Spatiotemporal Graph Neural Network for Performance Prediction of Photovoltaic Power Systems

Ahmad Maroof Karimi[1,2,4], **Yinghui Wu**[2], **Mehmet Koyuturk**[1,2], **Roger H. French**[1,2,3]
[1]SDLE Research Center, Case Western Reserve University, Cleveland, OH
[2]Department of Computer and Data Sciences, Case Western Reserve University, Cleveland, OH
[3]Department of Materials Science and Engineering, Case Western Reserve University, Cleveland, OH
[4]Oak Ridge National Lab, Oak Ridge, TN
{ahmadmaroof.karimi, yinghui.wu2, mehmet.koyuturk, roger.french}@case.edu

Abstract

In recent years, a large number of photovoltaic (PV) systems have been added to the electrical grid as well as installed as off-grid systems. The trend suggests that the deployment of PV systems will continue to rise in the future. Thus, accurate forecasting of PV performance is critical for the reliability of PV systems. Due to the complex non-linear variability in power output of the PV systems, forecasting PV power is a non-trivial task. This variability affects the stability and planning of a power system network, and accurate forecasting of the performance of the PV system can reduce the uncertainty caused during PV operation. In this work, we leverage spatial and temporal coherence among the power plants for PV power forecasting. Our approach is motivated by the observation that power plants in a region undergo similar environmental exposure. Thus, one power plant's performance can help improve the forecast of other power plants' power values in the region. We utilize the relationship between PV plants to build a spatiotemporal graph neural network (st-GNN) and train machine learning models to forecast the PV power. The computational experiments on large-scale data from a network of 316 systems show that spatiotemporal forecasting of PV power performs significantly better than a model that only applies temporal convolution to isolated systems or nodes. Furthermore, the longer the future forecast time, the difference between the spatiotemporal forecasting and the isolated system forecast when only temporal convolution is applied increases further.

Introduction

Due to the rapid increase in installation of commercial PV power plants having 25 years or longer lifetimes, the operation and planning for their reliable performance is a crucial challenge (Yang, French, and Bruckman 2019). Ensuring reliable performance includes monitoring the slow loss of electricity output and effectively planning based on the PV power output. We can achieve the required reliability when we accurately forecast power output.

To our knowledge, no work using st-GNNs is published for forecasting PV power output. However, st-GNNs have been used successfully in addressing other forecasting problems such as traffic forecasting (Guo et al. 2019; Yu, Yin, and Zhu 2018; Diao et al. 2019). Researchers' advantage in using st-GNNs for forecasting traffic variables is that the traffic features at a particular intersection are correlated with the nearby traffic intersections, thus resembling the spatial and temporal coherence, which can help improve the prediction of future values. Similarly, in PV power forecasting, the PV power plants located in the same region will have similar power output patterns due to commonalities in weather and solar irradiance. Thus, a GNN model in which nodes of the graph represents power plants and pair of nodes in the same region are connected such that connected power plants are useful in improving the power forecasting accuracy.

Related Work

A large number of studies have been published on forecasting PV power timeseries using various methods (Wu et al. 2020; Sobri, Koohi-Kamali, and Rahim 2018; Das et al. 2018). ARIMA, a conventional statistical method, is typically used to solve these kinds of problems. To make the statistical models reliable against non-stationary and highly nonlinear timeseries, recent forecasting approaches exploit neural network models (Cococcioni, D'Andrea, and Lazzerini 2012; Nespoli et al. 2019). The multi-layer perceptron artificial neural network (ANN) is more robust to non-linearity and sharp dips or rises in values. Radial basis function used in a neural network approach modeled the PV modules' electrical characteristics (Bonanno et al. 2012). The recurrent neural network (RNN) is a technique widely used in timeseries modeling, and it has been used in PV power output correction with high accuracy (Yona et al. 2013). Classification of timeseries into cloudy, rainy, and clear sky is an another improvement to enhance the accuracy of the learning model (Shi et al. 2012; Yang, Thatte, and Xie 2015).

Another set of approaches relies on functions and equations from physical models, clarifying the association between measurement values and environmental factors (Das et al. 2018). Numerical Weather Prediction (NWP) models are another approach for forecasting PV power output (Larson, Nonnenmacher, and Coimbra 2016). Some research on forecasting power generation uses ensemble approaches of the methods mentioned above. The hybrid models try to maximize the benefits of different models and min-

imize their limitations (Wu, Chen, and Abdul Rahman 2014; Gigoni et al. 2018). For example, hybrid models work better when a specific period of the year may cause more error depending on solar irradiance variability. Similarly, hybrid models perform better when some geographical locations have more fluctuating weather conditions leading to higher variability in solar irradiance.

The major limitation of all the existing PV forecasting models is that they only use the data from one single system at a time, while ignoring the information or the data recorded by the neighboring PV plants. The spatiotemporal graph neural network (st-GNN) based forecasting models have the potential to leverage temporal and spatial coherence among PV systems for the forecasting models. In graph theory, a graph is a structure in which individual pairs of objects or entities are related in some manner. A spatiotemporal graph is a structure that constitutes a temporal sequence of individual entities and interactions between them over a time period. In recent years, Graph Neural Network (GNN) techniques are starting to be widely used for use cases where signals from the neighboring nodes can be used to predict a particular node's values. GNNs are a generalized form of a structured two-dimension Convolutional Neural Network (CNN), which has successfully been used in the classification and regression models where the input data is structured, such as images (Alex Krizhevsky, Ilya Sutskever, and Hinton 2012; Lawrence et al. 1997; LeCun et al. 1989). CNNs have been successfully used to classify and quantify the degradation of PV cells exposed to harsh exposure conditions using electroluminescent images (Karimi et al. 2019, 2020). In the generalized form, each pixel in CNN can be represented as a node in the GNN. Several advances in GNN models have been proposed; some of the widely used GNN models are (Niepert, Ahmed, and Kutzkov 2016; Kipf and Welling 2017; Hamilton, Ying, and Leskovec 2017).

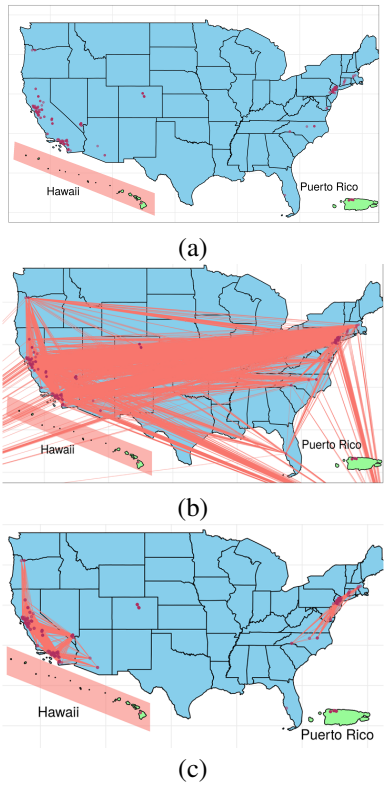

Figure 1: The distribution of PV systems across the United States and the graph model used to represent the relationships between the PV systems. (a) All PV systems in our dataset are shown by red circles on the map. (b, c) Graph showing the proximity of PV systems in which edges are included between PV systems that are proximate according to a low/moderate proximity threshold ($\epsilon = 0$, $\epsilon = 0.5$) is applied.

Dataset

The dataset consists of power timeseries of 316 PV systems represented by points on the map shown in Figure 1. The dataset's time interval is 5 minutes, and the length is two years, amounting to 210,240 datapoints for each timeseries. Some of the weather sensor data from the power plants are of 15 minute time intervals, so they are imputed with a cubic spline to increase data frequency to 5 minutes. There were few systems with one minute time interval datasets, and we sampled those to 5 minutes interval. The missing values of less than one hour were imputed using a linear fit model, and for gaps longer than 1 hour were imputed with month-by-month PV power prediction method (Curran et al. 2019). The dataset is partitioned into training, validation, and testing set as 690, 20, and 20 days respectively.

Architecture of st-Graph Neural Network

In this section, we will first describe a method to build a graph of otherwise isolated PV systems.

Graph is a prerequisite for the design of st-GNN model. The design of the GNN architecture is guided by the training method by which model will learn. There are several variants of GNN such as graph generative networks, graph auto-encoders, graph recurrent neural networks, spatial graph neural network, and spectral graph neural network. The spectral graph neural network applies graph convolution in spectral domain (Kipf and Welling 2017), and the st-GNN model described below is inspired by it.

PV Network Representation

Building a graph of PV systems is equivalent to assigning edges between the nodes. We use Gaussian kernel (Shuman et al. 2013) and a threshold value to assign edges. The value of the kernel function varies in between 0 and 1, and the threshold cutoff (ϵ_c) will determine the density or sparsity of a graph. The aim is to find the optimal neighborhood of a PV system or, in other words, find a value of ϵ_c such that the GNN model gives the highest forecasting accuracy. For a pair of PV systems "x" and "y" represented by two nodes on a graph, we calculate the distance between them. Let lat_x and lon_x represent latitude and longitude of node x respectively. Then for nodes "x" and "y", the distance between them is calculated by the haversine formula for a

sphere (Goodwin 1905) Equation 1 and 2, similar to the law of cosines equation:

$$a = sin^2((lat_y - lat_x)/2) + cos(lat_x)cos(lat_y) \; sin^2((lon_y - lon_x)/2) \quad (1)$$

$$d = 2 \times R \times arcsin(\sqrt{a}) \quad (2)$$

where $R = 6371$ is the radius of the earth. After the distance is calculated between all pair of nodes, according to Equation 3, nodes which are closer will have high $W_{i,j}$ value. For an edge to exist between two nodes the $W_{i,j}$ between them should be higher than ϵ. The reason for choosing distance as a metric for edge detection is that we know from the literature that the climate zone has one of the strongest effect on PV power output(Bonkaney, Madougou, and Adamou 2017). The nodes in a neighborhood will have similar climate zones thus their power output also has similar patterns and will be connected.

$$W_{i,j} = \begin{cases} exp\big(-\frac{dist(i,j)^2}{\sigma^2}\big), i \neq j \text{ and} \\ \qquad exp\big(-\frac{dist(i,j)^2}{\sigma^2}\big) \quad >= \epsilon \\ 0; \qquad \text{otherwise} \end{cases} \quad (3)$$

where, σ is normalizing constant and ϵ constants which controls the sparsity of the graph. The threshold value of ϵ is called threshold cutoff (ϵ_c). The value of ϵ varies between 0 and 1. Figure 1 (b) represents the graph for $\epsilon_c = 0$ while Figure 1 (c) represents the graph for $\epsilon_c = 0.5$ where the nodes in a region are connected by edges. There are several factors that affect the optimal value of ϵ_c such as dataset, number of PV plants, and their locations .

st-Graph Neural Network Model

GNNs have been proposed to extend deep learning for graph analysis (Wu et al. 2020; Hamilton, Ying, and Leskovec 2017). GNN takes an input feature representation $F = (X, A)$,where X is an n-dimensional feature matrix (each row is a feature vector for each node in a graph), and A is the graph's adjacency matrix. The goal is to transform F into a proper vector representation that minimizes a loss function L specified by downstream tasks. The GNN model is made up of convolutional layers and temporal layers. The convolutional layers in the GNN adopt a neighborhood aggregation architecture to learn a discriminative vector representation $h(v)$ for each node v (called "node embedding") across multiple transformation layers. The new layer $h^i(v)$ takes a node embedding $h^{i-1}(v)$ ($h^0(v)$ is $X(v)$ from input feature matrix) and updates its embedding $h^i(v)$ by aggregating the embeddings from its neighbors:

$$h^i(v) \leftarrow \eta \, \text{concat}(h^{i-1}(v), agg(h^{i-1}(v') : v\prime \in \mathcal{N}(v))) \quad (4)$$

where, η is a non-linear activation function.

In the spectral domain, the graph convolution of filter kernel g with signal x is represented as (Kipf and Welling 2017):

$$g \star x = U\hat{g}(\Lambda)U^T x \quad (5)$$

where, $U^T x$ is a Fourier transform of x, U is a matrix of eigenvectors of normalized graph Laplacian matrix $\Delta = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$. Eigen decomposition of $\Delta = U\Lambda U^T$, such that Λ is diagonal matrix of eigenvalues. Since calculating U is an expensive operation, Equation 5 can written as its approximate by Equation 6 (Hammond, Vandergheynst, and Gribonval 2011).

$$g \star x \approx \sum_{k=0}^{K-1} g_k T_k(\tilde{\Delta})x \quad (6)$$

where $\tilde{\Delta} = \frac{2}{\lambda_{max}}\Delta - I_N$, $g \in \mathbb{R}^K$ is a coefficient of $T_k(x)$ (polynomial of order k), such that, $T_k(\tilde{\Delta})(x) = 2T_{k-1}(\tilde{\Delta})(x) - T_{k-2}(\tilde{\Delta})(x)$, $T_0(\tilde{\Delta}) = x$, and $T_1(\tilde{\Delta}) = \tilde{\Delta}x$. In the above equation, the value of K shows the filters are localized up to K hops from the node and K is the size of the kernel for graph convolution.

The temporal convolution layers are 1-D convolutions with a one dimensional kernel filter followed by a sigmoid gated linear unit to provide non linearity. The sigmoid function chooses the relevant elements from the input for capturing complex structure and variances in the timeseries. The graph neural network shown in Figure 2 is a series of graph convolutional layers and temporal convolutional layers on timeseries which gives a building block of learning framework called spatiotemporal block (Yu, Yin, and Zhu 2018; Gehring et al. 2017), to capture the spatial and temporal dynamic behavior of each PV system. The spatiotemporal block is a technique which can extract the most useful temporal features and capture the most appropriate spatial features concurrently (Yu, Yin, and Zhu 2018). The representation of st-GNN architecture for training and forecasting the PV model is shown in Figure 2. We use $L2$ loss function as an objective function so that loss should be minimized over the period of time.

The network has a series of three blocks, two spatiotemporal (ST) blocks, and the third block is an output layer block. Each ST block has a spatial layer sandwiched between two temporal convolutional layers. The output block has two temporal layers connected in series, followed by a fully connected output layer. The GNN model's input data dimension is $\mathbb{R}^{H \times N \times C}$, where H is the number of previous data points of the timeseries used in the model, N is the total number of PV systems and C is the number of channels in the input. The output of the model is of the form $\mathbb{R}^{MXN}$ where M is the number of future time points for which model will forecast.

The TensorFlow2 (Abadi et al. 2016) library was used for building the GNN model and the model was trained and tested on Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz, 48 GB memory, 12 CPU cores, and 12 GB Nvidia GeForce RTX 2080 GPU card.

Local Solar Time Alignment

Local solar time is a term which corresponds to time according to the position of the sun in the sky relative to a specific position on the ground such that local solar noon always happens when the sun is directly overhead (see Figure 3 (a)). On the other hand, "clock time" is the standardized/artificial time that we use in our day-to-day life to standardize our

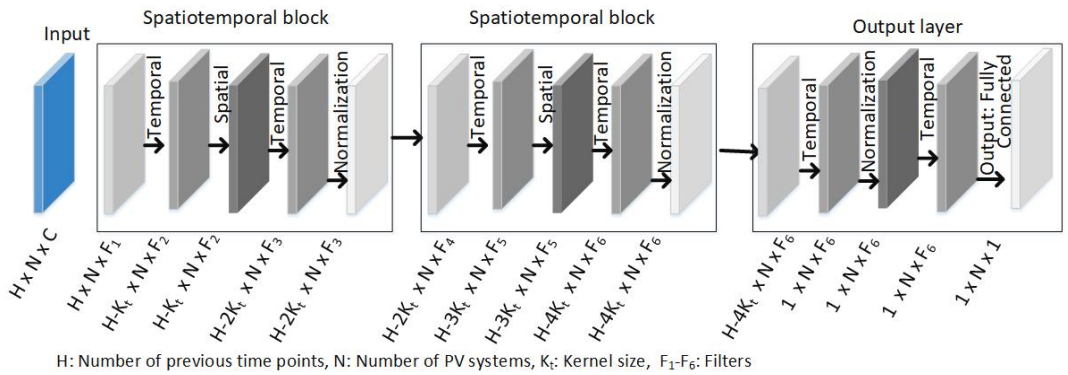

Figure 2: Architecture of spatiotemporal graph neural network. The dimension of input layer is $H \times N \times C$, where H is the number of previous data points or time lags in timeseries. If 2 hours of previous data points are used and the frequency is 5 minutes then the value of H becomes 24. N is number of PV systems or nodes in a graph (316 PV systems) and C is number of channels in the input dataset.

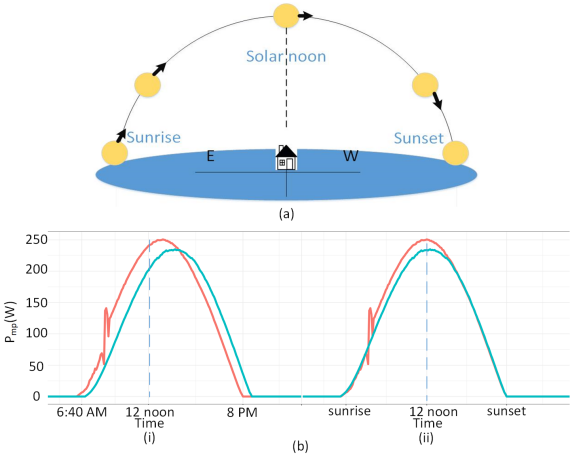

Figure 3: (a) Illustration of movement of sun in a day from sunrise to sunset. Solar noon is defined as the time when the sun is directly over the head at any specific location. (b)Figure shows the P_{mp} values of a two sample of PV systems for a day. (i) Plot of the timeseries of two PV systems before local solar noon alignment. (ii) Plot of the timeseries of two PV systems after local solar noon alignment.

time measurements. In clock time, nearby locations use the same time so the sun is not always at the top for all locations at noon time. This leads to different power timeseries curves to have peak at different clock times as shown in Figure 3 (b) (i). Clock time variation brings asynchronicity among the timeseries which affects the accuracy of prediction models.

The plots in Figure 3 (b) show an example of timeseries of two PV systems before and after local solar noon alignment. It can be seen in Figure 3 (b)(i), the timeseries peaks at some time after the 12 o'clock. Since both the PV systems are located in different areas, but in the same timezone, they peak at different clock times. The local solar time alignment as shown in Figure 3 (b)(ii) makes all timeseries peak at 12 o'clock which brings synchronicity across the timeseries. This approach induces a more linearity among the timeseries of the different PV systems.

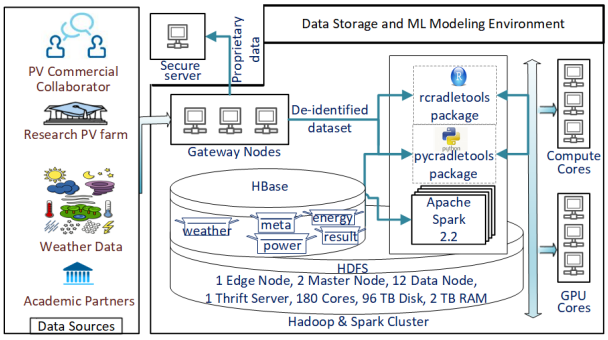

Figure 4: Schematic representation of the data workflow. The figure shows various data sources, and a program periodically executed by a cron job collects the datasets from these sources into the Hadoop cluster. rcradletools and pycradletools are two packages developed internally to provide the functionality to interact with HBase and perform data processing tasks in the R or Python environment, respectively.

Proposed Application Deployment

The path to application deployment involves two major steps: 1) Data management system, 2) Machine learning modeling and deployment.

Data Management: This step includes data acquisition, processing, and storage similar to the steps presented in our earlier work (Arash Khalilnejad et al. 2020). For data acquisition, the majority of data for this work comes from various commercial PV power plant companies, PV research sites, and weather data comes from SolarGIS. We

collect these datasets using web APIs, secure shell FTP, or receive them as CSV files over the cloud as encrypted zip files. The data from different sources have different formats, so the data sets are extracted using particular file parsers, which are scheduled to run as cron jobs. As soon as the data is collected, the first task is to anonymize the proprietary information and save the anonymized data in the Hadoop cluster, shown as a schematic representation in Figure 4. In the data processing step, timeseries from the HDFS are read and passed through validation, tidying, and uniform structuring. Numerical fields are checked for missing values or anomalies and assigned the quality score. The timeseries having a large number of missing values or anomalies receive a low-quality score. Low score timeseries are then passed through data imputation functions to improve data quality. Finally, all the PV power timeseries that receives high quality the score is ingested into HBase. Metadata of the PV systems are generated based on their properties and are also saved in HBase. Data in HBase are stored in a cell such that the value in a cell is uniquely identified by row, column family, column qualifier, and timestamp. Thus, every cell is freighted with a large overhead, and the problem becomes more intense when there are millions of rows. So, the database should be designed such that each cell contains a large amount of data compared to its unique identifier, and in our design, we keep one month of data as a string in a cell. Once the data is ingested in the database, we query the timeseries with over two years of data and are in the same timeperiod. For this work, we received 316 PV timeseries located across different regions in the U.S.

Machine Learning Modeling and Deployment: The project's goal is to build and deploy ML models that can enhance PV power forecasting for commercial power plants. Acquisition and processing of real-time data at regular intervals are critical for power forecasting, so the deployment of a production pipeline will need to have an integrated and automated connection with the data management and inference modules. Data from HBase will be retrieved using pycradletools, a python package developed internally providing an interface between the data management system and inference models. All the data analytics and ML training will happen on the high performance cluster (HPC) side, causing inherent lag from the real-time, so we aim for a near real-time solution. Having more features such as temperature and irradiance can enhance the performance of the forecasting model. We will add more channels corresponding to each feature in the production models. The deployed ML models will be based on the proposed st-GNN model's architecture presented in the above section. The ML models' forecasting will be saved in a result table, and the results will be shared with the power plants using a cloud-based API.

Results

In this section, we will present the results of our st-GNN models. We compare the spatiotemporal model with 1-D temporal convolution model and show that spatiotemporal models outperfoms models having only temporal convolution. The results of the GNN models are computed on 20 days of the test dataset. Mean absolute percentage error (MAPE) is calculated for each PV system using their observed value from the test data and forecasted value from the models. Table 1 compares the accuracy of the model for the spatiotemporal model ($\epsilon_c = 0.375$ or nodes in the neighborhood of the network are optimally connected) with the network when the nodes are isolated and only 1-D convolution, or temporal convolution is applied. Table 2 compare the two models when there are 316 PV systems and when only 77 PV systems were selected to show that the model performs better when there are more PV systems in the network. The tables show mean and standard deviation (sd) of MAPE values. The variability of the results can be ascertained by the violin/boxplots of Figure 5. The width of the violin plot shows the density of data points. Box plots inside the violin plots show median lines of MAPE score for 316 PV systems, and the height of boxes shows the range from first quartile to third quartile.

Plots in Figure 5 (a-d) shows that the optimum value of the network sparsity can be found empirically by varying the value of ϵ_c in the range of [0-1]. Consider Figure 5 (a), it shows 9 violin plots, corresponding to 9 different values of

Forecast (minute)	MAPE for 316 systems			
	s-t convolution		temporal convolution	
	ϵ_c=0.375		ϵ_c=1.0	
	mean	sd	mean	sd
120	11.01	5.04	18.98	5.15
105	9.31	4.36	15.63	4.57
90	8.39	3.87	13.62	4.07
75	7.67	3.36	11.78	3.54
60	7.24	2.87	10.12	2.96
45	6.28	2.61	8.42	2.45
30	4.68	2.48	6.65	2.13
15	2.75	2.37	3.92	2.01

Table 1: Mean and standard deviation of MAPE values for temporal convolution (standalone) vs spatiotemporal convolution for PV systems with optimum ϵ_c for st-GNN network.

Forecast (minute)	77 systems		316 systems	
	mean	sd	mean	sd
120	12.92	4.33	11.01	5.04
105	11.68	4.16	9.31	4.36
90	10.77	3.97	8.39	3.87
75	10.37	3.77	7.67	3.36
60	9.87	3.30	7.24	2.87
45	9.23	3.12	6.28	2.61
30	7.96	3.62	4.68	2.48
15	5.16	3.32	2.75	2.37

Table 2: Mean and standard deviation of MAPE value for forecasting 77 PV systems and 316 PV systems for optimum value of ϵ_c for 15 minutes to 120 minutes.

ϵ_c between 0 and 1. Each violin plot represents the MAPE values of 316 PV systems at a time point of 15 minutes in the future calculated for a specific value of ϵ_c. It is observed that the optimum value of ϵ_c for this model is 0.375, which is a value between 0 and 1. In Figure 5, we also observe that the difference between the error values of an optimum spatiotemporal network ($\epsilon_c = 0.375$) and only temporal network ($\epsilon_c = 1$) increases as the forecast time in future increased from 15 minutes to 120 minutes.

Figure 6 illustrates the MAPE results for the power prediction when the number of previous data points (H) used in the model is varied between 24 (2 hours) to 72 data points (6 hours) with the interval of 12 data points (1 hour). Figure 6 (a) shows the results of MAPE when the forecast future time is 15 minutes and Figure 6 (b) shows the MAPE value when the future time is 120 minutes.

Discussion

In this work, we leverage the st-GNN technique based on the motivation that spatial and temporal coherence among PV power plants can help improve forecasting. The local solar noon alignment method improved the power plants' temporal coherence by shifting all timeseries to local solar time. Solar noon alignment improves the model performance because it reduces the non-linearity among PV time-series. The result in Figure 5 shows the forecast values from 15 minutes to 120 minutes and ϵ_c from 0 to 1 for one-channel st-GNN model. The $\epsilon_c = 0.375$ is an optimum value which consistently gives the highest accuracy for the model or least MAPE values. The optimum value of $\epsilon_c = 0.375$ suggests that neither a graph with all pairs of nodes connected nor the ones having no edges are best, but a partially connected graph, in between the two extremes, gives the best performance. In other words, the signals from the nearby power plants are useful in predicting future values of a particular system, but not all PV systems are useful. In Figure 5, the difference between MAPE values of st-GNN network with isolated nodes ($\epsilon_c = 1$) and an optimal value of ϵ_c increases, indicating that the effect of spatiotemporal coherence is especially significant when predicting the values longer in the future. In the plots of Figure 6, we show the prediction values from the st-GNN models where time lag varies from 24 points (2 hours) up to 72 points (6 hours). Figure 6 (a) shows power forecast at 15 minutes & Figure 6 (b) shows the results for 120 minutes in the future. The prediction error or MAPE value shows little improvement after increasing the number of time lags from 24 points(2 hours) to 72 points (6 hours); this suggests that for more accurate results, time lag of 2 hours can be increased. We also observe that the MAPE are less scattered when we increase the number of time lags in the model.

Table 2 shows that increasing the number of PV plants helps improve the overall accuracy of the model by lowering the MAPE values. In Table 1, we see that the spatiotemporally coherent st-GNN model outperforms the st-GNN model in which the nodes are isolated.

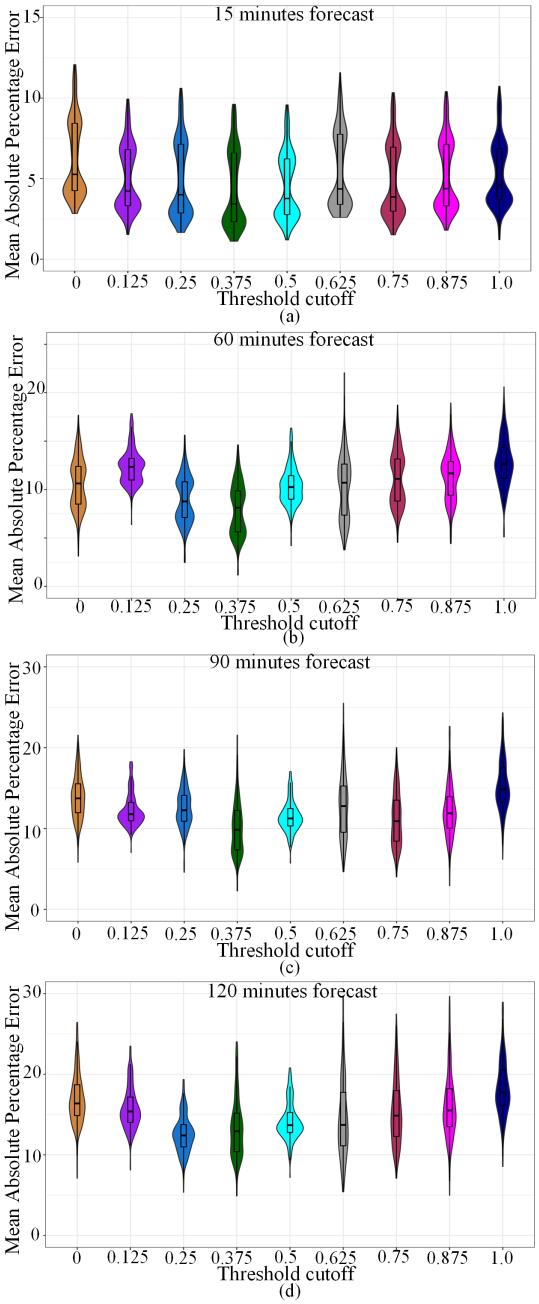

Figure 5: The four plots show performance of st-GNN forecasting models. Mean absolute percentage error (MAPE) on y-axis when predicting future values (a) 15 minutes to (d) 120 minutes. The x-axis corresponds to threshold cutoff ϵ value while building the graph of PV system locations, epsilon value of 0 indicates all nodes are connected and epsilon value of 1 indicates no edge exists in the graph and only temporal convolution is applied.

Conclusion

In this work, we have demonstrated the advantage of st-GNN network models that uses spatial and temporal coher-

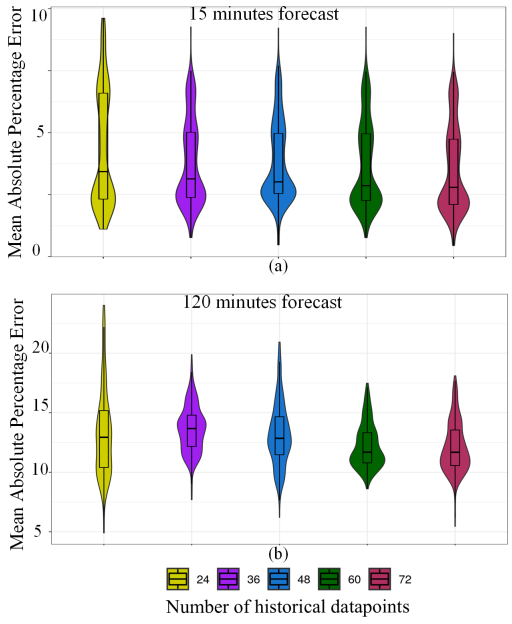

Figure 6: The plots show performance of spatiotemporal forecasting models when the number of time lags are increased from 24 (2 hours) to 72 (6 hours). (a) shows the plot when forecasting 15 minutes in the future and (b) shows the plot when forecasting 120 minutes in the future.

ence over the network in comparison to models in which only temporal or 1-D convolution is applied. Local solar time alignment improves the temporal coherence among the timeseries as the peak of the timeseries for every system occurs at 12 o'clock. The effect of spatiotemporal coherence is more evident when the future forecast time is longer, i.e., 120 minutes. Increasing the number of PV systems improved the overall accuracy of the systems, suggesting that increasing the nodes in the network contributes information that improves the overall accuracy of the model. The Error or the models' MAPE decreases from 19% in an isolated network to 11% for an optimally connected network having ϵ_c equals 0.375, when forecasted for 120 minutes in the future.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Number DE-EE-0008172. The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; Kudlur, M.; Levenberg, J.; Monga, R.; Moore, S.; Murray, D. G.; Steiner, B.; Tucker, P.; Vasudevan, V.; Warden, P.; Wicke, M.; Yu, Y.; and Zheng, X. 2016. TensorFlow: A System for Large-Scale Machine Learning. In *Twelth Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 265–283.

Alex Krizhevsky; Ilya Sutskever; and Hinton, G. E. 2012. Imagenet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger, ed., *Advances in Neural Information Processing Systems*, volume 25, 1106–1114. Lake Tahoe, NV: Curran Associates, Inc.

Arash Khalilnejad; Ahmad M. Karimi; Shreyas Kamath; Rojiar Haddadian; Roger H. French; and Alexis R. Abramson. 2020. Automated Pipeline Framework for Distributed Processing of Large-Scale Building Energy Time Series Data. *PLoS ONE 15(12): e0240461* doi:https://doi.org/10.1371/journal.pone.0240461.

Bonanno, F.; Capizzi, G.; Graditi, G.; Napoli, C.; and Tina, G. M. 2012. A radial basis function neural network based approach for the electrical characteristics estimation of a photovoltaic module. *Applied Energy* 97: 956–961. ISSN 0306-2619. doi:10.1016/j.apenergy.2011.12.085. URL http://www.sciencedirect.com/science/article/pii/S0306261911008919.

Bonkaney, A.; Madougou, S.; and Adamou, R. 2017. Impacts of cloud cover and dust on the performance of photovoltaic module in Niamey. *Journal of Renewable Energy* 2017.

Cococcioni, M.; D'Andrea, E.; and Lazzerini, B. 2012. One day-ahead forecasting of energy production in solar photovoltaic installations: An empirical study. *Intelligent Decision Technologies* 6(3): 197–210. ISSN 18758843, 18724981. doi:10.3233/IDT-2012-0136.

Das, U. K.; Tey, K. S.; Seyedmahmoudian, M.; Mekhilef, S.; Idris, M. Y. I.; Van Deventer, W.; Horan, B.; and Stojcevski, A. 2018. Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews* 81: 912–928. ISSN 1364-0321. doi:10.1016/j.rser.2017.08.017. URL http://www.sciencedirect.com/science/article/pii/S1364032117311620.

Diao, Z.; Wang, X.; Zhang, D.; Liu, Y.; Xie, K.; and He, S. 2019. Dynamic Spatial-Temporal Graph Convolutional Neural Networks for Traffic Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence* 33: 890–897. ISSN 2374-3468, 2159-5399. doi:10.1609/aaai.v33i01.3301890. URL https://aaai.org/ojs/index.php/AAAI/article/view/3877.

Gehring, J.; Auli, M.; Grangier, D.; Yarats, D.; and Dauphin, Y. N. 2017. Convolutional sequence to sequence learning. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, 1243–1252.

Gigoni, L.; Betti, A.; Crisostomi, E.; Franco, A.; Tucci, M.; Bizzarri, F.; and Mucci, D. 2018. Day-Ahead Hourly Forecasting of Power Generation From Photovoltaic Plants. *IEEE Transactions on Sustainable Energy* 9(2): 831–842. ISSN 1949-3037. doi:10.1109/TSTE.2017.2762435. Conference Name: IEEE Transactions on Sustainable Energy.

Goodwin, H. 1905. The Haversine in Nautical Astronomy. URL https://www.usni.org/magazines/proceedings/1905/october/new-method-nautical-astronomy. Library Catalog: www.usni.org.

Guo, S.; Lin, Y.; Feng, N.; Song, C.; and Wan, H. 2019. Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence* 33: 922–929. ISSN 2374-3468, 2159-5399. doi:10.1609/aaai.v33i01.3301922.

Hamilton, W. L.; Ying, R.; and Leskovec, J. 2017. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, 1025–1035. Red Hook, NY, USA: Curran Associates Inc. ISBN 978-1-5108-6096-4.

Hammond, D. K.; Vandergheynst, P.; and Gribonval, R. 2011. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis* 30(2): 129–150. ISSN 1063-5203. doi:10.1016/j.acha.2010.04.005.

Karimi, A. M.; Fada, J. S.; Hossain, M. A.; Yang, S.; Peshek, T. J.; Braid, J. L.; and French, R. H. 2019. Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification. *IEEE Journal of Photovoltaics* 1–12. ISSN 2156-3381. doi:10.1109/JPHOTOV.2019.2920732.

Karimi, A. M.; Fada, J. S.; Parrilla, N. A.; Pierce, B. G.; Koyutürk, M.; French, R. H.; and Braid, J. L. 2020. Generalized and Mechanistic PV Module Performance Prediction From Computer Vision and Machine Learning on Electroluminescence Images. *IEEE Journal of Photovoltaics* 1–10. ISSN 2156-3403. doi:10.1109/JPHOTOV.2020.2973448.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification With Graph Convolution Netwokrs. In *Semi-Supervised Classification with Graph Convolutional Networks*.

Larson, D. P.; Nonnenmacher, L.; and Coimbra, C. F. 2016. Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest. *Renewable Energy* 91: 11–20. ISSN 09601481. doi:10.1016/j.renene.2016.01.039. URL https://linkinghub.elsevier.com/retrieve/pii/S0960148116300398.

Lawrence, S.; Giles, C. L.; Tsoi, A. C.; and Back, A. D. 1997. Face recognition: a convolutional neural-network approach. *IEEE Transactions on Neural Networks* 8(1): 98–113. ISSN 1045-9227. doi:10.1109/72.554195.

LeCun, Y.; Boser, B.; Denker, J. S.; Henderson, D.; Howard, R. E.; Hubbard, W.; and Jackel, L. D. 1989. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* 1(4): 541–551. ISSN 0899-7667. doi:10.1162/neco.1989.1.4.541. Conference Name: Neural Computation.

Nespoli, A.; Ogliari, E.; Leva, S.; Massi Pavan, A.; Mellit, A.; Lughi, V.; and Dolara, A. 2019. Day-Ahead Photovoltaic Forecasting: A Comparison of the Most Effective Techniques. *Energies* 12(9): 1621. ISSN 1996-1073. doi:10.3390/en12091621. URL https://www.mdpi.com/1996-1073/12/9/1621.

Niepert, M.; Ahmed, M.; and Kutzkov, K. 2016. Learning Convolutional Neural Networks for Graphs. In *International Conference on Machine Learning*.

Shi, J.; Lee, W.-J.; Liu, Y.; Yang, Y.; and Wang, P. 2012. Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines. *IEEE Transactions on Industry Applications* 48(3): 1064–1069. ISSN 1939-9367. doi:10.1109/TIA.2012.2190816.

Shuman, D. I.; Narang, S. K.; Frossard, P.; Ortega, A.; and Vandergheynst, P. 2013. The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains. *IEEE Signal Processing Magazine* 30(3): 83–98. ISSN 1053-5888. doi:10.1109/MSP.2012.2235192.

Sobri, S.; Koohi-Kamali, S.; and Rahim, N. A. 2018. Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management* 156: 459–497. ISSN 0196-8904. doi:10.1016/j.enconman.2017.11.019. URL http://www.sciencedirect.com/science/article/pii/S0196890417310622.

Wu, Y.-K.; Chen, C.-R.; and Abdul Rahman, H. 2014. A Novel Hybrid Model for Short-Term Forecasting in PV Power Generation. doi:https://doi.org/10.1155/2014/569249.

Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; and Yu, P. S. 2020. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* 1–21. ISSN 2162-237X, 2162-2388. doi:10.1109/TNNLS.2020.2978386.

Yang, C.; Thatte, A. A.; and Xie, L. 2015. Multitime-Scale Data-Driven Spatio-Temporal Forecast of Photovoltaic Generation. *IEEE Transactions on Sustainable Energy* 6(1): 104–112. ISSN 1949-3029, 1949-3037. doi:10.1109/TSTE.2014.2359974. URL http://ieeexplore.ieee.org/document/6945846/.

Yang, H. E.; French, R. H.; and Bruckman, L. S., eds. 2019. *Durability and Reliability of Polymers and Other Materials in Photovoltaic Modules*. Plastics Design Library. Amsterdam: Elsevier, William Andrew Applied Science Publishers, 1st edition edition. ISBN 978-0-12-811545-9. doi:10.1016/C2016-0-01032-X.

Yona, A.; Senjyu, T.; Funabashi, T.; and Kim, C.-H. 2013. Determination Method of Insolation Prediction With Fuzzy and Applying Neural Network for Long-Term Ahead PV Power Output Correction. *IEEE Transactions on Sustainable Energy* 4(2): 527–533. ISSN 1949-3029, 1949-3037. doi:10.1109/TSTE.2013.2246591. URL http://ieeexplore.ieee.org/document/6473872/.

Yu, B.; Yin, H.; and Zhu, Z. 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 3634–3640. Stockholm, Sweden: International Joint Conferences on Artificial Intelligence Organization. ISBN 978-0-9992411-2-7. doi:10.24963/ijcai.2018/505. URL https://www.ijcai.org/proceedings/2018/505.