

Spatial-Temporal Solar Power Forecast through Use of Gaussian Conditional Random Fields

Bei Zhang, *Student Member, IEEE*, Payman Dehghanian, *Student Member, IEEE*,

Mladen Kezunovic, *Fellow, IEEE*

Department of Electrical and Computer Engineering
Texas A&M University
College Station, TX USA

Abstract—This paper introduces an application of the Gaussian Conditional Random Fields (GCRF) model for forecasting the solar power in electricity grids. The introduced forecasting technique is capable of modeling both the spatial and temporal correlations of various solar generation stations. It will be demonstrated in this paper how the suggested solution can significantly improve the forecast accuracy compared to the conventional forecasting models such as the persistent (PSS) model and the autoregressive with exogenous input (ARX) model. Besides, the GCRF model outperforms the other two models under the scenarios with unavailable or missing data. The suggested probabilistic model of the GCRF can also help better managing the existing uncertainties of the solar generations. The numerical experiments are conducted through which the effectiveness of the proposed approach is validated.

Index Terms—Forecast; solar power; spatial-temporal; Gaussian conditional random fields.

I. INTRODUCTION

Year 2014 has seen a new record in the U.S. solar industry, which has grown by 34% over year 2013, and the country's solar capacity is estimated to double by the end of year 2016 [1]. With the rapid growth of the solar industry, the variability and intermittency of this renewable source of energy brings about major challenges in energy balancing which may jeopardize the system reliability and flexibility [2]. Therefore, it is very critical to have an accurate real-time forecast of the solar generation so that both higher system operation efficiency and maximum solar utilization can be achieved.

There have been a lot of research efforts on the solar generation prediction techniques, which can be roughly clustered into three categories: 1) the numerical weather prediction (NWP)-based forecast [3]-[6]; 2) the data-driven methods [2], [7]-[16] which are, nowadays, the most popular methods; and 3) combinations of the NWP and the data-driven methods [17]-[19]. Recent studies have demonstrated the potential strong spatial correlations among geographically dispersed solar PV power plants [2], [12], [20]. Among the aforementioned studies, [2], [4], [11]-[13] consider the spatial correlations of solar sites, while the rest are proposed solely based on the local meteorological measurements. Studies show that the prediction accuracy can be significantly improved

when spatial correlations are considered [2], [4], [11]-[13].

Gaussian Conditional Random Field (GCRF) is a structured learning method which can well exploit the correlations among output variables, resulting in significant improvements of the prediction accuracy. Besides, its Gaussian nature facilitates the inference as well as the learning efficiency [21].

No studies have been devoted to the solar power forecast through application of the GCRF models. GCRF provides a probabilistic framework for exploiting complex dependence structure among output variables [22], [23], which can help model the spatial correlations among different solar generation stations more effectively. In this paper, the GCRF model is introduced to forecast the solar power considering both the spatial and temporal correlations. Different from the previous works, the proposed methodology captures the probabilistic nature of the GCRF model which will further help modeling the inherent uncertainties of the solar generation.

The rest of the paper is organized as follows: Section II provides the background information on the relationship between solar generation and solar irradiance; Section III introduces several solar power forecast models including the PSS, the ARX, and the GCRF models and the performance indices; Numerical experiments are conducted in Section IV to compare the forecast performances of different models under several scenarios; and finally Section V concludes with the contributions of this paper.

II. SOLAR GENERATION VS SOLAR IRRADIANCE

According to [12], the relationship between the solar power generation and the solar irradiance on a given material can be reasonably assumed as a linear relationship, as shown in (1).

$$P_{solar} = I_{solar} \times S \times \eta \quad (1)$$

where P_{solar} is the solar generation power; I_{solar} is the solar irradiance (kWh/m²); S is the area of the solar panel (m²); and η is the generation efficiency of a given material. With this assumed relationship, the solar generation can be easily predicted once the solar irradiance forecast is available.

III. SOLAR IRRADIANCE FORECAST MODELS

A. PSS Model

The basic idea of the PSS model is to use the solar irradiance at a given time interval as the forecast value at the next time interval, as described in (2):

This publication was made possible by NPRP 8-241-2-095 award from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

$$y'_k = y_k^{t-1} \quad (2)$$

where y'_k is the forecasted solar irradiance of the k^{th} solar station at time t and y_k^{t-1} is the measured value at time $t-1$. The PSS model is usually regarded as a benchmark forecast model. As a result, a forecast model performance is considered to be exceptional if it outperforms that of the PSS model.

B. ARX Model

The ARX model is a data-driven method and mathematically modeled in (3):

$$y'_k = c + \sum_{i=1}^p \varphi_i y_k^{t-i} + \sum_{j=1, j \neq k}^m \sum_{l=1}^{n_j} \beta_{jl} z_j^{t-l} + \varepsilon_t \quad (3)$$

where z_j^{t-l} is the historical measurement from station j at time $t-l$, which is the exogenous input as the measurements do not come from the target station k ; β_{jl} is the coefficient of the exotic input z_j^{t-l} ; φ_i is the coefficient associated with y_k^{t-i} which is the historical measurement within the target station k ; c is a defined constant; and ε_t is the white noise.

Other than the temporal correlation, the ARX model is also able to consider the spatial correlations, which is achieved by taking into account the measurements z from other solar generation stations.

C. GCRF Model

Similarly, the GCRF model takes into account both the spatial and temporal correlations. In this model, entire effort is to obtain the conditional distribution $P(\mathbf{y}|\mathbf{x})$ through a large volume of historical data, and then forecast the output variable based on the obtained distribution. $\mathbf{y}=[y'_1, y'_2, \dots, y'_N]^T$ is the forecasted solar output in multiple stations from 1 to N at the next time interval t ; vector \mathbf{x} represents the historical solar measurements in different stations, as shown in (4):

$$\mathbf{x} = [x_1^{t-1}, x_1^{t-2}, \dots, x_1^{t-m_1}, x_2^{t-1}, x_2^{t-2}, \dots, x_2^{t-m_2}, \dots, x_N^{t-1}, x_N^{t-2}, \dots, x_N^{t-m_N}] \quad (4)$$

The conditional distribution $P(\mathbf{y}|\mathbf{x})$ modeled in conditional random fields (CRF) is expressed in (5):

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x}, \mathbf{a}, \mathbf{\beta})} \exp\left(\sum_{i=1}^N A(\mathbf{a}, y_i, \mathbf{x}) + \sum_{j-i} I(\mathbf{\beta}, y_i, y_j, \mathbf{x})\right) \quad (5)$$

where $A(\mathbf{a}, y_i, \mathbf{x})$ is the association potential which associates the output y_i with input vector \mathbf{x} ; $I(\mathbf{\beta}, y_i, y_j, \mathbf{x})$ is the interaction potential to relate two output variables y_i and y_j [23]. In CRF applications, A and I could be approximated by linear combinations of pre-determined feature functions with corresponding parameters \mathbf{a} and $\mathbf{\beta}$, as denoted in (6) and (7).

$$A(\mathbf{a}, y_i, \mathbf{x}) = \sum_{k=1}^{K_i} \alpha_k f_k(y_i, \mathbf{x}) \quad (6)$$

$$I(\mathbf{\beta}, y_i, y_j, \mathbf{x}) = \sum_{l=1}^{L_{ij}} \beta_l g_l(y_i, y_j, \mathbf{x}) \quad (7)$$

If the feature functions f_k and g_l are defined in quadratic forms, as shown in (8) and (9), the association potential A and interaction potential I would also be quadratic functions of \mathbf{y} .

$$f_k(y_i, \mathbf{x}) = -(y_i - R_k(\mathbf{x}))^2, \quad k = 1, \dots, K \quad (8)$$

$$g_l(y_i, y_j, \mathbf{x}) = -e_{ij}^{(l)} S_{ij}^{(l)}(\mathbf{x}) (y_i - y_j)^2 \quad (9.a)$$

$$e_{ij}^{(l)} = \begin{cases} 1 & (i, j) \in G_l \\ 0 & \text{otherwise} \end{cases} \quad (9.b)$$

where G_l is the graph which imposes the relation between the output variables y_i and y_j . $S_{ij}^{(l)}(\mathbf{x})$ is the function representing similarity between outputs y_i and y_j . And $R_k(\mathbf{x})$ is a single prediction of y_i based on the input variables \mathbf{x} . Under this scenario, the conditional probability distribution $P(\mathbf{y}|\mathbf{x})$ becomes multivariate Gaussian distribution, namely $P(\mathbf{y}|\mathbf{x}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ is the mean vector and $\boldsymbol{\Sigma}$ is the covariance matrix. $P(\mathbf{y}|\mathbf{x})$ can be re-formed in (10) and (11).

$$P(\mathbf{y}|\mathbf{x}) = \frac{\exp\left(-\sum_{i=1}^N \sum_{k=1}^{K_i} \alpha_k (y_i - R_k(\mathbf{x}))^2 - \sum_{i,j} \sum_{l=1}^{L_{ij}} \beta_l e_{ij}^{(l)}(\mathbf{x}) (y_i - y_j)^2\right)}{Z(\mathbf{x}, \mathbf{a}, \mathbf{\beta})} \quad (10)$$

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})\right) \quad (11)$$

where $\boldsymbol{\Sigma}^{-1} = 2(\mathbf{Q}_1 + \mathbf{Q}_2)$, and \mathbf{Q}_1 , \mathbf{Q}_2 are expressed in (12), (13).

$$\mathbf{Q}_{1ij} = \begin{cases} \sum_{k=1}^{K_i} \alpha_k & i = j \\ 0 & i \neq j \end{cases} \quad (12)$$

$$\mathbf{Q}_{2ij} = \begin{cases} \sum_k \sum_{l=1}^{L_{ik}} \beta_l e_{ik}^{(l)} S_{ik}^{(l)}(\mathbf{x}) & i = j \\ -\sum_{l=1}^{L_{ij}} \beta_l e_{ij}^{(l)} S_{ij}^{(l)}(\mathbf{x}) & i \neq j \end{cases} \quad (13)$$

And $\boldsymbol{\mu} = \boldsymbol{\Sigma} \mathbf{b}$ in (10), and \mathbf{b} can be calculated in (14).

$$b_i = 2 \left(\sum_{k=1}^K \alpha_k R_k(\mathbf{x}) \right) \quad (14)$$

To obtain the conditional probability distribution, the parameters α_k and β_l need to be determined since $f_k(y_i, \mathbf{x})$ and $g_l(y_i, y_j, \mathbf{x})$ are both pre-determined. Such parameters can be learnt through maximizing the conditional log-likelihood of the training sets, as denoted in (15) and (16), which can be achieved by the application of the gradient descent algorithm.

$$L(\alpha, \beta) = \sum \log P(\mathbf{y}|\mathbf{x}) \quad (15)$$

$$(\mathbf{a}, \mathbf{\beta}) = \arg \max_{\mathbf{a}, \mathbf{\beta}} (L(\mathbf{a}, \mathbf{\beta})) \quad (16)$$

Since the model is Gaussian, $\boldsymbol{\mu}$ can be evaluated and used as the forecast of the output. Besides, according to the property of the multivariate Gaussian distribution, stating that its marginal distribution over a subset of its random variables is also a Gaussian distribution, the probability distribution of a single output y_i can be obtained as follows:

$$P(y_i | \mathbf{x}) \sim N(\mu_i, \Sigma_{ii}) \quad (17)$$

With such probability distributions, the forecast result can be more compatible with other system applications such as stochastic modeling, risk modeling, etc.

D. Forecast Performance Indices

Two indices are selected to evaluate the performance of the forecast models: the mean absolute errors (MAE) and the root

mean square error (RMSE), as denoted in (18) and (19).

$$\text{MAE} = \frac{1}{Z} \sum_{i=1}^Z |y_i - \hat{y}_i| \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{1}{Z} \sum_{i=1}^Z (y_i - \hat{y}_i)^2}, \quad (19)$$

where \hat{y}_i is the forecasted and y_i is the measured value.

IV. NUMERICAL EXPERIMENTS

A. System Description

Hourly solar irradiance data of 8 solar stations in the year 2010 are collected from the California Irrigation Management Information System (CIMIS) and adopted in our simulations. The locations of the selected solar stations are illustrated in Fig. 1 (more information is provided in [24]). The station No. 1 is selected as the target station in all the simulations, as denoted in green in Fig. 1. In order to enhance the spatial correlations in the system, two artificial stations (No. 9 & No. 10) are added very close to the target station. Therefore, it would seem reasonable to assume that the hourly solar irradiance measurements for such stations should be quite similar to those of station No. 1. Hence, the data from station No.1 plus some low level noise is used as the measurements of these two artificial stations.

B. Parameter Configuration of the ARX and GCRF Models

Based on the available data described above, parameter m introduced in the ARX model [see eq. (3)], which is a data-driven method, is selected to be 9. Besides, $p=10$, and $n_f=10$, which implies that the previous ten historical measurements of each solar station are considered for the forecasting purpose.

Regarding the GCRF model, entire attempt is devoted to model both the spatial and temporal correlations, as illustrated in Fig. 2. The temporal correlations are modeled through the association potential $A(\alpha, y_i, \mathbf{x})$, illustrated as the red dot line, by using the autoregressive (AR) model. Therefore, $K_f=1$ [see eq. (5)], and $R(x)$ in (8) can be calculated using (20):

$$R_i(x) = c + \sum_{m=1}^{p_i} \varphi_m y_i^{t-m} \quad (20)$$

where p_i is selected to be 10 to consider the previous 10 historical measurements. We model the spatial correlations through the interaction potential $I(\beta, y_i, y_j, \mathbf{x})$ in the GCRF

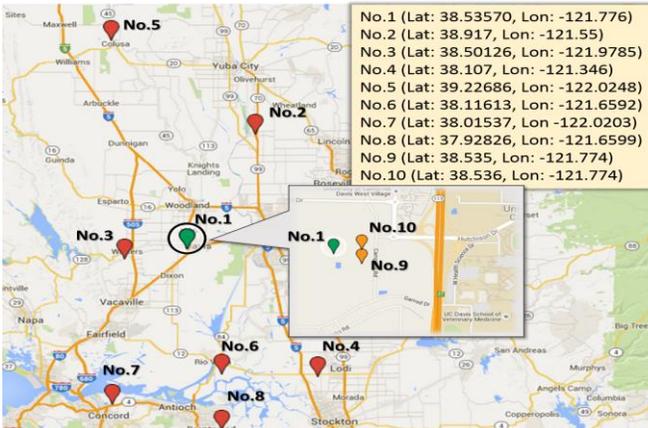


Figure 1. Location and spatial information of the studied solar stations.

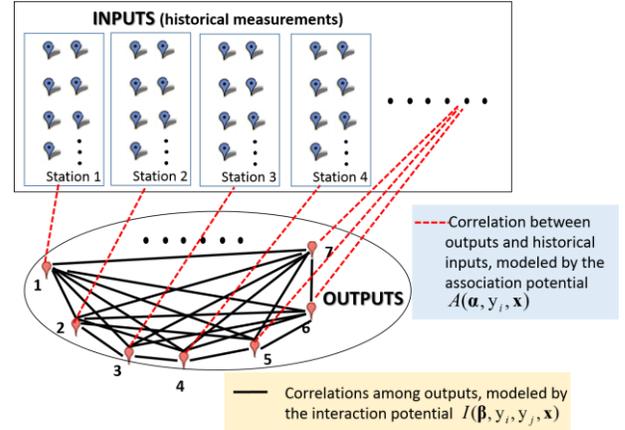


Figure 2. Spatial and temporal correlations in the GCRF model.

model, illustrated as the black line in Fig. 2. In (7), $L_f=1$ is selected and the graph in (9) is considered to be the geographical graph including all the solar stations. Therefore, S_{ij} in (9) is calculated in (21), where D_{ij} is the distance between station No. i and No. j . The reason to take the reciprocal is that the bigger the S_{ij} is, the closer y_i and y_j will be and vice versa.

$$S_{ij} = \frac{1}{D_{ij}^2} \quad (21)$$

C. Simulation Cases and Scenarios

Four cases are generated in the training and validation periods and are listed in Table I. Basically, two scenarios are considered in the numerical simulations: *Scenario 1* where there are no missing data, and *Scenario 2* where there are some missing data points (This scenario is tested since some measurements may not be available all the time due to various reasons such as equipment failure, communication issues, etc. Moreover, when a solar station is newly created, few historical data is available for training purposes).

There are several sub-scenarios considered in *Scenario 2* as follows: *Scenario 2-1* in which only one hourly data set happens to be missing in the target station; *Scenario 2-2* in which two successive hourly data sets happen to be missing in the target station; *Scenario 2-3* in which one hourly data set happen to be missing in several stations; and *Scenario 2-4* in which no data is available in one station, and therefore this station is excluded from the training process.

D. Simulation Results

The forecast results corresponding to the three studied models under *Scenario 1* are tabulated and compared in Table II. In addition, the detailed forecast performances of the studied techniques in Case 3 under *Scenario 1* are illustrated in Fig. 3 and Fig. 4. The green line denotes the ideal prediction performance, i.e., the closer the prediction result is to the line, the better the forecast is done. It can be seen that the GCRF model has the best performance in all cases under *Scenario 1*.

Figures 5-7 demonstrate the performance indices of PSS, ARX and GCRF models in the studied *Scenarios 2-1*, *2-2* and *2-3*, respectively. The results are demonstrated only for case 3 in all scenarios due to the page limitations. When one data is missing, we simply use the measurement of the previous hour to approximate that missing measurement. From Fig. 5 and Fig. 6, we can observe that the GCRF model is superior to the other two models when dealing with the cases with missing

TABLE I. TRAINING AND VALIDATION PERIODS

| Case | 1 | 2 | 3 | 4 |
|-------------------|-----------------|-------------|-----------------|-------------------|
| Training Period | January, March | May | July, September | November |
| Validation Period | February, April | April, June | August, October | October, December |

TABLE II. PERFORMANCE INDICES OF VARIOUS FORECASTING MODELS: SCENARIO 1

| Index | Cases | Forecast Model | | |
|-------|--------|----------------|---------|---------|
| | | PSS | ARX | GCRF |
| MAE | Case 1 | 90.3676 | 56.5334 | 55.1527 |
| | Case 2 | 98.1372 | 51.8562 | 40.4062 |
| | Case 3 | 96.6623 | 35.5478 | 25.5906 |
| | Case 4 | 92.8664 | 51.6816 | 29.6195 |
| RMSE | Case 1 | 111.9337 | 76.7467 | 74.4007 |
| | Case 2 | 116.5823 | 81.9164 | 60.6969 |
| | Case 3 | 111.6060 | 55.8073 | 40.6566 |
| | Case 4 | 108.1498 | 67.8648 | 43.7008 |

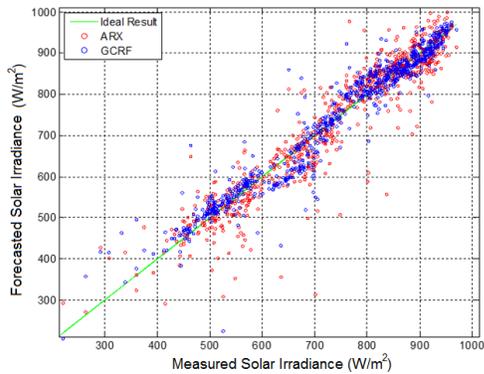


Figure 3. Prediction performance of ARX and GCRF models: Scenario 1 (Case 3)

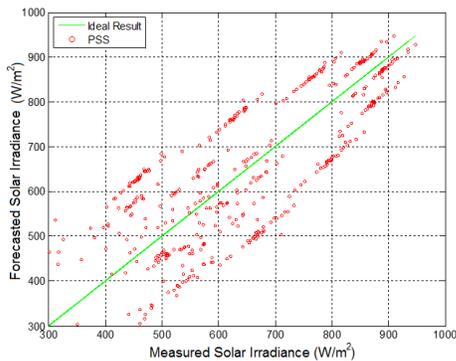


Figure 4. Prediction performance of PSS model: Scenario 1 (Case 3)

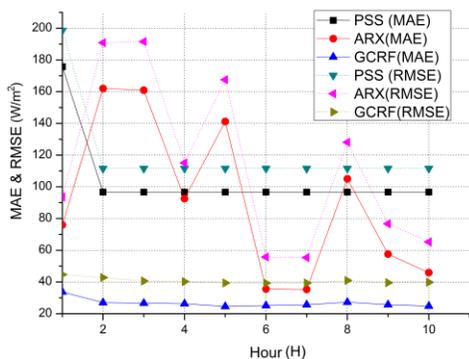


Figure 5. Performance indices of various forecasting models: Scenario 2-1 (Case 3)

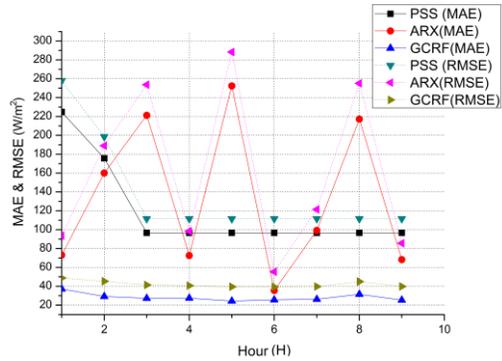


Figure 6. Performance indices of various forecasting models: Scenario 2-2 (Case 3)

data. The ARX model, however, may compromise a lot and its performance may become even worse than the PSS model especially when more missing information is encountered.

Fig. 7 illustrates the MAE performance indices of the three models when missing data occur in multiple solar stations. It can be observed that: 1) the GCRF model has the best results; 2) when there is no missing data in stations No. 9 & 10, the GCRF performs very well, and the less the missing information, the better its performance would be [see the first two sub-figures]; 3) by comparing the sub-figures in the first and second row, one can see that GCRF performance may also be compromised a bit when the data is missing in the added stations. However, it still outperforms the other two models most of the time.

The better performance of the GCRF model, especially when there is missing data, is due to its modeling of the spatial correlations among different solar stations. Here, the spatial correlations among stations No. 9, No. 10, and the target station No. 1 are relatively strong as they are physically very close to each other. Therefore, the measurements in station No. 9 & No. 10 can somehow serve as the backup when there is missing data in the target station. GCRF's performance also compromises a bit when missing data occur in station No. 9 or No. 10. Scenario 2-4 simulates this situation when the data from Station No. 5 are not available for training, and the

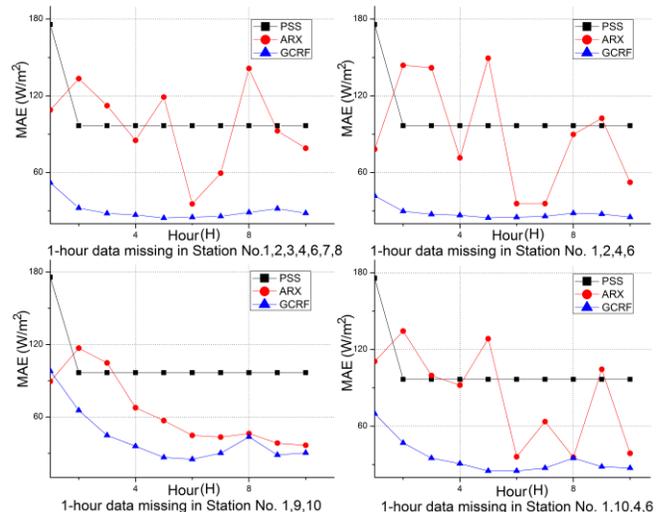


Figure 7. MAEs corresponding to various forecasting models: Scenario 2-3 (Case 3)

TABLE III. PERFORMANCE INDICES OF VARIOUS FORECASTING MODELS: SCENARIO 2-4

| Index | Cases | Forecast Model | | |
|-------|--------|----------------|---------|---------|
| | | PSS | ARX | GCRF |
| MAE | Case 1 | 96.6623 | 58.2602 | 55.9011 |
| | Case 2 | 96.6623 | 47.7145 | 43.0316 |
| | Case 3 | 96.6623 | 50.2712 | 26.8112 |
| | Case 4 | 96.6623 | 56.3702 | 30.7201 |
| RMSE | Case 1 | 111.6060 | 79.2030 | 75.6125 |
| | Case 2 | 111.6060 | 76.5143 | 63.7870 |
| | Case 3 | 111.6060 | 68.0714 | 41.8258 |
| | Case 4 | 111.6060 | 72.0233 | 44.8699 |

TABLE IV. GCRF'S PERFORMANCE FOR PROBABILITY DISTRIBUTION MODELING: SCENARIO 1 (STATION NO. 1)

| | Case 1 | Case 2 | Case 3 | Case 4 |
|----------------|--------|--------|--------|--------|
| $\pm 1 \sigma$ | 35.82% | 58.61% | 57.61% | 55.91% |
| $\pm 2 \sigma$ | 59.96% | 80.22% | 80.19% | 78.67% |
| $\pm 3 \sigma$ | 78.74% | 89.74% | 92.75% | 88.53% |

results are demonstrated in Table III. The GCRF model still has the best performance, although it is compromised a bit compared to that in *Scenario 1*. In contrast, the performance of the ARX model is not quite predictable and it is compromised a lot in Case 3.

By applying the property of multivariate Gaussian distribution, as described in (17), Table IV shows the GCRF's performance for the probability distribution modeling of the output from station No.1 under *Scenario 1*. The table shows the percentage of the real measurements that falls within the $\pm m \sigma$ ($m=1, 2, 3$) range of the forecast values. The uncertainty of the studied solar forecast basically fits the normal probability distribution, although the percentages are less than those of the standard normal distribution, especially in Case 1. Future work is to improve the performance of GCRF technique for modeling the probability distributions.

V. CONCLUSIONS

The contributions of this work are summarized as follows:

- Novel solar power forecast considering both the spatial and temporal correlations among different solar stations is introduced using the GCRF model
- Performance comparisons are conducted among the PSS, ARX and GCRF forecast models through extensive numerical experiments and various scenarios with and without missing data.
- Forecast uncertainty modeling is accomplished by utilizing the property of the multivariate Gaussian probability distribution.

VI. ACKNOWLEDGEMENT

Thanks are due to Dr. Zoran Obradovic and his team at Temple University for introducing us to their GCRF research.

VII. REFERENCES

- [1] Solar Energy Industries Association, "Solar Energy Facts: 2014 Year in Review", Dec. 2014. [Online]. Available: <http://www.seia.org/sites/default/files/Q4%202014%20SMI%20Fact%20Sheet.pdf>. [Accessed: Oct. 7th, 2015].
- [2] C. Yang, A. Thatte, and L. Xie, "Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation," *IEEE Trans. Sustainable Energy*, vol. 6, no. 1, pp. 104-112, Jan. 2015.
- [3] S. Jafarzadeh, M. S. Fadali, and C. Y. Evrenosoglu, "Solar power prediction using interval type-2 TSK modeling," *IEEE Trans. Sustainable Energy*, vol. 4, no. 2, pp. 333-339, Apr. 2013.
- [4] E. Lorenz, J. Hurka, D. Heinemann, and H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 2-10, Mar. 2009.
- [5] W. Ji., and K. C. Chee, "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN," *Solar Energy*, vol. 85, no. 5, pp. 808-817, May. 2011.
- [6] L. A. Fernandez-Jimenez, A. Muñoz-Jimenez, A. Falces, et al, "Short-term power forecasting system for photovoltaic plants," *Renewable Energy*, vol. 44, pp. 311-317, Aug. 2012.
- [7] A. Sfetsos, and A. H. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques," *Solar Energy*, vol. 68, no. 2, pp. 169-178, Feb. 2000.
- [8] H. Long, Z. Zhang, and Y. Su, "Analysis of Daily Solar Power Prediction with Data-driven Approaches," *Applied Energy*, vol. 126, pp. 29-37, Aug. 2014.
- [9] E. Izgi, A. Öztopal, B. Yerli, et al, "Short-mid-term solar power prediction by using artificial neural networks," *Solar Energy*, vol. 86, no. 2, pp. 725-733, Feb. 2012.
- [10] J. Shi, W. J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Trans. Industry Applications*, vol. 48, no. 3, pp. 1064-1069, May. 2012.
- [11] M. Lave, J. Kleissl, and J. S. Stein, "A wavelet-based variability model (WVM) for solar PV power plants," *IEEE Trans. Sustainable Energy*, vol. 4, no. 2, pp. 501-509, Apr. 2013.
- [12] C. Yang, and L. Xie, "A novel ARX-based multi-scale spatio-temporal solar power forecast model," in *2012 North American Power Symposium*, Urbana-Champaign, IL, USA, Sep. 9-11, 2012.
- [13] R. J. Bessa, A. Trindade, and V. Miranda, "Spatial-temporal solar power forecasting for smart Grids," *IEEE Trans. Industrial Informatics*, vol. 11, no. 1, pp. 232-241, Feb. 2015.
- [14] C. Chen, S. Duan, T. Cai, and B. Liu, "Online 24-h solar power forecasting based on weather type classification using artificial neural network," *Solar Energy*, vol. 85, no. 11, pp. 2856-2870, Nov. 2011.
- [15] H. T. Yang, C. M. Huang, Y. C. Huang, and Y. S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Trans. Sust. Energy*, vol. 5, no. 3, pp. 917-926, Jul. 2014.
- [16] G. Capizzi, C. Napoli, and F. Bonanno, "Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting," *IEEE Trans. Neural Networks and Learning Systems*, vol. 23, no. 11, pp. 1805-1815, Nov. 2012.
- [17] C. Tao, D. Shanxu, and C. Changsong, "Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement," in *2010 2nd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, pp. 773-777.
- [18] P. Bacher, H. Madsen, and H. A. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, no.10, pp.1772-1783, 2009.
- [19] M. G. Kratzenberg, S. Colle, and H. G. Beyer, "Solar radiation prediction based on the combination of a numerical weather prediction model and a time series prediction model," in *Proc. 1st Int. Congr. Heat Cool Build*, Lisbon, Portugal, Oct. 7-10, 2008, pp. 1-12.
- [20] J. Widen, "Correlations between large-scale solar and wind power in a future scenario for Sweden," *IEEE Trans. Sustain. Energy*, vol. 2, no. 2, pp. 177-184, Apr. 2011.
- [21] V. Radosavljevic, S. Vucetic, and Z. Obradovic, "Neural Gaussian Conditional Random Fields" In *Machine Learning and Knowledge Discovery in Databases*, Springer Berlin Heidelberg, pp.614-629, 2014.
- [22] J. Stojanovic, M. Jovanovic, Dj. Gligorijevic, Z. Obradovic, "Semi-supervised learning for structured regression on partially observed attributed graphs" in *Proc. the 2015 SIAM International Conference on Data Mining*, Vancouver, Canada, April 30 - May 02, 2015.
- [23] V. Radosavljevic, "Gaussian conditional random fields for regression in remote sensing," Ph.D. dissertation, Department of Computer and Information Sciences, Temple University, Philadelphia, 2012.
- [24] California Irrigation Management Information System (CIMIS), [Online]. Available: <http://www.cimis.water.ca.gov/>