Short-term Industrial Load Forecasting for Battery Energy Storage System Simulation

Zhichao Wu, Thomas Blank, Simon Bischof, Marc Weber
Institute for Data Processing and Electronics
Karlsruher Institut für Technologie
Karlsruhe, Germany
zhichao.wu@kit.edu

Abstract — **Load forecasting is an important research area for the operation of an energy system, especially when renewable generation and energy storage are introduced. In the context of industrial energy users with a high peak load during the production phases, load forecasting allows the system to utilize battery energy storage systems (BESS) and renewable energy, i.e. photovoltaics for peak shaving. In this case, the industrial electricity consumers and the network operators can both benefit from the reduced cost related to the high peak load charge and more reliable and efficient operation and management of the grid network respectively. However, the performance of peak shaving depends to a certain extent on the precision of the load curve prediction. This paper presents a short-term load prediction comparison of an industrial consumer based on a) a model-driven (time series) and b) a data-driven (machine learning) approach. The electric load data are collected from the fabrication plant at Karlsruher Institut für Technologie and used to train the model for load forecasting. The verification test was conducted by simulating the BESS in the energy system to compare its execution in peak shaving with different forecasting methods. The result proves that the machine learning method with recurrent neural network (RNN) provides more accurate and robust predictions so that the utilization of the capacity of BESS can be promoted to improve the peak shaving performance and cost reduction.**

Keywords-component; short-term industrial load forecasting; time-series model; battery energy storage system (BESS); machine learning; recurrent neural network (RNN); long short-term memory (LSTM)

I. INTRODUCTION

Renewable energy has been adopted to solve the contradiction between the growing economy and the increasing energy shortage on a global scale. The annual increase in renewable power generation capacity has been expanding rapidly in the last twenty years. And the renewable share of the annual installed capacity has increased from 25% in 2001 to over 80% in 2020 [1]. However, due to the simultaneity of power generation and consumption, increasing penetration of renewable energy brought challenges to the safe and stable operation of the grid. Therefore, large-scale and distributed energy storage systems are used for peak-shaving in different cases to eliminate the inhomogeneous distribution of the redundant renewable generation over time and space [2]. Among the different energy storage technologies, Battery Energy Storage System (BESS) is flexible to achieve the function of peak-shaving with high efficiency and fast response time, which makes it an ideal solution for small-scale renewable energy integration, e.g., residential and small-scale industrial users [3]. For the power supply and grid, the integration of BESS plays an important role in delaying the upgrade of the transmission equipment and smoothing out the power fluctuation in the distribution network. The instant charging and discharging characteristics of BESS can reduce the impact of renewable energy on the distribution network and enhance the controllability of the distribution network [4]. On the other hand for the network users, BESS improves the renewable energy share in the energy system when photovoltaics or other renewable energy is supplied and reduces the peak power to obtain economic benefits by saving network charge based on the peak usage. Thus, despite the large-scale BESS market in Germany and Europe are suffering from a decline due to the pandemic, the turnover for residential and commercial-scale BESS still rose dramatically in 2020 [5].

Typically, there are two types of real-time dispatch strategies for BESS when it is used for peak shaving. One is to set a pre-determined peak shaving threshold, and the other is to track a calculated load curve according to actual load. There is a lack of flexibility with the fixed threshold, some adaptive algorithms could be applied to partly make up the deficiency [1]–[3]. The tracking mode is more intelligent and utilizes the capacity of BESS more efficiently but the battery may not be able to keep charging or discharging when it is full or running low. Since the BESS is supposed to work at the "hourly" level in terms of time duration when it is used for peak-shaving, the accurate short-term load forecasting (STLF) of the next few hours is the essential element to plan the charging and discharging of the battery. The effect of peak-shaving using BESS depends on the accuracy of the forecasted load curve and renewable generation to a great extent [4], [5]. Regarding load forecasting, there is a wide availability of research on time series analysis or neural network models [6]–[10]. Most of the studies focus on the forecasting model to improve the accuracy of the prediction. However, there is limited analysis on the impact of different forecasting methods on BESS.

On that account, a comparison of the different load forecasting approaches is conducted in this work on the electric load of the fabrication plant at the Institute of Data Processing and Electronics (IPE) of Karlsruhe Institute of Technology (KIT). The load profile is collected as the High-resolution Industrial Production Energy dataset (HIPE) from the factory building which has been instrumented with smart meters[11]. This work aims to investigate the performance of peak shaving with BESS in an industrial application using different load forecast models. The load forecasts based on the time-series approach using the Autoregressive Integrated Moving Average model (ARIMA) and the recurrent neural network (RNN) approach using Long Short-Term Memory (LSTM) are conducted. The load prediction is used for BESS controlling in the simulation. The results prove that the accuracy improvement with load forecast is meaningful to BESS optimization. The main contributions and novelty of this work are the utilization of a high-resolution load profile for load forecasting and the comparative analysis of its impact on the BESS with renewable power integrated into the system.

II. Related Work and Background

The subject of this paper is the small industrial energy consumer. Typically, the load characteristics of these consumers are not steady like those of large industrial consumers. The peak power during the production process when the equipment is running goes to a significantly higher level than the annual average power. Besides, there is a lack of a fixed schedule of the daily production. Therefore, the moment when the peak power comes on a single day is uncertain. Usually, the grid usage cost of this kind of consumer can be reduced by installing optional smart meters to separate the charge into capacity price and energy price. The network usage fee in the fabrication plant under investigation in this work cut down more than 5% with just the smart meters mechanism. The integration of BESS and renewable energy is possible to further improve the cost reduction and renewable shares in the system. However, the BESS would have lacked impact if load forecasting is neglected or inaccurate[4].

Generally, load forecasting can be divided into short-term load forecasting (STLF), medium-term load forecasting (MTLF), and long-term load forecasting (LTLF). The prediction for the period that is less than one week is categorized as STLF[12]. Traditional time series models such as autoregressive (AR) [13], [14], autoregressive moving average (ARMA) [15] has been introduced for a long time. The autoregressive integrated moving average (ARIMA) was upgraded from ARMA through differencing process to deal with non-stationary time series [16]. Juberias et al. [17] produce an hourly forecast with the ARIMA model for the real-time control system. With the development of machine learning in recent years, some hybrid models have been proposed to take advantage of the strength of ARIMA and Artificial Neural Networks (ANN) models [18], [19]. A hybrid model with a K-means clustering algorithm was proposed to classify the annual data into clusters [20]. The performance of the forecast of the electricity peak load has been improved compared to using the ARIMA model alone.

In contrast, the data-driven approach is based on artificial intelligence. The ANN is one of the popular methods for load forecasting because of the ability to model nonlinearity. Park et al. [21] presented the ANN approach to interpolate training data and provide future load patterns. Hernandez et al. [22] presented an STLF with Artificial Neural Network (ANN) model based on a three-layer Multi-Layer Perceptron (MLP) to perform day-ahead load forecasting. Besides MLP, various structures of ANN have been used in load forecasting to improve the accuracy [23]. Ryu et al. [6] proposed a DNN-based load forecasting model and apply them to demand-side empirical load. The deep neural network (DNN) belongs to ANN but with a more hidden layer structure to improve its capability of feature abstraction.

Some other machine learning methods achieve decent results as well. Pereira et al. [24] apply a Fuzzy Inference System (FIS) to predict the electric load as time series. The FIS forecasting system delivers a low error but lost its interpretability due to the large number of rules it generated. Support Vector Machine (SVM) is a popular model used for non-linear time series modeling. Zhang et al. [25] and Jain et al. [26] present a load forecasting model based on Support Vector Regression (SVR) which achieves high accuracy in load forecasting. Wang et al. [27] conducted a comparative forecast of energy use in building with ANN, SVR, and long short-term memory (LSTM). The LSTM model which has the structure of a recurrent neural network achieves prediction with the smallest Mean Absolute Percentage Error (MAPE). Recently the LSTM model has been becoming attractive because it can learn long-term dependencies from the historic load and has been used in load forecasting for both residential households [28], [29] and industrial consumers [7], [30]. Several pieces of research were conducted to further improve the LSTM approach. Marino et al. [31] demonstrate an LSTM-based Sequence to Sequence (S2S) architecture to deal with data with a high resolution of one minute. Zheng et al. [32] improve LSTM by using Xgboost and clustering for features importance evaluation and merging similar days. Bouktif et al. [10] tailor the Genetic Algorithm (GA) and the Particle Swarm Optimization (PSO) algorithm to tune the hyperparameters of the LSTM model for load forecasting.

Although there is plenty of literature on load forecasting, the experimental validation of the load forecasting used for peak shaving with BESS in an energy system is quite limited. Barelli et al. [33] analyze the energy storage system in a micro-grid with the residential load predicted by ANN. Soman et al. [5] also present an LSTM model for peak shaving in campus micro-grid. However, both of these works still focus on the accuracy of the forecast. The only relevant literature available on the impacts of load forecasting on BESS is conducted by Papadopoulos et al. [4], where they simply applied the ANN model and compared it to the situation without prediction. Therefore, this paper presents a comparative assessment of peak shaving with load forecasting by time series analysis and data-driven model. Additionally, the BESS and renewable energy generation are also integrated into the system and model.

III. Data Analysis and Forecasting Framework

A. Dataset

The data used for training and simulation in this work includes the industrial load data collected from the electronics production sites at the Karlsruhe Institute of Technology (KIT) named as the HIPE dataset [11] and the weather information collected at the closest climate station

(Stations_ID: 4177) of German Weather Service (DWD) provided on the DWD open data server [34].

The factory building is instrumented with high-resolution smart meters. The measurement of the weekly load profile chosen randomly from the four seasons is compared in Fig. 1. From the figure, there is no clear seasonal variation during the peak hour. The original data is recorded with a resolution of about 5 seconds. Thus, due to the high resolution, the raw data is slightly noisy. Since the annual maximum power charged for grid usage generally refers to the aggregated smart meter readings with a 15-minutes resolution [35], the load forecasting resolution should be no less than 15-minutes. In this work, we implement the load profile with three different resolutions: 15-minutes, 5-minutes, and 1-minutes.

B. Time Series Analysis

1) ARIMA Models

Time series is one of the most popular fields in the research on load forecasting. Traditionally, the model most commonly used for time series analysis and forecasting are as follows, including

- Autoregressive (AR) models
- Moving Average (MA) models
- Autoregressive Moving Average (ARMA) models
- Autoregressive Integrated Moving Average (ARIMA) models

The AR models describe the linear relationship between current and historical values using the previous value of the time series, while the MA models represent the current value in terms of a linear combination of random disturbances or forecast errors from previous. The ARMA model is a combination of the AR and MA, where the current value of the time series is represented linearly as its previous value and disturbances. The ARIMA model applies an initial differencing step to eliminate the non-stationarity of the series. The full ARIMA model of order p, q, and d for a time series notated as ARIMA(p,d,q) is defined as (1):

$$\boldsymbol{y}_t' = \sum_{i=1}^{p} \phi_i \, \boldsymbol{y}'_{t-i} + \delta + \boldsymbol{\varepsilon}_t + \sum_{j=1}^{q} \theta_j \, \boldsymbol{\varepsilon}_{t-j} \qquad (1)$$

where $\boldsymbol{y}_t'$ is the differenced series, $\boldsymbol{\varepsilon}_t$ is the white noise error series, p is the order of the autoregressive part, q is the order of the moving average part, and d is the degree of first differencing involved, δ is the constant. Equation (1) can be further written in the backshift notation as (2) with lag operator L:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right)(1-L)^d \boldsymbol{y}_t = \delta + \left(1 + \sum_{j=1}^{q} \theta_j L^j\right) \boldsymbol{\varepsilon}_t \qquad (2)$$

The seasonal ARIMA (SARIMA) model includes seasonal terms into the ARIMA models. It can be written as ARIMA(p,d,q)$(P,D,Q)_m$. where $(P,D,Q)_m$ is the seasonal parts of the model, m is the number of time steps for a single seasonal period.

2) Model Fitting

The Augmented Dickey-Fuller (ADF) test is used to identify the order of non-seasonal and seasonal differencing. The ADF test shows that the time series is stationary. Therefore, the non-seasonal and seasonal differencing orders d and D are set to be zero. The autocorrelation function (ACF) and partial autocorrelation function (PACF) of the load data with 15-minutes resolution are illustrated as Fig.2. According to ACF, the periodical sinusoid reaches the peak at 96 lags, which is approximately the time for 1-day. The seasonal period m is considered to be 96. The ARIMA model used to fit the load data with a 15-minutes resolution can be written as ARIMA(p,0,q)$(P,0,Q)_{96}$.

Fig.3 shows the ACF and PACF after seasonal difference. From Fig.2 and Fig.3, both ACFs are decaying of sinusoidal,

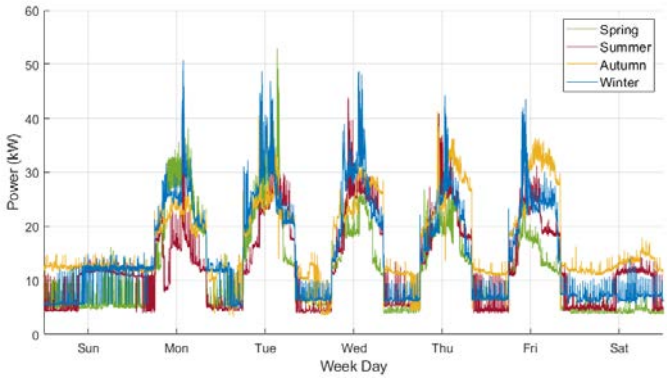

Figure 1. Weekly factory load measured by smart meters

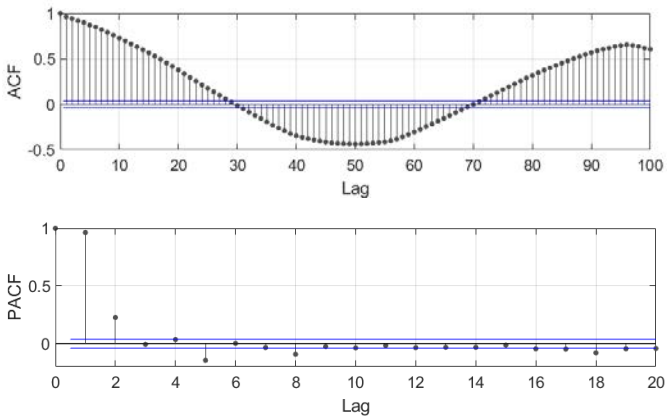

Figure 2. Autocorrelation function (ACF) and partial autocorrelation function (PACF) of the load

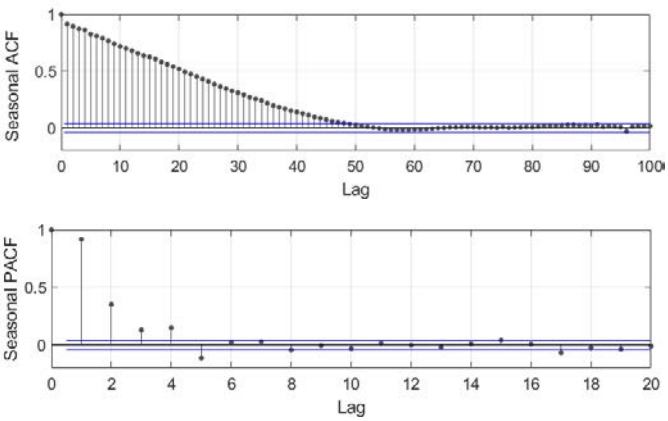

Figure 3. Autocorrelation function (ACF) and partial autocorrelation function (PACF) of the load after seasonal differencing

which suggests q and Q might be zero. The spike in PACF remains outside of the confidence bounds until lag 2 in Fig.2. The spike of PACF after lag 2 also drops significantly in Fig.3. Thus, ARIMA(2,0,0)(2,0,0)$_{96}$ is possible to be a suitable model. The order of the model for fitting the load profile is determined by minimizing the Akaike information criteria (AIC) value, which is an estimator of the relative quality of statistical models for time-series data. Table I presents the AIC values of the potential ARIMA models. From Table I, the smallest AIC is obtained with ARIMA(2,0,1)(3,0,1)$_{96}$. However, the AIC difference between ARIMA(2,0,1)(2,0,1)$_{96}$ and ARIMA(2,0,1)(3,0,1)$_{96}$ is negligible, but requires much less computation. The residuals with the ARIMA(2,0,1)(2,0,1)$_{96}$ model are also within a certain level and conforms to a random distribution. Therefore, ARIMA(2,0,1)(2,0,1)$_{96}$ is used for 15-minutes resolution load forecasting. Similarly, we apply the ARIMA(2,0,1)(2,0,1)$_{288}$ model for the 5-minutes resolution load data. However, the load with 1-minutes resolution is not used for the time-series method, because the computation required from the ARIMA(2,0,1)(2,0,1)$_{1440}$ model is impractical to apply for load forecasting.

TABLE I. TABLE THE AIC OF THE ARIMA MODELS

ARIMA(p,d,q)(P,D,Q)$_m$	AIC
ARIMA(2,0,0)(2,0,0)$_{96}$	11075
ARIMA(2,0,1)(2,0,1)$_{96}$	10910
ARIMA(2,0,1)(2,0,0)$_{96}$	11064
ARIMA(2,0,0)(2,0,1)$_{96}$	10919
ARIMA(3,0,0)(2,0,0)$_{96}$	11067
ARIMA(3,0,1)(2,0,1)$_{96}$	10911
ARIMA(2,0,0)(3,0,0)$_{96}$	11099
ARIMA(2,0,0)(3,0,1)$_{96}$	10912
ARIMA(2,0,0)(3,0,1)$_{96}$	10998
ARIMA(2,0,1)(3,0,1)$_{96}$	10903

C. Recurrent neural network

1) Neural network model

Feedforward neural networks (FFNN) can be used for load forecasting [4], [36], but FFNN is a static network and does not have a memory function, which prevents it from making full use of previous load information.

In contrast, Recurrent Neural Network (RNN) is a sequence-based neural network with memory as shown in Fig.4. Neurons in the recurrent neural network can receive information from other neurons and themselves, which creates a loop structure including input layer hidden layer and output layer. U, W, and V are the weight matrices for the input-to-hidden, hidden-to-hidden, and hidden-to-output connection respectively, which are shared across time. Regarding the load forecasting problem, RNN can establish a temporal correlation between the current state and historical load. The parameters of recurrent neural networks can be trained using the Back Propagation Through Time algorithm (BPTT). However, the gradient in BPTT accumulates exponentially with time, which may lead to the Long-Term Dependencies Problem [37]. To avoid this problem from gradient explosion or vanishing due to the long input sequence for load forecasting, we use the Long Short-Term

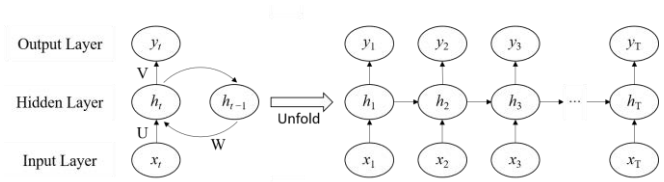

Figure 4. Structure of the Recurrent Neural Network.

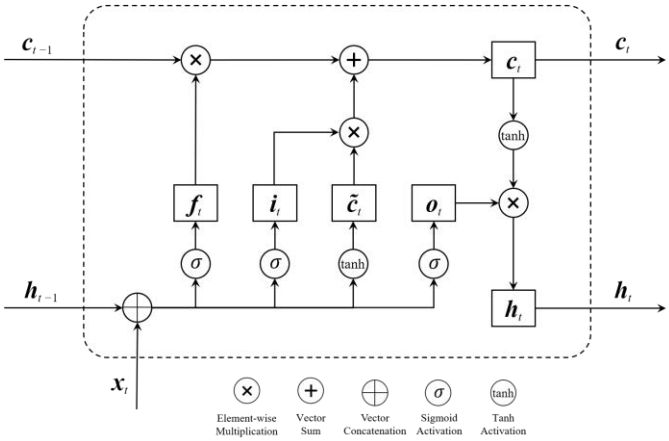

Figure 5. Structure of an LSTM cell.

Memory network (LSTM) model which introduces the gating mechanism to control the path of information transfer[38], [39].

LSTM introduces an internal state $\boldsymbol{c}_t$ and an external state $\boldsymbol{h}_t$ for cyclic message passing. The structure of an LSTM cell is illustrated in Fig.5. The internal state $\boldsymbol{c}_t$ and the external state $\boldsymbol{h}_t$ are calculated by (3) and (4).

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \tilde{\boldsymbol{c}}_t \tag{3}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot tanh(\boldsymbol{c}_t) \tag{4}$$

where $\boldsymbol{f}_t$, $\boldsymbol{i}_t$, and $\boldsymbol{o}_t$ are three gates to control the transfer of information, *tanh* denotes the tanh function, $\odot$ denotes the element-wise vector multiplication, $\boldsymbol{c}_{t-1}$ is the memory from the previous moment, $\tilde{\boldsymbol{c}}_t$ is a candidate state calculated from current input $\boldsymbol{x}_t$ with a nonlinear function as (5):

$$\tilde{\boldsymbol{c}}_t = tanh\,(\boldsymbol{W}_c \boldsymbol{x}_t + \boldsymbol{U}_c \boldsymbol{h}_{t-1} + \boldsymbol{b}_c) \tag{5}$$

where matrix $\boldsymbol{W}_c$, $\boldsymbol{U}_c$, and vector $\boldsymbol{b}_c$ are the parameters of the model.

The gates are designed to be a value between (0, 1) and allow a certain percentage of information to pass through:

- The forgetting gate $\boldsymbol{f}_t$ controls how much information needs to be forgotten about the previous state $\boldsymbol{c}_{t-1}$.
- The input gate $\boldsymbol{i}_t$ controls how much information needs to be stored in the candidate state $\tilde{\boldsymbol{c}}_t$ at the current moment.
- The output gate $\boldsymbol{o}_t$ controls how much information needs to be output to $\boldsymbol{h}_t$ of the internal state $\boldsymbol{c}_t$ at the current moment.

The gates can be calculated as (6)-(8):

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W}_f \boldsymbol{x}_t + \boldsymbol{U}_f \boldsymbol{h}_{t-1} + \boldsymbol{b}_f) \tag{6}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_i \boldsymbol{x}_t + \boldsymbol{U}_i \boldsymbol{h}_{t-1} + \boldsymbol{b}_i) \quad (7)$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_o \boldsymbol{x}_t + \boldsymbol{U}_o \boldsymbol{h}_{t-1} + \boldsymbol{b}_o) \quad (8)$$

where $\boldsymbol{W}_*$, $\boldsymbol{U}_*$, and $\boldsymbol{b}_*$ are also the network parameters that need to be learned from training, σ denotes the sigmoid activation function.

2) LSTM-based forecasting

LSTM layers can be used with other layers in a deep learning model. The architecture of the neural network used to train the load forecaster is a hybrid CNN-LSTM model proposed by Alhussein et al. [40], which uses a convolutional neural network (CNN) to extract the features from the input data. The model is simplified to use less convolutional and LSTM layers in our work as shown in Fig.6. The conv denotes the convolutional layer used to learn features. The max-pooling layer and Rectified Linear Unit (ReLU) layer are incorporated after the convolution layer. The dropout layers are used to prevent the overfitting issue. The fc denotes a fully connected layer to produce the final output. According to comparative results from Alhussein et al., the increase of lookback length in the CNN-LSTM model does not improve the forecasting performance. We adopt this assumption and set the length of the lookback sequences to be 6-step for all of the training in different time resolution. However, the lookback step is the same but the lookback period is different. The reference time is one and half hours for the 15-minutes resolution model, half an hour for the 5-minutes model, and six minutes for the 1-minutes model.

The exogenous input of the network, besides the load history, is a sequence of vectors. Minutes of the hour (Mi), hour of the day (Hi), day of the week (Di), weekend/holiday or not (Vi), and weather forecast (Wi) are used as the feature indicators to train the neural network. The indicators of daily time (Mi and Hi) are encoded by the one-hot encoder. The reason for encoding these continuous features using one-hot variables is to extend the non-linear capability of the network model.

3) Multi-step forecasting

The energy management controller needs a sequence of load forecasts for the following periods of time to decide on peak shaving. Therefore, a multi-step prediction is required instead of one step. Typically, there are several strategies to perform a multi-step ahead sequence forecasting:

- Recursive strategy

Only one predictor is trained to give a one-step-ahead forecast. The multi-step output sequence is generated by recursive callback the forecasted one-step value as the correct one.

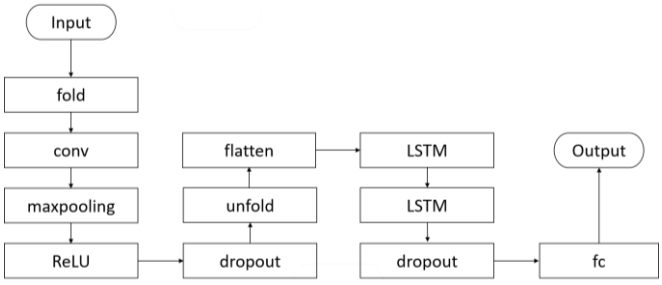

Figure 6. Structure of the neural network for load forecasting.

- Direct strategy

Direct strategy trains a sequence of predictors from identical input to predict the outputs for each time step in the future but each predictor still outputs a scalar value.

- MIMO strategy

The output is a sequence predicted by a single predictor. The difference from the previous cases is that the outputs of the model are vectors for the MIMO strategy.

- DIRMO strategy

DIRMO is the compromise strategy between the Direct and the MIMO strategy. The multi-step forecast problem is divided into smaller lengths.

Among the different strategies, the MIMO and recursive strategy are computationally cheaper solutions [41]. From the point of effect, Gasparin et al. compare the different strategies for the time series forecasting in electric load cases [42]. They found that the MIMO outperforms the recursive strategy when only load values are used in the training. If exogenous features are introduced into the augmented input vector, the MIMO strategy gains negligible improvement. The recursive strategy gets the better performance in this case.

4) Evaluation metrics

The performance of the forecasted load can be quantitatively assessed in different error metrics.

- Mean Absolute Percentage Error (MAPE)
- The Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)

MAPE is the sum of each absolute error divided by the actual value. It is one of the most commonly used metrics for assessing prediction accuracy. However, the problem with MAPE in our peak shaving consideration is that the MAPE divides each error value by the actual value. If the actual load at a given moment is very low and the error is very large, this will have a significant impact on the MAPE value. However, for peak shaving, we are more concerned about the error when the actual load is high. Therefore, MAPE is not an appropriate metric for our purpose.

MAE is an indicator of the average of the absolute errors and RMSE is the square root of the mean of the squared error defined as (9) and (10). From the definition, the errors in RMSE are squared before average, which makes it magnify the difference of the larger errors. For peak shaving use cases, MAE can be used to compare the average error. When the MAE is at the same level, RMSE can be used to distinguish the better model by applying the multiplicative penalty to the outliers.

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right| \quad (9)$$

$$\text{RMSE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{t}=1}^{\mathrm{n}}\left(\mathrm{y_t} - \hat{\mathrm{y}}_\mathrm{t}\right)^2} \quad (10)$$

IV. Simulation of Peak Shaving

A. Experiment Setup

The system of the factory with BESS is simulated for a week. A small amount of PV generation is integrated so that the BESS will conduct the peak shaving based on the residual load. However, the capacity of PV is limited compared to the load. There will be no PV surplus and feed-in issue in this work.

Fig. 7 shows a schematic of the energy system under investigation. The energy management controller sets the threshold of the peak shaving based on the forecast of load and PV generation and then determines the discharging power of the BESS according to the actual load.

B. Integration of PV generation

The power of PV generation can be empirically calculated as (11) [43]:

$$P_{PV} = P_{max} \cdot G_T/G_{STC} \cdot (1 + \gamma(T_c - T_{STC})) \cdot \eta \qquad (11)$$

where P_{max} is the nominal power of the PV module under standard test conditions (STC), G_T is the solar irradiance on the PV panel surface. G_{STC} and T_{STC} are the standard conditions of solar radiation and module temperature of the PV test. γ is the temperature coefficient of the PV module power, η is the overall efficiency. The actual cell temperature T_c can be calculated as (12) [43]:

$$T_c = T_{amb} + (\mathrm{NOCT} - T_{\mathrm{NOCT}}) \cdot G_T / G_{\mathrm{NOCT}} \qquad (12)$$

where T_{amb} is the ambient temperature, NOCT is the Nominal Operating Cell Temperature, which refers to a more realistic condition than STC. G_{NOCT} and T_{NOCT} are the reference irradiance and air temperature of NOCT measurement.

In the experiment, we consider a nominal power of 5kW for the integrated PV module. The PV power forecasting can be calculated from the radiation according to the weather forecast.

C. Energy management of BESS

The residual load is the difference between the load and the PV generation as (13). When the residual load has exceeded a threshold (P_{Thr}), the BESS starts to supply the power to the demand.

$$P_{res}(t) = P_{load}(t) - P_{PV}(t) \qquad (13)$$

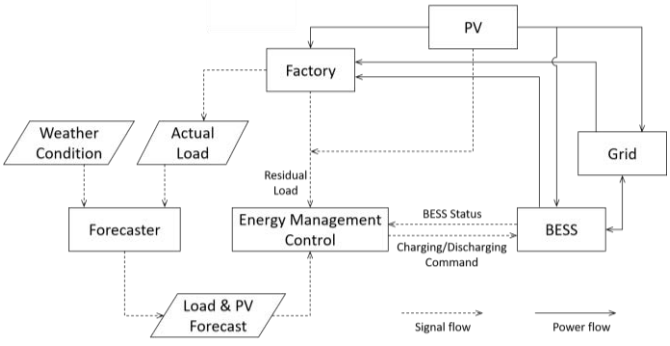

Figure 8. System configuration of the simulation.

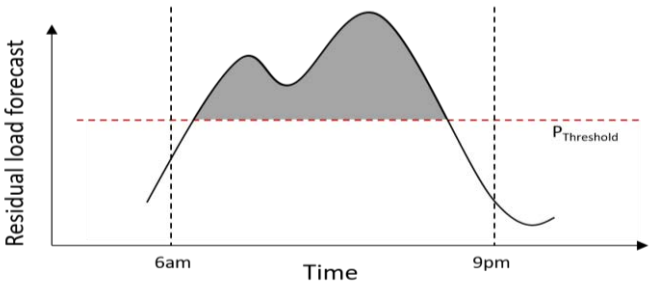

Figure 7. Description of threshold power and peak shaving window

1) Fixed-mode without load forecast

In the simulation, when no load forecasting is available, the threshold is set to be a fixed level such that the BESS is always sufficient during peak shaving. According to the load profile of the factory building, the window of the peak shaving is set to be from 6 am to 9 pm as shown in Fig.8. The battery is charged during the night and discharged when the residual load is above the threshold.

2) Tracking-mode with forecast

With the load forecast, the threshold can be updated during peak shaving according to the energy available in the BESS and potentially required energy. Fig. 9 shows the flowchart to update the threshold for peak shaving in our simulation. During the peak hour, the energy management controller first estimates the remaining energy in BESS and then chooses the appropriate threshold. The discharging power P_{BESS} is limited by the maximum power P_{max} of the BESS, which is dependent on the BESS capacity.

V. Results and Discussion

A. Load Forecasting Performance

The load data from the factory and weather data from DWD in 112 days (16 weeks from May to August) are used to train the ARIMA model and the CNN-LSTM network. The data is partitioned into training, validation, and test sets with 70%, 20%, and 10%. Since peak shaving operates on working days from Monday until Friday, the comparison of results focuses on the forecasting of the load on working days.

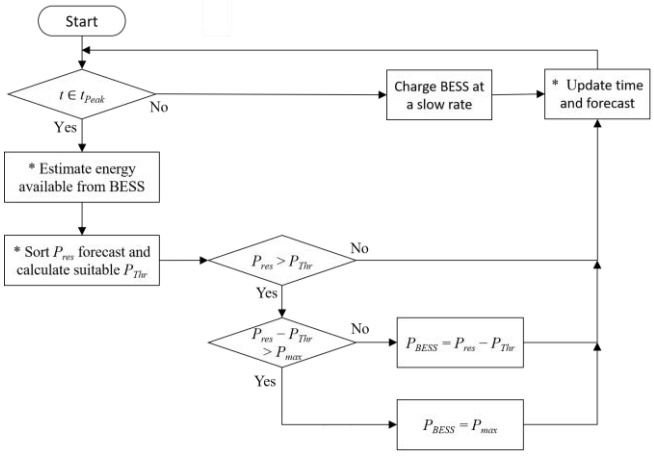

Figure 9. Flowchart of peak shaving algorithm.

a) Time series or neural network

First, we compare the load forecasting with the ARIMA model and the CNN-LSTM model based on the data with 15-minutes resolution. Since the window of peak shaving is 15 hours from 6 am to 9 pm, we compare the forecasting in 4 scenarios: 15-hours-ahead, 5-hours-ahead, 1-hour-ahead, and 1-step-ahead. E.g. in the 15-hours-ahead scenario, each forecasted value is predicted based on the load profile until 15-hours ago to the corresponding time.

Fig. 10 shows the forecasting results for a working day with the ARIMA model and the neural network model. Note that the ARIMA model has a high sensitivity to the variation of the peak load and better tracking performance to the trend of recent load. However, this good tracking performance presents a significant lag effect in forecasting beyond 5 hours. On contrary, the neural network model is more stable and has a relatively consistent forecast regardless of the forecast time in advance. The 5-hours ahead forecast with the CNN-LSTM model almost overlaps the 15-hours ahead forecast.

Fig.11 is a better illustration of the difference. Each solid line represents a forecast from the current time until the end of the peak shaving operation time (9 pm). When the load rises from 10 kW to more than 30 kW, the forecast from the ARIMA model responds rapidly to track the recent load. Although the ARIMA forecast provides a relatively reliable prediction for the overall trend after a few hours, however, the recent forecasts diffuse towards the current load. In contrast, the forecasts by the CNN-LSTM model almost converge to the main curve stream. But the performance of tracking a specific high peak load is lacking.

Table II shows the comparison of the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) of the forecast with the ARIMA and CNN-LSTM model. It turned out that the MAE and RMSE of the ARIMA forecast are at the same level, if not better, as the CNN-LSTM. The first reason is that the ARIMA model tracks the load much better during the valley hour when the load is low and stable,

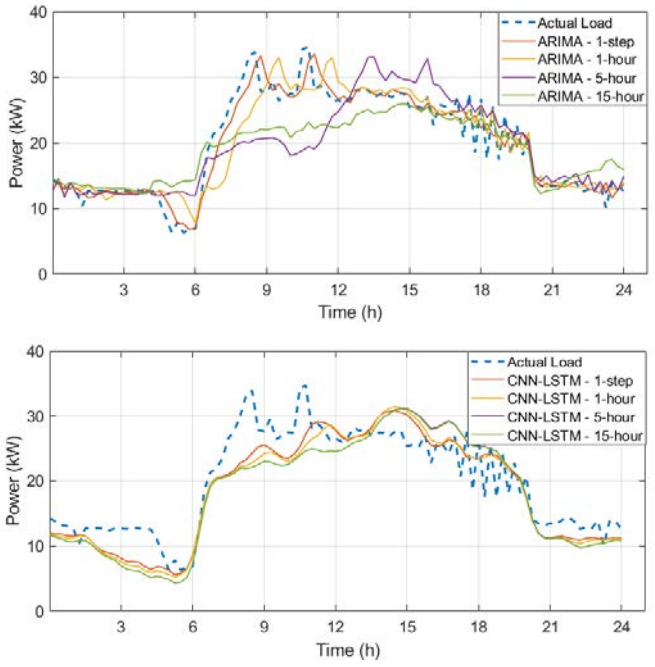

Figure 11. Comparison of load forecasting with ARIMA model and CNN-LSTM model with 15-minutes resolution load.

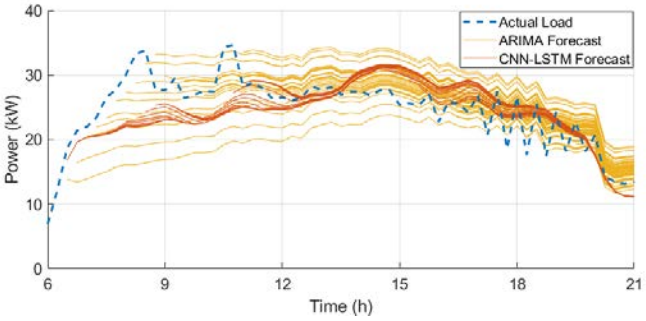

Figure 10. Comparison of load forecasting with the ARIMA model and the CNN-LSTM model in a peak shaving window.

which reduces the MAE and RMSE measured within the whole timespan. The second reason is that the ARIMA model does well in tracking the recent high peak as illustrated in Fig.10. Nevertheless, this does not mean that the ARIMA model outperforms the CNN-LSTM model in load forecasting for peak shaving. The ARIMA model has very different forecasts for 1-step ahead and 5-hours ahead scenarios. And the CNN-LSTM model provides a coherent and plausible description of the future load during the peak hour. Both models are further used for simulation.

There is another thing we notice from the table that the increase of the time resolution of data does not improve the performance of load forecasting. It might be just the opposite, the instability contains in the data with high time-resolution may reduce the effectiveness of the pattern learned from the training. This happens in both ARIMA and CNN-LSTM models in the forecast for the load with 5-minutes time resolution.

TABLE II. THE FORECAST COMPARISON OF THE MODELS

Method	Data Resolution	Forecast Scenario	MAE of the Forecast	RMSE of the Forecast
ARIMA	15-minutes	1-Step	1.613	2.299
		1-hour	2.116	3.226
		5-hours	3.489	4.934
		15-hours	3.218	4.279
	5-minutes	1-Step	1.825	2.049
		1-hour	2.851	3.991
		5-hours	4.348	5.860
		15-hours	4.332	5.271
CNN-LSTM	15-minutes	1-Step	2.801	3.489
		1-hour	3.107	3.854
		5-hours	3.625	4.362
		15-hours	3.613	4.350
	5-minutes	1-Step	4.128	5.571
		1-hour	5.253	7.158
		5-hours	5.275	7.170
		15-hours	5.275	7.170
	1-minutes	1-Step	1.624	2.336
		1-hour	3.757	5.013
		5-hours	3.757	5.013
		15-hours	3.757	5.013

b) Multi-step forecasting strategy

We verify the assumption that the recursive strategy provides better results in multi-step forecasting than the MIMO strategy. Table III gives the MAE and RMSE of the two strategies. The recursive strategy is slightly better considering the two metrics together. Fig.12 shows a forecast conducted at 6 am for the load from 6 am until 9 pm. The recursive forecast is also slightly more accurate than the MIMO forecast. Besides, the recursive forecast is more flexible to implement in the simulation.

c) Reference time step

The assumption about the length of the reference time sequence is also verified in our results. Table IV summarizes the MAE and RMSE of the forecasting results with different look-back steps. From the table, we can conclude that adding the length of the look-back sequence does not have a significant impact on the performance of prediction. However, the increase of the reference time steps does have a negative effect and reduces the accuracy of the forecasting.

B. Peak Shaving Performance

We conducted the peak shaving simulation for one week with the 15-minutes time resolution data in 3 cases to compare the load forecasting models for peak shaving use case: a) without forecast, b) forecast with the ARIMA model, c) forecast with the CNN-LSTM model. The setup of the simulation is described in section IV.

The BESS capacity is considered as the variables with values in the range of 5 to 200 kWh. The performance of peak shaving is evaluated based on the reduction of peak power and peak-hour energy. Fig.13 shows the comparison of the simulation results with 50 kWh capacity BESS. It can be noticed from the SOC curve that the ARIMA model always runs out of the capacity of BESS sooner than the CNN-LSTM model. This can be explained by the forecasting results presented in Fig.11. The ARIMA model tends to set a low threshold at the beginning due to underestimation of the load.

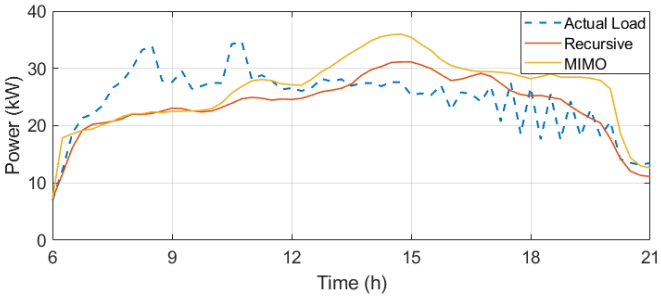

Figure 12. Comparison of multi-step strategies in a peak shaving window.

TABLE III. COMPARISON OF THE MULTI-STEP STRATEGIES

Method	Data Resolution	Forecast Scenario	Multi-step Strategy	MAE	RMSE
CNN-LSTM	15-minutes	15-hour	Recursive	3.744	4.631
			MIMO	3.177	6.059

TABLE IV. COMPARISON OF THE REFERENCE SEQUENCE LENGTH

Method	Reference Steps	Data Resolution	Forecast Scenario	MAE	RMSE
CNN-LSTM	6-steps	15-minutes	1-step	2.801	3.489
			1-hour	3.107	3.854
			5-hours	3.625	4.362
			15-hours	3.613	4.350
	12-steps	15-minutes	1-step	2.972	3.527
			1-hour	3.288	3.951
			5-hours	3.543	4.439
			15-hours	3.589	4.409
	24-steps	15-minutes	1-step	3.265	4.067
			1-hour	3.478	4.318
			5-hours	3.898	4.870
			15-hours	3.977	4.631

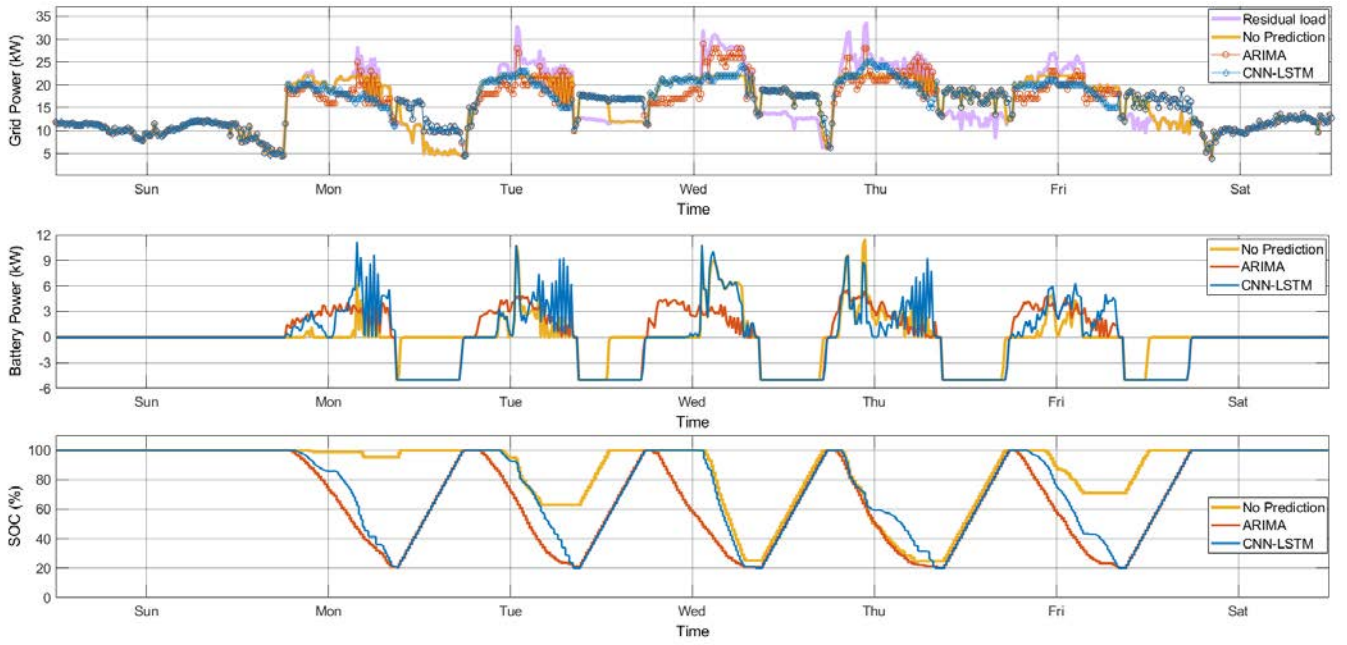

Figure 13. Comparison of load forecasting models in BESS peak shaving simulation with 50 kWh capacity.

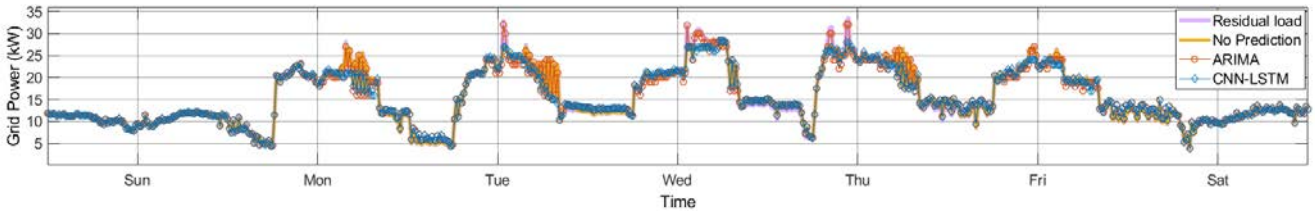

Figure 14. Grid power curve of the BESS peak shaving simulation with 10 kWh capacity.

Afterward, the ARIMA model moves to a high threshold as the result of the decrease in BESS remaining energy and the overestimation of the future load. Therefore, the peak power in the shaved curve of the ARIMA model is 29kW, which is higher than the 25kW with the CNN-LSTM model. Since both forecasting models deplete the available capacity of BESS in each cycle, we consider them as a tie in the reduction of peak-hour energy.

By contrast, peak shaving without prediction uses a fixed threshold. If the threshold is cleverly selected, it is possible to reduce the peak power of the grid load to 22 kW in the case of 50 kWh BESS. However, this is from the perspective of looking back in hindsight. It is hard to determine the most suitable threshold in advance. Even if we assume that the appropriate threshold can be picked beforehand, the peak shaving without load forecast is deficient in the respect of peak energy reduction. Alternatively, the reduction of peak power and peak energy can not be optimally achieved at the same time without a load forecast. The 50 kWh BESS discharges 118.4 kWh electricity in the weekly peak shaving simulation without forecast compared to 194.7 kWh (inverter loss included) with the forecast. This means that the load forecasting significantly improves the utilization of BESS by more than 60% in comparison with a fixed threshold.

The performances are similar for other BESS capacities. Fig.14 shows the grid power with a smaller BESS of 10 kWh. The reduced peak power of the no forecast model with carefully selected threshold, the ARIMA model, and the CNN-LSTM model are 28 kW, 32 kW, and 29 kW respectively. But the peak energy reduction without prediction is even less than half of the energy with the forecasting models. Fig. 15 provides an illustrative description of the peak power and the shifted peak energy as functions of the capacity of the BESS. The solid lines represent the maximum of the peak power in the simulation for each BESS capacity. The peak power falls first and then rises as the BESS capacity grows. The reason for the rebound

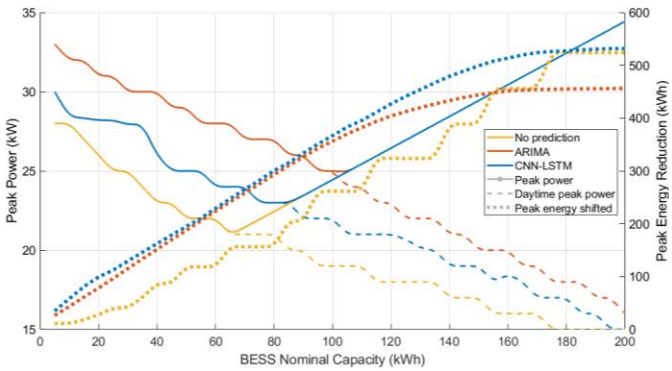

Figure 15. Peak power and peak energy reduction of the peak shaving with different BESS capacity

of peak power is that the charging power at night dominates the peak power when the BESS becomes larger. The dashed lines represent the maximum of the peak power during the daytime. And the dotted lines show the sum of the discharged electricity during peak shaving. From the perspective of peak power reduction, an optimally selected fixed threshold is possible to give the lowest peak power, followed by the CNN-LSTM model. But the forecast can dramatically improve energy shifting until the BESS capacity reaches a limit where further energy shifting requires much more BESS capacity. We note that from both the power and energy aspects, the neural network model always outperforms the time-series model with various BESS capacities.

VI. Conclusion

In this paper, we presented an investigation of the load forecasting models for industrial peak shaving with a battery energy storage system. An ARIMA model based on time series analysis and a deep learning model based on a hybrid CNN-LSTM network structure is developed. The CNN-LSTM network is trained with the historical load profile and exogenous time and weather information. The aim of implementing load forecasting in peak shaving is to provide a reference threshold during the peak hour. The load forecasting results generated with different models (the ARIMA model and the CNN-LSTM model), different time resolutions (15-minutes, 5-minutes, and 1-minutes), different multi-step forecasting strategies (recursive and MIMO), and different length of the reference sequence (6-steps, 12-steps, and 24-steps) are compared. Data with higher time resolution does not provide a convincible forecast compared to the 15-minutes profile. The neural network model with a recursive strategy based on the last 6 steps is chosen for the comparison simulation with the time series model. The results indicate that the load forecast plays an important role in improving the utilization of BESS capacity in peak energy reduction. The simulation results demonstrate that the deep learning model will provide a more reasonable reference threshold in peak shaving than the time series model throughout the process due to its robust prediction of the overall trend. Future research should investigate the optimization of the dispatching strategies with surplus renewable energy available. Another thing that can be considered in BESS peak shaving is flexible pricing schemes, where the energy shifting can be managed to deliver further values.

Acknowledgment

The authors are grateful to the Karlsruhe Institute of Technology for funding the presented work and the climate dataset from the German Weather Service.

References

[1] J. Sa'ed, S. Favuzza, F. Massaro, and E. Telaretti, "Optimization of BESS Capacity Under a Peak Load Shaving Strategy," in *2018 IEEE*

International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I CPS Europe), Jun. 2018, pp. 1–4, doi: 10.1109/EEEIC.2018.8494008.

[2] M. Zheng, C. J. Meinrenken, and K. S. Lackner, "Smart households: Dispatch strategies and economic analysis of distributed energy storage for residential peak shaving," *Applied Energy*, vol. 147, pp. 246–257, Jun. 2015, doi: 10.1016/j.apenergy.2015.02.039.

[3] K. H. Chua, Y. S. Lim, and S. Morris, "Energy storage system for peak shaving," *International Journal of Energy Sector Management*, vol. 10, no. 1, pp. 3–18, Jan. 2016, doi: 10.1108/IJESM-01-2015-0003.

[4] V. Papadopoulos, T. Delerue, J. V. Ryckeghem, and J. Desmet, "Assessing the impact of load forecasting accuracy on battery dispatching strategies with respect to Peak Shaving and Time-of-Use (TOU) applications for industrial consumers," in *2017 52nd International Universities Power Engineering Conference (UPEC)*, Aug. 2017, pp. 1–5, doi: 10.1109/UPEC.2017.8231939.

[5] "Peak Forecasting for Battery-based Energy Optimizations in Campus Microgrids | Proceedings of the Eleventh ACM International Conference on Future Energy Systems." https://dl.acm.org/doi/abs/10.1145/3396851.3397751 (accessed Apr. 13, 2021).

[6] S. Ryu, J. Noh, and H. Kim, "Deep Neural Network Based Demand Side Short Term Load Forecasting," *Energies*, vol. 10, no. 1, Art. no. 1, Jan. 2017, doi: 10.3390/en10010003.

[7] "Short-Term Non-Residential Load Forecasting Based on Multiple Sequences LSTM Recurrent Neural Network | IEEE Journals & Magazine | IEEE Xplore." https://ieeexplore.ieee.org/abstract/document/8481348 (accessed Apr. 14, 2021).

[8] J. Walther, D. Spanier, N. Panten, and E. Abele, "Very short-term load forecasting on factory level – A machine learning approach," *Procedia CIRP*, vol. 80, pp. 705–710, Jan. 2019, doi: 10.1016/j.procir.2019.01.060.

[9] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches †," *Energies*, vol. 11, no. 7, Art. no. 7, Jul. 2018, doi: 10.3390/en11071636.

[10] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Multi-Sequence LSTM-RNN Deep Learning and Metaheuristics for Electric Load Forecasting," *Energies*, vol. 13, no. 2, Art. no. 2, Jan. 2020, doi: 10.3390/en13020391.

[11] S. Bischof, H. Trittenbach, M. Vollmer, D. Werle, T. Blank, and K. Böhm, "HIPE: An Energy-Status-Data Set from Industrial Production," in *Proceedings of the Ninth International Conference on Future Energy Systems - e-Energy '18*, Karlsruhe, Germany, 2018, pp. 599–603, doi: 10.1145/3208903.3210278.

[12] Xinxing Pan and B. Lee, "A comparison of support vector machines and artificial neural networks for mid-term load forecasting," in *2012 IEEE International Conference on Industrial Technology*, Mar. 2012, pp. 95–101, doi: 10.1109/ICIT.2012.6209920.

[13] K. Liu *et al.*, "Comparison of very short-term load forecasting techniques," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 877–882, May 1996, doi: 10.1109/59.496169.

[14] S. R. Huang, "Short-term load forecasting using threshold autoregressive models," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 144, no. 5, pp. 477–481, Sep. 1997, doi: 10.1049/ip-gtd:19971144.

[15] J.-F. Chen, W.-M. Wang, and C.-M. Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting," *Electric Power Systems Research*, vol. 34, no. 3, pp. 187–196, Sep. 1995, doi: 10.1016/0378-7796(95)00977-1.

[16] M. Y. Cho, J. C. Hwang, and C. S. Chen, "Customer short term load forecasting by using ARIMA transfer function model," in *Proceedings 1995 International Conference on Energy Management and Power Delivery EMPD '95*, Nov. 1995, vol. 1, pp. 317–322 vol.1, doi: 10.1109/EMPD.1995.500746.

[17] G. Juberias, R. Yunta, J. G. Moreno, and C. Mendivil, "A new ARIMA model for hourly load forecasting," in *1999 IEEE Transmission and Distribution Conference (Cat. No. 99CH36333)*, Apr. 1999, vol. 1, pp. 314–319 vol.1, doi: 10.1109/TDC.1999.755371.

[18] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, Jan. 2003, doi: 10.1016/S0925-2312(01)00702-0.

[19] Jian-Chang Lu, Dong-Xiao Niu, and Zheng-Yuan Jia, "A study of short-term load forecasting based on ARIMA-ANN," in *Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826)*, Aug. 2004, vol. 5, pp. 3183–3187 vol.5, doi: 10.1109/ICMLC.2004.1378583.

[20] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings," *JAPAN ARCHITECTURAL REVIEW*, vol. 3, no. 1, pp. 62–76, 2020, doi: https://doi.org/10.1002/2475-8876.12135.

[21] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 442–449, May 1991, doi: 10.1109/59.76685.

[22] L. Hernandez, C. Baladrón, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, and J. Lloret, "Short-Term Load Forecasting for Microgrids Based on Artificial Neural Networks," *Energies*, vol. 6, no. 3, Art. no. 3, Mar. 2013, doi: 10.3390/en6031385.

[23] 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, Feb. 2001, doi: 10.1109/59.910780.

[24] C. M. Pereira, N. N. de Almeida, and M. L. F. Velloso, "Fuzzy Modeling to Forecast an Electric Load Time Series," *Procedia Computer Science*, vol. 55, pp. 395–404, Jan. 2015, doi: 10.1016/j.procs.2015.07.089.

[25] F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah, "Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique," *Energy and Buildings*, vol. 126, pp. 94–103, Aug. 2016, doi: 10.1016/j.enbuild.2016.05.028.

[26] R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy," *Applied Energy*, vol. 123, pp. 168–178, Jun. 2014, doi: 10.1016/j.apenergy.2014.02.057.

[27] W. Wang, T. Hong, X. Xu, J. Chen, Z. Liu, and N. Xu, "Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm," *Applied Energy*, vol. 248, pp. 217–230, Aug. 2019, doi: 10.1016/j.apenergy.2019.04.085.

[28] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: 10.1109/TSG.2017.2753802.

[29] M. Imani and H. Ghassemian, "Residential load forecasting using wavelet and collaborative representation transforms," *Applied Energy*, vol. 253, p. 113505, Nov. 2019, doi: 10.1016/j.apenergy.2019.113505.

[30] 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)*, Oct. 2018, pp. 1–6, doi: 10.1109/SmartGridComm.2018.8587554.

[31] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using Deep Neural Networks," in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, Oct. 2016, pp. 7046–7051, doi: 10.1109/IECON.2016.7793413.

[32] H. Zheng, J. Yuan, and L. Chen, "Short-Term Load Forecasting Using EMD-LSTM Neural Networks with a Xgboost Algorithm for Feature Importance Evaluation," *Energies*, vol. 10, no. 8, Art. no. 8, Aug. 2017, doi: 10.3390/en10081168.

[33] L. Barelli, G. Bidini, F. Bonucci, and A. Ottaviano, "Residential micro-grid load management through artificial neural networks," *Journal of Energy Storage*, vol. 17, pp. 287–298, Jun. 2018, doi: 10.1016/j.est.2018.03.011.

[34] "Index of /climate_environment/CDC/." https://opendata.dwd.de/climate_environment/CDC/ (accessed Apr. 19, 2021).

[35] "Smart Meter Roll-out: The German Case," *bne - Bundesverband Neue Energiewirtschaft e.V.* https://www.bne-online.de/en/news/article/smart-meter-roll-out-the-german-case/ (accessed Apr. 14, 2021).

[36] L. Zhuang, H. Liu, J. Zhu, S. Wang, and Y. Song, "Comparison of forecasting methods for power system short-term load forecasting based on neural networks," in *2016 IEEE International Conference on Information and Automation (ICIA)*, Aug. 2016, pp. 114–119, doi: 10.1109/ICInfA.2016.7831806.

[37] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, Mar. 1994, doi: 10.1109/72.279181.

[38] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[39] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," pp. 850–855, Jan. 1999, doi: 10.1049/cp:19991218.

[40] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting," *IEEE Access*, vol. 8, pp. 180544–180557, 2020, doi: 10.1109/ACCESS.2020.3028281.

[41] S. B. Taieb, G. Bontempi, A. Atiya, and A. Sorjamaa, "A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition," *arXiv:1108.3259 [cs, stat]*, Aug. 2011, Accessed: Apr. 23, 2021. [Online]. Available: http://arxiv.org/abs/1108.3259.

[42] A. Gasparin, S. Lukovic, and C. Alippi, "Deep Learning for Time Series Forecasting: The Electric Load Case," *arXiv:1907.09207 [cs, stat]*, Jul. 2019, Accessed: Apr. 07, 2021. [Online]. Available: http://arxiv.org/abs/1907.09207.

[43] A. Sarhan, H. Hizam, N. Mariun, and M. E. Ya'acob, "An improved numerical optimization algorithm for sizing and configuration of standalone photo-voltaic system components in Yemen," *Renewable Energy*, vol. 134, pp. 1434–1446, Apr. 2019, doi: 10.1016/j.renene.2018.09.069.