LASSO-based Single Index Model for Solar Power Generation Forecasting

Ningkai Tang, Shiwen Mao
Dept. ECE, Auburn Univ.
Auburn, AL, U.S.A.
Email: nzt0007@auburn.edu, smao@ieee.org

Yu Wang
Dept. EE, Nanjing Univ. Aeronautics & Astronautics, Nanjing, China
Email: yuwang15@nuaa.edu.cn

Mark Nelms
Dept. ECE, Auburn Univ.
Auburn, AL, U.S.A.
Email: nelmsrm@auburn.edu

***Abstract*—Despite the high promises of renewable energy, it brings great challenges to the existing power grid due to its nature of intermittent and uncontrollable generation. In order to fully harvest its potential, accurate forecasting of renewable power generation is indispensable for effective power management. In this paper, we propose a LASSO-based forecasting model and algorithm for solar power generation forecasting. We compare the proposed scheme with two representative schemes with a real world dataset. We find that the LASSO-based algorithm achieves a considerably higher accuracy comparing to the existing methods, using fewer training data and being robust to anomaly data points in the training data. Its variable selection capability also offers a trade-off between complexity and accuracy, which all make it a highly competitive solution to forecasting of solar power generation.**

***Index Terms*—Smart grid; renewable energy; generation forecasting; machine learning; LASSO.**

I. Introduction

With technology developments and evolution of the power grid, the concept of smart grid (SG) has emerged and is regarded as the next generation power grid [1]–[7]. A smart grid is an electricity network that can intelligently integrate the interactions of all users connected to generators, consumers, and those that assume both roles, in order to efficiently deliver sustainable, economic and secure electricity supplies [8]. Such capabilities are enabled by the computation, communication, and control mechanisms that are incorporated with the power grid. The smart grid is characterized with the two-way flow of power and information, microgrid, and distributed renewable energy resources (DRERs) [9]. In the meantime, the rise of new energy (e.g., photovoltaic power such as solar power) has brought new challenges to such unconventional power networks. Although integrating power charge from solar power generator could reinforce the macro grid, a large and uncertain amount of power generated by micro solar grids could lead to severe energy management problems [10].

In order to fully harvest the potential of DRERs, two key techniques, load forecasting (i.e., to predict the amount of power needed to meet the demand and supply equilibrium) and power generation forecasting (i.e., to predict how much power will be generated at a future time), are indispensable. Load forecasting has been well studied in the literature [11]–[13], with different statistics and machine learning approaches, such as nonparametric functional time series analysis, state space models, and artificial neural networks. Similarly, generation forecasting has been investigated with various models and methods as well.

Since solar power generation is linked directly with solar intensity, the power forecasting problem naturally translates to a weather forecasting problem. In [14] and [15], support vector machine (SVM) and nonlinear time series are used to predict solar intensity, respectively. Other prior works such as [16]–[18] also provided various effective solutions to the solar intensity prediction problem. Although the prior works have done a good job on achieving a low error rate, a deeper analysis will be helpful to gain a good understanding of the problem. For example, the SVM-based method [14] achieves a low error rate, but the selection of kernel is usually based on trial and error. For neural network based technologies [18], the neural network structure needs to be pre-designed and a quite complicated structure is needed to achieve a good precision, which, however, leads to a high computational cost. In rainy or cloudy days, the time-series based method [19] are usually not effective to capture the high variations in data. In [15], the authors present a locally linear model for nonlinear time series, which leads to an accurate approximation and an analysis on the relationship between the renewable power generation process and the weather processes. However, the importance of each variable is yet to be identified.

In this paper, we investigate the solar power generation forecasting problem, aiming to develop an effective method that not only achieve a high forecasting accuracy, but also helps to reveal the significance of weather variables. To this end, we propose a least absolute shrinkage and selection operator (LASSO) based method for solar power generation forecasting using historical weather data. Based on a single index model and LASSO, we develop an effective algorithm that maximizes Kendall's tau coefficient, a statistic that is used to measure the ordinal association between two measured quantities [20], to estimate the prediction model coefficients. The goal of variable selection is achieved naturally by the nature of LASSO, which automatically reduces the weights of less important variables and increases the sparsity of the overall coefficient vector. With the proposed algorithm, we can either maximize the prediction accuracy using all the weather data/variables, or achieve a trade-off between accuracy and complexity by using a limited, smaller number of variables. The proposed scheme is evaluated with the real dataset collected from a weather station [21], and comparison to two representative benchmark schemes. The proposed LASSO-based scheme outperforms both existing schemes with considerable reduction in prediction error.

The remainder of this paper is organized as follows. We formulate the forecasting problem and introduce LASSO preliminaries in Section II. We then present our LASSO-based algorithm in Section III and validate its performance in Section IV. We conclude the paper in Section V.

II. Forecasting Model and Problem Statement

A. System Model

For problem formulation, we adopt the single index model as follows.

$$Y|\mathbf{X} \sim P(\cdot, f(\mathbf{X}^T\beta)), \tag{1}$$

where $P(\cdot, \theta)$ represents a stochastically increasing family of functions with parameter θ and $\mathbf{X}$ as covariates, β is the coefficient vector of X and is unit normed, and $f(\cdot)$ is an unknown strictly smooth, increasing link function.

To relate the model (1) to our problem, Y is our desired estimation of future solar intensity and $\mathbf{X}$ is the weather data collected from a weather station. In our forecasting algorithm, we use a special case of the model as follows.

$$Y = f(\mathbf{X}^T\beta) + z, \tag{2}$$

where z is a zero mean variable with a finite variance representing forecasting error. Specifically, the weather data $\mathbf{X}$ consists of five weather data variables, including temperature, humidity, dew point, wind speed, and precipitation, which compose a 5-dimensional dataset.

B. LASSO Preliminaries

In machine learning and statistics, least absolute shrinkage and selection operator (LASSO) has become a popular method for regression analysis, ever since it was first introduced by Robert Tibshirani in 1996 [22]. Due to the nature of LASSO, where l_1 penalty is used, it performs both variable selection and regularization, and usually achieves a sparse solution. In recent years, LASSO has been successfully applied to single index models [23]–[25] due to the above mentioned capability.

We propose to use LASSO for solar power generation with high accuracy. In addition, since weather data gathered from the local weather station can vary in different types of weather parameters to monitor, it is important to identify the variables that are more important on solar power generation, especially when lacking of sufficient weather information. As discussed, linear regression, neural networks, and SVM based algorithms have already been applied to the solar power generation forecasting problem. To the best of our knowledge, this is the first application of LASSO to the problem, to achieve high prediction precision as well as variable selection.

C. Kendall's tau Coefficient

With a set of i.i.d. data samples, the proposed algorithm should be capable of simultaneous variable selection and generation forecasting through optimizing the relationship between Y and $\mathbf{X}^T\beta$. Since it is not clear whether the problem is linear or not, we propose to use Kendall's tau coefficient between Y and $\mathbf{X}^T\beta$ [20] instead of Pearson's correlation coefficient. Thus, we can precisely estimate β by maximizing the following τ coefficient between Y and $\mathbf{X}^T\beta$.

For discontinuous β, Kendall's tau coefficient is expressed as

$$\tau_n(\beta) = \frac{1}{n(n-1)} \sum_{1\le i_1 \ne i_2 \le n} \text{sign}(Y_{i_2} - Y_{i_1}) \cdot \text{sign}\left(\mathbf{X}_{i_2}^T\beta - \mathbf{X}_{i_1}^T\beta\right). \tag{3}$$

For continuous forms of β, we have

$$\tau_n^*(\beta) = \frac{1}{n(n-1)} \sum_{1\le i_1 \ne i_2 \le n} \text{sign}(Y_{i_2} - Y_{i_1}) \cdot \tanh\left(\frac{\mathbf{X}_{i_2}^T\beta - \mathbf{X}_{i_1}^T\beta}{c}\right), \tag{4}$$

where c is a small constant which can be seen as given.

III. Solution Algorithm

A. Proposed Algorithm

With the definitions of Kendall's tau coefficient in (3) and (4), the proposed solution algorithm consists of two parts, i.e., coefficient estimation and link function estimation. We have the algorithm executed with the following steps:

1) Coefficient Estimation: First we need to find an index j that can maximize the following value ρ_j, $j = 1, 2, ..., p$, where p is the dimension of the variables.

$$\rho_j = \frac{1}{n(n-1)} \sum_{1\le i_1 \ne i_2 \le n} \text{sign}(Y_{i_2} - Y_{i_1}) \cdot \text{sign}(\mathbf{X}_{i_2 j} - \mathbf{X}_{i_1 j}). \tag{5}$$

Next we set $\hat{\beta}_{(1)} = \text{sign}(\rho_{j_1})e_{j_1}$, where $e_j = [0, ..., 1, ..., 0]^T$ is a $p\times 1$ vector with 1 at the jth position and 0 at all other positions. Suppose we have $X_{j_1}, X_{j_2}, ..., X_{j_{k-1}}$ as the selected variables, and the current optimized coefficient is $\hat{\beta}_{(k-1)}$. For the remaining $j \notin \{j_1, j_2, ..., j_{k-1}\}$, we continue our procedure in parallel solving the following problem.

$$\hat{\beta}_j = \arg\max_{\beta_j} \left\{\tau_n^*(\hat{\beta}_{(k-1)} + \beta_j e_j) - \lambda|\beta_j|\right\}, \tag{6}$$

where λ is a system parameter. We will discuss its selection in detail in Section III-B. We then set j_k as

$$j_k = \arg\max_{j}\{\tau_n^*(\hat{\beta}_{(k-1)} + \hat{\beta}_j e_j), j \notin \{j_1, \ldots, j_{k-1}\}\}. \tag{7}$$

The algorithm will terminate if the following condition is satisfied, where ϵ is a predefined small positive value.

$$\tau_n^*(\hat{\beta}_{(k-1)} + \hat{\beta}_{j_k} e_{j_k}) - \tau_n^*(\hat{\beta}_{(k-1)}) < \epsilon. \tag{8}$$

Otherwise set $\hat{\beta}_{(k)}$ as

$$\hat{\beta}_{(k)} = \frac{\hat{\beta}_{(k-1)} + \hat{\beta}_{j_k} e_{j_k}}{||\hat{\beta}_{(k-1)} + \hat{\beta}_{j_k} e_{j_k}||_2}, \tag{9}$$

and repeat the above steps until the stop condition is satisfied. Then we obtain the estimated coefficient vector $\hat{\beta}$.

2) Link Function Estimation: Now we perform isotonic regression to estimate the link function $f(t)$, which remains unknown. First, we define

$$Z_i = \mathbf{X}_i^T \hat{\beta}. \tag{10}$$

We then sort $\{Z_1, Z_2, ..., Z_n\}$ in increasing order, and denote the result as $\{Z_{(1)}, Z_{(2)}, ..., Z_{(n)}\}$. We also sort $\{Y_1, Y_2, ..., Y_n\}$ according to $\{Z_{(1)}, Z_{(2)}, ..., Z_{(n)}\}$. We next execute the pool-adjacent-violators algorithm (as defined in [26]) on the sorted Y's, and mark the result as $\{Y_{(1)}, Y_{(2)}, ..., Y_{(n)}\}$

By choosing a symmetric and smooth kernel function $\text{Ker}(t)$, we estimate the link function $f(t)$ as

$$\hat{f}(t) = \frac{\sum_{j=1}^{n} \text{Ker}\left(\frac{1}{b}(t - Z_{(j)})\right) \times Y_{(j)}}{\sum_{j=1}^{n} \text{Ker}\left(\frac{1}{b}(t - Z_{(j)})\right)}, \tag{11}$$

where b is chosen by applying the cross validation technique. If there are at least two different Y's, the estimated function $f(t)$ will be strictly smooth and monotone.

3) Threshold and Prediction: With the estimated coefficient vector β, estimated link function $f(t)$, and new observation $\mathbf{X}'$, we can predict the solar intensity by

$$\hat{Y}^* = \hat{f}(\mathbf{X}'^T \hat{\beta}). \tag{12}$$

We manually set a threshold T for the minimum solar intensity. For example, we can set $T = 0$ by default since solar intensity cannot have a negative value. The final prediction result $\hat{Y}'$ is computed as

$$\hat{Y}' = \begin{cases} \hat{Y}^*, & \text{if } \hat{Y}^* > T \\ T, & \text{otherwise.} \end{cases} \tag{13}$$

With the estimated vector β, estimated link function $f(t)$, new observation $\mathbf{X}'$, and threshold T, we can precisely predict solar intensity $\hat{Y}'$.

B. Selection of λ

The parameter λ in (6) is one of the most important parameters in the proposed algorithm. It is sensitive to various problems and should be carefully tuned. In this section, we present two basic methods on how to choose λ.

The first method is the *cross validation technique*. For initialization, we need to shuffle the dataset and randomly split the samples into five subsets (for a five-fold cross validation) with equal size. Then we pick a set of possible λ values and $\tau_n(\beta)$ is calculated on a one-fifth data subset by using the estimated β from the remaining data. After repeating the process until all five parts have been calculated (i.e., to avoid the possible unbalanced results caused by the randomness in data), we could choose the λ that maximizes the average estimation precision in the cross validation process. The proposed algorithm is amenable to parallel computation and the processing speed can be greatly increased. The cross validation technique is used when we have abundant time and information, and are in need of best estimation precision.

Alternatively, we can use the *regularization path method* to achieve a tradeoff between precision and speed. When we emphasize more on speed and an acceptable precision is specified, we can choose the λ value with the following process.

1) First, choose a set of possible λ values and sort them by increasing order;
2) Then execute the proposed algorithm for each λ and record their performance;
3) Plot the achieved precision performance versus the values of λ;
4) Choose an acceptable point on the curve to guarantee the performance while achieving the maximized estimation speed due to sparsity.

The regularization path method also has the potential to achieve high estimation accuracy even when lacking of information.

C. Prediction Model

It is noticed that both the observational and forecasted weather dataset are time-series datasets that change over weather patterns and time periods. As the result shown in [14], solar intensity depends on multiple weather variables, which could help us to construct an accurate prediction model. The structure of the dataset and the possible relationship among the weather variables motivate our proposed LASSO-based method for developing solar intensity prediction models. To construct the model, we utilize historical weather data as input, which include several forecasting data parameters and the actual solar intensity, with totally six weather variables. The proposed algorithm constructs a function that computes solar intensity from the five forecasting weather variables. Thus we could use the function as the prediction model for future solar power generation. We also use part of the remaining data to test the model's accuracy. One unique benefit of using our proposed technique is the relatively low requirement for data size. In general, not too much data is needed; usually historical data over a $15 - 30$ day period will be sufficient.

We focus our study on short-term forecasting for the next few days. We develop a model that shows a relationship between solar intensity and forecasted weather data. For any time t, we form the model by using the historical data of the past 30 days as input, which means the data from $(t - 30)$ to $(t - 1)$. Using the proposed model, we then predict the solar intensity at time t. In Section IV, we also compare the accuracy of our proposed model with different popular and efficient models, including an SVM-based model [14] and a time-series based model [15].

Using the basic single index model presented in Section II, the general prediction model is given as

$$Y \sim P(\cdot, f(\mathbf{Temperature}, \mathbf{DewPoint}, \mathbf{WindSpeed}, \mathbf{Precipitation}, \mathbf{Humidity})), \quad (14)$$

where $f(\cdot)$ is the function that we determine using different prediction methods. The unit of each parameter in the model are: temperature in degrees of Fahrenheit, dew point in degrees of Fahrenheit, wind speed in miles per hour, precipitation in inches, and humidity in percentage between 0% and 100%. However, to avoid potential problems, before applying any selected algorithms, we normalize all the feature data to have a zero mean and unit variance.

To quantify the accuracy of each model, we compute the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) between our predicted solar intensity and the actual solar intensity observed. RMSE and MAPE are well-known statistical measurements of the accuracy of values predicted by models with respect to the observed values. RMSE and MAPE of zero indicate that the model exactly predicts solar intensity with no error (although this is impossible in reality). The closer the RMSE and MAPE values are to zero, the more accurate the model's prediction is. RMSE and MAPE are calculated as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{y}_t - y_t)^2} \quad (15)$$

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_t - y_t}{y_t}\right|, \quad (16)$$

where n represents the number of predicted values, $\hat{y}_t$ stands for the prediction result at time t, and y_t is the actual value of time t.

IV. Simulation Validation

In this section, we present our simulation settings and results. We use the dataset collected at the Davis Weather station, which is available at the UMass Trace Repository [21].

A. Dataset Description

The dataset [21] used in our simulation study was collected every five minutes at a weather station, which contains various sensors that measure temperature, wind chill, humidity, dew-point, wind speed, wind direction, rainfall, barometric pressure, sunlight, and UV. The dataset was recorded over quite a long period of time from February 2006 to January 2013. However, the dataset contains errors, which are indicated by a value of -100000, and missing data for some periods. In the simulation study, we excluded such errors and missing data. In accordance with our general knowledge, we observe peaks in hots summer days and valleys, in contrast, in cold winter time. Also, we find strong correlation between consecutive days. Thus we try to use seasons and months as additional parameters and use historical data of the past 30 days as training samples.

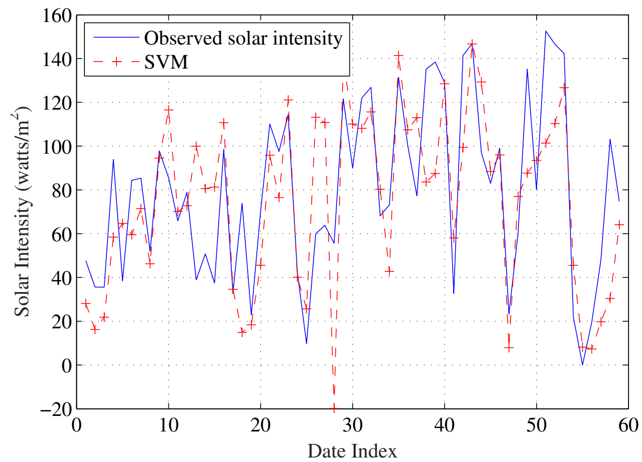

Fig. 1. Solar power generation prediction using the SVM-based method.

B. Simulation Results and Discussions

1) SVM-based Method: With the UMass dataset [21], we first apply the SVM method since it is shown to be effective and widely used in prediction and classification [14]. Here we use historical weather data as training samples and aim to predict the solar intensity data through January 1st, 2013 to February 28th, 2013. In the simulations, we found that different sets of training data have considerable effects on the estimation accuracy. Experimenting with all the data that is available, we achieve the optimal accuracy with the historical data from January 1st, 2012 to February 28th, 2012, which is exactly one year ahead of the target period for prediction. The predicted solar intensity and the real recorded solar intensity are plotted in Fig. 1, while the best RMSE achieved by the SVM-based method is 30.1524 watts/m^2. However the MAPE for the dataset is as high as 468.283, which is largely because the large deviation of the 55th day data, which has very low sunlight. If we exclude that day, the MAPE of the SVM-based method will be reduced to 39.2063.

2) TLLE-Based Method: In [15], we propose a time series based method named TLLE, which greatly improved the accuracy compared to SVM and Multi-linear Regression (MLR) based approaches. The same UMass Trace Repository data [21] is used in [15]. The historical data from January 1st, 2012 to February 28th, 2012 is used to construct the model, and solar power generation is predicted for the period from January 1st, 2013 to February 28th, 2013.

The predicted solar intensity and the real recorded solar intensity are plotted in Fig. 2. The best RMSE we obtained with the TLLE-based method is 23.1464 watts/m^2, which is about the same as that reported in [15] and a 23.2% reduction over the SVM-based method. This result validate the advantage of the TLLE-based method compared to the SVM-based method. We also find a high MAPE value, also due to the anomaly data of the 55th day. The MAPE for the remaining data, by excluding the 55th day data, is reduced to 29.0174, which is a 26.0% reduction over the SVM-based method.

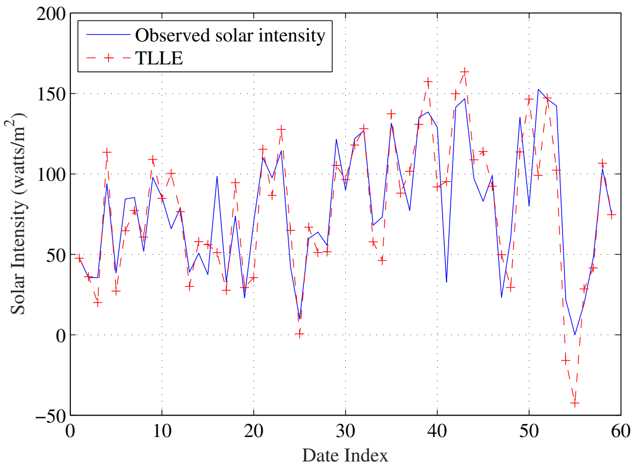

Fig. 2. Solar power generation prediction using the TLLE-based method.

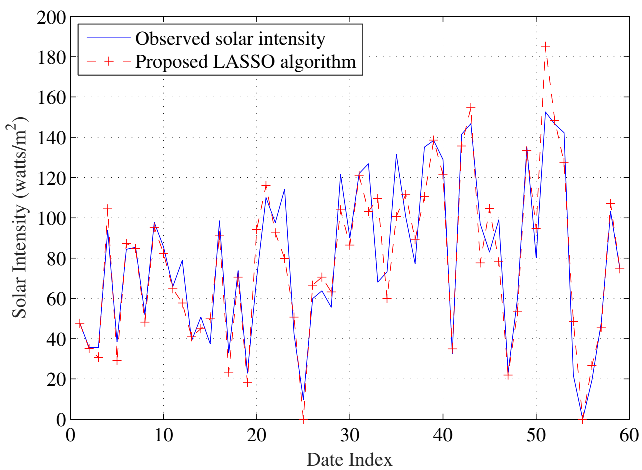

Fig. 3. Solar power generation prediction using the proposed LASSO-based method.

3) Proposed LASSO-based Algorithm: Now we apply the proposed LASSO-based method to predict solar intensity. In the simulation, we use a relatively smaller training sample size of 30, which means, the training data here is gathered from the past 30 days of the targets. Applying the proposed LASSO-based method to training data yields the prediction results that is plotted in Fig. 3.

For the LASSO-based prediction curve in Fig. 3, the RMSE is 14.0262 watts/m^2 and the MAPE is 17.817, representing a further 39.4% and 60.1% reduction over the TLLE-based approach, respectively. More important, *these results are achieved with the entire* 30*-day original data, i.e., without excluding the* 55*th day anomaly data in the dataset*. Our method also achieves a very stable performance in MAPE. Furthermore, even if we reduce the number of training data for 30 days to 15 days, the proposed LASSO-based algorithm still achieves a reliable result, with an RMSE value lower than 20 watts/m^2.

Clearly, our LASSO-based algorithm has achieved considerably higher accuracy compared to the two existing methods. In addition, it requires fewer training data and is robust to anomaly data points in the training data, which make it a high competitive solution to practical problems such as forecasting of solar power generation.

TABLE I
CORRELATION MATRIX OF MASSACHUSETTS DATASET

	Temp	Humid	Dew	Wind	Rain	Solar
Temp	1.000	0.219	0.960	−0.282	0.099	0.594
Humid	0.219	1.000	0.473	−0.392	0.423	−0.447
Dew	0.960	0.473	1.000	−0.351	0.203	0.403
Wind	−0.282	−0.392	−0.351	1.000	0.009	−0.088
Rain	0.099	0.423	0.203	0.009	1.000	−0.277
Solar	0.594	−0.447	0.403	−0.088	0.277	1.000

TABLE II
OPTIMIZED β WITH 3 VARIABLES

	Temp	Humid	Dew	Wind	Rain
β_j	0.4676	−0.3878	0	0	−0.7938

C. Variable Selection

A notable advantage of the LASSO-based algorithm is its ability of variable selection. By tuning the loss function with λ, it allows to identify which variable(s) are more "important" to the prediction result. The process is to tune λ in condition of an acceptable prediction accuracy, until any β_j has reached 0. Then we can treat the corresponding variable X_j as the least important one. Repeating this procedure, we can identify the second least important variable, and so forth.

Variable selection will at least provide us with two fascinating advantages: (i) reducing the computational complexity and (ii) simplifying the prediction model. Due to the structure of our proposed algorithm, historical data will also be used in the prediction stage. So when less parameters are used in the prediction model, less data will be used and the computational cost will be greatly reduced. In addition, a simplified model can provide us a clearer understanding of the relationship between solar power generation and all the parameters. We could use the reduced model to estimate solar power generation when the dataset is incomplete, or to reduce the computation time when necessary.

As an example, we use the UMass dataset to illustrate the variable selection procedure. Table I shows the correlation matrix computed with the dataset, where each element is the correlation of the row variable and the column variable. Although the problem cannot be simply defined as a linear one, we could still use the correlation matrix to obtain an intuitive observation. As the matrix shows, Temperature and Humidity are more closely correlated to Solar intensity, while Rain is quite independent with most other parameters. After adjusting the λ value to reduce the model to a 3-variable model, we have the optimized β values listed in Table II.

From Table II, we find that Dew point and Wind speed are identified as less important variables. The result for Wind speed coincide with the correlation matrix but Dew point and Rain have a conflict. However, noticing from Table I that Dew

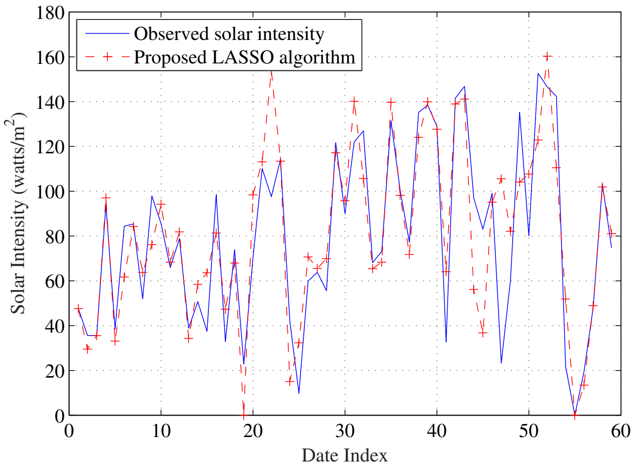

Fig. 4. Prediction with the three selected variables.

point is tightly correlated with Temperature, while Rain is quite independent, the result for β becomes highly reasonable. Fig. 4 provides the prediction result with the 3-parameter model. The RMSE in this case is 21.2468 watts/m^2, which is still better than both the SVM-based and the TLLE-based methods, but the MAPE has increased to 68.5998 due to the inaccuracy on some small values. Such phenomenon can be explained by the intrinsic characteristics of variance. When we use fewer variables, we actually lose a certain amount of variance and the prediction will become more unbiased to maintain the accuracy. Thus, it is expected to have a lower accuracy and the absolute percentage error on certain small values could become high. The variable selection capability provides a useful trade-off between computational/model complexity and accuracy.

V. Conclusion

In this paper, we proposed a LASSO-based algorithm that accurately predict solar power generation with a small amount of historical data. After presenting the detailed algorithm design, we compared the proposed scheme with two representative existing schemes with the UMass dataset. We found that the LASSO-based algorithm achieved considerably higher accuracy compared to two existing methods, using fewer training data and being robust to anomaly data points in the training data, and the variable selection capability offered a trade-off between complexity and accuracy, which all made it a high competitive solution to practical problems such as forecasting of solar power generation.

Acknowledgment

This work was supported in part by the NSF under Grants CNS-1702957 and DMS-1736470, and the Wireless Engineering Research and Education Center at Auburn University.

References

[1] Y. Wang, S. Mao, and R. Nelms, *Online Algorithms for Optimal Energy Distribution in Microgrids*, 1st ed. New York, NY: Springer, 2015.

[2] Y. Huang, S. Mao, and R. M. Nelms, "Adaptive electricity scheduling in microgrids," in *Proc. IEEE INFOCOM 2013*, Turin, Italy, Apr. 2013, pp. 1142–1150.

[3] ——, "Adaptive electricity scheduling in microgrids," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 270–281, Jan. 2014.

[4] Y. Wang, S. Mao, and R. Nelms, "Distributed online algorithm for optimal real-time energy distribution in the smart grid," *IEEE Internet of Things J.*, vol. 1, no. 1, pp. 70–80, Feb. 2014.

[5] Y. Huang, S. Mao, and R. Nelms, "Smooth scheduling for electricity distribution in the smart grid," *IEEE Sysetms J.*, vol. 9, no. 3, pp. 966–977, Sept. 2015.

[6] Y. Wang, S. Mao, and R. M. Nelms, "An online algorithm for optimal real-time energy distribution in smart grid," *IEEE Trans. Emerging Topics Comput.*, vol. 1, no. 1, pp. 10–21, July 2013.

[7] Y. Huang and S. Mao, "On quality of usage provisioning for electricity scheduling in microgrids," *IEEE Systems J.*, vol. 8, no. 2, pp. 619–628, June 2014.

[8] N. Hatziargyriou, *Micrograds: Architectures and Control*, 1st ed. Wiley-IEEE press, 2014.

[9] Y. Wang, S. Mao, and R. Nelms, "On hierarchical power scheduling for the macrogrid and cooperative microgrids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1574–1584, Dec. 2015.

[10] M. Alam, K. Muttaqi, and D. Sutanto, "A comprehensive assessment tool for solar PV impacts on low voltage three phase distribution networks," in *Proc. IEEE ICDRET'12*, Dhaka, Bangladesh, Jan. 2012, pp. 1–5.

[11] M. Chaouch, "Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 411–419, Jan. 2014.

[12] V. Dordonnat, S. J. Koopman, and M. Ooms, "Dynamic factors in periodic time-varying regressions with an application to hourly electricity load modelling," *Elsevier Comput. Stat. & Data Analy.*, vol. 56, no. 11, pp. 3134–3152, Nov. 2012.

[13] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001.

[14] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in *Proc. IEEE SmartGridComm'11*, Brussels, Belgium, Oct. 2011, pp. 528–533.

[15] Y. Wang, G. Cao, S. Mao, and R. Nelms, "Analysis of solar generation and weather data in smart grid with simultaneous inference of nonlinear time series," in *Proc. IEEE INFOCOM WKSHPS'15*, Hong Kong, China, Apr./May 2015, pp. 600–605.

[16] R. J. Bessa, A. Trindade, and V. Miranda, "Spatial-temporal solar power forecasting for smart grids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 1, pp. 232–241, Feb. 2015.

[17] Y. Zhang, M. Beaudin, R. Taheri, H. Zareipour, and D. Wood, "Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2253–2262, Sept. 2015.

[18] 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 Sys.*, vol. 1, no. 4, pp. 38–46, Dec. 2015.

[19] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, and M. Cony, "Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning," *Elsevier Solar Energy*, vol. 84, no. 10, pp. 1772–1781, Oct. 2010.

[20] M. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–89, 1938.

[21] University of Massachusetts, "The UMass trace repository," [online] Available: http://traces.cs.umass.edu/index.php/Sensors/Sensors/.

[22] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. Royal Stat. Soc., Series B (Methodological)*, vol. 73, no. 3, pp. 267–288, June 1996.

[23] P. Zeng, T. He, and Y. Zhu, "A Lasso-type approach for estimation and variable selection in single index models," *J. Comput. Graph. Stat.*, vol. 21, no. 1, pp. 92–109, Apr. 2012.

[24] W. Y. Hwang, H. H. Zhang, and S. Ghosal, "FIRST: Combining forward iterative selection and shrinkage in high dimensional sparse linear regression," *Statistics and its Interface*, vol. 2, no. 3, pp. 341–348, 2009.

[25] S. Luo and S. Ghosal, "Forward selection and estimation in high dimensional single index models," *Statistical Methodology*, vol. 33, pp. 172–179, Dec. 2016.

[26] O. Burdakov, A. Grimvall, and M. Hussian, "A generalised PAV algorithm for monotonic regression in several variables," in *Proc. COMPSTAT'04*, Prague, Czech Republic, 2004, pp. 761–767.