Very Short-Term Solar Power Forecasting Using a Frequency Incorporated Deep Learning Model

HOSSEIN PANAMTASH[1] **(Graduate Student Member, IEEE),**
SHAHRZAD MAHDAVI[1] **(Graduate Student Member, IEEE),**
QUN ZHOU SUN[1] **(Member, IEEE), GUO-JUN QI**[1] **(Fellow, IEEE), HONGRUI LIU**[2],
AND ALEKSANDAR DIMITROVSKI[1] **(Senior Member, IEEE)**

[1]Department of Electrical Engineering, University of Central Florida, Orlando, FL 32816 USA
[2]Department of Industrial and Systems Engineering, San José State University, San Jose, CA 95192 USA

CORRESPONDING AUTHOR: H. PANAMTASH (h.panamtash@Knights.ucf.edu)

ABSTRACT This paper aims to forecast solar power in very short horizons to assist in real-time distribution system operations. Popular machine learning methods for time series forecasting are studied, including recurrent neural networks with Long Short-Term Memory (LSTM). Although LSTM networks perform well in different applications by accounting for long-term dependencies, they do not consider the frequency domain patterns, especially the low frequencies in the solar power data compared to the sampling frequency. The State Frequency Memory (SFM) model in this paper extends LSTM and adds multi-frequency components into memory states to reveal a variety of frequency patterns from the data streams. To further improve the forecasting performance, the idea of Fourier Transform is integrated for optimal selection of the frequency bands by identifying the most dominant frequencies in solar power output. The results show that although the SFM model with uniform frequency selection does not significantly improve upon the LSTM model, the proper selection of frequencies yields overall better performances than the LSTM and 27% better than the persistent forecasts for forecast horizons up to one minute. Furthermore, a predictive voltage control based on solar forecasts is implemented to demonstrate the superior performance of the proposed model.

INDEX TERMS Solar forecasting, state frequency model, discrete Fourier transform (DFT), long short term memory (LSTM), very short term solar forecasting, predictive voltage control.

I. INTRODUCTION

THE power system industry has gradually shifted towards renewable generation in recent years. Specifically, the installation of solar generation units has been booming for the last decade. In the United States with the Investment Tax Credit established in 2006, an average growth rate of 50% annually is observed in the installation of solar power panels [1]. Although this rapid growth helps reduce generation cost and pollution compared to conventional fossil fuel units, the uncertainty associated with renewable generation introduces new challenges to the operation and control of power systems [2], [3]. Accurate and reliable forecasts play a key role in handling the uncertainties associated with solar generation [4]. Therefore, forecasting solar generation is vital for systems with high penetration of solar energy [5].

Solar power forecasting horizon can range from a few seconds to several days and months ahead [6], [7]. The length of the forecast horizon depends on the application of the forecast in power systems. Long-term solar forecasts are critical in power system planning and provide seasonal and yearly trends in solar power [5]. Solar forecasts with medium horizons are used for operational decisions such as unit commitment and day-ahead planning [8], [9], [10], [11]. Short-term solar forecasts are employed for system operation from minutes to a few hours ahead [12], [13], [14]. For example, short-term forecasts are utilized to cope with renewable power generation uncertainty in real-time economic dispatch [15]. Lastly, very short-term solar power forecasting ranges from a few seconds to a few minutes in terms of the horizon.

Very short-term forecasting is not given as much attention in the literature compared to longer forecast horizons. However, it can be an important asset in power system control such as voltage and frequency control applications [16], [17]. For example, solar power forecasts from 1 second up to 1 minute ahead are used to estimate the voltage changes in a power system with high penetration of PV for coordinated voltage regulator control [18], [19]. While other renewable energy resources such as wind power in general are more uncertain than solar power generation and exhibit more intermittency due to weather conditions [20], solar power is more volatile in small time windows of less than a minute. Our prior study shows that the PV power output can change more than 50% of the capacity in less than a minute [21]. Such large changes in solar power output can present a challenge for power balance and affect the voltage and frequency quality in the system. Therefore, solar forecasts with intra-minute horizons are very important, and in this study, we focus on improving very short-term solar power forecasting.

Solar forecasting models can be categorized into four groups in terms of their input type. The first group includes physical models that employ the underlying physical relationship between solar power generation and variables such as solar irradiance, zenith angle, temperature, and other weather data. These models usually depend on numerical weather predictions or satellite and sky images. Models depending on numerical weather predictions are generally used for long-term forecasts, while satellite-based models are mostly utilized in short-term forecasts [22], [23], [24]. The forecast models that rely on sky images, however, are applied to short-term and very short-term forecast horizons [25]. The second group consists of statistical models applied to historical data for generating solar power forecasts. Linear regression and Auto-Regressive Integrated Moving Average (ARIMA) models and their variations are examples of statistical models [26], [27], [28]. Statistical forecasting models are applied to intra-day and intra-hour solar forecasts for best results [6]. The third group employs machine learning methods to learn the relation between input and output variables to forecast solar power, including Random Forecasts (RF), Support Vector Regression (SVR), Gradient Boosting (GB), and Artificial Neural Networks (ANN) [24], [29]. The machine learning models typically outperform other models in short-term forecasts. The fourth group consists of the hybrid models that combine the previously mentioned models for solar power forecasting [30].

One of the important factors in model selection is the forecast horizon. For example, long-term forecasts typically rely on numerical weather predictions while short-term forecasts more commonly depend on statistical models and exogenous data such as cloud movement. However, very short-term forecasts with intra-minute horizons cannot rely as much on the exogenous data and the persistent forecast model usually performs best in such short forecast horizons [6]. The availability

of exogenous data such as sky images and satellite data is also limited in many cases.

Many studies have renewed the attention to neural networks in recent decades, with overall better performance than statistical models in solar forecasting literature [31], [32]. Although general neural network models can learn the relation between input and output arguments, they fail to capture the dynamic temporal dependencies within the data. Recurrent Neural Networks (RNNs) tackle this problem by introducing hidden states and using them alongside the input data to generate forecasts [33]. However, the hidden state in RNNs is unable to capture long-term dependencies due to vanishing gradients in back-propagation. The Long Short-Term Memory (LSTM) model first introduced in [34] is one of the commonly used solutions to this problem. The LSTM model adds several memory gates to the RNN structure which enables the model to capture long-term dependencies in the data. The LSTM is considered a state-of-the-art model and has been used in many studies for solar power forecasting [35], [36], [37]. An example of LSTM deep learning implementation is [38] which utilizes a deep LSTM-RNN network for reliable photovoltaic power forecasting. Also, there are several papers in the literature that apply ensemble-based methods to LSTM models [39]. For example, [40] utilizes an aggregation function based on the Choquet integral to improve forecast reliability by finding the largest consistency among the conflicting forecast results.

While the LSTM model is more capable in time series forecasting by preserving the trend information contained in the long-term sequences [37], it only deals with the data patterns in the time domain. The State Frequency Model (SFM) model introduced by [41] adds a decomposition step to the LSTM structure that extracts the frequency components with an additional dimension to the memory state representing different frequencies. This model has been used for stock price forecasting in [42] to capture the patterns corresponding to different frequencies in trading.

In this paper, the SFM model is applied to very short-term solar power forecasting. Surprisingly, the SFM model does not improve upon the LSTM model partly due to the fact that frequency patterns in the solar power data are much smaller compared to the sampling frequency. In order to optimize the SFM model, a Discrete Fourier Transform (DFT) is applied to the historical data, and the SFM frequency band is adjusted accordingly. This method is applied to solar power forecasts with horizons ranging from 1 second to 60 seconds. The PV data from a site in Florida, US is used to generate forecasts. The results show a significant improvement over the state-of-the-art LSTM model and the other benchmark models.

An example of a use case for intra-minute forecasts of solar power is represented in [19] for voltage regulator control. The predictive voltage regulator control is utilized in this study to show the benefits of improving the accuracy of very short-term solar forecasts. The voltage control results based on the
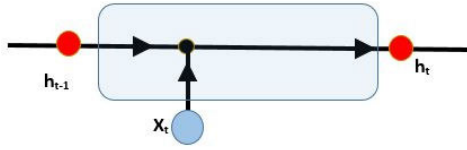
FIGURE 1. RNN structure.

FIGURE 2. LSTM structure.

proposed SFM model show a decrease in voltage deviations compared to the benchmark forecast models.

The remainder of this paper is organized as follows. Section II reviews the framework of RNN and LSTM models. Section III presents the SFM model. Section IV discusses the shortcoming of the SFM model and describes the model optimization using Discrete Fourier Transform (DFT). Section V discusses performance analysisbsphackesphack. Section VI presents the testing procedure and numerical results. Section VII provides concluding remarks and future work.

II. RELATED WORK

One of the widely used models for solar power forecasting is ANN. While feed-forward ANNs have good results in general modeling problems, they do not perform as well in time-series forecasting because the dynamics in the dataset are difficult to capture with a feed-forward structure. RNNs are designed to use past states in the network and the state-of-the-art LSTMs are successful examples of these efforts that have proved capable of improving the forecast results using long-term dependencies with data temporal dynamics in different applications.

RNNs are a class of neural networks that attempt to capture temporal dynamic patterns in the data by adding an internal state h_{t-1} to the feed-forward ANNs. The internal state acts as a memory for the model and enables the use of past states in addition to the inputs of the network at the current time. Fig. 1 presents the structure of a neuron in RNN models. An RNN model consists of successive layers that use the inputs and the state from the previous layer to generate the new state and the output.

RNNs can adapt to the dynamic patterns in the data but they fail to capture long-term dependencies due to vanishing gradients in the back-propagation process. Long Short-term Memory (LSTM) networks are one of the most effective solutions to this problem. The LSTM structure receives the input vector x_t, the memory state c_{t-1} and hidden state h_{t-1} from the previous layer then generates the memory state c_t and the hidden state h_t for time t as shown in Fig. 2 [34]. An LSTM neuron consists of a cell that acts as the memory state of the model and three gates that control the flow of information. The input gate i controls the amount of new data going into the cell, the forget gate f controls the information flow from the previous cell, and the output gate o controls the extent to which the data in the cell is used to generate the model output. The memory gates that are incorporated into
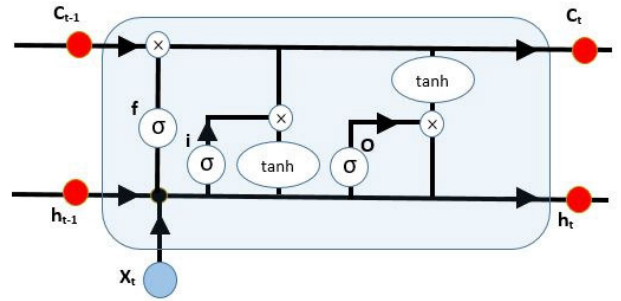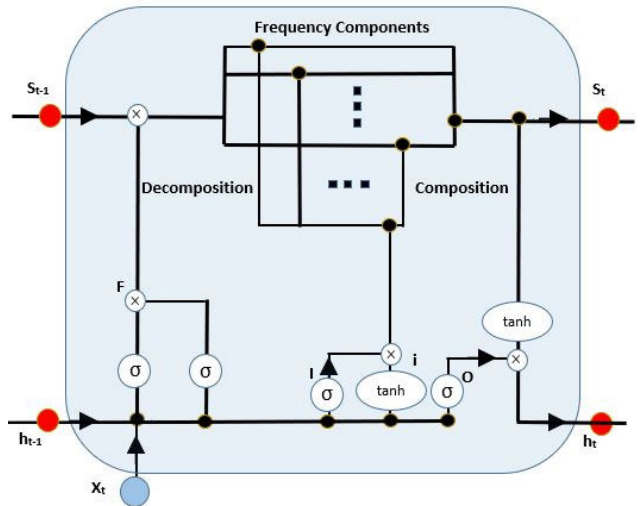

FIGURE 3. SFM structure.

the RNN structure control the flow of information and shape the cell at each layer. The LSTM network is capable of capturing long-term dependencies through the gate mechanism to resolve the vanishing gradient problem.

III. STATE FREQUENCY MEMORY (SFM)

The State Frequency Memory (SFM) model is an extension of LSTM networks that incorporates the data patterns in the frequency domain. This is particularly useful in very short-term forecasting where the patterns in the data consist of much smaller frequencies compared to the forecasting frequency. The long-term patterns (slow dynamics) in the data compared to the short-term (fast dynamics) of the forecast increases the number of memory gates that the past data has to go through. The SFM method is based on the existing LSTM model and adds the frequency patterns by integrating the Fourier transform of the inputs in shaping the output. Fig. 3 depicts the structure of an SFM cell.

The overall structure is similar to the LSTM model, but the state is extended to a matrix rather than a single value. The state matrix used in the SFM method uses a $D \times N$ matrix S corresponding to each feature and frequency. This matrix represents the state of the neural network at time t.

The matrix S is updated at each time using the previous state and the input matrix as described below [41]:

$$S_t = F_t \circ S_{t-1} + \begin{pmatrix} I_{t,1} i_{t,1} \\ I_{t,2} i_{t,2} \\ \vdots \\ I_{t,D} i_{t,D} \end{pmatrix} \begin{pmatrix} e^{j\omega_1 t} & e^{j\omega_2 t} & \dots & e^{j\omega_N t} \end{pmatrix} \tag{1}$$

where S_t represents the state of the network at time t, F_t is the forget gate matrix at time t, $I_{t,1}, I_{t,2}, \dots, I_{t,D}$ are the arguments from the input gate corresponding to each dimension, $i_{t,1}, i_{t,2}, \dots, i_{t,D}$ are the input modulations which is an aggregate of the current inputs and the last output as shown in Eq. (6), $e^{j\omega_1 t}, e^{j\omega_2 t}, \dots, e^{j\omega_N t}$ are the components for the Fourier transform into N frequencies and $\circ$ denotes element-wise multiplication. Each gate is described in detail below.

Forget Gate: The forget gate is a $D \times N$ matrix responsible for the decision that how much of the data from the past state is used in forming the current state. The forget gate F_t is defined as the outer product of two forget gate vectors:

$$F_t = F_t^S \times F_t^F \tag{2}$$

where F_t^S is the state forget gate vector and F_t^F is the frequency forget gate vector.

The state forget gate F_t^S decides how much information from each dimension of the past state is used in shaping the current state. F_t^S is defined as an element-wise sigmoid applied to the linear combination of the current input and the past output:

$$F_t^S = \sigma(\begin{pmatrix} w_1^{Sx} \\ w_2^{Sx} \\ \vdots \\ w_D^{Sx} \end{pmatrix} \circ \begin{pmatrix} x_{t,1} \\ x_{t,2} \\ \vdots \\ x_{t,D} \end{pmatrix} + \begin{pmatrix} w_1^{Sh} \\ w_2^{Sh} \\ \vdots \\ w_D^{Sh} \end{pmatrix} \circ \begin{pmatrix} h_{t-1,1} \\ h_{t-1,2} \\ \vdots \\ h_{t-1,D} \end{pmatrix} + \begin{pmatrix} b_1^S \\ b_2^S \\ \vdots \\ b_D^S \end{pmatrix}) \tag{3}$$

where σ is an element-wise sigmoid function, W^{Sx} and W^{Sh} are input and output weight vectors, x_t is the D dimensional input vector at time t, h_{t-1} is the output vector at time $t-1$, and b^S is a bias argument vector.

The frequency forget gate decides how much information from each frequency of the past state is used in shaping the current state. The frequency forget gate is similarly defined as below:

$$F_t^F = \sigma(\begin{pmatrix} w_1^{Fx} \\ w_2^{Fx} \\ \vdots \\ w_D^{Fx} \end{pmatrix} \circ \begin{pmatrix} x_{t,1} \\ x_{t,2} \\ \vdots \\ x_{t,D} \end{pmatrix} + \begin{pmatrix} w_1^{Fh} \\ w_2^{Fh} \\ \vdots \\ w_D^{Fh} \end{pmatrix} \circ \begin{pmatrix} h_{t-1,1} \\ h_{t-1,2} \\ \vdots \\ h_{t-1,D} \end{pmatrix} + \begin{pmatrix} b_1^F \\ b_2^F \\ \vdots \\ b_D^F \end{pmatrix}) \tag{4}$$

where W^{Fx} and W^{Fh} are input and output weight vectors, and b^F is a bias argument vector.

Input Gate: The input gate is used to determine how much of the input data from the current time is used in shaping the

current state in the SFM method. The input gate I_t used in Eq. (1) is defined as below:

$$I_t = \sigma(\begin{pmatrix} w_1^{Ix} \\ w_2^{Ix} \\ \vdots \\ w_D^{Ix} \end{pmatrix} \circ \begin{pmatrix} x_{t,1} \\ x_{t,2} \\ \vdots \\ x_{t,D} \end{pmatrix} + \begin{pmatrix} w_1^{Ih} \\ w_2^{Ih} \\ \vdots \\ w_D^{Ih} \end{pmatrix} \circ \begin{pmatrix} h_{t-1,1} \\ h_{t-1,2} \\ \vdots \\ h_{t-1,D} \end{pmatrix} + \begin{pmatrix} b_1^I \\ b_2^I \\ \vdots \\ b_D^I \end{pmatrix}) \tag{5}$$

where I_t is the input gate, W^{Ix} and W^{Ih} are input and output weight vectors, and b^I is a bias argument vector.

The input vector is shaped by the application of a hyperbolic tangent to the weighted combination of the input arguments and the previous output:

$$i_t = tanh(\begin{pmatrix} w_1^{ix} \\ w_2^{ix} \\ \vdots \\ w_D^{ix} \end{pmatrix} \circ \begin{pmatrix} x_{t,1} \\ x_{t,2} \\ \vdots \\ x_{t,D} \end{pmatrix} + \begin{pmatrix} w_1^{ih} \\ w_2^{ih} \\ \vdots \\ w_D^{ih} \end{pmatrix} \circ \begin{pmatrix} h_{t-1,1} \\ h_{t-1,2} \\ \vdots \\ h_{t-1,D} \end{pmatrix} + \begin{pmatrix} b_1^i \\ b_2^i \\ \vdots \\ b_D^i \end{pmatrix}) \tag{6}$$

where i_t is the input vector, W^{ix} and W^{ih} are input and output weight vectors, and b^i is the bias argument vector.

Output Gate: The output gate decides how much of the data from the current state is used for each frequency to calculate the final output of the model. The output gate is a linear combination of the current state, the previous output, and the current input vector. Similar to the input gate, the output gate is defined below:

$$O_t = \sigma(W_n^O |S_{t,n}| + W_{x,n}^O x_t + W_{h,n}^O h_{t-1} + b_n^O) \tag{7}$$

where O_t is the output gate, W_n^O, $W_{x,n}^O$ and $W_{h,n}^O$ are set of weights, and b_n^O is a bias vector. Note that $|S_{t,n}|$ is the amplitude of the nth frequency component in the state matrix.

The final output of the SFM model is derived from the sum of outputs corresponding to each frequency. Furthermore, the output for each frequency is obtained using an output gate and an activation function applied to a linear function of the amplitude of the state matrix for that frequency [41]:

$$h_t = \sum_{n=1}^{N} O_{t,n} \circ f_o(w_n^o |S_{t,n}| + b_n^o) \tag{8}$$

where h_t is the output vector at time t, $O_{t,n}$ is the output gate co-responding to the nth frequency at time t, f_o is the output activation function, w_n^o is a vector of weights, $S_{t,n}$ is the amplitude of the nth column of the state matrix and b_n^o is a vector of biases for time t and the nth frequency.

IV. OPTIMIZING SFM USING DISCRETE FOURIER TRANSFORM

Although the original SFM model can capture the frequency patterns in the dataset, the frequencies are uniformly included from ω_1 to ω_N. In this paper, we propose to select frequencies that are most important to solar power forecasting. Although including all frequencies may help, it also complicates the
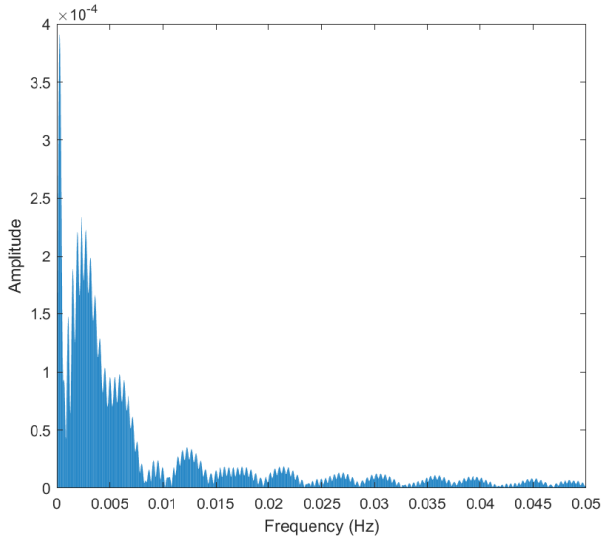
FIGURE 4. The DFT of historical data from the dataset.

model by adding correlated data corresponding to different frequencies. Specifically, in very short-term solar power forecasting, the frequency elements must reflect the low rate of change in the solar power output compared to the fast rate of sampling (every second). If the frequencies are not selected properly, the SFM model can end up complicating the LSTM without improvement in the accuracy of the forecast. Therefore, the paper proposes to select the most important frequencies based on the amplitude in Discrete Fourier Transform (DFT). The DFT of a vector consisting of N numbers $(x_0, x_1, \ldots, x_{N-1})$ which are sampled in equal spaces in the time domain converts them to another vector of complex numbers $(X_k = X_0, X_1, \ldots, X_{N-1})$ which represent the data in the frequency domain. The DFT is defined as below:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{\frac{-j2\pi}{N} kn} \qquad (9)$$

Using Euler's formula the DFT can be expressed as:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot [cos(\frac{2\pi}{N} kn) + j \cdot sin(\frac{2\pi}{N} kn)] \qquad (10)$$

The application of DFT can guide the frequency selection towards bandwidths with the most prominent frequencies. Fig. 4 shows the DFT result of the training data from the dataset under study in section VI-A. The figure only shows the frequency components under 0.05 Hz so that the largest frequency components are visible. The frequency components associated with frequencies higher than 0.05 Hz pose very low amplitudes and therefore are neglected.

It is visible that most of the high amplitudes appear in very low frequencies compared to the sampling frequency (1 second). The fact that solar power does not change dramatically in very short periods makes that obvious but the DFT can help select frequency bands.
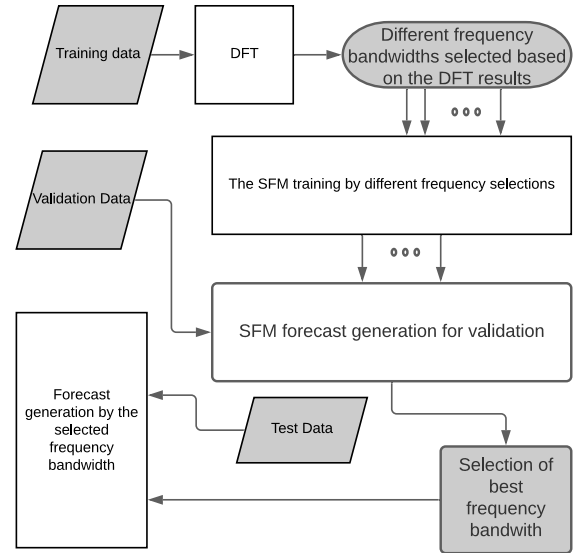

FIGURE 5. The frequency selection flowchart.

A flowchart is devised to select the optimal frequency band and the number of frequencies as shown in Fig. 5. It involves the following steps:

- The DFT is applied to historical data to extract the amplitude corresponding to different frequencies in the data.
- Different frequency bands are selected based on the areas with the highest amplitude in the frequency domain.
- An SFM model is trained based on each frequency band selection with a different number of frequencies.
- The models based on different frequency selections are used for forecasting the validation data and the optimal model is selected based on the results.
- The selected SFM model is applied to the test dataset to generate the final forecast results.

V. PERFORMANCE ANALYSIS

Generally, there are two approaches for comparison: a direct measure of forecast accuracy and a measure of benefit in the context of an application utilizing the forecasts. For the second approach, we investigate the benefit of very short-term solar power forecasting in predictive voltage control in a distribution network. The two approaches are described in detail below.

For the first approach, performance metrics are introduced to compare the forecasts generated using the proposed SFM-DFT model against benchmark models. We consider the accuracy of the forecasts as the performance metric utilizing the Root Mean Square Error (RMSE) compared to the actual solar power output, as described below:

$$RMSE = \sqrt{\frac{\sum_{t=t_1}^{t_2} (y_t - \hat{y}_t)^2}{n}} \qquad (11)$$

where y_t is the actual output, $\hat{y}_t$ is the forecast for time t and n is the number of data points between t_1 and t_2. The method with the smallest RMSE is the most accurate and therefore the best-performing method.

The second approach applies the forecasts acquired from different models to a predictive voltage regulator control algorithm to demonstrate the benefits brought by improved forecasts. The predictive voltage control algorithm is initially proposed in our previous study [19]. First, we generate solar power forecasts for the next 60 seconds. Then, the changes in the power are used to predict the voltage changes throughout a distribution system. Given the impedance matrix of the system, we can find the voltage changes for each node based on the changes in solar power injections. The average predicted voltage changes for the next 60 seconds are then used to determine voltage deviations from the nominal voltage under each voltage regulator.

Based on the predicted voltage deviations, the tap changes of each regulator are decided by prioritizing the voltage regulator with the most severe voltage difference. The signals are then sent to the voltage regulator to increase or decrease the taps if needed. The control signals are applied to a time delay intending to prevent excessive tap changes. More details can be found in [19].

The voltage deviation in the system is calculated using the Voltage Deviation Index:

$$VDI = \sum_{i=1}^{n} |1 - V_i|, \quad if \ |1 - V_i| > 0.01 p.u. \quad (12)$$

The VDI results are used to compare the effectiveness of different forecast models when applied to the predictive voltage control algorithm.

VI. RESULTS ANALYSIS
In this section, the results of the SFM model with and without frequency selection are compared to other models to show the effectiveness of the proposed model in this study.

A. FORECASTING SETTINGS
The dataset used in this study is the solar generation data from a 50kW PV site at the University of Central Florida operated by Duke Energy. The dataset contains the solar generation for 8 days (from 17th to 24th of July 2017). The data points are in 1-second intervals. The power output from the dataset is shown for one day in Fig. 6. Zooming in on the data shows the drastic changes within a small time window. The figure shows that in a 3-minute window the solar power output drops more than 20 kW and rises back up again. The sudden drops in solar power output are due to the rapid changes in cloud coverage. However, compared with the sampling frequency of one second, this change at the minute level is considered a slow dynamic.

Several models have been selected in this study as benchmarks to compare the results against the proposed model. The LSTM model and the original SFM model with uniform
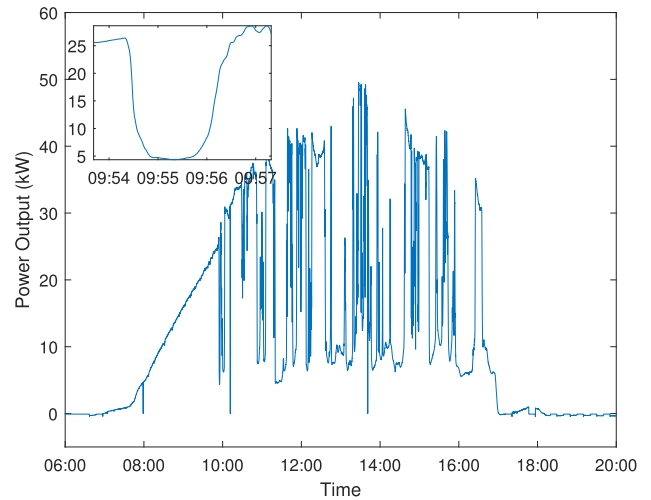

FIGURE 6. The solar power generation on 07/17/2017.

frequency selection are used to compare against the SFM model with frequency selection. The persistent model that normally performs well in the very short forecasting horizon is also selected as a benchmark. Since the forecast horizon is very small, achieving a better performance than the persistent model is extremely challenging. Data from different days and different times of the day are selected for this purpose. The first 80% of the data in each window is used for training and the remainder has been divided equally between validation and testing sets, so 10% of the selected data is dedicated to each one. The data windows are selected throughout the dataset to represent different patterns in the data without overlapping.

B. FORECAST RESULT COMPARISON
The RMSE is calculated for each model during the training, validation, and testing stages to ensure optimized model configuration. The training, validation, and test errors in different numbers of iterations for the SFM model with frequency selection are shown in Fig for one of the time windows selected. 7. We continue the iterations up to 4000 and calculate the error at 20 intervals. While the training error decreases by increasing the number of iterations, the validation error may be minimized at a lower iteration level. For example, the minimum validation and test error for an example time window happen at 280 iterations as shown in Fig. 7.

The optimal number of inputs to the forecast model (past instances of solar power) is selected based on the performance of the validation set.

There are several factors in the forecasting problem that can affect the performance of the model. To comprehensively analyze the forecast accuracy, we divide the result evaluation into three sections. In the first section, different frequency selections are examined. In the second section, different times of the day with different solar patterns alongside the overall
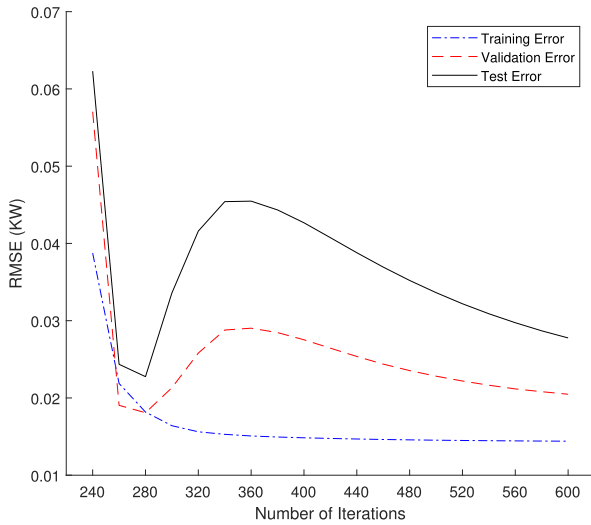
FIGURE 7. Training, validation and test error comparison.

FIGURE 8. Results of solar forecasting in the morning.

TABLE 1. The RMSE results for different frequencies of the SFM model.

Number of frequencies	Frequency area (Hz)		
	(0:1)	(0:0.1)	(0:0.01)
1	0.026792	0.026973	0.022758
2	0.027541	0.027842	0.022687
3	0.028145	0.026690	0.02414
4	0.027453	0.02738	0.02310
5	0.027660	0.028264	0.02371

results for an entire day are compared. The third section compares the accuracy for longer forecasting horizons.

1) FREQUENCY SELECTION

The number of frequencies and the frequency band that are selected for the SFM model are optimized for the best results based on the DFT of the training data. The frequency selection starts with a wide selection of frequency areas and gradually narrows down. Next, an SFM model is developed based on each of the frequency settings. The optimal area and number for the frequencies are determined by applying each individual model to predict the validation set data. Then the accuracy results from the different forecasts are compared to find the frequencies associated with the best forecast results.

Table 1 shows the results of the SFM using a different number of frequencies and different selected areas between zero and sampling frequency. It is visible that while the number of frequencies selected has an impact on the results, the most important factor affecting accuracy is the area from which the frequencies are selected. Based on the DFT of the data, the frequencies with the most amplitudes are in the very low-frequency range. The results with the frequencies selected from the low bandwidths are more accurate compared to the others.

The optimal number of frequencies selected for the SFM model in most cases is just one frequency, and in other cases,
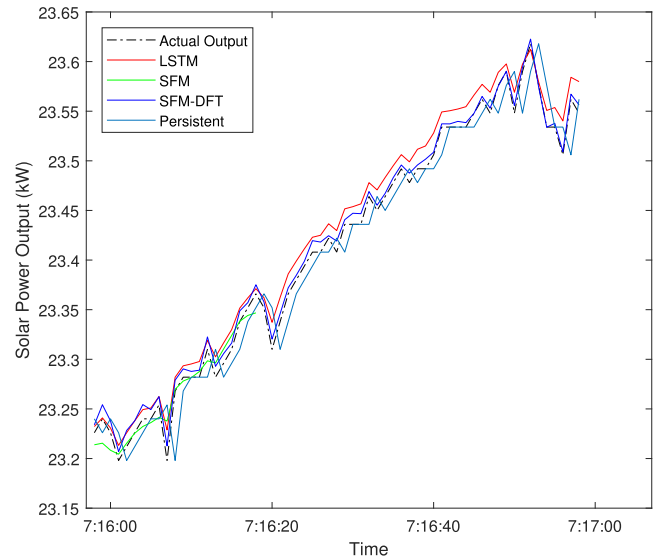
the difference in the accuracy between one frequency and the optimal model is negligible. Since increasing the number of frequencies results in a more complicated model and increases the training time for the SFM model, we select only one frequency for this dataset. The optimal frequencies based on the Fourier transform and the results from the validation sets for different windows in the dataset are between 0 and 0.01 times the sampling frequency (1 Hz).

Note that the frequency area selection does not affect the model training time, but an additional number of frequency selections will increase the training time for the SFM model. Therefore when the accuracy of the model doesn't improve noticeably with a higher number of frequencies, the number of frequencies selected for the model is not increased. For a 45-minute time period of training window, the average training time for linear models is about 0.3 seconds while the LSTM model takes 352 seconds for training. For the SFM model, it depends on the number of frequencies selected. The SFM training time increases by around 10% for each additional frequency selected. For example, the average training time of the SFM model with one frequency is 584 seconds, and using two frequencies the time is increased to 662 seconds. The model training is conducted offline and does not affect the online forecasting time. The forecast time for all the models is at millisecond-level and hence all models can produce forecasts in a timely manner.

2) DIFFERENT TIMES OF THE DAY

The frequency patterns in the data can differ based on the time of the day and the irradiance level. Therefore the forecast results are compared in time windows during different periods of the day. Fig. 8 presents the forecast results for LSTM, original SFM, and SFM with DFT compared to the actual output in the morning.

TABLE 2. The RMSE comparison between Different models (1-step ahead forecasting) in the morning.

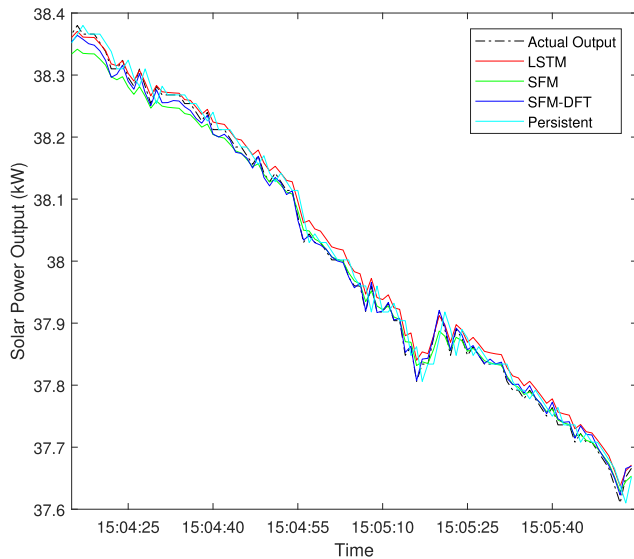| Model	SFM with DFT	SFM	LSTM	Persistent
RMSE (KW)	0.0105	0.0135	0.0117	0.0186

FIGURE 9. Results of solar forecasting in the afternoon.

TABLE 3. The RMSE comparison between Different models (1-step ahead forecasting) in the afternoon.

Model	SFM with DFT	SFM	LSTM	Persistent
RMSE (KW)	0.0250	0.0595	0.0303	0.0223

The performance comparison between the different forecasting models in terms of RMSE is presented in Table 2.

The result shows that all the models with memory states performed better than the Persistent forecast. Although the SFM model was not able to improve the LSTM model, the application of DFT and optimized frequency selection resulted in the best performance for the SFM model with DFT. The SFM with DFT model during the morning session performs 43% better than the persistent model, 10% better than the LSTM model, and 22% better than the SFM model.

The comparison between forecast models during a time window in the afternoon is shown in Fig. 9. The figure depicts the solar power output and forecasts from different models. The performance comparison between the different forecast models in terms of RMSE is presented in Table 3.

Given the very low volatility during this 1-minute period in the afternoon, the persistent model generates the best results. The SFM model without frequency selection has the worst performance among the forecast models. Although the SFM model with frequency selection did not perform as well as the persistent model during this particular time window, the results are much improved compared to LSTM.

TABLE 4. The RMSE comparison between Different models (1-step ahead forecasting).

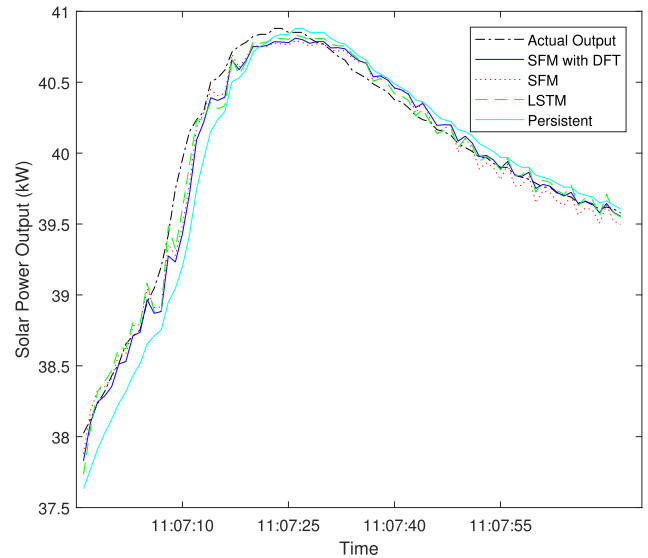| Model	SFM with DFT	SFM	LSTM	Persistent
RMSE (KW)	0.0228	0.0268	0.0266	0.0245

FIGURE 10. Results of SFM with DFT compared to SFM model (3-step ahead forecasting).

Despite a slight under-performance compared to the persistent model in the afternoon, SFM with DFT has an overall improvement in the entire day period. The aggregate results of forecast models during the entire day are presented in Table 4.

Overall the comparison between the results of forecast models during the day shows that the SFM model with frequency selection based on DFT is the only model that outperforms the persistent forecast. While the SFM model performed worse than LSTM, frequency selection based on DFT helped in capturing the underlying frequency patterns in the data and improving the LSTM model. The proposed SFM model with DFT compared to the benchmark models Persistent, LSTM, and SFM show 6.94%, 14.29%, and 14.92% improvement respectively.
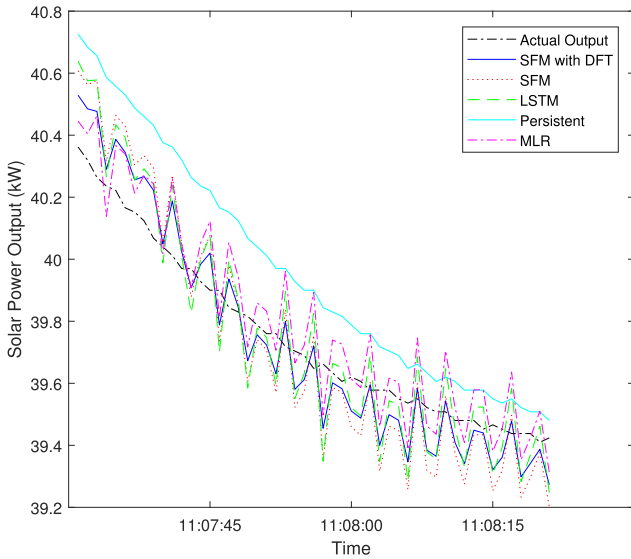
3) FORECAST HORIZON

The forecast horizon is another factor that can affect the performance of different models. To show the effectiveness of the proposed model in longer forecast horizons, the SFM model with frequency selection based on DFT is compared to the benchmark models for 3-second ahead forecasting. Fig. 10 shows the results of the SFM forecast model with and without DFT alongside the actual solar power output and the persistent forecast.

Understandably, the accuracy of the persistent forecast method is affected by increasing the forecast horizon. In the 3-step ahead forecasting, the SFM model outperforms the persistent model noticeably, especially during large changes

TABLE 5. The RMSE results of SFM with DFT compared to SFM and persistent models (3-step ahead forecasting).

| Forecast Model	Persistent	LSTM	SFM	SFM with DFT
RMSE (KW)	0.143209	0.105961	0.113703	0.087254

FIGURE 11. Results of SFM with DFT compared to other forecast models (10 seconds ahead forecasting).

TABLE 6. The RMSE results of SFM with DFT compared to other benchmark forecasts for longer horizons.

Forecast Model	Persistent	ANN	MLR	LSTM	SFM	SFM-DFT
10-sec horizon	0.4183	0.3039	0.3186	0.2790	0.2874	0.2443
30-sec horizon	1.0546	1.0206	1.0090	0.9411	0.9267	0.8766
60-sec horizon	1.5168	1.1836	1.3035	1.0568	1.3874	1.0405

in solar power. The SFM model with the application of DFT however improves the SFM model and has the best forecast accuracy. The RMSE comparison between the three methods is presented in Table 5. The overall results show the SFM model with DFT performs 23% better than the SFM model and 39% better than the persistent model.

Furthermore, the solar power forecasts are performed for longer horizons to show the extent of effectiveness of the proposed SFM model with DFT. We generate forecasts for 10-second, 30-second, and 60-second horizons. The results of 10-second ahead forecasts are shown in Fig. 11. In addition, two more benchmark models that are commonly used in the forecasting literature are included: An ordinary Artificial Neural Network (ANN) and a Multiple Linear Regression (MLR) model. The results of the forecasts with longer horizons are presented in Table 6.

The results of longer forecast horizons show that the SFM-DFT outperforms the other benchmark forecasts up to 60 seconds. The improvement of the SFM-DFT model compared to the persistent benchmark is 41.6%, 16.88%, and 31.4% for 10-second, 30-second, and 60-second forecast horizons respectively. Nevertheless, for forecast horizons
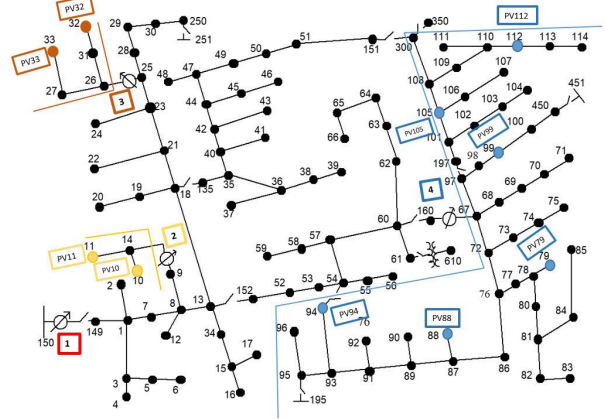
FIGURE 12. The IEEE 123 feeder with the location of voltage regulators and PV stations.

TABLE 7. The VDI results for voltage control using different forecast models.

Forecast Model	ANN	MLR	LSTM	SFM	SFM-DFT
average VDI (V)	1.114	2.040	1.514	0.981	0.2219

longer than 60 seconds, the performance of the proposed model is on par with the LSTM model. Therefore, the longer training time for the SFM model with DFT does not constitute selecting it over the LSTM model for horizons longer than 60 seconds.

While training the SFM model based on historical data takes time, since the training process is implemented offline and the forecast generation is very fast (within a few milliseconds), the real-time application of the proposed forecast model for model predictive control in the power system is possible.

We should also make a note that the proposed SFM model with frequency band selection does not outperform the LSTM model by a significant margin for forecast horizons of several minutes or longer. Therefore, this model is not recommended for longer forecast horizons.

C. PREDICTIVE VOLTAGE CONTROL WITH FORECASTS

he predictive voltage regulator control described in section V is applied to the IEEE 123 node test feeder with 4 voltage regulators. There are 10 PV stations positioned throughout the system. The solar power forecasts are generated using the SFM-DFT model as well as the benchmark models to compare the overall voltage profiles. The test feeder including the locations of voltage regulators and PV stations is shown in Fig. 12. The data from the PV site is scaled to achieve 100% penetration in the test system.

The VDI for the system is calculated under the predictive voltage regulator control using different forecasts of solar power output. The results are presented in Table 7. The persistent forecast is not included since it does not predict any changes in the solar output.

The VDI results from control algorithms based on different forecasts show significant improvement in the case of the proposed SFM-DFT forecast model.

VII. CONCLUSION

This paper applies an SFM recurrent neural network to the problem of very short-term solar forecasting. It is found that the SFM model which decomposes the memory state in the LSTM construction to different frequencies was not able to outperform the LSTM model due to low frequencies in solar power data compared to the high sampling frequency. However, the SFM model with the optimal selection of frequencies based on the Fourier transform can capture the different frequency patterns in this dataset and perform dramatically better than the LSTM model. The simulation results for an entire day show an improvement of 6.94% compared to the persistent forecast in 1-second ahead forecasting and a significant improvement of 39% for 3-second ahead forecasting. Additionally, forecasts up to 60 seconds ahead are shown to be more accurate than the benchmark models. Furthermore, the proposed forecast model is applied to a predictive voltage regulator control algorithm and results in better voltage profiles compared to the benchmarks, proving the benefit of accuracy improvement in very short horizons. The future work will include exogenous data such as sky images and irradiance to improve the forecast accuracy in longer forecast horizons.

REFERENCES

[1] *Solar Industry Research Data*. Accessed: Apr. 14, 2019. [Online]. Available: https://www.seia.org/solar-industry-research-data

[2] T. Mai et al., "Renewable electricity futures for the United States," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 372–378, Apr. 2014.

[3] X. Liang, "Emerging power quality challenges due to integration of renewable energy sources," *IEEE Trans. Ind. Appl.*, vol. 53, no. 2, pp. 855–866, Mar. 2017.

[4] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access J. Power Energy*, vol. 7, pp. 376–388, 2020.

[5] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin, and Z. Hu, "Photovoltaic and solar power forecasting for smart grid energy management," *CSEE J. Power Energy Syst.*, vol. 1, no. 4, pp. 38–46, Dec. 2015.

[6] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renew. Sustain. Energy Rev.*, vol. 27, pp. 65–76, Nov. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032113004334

[7] C. Voyant et al., "Machine learning methods for solar radiation forecasting: A review," *Renew. Energy*, vol. 105, pp. 569–582, May 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0960148116311648

[8] S. Córdova, H. Rudnick, Á. Lorca, and V. Martínez, "An efficient forecasting-optimization scheme for the intraday unit commitment process under significant wind and solar power," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1899–1909, Oct. 2018.

[9] H. Sangrody, N. Zhou, and Z. Zhang, "Similarity-based models for day-ahead solar PV generation forecasting," *IEEE Access*, vol. 8, pp. 104469–104478, 2020.

[10] X. Zhang, Y. Li, S. Lu, H. F. Hamann, B.-M. Hodge, and B. Lehman, "A solar time based analog ensemble method for regional solar power forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 268–279, Jan. 2019.

[11] Y. S. Manjili, R. Vega, and M. M. Jamshidi, "Data-analytic-based adaptive solar energy forecasting framework," *IEEE Syst. J.*, vol. 12, no. 1, pp. 285–296, Mar. 2018.

[12] X. Huang, J. Shi, B. Gao, Y. Tai, Z. Chen, and J. Zhang, "Forecasting hourly solar irradiance using hybrid wavelet transformation and Elman model in smart grid," *IEEE Access*, vol. 7, pp. 139909–139923, 2019.

[13] L. Cheng, H. Zang, Z. Wei, T. Ding, R. Xu, and G. Sun, "Short-term solar power prediction learning directly from satellite images with regions of interest," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 629–639, Jan. 2022.

[14] C. Feng, M. Cui, B.-M. Hodge, S. Lu, H. F. Hamann, and J. Zhang, "Unsupervised clustering-based short-term solar forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 4, pp. 2174–2185, Oct. 2019.

[15] S. Surender Reddy, P. R. Bijwe, and A. R. Abhyankar, "Real-time economic dispatch considering renewable power generation variability and uncertainty over scheduling period," *IEEE Syst. J.*, vol. 9, no. 4, pp. 1440–1451, Dec. 2015.

[16] Y. Ding, X. Cheng, F. Cui, X. Zhu, and H. Zhou, "Very-short term forecast of global horizontal irradiance based on ground-based sky imager and lifted condensation level calculation," in *Proc. China Int. Conf. Electr. Distrib. (CICED)*, Sep. 2014, pp. 973–978.

[17] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation—With application to solar energy," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850–3863, Sep. 2016.

[18] S. Mahdavi, H. Panamtash, A. Dimitrovski, and Q. Zhou, "Predictive and cooperative voltage control with probabilistic load and solar generation forecasting," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Aug. 2020, pp. 1–6.

[19] S. Mahdavi, H. Panamtash, A. Dimitrovski, and Q. Zhou, "Predictive coordinated and cooperative voltage control for systems with high penetration of PV," *IEEE Trans. Ind. Appl.*, vol. 57, no. 3, pp. 2212–2222, May 2021.

[20] K. Zou, A. P. Agalgaonkar, K. M. Muttaqi, and S. Perera, "Distribution system planning with incorporating DG reactive capability and system uncertainties," *IEEE Trans. Sustain. Energy*, vol. 3, no. 1, pp. 112–123, Jan. 2012.

[21] H. Panamtash, Q. Z. Sun, R. York, P. Brooker, and J. Kramer, "Solar power smoothing in a nanogrid testbed," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Exposit.*, Apr. 2022, pp. 1–5.

[22] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1571–1580, Oct. 2017.

[23] M. Kakimoto, Y. Endoh, H. Shin, R. Ikeda, and H. Kusaka, "Probabilistic solar irradiance forecasting by conditioning joint probability method and its application to electric power trading," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 983–993, Apr. 2019.

[24] H. S. Jang, K. Y. Bae, H.-S. Park, and D. K. Sung, "Solar power prediction based on satellite images and support vector machine," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1255–1263, Jul. 2016.

[25] Z. Zhen et al., "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. 2019.

[26] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, and S. Wang, "Short-term solar power forecasting based on weighted Gaussian process regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 300–308, Jan. 2018.

[27] D. van der Meer, G. R. C. Mouli, G. M.-E. Mouli, L. R. Elizondo, and P. Bauer, "Energy management system with PV power forecast to optimally charge EVs at the workplace," *IEEE Trans. Ind. Informat.*, vol. 14, no. 1, pp. 311–320, Jan. 2018.

[28] H. Panamtash and Q. Zhou, "Coherent probabilistic solar power forecasting," in *Proc. IEEE Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Jun. 2018, pp. 1–6.

[29] H. Panamtash, Q. Zhou, T. Hong, Z. Qu, and K. O. Davis, "A copula-based Bayesian method for probabilistic solar power forecasting," *Sol. Energy*, vol. 196, pp. 336–345, Jan. 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038092X1931179X

[30] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, "Solar forecasting methods for renewable energy integration," *Prog. Energy Combustion Sci.*, vol. 39, no. 6, pp. 535–576, Dec. 2013.

[31] G. Perveen, M. Rizwan, and N. Goel, "Comparison of intelligent modelling techniques for forecasting solar energy and its application in solar PV based energy system," *IET Energy Syst. Integr.*, vol. 1, no. 1, pp. 34–51, Mar. 2019.

[32] L. C. Parra Raffán, A. Romero, and M. Martinez, "Solar energy production forecasting through artificial neuronal networks, considering the Föhn, north and south winds in San Juan, Argentina," *J. Eng.*, vol. 2019, no. 18, pp. 4824–4829, Jul. 2019.

[33] S. M. Awan, Z. A. Khan, and M. Aslam, "Solar generation forecasting by recurrent neural networks optimized by Levenberg–Marquardt algorithm," in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2018, pp. 276–281.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[35] Y.-Y. Hong, J. J. F. Martinez, and A. C. Fajardo, "Day-ahead solar irradiation forecasting utilizing Gramian angular field and convolutional long short-term memory," *IEEE Access*, vol. 8, pp. 18741–18753, 2020.

[36] Y. Wang, Y. Shen, S. Mao, X. Chen, and H. Zou, "LASSO and LSTM integrated temporal model for short-term solar intensity forecasting," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2933–2944, Apr. 2019.

[37] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019.

[38] M. Abdel-Nasser and K. Mahmoud. (Oct. 2017). *Accurate Photovoltaic Power Forecasting Models Using Deep LSTM-RNN—Neural Computing and Applications*. [Online]. Available: https://link.springer.com/article/10.1007/s00521-017-3225-z

[39] A. M. Pirbazari, E. Sharma, A. Chakravorty, W. Elmenreich, and C. Rong, "An ensemble approach for multi-step ahead energy forecasting of household communities," *IEEE Access*, vol. 9, pp. 36218–36240, 2021.

[40] M. Abdel-Nasser, K. Mahmoud, and M. Lehtonen, "Reliable solar irradiance forecasting approach based on Choquet integral and deep LSTMs," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1873–1881, Mar. 2021.

[41] H. Hu and G. Qi, "State-frequency memory recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, vol. 70, 2017, pp. 1568–1577.

[42] L. Zhang, C. Aggarwal, and G.-J. Qi, "Stock price prediction via discovering multi-frequency trading patterns," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2017, pp. 2141–2149, doi: 10.1145/3097983.3098117.

HOSSEIN PANAMTASH (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Tehran and the M.Sc. degree in electrical engineering from the University of Central Florida, Orlando, FL, USA, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering. His research interests include solar power forecasting, renewable integration, microgrid control, and optimization.

SHAHRZAD MAHDAVI (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, and the M.Sc. degree in electrical engineering from the University of Central Florida, Orlando, FL, USA, where she is currently pursuing the Ph.D. degree in electrical engineering. Her research interests include renewable energy integration, distribution system modeling, and power system control.

QUN ZHOU SUN (Member, IEEE) received the Ph.D. degree in electrical engineering from Iowa State University. She is currently an Assistant Professor with the University of Central Florida (UCF). She is also the Director of the UCF Smart Infrastructure Data Analytics Laboratory. Before joining UCF, she was with Genscape Inc. and GE Grid Solutions as a Power System Engineer. She is dedicated to research that improves energy sustainability, resiliency, and security. She focuses on grid-edge resources, including smart buildings, rooftop PVs, and batteries, and their interactions with the grid.

GUO-JUN QI (Fellow, IEEE) received the Ph.D. degree from the Beckman Institute and the Department of Electrical and Computer Engineering, University of Illinois at Urbana–Champaign. He is currently an Assistant Professor with the Department of Computer Science, University of Central Florida. His current research interests include machine learning and data mining, as well as developing reliable information systems.

HONGRUI LIU received the Ph.D. degree in industrial and systems engineering from the University of Washington, Seattle, WA, USA, in 2010. She is currently an Assistant Professor in industrial and systems engineering with San José State University, San Jose, CA, USA. Her primary research interests include optimization modeling, computing algorithms, data analytics, machine learning, and artificial intelligence and their applications in the supply chain and energy industry.

ALEKSANDAR DIMITROVSKI (Senior Member, IEEE) received the B.Sc. degree in electrical engineering with an emphasis in power, the M.Sc. degree in applied computer sciences, and the Ph.D. degree in electrical engineering with an emphasis in power from Europe. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Central Florida (UCF). Before joining UCF, he had worked in research institutions, the power industry, and academia, both in the U.S. and Europe. His research interests include modeling, analysis, protection, and control of uncertain power systems.

• • •