Monthly Net Electricity Consumption Prediction Under High Penetration of Distributed Photovoltaic System

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

Kangping Li
Department of Electrical Engineering
Tsinghua University
Beijing 100084, China
kangpingli@tsinghua.edu.cn

João P. S. Catalão
Faculty of Engineering
University of Porto and INESC TEC
Porto 4200-465, Portugal
catalao@fe.up.pt

2021 International Conference on Smart Energy Systems and Technologies (SEST) | 978-1-7281-7660-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/SEST50973.2021.9543320

***Abstract*—Net electricity consumption (NEC) is the result of the joint action of actual electricity consumption (AEC) and distributed photovoltaic (DPV) generation. The accuracy of NEC prediction affects the retailers' gaming and ultimate interests in electricity market. The ascending DPV installations present new challenges, with the modifications to the NEC curve becoming greater as DPV penetration increases. To track the changes in DPV penetration and improve the prediction accuracy under high penetration of DPV, a monthly NEC prediction model assembled by support vector regression and time series modeling under an online update framework is proposed. First, the DPV features are extracted from a few known solar customers' information to identify whether other customers install DPV or not. Second, an online update framework is proposed and its accuracy is verified by two validations regarding the conversion of non-solar customers to solar customers (namely the change of DPV penetration). Third, a NEC decoupling model based on historical NEC data of months without DPV installation is established. Finally, a monthly NEC prediction model under different DPV penetrations is proposed. Simulation results show that the proposed prediction method with an online update is more accurate than the individual time series model, and the performance of the prediction model is getting better with the increasing DPV penetration.**

Keywords—Retailers, Electricity Market, Distributed Photovoltaic Penetration, Monthly Net Electricity Consumption Prediction, Online Update Model

NOMENCLATURE

A. Abbreviations

DPV	Distributed photovoltaic
AEC	Actual electricity consumption
NEC	Net electricity consumption
SVR	Support vector regression
ARIMA	Auto-regressive integrated moving average

B. Symbols

T	Set of timeslot
P_t, P_t^0	Set of electricity consumption of each customer whose DPV information is unknown/known
$\boldsymbol{x}$, $\boldsymbol{y}$	Input/Output vector of SVR
$f(\cdot)$	Function of SVR
w, b	Parameters of SVR
t	Index of timeslot
g, i, j	Number of the hour, month and year
H_g, M_i, Y_j	The g th hour, i th month and j th year
$P_{Net,t}$, $P_{Net,t}^0$	NEC of each customer whose DPV information is unknown/known
$P_{Load,t}$, $P_{Load,t}^0$	AEC of each customer whose DPV information is unknown/known
$P_{DPV,t}$, $P_{DPV,t}^0$	DPV generation of each customer whose DPV information is unknown/ known
$s_{cl,g}$	Slope of the cut line at H_g
$r_{tp,g}$	Degree of concavity at H_g
$r_{s,g}$	Average rate of change of the slope
ε_k	Prediction error of the prediction of P_{Load,H_g,M_i,Y_j} at k th time
ε_{av}	Average prediction error
$\hat{y}_{Load,H_g,M_i,Y_j}$	Real value of AEC of P_{Load,H_g,M_i,Y_j}
y_{Load,H_g,M_i,Y_j}	Prediction result of AEC of P_{Load,H_g,M_i,Y_j}
y_{DPV,H_g,M_i,Y_j}	Prediction result of DPV generation of P_{Load,H_g,M_i,Y_j}
$P_{Net,fluctuation}$	Fluctuation of NEC
$P_{Net,trend}$	Trend of NEC
p	Partial correlation order of ARIMA
d	Differential order of ARIMA
q	Auto-correlation order of ARIMA
B, C	Value of the Bayesian/Akaike information validation
n_{cor}	Correct identification number of non-solar customers converted to solar customers
n_{tol}	Real number of all non-solar customers to be identified that are converted to solar customers
a_{update}	Accuracy of the DPV installation update check validations
y_{ac_t}, y_{pre_t}	Real/predicted value of NEC

I. INTRODUCTION

Driven by national policies and other factors, a growing number of customers have a preference to install DPV. The popularity of DPV might as well be considered as a boon and a challenge [1]. Retailers who need to predict the monthly net electricity consumption (NEC) (namely the difference between monthly actual electricity consumption (AEC) and monthly DPV generation) of their customers before participating in the forward market are required to take the introduction of DPV into account [2-4]. With the rapid growth of DPV penetration, compared with the monthly AEC characteristics, the monthly NEC characteristics are different in trend, seasonality, and randomness, which may lead the existing monthly AEC prediction models are not available. Thus, it is of great urgency to improve the prediction accuracy of NEC so that to ensure the interests of retailers and customers [5-6].

The NEC prediction ideas can be divided into two categories: direct ideas and indirect ideas. The direct idea, whose prediction target is the NEC, involves using the DPV generation data monitored by the additional metering equipment installed, but this will significantly increase the cost due to the large number of DPV. The indirect idea, whose prediction targets are the AEC and DPV generation obtained after NEC decomposition, involves decomposing the net load into AEC and DPV generation by mining the hidden information contained in various data sources with sufficient temporal and spatial granularity [7-10]. In [11], Bayesian in-depth learning is used to capture uncertainty for residential net load prediction. Reference [12] uses the Gaussian process to predict the probability of residential electrical energy, DPV generation, and individual household net demand. Reference [13] details the application of probabilistic prediction methods in solar irradiance prediction and load prediction. In [14], it is found that increasing the proportion of DPV output power in the net load of residential users can improve the clarity and reliability of probability prediction in spring and winter. The photovoltaic generation prediction method proposed in [15] aiming for invisible photovoltaic sites is different from [16] in that it belongs to unsupervised learning. So, the historical data of the site is not required for model training, and it is suitable for high penetration.

For the above exiting researches, there are still some issues that need to be further addressed. First, it is difficult to distinguish the installation of DPV just based on the net load curve under the changeable weather scenarios. For example, when the net load curve is a clearly "concave" type, it is difficult to distinguish whether the customers' actual electrical consumption has reduced or DPV generation has increased. Second, they do not take into account the DPV installation information updates for non-solar customers, which leads to a part of DPV generation cannot be calculated in the NEC. Generally, it is common to regard this part as the prediction error of the AEC, which may cause deviations in the extraction of customer load characteristics and turn affects the net load prediction accuracy.

The main contributions of this paper are as follows:

(1) A novel monthly NEC prediction method utilizes decoupling technology is proposed.

(2) An online update framework of customer DPV installation information is proposed to improve the performance of the monthly NEC prediction model further.

(3) Based on the differences of fluctuation between AEC and DPV generation, it is judged whether it is the error of the AEC prediction or the installation of DPV.

The framework and specific process will be introduced in Section II. Next, the data set, case study, and performance evaluation will be proposed in Section III. Finally, the conclusion and future work will be written in Section IV.

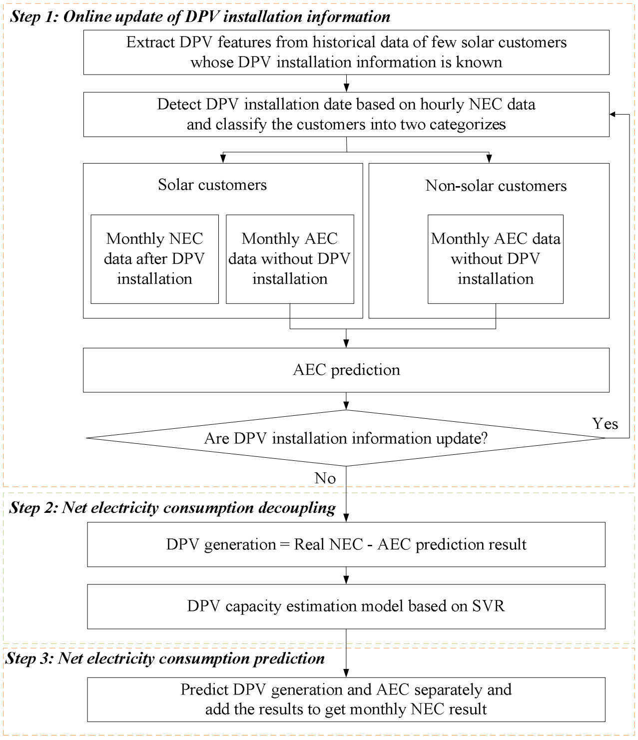

Fig. 1. Data updates and prediction framework

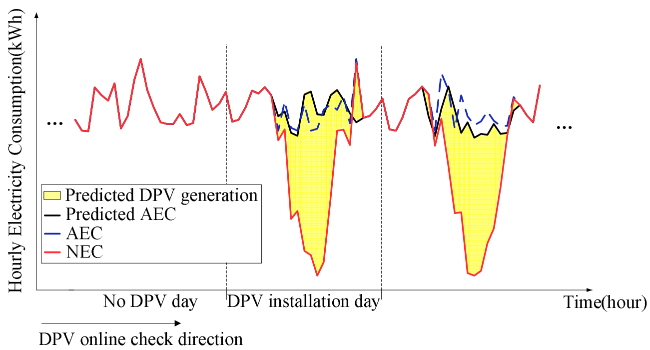

Fig. 2. Online update check diagram

II. PROPOSED METHODOLOGY

A. Framework

The work is to predict the monthly NEC of the customers for a retailer. The DPV penetration varies with the installation of DPV. To distinguish the different characteristics of monthly AEC and DPV generation and update DPV installation information in time, this paper decouples monthly NEC into monthly AEC and DPV generation and predicts them respectively under an online update data framework. This process consists of three steps (namely online update, decoupling, and prediction), as shown in Fig. 1.

Step 1: Online update procedure. When the customer does not install DPV, the NEC is equal to the AEC. When the customer is installed DPV, the NEC is equal to the AEC minus the amount of DPV generation. The difficulty is that customer DPV installation is time-varying. Thus, it is necessary to have a constant check whether a non-solar customer has newly installed DPV. The accuracy of the prediction can be improved by subtracting the DPV generation of customers who have newly installed DPV in time during the NEC prediction. The accuracy of the monthly electricity consumption prediction can be improved by using hourly data besides monthly data [17]. Above all, to master the DPV installation information of all customers, the DPV features are extracted from the historical hourly data of few solar customers whose DPV installation information is known. Then according to the features, other customers whose DPV installation information is unknown are classified into two categories: solar customers and non-solar customers.

As for solar customers, the DPV installation date can also be detected during detecting the DPV features. Besides, the online update of the solar customer database is targeted at non-solar customers. The diagram introducing a check for the new installation of DPV is shown in Fig.2. For non-solar customers, the monthly NEC is equal to the monthly AEC. Taking into account the prediction error, expect for the error value, validation used to determine whether a non-solar customer has become a solar customer has been added the shape of error by comparing random fluctuations in AEC with random fluctuations after the overlay of DPV generation.

Step 2: Decoupling procedure. Before the installation of DPV, the hourly AEC and the hourly NEC are equal. Based on this, the AEC on the day of the predicted DPV installation date is subtracted from the NEC on that day to obtain the DPV generation for that day, which is used as an output to the model for DPV capacity prediction. The model is based on support vector regression (SVR) considering the advantages of SVR model in handling small sample data and excellent generalization.

Step 3: Prediction procedure. The monthly AEC and DPV generation are predicted respectively, and their prediction results are added to get the monthly NEC prediction result.

B. *Feature Exaction and Classification*

Let T denote time and consist of the hour H_g, month M_i, and year Y_j, namely $T=\{H_g,M_i,Y_j\}, t\in T$. Among them, *g*, *i* and *j* represent the number of hour *H*, month *M* and year *Y* respectively. Let P_t denote the electricity consumption of each customer whose DPV information is unknown and define $P_t=\left\{P_{Net,t},P_{Load,t},P_{DPV,t}\right\}$. Let P_t^0 be the electricity consumption of each customer whose DPV information is known and define $P_t^0=\left\{P_{Net,t}^0,P_{Load,t}^0,P_{DPV,t}^0\right\}$. The relationship between AEC $P_{Load,t}$ and $P_{Load,t}^0$, NEC $P_{Net,t}$ and $P_{Net,t}^0$, and DPV generation $P_{DPV,t}$ and $P_{DPV,t}^0$ are described as following:

$$P_{Net,t}=P_{Load,t}-P_{DPV,t} \tag{1}$$

$$P_{Net,t}^0=P_{Load,t}^0-P_{DPV,t}^0 \tag{2}$$

Combining the NEC data of solar customers in different locations and the historical weather data, the NEC curves before and after DPV installation are compared so as to extract the typical features. In light of the characteristics of DPV panel generation under different weather scenarios, the corresponding feature extraction period and key indicators are mentioned in [18]. Besides, the daily net electricity curve may look distinctly low during noon due to the peak in DPV generation. Since there is a significant turning from about 8 o'clock to 11 o'clock, the magnitude of the slope and the speed of change during this period are expressed as indicators of the DPV installation and are calculated by the following formulas. The remaining unknown customers are divided into solar and non-solar customers according to the above indicators.

$$s_{cl,g}=\frac{P_{Net,H_{g+\Delta g},M_i,Y_j}-P_{Net,H_g,M_i,Y_j}}{H_{g+\Delta g}-H_g} \tag{3}$$

$$r_{tp,g}=\frac{P_{Net,H_{g2},M_i,Y_j}-P_{Net,H_{g1},M_i,Y_j}}{H_{g_2}-H_{g_1}}-P_{Net,H_g,M_i,Y_j}, g\in[g_1,g_2] \tag{4}$$

$$r_{s,g}=\frac{s_{cl,g+\Delta g}-s_{cl,g}}{H_{g+\Delta g}-H_g} \tag{5}$$

where $s_{cl,g}$ is the slope of the cut line at H_g. $r_{tp,g}$ represents the degree of concavity during H_{g_1} and H_{g_2}. $r_{s,g}$ is the average rate of change of the slope during H_g and $H_{g+\Delta g}$.

C. *DPV Installation Update Check*

The hourly AEC prediction error can be used to check the non-solar customers whether installs the DPV newly or not. The hourly AEC data of non-solar customers are predicted per day based on an artificial neuron network [19] and the prediction errors per time are recorded. The non-solar customer whose hourly AEC prediction error is higher than the average of the former error ε_{av} is suspected of installing DPV newly and the formulas are as follows:

$$\varepsilon_{av}=\frac{1}{m}\sum_{k=1}^{m}\varepsilon_k \tag{6}$$

$$\left|\hat{y}_{Load,H_g,M_i,Y_j}-y_{Load,H_g,M_i,Y_j}\right|>\varepsilon_{av} \tag{7}$$

where ε_k is the error of the prediction of P_{Load,H_g,M_i,Y_j} and *k* is the time of the prediction. $\hat{y}_{Load,H_g,M_i,Y_j}$ is the real AEC of P_{Load,H_g,M_i,Y_j} and y_{Load,H_g,M_i,Y_j} is the prediction result of AEC P_{Load,H_g,M_i,Y_j}.

Furthermore, to make sure that the sudden increase prediction error does not come from the prediction method itself, rather than from the installation of DPV by non-solar customers, validation has been added to compare random fluctuations in AEC with random fluctuations after the overlay of DPV generation. The stochastic fluctuations $P_{Net,fluctuation}$ in the AEC series are obtained by removing the trend $P_{Net,trend}$ from the actual consumption series, as shown in following:

$$P_{Net,fluctuation}=P_{Net,t}-P_{Net,trend} \tag{8}$$

The mathematical details of this calculation can be found in [20]. Comparing it with the NEC series fluctuation of solar customers whose DPV installation information is known on the same day, which is obtained by excluding the trend of AEC and the DPV generation from the NEC series. If both of these validations are met, this non-solar customer is converted to be a solar customer.

D. Net Electricity Consumption Decoupling

The difference between AEC and NEC is DPV generation. Before installing DPV, the curves of hourly AEC and the hourly NEC are overlapped. According to the hourly AEC prediction of the day before the date of installation of DPV, the amount of DPV generation is obtained by using the hourly AEC on the day of installation of DPV minus the hourly AEC prediction results, as shown in the following formula:

$$P_{Load,M_i,Y_j} = \sum_{q=1}^{k_2} P_{Net,H_q,M_i,Y_j} + \sum_{q=k_1}^{k_2} y_{DPV,H_q,M_i,Y_j} \tag{9}$$

where k_2 is the total number of the hours in DPV installation month M_i *and* k_1 is the number of the hours after installing DPV in month M_i.

The predicted DPV generation after decoupling is shown as:

$$y_{DPV,H_q,M_i,Y_j} = y_{Load,H_q,M_i,Y_j} - P_{Net,H_q,M_i,Y_j} \tag{10}$$

E. DPV Capacity Estimation

Let $\boldsymbol{x}$ be the vector of the inputs of SVR, namely $\boldsymbol{x} = \left[x_{wc}, x_{te}, x_g\right]$. It includes weather conditions x_{wc} (such as sunny, rainy, and cloudy), temperature x_{te}, and DPV generation x_g. Let $\boldsymbol{y}$ denote the DPV capacity information of few solar customers whose DPV information is known. A small number of the known solar customer hourly data is set as the training set, and the rest of the solar customer hourly data is set as the test set. Targeted training samples are as follows:

$$D = \left\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \cdots, (x_n, y_n)\right\} \tag{11}$$

$$f(x_i) = w^T x_i + b \tag{12}$$

where $f(x)$ is a function expected to be obtained through learning, and w and b are the parameters to be determined.

F. Monthly Net Electricity Consumption Prediction

To make the output of the prediction model better follow the actual results, the seasonal trend is removed. Differential operations can smooth a class of non-smooth sequences (i.e., sequences with a trend). The essence of the differential operator is to extract determination information by self-regression:

$$\nabla^d x_t = (1-B)^d x_t = \sum_{i=0}^{d} (-1)^i C_d^i x_{t-1} \tag{13}$$

where d is the differential order. B is the value of the Bayesian information criterion and C is the value of the Akaike information criterion.

Then the monthly AEC and DPV generation are predicted separately based on auto-regressive integrated moving average (ARIMA) of order (p,d,q):

$$\nabla^d X(t) = \sum_{i=1}^{p} \phi_i \nabla^d X(t-i) + a_t - \sum_{j=1}^{q} \theta_j a_{t-j} \tag{14}$$

where ∇^d is the dth difference operator, a_t is white noise with variance σ_a^2, ϕ_i and θ_j are model parameters [21]. The AEC and DPV generation obtained by NEC decomposition have seasonal cycle variations to different degrees. A differential operation with a step size of the cycle length for series with a fixed cycle can usually extract the cycle information better. The ARIMA model parameters p and q are determined by the trailing and truncated tails of the auto-correlation and partial correlation coefficients respectively.

G. Evaluation Criteria

To quantify the accuracy of the DPV installation update check validations, the accuracy a_{update} is calculated by following:

$$a_{update} = \frac{n_{cor}}{n_{tol}} \tag{15}$$

where n_{cor} is the correct identification number of non-solar customers converted to solar customers and n_{tol} is the real number of all non-solar customers to be identified that are converted to solar customers.

To quantify the accuracy of the proposed method, the error is evaluated using the mean absolute percentage error (MAPE) and root mean squared error (RMSE):

$$MAPE = \frac{1}{m} \sum_{t=1}^{m} \frac{\left|y_{pre_t} - y_{ac_t}\right|}{y_{ac_t}} \tag{16}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{t=1}^{m} \left|y_{pre_t} - y_{ac_t}\right|^2} \tag{17}$$

where y_{pre_t} is the predicted value of NEC and y_{ac_t} is the real value of NEC. m is the number of the sequence y_{pre_t} and y_{ac_t}.

III. Case Study

A. Dataset and Parameter Settings

The dataset used in our work is from an Australian grid named Ausgrid [22]. The framework proposed in this paper is for a real-life situation where DPV penetration is constantly changing. Based on a half-hourly dataset of 300 known solar households for three years, including hourly DPV generation and hourly AEC, DPV installation dates and DPV generation have been randomly selected to be added to the non-solar customers' AEC series. Thus, DPV penetration is increasing by adding solar customers. Three scenarios have been set up as shown in Fig. 3. As the penetration of DPV changes, it is obvious that the difference between monthly NEC and actual monthly electricity consumption becomes larger. Scenario 1: three hundred customers without DPV for three years; Scenario 2: One hundred customers without DPV installation and two hundred solar customers, of which two hundred solar customers' DPV installation dates are randomly determined over a three-year period, reaching 27% penetration by the end of the third year; Scenario 3: three hundred solar customers by the end of the third year, reaching a penetration of 39%. The DPV installation dates for three hundred solar customers have been randomly set over the three years.

B. Results and Analysis

Fig. 4 illustrates the comparison of the fluctuation of electricity consumption sequence before and after the installation of DPV by one customer. Although the AEC curve and the NEC curve are very similar due to the small output of DPV, the characteristics of the AEC series are found to be changed during the noon hour after the installation of DPV by observing the fluctuation of the two curves. As seen in Fig. 5, adding the validation of random fluctuations in the electricity consumption series has a positive effect on improving the accuracy of DPV installation updates. This is because when the weather is bad or the DPV capacity is small, it is somewhat strained to determine whether to install DPV only by the average error of AEC prediction.

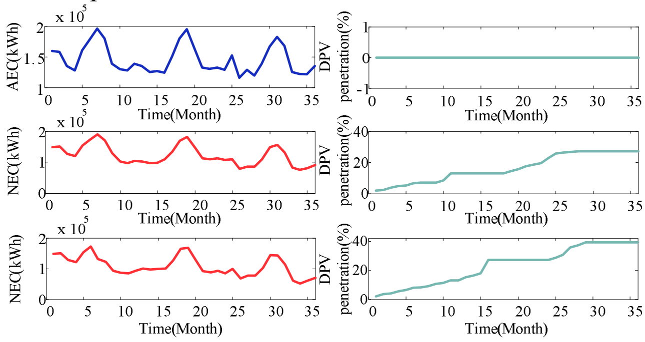

Fig. 3. Three scenarios of DPV penetration setting

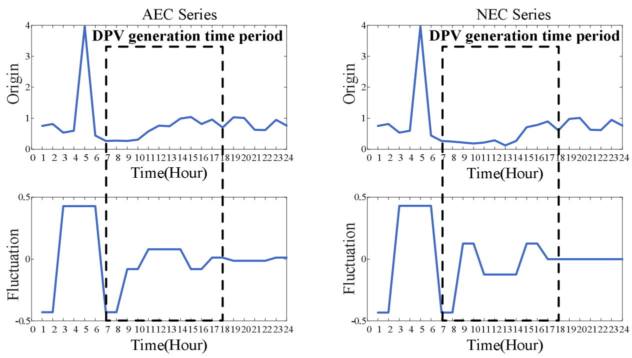

Fig. 4. Comparison of the fluctuation of electricity consumption sequence before and after the installation of DPV by one customer

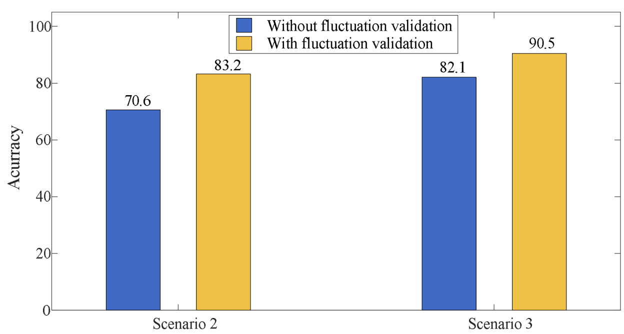

Fig. 5. Comparison of the accuracy of the DPV installation update check with and without fluctuation validation in scenario 2 and scenario 3

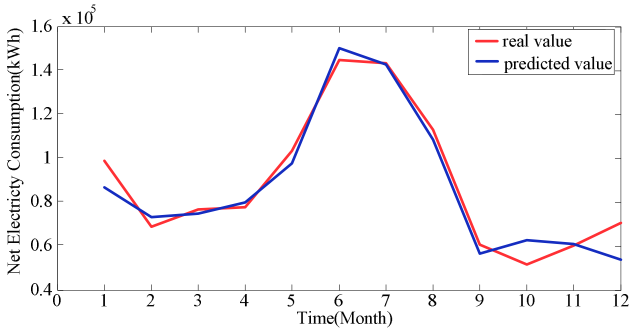

Fig. 6. Monthly NEC prediction results under scenario 3

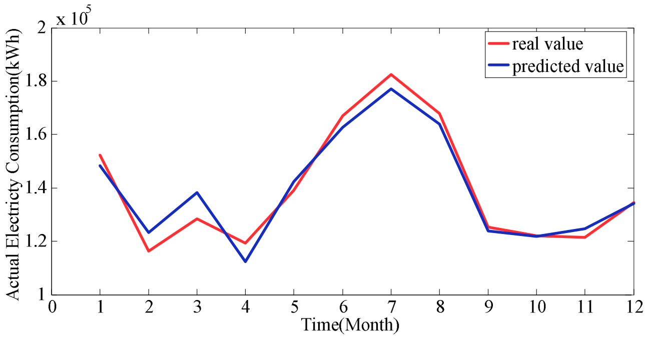

Fig. 7. Monthly AEC prediction results under scenario 3

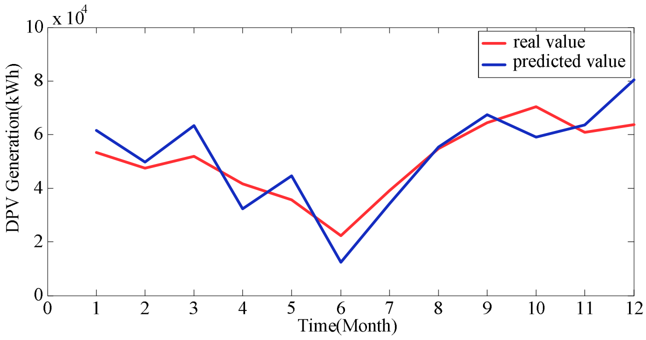

Fig. 8. Monthly DPV generation prediction results under scenario 3

The monthly prediction results of NEC, AEC and DPV generation based on the method proposed in this paper under scenario 3 are respectively shown in Fig. 6, Fig. 7 and Fig. 8. Although the proposed method works well in summer, performs poorly in winter due to strong volatility. Above all, the performance evaluation of proposed method and direct NEC prediction based on ARIMA under different scenarios is shown in table 1. When in the absence of DPV installations, the direct forecasting method is slightly better than the method proposed in this paper, because the cumulative error of decomposition and reorganization is larger in this case. However, when the simulated scenario is that the DPV penetration rate changes over time, the method proposed in this paper has the advantage of detecting the installation of DPV and updating the database of solar customers, making the monthly NEC prediction by the decomposition better than the direct prediction. This advantage becomes obvious as the DPV penetration increases.

TABLE I. PERFORMANCE EVALUATION OF THE PROPOSED METHOD AND ARIMA UNDER DIFFERENT DPV PENETRATION

Scenario	Method	MAPE	RAME(kW)
1	Proposed Method	0.1174	14372.59
	ARIMA	0.0983	14461.18
2	Proposed Method	0.0927	11892.62
	ARIMA	0.1094	13574.31
3	Proposed Method	0.0825	10876.52
	ARIMA	0.1126	13892.45

IV. CONCLUSION

The monthly electrical energy sequence is nonlinear and complicated. In this paper, to decouple DPV generation from the NEC, a classification method combining monthly and fine-grained data is proposed, which is both efficient and accurate. Then the SVR model is used to estimate the DPV panel parameters of solar customers. Next combined with different weather patterns and data updates, DPV generation is predicted for customers who install DPV. In view of the continuous updating of the DPV database, the process of change in DPV penetration is simulated in this paper, making the prediction of DPV generation more realistic. The validation that the fluctuations of monthly AEC series are injected new characteristics of DPV generation is used to determine the former non-solar customers have installed DPV, which ensures that the newly installed DPV generation is subtracted. The proposed monthly NEC prediction model taking into account DPV installation information update outperforms the direct model based on ARIMA under high DPV penetration.

Further research could focus on improving the performance in winter and extending this work to probability prediction based on weather data that are available in real time. Furthermore, retailers will face more challenges in predicting the NEC with the development of advanced metering infrastructure and the popularization of various distributed renewable energy sources [23].

REFERENCES

[1] L. Freire, A. Street, D. A. Lima, and L. A. Barroso, “A Hybrid MILP and Benders Decomposition Approach to Find the Nucleolus Quota Allocation for a Renewable Energy Portfolio,” IEEE Trans. Power Syst., vol. 30, no. 6, pp. 3265-3275, Nov. 2015.

[2] J. Yang, J. Zhao, F. Luo, F. Wen and Z. Y. Dong, "Decision-Making for Electricity Retailers: A Brief Survey," IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 4140-4153, Sept. 2018.

[3] M. Carrion, A. J. Conejo and J. M. Arroyo, “Forward Contracting and Selling Price Determination for a Retailer,” IEEE Trans. Power Syst., vol. 22, no. 4, pp. 2105-2114, Nov. 2007.

[4] A. Kaur, L. Nonnenmacher, C. F.M. Coimbra, “Net load forecasting for high renewable energy penetration grids.” Energy, vol. 114. pp:1073-1084, Nov. 2016.

[5] J. Xie, T. Hong and J. Stroud, “Long-Term Retail Energy Forecasting with Consideration of Residential Customer Attrition,” IEEE Trans. Smart Grid, vol. 6, no. 5, pp. 2245-2252, Sept. 2015.

[6] Y. Wang, Q. Chen, T. Hong and C. Kang, “Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges,” IEEE Transactions on Smart Grid, vol. 10, no. 3, pp. 3125-3148, May 2019.

[7] F. Sossan, L. Nespoli, V. Medici, and M. Paolone, “Unsupervised disaggregation of photovoltaic production from composite power flow measurements of heterogeneous prosumers,” IEEE Trans. Ind. Informat., vol. 14, no. 9, pp. 3904–3913, Sep. 2018.

[8] F. Bu, K. Dehghanpour, Y. Yuan, Z. Wang, and Y. Zhang, “A Data-Driven Game-Theoretic Approach for Behind-the-Meter PV Generation Disaggregation,” IEEE Trans. Power Syst., vol. 35, no. 4, pp. 3133-3144, Jul. 2020.

[9] D.W. van der Meer, J. Munkhammar, J. Widén, “Probabilistic forecasting of solar power, electricity consumption and net load: Investigating the effect of seasons, aggregation and penetration on prediction intervals,” Sol. Energy, vol. 171, pp. 397-413, Sept. 2018.

[10] W. Stainsby, D. Zimmerle, G. P. Duggan, “A method to estimate residential PV generation from net-metered load data and system install date,” Appl. Energy, vol. 267, Jun. 2020.

[11] M. Sun, T. Zhang, Y. Wang, G. Strbac and C. Kang, “Using Bayesian Deep Learning to Capture Uncertainty for Residential Net Load Forecasting,” IEEE Trans. Power Syst., vol. 35, no. 1, pp. 188-201, Jan. 2020.

[12] D.W. van der Meer, M. Shapero, A. Svensson, J. Widén, J. Munkhammar, “Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes,” Appl. Energy, vol. 213, pp. 195-207, Mar. 2018.

[13] D.W. van der Meer, J. Widén, J. Munkhammar, “Review on probabilistic forecasting of photovoltaic power production and electricity consumption,” Renewable Sustain. Energy Rev., vol. 81, pt. 1, pp. 1484-1512, Jan. 2018.

[14] D.W. van der Meer, J. Munkhammar, J. Widén, “Probabilistic forecasting of solar power, electricity consumption and net load: Investigating the effect of seasons, aggregation and penetration on prediction intervals,” Sol. Energy, vol. 171, pp. 397-413, Sept. 2018.

[15] H. Shaker, H. Zareipour and D. Wood, “Estimating Power Generation of Invisible Solar Sites Using Publicly Available Data,” IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 2456-2465, Sept. 2016.

[16] H. Shaker, H. Zareipour and D. Wood, “A Data-Driven Approach for Estimating the Power Generation of Invisible Solar Sites,” IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 2466-2476, Sept. 2016.

[17] T. Hong, J. Wilson and J. Xie, “Long Term Probabilistic Load Forecasting and Normalization With Hourly Information,” IEEE Trans. Smart Grid, vol. 5, no. 1, pp. 456-462, Jan. 2014.

[18] F. Wang, K. Li, X. Wang, et al. “A Distributed PV System Capacity Estimation Approach Based on Support Vector Machine with Customer Net Load Curve Features,” Energies, vol. 11, no. 7, Jul. 2018.

[19] T. K. Chau, S. S. Yu, T. Fernando, H. H. Iu, and M. Small, “A Load-Forecasting-Based Adaptive Parameter Optimization Strategy of STATCOM Using ANNs for Enhancement of LFOD in Power Systems,” IEEE Transactions on Industrial Informatics, vol. 14, no. 6, pp. 2463-2472, Jun. 2018.

[20] E. Gonzalez-Romera, M. A. Jaramillo-Moran, and D. Carmona-Fernandez, “Monthly Electric Energy Demand Forecasting Based on Trend Extraction,” IEEE Transactions on Power Systems, vol. 21, no. 4, pp. 1946-1953, Nov. 2006.

[21] https://www.ausgrid.com.au/Common/About-us/Corporate information /Data-to-share/Solar-home-electricity-data.aspx, accessed Mar.18, 2021.

[22] G.E.P.Box, G.M. Jenkins, *Time Series Analysis*. *Forecasting and Control*. San Francisco: Holden-Day, 1976.

[23] A. M. Fathabad, J. Cheng, K. Pan and F. Qiu, "Data-Driven Planning for Renewable Distributed Generation Integration," IEEE Transactions on Power Systems, vol. 35, no. 6, pp. 4357-4368, Nov. 2020.