Auto-encoder Neural Network-Based Monthly Electricity Consumption Forecasting Method Using Hourly Data

Zhenghui Li, Kangping Li, Fei Wang, Zengqiang Mi
Department of Electrical Engineering
North China Electric Power University
Baoding 071003, China
feiwang@ncepu.edu.cn

Wanwei Li
Development and Planning Department
State Grid Gansu Electric Power Company
Lanzhou 730046, China
liwanw@gs.sgcc.com.cn

Payman Dehghanian
Department of Electrical and Computer Engineering
George Washington University
Washington DC, USA
payman@gwu.edu

***Abstract*—The effectiveness of monthly electricity consumption forecasting (ECF) directly affects the profitability of electricity retailers in a deregulated market. The monthly ECF using fine-grained hourly data based on the multi-step forecasting strategy normally shows unsatisfactory performance, given the fact that it contains numerous forecasting steps. Aggregating the data points is a common approach which can reduce the forecasting steps by compressing the data series. However, the information loss caused by the additive aggregation method generally leads to low predictability of the compressed data series. To address this challenge, we propose an auto-encoder neural network (AENN) based data compression method. Specifically, an AENN with a small central layer is first trained to reconstruct the fine-grained hourly electricity consumption input data. Subsequently, the former part of the trained AENN is used to compress the hourly data into the coding series. Then, the multi-step forecasting model is trained based on the coding series. Finally, the forecast result of the coding is decoded using the latter part of the trained AENN to form the electricity consumption forecast. Numerical experiments demonstrate the superiority of the proposed method while combined with three representative AI forecasting algorithms.**

Keywords—Monthly electricity consumption forecasting, multi-step forecast, data compression, auto-encoder neural network, electricity market.

I. Introduction

In deregulated electricity markets, numerous electricity retailers participate in electricity trading directly [1]-[3]. They make a profit from purchasing electricity in wholesale markets and then selling it to customers. Typically, the customer's electricity needs of an electricity retailer are satisfied by sign mid- and long-term electricity purchasing contracts in futures markets, while the real-time electricity deviations are balanced through spot trading [4]. In the competitive electricity markets, however, the huge fluctuations in the spot price of electricity put many of them at the risk of financial loss. Electricity retailers are eager to accurately grasp the customers' mid- or long-term electricity consumption so that they can make reasonable negotiations or purchase decisions in the wholesale markets to avoid the risks in the spot market. As one of the main mid-term electricity trading modes, the accurate forecasting of monthly electricity consumption can provide a basic strategic guarantee for electricity retailers to negotiate contracts. Consequently, the necessity of developing an accurate monthly electricity consumption forecasting(ECF) method becomes imperative for these retailers.

There are dozens of different power forecasting methods [5]-[8] that have been used and documented during the past few decades. In the field of electricity consumption forecasting, most of which focus on short-term ECF, while few of them investigate into monthly ECF [9]. Different from short-term ECF, monthly ECF is characterized by the long forecasting time scale. [10] reviewed the most relevant studies to define and classify the various electricity consumption forecasting techniques developed for mid- and long-term horizons. The most widely implemented techniques generally contain two broad categories: parametric methods and artificial intelligence (AI) methods. Parametric methods mainly include some mathematical or statistical models such as autoregressive integrated moving average (ARIMA) [11] and grey dynamic models [12]. AI methods mostly include artificial neural network (ANN) [13], [14], support vector regression (SVR) [15]. Recent studies have shown that the statistical methods are lack of accuracy and not flexible [16], while the AI methods have established themselves as serious contenders because they are more suitable to complicated nonlinear electricity consumption forecasting problem. To avoid trapping into a local optimum of the general SVR, the chaotic genetic algorithm is used in [17] to optimize the parameter. The input and output of the model are monthly data of electricity consumption, and the forecast results show the high forecasting accuracy of the method. [18] also uses monthly data to build an MLP neural network model to forecast the monthly electricity consumption. To avoid introducing exogenous variables such as weather or economic data, [19] splits the monthly electricity consumption series into trend and fluctuation series and uses two neural networks to forecast them separately. The overall forecast results are obtained by adding up these two forecast results and shows better performance.

It is worth noting that all of these existing studies and their variants mainly focus on a one-step-ahead forecast problem. The historical monthly electricity consumption data is used to forecast the electricity demand in the next month, and the training data used are generally coarse-grained monthly data. Limited monthly data are available (only 12 data points a year), which means that only a few training samples can be built to train the model. The lake of sufficient training samples is one of the main reasons for the overfitting of AI models, and it directly leads to the low accuracy of forecast results.

With the widespread popularity of smart meters, more and more fine-grained data, e.g., hourly data, can be measured and collected [20]-[22]. It is generally used in the training of short-term forecasting in the past studies since it contains large number of detailed historical data that can be used to prevent overfitting. Nevertheless, it is rarely used in monthly ECF. If hourly data is used to build a monthly ECF model, the multi-step forecasting strategy needs to be utilized to forecast all hours of electricity consumption in the next month. Then, all forecast results should be added up to obtain the overall forecasting. Compared with the one-step forecasting model, the multi-step forecasting model shows poor accuracy since the long forecasting horizon will lead to the uncertainty of forecast results.

One of the solutions to solve the above problem is to develop a forecasting frame to reduce the step size of multi-step forecasting, which has been investigated in our previous work [23]. On this basis, we consider compressing the data appropriately to further reduce the forecasting steps while ensuring the sufficient forecasting sample. A simple method is to add several data together. However, the information loss caused by the additive data compression method leads to low predictability of the compressed series. In this paper, auto-encoder neural network (AENN), a robust unsupervised learning algorithm widely adopted in many other fields [24]-[26], is used to provide an automated way of data compression. It encodes the input into low-dimension coding and decodes it to reconstruct the input so that the coding can represent input data accurately with lower information loss. What's more, the coding can be forecasted and then decoded to get the forecast results of the original data. Justification of why AENN is suitable for this task is given. In the case study, we use the Sample Entropy theory to demonstrate the higher predictability of the coding than the data compressed by the additive data compression method. As stated above, the key contributions can be summarized as twofold:

1) A novel approach using AENN to compress the electricity consumption series to reduce the forecasting step. The time series compressed by AENN shows better predictability than that using the additive aggregation method, thus can provide a higher forecasting accuracy.

2) The Sample Entropy theory is used to demonstrate the high predictability of the coding, and detailed simulation experiments also prove the superiority of the proposed method.

The rest of this paper is organized as follows. In Section II, we present the overall forecasting framework. Then, Empirical tests and results based on real data are presented in Section III, and the conclusion is made in Section IV.

II. Methodologies

In this section, we first introduce the general framework of the proposed method, then the two steps of the framework are illustrated in detail, respectively.

A. *Framework of the Method*

In the conventional monthly ECF method, the historical data used is generally monthly data. It is a one-step forecasting problem consists of forecasting the next $D^M_{N_1+1}$ of a historical monthly electricity consumption time series $[D^M_1,...,D^M_{N_1}]$, where D^M is the electricity consumption of a month. The limitation on the number of training samples that can be constructed from N_1 observations to train a forecasting model generally leads to the overfitting of the model. Therefore, it cannot provide an accurate forecast result.

To overcome the above issue, we tend to to use fine-grained hourly data to construct more training samples. The total electricity consumption for a month can be expressed as $D^M = \sum_{t=1}^{T} D^H_t$, where D^H is the electricity consumption of an hour and *T* is the number of hours in a month. Therefore, the monthly ECF turns into a multi-step forecasting problem. We need to forecast the next *T* values $[D^H_{N_2+1},...,D^H_{N_2+T}]$ of a hourly electricity consumption time series $[D^H_1,...,D^H_{N_2}]$. It is obvious that the number of observations N_2 is much larger than N_1 so that more training samples with abundant information can be used as a support to the forecasting model. Nevertheless, the *T* forecasting steps (about 720) are too large to ensure the effectiveness of multi-step forecasting models. In order to overcome the above problem, this paper aims to develop a method that can reduce the forecasting steps of the hourly data based multi-step forecasting model.

The proposed method includes two major steps is illuminated in Fig. 1. The first step investigated in our previous work aims to decompose the forecasting steps of hourly data based monthly ECF preliminarily. The historical electricity consumption hourly data $[D^H_1,...,D^H_{N_2}]$ is decomposed into *Series1* to *Series7* based on the week label (Monday to Sunday), denoted as $[D^{W_i}_1,...,D^{W_i}_{w_i}], i=1,...,7, \sum_{i=1}^{7} w_i = N_2$, where W_i represents week " i ", w_i represents the total number of hours of week " i ". As a result, the *T* forecasting steps are divided into seven parts. The details of first step are described in Section II-B.

The next step proposed in this paper aims to further reduce the forecasting steps of each decomposed series. Adding several data points together to one seems to be a common approach to compress the forecasting steps, whereas it may cause large information loss so that lead to poor predictability of the compressed data.

In this paper, we develop a novel data compression method using AENN to reduce the forecasting steps. Herein, taking *Series1* $[D^{W_1}_1,...,D^{W_1}_{w_1}]$ as an example, we first train an auto-encoder model based on it. Subsequently, the encoder model (the former part of the auto-encoder model) is used to compress *Series1* to a new series, denoted as *AE Coding Series11* $[AE\_D^{W_1}_1,...,AE\_D^{W_1}_{w_1/k}]$, where *k* is the compression scale. The compression process is similar to data dimensionality reduction. Then, we develop a multi-step forecasting model using *AE Coding Series11* based on an iterated strategy. It is a multi-step forecasting strategy, which using the result forecasted at one iteration as the model input in the next iteration [24]. Finally, the forecast results are decoded using the decoder model (the latter part of the auto-encoder model) to obtain the electricity consumption forecast. The details of second step are described in Section II-C.

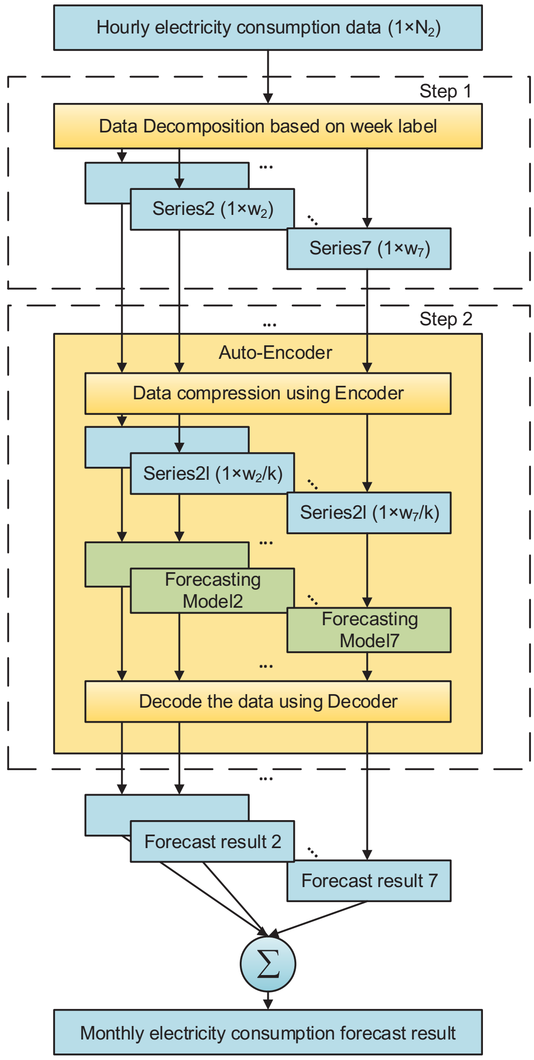

Fig. 1. Framework of proposed method.

B. Data Decomposition-Accumulation Approach

Fig. 2 illustrates principle of the decomposition-accumulation approach. Suppose that under the influence of calendar factors, the electricity consumption behaviors on the same day of each week are similar. As shown in Fig. 2, the hourly electricity consumption data is divided into seven series according its week label (the day of the week), denoted as *Series1* $[D_1^{W_1},...,D_{w_1}^{W_1}]$ to *Series7* $[D_1^{W_7},...,D_{w_7}^{W_7}]$. Therefore, the electricity consumption of a month can be expressed as $D^M = \sum_{i=1}^{7}\sum_{t=1}^{T_i} D_t^{W_i}$, where T_i is the total number of hours of the week " i " in the month. As a result, the T forecasting steps are decomposed into seven parts, denoted as $T_i, i = 1,...,7$. Each T_i is about one seventh of T. Then, seven forecasting models are built based on each series respectively. Each forecasting model is used to forecast the hourly electricity consumption of all Mondays to all Sundays in the next month. Taking January 2018 as an example, it contains five Mondays with a total of 120 hours. the first model is used to forecast the next 120 values $[D_{w_1+1}^{W_1},...,D_{w_1+120}^{W_1}]$ of *Series1* $[D_1^{W_1},...,D_{w_1}^{W_1}]$. The same procedure can be adapted to obtain all Tuesdays to all Sundays forecasts. Finally, the seven forecast results can be accumulated to obtain the monthly electricity consumption forecast. Since the performance of multi-step forecasting models is generally inversely proportional to the forecasting steps, the decomposition-accumulation based forecasting approach can reduce the forecasting steps initially to guarantee the accuracy of multi-step forecasting.

C.

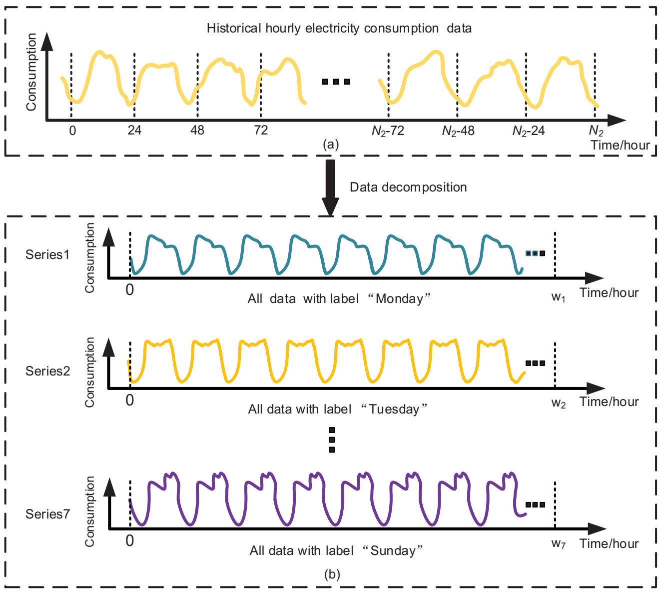

Fig. 2. Schematic diagram of data decomposition method.

D. Data Compression Using Auto-encoder Neural Network

According to Section II-B, the forecasting steps of each model are $T_i, i = 1,...,7$. To further reduce the forecasting steps, the data resolution of *Series1* to *Series7* should be decreased appropriately. Adding several data points to one to compress the series is a common method shown in Fig. 3a. It is worth noting that finding a specific compression scale to ensure optimal predictability of the compressed data tends to be difficult and is not the main concern of this paper, since it is generally related to the selection of the data set. The *Aggregated Series1l* to *Aggregated Series7l* are denoted as $[Agg\_D_1^{W_i},...,Agg\_D_{w_i/k}^{W_i}]$, $i = 1...7$, in this paper, we discuss two data compression scale:

1) $k = 6$: As shown in Fig. 3a, each six data points are aggregated to one, which means that the day will be divided into four periods (0:00 to 6:00, 6:00 to 12:00, 12:00 to 18:00 and 18:00 to 24:00).

2) $k = 8$: Each eight data points are aggregated to one, which means that the day will be divided into three periods (0:00 to

8:00, 8:00 to 16:00 and 16:00 to 24:00). We will discuss this situation in detail in Section III as a comparison.

Nevertheless, this kind of data compression cannot diminish the information loss during the procedure of data aggregation, hence directly leads to a decrease in the predictability of the compressed data.

In this paper, an auto-encoder neural network based data compression method is proposed as shown in Fig. 3b. AENN is a commonly used data compression model based on unsupervised deep learning technique. It consists of a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors which can convert high-dimensional data to low-dimensional codes [27]. The architecture of AENN used in this paper is shown in Fig. 4.

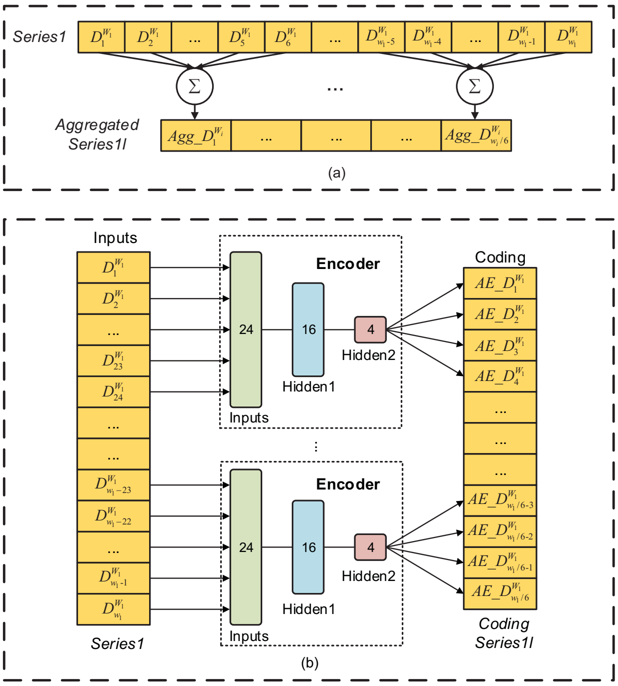

Fig. 3. Schematic diagram of data compression. (k = 6)

It is a five-layer neural network about the symmetry of the middle layer. The number of neurons in input layer, hidden layer1, hidden layer3, and output layer are 24, 16, 16 and 24, respectively. The number of neurons in hidden layer2 is 24/k. Activation function of hidden layer1, hidden layer2, hidden layer3, and output layer are rectified linear unit, linear, rectified linear unit and hyperbolic tangent, respectively. The stopping criterion of the AENN model is 200 maximum epochs. The proposed method is ideal for compressing the data since it has the following three characteristics:

(1) Nonlinear neural network based AENN can better extract nonlinear features in the input data, meanwhile coding can reconstruct the input data through the decoder, thus the coding can represent input data accurately with low information loss.

(2) As can be seen in Fig. 3b. The new coding series (*AE Coding Series1l* to *AE Coding Series7l*) are arranged according to the order of neurons in hidden layer. The regularity of the coding series enables its predictability.

(3) Forecast results of each part can be decoded through the decoder, so that we can get the electricity consumption forecast results rather than coding forecast results.

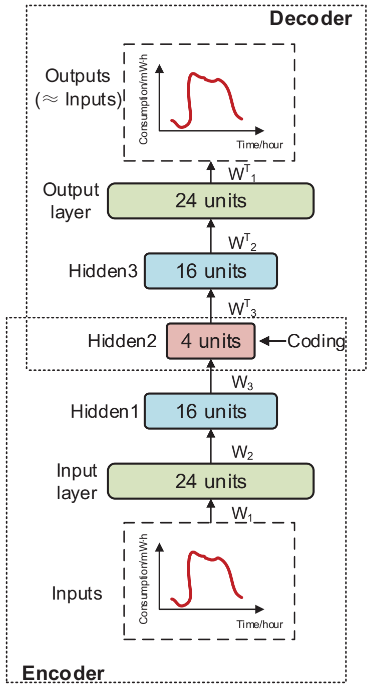

Fig. 4. AENN network used in this paper. ($k = 6$)

It is obvious that the regularity of a time series is proportional to its predictability. Sample Entropy is a measure of regularity and complexity of time series through measuring the probability of generating a new pattern in it [28]. The larger the sample entropy, the lower the predictability of time series. In this paper, sample entropy theory is used to demonstrate the effectiveness of the proposed method. For a fair comparison, the electricity consumption data are compressed at the same scale using aggregated method and AENN. The data with various granularity is recorded in Table I.

TABLE I. REPRESENTATION SYMBOLS FOR DATA WITH DIFFERENT GRANULARITY

Data granularity	Symbol	Data dimension
Monthly data	$[D_1^M,...,D_{N_1}^M]$	N_1
Hourly data	$[D_1^H,...,D_{N_2}^H]$	N_2
Decomposed data based on week label	$[D_1^{Wi},...,D_{w_i}^{Wi}]$	w_i
Compressed data using aggregated method	$[Agg\_D_1^{W_i},...,Agg\_D_{w_i/k}^{W_i}]$	w_i/k
Compressed data using AENN method	$[AE\_D_1^{W_i},...,AE\_D_{w_i/k}^{W_i}]$	w_i/k

III. COMPUTATION EXPERIMENTS & RESULTES

In this section, we first introduce the setup of the computational experiments including the data set, the

Identify applicable funding agency here. If none, delete this text box.

representative forecasting algorithms, the accuracy measures and the computational environment. Then two cases are conducted based on the collected data set.

A. Setup of Computational Experiments

The hourly electricity consumption data from Jan. 1, 2012 to Dec. 31, 2018 collected from PJM [29] are used to evaluate the superiority of the proposed method. In order to find an auto-encoder with better performance, we first picked five years (from Jan. 1, 2012 to Dec. 31, 2016) as the training set to train an auto-encoder model. Then, we use one year (from Jan. 1, 2017 to Dec. 31, 2017) as the validating set to verify the forecasting performance of the data compressed by trained encoder model, the auto-encoder model with high forecasting accuracy will be preserved. Finally, we use one year (from Jan. 1, 2018 to Dec. 31, 2018) as the test set to compare the forecast results.

In order to verify the practicability of the proposed method, three representative forecasting algorithms including BP neural network, support vector machines (SVM) and long short-term memory (LSTM) recurrent neural network sharing exactly the same variables were chosen. Moreover, three metrics, i.e. MAPE, MAE and RMSE are used to measure the forecasting accuracy, which are defined as follows:

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|D_t^M - \hat{D}_t^M\right|}{D_t^M}\times 100\% \tag{1}$$

$$RMSE = \frac{1}{T}\sqrt{\sum_{t=1}^{T}\left(D_t^M - \hat{D}_t^M\right)^2} \tag{2}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|D_t^M - \hat{D}_t^M\right| \tag{3}$$

where D_t^M and $\hat{D}_t^M$ are the actual and forecasted monthly electricity consumption, respectively. T is the time span of the forecasting.

The computational experiments in this paper are performed using MATLAB (R2016a) and Python 3.7 on a laptop equipped with Intel Core i5-7300HQ 2.50 GHz CPU, 12GB usable RAM and Microsoft Windows 10 Home Edition.

B. Comparison results between conventional method and multi-step forecasting method

Fig. 5 shows the forecast results of conventional ECF method using monthly data and multi-step forecasting method using hourly data without data compression. Here, the conventional method is composed of one-step forecasting model. It worth noting that the seasonal term is generally considered in monthly ECF. Since the season length of electricity usage pattern is verified as 12 [17], the input and output of the model are the monthly electricity consumption of previous twelve months $[D_{t-11}^M, \ldots, D_t^M]$ and the next month D_{t+1}^M. Similarly, considering the seasonal term, the input and output of the iterative multi-step forecasting model are the hourly electricity consumption $[D_{t-8759}^H, D_{t-167}^H, D_{t-23}^H, D_{t-11}^H, \ldots, D_t^H]$ and the next hour $[D_{t+1}^H]$. Then, the trained model is used to iteratively forecast T times to get per hour electricity consumption in the next month. T is determined by the number of hours in the next month.

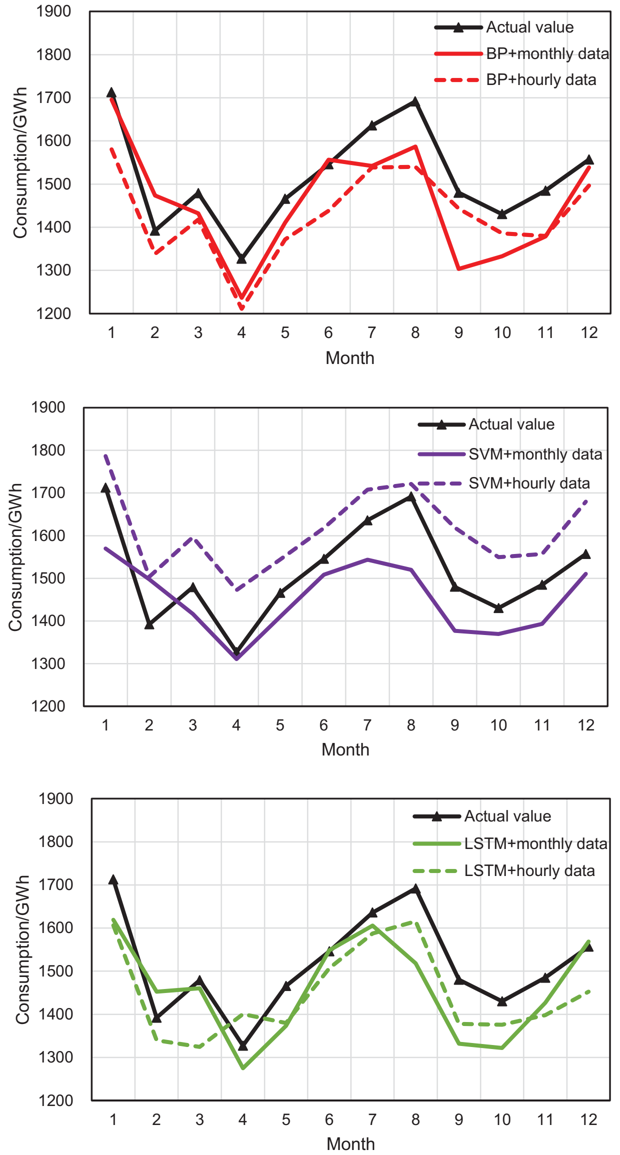

Fig. 5. Forecast results of conventional ECF method using monthly data and multi-step forecasting method using hourly data.

TABLE II. ACCURACY MEASURES OF CONVENTIONAL ECF METHODS AND MULTI-STEP FORECASTING METHODS

Forecasting model	Data granularity	Forecasting steps	MAPE(%)	MAE	RMSE
BP	Monthly	one-step	5.03	74.97	87.84
	Hourly	multi-step	5.78	88.46	95.19
SVM	Monthly	one-step	5.32	82.30	92.85
	Hourly	multi-step	6.48	95.94	101.55
LSTM	Monthly	one-step	4.66	70.91	87.79
	Hourly	multi-step	5.43	82.14	87.80

The identical measures of error are enumerated in Table II. All the three forecasting algorithms perform better when using monthly data than when using hourly data. The LSTM one-step forecasting model achieves the best MAPE of 4.66%, while the SVM multi-step forecasting model gets the worst 6.48%. It can be seen from Fig. 5 that, both of the conventional ECF method and multi-step forecasting method based on three representative algorithms cannot provide satisfactory forecast results. The reason for this phenomenon might be that, for the conventional ECF method, the low accuracy is cause by the over-fitting of the forecasting algorithms due to the small number of training samples. While for the multi-step forecasting method, the low accuracy is caused by the accumulation of errors in the process of iterative forecasting.

C. Comparison results of two data compression methods

Here, we compare the results between two data compression methods based on the data decomposition-accumulation approach. To evaluate the efficiency and performance of proposed AENN data compression method, two comparative cases are conducted. *Case 1*: Monthly ECF using data decomposition-accumulation approach combined with aggregated data compression method. *Case 2*: Monthly ECF using data decomposition-accumulation method combined with AENN data compression method.

For a fair and comprehensive comparison, two compression scales ($k = 6$ and $k = 8$) are discussed in both cases, which share the same input and output as shown in Table III.

TABLE III. INPUT AND OUTPUT OF THE MODEL IN TWO CASES

Scale	Case	Input	Output
$k = 6$	*Case1*	$[Agg\_D^{Wi}_{t-207}, Agg\_D^{Wi}_{t-31}, Agg\_D^{Wi}_{t-15}, Agg\_D^{Wi}_{t-7}, ..., Agg\_D^{Wi}_{t}]$	$[Agg\_D^{W_i}_{t+1}]$
	Case2	$[AE\_D^{Wi}_{t-207}, AE\_D^{Wi}_{t-31}, AE\_D^{Wi}_{t-15}, AE\_D^{Wi}_{t-7}, ..., AE\_D^{Wi}_{t}]$	$[AE\_D^{W_i}_{t+1}]$
$k = 8$	*Case1*	$[Agg\_D^{Wi}_{t-155}, Agg\_D^{Wi}_{t-23}, Agg\_D^{Wi}_{t-11}, Agg\_D^{Wi}_{t-7}, ..., Agg\_D^{Wi}_{t}]$	$[Agg\_D^{W_i}_{t+1}]$
	Case2	$[AE\_D^{Wi}_{t-155}, AE\_D^{Wi}_{t-23}, AE\_D^{Wi}_{t-11}, AE\_D^{Wi}_{t-7}, ..., AE\_D^{Wi}_{t}]$	$[AE\_D^{W_i}_{t+1}]$

TABLE IV. SAMPLE ENTROPY OF VARIOUS DATA GRANULARITY

Data granularity	Sample entropy
D^H	0.8658
D^{W_i}	0.8891
$Agg\_D^{W_i}$ ($k = 6$)	1.7395
$AE\_D^{W_i}$ ($k = 6$)	0.9605
$Agg\_D^{W_i}$ ($k = 8$)	1.6625
$AE\_D^{W_i}$ ($k = 8$)	1.0044

The Sample Entropy theory is used to analyze the predictability of electricity consumption data of various granularity. Generally, the higher the sample entropy, the lower the predictability. Herein, for monthly data D^M and hourly data D^H, the sample entropy is calculated directly. For D^{W_i}, $Agg\_D^{W_i}$ and $AE\_D^{W_i}, i = 1, ..., 7$, the average sample entropy of seven series is calculated. As we can see in Table IV, the sample entropy of D^H is the minimum, which indicates that hourly fine-grained data indeed have high predictability; The average sample entropy of D^{W_i} is slightly larger than D^H, since the decomposition method interrupts the continuity of the original data series; The average sample entropy of $Agg\_D^{W_i}$ is about twice as large as $AE\_D^{W_i}$, which indirectly proves the effectiveness of the proposed method in improving the predictability of compressed electricity consumption data.

The forecast results using different data compression methods are recoded in Table V ($k = 6$) and Table VI ($k = 8$), while the best results of the same forecasting algorithm are in grey. It is obviously that the forecasting accuracy based on proposed AENN data compression method tends to be higher than that of additive aggregation data compression method in most months. Table VII gives the three forecasting metrics correspondingly. Results also show that in both compression scales, the three forecasting algorithms performs better while using $AE\_D^{W_i}$. The LSTM using $AE\_D^{W_i}$ ($k = 8$) achieves the best MAPE of 2.56%. Nevertheless, it is noting that not all forecasting algorithms perform well when k = 8, which indicates that it is indeed difficult to find a specific compression scale to ensure optimal predictability of the compressed data.

TABLE VII. ACCURACY MEASURES OF FORECAST RESULTS OF DIFFERENT DATA COMPRESSION METHODS

Scale	Forecasting model	MAPE(%)	MAE	RMSE
$k = 6$	BP using $Agg\_D^{W_i}$	4.08	63.331	75.852
	BP using $AE\_D^{W_i}$	2.99	46.852	60.787
	SVM using $Agg\_D^{W_i}$	3.94	60.110	68.468
	SVM using $AE\_D^{W_i}$	2.74	43.099	58.628
	LSTM using $Agg\_D^{W_i}$	3.88	58.767	67.181
	LSTM using $AE\_D^{W_i}$	2.86	44.292	51.867
$k = 8$	BP using $Agg\_D^{W_i}$	4.37	67.349	75.034
	BP using $AE\_D^{W_i}$	3.10	49.123	65.801
	SVM using $Agg\_D^{W_i}$	3.99	61.768	73.192
	SVM using $AE\_D^{W_i}$	2.79	43.850	59.296
	LSTM using $Agg\_D^{W_i}$	3.22	49.617	57.869
	LSTM using $AE\_D^{W_i}$	2.56	39.564	50.713

IV. CONCLUSION

An AENN data compression based monthly ECF method is proposed to reduce the forecasting steps while using hourly data, so that can improve monthly ECF results. We combined the proposed method with BPNN, SVM and LSTM forecasting algorithms to evaluated the performance with the datasets from PJM. Simulation results show that the strategy performs better than additive aggregation data compression method. It is worth noting that AENN is more conducive to guaranteeing the predictability of the series in data compression than simple additive method, which has been rarely studied in other papers. The research is important for load aggregator [30, 31] and micro-grid operation [32, 33] with distributed PVs [34, 35]. In the future, we intend to develop an optimization problem to study the balance between data compression using AENN and guaranteeing the predictability of series.

TABLE V. FORECAST RESULTS OF DIFFERENT DATA COMPRESSION METHODS ($k = 6$)

Month	True Value/GWh	Forecast results					
		BP using $Agg\_D^{W_i}$	BP using $AE\_D^{W_i}$	SVM using $Agg\_D^{W_i}$	SVM using $AE\_D^{W_i}$	LSTM using $Agg\_D^{W_i}$	LSTM using $AE\_D^{W_i}$
Jan.	1712.78	1614.51	1604.84	1597.44	1599.03	1624.75	1639.09
Feb.	1392.15	1429.07	1386.83	1424.67	1389.33	1445.08	1412.91
Mar.	1479.13	1526.23	1456.38	1525.71	1448.54	1589.06	1524.27
Apr.	1327.16	1317.83	1327.20	1402.17	1327.43	1348.67	1327.91
May	1465.89	1348.92	1393.25	1441.39	1355.95	1397.35	1400.95
Jun.	1545.85	1501.02	1531.25	1487.08	1519.93	1464.70	1478.33
Jul.	1635.92	1593.11	1574.13	1607.25	1606.34	1635.34	1616.37
Aug.	1691.67	1547.16	1591.80	1590.76	1593.98	1627.06	1610.32
Sep.	1480.04	1458.48	1496.60	1504.66	1451.82	1466.50	1459.47
Oct.	1430.35	1450.89	1468.94	1497.47	1436.79	1515.22	1468.60
Nov.	1484.89	1377.16	1381.16	1372.00	1429.42	1401.95	1405.22
Dec.	1556.82	1626.22	1575.21	1591.21	1573.32	1593.39	1576.13

TABLE VI. FORECAST RESULTS OF DIFFERENT DATA COMPRESSION METHODS ($k = 8$)

Month	True Value/GWh	Forecast results					
		BP using $Agg\_D^{W_i}$	BP using $AE\_D^{W_i}$	SVM using $Agg\_D^{W_i}$	SVM using $AE\_D^{W_i}$	LSTM using $Agg\_D^{W_i}$	LSTM using $AE\_D^{W_i}$
Jan.	1712.78	1572.73	1588.50	1587.87	1612.42	1600.25	1601.61
Feb.	1392.15	1447.77	1389.39	1366.98	1388.77	1437.41	1397.57
Mar.	1479.13	1410.84	1443.90	1471.33	1450.62	1509.64	1504.54
Apr.	1327.16	1264.78	1376.28	1285.45	1325.08	1307.44	1336.33
May	1465.89	1405.47	1424.30	1363.66	1355.37	1436.91	1411.74
Jun.	1545.85	1489.87	1501.68	1514.35	1514.41	1466.30	1514.31
Jul.	1635.92	1564.94	1540.28	1593.39	1608.76	1629.54	1683.40
Aug.	1691.67	1568.67	1553.41	1572.94	1582.51	1617.04	1656.35
Sep.	1480.04	1497.48	1495.95	1453.83	1477.98	1455.75	1495.36
Oct.	1430.35	1482.95	1439.25	1472.91	1407.20	1500.74	1520.83
Nov.	1484.89	1410.38	1470.17	1378.86	1418.46	1417.33	1441.83
Dec.	1556.82	1583.74	1537.92	1628.65	1578.78	1592.42	1563.07

ACKNOWLEDGMENT

This work was supported by the National Key R&D Program of China (2018YFE0122200) and the Science and Technology Project of State Grid Gansu Electric Power Company (522730190009).

REFERENCES

[1] J. Yang, J. Zhao, F. Lup, F. Wen and Z. Dong, "Decision-Making for Electricity Retailers: A Brief Survey," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4140-4153, Sept. 2018.

[2] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-khah and J.P.S. Catalão, "Synchronous pattern matching principle based residential demand response baseline estimation: mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972-6985, Nov. 2018.

[3] W. Charytoniuk, E. Don Box, W. J. Lee, M. S. Chen, P. Kotas, and P. Van Olinda, "Neural-network-based demand forecasting in a deregulated environment," *IEEE Trans. Ind. Appl.*, vol. 36, no. 3, pp. 893–898, May/Jun. 2000.

[4] F. Wang, X. Ge, K. Li and Z. Mi, "Day-ahead market optimal bidding strategy and quantitative compensation mechanism design for load aggregator engaging demand response," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp.5564 – 5573, Nov./Dec. 2019.

[5] M. Yang, Y. Lin and X. Han, "Probabilistic wind generation forecast based on Sparse Bayesian Classification and Dempster-Shafer Theory," *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 1998–2005, May/Jun. 2016.

[6] F. Wang, Z. Zhang, C. Liu, Y. Yu, S. Pang, N. Duić, M. Shafie-khah and J.P.S. Catalão, "Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting," *Energy Convers. Manage.*, vol. 181, pp. 443–462, Feb. 2019.

[7] S. Fan, L. Chen, and W. J. Lee, "Short-term load forecasting using comprehensive combination based on multimeteorological information," *IEEE Trans. Ind. Appl.*, vol. 45, no. 4, pp. 1460–1466, Jul./Aug. 2009.

[8] F. Wang, Z. Zhen, C. Liu, Z. Mi, B. Hodge, M. Shafie-khah and J.P.S. Catalão, "Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting," *Energy Convers. Manag.*, vol. 157, pp. 123–135, Feb. 2018.

[9] C. Kuster, Y. Rezgui and M. Mourshed, "Electrical load forecasting models : A critical systematic review," *Sustain. Cities Soc.*, vol. 35, pp. 257–270, Aug. 2017.

[10] S. R. Khuntia, J. L. Rueda and M. A. M. M. van der Meijden, "Forecasting the load of electrical power systems in mid- and long-term horizons: a review," *IET Gener. Transm. Distrib.*, vol. 10, no. 16, pp. 3971–3977, Aug. 2016.

[11] C. Yuan, S. Liu and Z. Fang, "Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model," *Energy*, vol. 100, pp. 384–390, Feb. 2016.

[12] L. Wu, X. Gao, Y. Xiao, Y. Yang, and X. Chen, "Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China," *Energy*, vol. 157, pp. 327–335, 2018.

[13] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen and K. Li, "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting," *Appl. Sci.*, vol. 8, no. 8, pp. 1286, Aug. 2018.

[14] M. Ghiassi, D. K. Zimbra and H. Saidane, "Medium term system load forecasting with a dynamic artificial neural network model," *Electr. Power Syst. Res.*, vol. 76, no. 5, pp. 302–316, Jun. 2006.

[15] Z. Hu, Y. Bao, R. Chiong and T. Xiong, "Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection," *Energy*, vol. 84, pp. 419–431, Apr. 2015.

[16] H. X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev*., vol. 16, no. 6, pp. 3586–3592, Apr. 2012.

[17] W. C. Hong, Y. Dong, W. Y. Zhang, L. Y. Chen and B. K. Panigrahi, "Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 44, no. 1, pp. 604–614, Sep. 2012.

[18] T. Yalcinoz and U. Eminoglu, "Short term and medium term power distribution load forecasting by neural networks," *Energy Convers. Manag.*, vol. 46, no. 9–10, pp. 1393–1405, Sep. 2004.

[19] E. González-romera, M. Á. Jaramillo-morán and D. Carmona-fernández, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Trans. Power Syst*, vol. 21, no. 4, pp. 1946–1953, Nov. 2006.

[20] F. Wang, K. Li, L. Zhou, H. Ren, J. Contreras, M. Shafie-khah and J.P.S. Catalão, "Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 105, pp. 529–540, Feb. 2019.

[21] Z. Zhen, S. Pang, F. Wang, K. Li, H. Ren, M. Shafie-khah and J. P. S. Catalão, "Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting," *IEEE Trans. Ind. Appl*., vol. 55, no. 4, pp. 3331–3342, Jul./Aug. 2019.

[22] Q. Chen, F. Wang, B. Hodge, J. Zhang, Z. Li, M. Shafie-khah and J.P.S. Catalão, "Dynamic price vector formation model-based automatic demand response strategy for PV-assisted EV charging stations," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2903–2915, Nov. 2017.

[23] Z. Li, K. Li, Z. Xuan, Z. Zhen and F. Wang, "Decomposition-accumulation principle based monthly electricity consumption forecasting approach using EMD-XGBoost model," in *2019 IEEE Sustainable Power & Energy Conference*, Beijing, China, Nov. 21–23, 2019, pp. 2425-2430.

[24] H. Jahangir, H. Tayarani, S. Baghali, A. Ahmadian, A. Elkamel, M. Aliakbar and M. Castilla, "A novel electricity price forecasting approach based on dimension reduction strategy and rough artificial neural networks," *IEEE Trans. Ind. Informatics*, vol. 3203, no. c, pp. 1–1, 2019.

[25] F. J. Pulgar, F. Charte, A. J. Rivera and M. J. del Jesus, "Choosing the proper autoencoder for feature fusion based on data complexity and classifiers: Analysis, tips and guidelines," *Inf. Fusion*, vol. 54, pp. 44–60, Jul. 2019.

[26] D. Del Testa and M. Rossi, "Lightweight lossy compression of biometric patterns via denoising autoencoders," *IEEE Signal Process. Lett*., vol. 22, no. 12, pp. 2304–2308, Dec. 2015.

[27] Hinton G. E. and Salakhutdinov R. R., "Reducing the dimensionality of data with neural networks," *Science*, vol 313, no. 5786, pp. 504–507. Jul. 2006.

[28] H. B. Xie, W. X. He and H. Liu, "Measuring time series regularity using nonlinear similarity-based sample entropy," *Phys. Lett. Sect. A Gen. At. Solid State Phys.*, vol. 372, no. 48, pp. 7140–7146, Oct. 2008.

[29] Data Miner 2, PJM. Available online: http://dataminer2.pjm.com/list (accessed on 29 November 2019).

[30] F. Wang, B. Xiang, K. Li, X. Ge, H. Lu, J. Lai, P. Dehghanian, "Smart Households' Aggregated Capacity Forecasting for Load Aggregators under Incentive-based Demand Response Programs." *IEEE Trans. Ind. Appl.*, Early Access, Jan. 2020, DOI: 10.1109/TIA.2020.2966426.

[31] F. Wang, H. Xu, T. Xu, K. Li, M. Shafie-khah, and J. P. S. Catalão, "The values of market-based demand response on improving power system reliability under extreme circumstances," *Appl. Energy*, vol. 193, pp. 220–231, May 2017, DOI: 10.1016/j.apenergy.2017.01.103.

[32] F. Wang, L. Zhou, H. Ren, X. Liu, M. Shafie-khah, and J.P.S. Catalão, "Multi-objective optimization model of source-load-storage synergetic dispatch for a building energy management system based on TOU price demand response," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1017–1028, Mar.-Apr. 2018, DOI: 10.1109/TIA.2017.2781639.

[33] J. Lai, X. Lu, F. Wang, P. Dehghanian, and R. Tang, "Broadcast gossip algorithms for distributed peer-to-peer control in AC microgrids," *IEEE Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2241-2251, May-Jun. 2019, DOI: 10.1109/TIA.2019.2898367.

[34] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duic, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, Art. no. 113595, 2019, DOI: 10.1016/j.apenergy.2019.113595.

[35] Z. Zhen, Z. Xuan, F. Wang, R. Sun, N. Duic and T.Jin, "Imagephase shift invariance based multi-transform-fusion method for cloud motion displacement calculation using sky images," *Energy Convers. Manag.*, vol. 197, p. 111853, Oct. 2019.