Using Machine Learning to Assess Solar Energy Grid Disturbances

Jose Ramirez[1], Esteban A. Soto[1], Ebisa Wollega[2], and Lisa B. Bosman[1]

Purdue Polytechnic Institute
[1]Purdue University
West Lafayette, IN 47907, USA
ramir241@purdue.edu, soto34@purdue.edu, and lbosman@purdue.edu

Department of Engineering
[2]Colorado State University – Pueblo
Pueblo, CO 81001, USA
ebisa.wollega@csupueblo.edu

Abstract

Energy generation, sources and distribution methods have been continuously evolving over the past decade. With the increased efficiency associated with solar energy production and distribution, local homeowners have also assumed the role of energy generators, even getting credit for access electricity supplied to the grid given the policy around net-metering. When planning their energy distribution frameworks, electricity providers have to take these changes in energy consumption and generation into account. However, little is known about how solar energy systems impact the demand and supply of grid electricity managed by utility companies. This study proposes a new approach to solar energy predictive modeling which combines machine learning and a variety of publicly available data sources to predict site-specific temperature and solar irradiance (the two primary "missing ingredients"). The preliminary findings show a decreased error when using the new approach (near-future data) in comparison to the traditional approach (historical data) for predicting solar energy generation. As the adoption of solar energy increases, so will potential disruptions to the grid. These preliminary findings show the potential for aggregating individual site-specific predictions to the regional level for the purpose of estimating area-specific solar energy disturbances and moving efforts towards predictive grid optimization.

Keywords

Photovoltaic, Prediction, and Machine Learning

1. Introduction

1.1 Problem Identification

With the increased efficiency associated with solar energy production and distribution, local homeowners have also assumed the role of energy generators, even getting credit for access electricity supplied to the grid given the policy around net-metering. Currently, 40 states plus DC and 4 territories of the United States have net-metering programs or laws (DSIRE, 2020). When planning their energy distribution frameworks, decision makers (Regional Transmission Organizations, Independent System Operators, Electric Markets, and Utility Companies) have to take these changes in energy consumption and generation into account. However, little is known about how solar energy systems impact the demand and supply of grid electricity managed by utility companies.

1.2 Current Approaches to the Problem

Various approaches have been proposed for the predictive maintenance of photovoltaic systems, including manual diagnostics, failure modes, and effects analysis, machine learning, and analysis of real-time sensors (L. B. Bosman, Leon-Salas, Hutzel, & Soto, 2020). Climatic variations significantly affect the energy produced by photovoltaic systems causing voltage variations and low quality in energy production (Ghiani, Pilo, & Cossu, 2013; Kalogirou, 2001; Obando, Carvajal, & Pineda, 2019). Solar irradiance is one of the critical variables to predict the production of photovoltaic energy; currently, there are various methods and models to forecast solar irradiance. One way to forecast solar irradiation is through numerical weather prediction methods. There are regional or global mathematical models that use current weather conditions to predict the weather for a given zone (Krishnamurti, 1995; Reikard & Hansen, 2019).

According to West et al. (2014), movement patterns and numbers of clouds in the sky are a critical factor in predicting irradiance. Physical techniques use satellite imagery as another model to generate data on the movement of clouds. With the cloud movement data, models can predict solar irradiation in the short term, up to 6 hours ahead (Arbizu-Barrena, Ruiz-Arias, Rodríguez-Benítez, Pozo-Vázquez, & Tovar-Pescador, 2017).

Statistical models have been widely used for the forecasting of variables. Mainly, time series have shown excellent performance when historical data needs to be collected at short time intervals for predicting irradiance (Reikard & Hansen, 2019). On the other hand, Box-Jenkins-based autoregressive methods are the most widely used models for short-term prediction of solar irradiation. The method is statistically reliable in predicting meteorological variables. The ARMAX and ARIMAX models are the most used for the prediction of solar irradiance due to the flexibility that the technique has to model complex patterns (Bin Shams, Haji, Salman, Abdali, & Alsaffar, 2016).

1.3 Gaps in Current Approaches

Traditional models based on linear models do not consider the stochastic behavior of variables that can affect the prediction of solar irradiance (Zeng & Qiao, 2013). The use of machine learning models is proposed to solve this problem (Husein and Chung, 2019). Mellit and Pavan (2010) propose three approaches to use machine learning in predicting solar irradiation: (1) represent solar irradiation as a function of meteorological variables; (2) predict irradiance using historical data; (3) a combination of both approaches. Currently, the most widely used machine learning techniques for predicting solar irradiation are feed-forward neural networks, recurrent neural networks, and support vector machines (Husein & Chung, 2019).

Another of the critical variables in the prediction of the performance of photovoltaic systems is the ambient temperature and the cell temperature. The temperature around the cell is a crucial parameter that significantly impacts the performance of photovoltaic systems (Routh et al., 2012). It is critical to have as input data a forecast temperature to control the operation of photovoltaic systems (Ramakrishna, Bernstein, Dall'Anese, Scaglione, & Ieee, 2018). The efficient prediction of ambient temperature is essential for decision-making in the management of photovoltaic systems and the characterization of the systems (Chaabene & Ben Ammar, 2008). Ambient temperature is a critical parameter for predicting the behavior and performance of solar systems (Hassan, Youssef, Mohamed, Ali, & Hanafy, 2016; Hernandez, Gordillo, & Vallejo, 2013). Currently, machine learning models are the most used to predict ambient temperature (Ferreira, Gomes, Martins, & Ruano, 2012). Module temperature plays a crucial role in predicting energy production in a photovoltaic system (TamizhMani, Ji, Tang, Petacci, & Osterwald, 2003). According to Ju et al. (2013) the cell temperature is directly related to the performance of the photovoltaic module. In the literature, there is a comprehensive set of models and techniques to estimate the cell temperature: Analytical models, empirical models, and machine learning models.

1.4 Proposed Solution and Motivation for the Study

The goal of the research is to use machine learning of OFF-SITE weather parameters to predict ACTUAL ON-SITE solar irradiance and ambient temperature. Practically speaking, we would like to be able to watch the news (or lookup weather predictions) based on nearby weather stations and use that information to forecast weather ON-SITE. When the solar irradiance and ambient temperature are accurately estimated, expected power can be simply calculated based on a formula (Lisa Bosman, 2014; L. Bosman & Darling, 2016; L. B. Bosman & Darling, 2018). Solar energy system performance for fixed flat-plate panels can be calculated using Equation 1 (Dobos, 2014), including the given variables.

$$P_{mod} = \frac{I_M}{I_0} * P_{DC} * [1 + \gamma * (T_M - T_0)] * \delta \quad (1)$$

- P_{mod} = module estimated AC power generation, W
- I_M = module plane-of-array irradiance, W/m^2
- I_0 = STC solar irradiance, W/m^2
- P_{DC} = module rated maximum DC power, W
- γ = module temperature coefficient, %/˚C
- T_M = module temperature, ˚C
- T_0 = STC temperature, ˚C

- δ = derate factor

Within Equation 1, the greatest uncertainty, resulting in the greatest solar energy estimation inaccuracies, is limited to two variables including I_M (module plane-of-array irradiance, W/m^2) and T_M (module temperature, ˚C). Currently, for the purpose of prediction, the module plane-of-array irradiance (W/m^2) and ambient temperature, used to estimate module temperature (˚C), are commonly derived from TMY (Typical Metrological Year) data (Wilcox & Marion, 2008). TMY data sets are commonly used to design renewable energy performance models. However, the data sets are not interchangeable due to the variation in data structures and variables collected (Hong, Change, & Lin, 2013). The TMY data sets offer hourly values of solar irradiance and meteorological parameters for 1-year periods for locations within the United States and existing territories. Because of the "typical" nature of the data sets, they are not designed for worst-case conditions or even present day scenarios, and are commonly limited to major cities within the United States. As this is a conference proceeding, this study is limited in scope and will show how machine learning can use local weather station data to predict on-site temperature. Future work will integrate solar irradiance prediction. Figure 1 shows the overall purpose of the research. The first step is to obtain real-time data from weather stations. Subsequently, the data from the weather stations will be used for the on-site forecast of variables as temperature. Having predictive models for variables on-site will allow predicting solar electricity production. In the future, data from different resources will help electricity providers to level the demand and supply.

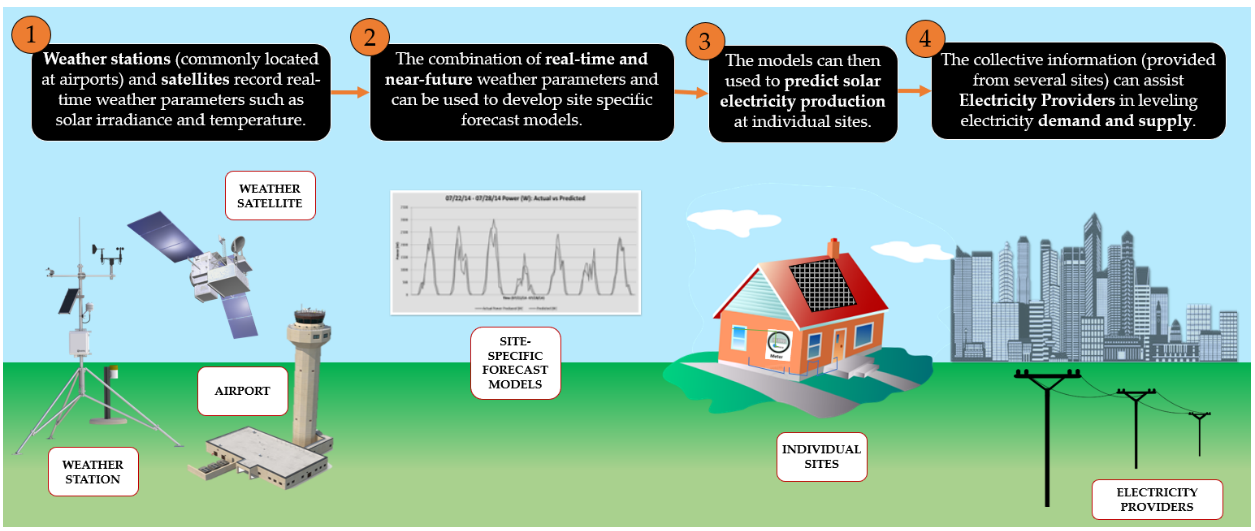

Figure 1: Purpose of research is to use local weather station data to predict near-future weather parameters

2. Methods

The researchers propose a new approach to solar energy prediction which combines machine learning and a variety of data sources. A data set of power production and real-time weather was collected from the College of Menominee Nation's Solar Energy Research Institute through the Sunny Portal and Enphase Energy interface. Also, local historical weather was downloaded from Dark Sky and World Weather Online. The data set consists of 5 years' worth of data from the start of May 2014 to the end of May 2019 at hourly increments. This data set is complimented with NREL's PVWatts for the purpose of comparison.

The College of Menominee Nation's Solar Energy Research Institute (SERI) was established in 2014 and consists of two main systems, a 3.0 kW micro inverter system and a 13.2 kW central inverter system, in addition to performance and weather data collection systems. The 3.0 kW micro inverter system was installed in mid-April 2014 and consists of twelve 250 W Solar World standard crystalline silicon panels each with its own Enphase micro inverter. The panels are positioned at a fixed tilt on a metal roof 6/12 pitch of 26.56 degrees at a south facing orientation of 180 degrees. The 13.2 kW central inverter system was installed in mid-September 2014 and consists of two SMA central inverters and forty-eight 275 W Solar World standard crystalline silicon panels. The panels are positioned at a fixed tilt on a metal roof 6/12 pitch of 26.56 degrees at a south facing orientation of 180 degrees. The performance related data collection includes individual solar energy panel generation and inverter output in 1-hour time increments. The weather related data collection is also available in 1-hour time increments and includes plane-of-array solar irradiance, module temperature, ambient temperature, wind direction, and wind speed.

World Weather Online offers a free weather API to access unlimited historical data and up to 5 days forecast weather data. It has its own weather forecasting model which integrates algorithms and data from (1) European Centre for Medium-Range Weather Forecasts, (2) World Meteorological Organization, (3) NASA weather satellite imagery, (4) NOAA GFS2 model and (5) JMA research models. DarkSky offers a free weather API developed using a wide range of data sources to including (1) NOAA, (2) Canadian Meteorological Center, (3) German Meteorlogical Center, (4) Icelandic Meteoroloical Center, and (5) EUMETNET's Meteoalarm weather alerting system. DarkSky offers both forecasts and historical data in 1-hour increments.
PVWatts was developed in 1999 by National Renewable Energy Laboratory, and is the standard industry tool used to estimate PV system energy production and resulting cost of energy (Darling, You, Veselka, & Velosa, 2011). Upon identifying a location to get started, the user must enter System Info, including DC System Size, Module Type, Array Type, System Losses, Tilt, and Azimuth. NREL offers an API website to assist in the development of a software application for model analysis using larger data sets. The results provide a monthly and hourly breakdown of AC energy production and the associated AC energy value. The TMY2 data set for Green Bay, WI, location ID 14898, was used because it was the closest location to the test facility in Keshena, WI.

3. Results and Discussion

The data comprise of 43684 samples and 78 variables including the binary encoding of the categorical independent variables. The inferential analysis shows that four independent variables are highly correlated with the response variable, ambient temperature. These variables are temperature with correlation coefficient of r = 0.963, heat index r = 0.963, dew point r = 0.901, wind chill r = 0.957. Detail analysis of coefficient of determination (r^2) shows that the four predictors combined result in r^2 of 0.932, which essentially equivalently explains the variability in the response variable compared to all the 77 predictors with r^2 of 0.935. For the purpose of this paper, the temperature variable with r^2= 0.927 is utilized to illustrate the selection of a predictive model by using machine learning technique. Since the underlying response variable is a continuous variable, regression model is proposed as the predictive model selection domain. The regression equation consisting of all the predictors is given as follows:

Ambtemp = 14.82 + 1.0194 Temp - 0.09093 Windspeed - 0.2464 Precip - 0.06274 Humidity
- 0.0308 Visibility - 0.00755 Pressure - 0.00442 Cloudcover + 0.1341 Heatindex
+ 0.0623 Dewpoint - 0.2700 Windchill + 0.01560 Windgust + 0.1239 Feelslike
+ 0.3300 UVindex + 0.0 Wind16point_E - 0.137 Wind16point_ENE
- 0.055 Wind16point_ESE - 0.034 Wind16point_N + 0.112 Wind16point_NE
- 0.019 Wind16point_NNE + 0.165 Wind16point_NNW + 0.088 Wind16point_NW
- 0.293 Wind16point_S - 0.003 Wind16point_SE + 0.251 Wind16point_SSE
- 0.2637 Wind16point_SSW - 0.1936 Wind16point_SW + 0.1715 Wind16point_W
+ 0.0365 Wind16point_WNW + 0.0952 Wind16point_WSW + 0.0 Weathercode_113
+ 0.2582 Weathercode_116 + 0.5934 Weathercode_119 + 0.557 Weathercode_122
+ 0.478 Weathercode_143 + 2.303 Weathercode_176 + 6.51 Weathercode_179
- 0.47 Weathercode_182 + 0.221 Weathercode_200 - 0.109 Weathercode_227
+ 0.506 Weathercode_230 + 0.684 Weathercode_248 + 0.890 Weathercode_260
- 0.129 Weathercode_263 + 0.034 Weathercode_266 + 0.62 Weathercode_281
- 1.365 Weathercode_284 - 0.205 Weathercode_293 + 0.394 Weathercode_296
+ 4.597 Weathercode_299 + 0.187 Weathercode_302 + 3.401 Weathercode_305
+ 1.718 Weathercode_308 + 1.958 Weathercode_311 - 0.456 Weathercode_314
+ 0.172 Weathercode_317 + 0.564 Weathercode_320 + 1.408 Weathercode_323
+ 0.803 Weathercode_326 + 1.383 Weathercode_329 + 0.587 Weathercode_332
+ 0.259 Weathercode_335 + 0.525 Weathercode_338 + 0.246 Weathercode_350
+ 0.560 Weathercode_353 + 1.163 Weathercode_356 + 0.219 Weathercode_359
+ 3.26 Weathercode_362 + 4.949 Weathercode_368 + 3.544 Weathercode_371
+ 1.961 Weathercode_386 - 0.611 Weathercode_389

The selection of a predictive model lies in minimizing the expected mean square error (MSE) of the candidate models, where the MSE is the square of the difference between the true value and the predicted response of the given data. The MSE is composed of three terms: the variance, the squared bias, and the square of the inherent error. The inherent error is irreducible; the variance and bias are reducible errors. Variance is the extent of change in a predictive model if a different dataset was used for training the model. It is common that training a predictive model on different datasets gives different MSE values. However, the difference is not significant if the model has low variance. For a model that has high variance, small change in training data will cause large shift in the trajectory of the predictive model. High variance in a model is generally due to overfitting the dataset by using a flexible predictive model such as high degree polynomial regression. On the contrary, bias error is the reflection of the use of an overly simplified rigid predictive model. High bias is related to a large gap between the predictive model response and the true value of the response. Minimizing the bias-variance trade-off is not a trivial task since bias decreases as variance increases and vice-versa. Resampling methods such as cross-validation techniques are commonly deployed to evaluate the MSE by balancing the trade-offs.

Experiments are run on four cross-validation approaches and six predictive models. The predictive models are polynomial regressions of degree 1, 2, 3, 4, 5, and 6. The cross-validation approaches are leave one out cross validation (LOOCV), 2-fold, 5-fold, and 10-fold. In LOOCV, all but one sample are used to train the predictive model; the model is tested on the left-out sample. The samples are divided into halves in 2-fold approach, where one-half is used to train the model and the other half is for testing. In 5-fold, the samples are divided into five equal portions; 80% is used to train the model and 20% for testing. The 80% to 20% approach is repeated until all folds are covered. The 10-fold approach is similar to 5-fold other than that 90% is used for training and 10% for testing.

Figure 2 and 3 depicts the performance of the expected MSE experimental results on the test datasets. An MSE result data point in each graph represents the expected value of a predictive model's performance on a test dataset; the training and testing datasets are randomly selected for each run. The predictive model developed (trained) by 2-fold method gives smaller variance and high bias when tested on datasets that are different from the training sets, since the method is trained on smaller training dataset compared to the other three approaches. On the contrary, the predictive model trained by LOOCV method shows high variation and low bias as shown in the top chart of Figure 2. The predictive models trained by the 5-fold and 10-fold methods provide a balanced variance-bias tradeoffs as shown by bottom of Figure 2. From the top chart of Figure 3, the variance of the 5-fold is lower than the LOOCV, while the bias is the other way around. The high variance is due to high correlation between the training samples for the different random runs since the training data are identical except for one sample which is left out for testing. The behavior of the 5-fold and the 10-fold is essentially similar as can be seen in right bottom corner chart. The 5-fold shows higher variance and lower bias compared to 2-fold as shown by the bottom of Figure 3. The comparative summary of the approaches is given in Table 1 below.

Table 1: Comparative Summary of Machine Learning Approaches

Size of training data	LOOCV	>	10-fold / 5-fold	>	2-fold
Size of testing data	LOOCV	<	10-fold / 5-fold	<	2-fold
Bias	LOOCV	<	10-fold / 5-fold	<	2-fold
Variance	LOOCV	>	10-fold / 5-fold	>	2-fold

It is imperative to select either 5-fold or 10-fold to train a predictive model to balance the trade-off between bias and variance. The 10-fold approach is selected for the purpose of this paper.

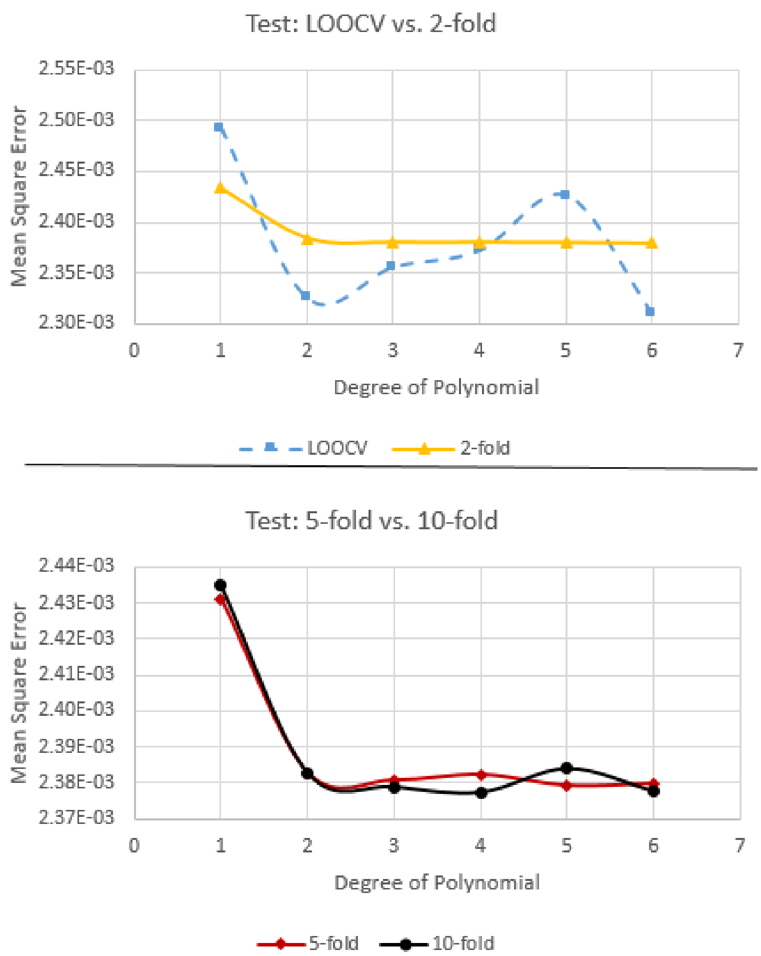

Figure 2: Cross-validation approaches MSE comparisons of regression polynomial predictive models (part 1)

The MSE of the 10-fold method decreases rapidly from the polynomial of degree one (linear regression) to the polynomial of degree two (quadratic regression) as shown by the solid black line. There is not much improvement on the MSE regardless of the increase in the degree of the polynomial from two to six. Therefore, quadratic regression predictive model can be used as the minimizer of the MSE in this experiment.

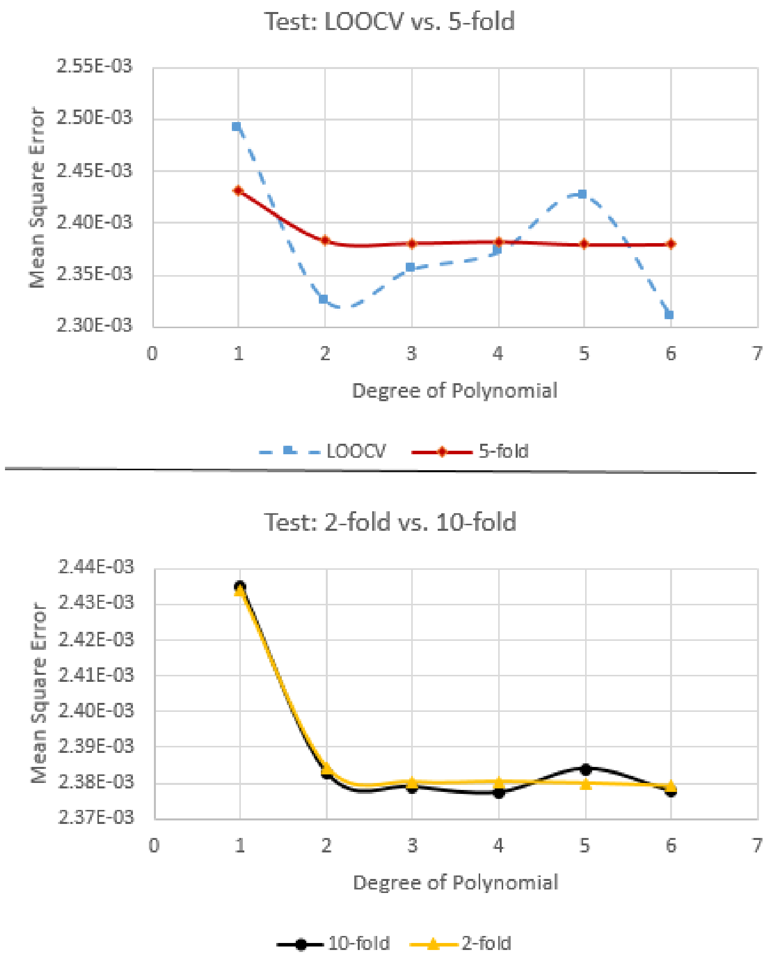

Figure 3: Cross-validation approaches MSE comparisons of regression polynomial predictive models (part 2)

4. Conclusion

The preliminary findings show a decreased error when using the new approach (near-future data) in comparison to the traditional approach (historical data) for predicting solar energy generation. Findings show near-future (5-7 days) weather prediction data associated with local weather stations can be used in combination with machine learning techniques to accurately estimate potential solar energy generated electricity disruptions to the grid. These results show positive impacts towards electricity forecasting, which can reduce stakeholder costs in high solar energy penetration areas.

Energy generation, sources and distribution methods have been continuously evolving over the past decade. In 2019, solar-produced energy accounted for 1.8% of the electricity generated in the U.S., up from 0.11% in 2012. In 2020, solar-produced energy is forecasted to increase by 20% (EIA, 2020). As the adoption of solar energy increases, so will potential disruptions to the grid. These preliminary findings show the potential for aggregating individual site-specific predictions to the regional level for the purpose of estimating area-specific solar energy disturbances and moving efforts towards predictive grid optimization.

References

Arbizu-Barrena, C., Ruiz-Arias, J. A., Rodríguez-Benítez, F. J., Pozo-Vázquez, D., & Tovar-Pescador, J. (2017). Short-term solar radiation forecasting by advecting and diffusing MSG cloud index. *Solar Energy, 155*, 1092-1103.

Bin Shams, M., Haji, S., Salman, A., Abdali, H., & Alsaffar, A. (2016). Time series analysis of Bahrain's first hybrid renewable energy system. *Energy, 103*, 1-15. doi:10.1016/j.energy.2016.02.136

Bosman, L. (2014). A Decision Support System to Analyze, Predict, and Evaluate Solar Energy System Performance: PVSysCO (photovoltaic System Comparison).

Bosman, L., & Darling, S. (2016, 20-23 Nov. 2016). *Difficulties and recommendations for more accurately predicting the performance of solar energy systems during the snow season.* Paper presented at the 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA).

Bosman, L. B., & Darling, S. B. (2018). Performance modeling and valuation of snow-covered PV systems: examination of a simplified approach to decrease forecasting error. *Environmental Science and Pollution Research*, 1-8.

Bosman, L. B., Leon-Salas, W. D., Hutzel, W., & Soto, E. A. (2020). PV System Predictive Maintenance: Challenges, Current Approaches, and Opportunities. *Energies, 13*(6), 1398.

Chaabene, M., & Ben Ammar, M. (2008). Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems. *Renewable Energy, 33*(7), 1435-1443. doi:10.1016/j.renene.2007.10.004

Darling, S., You, F., Veselka, T., & Velosa, A. (2011). Assumptions and the levelized cost of energy for photovoltaics. *Energy & Environmental Science, 4*(9), 3133-3139.

Dobos, A. P. (2014). *PVWatts version 5 manual*. Retrieved from

DSIRE. (2020). *Net Metering*. Retrieved from https://www.dsireusa.org/

EIA. (2020). What is U.S. electricity generation by energy source? Retrieved from https://www.eia.gov/tools/faqs/faq.php?id=427&t=3

Ferreira, P. M., Gomes, J. M., Martins, I. A. C., & Ruano, A. E. (2012). A Neural Network Based Intelligent Predictive Sensor for Cloudiness, Solar Radiation and Air Temperature. *Sensors, 12*(11), 15750-15777. doi:10.3390/s121115750

Ghiani, E., Pilo, F., & Cossu, S. (2013). Evaluation of photovoltaic installations performances in Sardinia. *Energy Conversion and Management, 76*, 1134-1142. doi:10.1016/j.enconman.2013.09.012

Hassan, G. E., Youssef, M. E., Mohamed, Z. E., Ali, M. A., & Hanafy, A. A. (2016). New Temperature-based Models for Predicting Global Solar Radiation. *Applied Energy, 179*, 437-450. doi:10.1016/j.apenergy.2016.07.006

Hernandez, J., Gordillo, G., & Vallejo, W. (2013). Predicting the behavior of a grid-connected photovoltaic system from measurements of solar radiation and ambient temperature. *Applied Energy, 104*, 527-537. doi:10.1016/j.apenergy.2012.10.022

Hong, T., Change, W.-K., & Lin, H.-W. (2013). A fresh look at weather impact on peak electricity demand and energy use of buildings using 30-year actual weather data. *Applied energy, 111*(November 2013), 333-350.

Husein, M., & Chung, I. Y. (2019). Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach. *Energies, 12*(10), 21. doi:10.3390/en12101856

Ju, X., Vossier, A., Wang, Z. F., Dollet, A., & Flamant, G. (2013). An improved temperature estimation method for solar cells operating at high concentrations. *Solar Energy, 93*, 80-89. doi:10.1016/j.solener.2013.02.028

Kalogirou, S. A. (2001). Artificial neural networks in renewable energy systems applications: a review. *Renewable & Sustainable Energy Reviews, 5*(4), 373-401. doi:10.1016/s1364-0321(01)00006-5

Krishnamurti, T. N. (1995). Numerical weather prediction. In J. L. Lumley & M. Van Dyke (Eds.), *Annual Review of Fluid Mechanics* (Vol. 27, pp. 195-224): Annual Reviews Inc. {a}, P.O. Box 10139, 4139 El Camino Way, Palo Alto, California 94306, USA.

Mellit, A., & Pavan, A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy, 84*(5), 807-821. doi:10.1016/j.solener.2010.02.006

Obando, E., Carvajal, S., & Pineda, J. (2019). Solar Radiation Prediction Using Machine Learning Techniques: A Review. *Ieee Latin America Transactions, 17*(4), 684-697. doi:10.1109/tla.2019.8891934

Ramakrishna, R., Bernstein, A., Dall'Anese, E., Scaglione, A., & Ieee. (2018). Joint Probabilistic Forecasts of Temperature and Solar Irradiance. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (Icassp)*, 3819-3823. Retrieved from <Go to ISI>://WOS:000446384603197

Reikard, G., & Hansen, C. (2019). Forecasting solar irradiance at short horizons: Frequency and time domain models. *Renewable Energy, 135*, 1270-1290. doi:10.1016/j.renene.2018.08.081

Routh, T. K., Bin Yousuf, A. H., Hossain, M. N., Asasduzzaman, M. M., Hossain, M. I., Husnaeen, U., . . . Ieee. (2012). *Artificial Neural Network Based Temperature Prediction and Its Impact On Solar Cell.*

TamizhMani, G., Ji, L., Tang, Y., Petacci, L., & Osterwald, C. (2003). *Photovoltaic module thermal/wind performance: long-term monitoring and model development for energy rating.* Paper presented at the

NCPV and Solar Program Review Meeting Proceedings, 24-26 March 2003, Denver, Colorado (CD-ROM).
West, S. R., Rowe, D., Sayeef, S., & Berry, A. (2014). Short-term irradiance forecasting using skycams: Motivation and development. *Solar Energy, 110*, 188-207.
Wilcox, S., & Marion, W. (2008). *Users Manual for TMY3 Data Sets, Technical Report NREL/TP-581-43156*. Retrieved from
Zeng, J. W., & Qiao, W. (2013). Short-term solar power prediction using a support vector machine. *Renewable Energy, 52*, 118-127. doi:10.1016/j.renene.2012.10.009

Biographies

Jose Ramirez is an industrial engineer who graduated from the University of Magdalena in Santa Marta, Colombia. He is currently pursuing his master's degree in engineering technology at Purdue University. His experience includes working at the Office of Intellectual Property and Knowledge Transfer, and the Centre for Entrepreneurship and Innovation, both at the University of Magdalena in Colombia. With renewable energy and sustainability as his main research interests, Ramirez has worked across multiple subjects that include reviews and analysis of renewable energy sources, including microalgae, piezoelectric material, and solar energy; his work also includes mathematical modeling and optimization. Current research projects focus on applying machine learning models to forecast climate parameters (thesis research through Purdue University) and using machine learning to estimate power systems planning against rising temperatures (research internship at Arizona State University).

Esteban Soto has a master's degree in Industrial Management and a bachelor's degree in Industrial Engineering from the University of Concepcion, Chile. He is currently pursuing a Ph.D. in Technology at Purdue University. Before beginning his doctoral studies, he worked at the Ministry of Energy of Chile in charge of the implementation of the Energetic Efficiency program in the Biobio Region. Soto is a co-founder and former CEO of a startup that works on the development of new technologies to reduce emissions of particulate matter produced by the combustion of biomass. His research areas of interest include renewable energy, solar energy, and new models for exchanging energy among peers.

Dr. Ebisa Wollega is an assistant professor of engineering at Colorado State University Pueblo**.** Dr. Wollega's research areas include large scale optimization, AI algorithms, and data analytics with applications in energy systems. Dr. Wollega serves as an officer in American Society for Engineering Education (ASEE) Industrial Engineering Division.

Dr. Lisa Bosman, PhD in Industrial Engineering, is an Assistant Professor within the Purdue Polytechnic Institute (formerly, the College of Technology) at Purdue University. Dr. Bosman's engineering research focuses on the development of information systems to enable the integration of grid optimization, solar energy performance modeling, and decision making. Prior to joining higher education, Dr. Bosman spent several years working in industry as a manufacturing engineer with well-known companies including Harley-Davidson Motor Company, John Deere, and Oshkosh Truck. Dr. Bosman has authored over 50 publications in international and national journals and conferences. In addition, she has obtained over $1M USD in research funding from agencies including the National Foundation (NSF), Environmental Protection Agency (EPA), and the National Aeronautics and Space Administration (NASA). She has been an invited speaker and workshop facilitator for over 20 engagements.