.10 and 6.14, the following equation can be derived (maximising the power supplied, $G(t)$ by the company) [187]:

$$B + Z\sum_{h\in\mathscr{T}} p(h) = KT\,p(t)(G(t)(1-e(t)) + Z) \;\; \forall\, t \in \mathscr{T} \tag{6.15}$$

Where $B = \sum_{n\in\mathscr{N}} B_n$ and $Z = \sum_{n\in\mathscr{N}} \zeta_n$. By solving equation 6.15, the following equation 6.16 is derived. By letting $\zeta_n = 1$ for each consumer, values of Z coincides with N. It can be found that by given G, the price $p^*(t)(G(t)(1-e(t)) + N)$ is a constant for all time slots. Thus, the power availability is inversely proportional to the prices.

$$p^*(t) = \frac{B}{G(t)(1-e(t)) + Z}\left(\frac{1}{KT - \sum_{h\in\mathscr{T}} \frac{Z}{G(h)(1-e(h))+Z}}\right) \tag{6.16}$$

6.3 Case Study

6.3.1 DSM with Scenario Year 2040

This Stackelberg game offers consumers to utilise a minimum amount of power from the power grid and offer the maximum amount of power to the system operators for load-shifting purposes. The first cast study intends to find the relationship between forecast accuracy and the performance of DSM. In this case study, the equilibrium point is achieved under the following circumstances and assumptions:

- Equation 6.14 and 6.16 are utilised to find the optimal solution.

- The aim is to evaluate the relationship between forecast accuracy and DSM outcomes. Therefore, the real load used is from Section 5.4.3.
- K = 1, which represents only one company involved.
- The load profile inherits from 5.4.3 assuming 220 households are homogeneous.
- Electric load consists of base, EV and heat loads. The base load is separated from the profile. EV and heat loads are used for management with 100% control availability, which denotes that $\sum_{n\in\mathcal{N}} d_n(\mathbf{p},t) = G(t)(1-e(t))$.
- The original price applied to the base load while incentives only provided the control-available power from consumers.
- Total power supplied by the system operator throughout a 24-hour day equals the requirements (base + EV + heat).
- The EV load is assumed to be shiftable throughout the 24-hour day as consumers may drive the car in/out of the monitoring region and receive similar control signals.

The following Figure 6.1 indicates the Stackelberg games management result. The right plot indicates real aggregated consumers' half-hourly demand, the mid plot indicates the original pricing signal and the Stackelberg game price signal, and the right plot indicates the summed billing savings for 220 households. This result is based on real load which is regarded as the baseline for the next comparison. It can be found that prices are inversely proportional to power availability.

Furthermore, the savings are compared among DAC-LSTM, ARIMA and Persistence methods. The pricing signal remains unchanged, the load profiles are derived from 5.4.3, and the control window is 30 minutes. Table 6.1 shows simulation results from various methods.

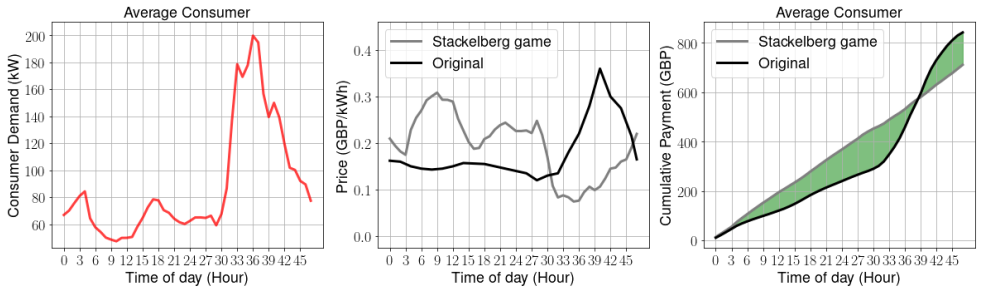


Figure 6.1: Consumer's demand (left), Stackelberg game and original prices (middle), and the cumulative original and Stackelberg games payments(right)

Table 6.1: Stackelberg game results from DAC-LSTM, ARIMA and Persistence (the scenario year 2040)

Method	Original	DAC-LSTM	ARIMA	Persistence
Billings-Max (£)	843.0	712.1	690.1	705.7
Savings-Max (£)	-	130.9	152.3	137.3
MAPE (%)	-	3.79	7.71	7.04
Savings-Opt (£)	-	87.9	52.6	58.3
Savings (%)	-	10.4%	6.2%	6.9%

Table 6.1 illustrates the original billings, which is £843 per day for 220 households. The Savings-Max represents bill reductions of each method, assuming the forecast result is 100% accurate. The MAPE is referred to as the error rate in this case study. The MAPE value of DAC-LSTM is 3.79%, which represents that the power availability of the forecast load reduces by 3.79%. As equation 6.16 shows, the term $G(t)$ and $G(h)$ reduces by corresponding MAPE value, $(1 - e(t))$. Thus, the Saving-opt is derived for each method, £87.9 (DAC-LSTM), £52.6 (ARIMA), and £58.3 (Persistence). In this study, it can be found that the ARIMA method provides the most savings if the error is not considered. However, the DAC-LSTM outperforms the other two methods when introducing the error rate because of the smaller MAPE.

6.3.2 DSM with Scenario Year 2020 and 2030

The case studies in this section intend to analysis the influences of controllable power offered by the end-users to the DSM. The following table 6.2 and 6.3 list the Stackelberg game management results with the scenario year 2020 and 2030 datasets.

Table 6.2: Stackelberg game results from DAC-LSTM, ARIMA and Persistence (the scenario year 2020)

Method	Original	DAC-LSTM	ARIMA	Persistence
Billings-Max (£)	553.3	541.1	539.6	542.3
Savings-Max (£)	-	12.2	13.7	11
MAPE (%)	-	3.67	5.38	6.89
Savings-Opt (£)	-	7.7	5.4	5.3
Savings (%)	-	1.4%	1.0%	0.9%

Table 6.3: Stackelberg game results from DAC-LSTM, ARIMA and Persistence (the scenario year 2030)

Method	Original	DAC-LSTM	ARIMA	Persistence
Billings-Max (£)	701.1	634	642.7	638.2
Savings-Max (£)	-	67.1	58.4	62.9
MAPE (%)	-	3.72	6.44	7.08
Savings-Opt (£)	-	44.1	21.5	19.8
Savings (%)	-	6.3%	3.1%	2.8%

From table 6.2 - 6.1 and table 5.9, the percentages of bill savings of DAC-LSTM methods increase from 1.4% to 10.4% and the percentages of controllable power increase from 2.6% to 36.1%. As for the ARIMA and Persistence methods, the results follow a similar trend. This result illustrates that the performance of DSM proportional to the controllable power offered by the consumers. The DSM could achieve better performance with accurate forecast methods with the increased EV penetration levels in the future power system.

6.4 Conclusion

In this Chapter, a Stackelberg Game method is implemented based on Game Theory to evaluate the value of improving forecast accuracy in deploying a demand-side management strategy. Although current research on DR is developed based on the assumption that the forecast load is 100% accurate, this work intends to study the influence factor of forecast accuracy by modifying existing DR methods' optimisation functions with more elements and constraints.

According to the Stackelberg Game proposed in [187], the equilibrium is achieved where the company reaches the most revenue, and the consumer offers the maximum power availability to the company. Further, based on the existing Stackelberg Game, this work extends the original optimisation functions and constraints to fit the experiment purposes, including first, the amount of electricity supplied must meet the consumer's requirement each day. Second, the forecast error (MAPE) reduces the controllable power offered by the consumers. Third, the power supplied by the company must meet the base load requirement for each control period.

When considering the forecast error, if the maximum power availability from consumers reduces, the price increases. Based on this fact, the case study in this chapter shows that the forecast error (in this work measured in MAPE) significantly affects the demand-side management performance. Moreover, as the equations and the second case study illustrates, the utility bill saving is proportional to the controllable power offered by the consumers.

In conclusion, the achievements in this chapter are:

1. A Stackelberg Game DR strategy is introduced, considering the forecast error as a factor that affects the DR performance.

2. The case study proves that forecast accuracy affects the performance of the DR strategy. For example, in case study one, the forecast error (3.79%) reduces the utility savings from £130.9 to £87.9.
3. The portion of controllable power offered by the consumers influences the utility bill savings proportionally. For example, the percentages of bill savings of DAC-LSTM methods increase from 1.4% to 10.4% and the percentages of controllable power increase from 2.6% to 36.1%.

Chapter 7

Conclusions

7.1 Thesis Summary

This thesis first presents an adaptive DAC-LSTM forecast method, validated with the dataset from Thames Valley Vision Project, for short-term electricity demand forecast. Compared with other existing methods, this DAC-LSTM method solves the following problems, including the forecast results are not accurate for peak and off-peak periods simultaneously, the datasets utilised for training the model are sufficient while the collectable datasets in the real world are usually limited. To be more specific, the DAC-LSTM method dynamically distinguishes peak and off-peak hours using the Z-score method, applying different module parameters to the separated peak and off-peak loads, which other methods barely have. Further, the proposed method introduces the novel DAC block to compensate for forecast errors with various activation functions and dynamic parameters according to different forecast steps. The greater the current-to-average forecast error ratio or closer to present time stamps, the larger the compensation factors. Besides, the factor caps are set to prevent the model from over-compensation conditions. Finally, the sensitivity of introduced parameters is analysed, providing the performance of the developed method under different parameter values.

In chapter 4, the DAC-LSTM method is first compared with widely utilised methods, ARIMA and Persistence, and then with other state-of-art machine learning (hybrid) methods. In the case study section, the proposed method is first evaluated with ARIMA and LSTM methods for different forecast steps, from 30 minutes to 24 hours. The result illustrates the advantages and improvements of DAC-LSTM in short-term load forecast. Second, the proposed DAC-LSTM method is evaluated by varying the input historical dataset length from 550 days to 14 days and the input dataset collected from 220 households to 15 households. These case studies demonstrate the adaptation of the DAC-LSTM in STLF when the training dataset is insufficient or incoherent. First, the DAC-LSTM method can provide accurate forecast results when the trained model lacks historical data or sources or unfits the load pattern. Further, the proposed method is applied to UK domestic load and shows convincing results compared with the UK half-hourly forecast result provided by ELEXON, and this proves that the application of the DAC-LSTM method ranges from distribution level network (highly non-linear) to domestic load level. Finally, the proposed method is evaluated qualitatively with other machine learning methods by comparing the accuracy improvement based on LSTM. These features offer opportunities for the electricity supplier and the grid operator to improve the existing services, DR plan, real-time pricing schemes, etc. For example, in a P2P trading system, more accurate STLF methods could reduce the Ethereum gas trading cost, therefore reducing the operation cost and improving efficiency. Also, in a network with high DG penetration, the share of renewable generation could be increased with more accurate forecast methods to reduce carbon footprint.

In chapter 5, as demand profiles in the future are altered because of the net zero carbon roadmap announced by governments, the proposed DAC-LSTM method should be evaluated with future load scenarios. In this chapter, several scenarios are built based on published research and policies for the UK distribution network in the year

2020, 2030 and 2040. These scenarios consist of bottom-up base, EV and e-heating load profiles at different EV and heat penetration levels. For each simulation year, the forecast result accuracy reduces as the EV and e-heating penetration level increase, while the DAC-LSTM obtains the highest accuracy because it compensates for the most error generated at the peak due to EV charging loads.

In chapter 6, the value of utilising the forecast method with higher accuracy is assessed. A Stackelberg Game is introduced to find the equilibrium point for suppliers and consumers. Equations prove that the price is inversely proportional to the power availability. In this study, power availability represents the power offered by consumers to suppliers for demand-side management. In addition, this work also considers forecast error as a vital factor in the game. As the result demonstrates, the lower forecast error rate significantly reduces the utility bill for consumers.

Overall, this thesis starts with the method development, followed by comprehensive case studies to evaluate the performance of the DAC-LSTM method under several situations. The conclusion is that the DAC-LSTM method shows its values in the distribution network short-term load forecast.

7.2 Thesis Statement

Machine learning techniques can be utilised to develop robust forecast methods which provide accurate forecast results for peak and off-peak load, and adapt to smaller network applications. It is proved to be correct by the DAC-LSTM method developed in chapter 3 and evaluations in chapters 4-6.

7.3 Potential Impact of the Research

This thesis presents a complete, step-by-step modelling and evaluation framework to build a short-term load forecast method, which can be utilised in power networks (especially distribution networks) where the historical training dataset is limited or incoherent for various reasons. Each chapter presents a different part of the method development or evaluation procedure, and they are all linked together, presenting an adaptive forecast method with wide application purposes and for further research.

First is the DAC-LSTM short-term load forecast method. The DAC-LSTM forecast method is developed based on the LSTM network and provides dynamic error compensations in forecasting load, especially in peak and off-peak load simultaneously forecasting. The novel method utilises historical data collected from the network and enhances the ability to forecast when the historical dataset is insufficient. Moreover, the DAC-LSTM method distinguishes the peak and off-peak periods dynamically and provides accurate forecast results for both periods. The sensitivity analysis validates the method under various parameter sets with TVVP and ACN datasets. With sufficient case studies, the proposed method outperforms widely used ARIMA and Persistence forecast methods and is competitive compared with other hybrid machine learning methods. As system operators such DNOs, the accurate peak-hour forecast result will reduce the backup generation cost. Moreover, the DAC-LSTM gives accurate forecast results when the training dataset is limited, therefore, DSM alternatives can be deployed to smaller networks, such as micro-grid, community P2P trading networks, etc. This will bring more options for system operators to improve their alternatives.

The second is the UK distribution network scenarios. Bottom-up UK distribution network load profiles are first generated for the scenario years 2020, 2030 and 2040. The purposes are to evaluate the developed DAC-LSTM method with future load profiles that are highly non-linear, with higher EV and e-heating penetration levels. These scenarios are built based on existing research and policies and consist of base, EV and e-heating loads. However, the implications of this work extend much further. For example, more elements can be added, and the wetload can be separated from the base load. Academic researchers could be benefited in control algorithm development by these profiles.

The third is the value of utilising an accurate load forecast method in demand-side management. The novelty of this work is to first introduce forecast error factors and future load scenarios into demand-side management. This work studies the implementation value of utilising better forecast methods. Moreover, the utility bill reduction proves that DAC-LSTM outperforms other forecast methods. When the control centre deploys the DSM alternatives, the consideration of forecast errors will improve the outcomes, therefore reducing the unexpected cost, for example, increased uncontrollable power will weaken the preset control algorithms.

7.4 Limitations

Although the proposed short-term load forecast method can provide accurate forecast results in various situations, there are still a few limitations of the work. The forecast method is, first, validated with Thames Valley Vision Project and bottom-up built future scenario load profiles, which is sufficient for novel method development but still requires massive real-world tests before application. Second, the proposed DAC-LSTM method uses big data and machine learning techniques, which require massive

data collected from advanced monitoring devices and massive computation ability for model training. Finally, the method to create a high-resolution dataset is linear interpolation and adding random errors. This actual real-world 5-minute resolution dataset might slightly deviate from this research.

When evaluating the performances of the proposed method, the residential load demand is the most complex and contains plenty of uncertainties and variables. Future load scenarios are obtained based on existing research and policies, which cannot consider all factors and make precise forecasts. In addition, future scenarios only consist of base, EV and e-heating load as load profile elements, assuming the user energy consumption behaviours are slightly altered. However, built scenarios still can present the overall trend of the future demand profiles, which contain high EV and e-heating penetration levels.

Moreover, the value of DAC-LSTM for demand-side management is demonstrated by a Stackelberg Game which achieves a Nash equilibrium for only one company and several consumers by optimising the available power offered by consumers and energy prices from the company. However, the real-world situation is much more complicated. The real-world game might involve multiple companies which supply electricity generated by different types of renewable energy and price, and different shiftable power availability. In this work, only one company with a day-ahead price signal is involved, and the shiftable power percentage is assumed to be 100%. However, the model simplification has minor effects in yielding the experimental result. The utility bill saving illustrates that improving forecast accuracy helps bill reduction.

7.5 Scope for Expansion of Research

Though the proposed method has good performance and values for network management, the research can be expanded:

1. The proposed method is designed for short-term load forecast, especially 30 minutes. The forecast steps and accuracy can be varied and improved with further research, such as forming a hybrid forecast method with other approaches, taking different input dataset analysing procedures, etc.
2. This thesis focuses on the short-term load forecast, while the mid and long-term load forecast require further work.
3. When considering the impact on the network, for example, high EV penetration may cause voltage regulation problems that require further demand response actions. The forecast method should consider these actions as input, leaving room for improvement in future work.
4. The load profile of future scenarios might contain more elements further base, EV and e-heating loads. The wetload can be separated from the base load, which contains the controllable home appliances. This requires more detailed available system operation data or simulation models.
5. The demand-side management might consider the effects and profits of distributed generations or other renewable energy sources. For example, the value of higher forecast accuracy could be investigated with various penetration levels of distributed generations, the share of different renewable energy sources, etc.
6. Further, a power network model could be introduced for network simulation. Therefore, research such as voltage regulation, optimal power flow, fault analysis, etc., could be carried out with an accurate forecast method and demand-side management.

Appendix A

Appendix-1

Table A.1: Case 1: comparison Between DAC-LSTM and LSTM Methods with forecast steps from half hour to 24 hours

Step (0.5 hour)	R2	Corr	MAPE	R2	Corr	MAPE
1	0.991	0.994	4.04	0.901	0.956	7.92
2	0.958	0.979	4.6	0.854	0.937	9.91
3	0.898	0.949	7.61	0.826	0.929	10.87
4	0.861	0.933	9.27	0.795	0.922	11.95
5	0.829	0.92	10.23	0.76	0.908	12.7
6	0.804	0.91	11.43	0.734	0.904	14.18
7	0.786	0.901	11.78	0.728	0.9	13.99
8	0.772	0.894	12.35	0.719	0.895	14.2
9	0.76	0.887	12.92	0.719	0.896	14.49
10	0.735	0.876	13.71	0.682	0.879	15.19
11	0.743	0.875	13.39	0.717	0.889	14.23
12	0.738	0.873	13.61	0.703	0.882	14.56
13	0.719	0.862	14.34	0.679	0.867	15.22
14	0.726	0.866	14.14	0.691	0.873	14.79
15	0.737	0.869	14.01	0.71	0.877	14.4
16	0.716	0.86	14.17	0.695	0.866	13.99
17	0.733	0.871	13.62	0.712	0.88	13.66
18	0.733	0.871	14.24	0.711	0.884	14.19
19	0.738	0.874	13.73	0.723	0.887	13.41

20	0.738	0.875	13.74	0.729	0.894	13.36
21	0.738	0.878	13.21	0.732	0.898	12.93
22	0.729	0.879	12.92	0.73	0.901	12.7
23	0.706	0.871	13.4	0.717	0.899	13.11
24	0.713	0.872	13.25	0.73	0.899	12.73
25	0.702	0.873	13.02	0.734	0.904	12.05
26	0.71	0.874	12.92	0.738	0.899	11.98
27	0.709	0.871	13.01	0.755	0.903	11.52
28	0.721	0.88	12.27	0.749	0.902	11.66
29	0.714	0.879	12.5	0.759	0.905	10.93
30	0.711	0.875	12.9	0.725	0.894	12.49
31	0.718	0.877	12.5	0.757	0.905	11.31
32	0.705	0.866	13.34	0.762	0.908	11.14
33	0.688	0.86	13.91	0.741	0.906	11.66
34	0.693	0.857	14.33	0.748	0.903	11.46
35	0.695	0.855	14.83	0.76	0.903	10.94
36	0.713	0.864	14.31	0.763	0.906	11.1
37	0.703	0.858	15.18	0.764	0.903	10.95
38	0.719	0.867	14.63	0.77	0.905	10.73
39	0.724	0.868	14.73	0.782	0.907	10.33
40	0.731	0.871	14.05	0.766	0.906	11.03
41	0.729	0.871	14.26	0.769	0.905	10.69
42	0.737	0.873	14.31	0.779	0.906	10.58
43	0.731	0.873	14.04	0.772	0.906	10.36
44	0.735	0.877	13.54	0.772	0.91	10.38
45	0.747	0.879	13.08	0.779	0.913	10.58
46	0.745	0.879	13.39	0.787	0.912	10.19
47	0.743	0.877	13.42	0.798	0.915	9.88
48	0.76	0.882	12.4	0.788	0.914	10.75

Table A.2: Case 2: Comparison Between DAC-LSTM and ARIMA Methods with forecast steps from half hour to 12 hours

Step (0.5 hour)	DAC-LSTM			ARIMA		
	R2	Corr	MAPE (%)	R2	Corr	MAPE (%)
1	0.991	0.994	4.04	0.92	0.923	6.82
2	0.958	0.979	4.6	0.912	0.918	7.01
3	0.898	0.949	7.61	0.894	0.904	8.87
4	0.86	0.933	9.27	0.815	0.907	9.78
5	0.829	0.92	10.23	0.799	0.904	11.42
6	0.804	0.91	11.43	0.786	0.885	12.5
7	0.786	0.901	11.78	0.777	0.881	14.1
8	0.772	0.894	12.35	0.752	0.876	14.6
9	0.76	0.888	12.93	0.759	0.864	15.2
10	0.735	0.877	13.71	0.744	0.877	14.5
11	0.743	0.876	13.39	0.717	0.867	14.7
12	0.738	0.873	13.62	0.734	0.869	14.9
13	0.719	0.862	14.34	0.741	0.857	15.3
14	0.726	0.867	14.15	0.665	0.863	14.4
15	0.737	0.87	14.02	0.703	0.85	14.7
16	0.716	0.86	14.17	0.712	0.865	15.3
17	0.733	0.871	13.63	0.712	0.86	14.9
18	0.733	0.872	14.25	0.743	0.859	15.3
19	0.738	0.874	13.73	0.705	0.866	14.2
20	0.739	0.875	13.74	0.7	0.865	15.5
21	0.738	0.879	13.22	0.721	0.862	13.5
22	0.73	0.88	12.92	0.701	0.858	14.2
23	0.707	0.871	13.41	0.683	0.873	13.7
24	0.713	0.873	13.26	0.695	0.849	14.9
25	0.702	0.873	13.03	0.681	0.86	15.3
26	0.71	0.874	12.93	0.697	0.864	17.3
27	0.709	0.872	13.02	0.711	0.862	17.5

28	0.722	0.881	12.28	0.703	0.862	17.4
29	0.715	0.879	12.5	0.712	0.868	18.4
30	0.712	0.876	12.91	0.693	0.863	19.3
31	0.719	0.877	12.53	0.667	0.862	20.6
32	0.705	0.867	13.35	0.689	0.859	19.8
33	0.689	0.86	13.91	0.673	0.877	21.7
34	0.694	0.857	14.34	0.642	0.854	20.9
35	0.696	0.856	14.84	0.627	0.858	21.5
36	0.714	0.865	14.31	0.658	0.849	22.5
37	0.704	0.859	15.18	0.633	0.865	23.8
38	0.719	0.868	14.63	0.619	0.851	22.2
39	0.725	0.869	14.73	0.588	0.853	23.4
40	0.731	0.872	14.05	0.591	0.853	23.7
41	0.729	0.871	14.27	0.489	0.855	24.1
42	0.737	0.873	14.32	0.57	0.852	25.3
43	0.732	0.874	14.05	0.518	0.848	25.9
44	0.735	0.877	13.55	0.511	0.85	26.4
45	0.747	0.88	13.09	0.525	0.849	26.8
46	0.745	0.879	13.39	0.471	0.855	30.1
47	0.744	0.877	13.43	0.415	0.847	29.2
48	0.76	0.882	12.4	0.43	0.843	29.8

Thousand tonnes of oil equivalent

End use	Gas	Oil	Solid fuel	Electricity	Heat sold[1]	Bioenergy & Waste[1]	Total
Space heating	22,865	2,314	667	2,148	52	682	28,728
Water heating	6,139	454	45	653	-	203	7,494
Cooking/catering	617	-	-	490	-	-	1,108
Heat total	29,622	2,769	712	3,291	52	884	37,330
Lighting and appliances	-	-	-	6,464	-	-	6,464
Overall total	29,622	2,769	712	9,755	52	884	43,794

[1] Heat sold and waste is included in this table. Assumptions have been made that, in the domestic and industry sector, all uses for these two sources is for space and water heating.

Figure A.1: Domestic energy consumption by fuel and end use, 2013 [179]

Table A.3: Case 3: R2

	R2		
Amount of days	Per.	ARIMA	DAC-LSTM
550	0.881	0.920	0.991
500	0.867	0.891	0.990
450	0.913	0.909	0.991
400	0.864	0.833	0.991
350	0.898	0.892	0.990
300	0.898	0.874	0.975
250	0.883	0.897	0.989
200	0.893	0.894	0.990
150	0.898	0.891	0.998
100	0.876	0.873	0.958
50	0.850	0.825	0.926
30	0.795	0.794	0.904
14	0.824	0.820	0.875

Table A.4: Case 3: MAPE

	MAPE(%)		
Amount of days	Per.	ARIMA	DAC-LSTM
550	8.71	6.82	4.04
500	8.05	7.76	4.46
450	9.29	7.02	4.49
400	8.76	8.83	4.72
350	10.41	7.61	4.77
300	8.00	7.92	5.00
250	7.91	7.36	5.84
200	8.50	7.42	5.95
150	8.29	7.77	6.12
100	7.21	6.95	5.78
50	8.81	8.38	6.86
30	9.95	9.87	7.62
14	7.75	7.75	7.91

Table A.5: Case 3: Correlation

	Correlation		
Amount of days	Per.	ARIMA	DAC-LSTM
550	0.940	0.959	0.994
500	0.934	0.942	0.993
450	0.956	0.961	0.996
400	0.932	0.914	0.995
350	0.949	0.946	0.995
300	0.949	0.939	0.993
250	0.941	0.947	0.994
200	0.947	0.944	0.988
150	0.949	0.953	0.991
100	0.938	0.943	0.992
50	0.925	0.913	0.997
30	0.897	0.899	0.997
14	0.913	0.913	0.995

Table A.6: Case 4: R2

	R2		
Amount of households	Per.	ARIMA	DAC-LSTM
220	0.881	0.920	0.991
150	0.896	0.867	0.991
100	0.860	0.855	0.987
50	0.736	0.743	0.986
30	0.636	0.658	0.970
15	0.591	0.587	0.934

Table A.7: Case 4: MAPE

	MAPE(%)		
Amount of households	Per.	ARIMA	DAC-LSTM
220	8.71	6.82	4.04
150	10.12	11.24	4.20
100	11.55	16.62	4.75
50	15.55	19.49	6.47
30	19.98	19.84	8.32
15	22.34	21.23	9.12

Table A.8: Case 4:Correlation

	Correlation		
Amount of households	Per.	ARIMA	DAC-LSTM
220	0.940	0.923	0.994
150	0.948	0.932	0.993
100	0.930	0.925	0.992
50	0.868	0.864	0.989
30	0.818	0.822	0.989
15	0.792	0.714	0.981

Bibliography

[1] J. J. Grainger, *Power system analysis*. McGraw-Hill, 1999.

[2] V. A. Evangelopoulos, P. S. Georgilakis, and N. D. Hatziargyriou, "Optimal operation of smart distribution networks: A review of models, methods and future research," *Electric Power Systems Research*, vol. 140, pp. 95–106, 2016.

[3] K. E. Antoniadou-Plytaria, I. N. Kouveliotis-Lysikatos, P. S. Georgilakis, and N. D. Hatziargyriou, "Distributed and decentralized voltage control of smart distribution networks: Models, methods, and future research," *IEEE Transactions on smart grid*, vol. 8, no. 6, pp. 2999–3008, 2017.

[4] L. Ge, Y. Li, Y. Li, J. Yan, and Y. Sun, "Smart distribution network situation awareness for high-quality operation and maintenance: A brief review," *Energies*, vol. 15, no. 3, p. 828, 2022.

[5] S. Ma, B. Xu, H. Gao *et al.*, "An earth fault locating method in feeder automation system by examining correlation of transient zero mode currents," *Automation of Electric Power Systems*, vol. 32, no. 7, pp. 48–52, 2008.

[6] S. Ruj and A. Nayak, "A decentralized security framework for data aggregation and access control in smart grids," *IEEE transactions on smart grid*, vol. 4, no. 1, pp. 196–205, 2013.

[7] A. Muñoz, E. F. Sánchez-Úbeda, A. Cruz, and J. Marín, "Short-term forecasting in power systems: a guided tour," in *Handbook of power systems II*. Springer, 2010, pp. 129–160.

[8] M. Kanagawa and T. Nakata, “Assessment of access to electricity and the socio-economic impacts in rural areas of developing countries,” *Energy policy*, vol. 36, no. 6, pp. 2016–2029, 2008.

[9] S. Liu, A. Zeng, K. Lau, C. Ren, P.-w. Chan, and E. Ng, “Predicting long-term monthly electricity demand under future climatic and socioeconomic changes using data-driven methods: A case study of hong kong,” *Sustainable Cities and Society*, vol. 70, p. 102936, 2021.

[10] F. Riva, H. Ahlborg, E. Hartvigsson, S. Pachauri, and E. Colombo, “Electricity access and rural development: Review of complex socio-economic dynamics and causal diagrams for more appropriate energy modelling,” *Energy for Sustainable Development*, vol. 43, pp. 203–223, 2018.

[11] J. Hu and A. V. Vasilakos, “Energy big data analytics and security: challenges and opportunities,” *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2423–2436, 2016.

[12] D. S. Markovic, D. Zivkovic, I. Branovic, R. Popovic, and D. Cvetkovic, “Smart power grid and cloud computing,” *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 566–577, 2013.

[13] X. He, T. Huang, C. Li, H. Che, and Z. Dong, “A recurrent neural network for optimal real-time price in smart grid,” *Neurocomputing*, vol. 149, pp. 608–612, 2015.

[14] M. Q. Raza and A. Khosravi, “A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings,” *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, 2015.

[15] M. Giriraj and S. Muthu, “A cloud computing methodology for industrial automation and manufacturing execution system.” *Journal of Theoretical & Applied Information Technology*, vol. 52, no. 3, 2013.

[16] D. Allinson, K. N. Irvine, J. L. Edmondson, A. Tiwary, G. Hill, J. Morris, M. Bell, Z. G. Davies, S. K. Firth, J. Fisher *et al.*, "Measurement and analysis of household carbon: The case of a uk city," *Applied Energy*, vol. 164, pp. 871–881, 2016.

[17] J. M. Chen, "Carbon neutrality: toward a sustainable future," *The Innovation*, vol. 2, no. 3, 2021.

[18] S. A. A. Kazmi, M. K. Shahzad, A. Z. Khan, and D. R. Shin, "Smart distribution networks: A review of modern distribution concepts from a planning perspective," *Energies*, vol. 10, no. 4, p. 501, 2017.

[19] C. Lueken, P. M. Carvalho, and J. Apt, "Distribution grid reconfiguration reduces power losses and helps integrate renewables," *Energy Policy*, vol. 48, pp. 260–273, 2012.

[20] U. Sultana, A. B. Khairuddin, M. Aman, A. Mokhtar, and N. Zareen, "A review of optimum dg placement based on minimization of power losses and voltage stability enhancement of distribution system," *Renewable and Sustainable Energy Reviews*, vol. 63, pp. 363–378, 2016.

[21] X. Shen, M. Shahidehpour, Y. Han, S. Zhu, and J. Zheng, "Expansion planning of active distribution networks with centralized and distributed energy storage systems," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 1, pp. 126–134, 2016.

[22] I. Konstantelos, S. Giannelos, and G. Strbac, "Strategic valuation of smart grid technology options in distribution networks," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1293–1303, 2016.

[23] S. M. Nosratabadi, R.-A. Hooshmand, and E. Gholipour, "A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 341–363, 2017.

[24] T. Bashir, C. Haoyong, M. F. Tahir, and Z. Liqiang, "Short term electricity load forecasting using hybrid prophet-lstm model optimized by bpnn," *Energy Reports*, vol. 8, pp. 1678–1686, 2022.

[25] A. Y. Saber and A. R. Alam, "Short term load forecasting using multiple linear regression for big data," in *2017 IEEE symposium series on computational intelligence (SSCI)*. IEEE, 2017, pp. 1–6.

[26] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on power systems*, vol. 16, no. 1, pp. 44–55, 2001.

[27] J. Mohammed, S. Bahadoorsingh, N. Ramsamooj, and C. Sharma, "Performance of exponential smoothing, a neural network and a hybrid algorithm to the short term load forecasting of batch and continuous loads," in *2017 IEEE Manchester PowerTech*. IEEE, 2017, pp. 1–6.

[28] W. J. Stevenson, M. Hojati, and J. Cao, *Operations management*. McGraw-Hill Education Chicago-USA, 2014.

[29] Z. Zheng, H. Chen, and X. Luo, "A kalman filter-based bottom-up approach for household short-term load forecast," *Applied Energy*, vol. 250, pp. 882–894, 2019.

[30] S. Sharma, A. Majumdar, V. Elvira, and E. Chouzenoux, "Blind kalman filtering for short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4916–4919, 2020.

[31] G. Shilpa and G. Sheshadri, "Arimax model for short-term electrical load forecasting," *International Journal Of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 4, 2019.

[32] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[33] E. Mele *et al.*, "A review of machine learning algorithms used for load forecasting at microgrid level," in *Sinteza 2019-International Scientific Conference on Information Technology and Data Related Research*. Singidunum University, 2019, pp. 452–458.

[34] L. Nguyen, "Tutorial on support vector machine," *Appl. Comput. Math*, vol. 6, pp. 1–15, 2017.

[35] M. Mohandes, "Support vector machines for short-term electrical load forecasting," *International Journal of Energy Research*, vol. 26, no. 4, pp. 335–345, 2002.

[36] M. AboGaleela, M. El-Marsafawy, M. El-Sobki *et al.*, "Optimal scheme with load forecasting for demand side management (dsm) in residential areas," *Energy and Power Engineering*, vol. 5, no. 04, p. 889, 2013.

[37] S. Sharma, R. K. Agrawal, and M. M. Tripathi, "Synergism of recurrent neural network and fuzzy logic for short term energy load forecasting," in *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2020, pp. 165–169.

[38] G. J. Tsekouras, N. D. Hatziargyriou, and E. N. Dialynas, "An optimized adaptive neural network for annual midterm energy forecasting," *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 385–391, 2006.

[39] J. Llanos, D. Sáez, R. Palma-Behnke, A. Núñez, and G. Jiménez-Estévez, "Load profile generator and load forecasting for a renewable based microgrid using self organizing maps and neural networks," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–8.

[40] L. Wu, C. Kong, X. Hao, and W. Chen, "A short-term load forecasting method based on gru-cnn hybrid neural network model," *Mathematical Problems in Engineering*, vol. 2020, 2020.

[41] Y. Xuan, W. Si, J. Zhu, Z. Sun, J. Zhao, M. Xu, and S. Xu, "Multi-model fusion short-term load forecasting based on random forest feature selection and hybrid neural network," *IEEE Access*, vol. 9, pp. 69 002–69 009, 2021.

[42] C. Wang, J. Zhang, L. Tian, L. Xue, Y. Zheng, and L. Liu, "Short-term load forecasting based on kprototypes clustering and random forest," in *2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2)*. IEEE, 2021, pp. 1226–1230.

[43] A. Moradzadeh, S. Zakeri, M. Shoaran, B. Mohammadi-Ivatloo, and F. Mohammadi, "Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms," *Sustainability*, vol. 12, no. 17, p. 7076, 2020.

[44] J. Moon, Y. Kim, M. Son, and E. Hwang, "Hybrid short-term load forecasting scheme using random forest and multilayer perceptron," *Energies*, vol. 11, no. 12, p. 3283, 2018.

[45] G. Heinemann, D. Nordmian, and E. Plant, "The relationship between summer weather and summer loads-a regression analysis," *IEEE Transactions on Power Apparatus and Systems*, no. 11, pp. 1144–1154, 1966.

[46] J. Xie, T. Hong, and J. Stroud, "Long-term retail energy forecasting with consideration of residential customer attrition," *IEEE transactions on smart grid*, vol. 6, no. 5, pp. 2245–2252, 2015.

[47] M. A. Hammad, B. Jereb, B. Rosi, D. Dragan *et al.*, "Methods and models for electric load forecasting: a comprehensive review," *Logist. Sustain. Transp*, vol. 11, no. 1, pp. 51–76, 2020.

[48] S. R. Khuntia, J. L. Rueda, and M. A. van Der Meijden, "Forecasting the load of electrical power systems in mid-and long-term horizons: A review," *IET Generation, Transmission & Distribution*, vol. 10, no. 16, pp. 3971–3977, 2016.

[49] R. Billinton and D. Huang, "Effects of load forecast uncertainty on bulk electric system reliability evaluation," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 418–425, 2008.

[50] L. Wu, M. Shahidehpour, and T. Li, "Stochastic security-constrained unit commitment," *IEEE Transactions on power systems*, vol. 22, no. 2, pp. 800–811, 2007.

[51] T. Hong, "Spatial load forecasting using human machine co-construct intelligence framework," *Master thesis*, 2008.

[52] H. L. Willis, *Spatial electric load forecasting*. CRC Press, 2002.

[53] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *International Journal of Forecasting*, vol. 30, no. 2, pp. 357–363, 2014.

[54] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, vol. 1, pp. 125–151, 2014.

[55] P. Pinson, "Wind energy: Forecasting challenges for its operational management," *Statistical Science*, vol. 28, no. 4, pp. 564–585, 2013.

[56] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[57] R. Patel, M. R. Patel, and R. V. Patel, "A review: Introduction and understanding of load forecasting," *J. Appl. Sci. Comput.*, vol. 4, no. 4, pp. 1449–1457, 2019.

[58] T. Hong, *Short term electric load forecasting*. North Carolina State University, 2010.

[59] D. Bunn and E. D. Farmer, "Comparative models for electrical load forecasting," 1985.

[60] P. Singh, P. Dwivedi, and V. Kant, “A hybrid method based on neural network and improved environmental adaptation method using controlled gaussian mutation with real parameter for short-term load forecasting,” *Energy*, vol. 174, pp. 460–477, 2019.

[61] Y.-S. Lee and L.-I. Tong, “Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming,” *Knowledge-Based Systems*, vol. 24, no. 1, pp. 66–72, 2011.

[62] T. Ahmed, D. H. Vu, K. M. Muttaqi, and A. P. Agalgaonkar, “Load forecasting under changing climatic conditions for the city of sydney, australia,” *Energy*, vol. 142, pp. 911–919, 2018.

[63] N. Amjady, “Short-term hourly load forecasting using time-series modeling with peak load estimation capability,” *IEEE Transactions on power systems*, vol. 16, no. 3, pp. 498–505, 2001.

[64] G. P. Zhang, “Time series forecasting using a hybrid arima and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[65] M. Khashei and M. Bijari, “A novel hybridization of artificial neural networks and arima models for time series forecasting,” *Applied Soft Computing*, vol. 11, no. 2, pp. 2664–2675, 2011.

[66] K.-Y. Chen and C.-H. Wang, “A hybrid sarima and support vector machines in forecasting the production values of the machinery industry in taiwan,” *Expert Systems with Applications*, vol. 32, no. 1, pp. 254–264, 2007.

[67] H. J. Sadaei, P. C. d. L. e Silva, F. G. Guimaraes, and M. H. Lee, “Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series,” *Energy*, vol. 175, pp. 365–377, 2019.

[68] N. Amral, C. Ozveren, and D. King, “Short term load forecasting using multiple linear regression,” in *2007 42nd International universities power engineering conference*. IEEE, 2007, pp. 1192–1198.

[69] T. Hong, M. Gui, M. E. Baran, and H. L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," in *Ieee pes general meeting*. IEEE, 2010, pp. 1–8.

[70] K. M. Kam, *Stationary and non-stationary time series prediction using state space model and pattern-based approach*. The University of Texas at Arlington, 2014.

[71] M. Lineesh, K. Minu, and C. J. John, "Analysis of nonstationary nonlinear economic time series of gold price: A comparative study," in *International Mathematical Forum*, vol. 5, no. 34. Citeseer, 2010, pp. 1673–1683.

[72] D. Janardhanan and E. Barrett, "Cpu workload forecasting of machines in data centers using lstm recurrent neural networks and arima models," in *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, 2017, pp. 55–60.

[73] A. Dubey and S. Santoso, "Electric vehicle charging on residential distribution systems: Impacts and mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

[74] S. Kumar, S. Mishra, and S. Gupta, "Short term load forecasting using ann and multiple linear regression," in *2016 second international conference on computational intelligence & communication technology (cict)*. IEEE, 2016, pp. 184–186.

[75] Y. Li, J. Che, and Y. Yang, "Subsampled support vector regression ensemble for short term electric load forecasting," *Energy*, vol. 164, pp. 160–170, 2018.

[76] A. Lahouar and J. B. H. Slama, "Day-ahead load forecast using random forest and expert input selection," *Energy Conversion and Management*, vol. 103, pp. 1040–1051, 2015.

[77] L.-C. Ying and M.-C. Pan, "Using adaptive network based fuzzy inference system to forecast regional electricity loads," *Energy Conversion and Management*, vol. 49, no. 2, pp. 205–211, 2008.

[78] M. Barman, N. D. Choudhury, and S. Sutradhar, "A regional hybrid goa-svm model based on similar day approach for short-term load forecasting in assam, india," *Energy*, vol. 145, pp. 710–720, 2018.

[79] J. Faraji, A. Ketabi, H. Hashemi-Dezaki, M. Shafie-Khah, and J. P. Catalao, "Optimal day-ahead scheduling and operation of the prosumer by considering corrective actions based on very short-term load forecasting," *IEEE Access*, vol. 8, pp. 83 561–83 582, 2020.

[80] Z. Xiao, S.-J. Ye, B. Zhong, and C.-X. Sun, "Bp neural network with rough set for short term load forecasting," *Expert Systems with Applications*, vol. 36, no. 1, pp. 273–279, 2009.

[81] S. Panigrahi, Y. Karali, and H. Behera, "Time series forecasting using evolutionary neural network," *International Journal of Computer Applications*, vol. 75, no. 10, 2013.

[82] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," *International journal of systems science*, vol. 33, no. 1, pp. 23–34, 2002.

[83] C. Zhuo and S. Long-Xiang, "Short-term electrical load forecasting based on deep learning lstm networks," *Electronic Technology*, vol. 10, no. 39, pp. 122–125, 2018.

[84] A. Baliyan, K. Gaurav, and S. K. Mishra, "A review of short term load forecasting using artificial neural network models," *Procedia Computer Science*, vol. 48, pp. 121–125, 2015.

[85] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[86] C. Yang, W. Jiang, and Z. Guo, "Time series data classification based on dual path cnn-rnn cascade network," *IEEE Access*, vol. 7, pp. 155 304–155 312, 2019.

[87] W. Zhang, H. Quan, O. Gandhi, R. Rajagopal, C.-W. Tan, and D. Srinivasan, "Improving probabilistic load forecasting using quantile regression nn with skip connections," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5442–5450, 2020.

[88] M. Dong and L. Grumbach, "A hybrid distribution feeder long-term load forecasting method based on sequence prediction," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 470–482, 2019.

[89] S. Basodi, C. Ji, H. Zhang, and Y. Pan, "Gradient amplification: An efficient way to train deep neural networks," *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 196–207, 2020.

[90] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network," *Energies*, vol. 11, no. 12, p. 3493, 2018.

[91] M. Z. A. Bhotto, R. Jones, S. Makonin, and I. V. Bajić, "Short-term demand prediction using an ensemble of linearly-constrained estimators," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3163–3175, 2021.

[92] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, "An overview and comparative analysis of recurrent neural networks for short term load forecasting," *arXiv preprint arXiv:1705.04378*, 2017.

[93] Z. Wang and R. S. Srinivasan, "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 796–808, 2017.

[94] L. Xiao, J. Wang, Y. Dong, and J. Wu, "Combined forecasting models for wind energy forecasting: A case study in china," *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 271–288, 2015.

[95] A. Abdoos, M. Hemmati, and A. A. Abdoos, "Short term load forecasting using a hybrid intelligent method," *Knowledge-Based Systems*, vol. 76, pp. 139–147, 2015.

[96] J.-r. Dong, C.-y. Zheng, G.-y. Kan, M. Zhao, J. Wen, and J. Yu, "Applying the ensemble artificial neural network-based hybrid data-driven model to daily total load forecasting," *Neural Computing and Applications*, vol. 26, no. 3, pp. 603–611, 2015.

[97] F. He, J. Zhou, Z.-k. Feng, G. Liu, and Y. Yang, "A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with bayesian optimization algorithm," *Applied energy*, vol. 237, pp. 103–116, 2019.

[98] U. B. Tayab, A. Zia, F. Yang, J. Lu, and M. Kashif, "Short-term load forecasting for microgrid energy management system using hybrid hho-fnn model with best-basis stationary wavelet packet transform," *Energy*, vol. 203, p. 117857, 2020.

[99] S. Kouhi and F. Keynia, "A new cascade nn based method to short-term load forecast in deregulated electricity market," *Energy Conversion and Management*, vol. 71, pp. 76–83, 2013.

[100] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 392–401, 2006.

[101] H. Takeda, Y. Tamura, and S. Sato, "Using the ensemble kalman filter for electricity load forecasting and analysis," *Energy*, vol. 104, pp. 184–198, 2016.

[102] M. Barman and N. B. D. Choudhury, "Season specific approach for short-term load forecasting based on hybrid fa-svm and similarity concept," *Energy*, vol. 174, pp. 886–896, 2019.

[103] N. B. Adam, M. Elahee, and M. Dauhoo, "Forecasting of peak electricity demand in mauritius using the non-homogeneous gompertz diffusion process," *Energy*, vol. 36, no. 12, pp. 6763–6769, 2011.

[104] Y. Li, Y.-Q. Bao, B. Yang, C. Chen, and W. Ruan, "Modification method to deal with the accumulation effects for summer daily electric load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 913–918, 2015.

[105] M. Dahl, A. Brun, O. S. Kirsebom, and G. B. Andresen, "Improving short-term heat load forecasts with calendar and holiday data," *Energies*, vol. 11, no. 7, p. 1678, 2018.

[106] S. B. Taieb and R. J. Hyndman, "A gradient boosting approach to the kaggle load forecasting competition," *International journal of forecasting*, vol. 30, no. 2, pp. 382–394, 2014.

[107] Z. Guo, K. Zhou, X. Zhang, and S. Yang, "A deep learning model for short-term power load and probability density forecasting," *Energy*, vol. 160, pp. 1186–1200, 2018.

[108] A. Kaur, H. T. Pedro, and C. F. Coimbra, "Impact of onsite solar generation on system load demand forecast," *Energy conversion and management*, vol. 75, pp. 701–709, 2013.

[109] H. Takeda, "Short-term ensemble forecast for purchased photovoltaic generation," *Solar Energy*, vol. 149, pp. 176–187, 2017.

[110] H. Hasan, M. R. Munawar, and R. H. Siregar, "Neural network-based solar irradiance forecast for peak load management of grid-connected microgrid with photovoltaic distributed generation," in *2017 International Conference on Electrical Engineering and Informatics (ICELTICs)*. IEEE, 2017, pp. 87–90.

[111] T. Anwar, B. Sharma, K. Chakraborty, and H. Sirohia, "Introduction to load forecasting," *Int. J. Pure Appl. Math*, vol. 119, no. 15, pp. 1527–1538, 2018.

[112] S. R. Khuntia, B. W. Tuinema, J. L. Rueda, and M. A. van der Meijden, "Time-horizons in the planning and operation of transmission networks: an overview," *IET Generation, Transmission & Distribution*, vol. 10, no. 4, pp. 841–848, 2016.

[113] N. A. Mohammed and A. Al-Bazi, "An adaptive backpropagation algorithm for long-term electricity load forecasting," *Neural Computing and Applications*, vol. 34, no. 1, pp. 477–491, 2022.

[114] J. Sowinski, "The impact of the selection of exogenous variables in the anfis model on the results of the daily load forecast in the power company," *Energies*, vol. 14, no. 2, p. 345, 2021.

[115] K. Zhu, J. Geng, and K. Wang, "A hybrid prediction model based on pattern sequence-based matching method and extreme gradient boosting for holiday load forecasting," *Electric Power Systems Research*, vol. 190, p. 106841, 2021.

[116] S. Taheri, M. Jooshaki, and M. Moeini-Aghtaie, "Long-term planning of integrated local energy systems using deep learning algorithms," *International Journal of Electrical Power & Energy Systems*, vol. 129, p. 106855, 2021.

[117] M. Ghiassi, D. K. Zimbra, and H. Saidane, "Medium term system load forecasting with a dynamic artificial neural network model," *Electric power systems research*, vol. 76, no. 5, pp. 302–316, 2006.

[118] F. S. Oliveira, C. Ruiz, and A. J. Conejo, "Contract design and supply chain coordination in the electricity industry," *European Journal of Operational Research*, vol. 227, no. 3, pp. 527–537, 2013.

[119] S. M. Islam, S. M. Al-Alawi, and K. A. Ellithy, "Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network," *Electric Power Systems Research*, vol. 34, no. 1, pp. 1–9, 1995.

[120] E. Gonzalez-Romera, M. A. Jaramillo-Moran, and D. Carmona-Fernandez, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Transactions on power systems*, vol. 21, no. 4, pp. 1946–1953, 2006.

[121] N. Amjady, D. Farrokhzad, and M. Modarres, "Optimal reliable operation of hydrothermal power systems with random unit outages," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 279–287, 2003.

[122] J. P. Carvallo, P. H. Larsen, A. H. Sanstad, and C. A. Goldman, "Long term load forecasting accuracy in electric utility integrated resource planning," *Energy Policy*, vol. 119, pp. 410–422, 2018.

[123] T. A. Farrag and E. E. Elattar, "Optimized deep stacked long short-term memory network for long-term load forecasting," *IEEE Access*, vol. 9, pp. 68 511–68 522, 2021.

[124] R. K. Agrawal, F. Muchahary, and M. M. Tripathi, "Long term load forecasting with hourly predictions based on long-short-term-memory networks," in *2018 IEEE Texas Power and Energy Conference (TPEC)*. IEEE, 2018, pp. 1–6.

[125] S. Tondolo de Miranda, A. Abaide, M. Sperandio, M. M. Santos, and E. Zanghi, "Application of artificial neural networks and fuzzy logic to long-term load forecast considering the price elasticity of electricity demand," *International Transactions on Electrical Energy Systems*, vol. 28, no. 10, p. e2606, 2018.

[126] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and dsm in smart grid: A review," *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 1311–1322, 2016.

[127] H. T. Haider, O. H. See, and W. Elmenreich, "A review of residential demand response of smart grid," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 166–178, 2016.

[128] U. Energy, "Benefits of demand response in electricity markets and recommendations for achieving them," *no. February*, pp. 1–122, 2006.

[129] R. de Sá Ferreira, L. A. Barroso, P. R. Lino, M. M. Carvalho, and P. Valenzuela, "Time-of-use tariff design under uncertainty in price-elasticities of electricity demand: A stochastic optimization approach," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2285–2295, 2013.

[130] K. Herter, "Residential implementation of critical-peak pricing of electricity," *Energy policy*, vol. 35, no. 4, pp. 2121–2130, 2007.

[131] S. Iqbal, M. Sarfraz, M. Ayyub, M. Tariq, R. K. Chakrabortty, M. J. Ryan, and B. Alamri, "A comprehensive review on residential demand side management strategies in smart grid environment," *Sustainability*, vol. 13, no. 13, p. 7170, 2021.

[132] S. A. Mansouri, E. Nematbakhsh, M. S. Javadi, A. R. Jordehi, M. Shafie-khah, and J. P. Catalão, "Resilience enhancement via automatic switching considering direct load control program and energy storage systems," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2021, pp. 1–6.

[133] Y. Chen, S. Mei, F. Zhou, S. H. Low, W. Wei, and F. Liu, "An energy sharing game with generalized demand bidding: Model and properties," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2055–2066, 2019.

[134] S. Mohseni, A. C. Brent, S. Kelly, W. N. Browne, and D. Burmester, "Modelling utility-aggregator-customer interactions in interruptible load programmes using non-cooperative game theory," *International Journal of Electrical Power & Energy Systems*, vol. 133, p. 107183, 2021.

[135] T. Vesselényi, O. Moldovan, C. Bungau, and L. Csokmai, "A survey on soft computing techniques used in intelligent building control," *Recent Innov. Mechatron*, vol. 1, 2014.

[136] G. Feng, “A survey on analysis and design of model-based fuzzy control systems,” *IEEE Transactions on Fuzzy systems*, vol. 14, no. 5, pp. 676–697, 2006.

[137] R. Lu, S. H. Hong, and M. Yu, “Demand response for home energy management using reinforcement learning and artificial neural network,” *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6629–6639, 2019.

[138] G. Hafeez, K. S. Alimgeer, Z. Wadud, I. Khan, M. Usman, A. B. Qazi, and F. A. Khan, “An innovative optimization strategy for efficient energy management with day-ahead demand response signal and energy consumption forecasting in smart grid using artificial neural network,” *IEEE Access*, vol. 8, pp. 84 415–84 433, 2020.

[139] M. Yousefi, A. Hajizadeh, M. N. Soltani, and B. Hredzak, “Predictive home energy management system with photovoltaic array, heat pump, and plug-in electric vehicle,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 430–440, 2020.

[140] M. H. K. Tushar, A. W. Zeineddine, and C. Assi, “Demand-side management by regulating charging and discharging of the ev, ess, and utilizing renewable energy,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 117–126, 2017.

[141] Z. Chen, L. Wu, and Y. Fu, “Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization,” *IEEE transactions on smart grid*, vol. 3, no. 4, pp. 1822–1831, 2012.

[142] C. Vivekananthan, Y. Mishra, and F. Li, “Real-time price based home energy management scheduler,” *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2149–2159, 2014.

[143] C. M. Affonso and R. V. da Silva, “Demand side management of a residential system using simulated annealing,” *IEEE Latin America Transactions*, vol. 13, no. 5, pp. 1355–1360, 2015.

[144] S. Bahrami, M. Parniani, and A. Vafaeimehr, “A modified approach for residential load scheduling using smart meters,” in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*. IEEE, 2012, pp. 1–8.

[145] J. Ponoćko and J. V. Milanović, “Multi-objective demand side management at distribution network level in support of transmission network operation,” *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1822–1833, 2019.

[146] J. Wang, Y. Li, and Y. Zhou, “Interval number optimization for household load scheduling with uncertainty,” *Energy and Buildings*, vol. 130, pp. 613–624, 2016.

[147] A. Jindal, B. S. Bhambhu, M. Singh, N. Kumar, and K. Naik, “A heuristic-based appliance scheduling scheme for smart homes,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3242–3255, 2019.

[148] C. Cecati, C. Citro, and P. Siano, “Combined operations of renewable energy systems and responsive demand in a smart grid,” *IEEE transactions on sustainable energy*, vol. 2, no. 4, pp. 468–476, 2011.

[149] N. Ding, C. Benoit, G. Foggia, Y. Bésanger, and F. Wurtz, “Neural network-based model design for short-term load forecast in distribution systems,” *IEEE transactions on power systems*, vol. 31, no. 1, pp. 72–81, 2015.

[150] G. Wang, A. Gunasekaran, and E. W. Ngai, “Distribution network design with big data: model and analysis,” *Annals of Operations Research*, vol. 270, no. 1, pp. 539–551, 2018.

[151] I. Asenova and D. Georgiev, “Short-term load forecast in electric energy system in bulgaria,” *Advances in Electrical and Electronic Engineering*, vol. 8, no. 4, pp. 102–106, 2011.

[152] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *arXiv preprint arXiv:1409.3215*, 2014.

[153] Y. Hu, A. Huber, J. Anumula, and S.-C. Liu, "Overcoming the vanishing gradient problem in plain recurrent networks," *arXiv preprint arXiv:1801.06105*, 2018.

[154] S. Zhang, Y. Wang, M. Liu, and Z. Bao, "Data-based line trip fault prediction in power systems using lstm networks and svm," *Ieee Access*, vol. 6, pp. 7675–7686, 2017.

[155] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "A day-ahead pv power forecasting method based on lstm-rnn model and time correlation modification under partial daily pattern prediction framework," *Energy Conversion and Management*, vol. 212, p. 112766, 2020.

[156] "Thames valley vision project." [Online]. Available: http://www.thamesvalleyvision.co.uk/

[157] Z. J. Lee, T. Li, and S. H. Low, "ACN-Data: Analysis and Applications of an Open EV Charging Dataset," in *Proceedings of the Tenth International Conference on Future Energy Systems*, ser. e-Energy '19, Jun. 2019.

[158] M. Aunedi, M. Woolf, M. Bilton, and G. Strbac, "Impact & opportunities for wide-scale ev deployment," *Report B1 for the "Low Carbon London" LCNF project: Imperial College London*, 2014.

[159] M. Bermudez-Edo, P. Barnaghi, and K. Moessner, "Analysing real world data streams with spatio-temporal correlations: Entropy vs. pearson correlation," *Automation in Construction*, vol. 88, pp. 87–100, 2018.

[160] J. van Brakel, "Robust peak detection algorithm (using z-scores)," *URL https://stackoverflow. com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data*, 2019.

[161] S. Pratik and I. Roberto. Building neural networks with python code and math in detail. [Online]. Available: https://pub.towardsai.net/building-neural-networks-with-python-code-and-math-in-detail-ii-bbe8accbf3d1

[162] S. Tewari, C. J. Geyer, and N. Mohan, "A statistical model for wind power forecast error and its application to the estimation of penalties in liberalized markets," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2031–2039, 2011.

[163] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of arima and lstm in forecasting time series," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2018, pp. 1394–1401.

[164] M. Tan, S. Yuan, S. Li, Y. Su, H. Li, and F. He, "Ultra-short-term industrial power demand forecasting using lstm based hybrid ensemble learning," *IEEE transactions on power systems*, vol. 35, no. 4, pp. 2937–2948, 2019.

[165] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting," *IEEE Access*, vol. 7, pp. 88 058–88 071, 2019.

[166] H. Choi, S. Ryu, and H. Kim, "Short-term load forecasting based on resnet and lstm," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2018, pp. 1–6.

[167] H. Dali and L. Chengcheng, "Demand forecast of equipment spare parts based on eemd-lstm," in *2021 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, vol. 6. IEEE, 2021, pp. 230–234.

[168] X. Wu, Z. Tian, and J. Guo, "A review of the theoretical research and practical progress of carbon neutrality," *Sustainable Operations and Computers*, vol. 3, pp. 54–66, 2022.

[169] L. Chen, “Are emission trading schemes a pathway to enhancing transparency under the paris agreement,” *Vt. J. Envtl. L.*, vol. 19, p. 306, 2018.

[170] Z. Ya-Xin, L. Hui-Lin, and W. Can, “Progress and trends of global carbon neutrality pledges,” *Advances in Climate Change Research*, vol. 17, no. 1, p. 88, 2021.

[171] M. Salvia, D. Reckien, F. Pietrapertosa, P. Eckersley, N.-A. Spyridaki, A. Krook-Riekkola, M. Olazabal, S. D. G. Hurtado, S. G. Simoes, D. Geneletti *et al.*, “Will climate mitigation ambitions lead to carbon neutrality? an analysis of the local-level plans of 327 cities in the eu,” *Renewable and Sustainable Energy Reviews*, vol. 135, p. 110253, 2021.

[172] M. Liu, Y. Yuan, X. Huo, M. Li, and Y. Chai, “Simultaneous estimation of gps p1-p2 differential code biases using low earth orbit satellites data from two different orbit heights,” *Journal of Geodesy*, vol. 94, no. 12, pp. 1–18, 2020.

[173] A. Nilsson, “Remissyttrande: Miljömålsberedningens betänkande om ett klimatpolitiskt ramverk för sverige (sou 2016: 21),” 2016.

[174] T. Gál, S. I. Mahó, N. Skarbit, and J. Unger, “Numerical modelling for analysis of the effect of different urban green spaces on urban heat load patterns in the present and in the future,” *Computers, Environment and Urban Systems*, vol. 87, p. 101600, 2021.

[175] M. Child, R. Ilonen, M. Vavilov, M. Kolehmainen, and C. Breyer, “Scenarios for sustainable energy in scotland,” *Wind Energy*, vol. 22, no. 5, pp. 666–684, 2019.

[176] M. Noussan and S. Tagliapietra, “The effect of digitalization in the energy consumption of passenger transport: An analysis of future scenarios for europe,” *Journal of Cleaner Production*, vol. 258, p. 120926, 2020.

[177] A. Colmenar-Santos, A.-M. Muñoz-Gómez, E. Rosales-Asensio, and Á. López-Rey, “Electric vehicle charging strategy to support renewable energy sources in europe 2050 low-carbon scenario,” *Energy*, vol. 183, pp. 61–74, 2019.

[178] “The 2018 ten year network development plan.” [Online]. Available: https://tyndp.entsoe.eu/tyndp2018/

[179] U. Government. Estimates of heat use in the uk. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/386858/Estimates_of_heat_use.pdf

[180] “Spatio-temporal heat demand for lsoas in england and wales.” [Online]. Available: https://doi.org/10.5286/ukerc.edc.000944

[181] W. Sun, F. Neumann, and G. P. Harrison, “Robust scheduling of electric vehicle charging in lv distribution networks under uncertainty,” *IEEE Transactions on Industry Applications*, vol. 56, no. 5, pp. 5785–5795, 2020.

[182] K. Lepanjuuri, P. Cornick, C. Byron, I. Templeton, and J. Hurn, “National travel survey 2016,” Technical Report Prepared for the Department for Transport, Tech. Rep., 2017.

[183] “Office for national statistics - household projections.” [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/householdprojectionsforengland/2018based

[184] ENTSO-E. Maps and data. [Online]. Available: https://tyndp.entsoe.eu/maps-data/

[185] K. Spees and L. B. Lave, “Demand response and electricity market efficiency,” *The Electricity Journal*, vol. 20, no. 3, pp. 69–85, 2007.

[186] S. Nolan and M. O’Malley, “Challenges and barriers to demand response deployment and evaluation,” *Applied Energy*, vol. 152, pp. 1–10, 2015.

[187] K. Alshehri, J. Liu, X. Chen, and T. Başar, “A game-theoretic framework for multiperiod-multicompany demand response management in the smart grid,” *IEEE Transactions on Control Systems Technology*, vol. 29, no. 3, pp. 1019–1034, 2020.

Published Work

2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER)

Development of a Hardware in-the-Loop Co-Simulation Platform for Smart Distribution Networks

Yuanda Gao[1] Desen Kirli[1] Mehdi Zeinali[1] Shubhankan Mukherjee [1] Aziza Birzhanova[1]
John Thompson[1] Naran Pindoriya[2] Aristides Kiprakis[1]
[1]*University of Edinburgh, United Kingdom*
[2]*Indian Institute of Technology Gandhinagar, India*
yuanda.gao@ed.ac.uk, desen.kirli@ed.ac.uk, m.zeinali@ed.ac.uk, s.mukherjee-5@sms.ed.ac.uk,
a.birzhanova@sms.ed.ac.uk, j.s.thompson@ed.ac.uk, naran@iitgn.ac.in, kiprakis@ed.ac.uk

***Abstract*—This paper describes a hardware in-the-loop real-time co-simulation platform for smart distribution networks incorporating Phasor Measurement Units (PMUs) and communications. The platform is designed to simulate and collect data from different network topologies, penetrations of different types of generations and loads and communication topologies. These capabilities make the platform a suitable tool for the deployment and test of smart control algorithms and energy trading strategies. This paper also identifies and addresses a communication infrastructure which provides the right trade-offs for reliability, low-latency and low-cost communication. Moreover, the various applications of PMU in distribution networks are discussed and the potency of PMU being the custodian of the electrical grid is elucidated.**

***Index Terms*—Power system simulation, smart grid, real-time systems, phasor measurement units, energy management, communication network.**

I. Introduction

With human progress and social development, the traditional electrical grid is changing. On the consumer side, the integration of electric vehicles, storage devices and other intelligent appliances make the load more complicated while the unexpected consumer reaction to the Demand-Side Management (DSM) programs leads to more unpredictable nature of the load. Further, the significant increase of distributed generations (DGs) also adds more stresses on the existing power grid. Therefore, the traditional power grid has been upgraded and become smarter by applying Wide Area Monitoring System (WAMS) technology with sophisticated control systems. In WAMS, PMUs are commonly deployed among the system to achieve a wide area monitoring with time-synchronised and high-resolution voltage and current phasors measurement, which can help in the prediction of voltage violations and prevent the unstable operation of the grid.

Meanwhile, due to the measurement precision and frequency which generate big data, one of the most important parts of the of WAMS setup is the communication network in addition to PMUs, Phasor Data Concentrators (PDCs) and super PDCs [1]. Throughout the recent investigation of different PMU applications, a key assumption is that there is a proper communication network, which can fulfil the requirement for timely and reliable sharing of PMU data. Several challenges need to be considered to evaluate and improve the performance of the required communication system such as impact of the transport layer protocol (TCP/IP protocols), the effects of data compression techniques to deal with the large amount of data produced in different PMUs within the smart grid, and mitigating potential cybersecurity-related problems.

Before the actual deployment of hardware and software platforms, the success of pre-simulation is necessary. Depending on the application, there is a multitude of simulation tools commonly used like Simulink, OpenDSS, Siemens PSS/E, MATPOWER, ETAP, etc. Furthermore, some platforms are proposed for smart grid simulation, e.g., GridSpice, which is a scalable and cloud-based open-source simulation framework that enables large network modelling and blur boundaries between generation, transmission, distribution, and markets [2]. Additionally, MASGriP presents a multi-agent smart grid simulation platform, which implements several consumer and producer agents and considers real characteristics, different goals and actuation strategies to simulate technical and economical activities of several players [3].

In this paper, we describe a hardware in-the-loop platform for smart grid real-time simulation. The platform

makes provision for the utilisation of low-cost PMU modules as Hardware-In-the-Loop (HIL), allowing the implementation of advanced wide-area control strategies. Development of this low-cost hardware-in-the loop platform is of significant aid towards achieving real-time monitoring for optimised energy management, which is the central objective of the Data-Driven Intelligent Energy Management (D-DIEM) collaborative project between the Universities of Edinburgh and Queen's in Belfast from the United Kingdom, and the Indian Institutes of Technology in Gandhinagar and Kanpur and the Madan Mohan Malaviya University of Technology from India [4]. Use of PMUs in the proposed platform embraces the aforementioned points and accommodates the accomplishment of D-DIEM project.

The paper is organised as follows. Section II describes the hardware platform development. In Section III we present our practical evaluation of current communication networks when used for communication of PMU data. We also study the improvement of communication systems associated with using data compression techniques to reduce the volume of data associated with open PMU networks. In Section IV, some key applications of the developed platform are introduced. Finally, in Section V, two case studies are proposed based on the simulation platform.

II. Platform Development

The functional design of the co-simulation platform is illustrated in Figure. 1; the upper layer indicates the power flow in the system while the consumers can merchandise and sell electricity to the network operator; the lower optimisation and control layer can exchange information with the upper layer via the communication system. The platform includes a power system real-time simulator, the PMU and a workstation used to emulate the optimisation and control layer.

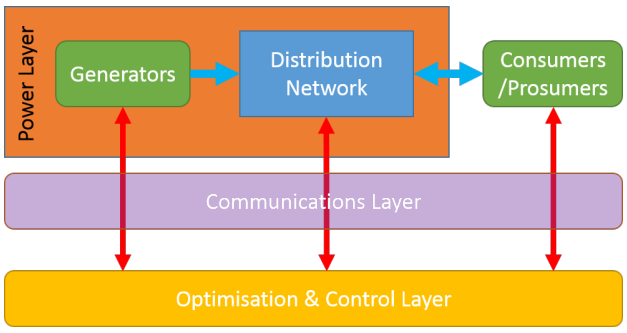


Fig. 1. The structure of the real-time simulation test rig

A. Real-time simulator: OPAL-RT

In power systems, real-time simulation is necessary for users to test their developed hardware devices and power system algorithms. The simulation result from the simulator represents conditions in a real network and thus provides opportunities for the developer to take delays from the transmission system, communication system and control centre into account, therefore, making improvements. Additionally, the behaviour of developed physical devices and their in-field interactions can be recorded in real-time simulations.

The selected network real-time simulator is OPAL-RT OP5600, which contains several I/O channels and is compatible with external modules. The network is developed in Simulink and embedded into OPAL-RT via RT-LAB. In this stage, the network can be formed as various topologies, penetrations of different types of generations and loads for different network simulations. As shown in Figure. 1, the OPAL-RT emulates the power layer and generates real-time bus voltage and current that can be measured by the PMU via the Analogue and Digital I/Os channels on the OPAL-RT. Besides, information can be exchanged with a computer to achieve bi-directional communication in the network via cable or wireless communication. These make the platform a dedicated tool to be deployed in the real network and execute tests.

B. Phasor Measurement Units

The PMU is the device that converts the three-phase voltage and current analogue signals into synchrophasors. In detail, the measured voltage and current from the potential transformer and the current transformer go through filters for anti-aliasing and then being passed to the analogue-to-digital (A/D) converter module. Meanwhile, the converted digital signal is appended with time tags which come from the GPS clock with the help of phase-locked oscillators. Afterwards, the data is sent to the microprocessor to generate the corresponding phasor, frequency and Rate of Change of Frequency (ROCOF). Finally, the processed data goes into the communication module for later transmission, which may involve data compression, data encryption, etc. technologies. Therefore, PMUs bring two key functionalities including generating high resolution measurement (50Hz and 60Hz typically) and time-synchronised voltage and current phasor, and these provide us the opportunity to deploy real-time protection and control strategies, and obtain accurate real-time network operation status. The structure of the PMU is shown in Figure. 2.

The PMU module used in this co-simulation platform is the OpenPMU, a low-cost, open source implementation of a Phasor Measurement Unit [5]. In addition to the OpenPMU, external plug-in devices are available to cooperate with the platform to provide functions like

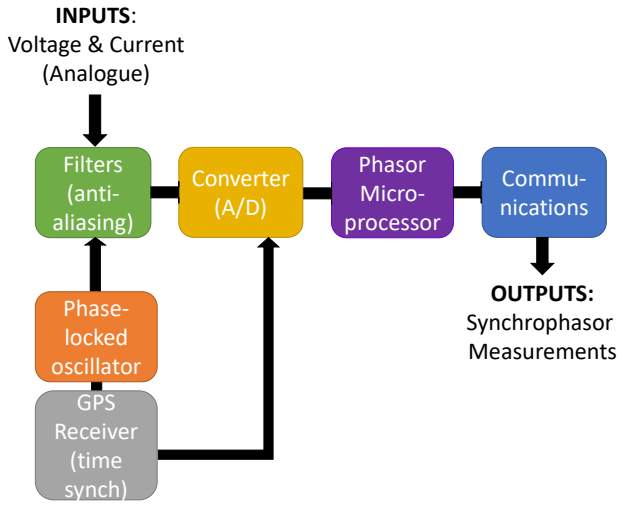


Fig. 2. The block diagram of a PMU.

wireless communication, data encryption and decryption, data compressing, etc.. The detailed communication systems and the field test is discussed in section III.

Figure. 3 illustrates the developed co-simulation platform. It shows the front view of the OPAL-RT real time simulator connected with a computer through an Ethernet connection, while Analogue and Digital I/Os are coming through the backside of the device.

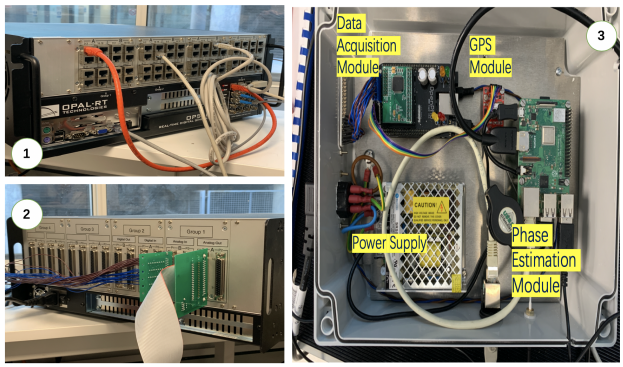


Fig. 3. [1] Front view of OPAL-RT, [2] Back view of OPAL-RT, [3] Structure of OpenPMU (with power supply)

C. Workstation

The data from the PMU and OPAL-RT can be exchanged with the computer, which behaves as the control centre in the real network (as Figure. 4 shows). Algorithms are deployed on the workstation to analyse the data collected from PMUs, monitor the system operation and, therefore, make a diagnosis. For example, using limited bus measurements to achieve high observability and network fault detection. Afterwards, control algorithms can be deployed through the communication system to consumers and the consumers' reactions can be sent directly into the control centre. Also, the behaviour of the consumers can be emulated on the computer. The application of the platform is discussed in the later paper. This makes the platform ready to test different control algorithms.

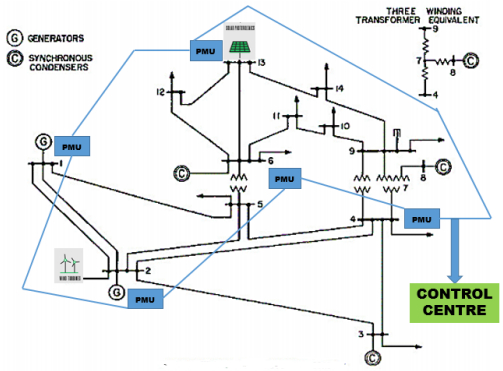


Fig. 4. Structure of the network (based on IEEE 14-Bus network)

III. Communication Systems

As shown in Figure. 1 the telecoms or communication module is a key element of a PMU. WAMS can make use of PMUs located across the smart grid to help monitor the network conditions and the status of the power grid in real-time. Due to the real-time monitoring, a fast, responsive, reliable and secure communication infrastructure is required. Effective information exchange between different entities in a WAMS particularly between PMUs and PDCs is one of the key requirements for a successful smart grid system. The following latency and compression schemes are carried on the Raspberry Pi (included in OpenPMU for communication) and the PC.

A. Requirements and challenges of PMU data communication

Achieving a high data throughput, with low communications latency and high reliability (low bit error rate) are the primary factors for choosing an appropriate communication system for the transfer of PMU data in a WAMS. The communication system for PMUs in a real-time WAMS, should support an overall time delay of less than 1 second [1]. Based on the cycle of 50 reports/sec for a 50-Hz power system with a packet size of 40 bytes mentioned in Section II.B, Table I shows this corresponds to a 2KB data packet. Also, we have considered packet sizes of up to 10 kbytes/sec to understand what happens if the reporting rate increases or higher resolution data is needed.

TABLE I
PMU DATA PACKET SIZE AND REPORTING RATE IN ONE SECOND.

Packet size (bytes)	Reporting rate (packet per second)
40	50

B. Last mile technologies

Communication links used by PMUs include both wired and wireless options. Some key options are discussed and highlighted as follows: • Cellular Mobile Networks: both 3G (data rates 1.5 Mbps - 6.8 Mbps) and 4G (data rates 12 Mbps up to 22 Mbps) can provide wireless access for wide-area coverage of smart grid systems to link to a WAMS; • Wired last mile connectivity: full optical fibre broadband at a University Network and Cable Broadband (data rates are 100Mbps or higher in both cases) are studied to link to the internet backbone and the WAMS. In order to evaluate the proposed model [6], the performance of two internet protocols (Transmission Control Protocol(TCP) / User Datagram Protocol (UDP)) has been tested in a PMU-PDC communication model to emulate real data packets exchanged in a WAMS system. A python script written to act as the PDC is run on a laptop PC, which communicates with the PMU code, which was installed on the Raspberry Pi 3B (the same platform as for the Open PMU) to emulate a low-cost PMU in a WAMS application. To provide increased security, data communication through the internet network is encrypted by a Virtual Private Network (VPN) service running on both the PC and the Raspberry Pi.

In this work, an accurate and reliable test set-up has been implemented for gathering end-to-end packet latency measurements, using high precision time-synchronisation with GPS receivers implemented both in the server and the client-side. We have investigated the one-way latency for remote control PMU connections and the results are shown in Figure. 5. Our experimental investigation reveals that the typical average latencies range from 100 msec to 600 msec for 1 kbyte to 10 kbyte short data packets. From Figure. 5 it can be seen that the wired connections have better latency value for different data packet sizes in comparison to wireless (cellular) links. Results also suggest that UDP packets experience 4 times more losses than TCP packets for wired connections. UDP may be preferable in terms of latency, provided the packet loss rate is not too high.

1) Compression Techniques: These work based on removing redundant data and encoding the data more effectively. This means that the data could be stored in a smaller memory space at the PDC and require less bandwidth on the communication channel for transmission. The packet sizes can be reduced significantly by compression techniques, which relaxes the requirements for the communication system. For PMU datasets we need to use lossless compression techniques [7], in order not to lose any important information at the PDC. In lossless schemes, statistical modelling approaches are used to find and reduce repeated information in a data set. Some typical methods that compression techniques are using include removing space characters and replacing repeated multiple character strings with a shorter data sequence. The Space Saving is a useful term to describe reduction size which can be defined in equation (1):

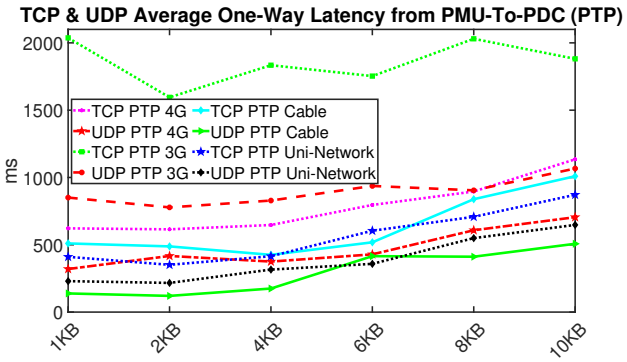


Fig. 5. Latency evaluation for wired and wireless Internet-based communication networks

$$S = 1 - \frac{C}{U} \tag{1}$$

Where S, U and C are the Space Saving, the uncompressed data size (in bytes) and the compressed data size (in bytes) respectively. In order to evaluate the impact of data compression on PMU data, different data packet sizes are studied. Also, two lossless compression techniques, Lempel-Ziv-Welch (LZW) and Adaptive Huffman (AH) have been implemented [7] to measure their performance on PMU data. As can be seen in Figure. 6, the LZW method has better performance on large dataset sizes, achieving a high space saving percentage of up to 87% for large data packets and therefore reducing the required communication bandwidth significantly. On the other hand, the AH method achieves a lower space saving percentage of around 75% but it takes advantage of a simpler and faster algorithm that makes it easier to implement on hardware. To select proper compression schemes for a different part of our communication scenario we have to make a trade-off between compression rate, processing time and hardware capabilities to meet the required requirement.

IV. APPLICATIONS OF THE PLATFORM

As the complexity of the electricity grid and market grows, the applications of PMU technologies are increasing, especially in the field of power system monitoring,

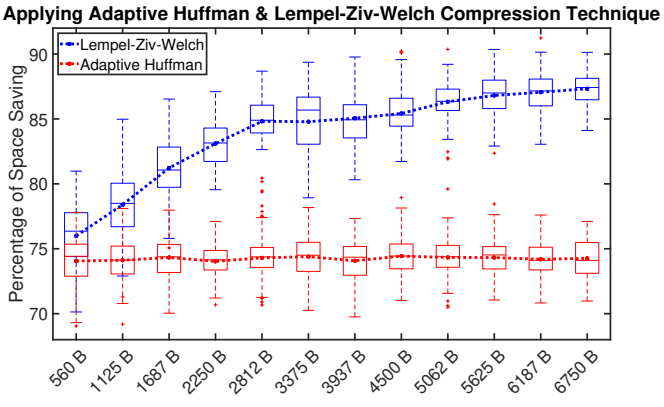


Fig. 6. Comparison of LZW and AH compression schemes on different data packet size

control and protection. In order to address both real-time and off-line capabilities of the PMU, its most significant applications in smart grids are categorised accordingly and presented in Table II.

TABLE II
APPLICATIONS OF THE PMU PLATFORM CATEGORISED INTO REAL-TIME AND OFF-LINE TASKS. (PARTIALLY FROM [8], [9], [10])

Real-time Applications	Off-line Applications
State estimation	Validation of system models
Power quality monitoring	Calibrating parameters
Unit monitoring and control	Post-disturbance analysis
Congestion management	
Adaptive protection	
Disturbance source identification	
Load shedding	
Real-time model validation	

In this paper an indicative set of smart grid applications that can be simulated in the developed platform are discussed.

A. Voltage Regulation

As the Supervisory Control and Data Acquisition (SCADA) system scanning rate is 1 sample in 2-4 seconds, it is restricted to identifying the dynamic changes in the power system [11]. Application of time-synchronised PMUs could address the existing challenges of real-time voltage stability monitoring and control.

To monitor the voltage stability margin and predict possible voltage collapse, phasor measurements of current and voltage from the load buses could be utilised to calculate the load impedance and estimate the equivalent Thevenin's impedance of the system. Applying the theorem of maximum power transfer, voltage stability margin can be supervised to predict the point of voltage collapse. The estimation of Thevenin's impedance could vary in methods and depend on the number of PMUs in the network [12]. The optimal placement of PMUs in the network should be established to achieve the cost-efficient solution that provides full observability of the system.

Upon the prediction of voltage stability margin or detection of possible voltage violations, PMUs can provide the synchronised voltage phasors, which bring an improvement in the control of state-of-the-art voltage control techniques. Placing the PMUs at active nodes, i.e. Distributed Generation (DG), Static VAR Compensator (SVC) or large-scale load buses, and applying the voltage-to-power sensitivities, the injected power from the DGs can be regulated [13]. The computation part is executed in the central PDC that provides either automatic regulation in case of voltage surpassed its limits or implements preventive actions to avoid the possible voltage instability. The involvement of reactive power compensators, such as shunt reactors or shunt capacitor banks, could be additional ease in real-time voltage regulation with voltage phasor measurements obtained from PMU [14]. Apart from the regulation of injected power from DGs, the voltage can be controlled by implementing demand-side management techniques that aid with matching supply and demand and consequently improve the voltage profile.

B. Fault analysis

One of the most important applications of the PMU is in the use of fault (FLT) location monitoring and protection. With the optimal position and number of PMUs deployment strategy, the system observability can be improved. With efficient system monitoring, the Synchrophasor unit reports any transient event accurately. The continuously monitored network status is compared with the PDC historian. A simple method of residual value comparison calculates the difference in estimated and real-time sampled value [15]. If the difference exceeds the threshold limit, then it is evident that the system has gone into a fault or unbalanced state. PMUs take about 40-60 samples/cycle at the nominal frequency (50Hz). So, within 2-3 cycles the PMUs detect transients and anomalies [16]. Within the next 2-3 cycles, the relays get energised and circuit breakers (CBs) react. When a fault is detected by PMU, relays in specific zones of the grid with specific Current Setting Threshold (CST) and Time Setting Multiplier (TSM) is energised to trip the CB either by using battery power or on the grid power supply. For the low voltage zones, we use the circuit breaker and the contactors for the high voltage areas. This takes about 750 msec to 1 sec to completely isolate the affected location. Specific Control Algorithm (CA) are set in place to initiate controlled fault isolation.

C. Demand-side management

The use of PMUs can improve DSM methods in terms of response to dynamic events in the system. If the frequency of the power system is seen to be close to its limits, automatic control of the non-priority loads can be launched to return the frequency within the stable state. The non-priority loads in a direct load control (DLC) program can act as virtual resources as they can be switched off during the overload in the power system. Due to the short response time of the DLC program, it can be added to primary frequency control methods [17]. If frequency data from PDC indicates that the frequency tends to fall though is still within its stable condition and requires supervision, the distribution system operator (DSO) can apply preventive demand control actions. As the whole monitoring and control are taking place in real-time, then only price-based or incentive-based programs can be implemented due to their ability to produce short-term solutions to the system. The real-time pricing scheme is based on real-time wholesale electricity prices, which are in turn related to real-time data from PMUs. A short notice an hour ahead from the utility allows the consumers to quickly respond and improve the load pattern of the whole system. By reaching the objectives of DSM such as peak clipping, valley filling, and load shifting, the flat pattern of the load can be observed in the distribution network. This can decrease the costs of investment in reinforcing the distribution lines, increase the output generation of DGs and improve the whole state of the network.

D. Validation of Local Energy Market and Other Dynamic Models

The innovate smart grid technologies such as local energy markets, that enable local transactions of electricity at the distribution level, require accurate monitoring, protection and control of power systems in real-time. Most of these emerging technologies aim to utilise the wide-area monitoring, protection and control (WAMPAC) [9]. The implementation of WAMPAC relies on the specific attributes of the PMU which are namely, real-time and time coordinated measurements of the distribution system using GPS.

Local energy market transactions usually take place amongst prosumers on the same designated distribution system [18]. The transactions are designed to happen at frequent intervals that range in duration. The volume of exported electricity from the distributed generation depends on whether their offer (i.e. the price in £/kWh that they are willing to sell their generation for) is matched with the bid of a buyer and their desired units of electricity. Additionally, the buyers and sellers could be allowed to change their prices and preferences in real-time. For example, the Brooklyn MicroGrid [19], which is a blockchain-based P2P energy trading platform, allows these real-time modifications and facilitates market-clearing in set intervals of 15 minutes [18]. This introduces a dynamic system that requires the real-time measurement, monitoring and protection capabilities of PMU.

Traditionally, off-line power flow models coupled with dynamic data are used for validation which may involve disturbance events [20]. [10] presents an example of the novel approach for testing dynamic real-time system models with a PMU, using North American Western Interconnection as a case study. Using a similar methodology to [20] and [10], PMUs can be used for validating the feasibility of local peer-to-peer electricity transaction models and calibrating system parameters in the transaction model.

E. Other Applications

Other applications of the PMU simulation platform involves the test of backup protection [21]:

1) Control of Backup Protection Distance Relays
2) Providing additional control on angular stability of the Power System
3) Advanced Forecastability and Dependable Supply.

These ancillary protections play an important role in keeping the grid healthy. At times there can be an unsuccessful attempt to disconnect the specific zonal CB, so in order to keep the system safe, we apply the philosophy next of kin method-based CB disconnection. Due to the disconnection of CB, there can be transient decrease/increase in load or generation, and this has temporal effects on the rotor angle stability of the generators supplying the load and gives rise to voltage instabilities. Lastly, PMUs allow state estimation and help to predict future demand which enables the utilities to match demand and supply in a better way.

V. Case studies

To apply the theoretical framework of the proposed platform on practice, two different applications were tested, and the results are provided in the following section.

A. Voltage control through DLC program

In the first case study, a modified version of the IEEE 15-bus network was modelled in Simulink with the help of Simscape Power Systems library package and the results were obtained in RT-LAB software of the OPAL-RT 5600 real-time digital simulator. A single line diagram is presented in Figure. 7. The distribution

network is operating at 11kV and has four DGs connected to the grid. The DGs are presented as photovoltaic (PV) solar systems rated at 100kWp at standard conditions (1000 W/m2 irradiance and 25° C ambient temperature). Residential loads are placed at buses with DG connected and are divided into two groups. The first group illustrates uncontrollable high-priority appliances (static load), while the second group represents the loads signed to the DLC program. In this case study, the role of controllable appliances is played by air-conditioning systems (AC).

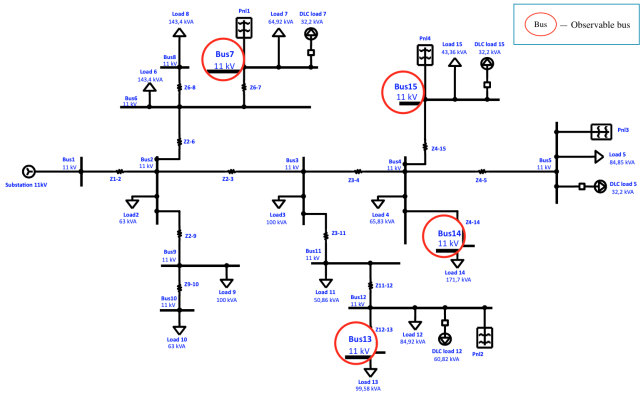


Fig. 7. Single line diagram of 15-bus network

The major challenge associated with the connection of PV solar systems is their intermittent nature caused by constantly changing solar irradiance. The insolation patterns of four PV arrays are presented in Figure. 8. This corresponds to two minutes of real-time simulation, with a cloud shading possibility.

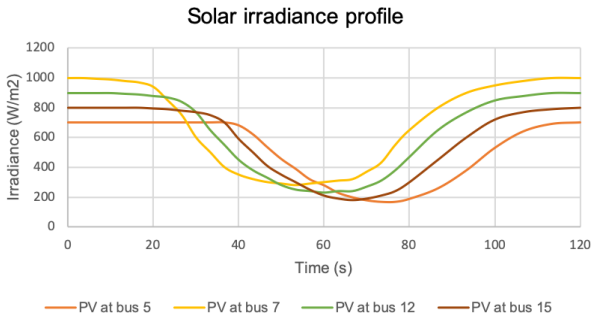


Fig. 8. Solar insolation patterns

Consequently, it leads to fluctuations in voltage and frequency levels in the network. To reduce the total cost of the grid, PMUs are only placed at buses with critical voltage levels. For this purpose, an offline simulation was carried out in Simulink to define the buses with critical voltage levels, omitting those that are close to the substation. During the passing cloud, generated power from solar systems decreased, and, therefore, led to voltage levels at buses 12, 13 and 14 to fall to critical points (Figure. 9). This is an expected outcome due to the larger distance of these buses from the substation combined with connected heavy loads. Hence, it is assumed that two PMUs are placed at buses 13 and 14, missing bus 12 due to the neighbouring connection of buses 12 and 13, i.e. it is sufficient to check only one of them. To increase the observability of the network, two additional PMUs are placed at buses 7 and 15, resulting voltage phasors can be further applied to the DLC program. If the magnitudes of voltage phasors at the corresponding buses are less than the statutory limit, which is ±6 % for the UK distribution networks, an instruction is sent to switch off the AC systems of participating customers.

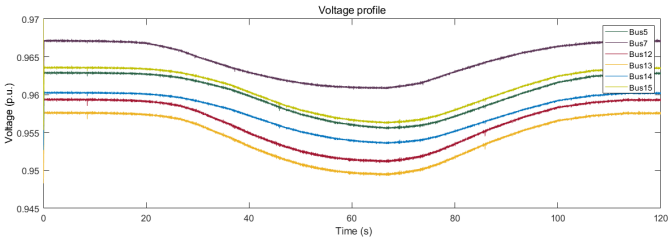


Fig. 9. Offline simulation of voltage levels without control

After receiving magnitudes of voltage phasors from PMUs, a control action of DLC program was simulated in real-time. Figure. 10 shows that the DLC was enabled at 55 seconds, after which the controller checked the voltage phasors at the observable buses. As voltage levels at buses 13 and 14 were at critical levels, the instruction was sent to switch off the AC loads at buses 5 and 12, due to the nearest location. By decreasing the volume of load, the voltage levels were increased to the acceptable range.

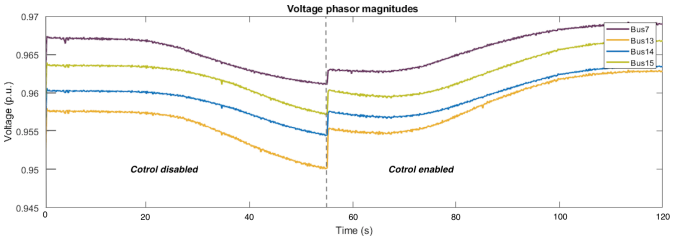


Fig. 10. Voltage level with DLC

B. *Transient fault analysis*

1) Model description and case scenarios: In this section, the transient fault analysis has been carried out. In Figure. 11, a Single Line Diagram depicting a distribution network scenario has been emulated. At the Bulk Supply Point, the nominal bus voltage is 1 kV with a system power factor of 0.8. As per IEC 60038:1983, any voltage up to 1000 volts (for AC Systems) is

classified as low voltage and low voltage circuit breakers have been used. At bus 7, a 5MWp Solar Farm has been connected and at bus 10, a 5MW (11 kV) Wind Farm has been connected. These two Distributed Generators support energy demand and help to ease the voltage sag.

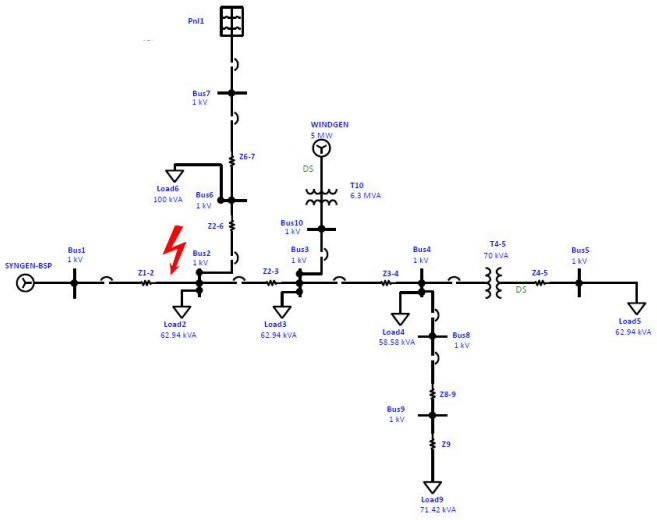


Fig. 11. Network model

To understand the effect of faults in the system and fault contribution by various generation elements in the grid, three-phase balanced line faults are created between bus 1 and 2 in the above network. This is chosen because the magnitude of the fault current is always high when it close to generating source as the network impedance is the least.

In order to analyse the fault transient in OPAL-RT, repetitive Line-Line-Line symmetrical faults are created manually between Bus 1 and Bus 2. In Figure. 12, fault transient occurs during the following intervals: 0.5s ,1.5s and 2.5s. This is to ensure that the faults are traceable in real-time. The fault is allowed to persist for about 500 milliseconds. After this, it is manually cleared without using the CB. This is done to identify the areas where the fault current is high and set relay grading and circuit breaker rating accordingly.

A characteristic of a radial network that voltage keeps decreasing down the line. So, distributed generation serves as a cardinal instrument to maintain the right voltage balance and account for power shortage and supply the residual/deficit power.

2) Result discussion: From the Table. III, we can see the buses 2, 3 and 6 are critically affected as they are closest to the fault location (marked as red). It is also observed that power supply has been constrained in buses 4, 5 and 8. Due to this, there is a growing need for a fast sensing device like the PMU. An optimal PMU placement algorithm is implemented here [22].

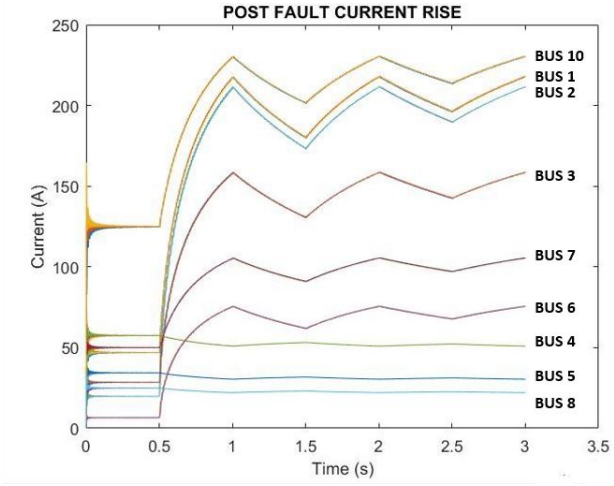


Fig. 12. RMS Post Fault Current Values

TABLE III
FAULT CURRENT VALUES (RMS)

BUS	CURRENT(Nom)	FLT CURRENT	CHANGE
BUS 1	46.43	217.047	79%
BUS 2	19.858	211.499	91%
BUS 3	28.368	158.23	82%
BUS 4	57.682	50.432	-14%
BUS 5	34.357	29.944	-15%
BUS 6	6.619	75.648	91%
BUS 7	50.117	105.277	52%
BUS 8	24.901	21.749	-14%
BUS 10	125.134	230.411	46%

According to this algorithm, PMUs are usually placed at buses with the largest interconnected branches or with generators or with heavy loads. Also, to maintain full observability and optimise cost function, the PMUs are placed one bus apart. So, in this system PMUs are placed at bus 1, 2, 3, 5 and 7. PMUs can be used to sense a sudden onset of the fault and since they have a very high sampling and reporting rate, they identify fault transient very quickly. The faults are generally addressed using negative or zero sequence, but PMUs can measure only positive sequence. So, it is observed that fault offsets the positive sequence voltage and current phasors and a specific control algorithm is in place to differentiate between the change in positive sequence current phasor due to overload and fault (i.e. current surge).

Accounting for the change in current phasor at bus two due to fault as shown in Figure. 13, we see a sharp rise rate of change of positive sequence current phasor both magnitude and phase angle. This is detected by the PMU and a control threshold is set to differentiate between a sudden load change and a fault. When this threshold is crossed, fault protection devices are triggered and this energises the circuit breakers and trips them.

In the Simulation, as shown in Figure. 14, at 0.5 seconds, first fault occurs. The PMUs monitor the change

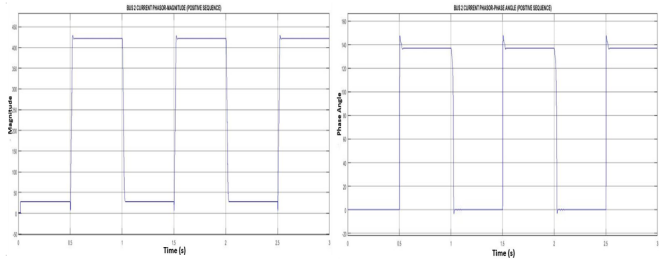


Fig. 13. Bus 2 Positive Sequence Current Phasor

in positive sequence current phasor and the control threshold is set at 60A. When this is crossed, a trip command is generated wherein a trip signal changes from high logic (1) changes to low logic (0) to open the circuit breaker. Once the circuit breaker is opened, the voltage returns to its nominal value and the rest of the circuit gets disconnected thereby stabilising the system. PMUs act as the custodian of the grid and they are very prompt and fast at reporting events. This forms the basic pillar to keep the power grid stable.

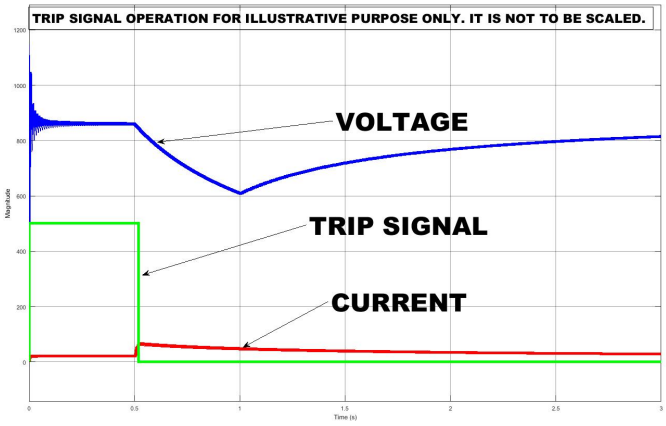


Fig. 14. Protection Logic Sequence

VI. Conclusion

This paper discusses a hardware in-the-loop real-time simulation system, which is built up based on the real-time simulator, OPAL-RT and a low-cost PMU platform. The proposed system is easily setup while it enables simulation and study of a wide range of power system applications. Moreover, this co-simulation platform allows the inclusion of power electronics, with low-cost ICT interfaces to enable network services such as demand response as well as peer-to-peer energy trading, and involves feasible communication systems configurations to meet the key requirements of a successful smart grid system. The platform also provides an opportunity not only for lab simulation but also for Hardware-In-the-Loop simulations for the development of controls and devices that can then be deployed in the distribution network. In the final section, a few key applications of the platform were addressed including voltage regulation, fault analysis, DSM and Peer-to-Peer energy trading. Finally, a few case studies are carried out upon the simulation platform, which shows the importance and flexibility of the hardware in-the-loop co-simulation platform.

VII. Acknowledgement

This work was partially supported by the EPSRC Centre for Energy System Integration (Grant ref. EP/P001173/1) and the UKIERI-DST funded Data-Driven Intelligent Energy Management for Environmentally Sustainable Energy Access project. The authors would also like to thank Dahunsi Okekunle for his assistance with setting up the co-simulation platform.

References

[1] C. Danielson, L. Vanfretti, M. S. Almas, Y. Choompoobutrgool, and J. Gjerde, "Analysis of communication network challenges for synchrophasor-based wide-area applications," in *2013 IREP Symposium Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid*. IEEE, 2013, pp. 1–13.

[2] K. Anderson, J. Du, A. Narayan, and A. El Gamal, "Gridspice: A distributed simulation platform for the smart grid," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2354–2363, 2014.

[3] P. Oliveira, T. Pinto, H. Morais, and Z. Vale, "Masgrip—a multi-agent smart grid simulation platform," in *2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012, pp. 1–8.

[4] N. Pindoriya, A. Kiprakis, C. K. Ajay, S. Singh, D. Garg, D. Padmanabhan, and J. Thompson, "Integrated energy management framework for environmentally sustainable energy access," in *2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*. IEEE, 2018, pp. 1–8.

[5] D. M. Laverty, R. J. Best, P. Brogan, I. Al Khatib, L. Vanfretti, and D. J. Morrow, "The openpmu platform for open-source phasor measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 4, pp. 701–709, 2013.

[6] M. Zeinali and J. Thompson, "Implementation of highly accurate test-bed for practical evaluation of wired and wireless internet based smart grid communications," 08 2019.

[7] M. Zeinali and J. S. Thompson, "Impact of compression and aggregation in wireless networks on smart meter data," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2016, pp. 1–5.

[8] S. S. Noureen, V. Roy, and S. B. Bayne, "Phasor measurement unit integration: A review on optimal pmu placement methods in power system," *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, 2017.

[9] M. K. Penshanwar, M. Gavande, and M. F. A. R. Satarkar, "Phasor measurement unit technology and its applications - a review," *2015 International Conference on Energy Systems and Applications*, 2015.

[10] Y. Lu, S. Kincic, H. Zhang, and K. Tomsovic, "Validation of real-time system model in western interconnection," *2017 IEEE Power Energy Society General Meeting*, 2017.

[11] C. Tu, X. He, Z. Shuai, and F. Jiang, "Big data issues in smart grid–a review," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 1099–1107, 2017.

[12] K. Alzaareer and M. Saad, "Real-time voltage stability monitoring in smart distribution grids," in *2018 International Conference on Renewable Energy and Power Engineering (REPE)*. IEEE, 2018, pp. 13–17.

[13] X. Wang, C. Wang, T. Xu, H. Meng, P. Li, and L. Yu, "Distributed voltage control for active distribution networks based on distribution phasor measurement units," *Applied energy*, vol. 229, pp. 804–813, 2018.

[14] P. Li, H. Su, L. Yu, Z. Liu, C. Wang, and J. Wu, "Voltage control method of distribution networks using pmu based sensitivity estimation," *Energy Procedia*, vol. 158, pp. 2707–2712, 2019.

[15] D. M. Laverty, H. Kirkham, D. J. Morrow, and X. Liu, "Estimation of goodness of fit of synchrophasors during transient faults," in *2017 IEEE Power & Energy Society General Meeting*. IEEE, 2017, pp. 1–5.

[16] P. Rajaraman, N. Sundaravaradan, B. Mallikarjuna, D. Mohanta *et al.*, "Robust fault analysis in transmission lines using synchrophasor measurements," *Protection and Control of Modern Power Systems*, vol. 3, no. 1, p. 14, 2018.

[17] K. Samarakoon, J. Ekanayake, and N. Jenkins, "Investigation of domestic load control to provide primary frequency response using smart meters," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 282–292, 2011.

[18] E. Mengelkamp, J. Gärttner, K. Rock, S. Kessler, L. Orsini, and C. Weinhardt, "Designing microgrid energy markets a case study: The brooklyn microgrid," *Applied Energy*, vol. 210, p. 870–880, 2018.

[19] "Community powered energy." [Online]. Available: https://www.brooklyn.energy/about

[20] K. S. Shetye, W. Jang, and T. J. Overbye, "System dynamic model validation using real-time models and pmu data," *2018 Clemson University Power Systems Conference (PSC)*, 2018.

[21] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Transactions on smart grid*, vol. 1, no. 1, pp. 20–27, 2010.

[22] K.-P. Lien, C.-W. Liu, C.-S. Yu, and J.-A. Jiang, "Transmission network fault location observability with minimal pmu placement," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1128–1136, 2006.

Long Short-Term Memory Based Dynamic Adaptive Compensation Method for Short-term Electric Load Forecast in Distribution Network

Yuanda Gao, *University of Edinburgh,* James Robertson, *University of Edinburgh*
Aristides Kiprakis, *University of Edinburgh*

***Abstract*—With the development of the power grid, the smart grid makes the system more intelligent, efficient, sustainable, and reliable with integrated Information and Communication Systems (ICTs) and big data. In this research, a Long Short-Term Memory (LSTM) based Dynamic Adaptive Compensation (DAC) forecast method is proposed. This method utilises high time-resolution data and LSTM network to provide short-term load forecast results with dynamic error correction ability. Compared with existing Autoregressive Integrated Moving Average Model (ARIMA) and Persistence forecast methods, the proposed features at providing robust and more accurate forecast results while the available historical training dataset or the data source is limited. Therefore, in the distribution network (DN) or microgrid, the DAC-LSTM method provides accurate short-term forecast result for the purpose of operation cost reduction, for example, Peer-to-Peer (P2P) trading, and Distributed Generation (DG) management to reduce carbon footprint, etc. Moreover, the proposed method gives high performance upon domestic load which expands the application range.**

***Index Terms*—Smart grid (SG), long short-term memory (LSTM), machine learning (ML), short-term load forecast (STLF), big data.**

I. Introduction

ELECTRIC load forecast is fundamental in smart grid as it can help suppliers to model and forecast load in advance, balance the demand and supply, adjust demand-side management (DSM) plan, implement real-time pricing schemes, etc. In this regard, the accuracy and robustness of load forecast method is important. In addition, the electricity market is motivated by economics while the increase in load forecasting accuracy reduces the negative impact on the economy [1], [2]. Generally, based on the forecast time step, the electrical load forecast is classified into three categories, including long-term, medium-term, and short-term [3].

In a distribution network (DN), the relationship between consumers, suppliers, and distribution system operators (DSO) becomes complex [4], [5]. The development of the smart grid has created massive real-time and historical data, which is collected from the monitoring devices, such as smart meters, Phasor Measurement Units (PMUs), and other user behaviour data [6]. In recent publications, short-term load forecast (STLF) methods have been proposed using hybrid methods to improve the forecast accuracy in different situations. These methods utilise sufficient historical data collected from the system from DN and even from the microgrid for the purpose of model training. But generally, the trained model fits only to the training historical dataset. Moreover, the DSM such as Peer-to-Peer (P2P) trading methods are usually utilised in small network. This indicates that the collectable historical data might be insufficient to train the model and the nonlinearity of load curve increases further. Based on this fact, papers published in improving the STLF accuracy while the historical data is insufficient are lacked.

This paper describes novel STLF techniques based on Long Short-Term Memory (LSTM) network. In this research, the STLF based model will be fully illuminated. The dataset from Thames Valley Vision (TVV) Project is utilised for model development [7]. The forecast model is featured in combining big data with the state-of-art machine learning methods to generate the STLF model for the distribution network. The forecast model adopts, first, multi-feature and high time resolution datasets to improve highly non-linear load forecast and, second, the dynamic adaptive compensation (DAC) method to improve the forecast accuracy while the input dataset is limited. Therefore, the forecast model can be unitised in smaller distribution network or communities, and improve control methods like P2P trading and other real-time control.

The remainder of this paper is organised as follow. In Section II, the present load forecast techniques are illustrated. In Section III, the detailed methodology is presented. In Section IV, the experimental results and their analysis are presented. Finally, the conclusion, limitation, and future work are discussed in Section V.

II. Load Forecasting Techniques

There are various types of techniques in load forecast that have been proposed in the last few decades to improve the forecast accuracy based on regression model [8], and machine learning (ML) [9].

A. Statistical Methods

The currently used methods to achieve load forecast mainly includes statistical forecast methods and machine learning models [10]. Commonly used statistical forecast methods include the persistence method, Autoregressive Integrated Moving Average Model (ARIMA), which can be split into Autoregressive and Moving Average models, and its variants [11] [12] [13]. According to [14] and [15], the stationary model shows good performance only over the stationary data, while the traditional electric load is simple, without elements like renewable generation, smart home appliances, storage devices, etc.

The ARIMA method analyses time-series data based on the assumption that the collected and forecast dataset are linearly related, while the actual load pattern is highly non-smooth and non-linear [16]. Nowadays, the load linearity at transmission network (TN) and distribution network level are greater than in the past. For example, the increase of electric vehicle (EV) and distributed generation (DG) penetration contributes a huge portion of uncertain electricity generation and usage across the day [17]. The linear methods (persistence, ARIMA, etc.) show weakness in load forecasting especially in DN compared with the ML methods.

B. Machine Learning Methods

1) Artificial Neural Networks: ML methods have advantages of analysing non-linear data especially the Artificial Neural Networks (ANNs) [18], [19]. In contrast to the ARIMA method and its variants, ANNs include several non-linear self-adaptive methods, such as feed-forward Multilayer Perceptron (MLP) [20], Support Vector Regression (SVR), Fuzzy Logic (FL), Genetic Algorithms (GA), Random Forest (RF), Grey Projection Network (GPN), Back Propagation (BP), etc. [21], [22]. However, different ML methods have varies application ranges. When processing the time-series data, of which the electric load profile represents, the ML models like BP algorithms, RF, GPN, etc. are lack the consideration of time correlation of time-series data [23]. Also, FL and GA are focusing on different aspects of data processing, like data classification and solution optimisation [24]. In solving STLF problems, the Long Short-Term Memory (LSTM) method, based on Recurrent Neural Networks (RNN), is introduced which provides the ability to analyse non-linear data, taking into account the time correlation of time series data [23]–[25].

2) Long Short-Term Memory: The RNN is an improved method using the temporal information of the input data, where connections between units form a directed cycle within the same layer, while the ANN cannot find the relationship between data and time. Therefore, the output of each time step in the RNN is affected by the input data from previous steps [26], [27]. But the vanishing gradient problem affects the RNN when the input time-series data becomes deep and complex [28], [29]. In some cases, the gradient information during backpropagation progress will be vanishingly small, preventing the weight stored in the network from changing its value [30].

To ease the vanishing gradient problem, LSTM blocks are introduced to remember values for either long or short duration of time [31]. Hidden units in LSTM blocks trap the coming input data depending on the weight at the input and output gate, as shown in Figure 1. Therefore, in the backpropagation progress, the gradient can be propagated back across several time steps without exploding and vanishing. This helps the LSTM network learn the long-range dependencies of time-series.

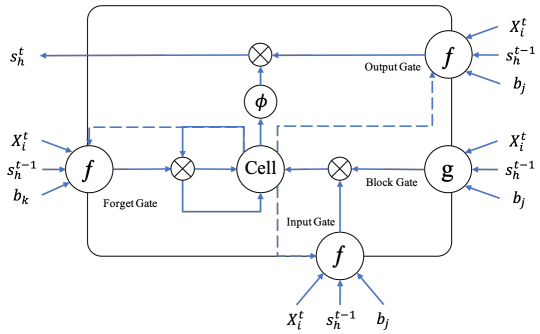


Fig. 1: The LSTM block

III. Model Development

This section presents the detailed methodology for developing the forecast model. It includes the data processing, peak detection method, and Dynamic Adaptive Compensation (DAC) module development. The block diagram is shown in Figure 3 and the structure of DAC-LSTM is shown in Figure 2. First, raw data is fed into the data processing module to generate the required data format. Then required dataset are passed to the LSTM forecast module. Because LSTM method cannot forecast both peak load and off-peak load accurately at once, the peak load detection module is added to distinguish peak and off-peak dynamically. According to the forecast error, the DAC module is used to improve forecast results for peak and off-peak periods separately.

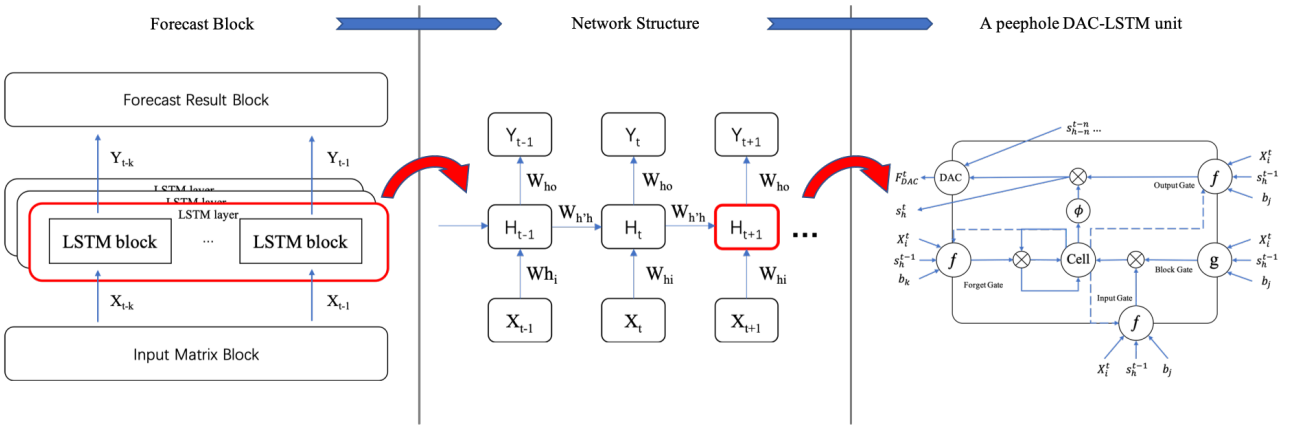


Fig. 2: DAC-LSTM network

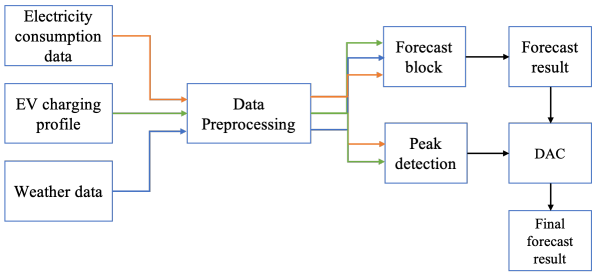


Fig. 3: Block diagram of the DAC-LSTM scheme

A. Data Description

There are three types of datasets used, including electrical load, ambient condition, and EV charging profile. The real-world historical load data is obtained from the Thames Valley Vision Project (TVVP) [7], which is half-hourly recorded, comes from 220 domestic properties, starts in February and March 2013 and ends in November 2014. Also, the weather data in the Thames Valley area is collected from the local weather station, from January 2013 to December 2014, including 16 features. The EV charging profiles is from Adaptive Charging Network (ACN), including charging profiles from charging stations [32]. We assume that EV users in the US are similar to them in the UK. According to the EV charging data source, the dataset is made by close collaboration with PowerFlex Systems around the US. To emulate the real load pattern scenarios with high EV penetrations, these two datasets (dataset from TVV and ACN) are combined. The detailed explanation can be found in section III-B.

B. Data input

The input dataset is pre-processed to clean the bad point and be expanded to a 5-min time step. Then the following steps are taken as the data processing progress.

1) Correlation analysis: This section presents the process to select features that are related to load forecast. The used feature data is recorded from the same area as the TVV Project and with the same period. The collected data includes 16 features while only 6 of them which show a potential relationship with load are selected.

2) Input features: The input weather data includes 6 features as shown in Table 1.1: load, tempC, weatherCode, humidity, cloudcover, wkorNot.

TABLE I: Input features and its represents

Term	Load	WeatherCode
Represent	Load(W)	Current weather
Term	tempC	Cloudcover
Represent	Temperature (°C)	Could cover rate (%)
Term	Humidity	wkorNot
Represent	Humidity	weekday or weekend

3) Correlation coefficient: The correlation analysis introduces the Pearson Correlation Analysis as it illustrates the correlation between time series data [33]. The following Table II is derived:

TABLE II: Numerical feature correlation coefficient

Index	Load	tempC	Humidity
Coefficient	1	-0.154	0.022
Index	Cloudcover	wkornot	weatherCode
Coefficient	-0.026	-0.02	-0.116

Table II shows the load has a positive correlation with humidity and a negative correlation with temperature. Because the weatherCode influences the weather, humidity, etc. features, the weatherCode feature is excluded.

C. Peak Detection

The LSTM forecast result cannot forecast both peak and off-peak load with high accuracy. Besides, when the

peak comes, the forecast result may exceed or below the actual peak value significantly due to the forecast hysteresis. Figure 4 shows the hysteresis of the forecast model and the large forecast error at the peak (compared with the off-peak). Therefore, before making the error correction upon the complete forecast result, the peak detection method is applied to distinguish peak and off-peak load. In the later section, methods are used to improve the average forecast error and minimise the forecast error during the peak.

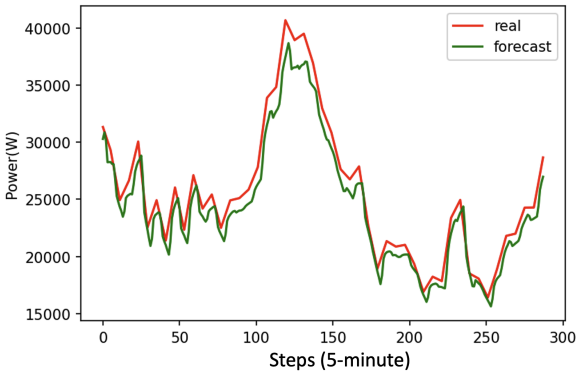


Fig. 4: Typical daily load and its forecast result based on LSTM method

Based on the principle of dispersion, Z-score is used [34]. According to Z-score method, the time-series will be regarded as peak only when:

$$|y - avg| > threshold * std \tag{1}$$

D. Dynamic Adaptive Compensation

This section describes the Dynamic Adaptive Compensation (DAC) methods, which are used to improve the forecast result dynamically. As Figure 5 shows, the DAC is an add-on function to the LSTM block. The overall block diagram is shown in Figure 6. The pre-forecast value is compared with the real load value, generating an error. According to the forecast step of the model and the repetitive forecast procedure, calculated errors are passed through a moving average error block which gives the average error value. Besides, the bias β_{bias} is generated, which gives a dynamic-adjust bias value. Also, the real-time forecast value passes through the Activation function, in this research sigmoid, to generate a parameter $\alpha_{sigmoid}$. The parameter $\alpha_{sigmoid}$ is fixed according to different forecast steps. Finally, Equation 2 is introduced to give the final forecast result.

$$F_{DAC}^{t+i} = \alpha_{sigmoid}^{i} \beta_{bias}^{i} E_i^t + S_h^{t+i} \tag{2}$$

$$E_i^t = S_h^t - R^t \tag{3}$$

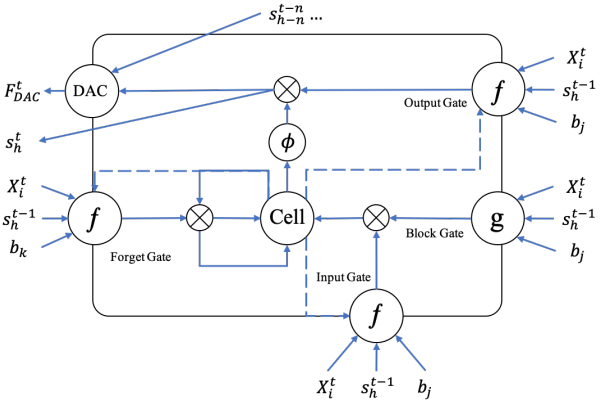


Fig. 5: A peephole of the DAC-LSTM cell unit

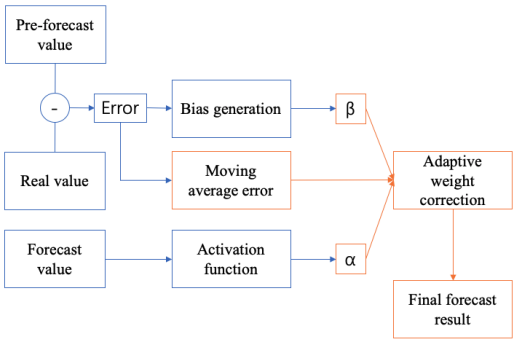


Fig. 6: Block diagram of DAC module

$$\alpha_{sigmoid}^{i} = f_\alpha(i) \tag{4}$$

$$\beta_{bias}^{i} = \frac{E_i^t}{E_{avg}} f_\beta(i) \tag{5}$$

In equations above, the present time is t and the forecast step is i. F_{DAC} is the forecast value from DAC-LSTM, $S_{h,i}$ is the forecast result from original LSTM block at i^{th} forecast step, R_i is the real recorded load data corresponding to i^{th} forecast step, E_i^t is the forecast error at time t, $\alpha_{sigmoid}^{i}$ is the forecast step parameter, fixed with step, and β_{bias}^{i} is the adaptive error parameter, $f_\alpha(i)$ is the sigmoid related function, $f_\beta(i)$ has positive correlation with $\frac{E_i^t}{E_{avg}}$ and $f_\beta(i)$ is a dynamic changing parameter. The maximum compensation value is preset to avoid infinity problems.

The equation of $f_\alpha(i)$ is described as follow:

$$f_\alpha(i) = 1.5 - \frac{1}{1 + e^{(-i+1)/2}} \tag{6}$$

The equation of $f_\beta(i)$ is described as follow:

$$f_\beta(i) = k\delta(\frac{1-i}{m} + 1) + 1 \tag{7}$$

$$\delta = \frac{2}{5 + 5e^{n(1-\frac{E_i^t}{E_{avg}})}} - \frac{1}{5} \tag{8}$$

In equation 8, the max/minimum value is limited to ± 0.2, which represents around 80-120% k error compensation (depends on which step), as shown in Figure **??**. m and n are constant, of which the test dataset determines values. k is the compensation factor, determining the upper and lower bound of the compensated P.U. value. If the E^t keeps increasing compared with the previous steps, k increases with a step of 0.01 and vice-versa. The default value of k is 1 and k is reset to default when $\frac{E_i^t}{E_{avg}}$ changes the positive ad negative. In this block, the value used to compensate for forecast error changes according to the previous forecast error and the forecast steps. $\alpha_{sigmoid\_i}$ helps avoid the non-convergence of the DAC method, and β_{bias_i} helps the model to increase forecast accuracy when the convergence is decreasing. Besides, in this procedure, the compensated forecast result exceeds or is less than the actual value significantly when the Error increase. Therefore, to improve the forecast accuracy during the peak, the peak detection and compensation cap methods are used to avoid over or under-compensation. The result with or without a cap is shown in Figure 7. The maximum/minimum compensation value is set to $\pm 0.2k\ E_i^t$.

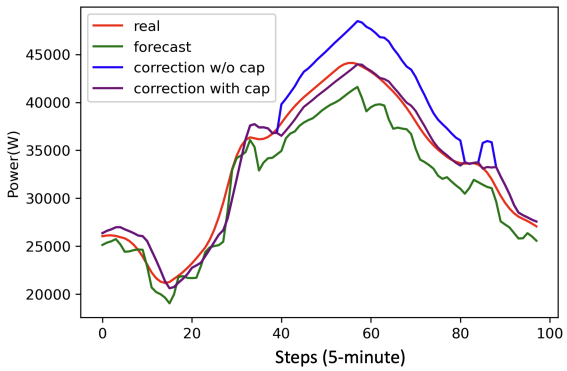


Fig. 7: The performance of applying peak cap

IV. CASE STUDY

In this section, the performance of the DAC-LSTM method is evaluated and carried out with case studies. Moreover, the ELEXON UK domestic load data is used to test the application of DAC-LSTM in a higher-level power network. When evaluating the DAC-LSTM, the following settings are made:

1) Each model is trained 10 times, and then takes the average error value to show a steady performance of DAC-LSTM.
2) When training the LSTM model, the model parameters are adjusted according to the training data pool size, and forecast steps.
3) Each model is trained for 500 epochs and then returns to the best epoch.
4) When needed, the input dataset is expanded to 5 min time step with linear interpolation and random errors ($\pm$ 1.5%).

A. Performance Matrices

The performance is evaluated by using the following values: R-squared (R2) and Mean Absolute Percentage Error (MAPE).

$$R^2 = 1 - \frac{UnexplainedVariation}{TotalVariation} \tag{9}$$

$$MAPE = \frac{1}{n}\sum |\frac{A_i - F_i}{A_i} \tag{10}$$

B. Input data description

1) Case 1-3: The data used to train the model is from TVVP, ACN and locl meteorological stations. The detailed description can be found in Section III-A.

2) Case 4: The data used to train the model is obtained from ELEXON Portal. The dataset contains the load profile from the UK domestic, from January 2013 to December 2014. The date is half-hourly recorded. Only one feature (load) is used for model training.

C. Case 1: Various length of historical data

The first case study compares forecast results with different amounts of days' data for training. First, the model is tested with different training pools, from 550 days of data to 14 days of data. The expected forecast step is 6, which represents 30 minutes forecast. This case study aims to test the model's robustness when input historical data is insufficient.

Table III illustrate the R2, MAPE, and correlation values according to different amounts of days for training by using Persistence, ARIMA, and DAC-LSTM methods. It shows with the reduction of training samples, the forecast accuracy of these methods reduces. But the DAC-LSTM method provides better forecast accuracy compared with the other two. From less than 150 days, the forecast accuracy of DAC-LSTM reduces significantly because of the lack of historical data for training. The forecast accuracy from DAC-LSTM remains over 0.95 by using more than 100-day historical data. This experiment gives that the DAC-LSTM methods can abstract load features with much less historical data compared with ARIMA and Persistence (from 550 to 150 days for this dataset).

TABLE III: Case 1: R2, MAPE and Correlation

	R2			MAPE			Correlation		
Amount of days	Per.	ARIMA	DAC-LSTM	Per.	ARIMA	DAC-LSTM	Per.	ARIMA	DAC-LSTM
550	0.881	0.920	0.991	8.71	6.82	4.04	0.940	0.959	0.994
500	0.867	0.891	0.990	8.05	7.76	6.46	0.934	0.942	0.993
450	0.913	0.909	0.991	9.29	7.02	5.49	0.956	0.961	0.996
400	0.864	0.833	0.991	8.76	8.83	5.65	0.932	0.914	0.995
350	0.898	0.892	0.990	13.41	7.61	4.45	0.949	0.946	0.995
300	0.898	0.874	0.975	8.00	7.92	5.01	0.949	0.939	0.993
250	0.883	0.897	0.989	7.91	7.36	3.84	0.941	0.947	0.994
200	0.893	0.894	0.990	8.50	7.42	4.43	0.947	0.944	0.991
150	0.898	0.891	0.998	8.29	7.77	3.86	0.949	0.953	0.991
100	0.876	0.873	0.958	7.21	6.95	3.78	0.938	0.943	0.992
50	0.850	0.825	0.926	8.81	8.38	5.83	0.925	0.913	0.997
30	0.795	0.794	0.904	9.95	9.87	6.61	0.897	0.899	0.997
14	0.824	0.820	0.875	7.75	7.75	6.52	0.913	0.913	0.995

TABLE IV: Case 2: R2, MAPE and Correlation

	R2			MAPE			Correlation		
Amount of households	Per.	ARIMA	DAC-LSTM	Per.	ARIMA	DAC-LSTM	Per.	ARIMA	DAC-LSTM
220	0.881	0.920	0.991	8.71	6.82	4.04	0.940	0.923	0.994
150	0.896	0.867	0.991	10.12	11.24	4.20	0.948	0.932	0.993
100	0.860	0.855	0.987	11.55	16.62	4.75	0.930	0.925	0.992
50	0.736	0.743	0.986	15.55	19.49	6.47	0.868	0.864	0.989
30	0.636	0.658	0.970	19.98	19.84	8.32	0.818	0.822	0.989
15	0.591	0.587	0.934	22.34	21.23	9.12	0.792	0.714	0.981

D. Case 2: Various sizes of data pool

The second case study compares forecast results with different amounts of households' data used for training. Historical days for training is fixed at 550. The expected forecast step is 6, which represents 30 minutes forecast. The trained models are used to forecast 220 households' electrical load.

Table IV illustrate the R2, MAPE, and correlation values according to different amounts of households for training by using Persistence, ARIMA, and DAC-LSTM methods. It shows with the reduction of training samples, the forecast accuracy reduces. The DAC-LSTM method provides better forecast accuracy when the data pool is reduced. The DAC-LSTM forecast accuracy remains over 0.95 by using more than 30-household historical data. For less than 30 households, the forecast accuracy of DAC-LSTM reduces significantly because the periodicity reduces and nonlinearity increases, which restricts the model training. This experiment gives that the DAC-LSTM method can abstract load features with much fewer data sources compared with ARIMA and Persistence (from 220 to 30 for this dataset)

E. Case 3: Comparison between DAC-LSTM and LSTM methods with various forecast steps

Third, the proposed DAC-LSTM method is compared with the original LSTM method. The forecast step varies from 30 minutes (1 step) to 24 hours (48 steps). The forecast results comparison is shown in Figure 8.

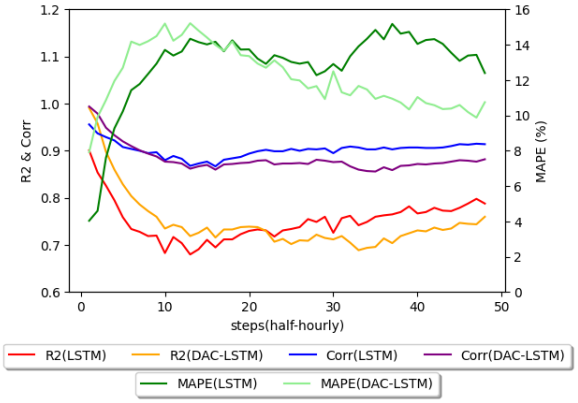


Fig. 8: The forecast results comparison between LSTM and DAC-LSTM methods (forecast step varies from 0.5 hour to 24 hours)

In Figure 8, the DAC-LSTM method shows significant improvement when the forecast step is less than 12 hours. When the forecast steps range from 1 (30 minutes) to 18 (9 hours), the DAC-LSTM method is better than the LSTM method, while the percentage of improvement varies from 12.2% (R2) and 53% (MAPE) to 3.1% (R2) and 2.7%.

F. Case 4: Compare with ELEXON UK domestic load

In this case study, the DAC-LSTM method is applied to the UK domestic load derived from ELEXON to expand the application range. Because the ELEXON data is derived from UK domestic, the load profile is more smooth and shows less distortion among days when compared with distribution network load data. The dataset is recorded half-hourly and without linear interpolation before training. Table V shows the real and forecast accuracy from ELEXON and DAC-LSTM. The R2 of the ELEXON forecast and DAC-LSTM forecast results are 0.97 and 0.99 separately, while the DAC-LSTM method gives less average error.

TABLE V: Forecast result: UK domestic load dataset

	R2	**MAPE (%)**	RMSE (%, normalised)
ELEXON	0.97	0.17	0.13
DAC-LSTM	0.99	0.13	0.11

V. Conclusion and Future Work

In this work, a robust DAC-LSTM forecast model is proposed for short-term electricity load forecast. The proposed method features a non-linear electric load forecast, using the big data collected from monitoring devices which gives the model a quicker response to load pattern changes, therefore higher forecast accuracy. Also, compared with existing forecast methods, the DAC-LSTM can provide accurate forecast results when the trained model is lack historical data and data sources, or the trained model unfits the load pattern. Moreover, the method is applied to the domestic load and shows convincing results compared with the dataset from ELEXON. The application of the method ranges from distribution level network (highly non-linear) to domestic load level. All these features offer opportunities for the electricity supplier and the grid operator to improve the existing services, DR plan, real-time pricing schemes, etc. For example, in a P2P trading system, more accurate STLF methods could reduce the Ethereum gas trading cost, therefore reducing the operation cost and improving efficiency. Also in the network with high DG penetration, the share of renewable generation could be increased with more accurate forecast methods to reduce carbon footprint.

The DAC-LSTM method has good performance compared with other widely used STLF methods, some works remain for the future. The application of the DAC-LSTM method can be extended to MTLF and LTLF which requires further experiment upon the existing model. Also, the forecast accuracy and robustness of the proposed method can be tested with more historical data collected from the various distribution networks.

In addition, the proposed method has limitations. Though the DAC-LSTM method utilises the advantage of big data and machine learning techniques, it requires massive data collected from advanced monitoring devices and massive computation ability. Moreover, Although the ML techniques have been developed for decades and applied in the power system operation, the real-world application of the DAC-LSTM method still needs time to be validated.

Acknowledgment

The authors would like to thank Thames Valley Vision Project for sharing the user electrical consumption data.

References

[1] J. Lago, F. De Ridder, P. Vrancx, and B. De Schutter, "Forecasting day-ahead electricity prices in europe: the importance of considering market integration," *Applied energy*, vol. 211, pp. 890–903, 2018.

[2] N. Singh, S. R. Mohanty, and R. D. Shukla, "Short term electricity price forecast based on environmentally adapted generalized neuron," *Energy*, vol. 125, pp. 127–139, 2017.

[3] C. Nataraja, M. Gorawar, G. Shilpa, and J. S. Harsha, "Short term load forecasting using time series analysis: a case study for karnataka, india," *Int. J. Eng. Sci. Innov. Technol*, vol. 1, no. 2, pp. 45–53, 2012.

[4] C. Cecati, C. Citro, and P. Siano, "Combined operations of renewable energy systems and responsive demand in a smart grid," *IEEE transactions on sustainable energy*, vol. 2, no. 4, pp. 468–476, 2011.

[5] N. Ding, C. Benoit, G. Foggia, Y. Bésanger, and F. Wurtz, "Neural network-based model design for short-term load forecast in distribution systems," *IEEE transactions on power systems*, vol. 31, no. 1, pp. 72–81, 2015.

[6] D. Syed, A. Zainab, A. Ghrayeb, S. S. Refaat, H. Abu-Rub, and O. Bouhali, "Smart grid big data analytics: Survey of technologies, techniques, and applications," *IEEE Access*, vol. 9, pp. 59 564–59 585, 2020.

[7] "Thames valley vision project." [Online]. Available: http://www.thamesvalleyvision.co.uk/

[8] S. Kumar, S. Mishra, and S. Gupta, "Short term load forecasting using ann and multiple linear regression," in *2016 second international conference on computational intelligence & communication technology (cict)*. IEEE, 2016, pp. 184–186.

[9] X. Tang, Y. Dai, Q. Liu, X. Dang, and J. Xu, "Application of bidirectional recurrent neural network combined with deep belief network in short-term load forecasting," *IEEE Access*, vol. 7, pp. 160 660–160 670, 2019.

[10] P. Zhang, X. Wu, X. Wang, and S. Bi, "Short-term load forecasting based on big data technologies," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 3, pp. 59–67, 2015.

[11] F. Wu, R. Jing, X.-P. Zhang, F. Wang, and Y. Bao, "A combined method of improved grey bp neural network and meemd-arima for day-ahead wave energy forecast," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 4, pp. 2404–2412, 2021.

[12] E. Gonzalez-Romera, M. A. Jaramillo-Moran, and D. Carmona-Fernandez, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Transactions on power systems*, vol. 21, no. 4, pp. 1946–1953, 2006.

[13] J. C. López, M. J. Rider, and Q. Wu, "Parsimonious short-term load forecasting for optimal operation planning of electrical distribution systems," *IEEE transactions on power systems*, vol. 34, no. 2, pp. 1427–1437, 2018.

[14] K. M. Kam, *Stationary and non-stationary time series prediction using state space model and pattern-based approach*. The University of Texas at Arlington, 2014.

[15] M. Lineesh, K. Minu, and C. J. John, "Analysis of nonstationary nonlinear economic time series of gold price: A comparative study," in *International Mathematical Forum*, vol. 5, no. 34. Citeseer, 2010, pp. 1673–1683.

[16] D. Janardhanan and E. Barrett, "Cpu workload forecasting of machines in data centers using lstm recurrent neural networks and arima models," in *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, 2017, pp. 55–60.

[17] A. Dubey and S. Santoso, "Electric vehicle charging on residential distribution systems: Impacts and mitigations," *IEEE Access*, vol. 3, pp. 1871–1893, 2015.

[18] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-term load forecasting in smart grids: an intelligent modular approach," *Energies*, vol. 12, no. 1, p. 164, 2019.

[19] P. Singh and P. Dwivedi, "Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem," *Applied energy*, vol. 217, pp. 537–549, 2018.

[20] J. Faraji, A. Ketabi, H. Hashemi-Dezaki, M. Shafie-Khah, and J. P. Catalao, "Optimal day-ahead scheduling and operation of the prosumer by considering corrective actions based on very short-term load forecasting," *IEEE Access*, vol. 8, pp. 83 561–83 582, 2020.

[21] Z. Tan, J. Zhang, Y. He, Y. Zhang, G. Xiong, and Y. Liu, "Short-term load forecasting based on integration of svr and stacking," *IEEE Access*, vol. 8, pp. 227 719–227 728, 2020.

[22] H. Jiang, Y. Zhang, E. Muljadi, J. J. Zhang, and D. W. Gao, "A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3341–3350, 2016.

[23] C. Zhuo and S. Long-Xiang, "Short-term electrical load forecasting based on deep learning lstm networks [j]," *Electronic Technology*, vol. 10, no. 39, pp. 122–125, 2018.

[24] M. Z. A. Bhotto, R. Jones, S. Makonin, and I. V. Bajić, "Short-term demand prediction using an ensemble of linearly-constrained estimators," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 3163–3175, 2021.

[25] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network," *Energies*, vol. 11, no. 12, p. 3493, 2018.

[26] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[27] C. Yang, W. Jiang, and Z. Guo, "Time series data classification based on dual path cnn-rnn cascade network," *IEEE Access*, vol. 7, pp. 155 304–155 312, 2019.

[28] W. Zhang, H. Quan, O. Gandhi, R. Rajagopal, C.-W. Tan, and D. Srinivasan, "Improving probabilistic load forecasting using quantile regression nn with skip connections," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5442–5450, 2020.

[29] M. Dong and L. Grumbach, "A hybrid distribution feeder long-term load forecasting method based on sequence prediction," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 470–482, 2019.

[30] S. Basodi, C. Ji, H. Zhang, and Y. Pan, "Gradient amplification: An efficient way to train deep neural networks," *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 196–207, 2020.

[31] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, "An overview and comparative analysis of recurrent neural networks for short term load forecasting," *arXiv preprint arXiv:1705.04378*, 2017.

[32] Z. J. Lee, T. Li, and S. H. Low, "ACN-Data: Analysis and Applications of an Open EV Charging Dataset," in *Proceedings of the Tenth International Conference on Future Energy Systems*, ser. e-Energy '19, Jun. 2019.

[33] M. Bermudez-Edo, P. Barnaghi, and K. Moessner, "Analysing real world data streams with spatio-temporal correlations: Entropy vs. pearson correlation," *Automation in Construction*, vol. 88, pp. 87–100, 2018.

[34] J. van Brakel, "Robust peak detection algorithm (using z-scores)," *URL https://stackoverflow.com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data*, 2019.

Yuanda Gao received the B.Eng. and M.Sc. degrees from the University of Edinburgh, Edinburgh, U.K. in 2017 and 2018, respectively. He is currently pursuing the Ph.D. degree in the University of Edinburgh. His research interests include power system analysis, machine learning, data analysis and demand-side management.

PLACE PHOTO HERE

James Robertson received the M.Eng and Ph.D. degrees from the University of Edinburgh, Edinburgh, U.K. in 2010 and 2015, respectively. From 2014 to 2022 he was a Research Associate with the University of Edinburgh. He is currently a European Power Analyst with Wood Mackenzie. His research interests include power systems analysis and the network integration of renewable energy resources.

PLACE PHOTO HERE

Aristides Kiprakis (Senior Member, IEEE) received the B.Eng. degree in electronic engineering from the Technological Educational Institute of Crete, Greece, in 1999, the Diploma degree in communications, control, and DSP from the University of Strathclyde, Glasgow, U.K., in 2000, and the Ph.D. degree in electrical power engineering from The University of Edinburgh, Edinburgh, U.K., in 2005. He is currently a Senior Lecturer in Power Systems with the School of Engineering, The University of Edinburgh, where he leads the Agile Energy Systems Group. He is also the Director of the M.Sc. in Electrical and Advanced Power Engineering Program and the School's Director for International Students. Dr. Kiprakis is a member of the UK Institution for Engineering and Technology (IET) and a STEM Ambassador for East Scotland.