# Improving Solar PV Scheduling using Statistical Techniques

by

Dhiwaakar Purusothaman Soundiah Regunathan Rajasekaran

A Thesis Presented in Partial Fulfilment
of the Requirements for the Degree
Master of Science

Approved April 2016 by the
Graduate Supervisory Committee:

Nathan G. Johnson, Co-Chair
George G. Karady, Co-Chair
Raja Ayyanar, Member

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

The inherent intermittency in solar energy resources poses challenges to scheduling generation, transmission, and distribution systems. Energy storage devices are often used to mitigate variability in renewable asset generation and provide a mechanism to shift renewable power between periods of the day. In the absence of storage, however, time series forecasting techniques can be used to estimate future solar resource availability to improve the accuracy of solar generator scheduling. The knowledge of future solar availability helps scheduling solar generation at high-penetration levels, and assists with the selection and scheduling of spinning reserves. This study employs statistical techniques to improve the accuracy of solar resource forecasts that are in turn used to estimate solar photovoltaic (PV) power generation. The first part of the study involves time series forecasting of the global horizontal irradiation (GHI) in Phoenix, Arizona using Seasonal Autoregressive Integrated Moving Average (SARIMA) models. A comparative study is completed for time series forecasting models developed with different time step resolutions, forecasting start time, forecasting time horizons, training data, and transformations for data measured at Phoenix, Arizona. Approximately 3,000 models were generated and evaluated across the entire study. One major finding is that forecasted values one day ahead are near repeats of the preceding day—due to the 24-hour seasonal differencing—indicating that use of statistical forecasting over multiple days creates a repeating pattern. Logarithmic transform data were found to perform poorly in nearly all cases relative to untransformed or square-root transform data when forecasting out to four days. Forecasts using a logarithmic transform followed a similar profile as the immediate

day prior whereas forecasts using untransformed and square-root transform data had smoother daily solar profiles that better represented the average intraday profile. Error values were generally lower during mornings and evenings and higher during midday. Regarding one-day forecasting and shorter forecasting horizons, the logarithmic transformation performed better than untransformed data and square-root transformed data irrespective of forecast horizon for data resolutions of 1-hour, 30-minutes, and 15-minutes.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

## Chapter 1. Introduction

Renewable power generation capacity in the United States is expected to increase by 25% between 2013 and 2018 [EIA Annual Energy Outlook 2015]. Solar photovoltaic (PV) resources are among the fastest growing renewable source of energy. Currently, solar PV increases at an annual average rate of 6.8% [EIA Annual Energy Outlook 2015]. The installed capacity of solar photovoltaics (PV) has been steadily increasing across the United States with an increase of more than 54% in 2014 alone [NREL data book 2014]. Utility-scale solar PV capacity is expected to increase by 123% between 2014 and 2016 and [EIA STEO report 2015]. Concentrating solar power (CSP) units are also becoming more prevalent, with the largest annual increase in capacity of 767 MWac recorded in 2014 [SEIA report 2015]. The growth in solar energy utilization has been prompted by several factors including a decrease in solar PV module costs, improvements in solar-storage for CSP, greater investment in renewables, tax incentives, renewable energy mandates, and a shift in society's viewpoint towards low-carbon energy [EIA Today in Energy 2015; Energy.gov solar 2015; NREL Renewable Portfolio Standards 2015; SEIA solar ITC 2015; Sunshot PV system pricing trends 2014; U.S. DOE CSP report 2014; U.S. DOE Investing in American Energy report 2015]. In 2016, the utility-scale solar generation is expected to average 89 GWh/day in the United States [EIA STEO report 2015].

Solar electricity is prone to intermittency arising from fluctuations in the solar resource due to cloud cover. Forecasting techniques employed to estimate future variability in solar resources can aid in planning and scheduling resources and spinning reserves to minutes or hours in advance. This helps to balance renewable energy supply and end-user

demand without use of excessive amounts of storage, demand response, or renewables curtailment. Power electronic technologies that can be deployed on a larger-scale are required to make the power grid more flexible and reliable to include more renewable energy generation in to the network [World Energy Outlook 2012].

## Chapter 2: Time Series Forecasting of GHI using ARIMA models

### 2.1 Introduction

Solar power generation has increased steadily over the past decade in the United States [EIA STEO report 2015]. In 2014 alone, solar photovoltaic (PV) generation capacity increased by 30% (6200 MW) and concentrating solar power (CSP) generation capacity increased by 54% (767 MW) [NREL 2014, SEIA report 2015]. This growth has been driven by several factors including lower solar PV module costs, advanced CSP technology such as thermal storage and dish-engine, greater investment in renewables, favorable tax incentives, renewable energy mandates, net metering requirements, and a gradual shift in public opinion towards low-carbon energy [EIA Today in Energy 2015; Energy.gov solar 2015; SEIA Solar ITC 2015; Sunshot PV System Pricing Trends 2014; U.S. DOE CSP Report 2014; U.S. DOE Investing in American Energy Report 2015]. Recent extensions of renewable energy tax credits originally granted under the Energy Policy Act of 2005 are expected to continue and thus increase the rate of solar PV capacity expansion through 2019 and beyond [Energy Policy Act 2005]. Estimates for 2016 indicate that utility-scale solar generation—PV and CSP—in the United States is expected to generate an average of 89 GWh/day, or about 0.8% of total energy supply [EIA STEO Report 2015; SEIA Report 2016].

The transition to higher penetration solar power in the United States energy portfolio faces several technical and economic challenges for existing infrastructure and business models. Ongoing technical concerns include voltage regulation, protection co-ordination, distributed energy resource (DER) control, and power quality control [Perez and Fthenakis 2013; Thongpron et al. 2004; Brisette et al. 2013; Martin et al. 2010; Hanna

et al. 2014; U.S. DOE High Penetration Report 2009; SEIA RES Report 2015]. In power electronics and controls, new devices and modifications are needed to mitigate real-time line voltage variations resulting from higher penetration of renewables [Perera et al 2013]. Unit commitment, contingency analysis, and economic dispatch procedures are also being affected by increasing amounts of uncontrolled DER, thus requiring significant adjustments to existing practices. Thermal power plants, for example, will need to remain online and operate at part-load in order to provide on-demand power in the event of clouding or other disturbances [NREL Hawaii Study 2013]. This has resulted in short periods of excess power that cannot be used within a utility service region. Multiple instances of negative electricity pricing have occurred in Germany during periods of high solar PV and wind power output yet load electrical demand [Fraunhofer 2015]. Similar concerns are being voiced in the United States regarding the "duck curve" that estimates a 13,000 MW increase in dispatchable generation of power over a three-hour period as solar PV drops off and consumer demand increases in the late afternoon [CalISO 2013]. Renewable intermittency also poses a challenge for transmission and distribution scheduling. The Federal Energy Regulatory Commission (FERC) requires that electric utilities provide 15-minute scheduling with a degree of certainty that is affected by uncontrollable DER [FERC scheduling report 2012]. Energy storage is one solution to improve power quality, provide operating reserve, and shift power between on-sun and off-sun periods of the day [NREL Impacts of Solar Power report 2012]. Yet widespread adoption of storage technologies may be several years away following further price reductions [U.S. DOE Quadrennial Technology Review 2015].

Advanced knowledge of solar resource availability may help mitigate power quality disturbances, improve scheduling, avoid negative power pricing, and reduce the cost of delivered electricity. A study by Martin et al. demonstrated that forecasted global horizontal irradiance (GHI) data could be used to improve estimates of future solar PV power output [Martin et al. 2010]. Forecasted data can also be used for optimizing energy dispatch by the utility to improve demand side management and load balancing [Hanna et al 2014]. Similarly, weather and atmospheric modelling, sky cameras, radar and lidar, as well as statistical methods can be used to estimate near-term solar resources [Wang F. et al. 2012; Barrett A. et al. 2012; Graham et al. 1988; Kang and Tam 2015; Heinemann et al. 1999]. This study uses statistical techniques and time series data to forecast GHI values that are in turn used to estimate solar PV power output.

**2.2 Time Series Forecasting**

Time series statistical forecasting studies commonly use the clearness index, GHI, and direct normal irradiation (DNI) as operational data [Heinemann et al. 1999; Maimouna et al. 2013; Mellit et al. 2005; Reikard et al. 2009; Sozen et al. 2005; Wang F. et al. 2012]. Most published works complete a single case study of a particular location and use a single data input resolution with one-hour time steps [Sozen et al. 2005; Heinemann et al. 1999]. These studies have been expanded through the inclusion of environmental factors such as ambient temperature and relative humidity that improve the accuracy of model predictions [Li et al. 2011]. A comprehensive summary of forecasting methods provided by Reikard in 2009 compared the performance of time series methods, neural networks, and hybrid models using data from six locations. Statistical forecasting methods had the least amount of error in most tests. A logarithmic transformation of the initial data gave further

reductions in model error [Reikard 2009]. This performance of models after logarithmic transformation with 60-minute and 30-minute resolution, however, did not hold with 15-minute or 5-minute data. Neural networks tended to perform better using higher-resolution data and a transfer function that incorporated causal inputs such as humidity or temperature. Hybrid models implementing a combination of methods also performed better than time series forecasting using 15-minute or 5-minute data. Reikard's study forecasted values out four hours and the error for ARIMA models reduced by 6% for Phoenix. This study, however, did not seek to examine if statistical models and regression coefficients change for a single location using training data from different periods of the year. A parallel line of inquiry yet untouched is an investigation of forecast performance using models and regression coefficients developed from training data taken from another period of the year, e.g., how does a model developed using data from May perform when applied in December? Another area left for further study was how the forecasting start time (e.g., 6:00 am, 12:00 pm) effected model error.

This article presents a comparative study of time series forecasting models developed with different time step resolutions, forecasting start time, forecasting time horizons, training data, and transformations for data measured at Phoenix, Arizona. In this study, forecasting algorithms were applied to 5-minute, 15-minute, 30-minute, and 60-minute solar resource data to examine how real-time solar resource data with various fidelities can be used to improve knowledge of the future solar resource over a 24-hour period. Similar analyses were completed with errors examined for one-step ahead across all hours of a one-day period. The effect of forecasting start time was explored by beginning

forecasts at 12:00 am, 6:00 am, 12:00 pm, and 6:00 pm. Data transformations included a square-root transformation and a log transform. Time series models were also generated using training data for each month of the year and applied to other months. The purpose of these models was to examine how fitting parameters and coefficients change over the year. Solar resource data was used from Phoenix, Arizona, in the United States with a measurement resolution of 1-minute [MIDC/NREL].

### *2.2.1 Time Series Regression and Forecasting Methods*

Regression methods are applied to a set of training data to generate a model that describes how solar irradiance changes over time. This model is then used to forecast solar irradiance in future time steps.

A stationary process is a process whose characteristics do not change with time. An auto-covariance function is used to identify the variation in a variable between two points in time. The two basic forms of time series models include an autoregressive model (AR) and a moving average model (MA). An AR model is suited to time series data that are linearly dependant on historical values, whereas an MA can be used to describe the behavior of time series data that is a function of the mean value from current and previous white noise error terms. The extent to which these models are dependent on the historical values are represented using their corresponding time step lags $p$ and $q$.

Time series data that exhibit stationarity and have a decreasing auto-covariance function are suitable for forecasting using Autoregressive Moving Average (ARMA) models [Brockwell 2006]. An ARMA($p,q$) model can be represented as

$$\phi_p(B)X_t = \theta_q(B)Z_t \quad (1)$$

$$\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p \quad (2)$$

$$\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q \quad (3)$$

where $\phi(B)$ represents the autoregressive (AR) portion of the model, $\theta(B)$ represents the moving average (MA) portion of the model, and $B$ stands for the backshift operator $B^j x_t = x_{t-j}$. The process $X_t$ is the time series data on which the forecast is performed. $X_t$ is defined here as an autoregressive process of order $p$ and a moving average process of order $q$ [Brockwell 2006; Wei 1990]. The term $Z_t$ represents a white noise process with $\{Z_t\} \sim \text{WN}(0,\sigma^2)$.

If either condition for an ARMA model is not satisfied then differencing can be applied to seek a process with stationarity and decreasing auto-covariance. Differencing methods are used in Autoregressive Integrated Moving Average (ARIMA) models. An ARIMA($p,d,q$) model for the process $X_t$ can be expressed as

$$\phi_p(B)(1-B)^d X_t = \theta_q(B)Z_t \quad (4)$$

where $d$ is the order of differencing used to process input data and trending in the time series data [Brockwell book 2006; Wei 1990]. An ARIMA model with no differencing ($d = 0$) simplifies to an ARMA($p,q$) model.

Different ARIMA models can be obtained by choosing different values of $p$, $d$, and $q$. Seasonal Autoregressive Integrated Moving Average (SARIMA) models can also be obtained using differencing methods across multiple time steps consistent with periodic

data. These periodic data repeat over a consistent interval such that differencing over that interval produces a consistent trend. For example, hourly outdoor ambient temperature data follows a fairly consistent pattern that repeats daily. Solar irradiation data follows the same pattern. A SARIMA model can be represented using the following equation

$$\Phi_P(B^s)\phi_p(B)Y_t = \Theta_Q(B^s)\theta_q(B)Z_t \tag{5}$$

where

$$Y_t = (1-B)^d(1-B^s)^D X_t \tag{6}$$

and $D$ is the seasonal order of differencing, $P$ is the seasonal autoregressive factor, and $Q$ is the seasonal moving average factor.

Model goodness of fit in this study is evaluated using the Akaike Information Criterion (AIC). The AIC is generally used to compare models with varying numbers of explanatory variables [Brockwell 2006]. This is because the $R^2$ value always increases with the addition of more explanatory data, even if the data is random and no physical significance. In contrast, the AIC is penalized by the number of parameters in the model

$$\text{AIC} = -\ln(L) + 2m \tag{7}$$

where, $L$ is the maximized likelihood function of the particular model and $m$ is the number of parameters in the model given by $(p + q + 1)$. Lower values of AIC indicate models with better fit.

Forecasting with a SARIMA process is completed using Eq. (8) with predictor $P_n$ for $h$ time steps into the future using coefficients from the regression analysis as follows:

$$P_n X_{n+h} = P_n Y_{n+h} + \sum_{j=1}^{d+Ds} a_j P_n X_{n+h-j} \quad (8)$$

## 2.3 Study Parameters and Data

### *2.3.1 Study Parameters*

A comparative study was completed of time series forecasting models developed using different time step resolutions, data transformations, forecasting start time, forecasting time horizons, and training data.

- Time step resolutions—The 1-minute input data was averaged into 5-minute, 15-minute, 30-minute, and 1-hour data.
- Data transformations—Two data transformations were applied to the input data to improve model stationarity using a square-root transformation and a natural logarithmic transformation.
- Forecasting start time—The simulation was run for solar irradiation data starting at different times of the day including 12:00 am, 6:00 am, 12:00 pm, and 6:00 pm.
- Forecast time horizon—Forecasting was completed out to periods of 5 minutes, 15 minutes, 30 minutes, 1 hour, and 24 hours.
- Training data—Model parameters and coefficients developed using training data from one period of the year (e.g., May) were applied to other periods of the year (e.g., December). This was repeated allowing the coefficients to change but keeping the same model parameters.

*2.3.2 Model Data*

High-resolution data available at 1-minute intervals were averaged to obtain 5-minute data, 15-minute data, 30-minute data, and 1-hour data as shown in Fig. 1.

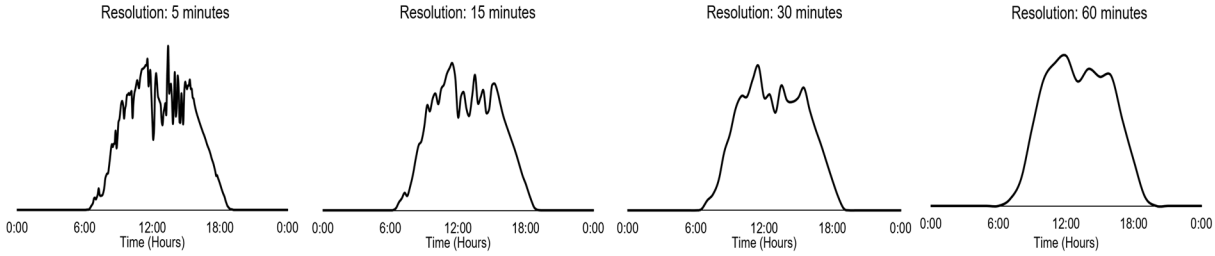


**Figure 1. Illustration of GHI averaging and smoothing for various time resolutions.**

*2.3.3 Performance Metrics*

Performance metrics used in this study include the AIC and other standard values such as the root mean square error (RMSE) and the mean percentage error. The formula for calculating the RMSE is given below.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{y}_t - y_t)^2}{n}} \tag{9}$$

where

$\hat{y}$ : measured data point

$y$ : forecasted data point

$n$ : number of data points

***2.4* Results and Analysis**

Time series regression methods were completed with parameters including $p = 2$, $d = 1$, $q = 2$, $P = 1$, $D = 1$, and $Q = 1$. This equates to 143 models tested for every time step resolution, data transformation, forecast start time, forecast horizon, and training data with

the single best fit model introduced in the results. A total number of approximately 3,000 models were generated and evaluated across the entire study.

*2.4.1 Comparison of GHI Forecasts Out Multiple Days in a Single Month*

Forecast simulations were first completed using different time step resolutions of input data as graphed in Fig. 3. The simulations were run for seven days of data (1st April – 07th April 2012) forecasted out to four days in the future (forecast horizon). The coarseness in historical and forecasted data decreases is smoothed with the averaging process, as expected. One interesting finding is that the 5-minute and 15-minute forecasted data gave higher peak GHI values for the selected data. This occurs because the intermittency during the seventh-day is smoothed during the averaging process and thus decreases the higher peak values observed. Yet the most influential finding is that forecasted values out past one day are near repeats of the preceding day, indicating that any forecast beyond the seasonal difference (24 hours) simply resolves into a static or near-static average value for that time of day and is repeated for each day into the future. Table 1 provides the RMSE for these models with the SARIMA parameters listed in Table 2.

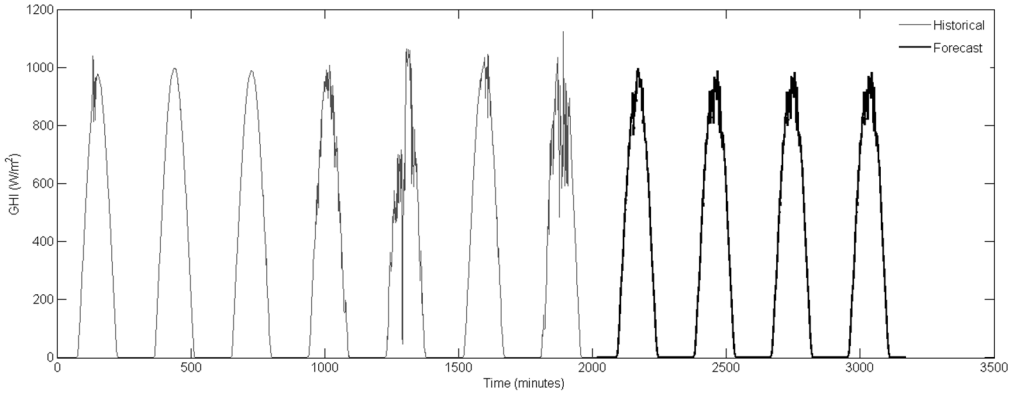


**Figure 2a. Four-day GHI forecast using 5-minute data (untransformed).**

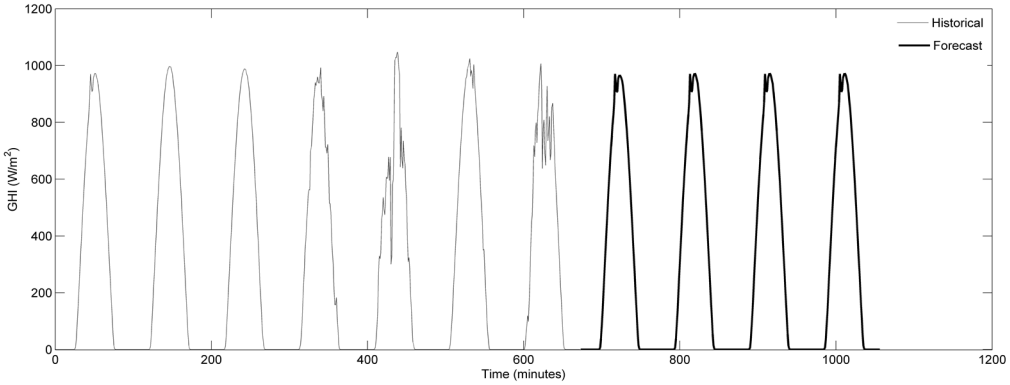


**Figure 2b. Four-day GHI forecast using 15-minute data (untransformed).**

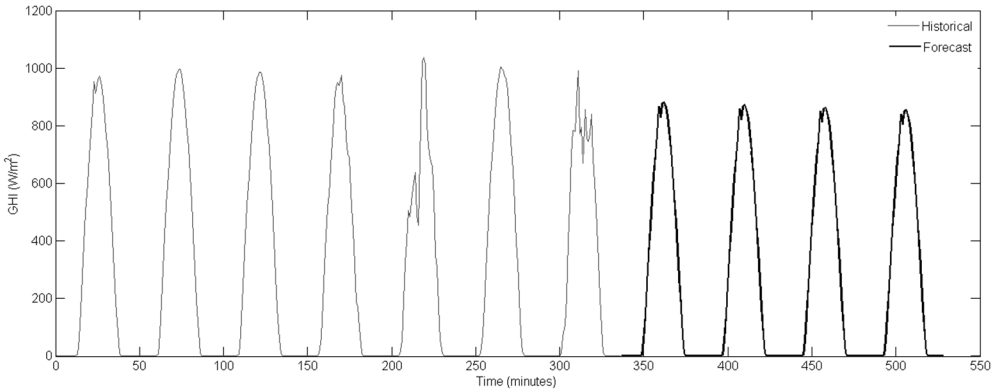


**Figure 2c. Four-day GHI forecast using 30-minute data (untransformed).**

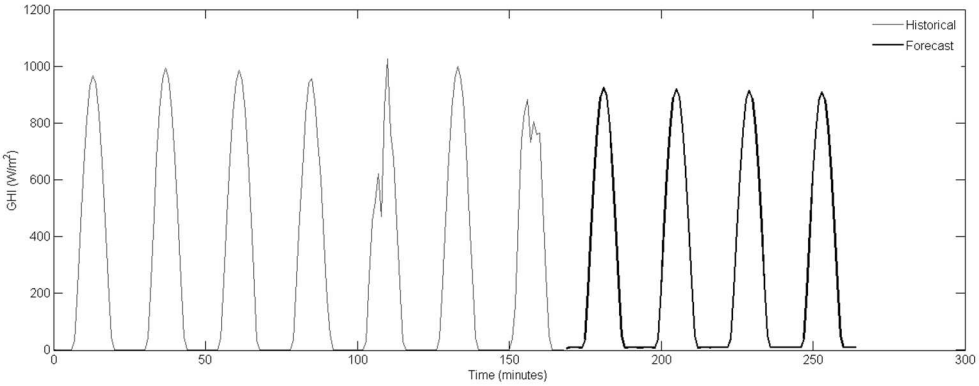


**Figure 2d. Four-day GHI forecast using 1-hour data (untransformed).**

Data transformations were then applied as discussed in section 3.1. Forecasting results obtained after square-root and logarithmic transformations are shown in Fig. 4a and Fig. 4b, respectively. The square-root transform (Fig. 4a) followed the same behavior as

the untransformed data (Fig. 3d) and smoothed-out the intermittency noticed in the last day before the forecasting horizon began. In contrast, the log transform forecasts (Fig. 4b) preserved the variability measured in the last day and replicated this variability across each day of the forecasting time horizon. Table 1 provides the RMSE for these models with the SARIMA parameters listed in Table 2.

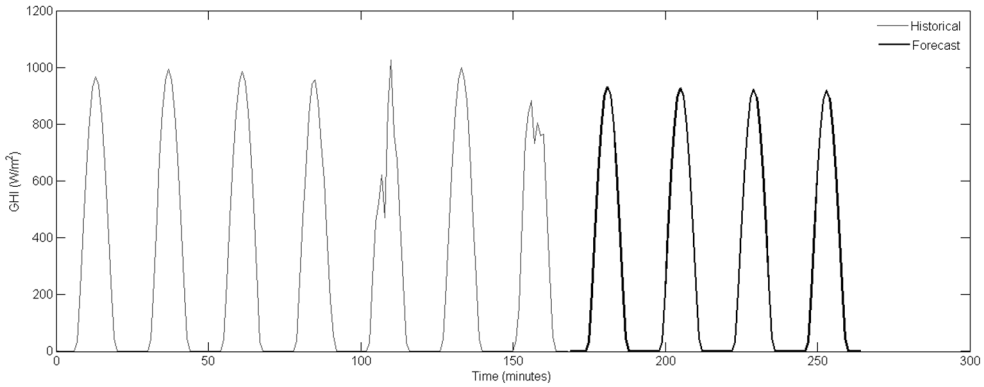


**Figure 3a. Four-day GHI forecast using 1-hour data (square-root transform).**

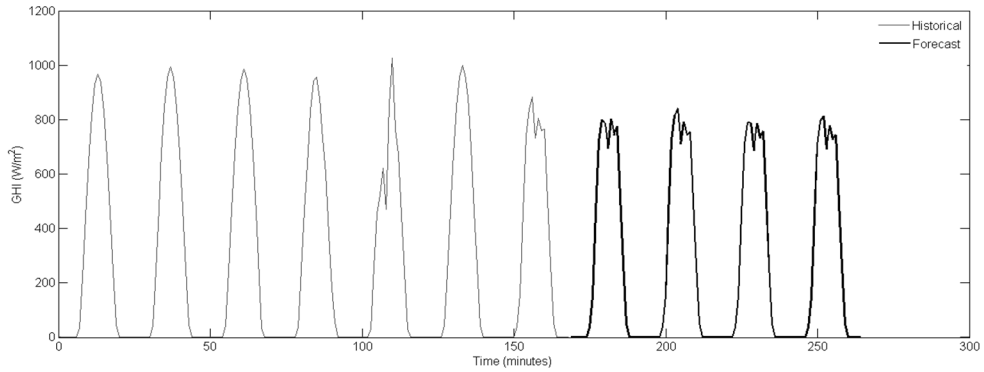


**Figure 3b. Four-day GHI forecast using 1-hour data (log transform).**

Simulations were also performed with the forecast beginning at different start times. Figure 5 shows these forecasts for different start times of 6:00 am, 12:00 pm, and 6:00 pm to compare with Fig. 3b starting at 12:00 am. It is clear that the seasonal behavior of the dataset is preserved regardless of the forecasting start time. One item of note is that

forecasts starting at 12:00 pm and 6:00 pm create small and continuing decreases in the shape of the forecasted daily profiles. This provides more evidence that longer-term forecasting outside of a single-day is unwise, particularly if transformations have not been applied to create a stationary process. Therefore the analyses of SARIMA models in future sections are completed with a maximum forecast horizon of one day. Table 1 provides the RMSE for these models with the SARIMA parameters listed in Table 2.

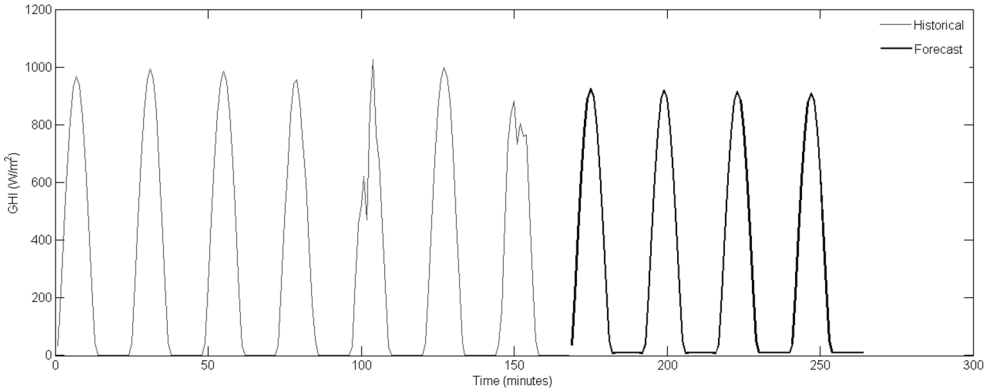


**Figure 4a. Four-day GHI forecast using 1-hour data at 6:00 am start (untransformed).**

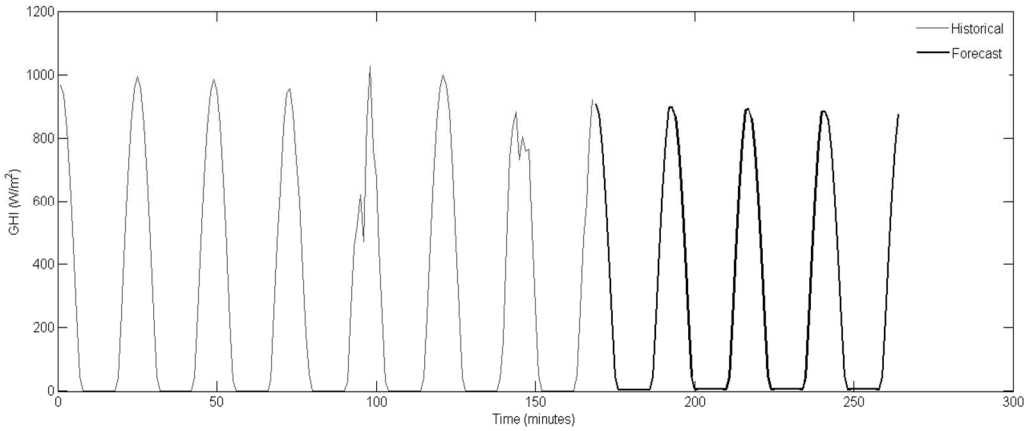


**Figure 4b. Four-day GHI forecast using 1-hour data at 12:00 pm start (untransformed).**

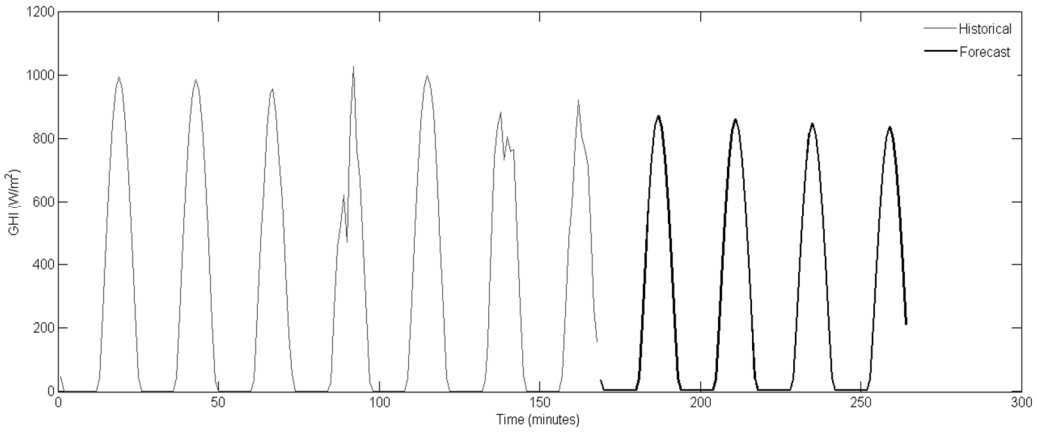


**Figure 4c. Four-day GHI forecast using 1-hour data at 6:00 pm start (untransformed).**

Table 1 summarizes the root mean square errors (RMSE) for time series data of different resolutions, start times, and data transformations. The RMSE tends to reduce with when using coarser data such as 1-hour time steps. This helps to explain the usage of hourly data for time series analysis. Transformations were observed to have a greater effect on the variability of RMSE between models created with higher-resolution data relative to models created from lower-resolution data. In general, models created from untransformed data gave smaller errors on a four-day forecast horizon when compared to transformed data, suggesting that the smoothed profile of untransformed data provides a better representation of the average daily solar profile across multiple days. This finding may not always hold, however, if a consistent intraday pattern is observed such as afternoon showers or morning fog that goes on for days, weeks, or months and that pattern would be lost if the entire day is smoothed into a consistent "ideal" profile as shown in Fig. 3d.

Table 2 lists the SARIMA model with the lowest AIC value from the 216 models evaluated for each study case. One observation is that for each data resolution, SARIMA

models generated using untransformed and square-root transformed data have the same number and type of fit parameters across all possible start times. However, models generated from log transform data do not yield a single model that is universally optimal across the start times tested. A deeper look at the fit parameters in Table 2 illustrates that the one-hour fits have a reduced incidence of a differencing term, suggesting that the time series data with shorter intervals is non-stationary or has increasing auto-covariance that can be resolved through single-lag or seasonal differencing. One clear finding is that no single model is optimal across all case studies evaluated, even though several similarities exist to form common subsets within the larger study.

**Table 1. Comparison of four-day GHI forecast results using RMSE ($W/m^2$).**

| Start Time | Time Step Resolution | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5-minute | | | 15-minute | | | 30-minute | | | 1-hour | | |
| | Untransformed | Square-root | Log | Untransformed | Square-root | Log | Untransformed | Square-root | Log | Untransformed | Square-root | Log |
| 12:00 AM | 128.97 | 125.93 | 386.50 | 116.87 | 116.62 | 145.11 | 79.96 | 110.88 | 130.48 | 85.84 | 88.06 | 125.51 |
| 6:00 AM | 122.44 | 144.77 | 221.37 | 117.49 | 140.21 | 133.21 | 79.89 | 88.52 | 123.99 | 85.87 | 88.00 | 108.80 |
| 12:00 PM | 116.69 | 105.41 | 149.95 | 136.07 | 134.22 | 122.48 | 123.77 | 103.31 | 114.71 | 78.28 | 78.61 | 110.07 |
| 6:00 PM | 90.24 | 100.70 | 231.40 | 74.21 | 91.88 | 279.33 | 104.93 | 126.21 | 107.92 | 77.70 | 79.04 | 109.22 |

**Table 2a. Parameters of the best fit SARIMA regression model evaluated using 5-minute and 15-minute data.**

| Start Time | Time Step Resolution | | | | | |
|---|---|---|---|---|---|---|
| | 5-minute | | | 15-minute | | |
| | Untransformed | Square-root | Log | Untransformed | Square-root | Log |
| 12:00 AM | (2,1,1)(1,1,1)[288] | (2,1,1)(1,1,1)[288] | (0,1,1)(1,0,1)[288] | (2,1,1)(1,1,1)[96] | (2,1,1)(1,1,1)[96] | (1,0,0)(1,0,0)[96] |
| 6:00 AM | (2,1,1)(1,1,1)[288] | (2,1,1)(1,1,1)[288] | (0,1,1)(1,0,1)[288] | (2,1,1)(1,1,1)[96] | (2,1,1)(1,1,1)[96] | (0,0,1)(1,1,1)[96] |
| 12:00 PM | (2,1,1)(1,1,1)[288] | (2,1,1)(1,1,1)[288] | (0,1,1)(1,0,1)[288] | (2,1,1)(1,1,1)[96] | (2,1,1)(1,1,1)[96] | (0,0,1)(1,1,1)[96] |
| 6:00 PM | (2,1,1)(1,1,1)[288] | (2,1,1)(1,1,1)[288] | (0,1,1)(1,0,1)[288] | (2,1,1)(1,1,1)[96] | (2,1,1)(1,1,1)[96] | (2,1,1)(1,1,1)[96] |



**Table 2b. Parameters of the best fit SARIMA regression model evaluated using 30-minute and 1-hour data.**

| Start Time | Time Step Resolution | | | | | |
|---|---|---|---|---|---|---|
| | 30-minute | | | 1-hour | | |
| | Untransformed | Square-root | Log | Untransformed | Square-root | Log |
| 12:00 AM | (0,1,1)(1,0,1)[48] | (0,1,1)(1,0,1)[48] | (2,0,1)(0,1,0)[48] | (0,0,1)(1,0,1)[24] | (0,0,1)(1,0,1)[24] | (0,0,1)(1,1,1)[24] |
| 6:00 AM | (0,1,1)(1,0,1)[48] | (0,1,1)(1,0,1)[48] | (0,0,1)(1,1,1)[48] | (0,0,1)(1,0,1)[24] | (0,0,1)(1,0,1)[24] | (0,0,1)(1,1,1)[24] |
| 12:00 PM | (0,1,1)(1,0,1)[48] | (0,1,1)(1,0,1)[48] | (0,0,1)(1,1,1)[48] | (0,0,1)(1,0,1)[24] | (0,0,1)(1,0,1)[24] | (2,0,1)(0,1,0)[24] |
| 6:00 PM | (0,1,1)(1,0,1)[48] | (0,1,1)(1,0,1)[48] | (0,0,1)(1,1,1)[48] | (0,0,1)(1,0,1)[24] | (0,0,1)(1,0,1)[24] | (1,0,0)(1,0,0)[24] |

*2.4.2 Comparison of GHI Forecasts Out a Single Day in Multiple Months*

The best SARIMA models developed for January, April, July, and October (one month in each quarter of the year) were applied across those same months to explore how a fit developed for one month—model parameters and coefficients—may fit the data observed in another month. Table 3 provides coefficients for 1-lag moving average, 24-lag seasonal autoregressive, and 24-lag seasonal moving average terms. The constant terms are clearly not statistically significant when noting that a single standard deviation will cross zero. The regression terms are significant and all models have the same parameters, except for the July regression that includes a single-lag differencing term. Table 4 takes another look shows how to use these models by applying the same model—parameters and coefficients—from one month to another. This exploration can help describe how well a model constructed using limited data from one time of the year might be applied to another time of the year in which data is unavailable. This was replicated in Table 5 but the model efficiencies were allowed to change. It was clear that allowing the model coefficients to change with the training data provided better results, as expected, and while it may be permissible to use the same model parameters across an entire year it is not permissible to use a single set of model coefficients across an entire year. The model the best fit July provided a higher RMSE compared to other models. This abnormality could have risen due to any inherent errors present in the data and is independent of the model itself.

**Table 3. SARIMA model coefficients based on minimum AIC.**

| Month Trained SARIMA Model | MA1 (std) | SAR24 (std) | SMA24 (std) | Constant (std) | Variance (std) |
|---|---|---|---|---|---|
| January (0,0,1)(1,0,1)[24] | 0.750 (0.164) | 0.997 (0.006) | -0.999 (0.099) | 0.076 (2.328) | 388.139 (19.248) |
| April (0,0,1)(1,0,1)[24] | 0.560 (0.039) | 0.994 (0.005) | -0.999 (0.051) | 0.794 (3.644) | 1820.020 (78.341) |
| July (0,1,1)(1,0,1)[24] | 0.444 (0.055) | 0.790 (0.027) | -0.412 (0.067) | 0.005 (6.513) | 5814.910 (431.779) |
| October (0,0,1)(1,0,1)[24] | 0.587 (0.037) | 0.987 (0.003) | -0.845 (0.035) | -0.894 (2.094) | 472.724 (21.372) |

**Table 4. Comparison of SARIMA model RMSE for selected months with model coefficients fixed by the originating month of training data.**

| Month Trained SARIMA Model | Month Applied | | | |
|---|---|---|---|---|
| | Jan | Apr | Jul | Oct |
| January (0,0,1)(1,0,1)[24] | 33.54 | 92.97 | 147.24 | 83.86 |
| April (0,0,1)(1,0,1)[24] | 33.55 | 87.89 | 152.51 | 77.86 |
| July (0,1,1)(1,0,1)[24] | 198.11 | 295.78 | 341.95 | 255.39 |
| October (0,0,1)(1,0,1)[24] | 42.18 | 88.40 | 138.38 | 55.91 |

**Table 5. Comparison of SARIMA model RMSE for selected months with model coefficients changing with training data from each month.**

| Month Trained SARIMA Model | Month Applied | | | |
|---|---|---|---|---|
| | Jan | Apr | Jul | Oct |
| January (0,0,1)(1,0,1)[24] | 33.55 | 87.86 | 124.99 | 55.90 |
| April (0,0,1)(1,0,1)[24] | 33.55 | 87.86 | 124.99 | 55.90 |
| July (0,1,1)(1,0,1)[24] | 33.32 | 85.39 | 341.90 | 57.13 |

| | | | | |
|---|---|---|---|---|
| October (0,0,1)(1,0,1)[24] | 33.55 | 87.86 | 124.99 | 55.90 |

*2.4.3 Comparison of GHI Forecasts of SARIMA Models across All Months*

The series of SARIMA models displayed in Table 6 are a subset of those evaluated for each case study application. The AIC and RMSE values corresponding to each model are given in Table 6 and Table 7, respectively. The lowest AIC and RMSE values are indicated in bold. It is interesting that while Model 7 provides the lowest AIC in most cases—best explanatory power for data used. Model 2 provides the least RMSE error for most of the models and hence greater accuracy. This rises a conflict between using AIC and RMSE for choosing the best fit model. Some of the months had models other than 7 and 2 that provided lesser AIC and RMSE respectively but it was also noted that those values were relatively closer to that of models 7 and 2. Hence, models 7 and 2 were chosen as best fit models based on AIC and RMSE respectively.

A scatter plot of the RMSE values for all 12 months of a year is plotted against the various SARIMA models in Fig. 6. January and December have the lowest RMSE for most of the models while September has the highest RMSE errors. This helps in analyzing the consistency in seasonality of the time series data and the effect of monthly data variations on model fitting. An important observation that can be made from this graph is that for all the 12 months, all the 11 SARIMA models have resulted in RMSE errors that are very close to each other. This justifies the usage of AIC factor to choose the best fit model instead of the RMSE thus ensuring higher accuracy. Also, it can be seen that the RMSE

errors for different months fall under one of the two ranges of values (0-50 and 70-120 $W/m^2$). This can be used to validate the performance of a model for any particular month based on the RMSE value it returns.

Figure 7 gives a box blot of the data in Fig. 6 but with the September data removed as an outlier. Model 11 is clearly not worth considering. An interesting finding from Figure 7 is that Models 1-10 have nearly the same minimum and maximum values for RMSE across remaining months of the year (September excluded). However, the average RSME changes significantly and easily shows that Model 2 provides the lowest RMSE with Model 1 and Model 10 as close seconds.

**Table 6. AIC values – all forecast models (lowest value in bold for each month).**

| Model No. | SARIMA Model | Jan | Feb | Mar | Apr | May | June | July | Aug | Sep | Oct | Nov | Dec | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | (1,0,0)(1,0,0) [24] | 1628.60 | 1572.05 | 1721.82 | 1848.70 | 2222.66 | 2230.02 | 1979.60 | 1935.05 | 1861.26 | 1598.80 | 1521.38 | 1431.42 | 1795.95 |
| 2 | (2,0,1)(0,1,0) [24] | 1605.40 | 1571.40 | 1723.97 | 1860.50 | 2227.08 | 2237.66 | 1995.40 | 1914.88 | 1874.34 | 1598.90 | 1514.14 | 1428.37 | 1796.00 |
| 3 | (2,1,1)(1,1,1) [24] | 1517.20 | **1549.44** | **1615.27** | 1789.90 | 2125.72 | 2116.93 | 1951.20 | 1812.76 | 1829.12 | 1528.90 | 1406.56 | 1400.36 | **1720.28** |
| 4 | (0,1,1)(1,1,1) [24] | 1553.80 | 1572.12 | 1626.91 | 1798.50 | 2123.93 | 2115.15 | 1950.20 | 1850.86 | **1827.12** | 1541.40 | 1414.04 | 1398.55 | 1731.05 |
| 5 | (0,0,1)(1,1,1) [24] | 1489.60 | 1566.46 | 1629.61 | 1752.90 | 2123.39 | 2115.02 | 2161.50 | 1803.13 | 1928.94 | 1525.90 | **1394.19** | 1388.72 | 1739.95 |
| 6 | (0,1,1)(1,0,1) [24] | 1553.30 | 1592.22 | 1626.76 | 1796.60 | 2123.11 | 2115.43 | **1943.00** | 1846.23 | 1843.00 | 1544.30 | 1430.01 | 1397.12 | 1734.26 |
| 7 | (0,0,1)(1,0,1) [24] | **1488.30** | 1589.48 | 1632.23 | **1747.90** | **2122.93** | 2115.29 | 2055.90 | **1797.41** | 1947.74 | **1521.40** | 1397.21 | **1381.54** | 1733.11 |
| 8 | (0,0,2)(1,0,1) [24] | 1576.10 | 1640.23 | 1683.17 | 1799.20 | 2123.64 | 2115.20 | 2139.70 | 1869.36 | 2005.73 | 1574.50 | 1456.63 | 1405.53 | 1782.42 |
| 9 | (0,0,2)(1,1,1) [24] | 1578.10 | 1624.42 | 1681.47 | 1805.50 | 2124.02 | **2114.93** | 2138.00 | 1878.47 | 1998.96 | 1584.70 | 1443.17 | 1411.81 | 1781.96 |
| 10 | (0,0,2)(1,0,0) [24] | 1602.70 | 1568.67 | 1719.82 | 1840.60 | 2222.24 | 2229.71 | 1978.40 | 1906.56 | 1861.26 | 1597.20 | 1512.66 | 1426.36 | 1788.85 |
| 11 | (0,0,2)(0,0,1) [24] | 2028.26 | 2044.34 | 2142.52 | 2203.60 | 2351.91 | 2350.43 | 2276.70 | 2244.48 | 2182.91 | 2121.30 | 2055.59 | 1977.95 | 2165.00 |

**Table 7. Root mean square error – all forecast models (lowest value in bold for each month).**

| Model No. | SARIMA Model | Jan | Feb | Mar | Apr | May | June | July | Aug | Sep | Oct | Nov | Dec | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | (1,0,0)(1,0,0) [24] | 11.39 | 104.20 | 31.88 | 111.63 | **98.88** | 33.62 | 118.82 | 34.63 | 287.94 | 23.88 | **72.08** | 15.27 | 78.69 |
| 2 | (2,0,1)(0,1,0) [24] | **7.51** | 105.66 | **25.77** | 117.86 | 98.09 | **4.08** | **86.02** | 9.79 | 302.10 | 22.71 | 73.59 | **12.74** | 72.16 |
| 3 | (2,1,1)(1,1,1) [24] | 12.53 | **94.84** | 30.56 | 104.10 | 115.73 | 44.54 | 93.86 | 9.84 | 170.59 | 58.94 | 89.04 | 15.66 | 70.02 |
| 4 | (0,1,1)(1,1,1) [24] | 12.48 | 97.65 | 30.61 | 104.56 | 115.74 | 44.37 | 94.12 | **8.86** | 170.58 | 49.28 | 88.97 | 15.63 | **69.41** |
| 5 | (0,0,1)(1,1,1) [24] | 13.45 | 102.43 | 31.07 | 98.78 | 115.23 | 37.02 | 88.98 | 23.17 | 325.53 | 66.36 | 82.04 | 21.01 | 83.76 |
| 6 | (0,1,1)(1,0,1) [24] | 19.01 | 99.90 | 39.18 | **78.69** | 116.89 | 42.06 | 235.67 | 34.91 | 175.43 | **18.48** | 76.79 | 23.13 | 80.02 |
| 7 | (0,0,1)(1,0,1) [24] | 17.33 | 102.21 | 31.82 | 85.83 | 116.96 | 36.05 | 118.24 | 44.50 | **158.34** | 23.84 | 73.09 | 26.36 | 69.55 |
| 8 | (0,0,2)(1,0,1) [24] | 17.31 | 105.52 | 32.11 | 85.77 | 117.01 | 36.04 | 115.15 | 44.10 | 181.06 | 23.46 | 73.20 | 25.90 | 71.39 |
| 9 | (0,0,2)(1,1,1) [24] | 13.52 | 107.07 | 30.77 | 99.84 | 115.18 | 37.04 | 100.85 | 23.73 | 329.11 | 63.30 | 82.61 | 20.36 | 85.29 |
| 10 | (0,0,2)(1,0,0) [24] | 10.07 | 104.09 | 31.45 | 112.44 | 99.43 | 36.32 | 113.15 | 28.91 | 287.89 | 23.79 | 72.18 | 15.42 | 77.93 |
| 11 | (0,0,2)(0,0,1) [24] | 172.75 | 118.69 | 264.96 | 248.37 | 301.32 | 281.29 | 296.28 | 275.42 | 344.84 | 171.05 | 157.26 | 143.71 | 231.33 |

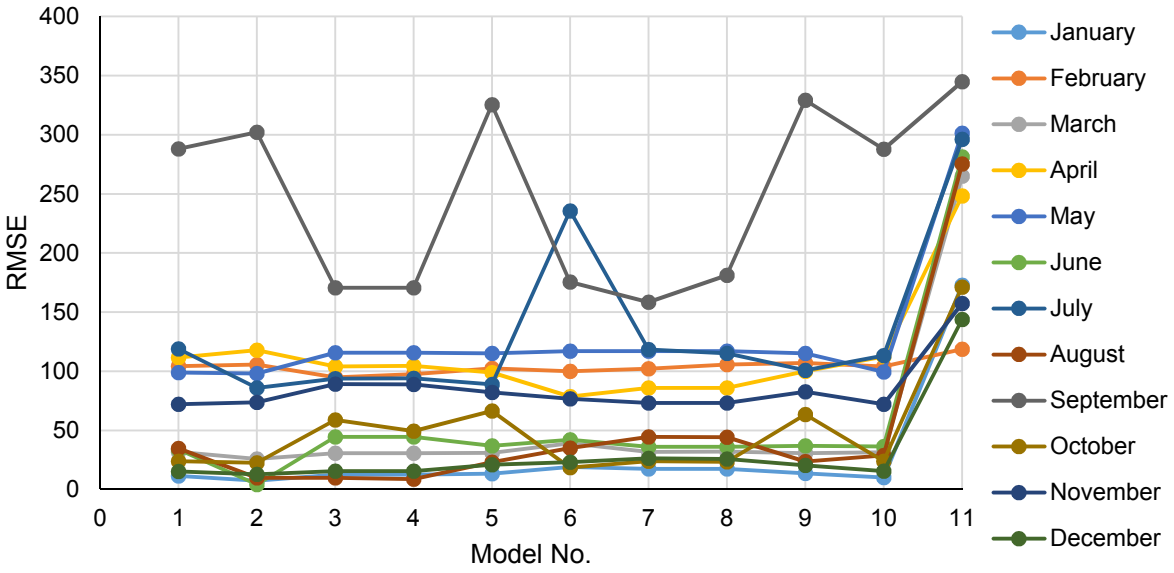


**Figure 5. Selected SARIMA Models with RMSE (all months).**

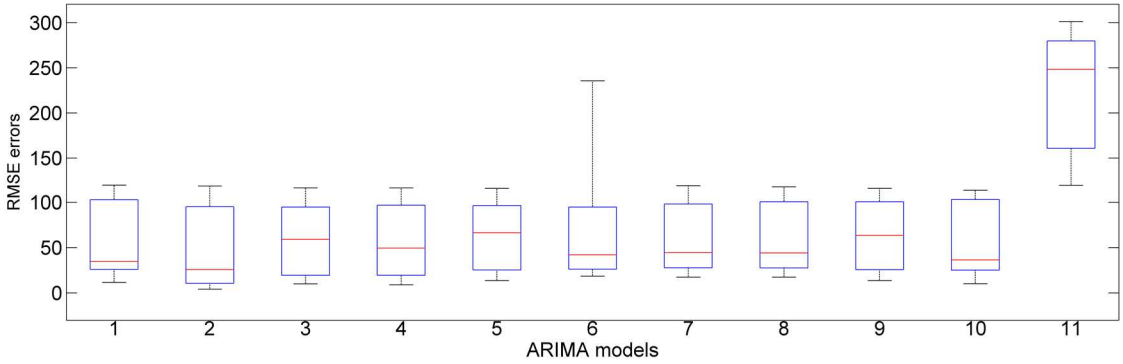


**Figure 6. Selected SARIMA Models with RMSE (September removed).**

*2.4.4 Comparison of GHI Forecasts to Peer Studies*

Time series results from Sections 2.4.1-2.4.3 are summarized here and compared against peer studies. Inferences can be made across studies though this is not a direct quantitative comparison because the peer studies are based on a different location and time series dataset that the regression models generated here. Table 8 contains models and RMSE errors from the Reikard et al. 2009 study that uses various regression and

forecasting methods such as log regression, ARIMA models, and neural networks [Reikard et al. 2009].

**Table 8. Forecast models and errors from reference study [Reikard et al. 2009].**

| Model No. | Resolution | Forecast Horizon | Model | Transformation | RMSE Error |
|---|---|---|---|---|---|
| a | 1-hour | 1-hour | Log regression | Log | 29.96 |
| b | 1-hour | 1-hour | Unobserved Components Model | None | 29.92 |
| c | 1-hour | 1-hour | ARIMA (1,0,0)(1,1,0) [24] | None | 23.60 |
| d | 1-hour | 1-hour | NN | None | 29.38 |
| e | 1-hour | 1-hour | Hybrid - ARIMA & NN | None | 23.67 |

Table 9 includes the best fit models obtained in this study for forecast horizons out to 1-hour at different time step resolutions. Table 10 provides similar information yet focuses on shorter forecast horizons including 5 minutes, 15 minutes, and 30 minutes. The comparison study in Table 8 includes only 1-hour resolution data. Models developed and shown in Table 9 provide similar results for 1-hour resolution data with a potential outlier in 15-minute resolution log transform data. The trend for no transformation as improvement over log transformation does hold between the comparison study and this study when looking at 1-hour resolution data. Although the RMSE errors change between Table 9 and Table 10, the model ranks as defined by RMSE error maintains the same order when the forecast horizon is reduced. Regarding transformations, the logarithmic transformation performed better as compared to untransformed and square-root transformed data irrespective of forecast horizon for data resolutions of 1-hour, 30-minutes, and 15-minutes. For data resolution of 5-minutes, square-root transformation provided lower RMSE values for all forecast horizons.

**Table 9. Forecast models and errors for 1-hour forecast horizon.**

| Model No. | Resolution | Forecast Horizon | SARIMA Model | Transformation | RMSE Error |
|---|---|---|---|---|---|
| 7 | 1-hour | 1-hour | (0,0,1)(1,0,1)[24] | None | 25.72 |
| 7 | 1-hour | 1-hour | (0,0,1)(1,0,1)[24] | Square-root | 35.81 |
| 5 | 1-hour | 1-hour | (0,0,1)(1,1,1)[24] | Log | 29.12 |
| 6 | 30-min | 1-hour | (0,1,1)(1,0,1)[48] | None | 70.17 |
| 6 | 30-min | 1-hour | (0,1,1)(1,0,1)[48] | Square-root | 65.18 |
| 2 | 30-min | 1-hour | (2,0,1)(0,1,0)[48] | Log | 34.46 |
| 3 | 15-min | 1-hour | (2,1,1)(1,1,1)[96] | None | 35.89 |
| 3 | 15-min | 1-hour | (2,1,1)(1,1,1)[96] | Square-root | 36.45 |
| 1 | 15-min | 1-hour | (1,0,0)(1,0,0)[96] | Log | 72.51 |
| 3 | 5-min | 1-hour | (2,1,1)(1,1,1)[288] | None | 37.52 |
| 3 | 5-min | 1-hour | (2,1,1)(1,1,1)[288] | Square-root | 34.61 |
| 6 | 5-min | 1-hour | (0,1,1)(1,0,1)[288] | Log | 82.14 |

**Table 10. Forecast models and errors for 5-minute, 15-minute, and 30-minute forecast horizons.**

| Model No. | Resolution | Forecast Horizon | SARIMA Model | Transformation | RMSE Error |
|---|---|---|---|---|---|
| 6 | 30-min | 30-min | (0,1,1)(1,0,1)[48] | None | 38.65 |
| 6 | 30-min | 30-min | (0,1,1)(1,0,1)[48] | Square-root | 31.71 |
| 2 | 30-min | 30-min | (2,0,1)(0,1,0)[48] | Log | 24.53 |
| 3 | 15-min | 15-min | (2,1,1)(1,1,1)[96] | None | 26.41 |
| 3 | 15-min | 15-min | (2,1,1)(1,1,1)[96] | Square-root | 26.38 |
| 1 | 15-min | 15-min | (1,0,0)(1,0,0)[96] | Log | 9.59 |
| 3 | 5-min | 5-min | (2,1,1)(1,1,1)[288] | None | 50.71 |
| 3 | 5-min | 5-min | (2,1,1)(1,1,1)[288] | Square-root | 49.10 |
| 6 | 5-min | 5-min | (0,1,1)(1,0,1)[288] | Log | 91.21 |

These findings, however, need to be validated against more case study applications and examples in order to develop generalizable findings. An important shortcoming of the peer time series studies is its focus on forecasting to only one data point in the future that is unchanging across all comparisons completed. This study mitigates that issue by running forecasting algorithms out to one-hour ahead during all hours of the day when the sun is shining. Table 11 provides these data for a selected SARIMA model (0,0,1)(1,0,1)[24]. This is 6:00 am to 6:00 pm with each forecast moving one hour into the future for simulation. This helps explore if a single model or method can be universally applied to achieve optimal or near-optimal performance for an entire day that also represents the entire month and that quarter of the year. Table 12 shows that no data transformation or square-root transformation continue to perform better than log transformation in nearly all cases (except October). This further corroborates earlier findings that the log transform should not be used in these case study data based on evidence from forecasting applications out four days, one day, or one hour. One could otherwise change the type of transformation based on the time of year, using either untransformed or square-root transformed data for the first three quarters of the year logarithmic transformation in the last quarter of the year. Another interesting finding from Tables 11a – 11d is that in most cases, irrespective of the data transformation used, RMSE values are lower during mornings and evenings and higher during midday. This can directly be related to the magnitude of solar irradiation and safely assumed that forecast models perform better when the solar irradiation is relatively low. Hence, it can be inferred that forecasting should be performed during mornings and evenings. From table 12, it can be seen that on comparing these results with that of tables

10 and 11, it can be seen that untransformed and square-root data perform better than logarithmic data

**Table 11a. SARIMA model RMSE error for one-hour ahead forecasting applied to a single day in January.**

| Data transformation | Forecast Time of Day | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **6:00 am** | **7:00 am** | **8:00 am** | **9:00 am** | **10:00 am** | **11:00 am** | **12:00 pm** | **1:00 pm** | **2:00 pm** | **3:00 pm** | **4:00 pm** | **5:00 pm** | **6:00 pm** |
| None | 0.01 | 2.29 | 22.29 | 20.70 | 2.25 | 6.49 | 215.22 | 260.75 | 38.55 | 13.84 | 20.28 | 74.66 | 14.77 |
| Square-root | 0.00 | 4.10 | 20.06 | 27.73 | 10.98 | 29.49 | 276.86 | 217.75 | 47.92 | 15.67 | 14.63 | 15.44 | 1.67 |
| Log | 1.05 | 8.45 | 41.72 | 98.94 | 97.25 | 186.62 | 508.02 | 135.36 | 276.36 | 136.02 | 92.77 | 5.84 | 1.01 |



**Table 11b. SARIMA model RMSE error for one-hour ahead forecasting applied to a single day in April.**

| Data transformation | Forecast Time of Day | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **6:00 am** | **7:00 am** | **8:00 am** | **9:00 am** | **10:00 am** | **11:00 am** | **12:00 pm** | **1:00 pm** | **2:00 pm** | **3:00 pm** | **4:00 pm** | **5:00 pm** | **6:00 pm** |
| None | 16.25 | 34.06 | 33.45 | 49.15 | 34.19 | 8.94 | 125.96 | 52.52 | 40.04 | 132.19 | 98.69 | 15.33 | 7.93 |
| Square-root | 17.06 | 22.97 | 33.15 | 57.17 | 32.34 | 2.04 | 129.93 | 63.99 | 43.85 | 139.79 | 110.80 | 20.97 | 17.06 |
| Log | 0.45 | 62.99 | 66.07 | 149.60 | 42.75 | 65.51 | 65.28 | 38.77 | 39.98 | 264.74 | 223.41 | 86.93 | 17.95 |

**Table 11c. SARIMA model RMSE error for one-hour ahead forecasting applied to a single day in July.**

| Data transformation | Forecast Time of Day | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **6:00 am** | **7:00 am** | **8:00 am** | **9:00 am** | **10:00 am** | **11:00 am** | **12:00 pm** | **1:00 pm** | **2:00 pm** | **3:00 pm** | **4:00 pm** | **5:00 pm** | **6:00 pm** |
| None | 23.78 | 12.91 | 14.94 | 293.42 | 89.92 | 202.62 | 129.18 | 37.51 | 44.60 | 23.93 | 52.89 | 38.48 | 63.58 |
| Square-root | 10.36 | 14.95 | 5.19 | 323.10 | 290.15 | 373.00 | 116.40 | 33.06 | 17.67 | 36.83 | 25.27 | 3.19 | 24.60 |
| Log | 32.71 | 27.02 | 67.49 | 283.58 | 65.11 | 238.41 | 83.14 | 24.98 | 34.24 | 25.04 | 5.59 | 1.40 | 0.42 |



**Table 11d. SARIMA model RMSE error for one-hour ahead forecasting applied to a single day in October.**

| Data transformation | Forecast Time of Day | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **6:00 am** | **7:00 am** | **8:00 am** | **9:00 am** | **10:00 am** | **11:00 am** | **12:00 pm** | **1:00 pm** | **2:00 pm** | **3:00 pm** | **4:00 pm** | **5:00 pm** | **6:00 pm** |
| None | 5.10 | 11.91 | 5.06 | 19.94 | 14.73 | 14.17 | 13.49 | 44.42 | 4.00 | 11.85 | 1.83 | 7.05 | 8.46 |
| Square-root | 2.97 | 38.70 | 32.77 | 27.72 | 17.24 | 8.04 | 12.80 | 8.33 | 2.63 | 0.91 | 2.92 | 3.26 | 0.05 |
| Log | 6.45 | 24.57 | 0.06 | 28.72 | 14.71 | 19.76 | 18.10 | 59.62 | 9.14 | 18.88 | 7.10 | 7.02 | 0.29 |

**Table 12. SARIMA model RMSE error comparison across forecast months in Table 11.**

| Data transformation | Forecast Month | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | January | | | April | | | July | | | October | | |
| | Min | Max | Average | Min | Max | Average | Min | Max | Average | Min | Max | Average |
| None | 0.01 | 260.75 | 53.24 | 7.93 | 132.19 | 49.90 | 12.92 | 293.43 | 79.06 | 1.83 | 44.43 | 12.47 |
| Square-root | 0.00 | 276.86 | 52.49 | 2.04 | 139.79 | 53.16 | 0.42 | 283.58 | 68.39 | 0.06 | 59.63 | 16.49 |
| Log | 1.01 | 508.02 | 122.27 | 0.45 | 264.74 | 86.49 | 3.19 | 373.01 | 97.99 | 0.06 | 38.71 | 12.18 |

## 2.5 Application of GHI forecasting to incident radiation forecasting

Forecasts GHI data can be used to estimate future solar PV power output by translating the GHI data into solar radiation incident on the PV array using a series of equations introduced in Duffie and Beckman (Duffie a Beckman 2013). The total incident solar radiation $G_T$ can then be used to calculate solar PV power output using the panel's rated efficiency or productivity data. The equation set 10-15 offers improved accuracy by accounting for the circumsolar diffuse radiation using the anisotropy index, $A_i$, and a horizontal brightening factor, $f$, that better describes the diffuse radiation when the sun is near the horizon.

$$G_T = \left(G_b + G_d A_i\right) R_b + G_d R_d \left(1 - A_i\right)\left[1 + f \sin^3\left(\frac{\beta}{2}\right)\right] + \rho R_a \left(G_b + G_d\right) \tag{10}$$

$$R_b = \frac{\cos\theta_i}{\cos\theta_z}; \quad R_d = \frac{1+\cos\beta}{2}; \quad R_a = \frac{1-\cos\beta}{2} \tag{11}$$

$$A_i = \frac{G_b}{G_o} \tag{12}$$

$$f = \sqrt{\frac{G_b}{G_b + G_d}} \tag{13}$$

$$GHI = G_d + G_b * \cos(\theta_z) \tag{14}$$

$$k_t = \frac{GHI}{G_o} \tag{15}$$

where,

$G_T$ : total incident solar radiation (W/m$^2$)

$G_b$ : beam radiation or Direct Normal Irradiation (W/m$^2$)

$G_d$ : diffuse sky radiation or Diffuse Horizontal Irradiation (W/m$^2$)

$GHI$ : Global Horizontal Irradiation (W/m$^2$)

$G_o$ : extraterrestrial horizontal irradiation (W/m$^2$)

$k_t$ : clearness index (-)

$A_i$ : anisotropy index (-)

$f$ : horizontal brightening factor (-)

$\beta$ : slope of the surface ($^o$)

$\rho$ : reflection coefficient (-)

$\theta_i$ : angle of incidence ($^o$)

$\theta_z$ : zenith angle ($^o$)

The total incident solar radiation was calculated for one day using hourly resolution data and SARIMA model (0,0,1)(1,0,1)[24]. Figure 8 shows a comparison between the measured and forecasted GHI that is then translated into Fig. 9 using Equations 10-15. It is clear that the incident radiation closely follows the global horizontal irradiation and the error between the measured and the forecasted values are reasonable in the absence of perfect information. This can be taken one step further to simulate the solar PV power output as a function of the incident radiation.

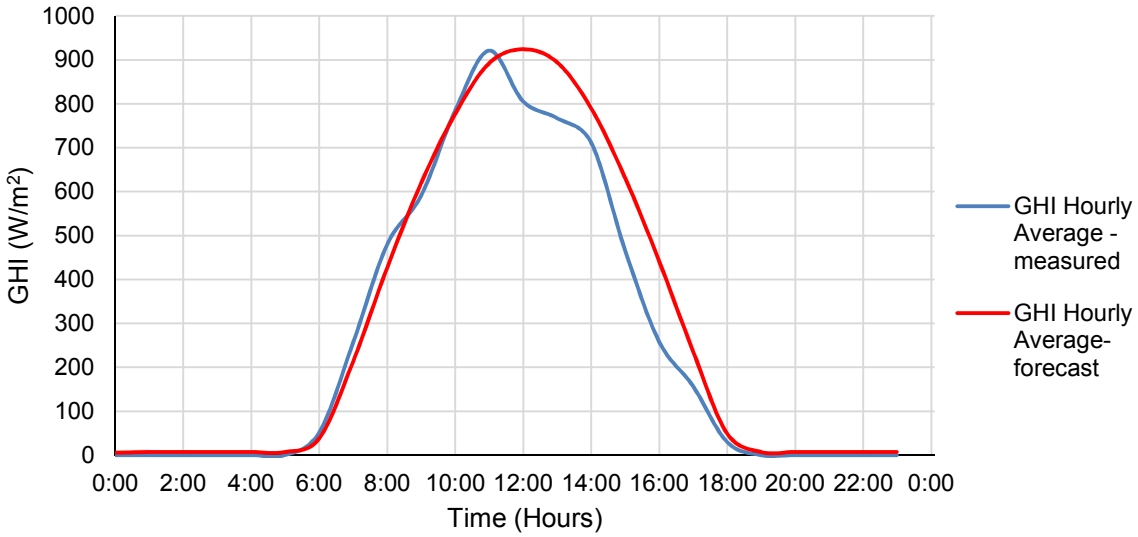


**Figure 7. Measured GHI data and forecasted GHI data.**

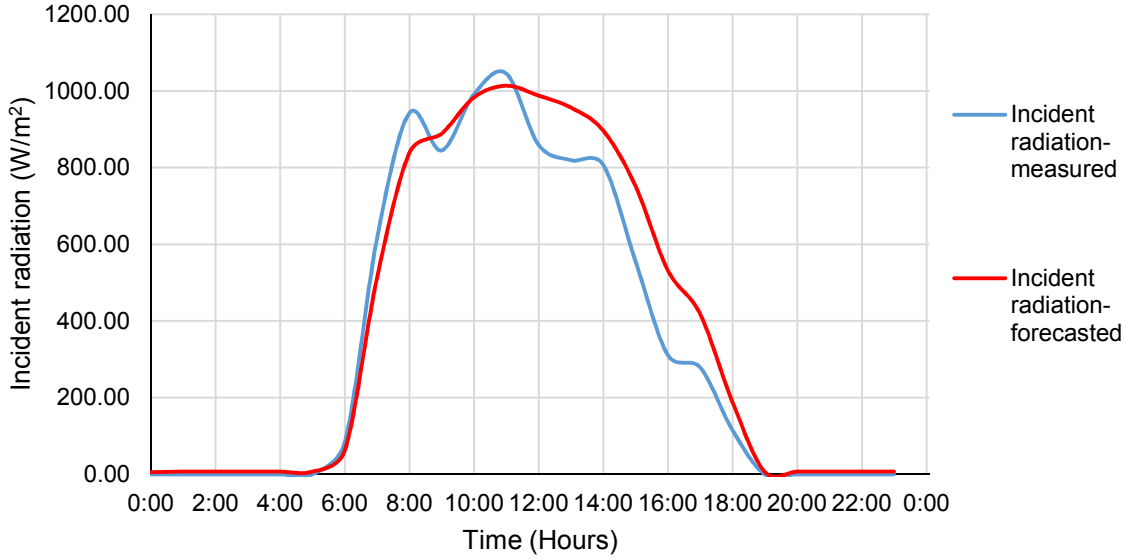


**Figure 8. Calculated incident radiation using measured GHI data and forecasted GHI data.**

## Chapter 3: Discussion

Time series forecasting techniques were applied to GHI data to estimate future solar resource availability. The forecasted solar resource data can be used to calculate solar PV power output to aid in scheduling generation, planning capacity reserve, and dispatching backup energy storage in case of shortage or supply-demand mismatch. These techniques can be applied to distributed or centralized solar generation assets. This work can be extended to forecast solar PV power output in real-time with comparisons completed between different field site locations.

Incident solar radiation closely follows the GHI, as mentioned earlier. Hence, any fluctuations in GHI will also be reflected in incident solar radiation and solar PV output. Existing control techniques like ramp rate control and active power curtailment can be used to reduce the intermittency in the output power. The most common technique is the use of an external battery storage system with solar PV. Based on set threshold values of PV output power, the storage system supplies the deficit power to the grid or charges from the solar PV system during that time of the day when solar power is excess. Forecasting techniques introduced here can be used to create probability density functions of future renewables availability for a more accurate picture of future energy storage availability. The combination of forecasted solar, forecasted load, and forecasted storage can yield data to control power output in the near term to maintain sufficient energy reserves for the future. This form of forecasting with adaptive control techniques may provide improved autonomy in micro-grids through curtailing non-critical loads.

## References


Act, E. P. (2005). 109th Congress. Public Law, 109, 54.

American Physical Society: Panel on Public Affairs. (2009). Integration Renewable Electricity on the grid.

Bank J., Mather B., Keller J. and Coddington M. (January 2013). High Penetration Photovoltaic Case Study Report. National Renewable Energy Laboratory.

Barrett, A., Hogan, R., & Forbes, R. Evaluation and improvement of mixed-phase cloud schemes using radar and lidar observations.

Brissette, A., Hoke, A., Traube, J., Lu, F., & Maksimovizc, D. (2013, August). Study on the effect of solar irradiance intermittency mitigation on electric vehicle battery lifetime. In Technologies for Sustainability (SusTech), 2013 1st IEEE Conference on (pp. 262-267). IEEE.

Brockwell, P. J., & Davis, R. A. (2006). Introduction to time series and forecasting. Springer Science & Business Media, 179.

Brockwell, P. J., & Davis, R. A. (2006). Introduction to time series and forecasting. Springer Science & Business Media, 84.

Brockwell, P. J., & Davis, R. A. (2006). Introduction to time series and forecasting. Springer Science & Business Media, 180.

Brockwell, P. J., & Davis, R. A. (2006). Introduction to time series and forecasting. Springer Science & Business Media, 173.

California ISO. Fast Facts: What the duck curve tells us about managing a green grid. Retrieved from: https://www.caiso.com/Documents/FlexibleResourcesHelpRenewables_FastFacts.pdf

Diagne, M., David, M., Lauret, P., Boland, J., & Schmutz, N. (2013). Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews, 27, 65-76.

Durkay J. (January 2016). State Renewable Portfolio Standards and Goals. National Conference on State Legislatures. Retrieved from: http://www.ncsl.org/research/energy/renewable-portfolio-standards.aspx

Eber K. and Corbus D. (June 2013). Hawaii Solar Integration Study: Executive Summary. National Renewable Energy Laboratory.

Federal Energy Regulatory Commission. (June 2012). Integration of Variable Energy Resources. Retrieved from: https://www.ferc.gov/whats-new/comm-meet/2012/062112/E-3.pdf

Graham, V. A., Hollands, K. G. T., & Unny, T. E. (1988). A time series model for K t with application to global synthetic weather generation. Solar Energy,40(2), 83-92.

Hammer, A., Heinemann, D., Lorenz, E., & Lückehe, B. (1999). Short-term forecasting of solar radiation: a statistical approach using satellite data. Solar Energy, 67(1), 139-150.

Hanna, R., Kleissl, J., Nottrott, A., & Ferry, M. (2014). Energy dispatch schedule optimization for demand charge reduction using a photovoltaic-battery storage system with solar forecasting. Solar Energy, 103, 269-287.

Huque A. (February 2015). Smart Inverter Grid Support Functions and Potential Impact on Reliability.

International Energy Agency. (2012). World Energy Outlook: Renewable Energy Outlook.

Jamehbozorg, A., Keshmiri, S. N., & Radman, G. (2011, March). PV Output Power Smoothing Using Energy Capacitor System. In Southeastcon, 2011 Proceedings of IEEE (pp. 164-168). IEEE.

Kang, B. O., & Tam, K. S. (2015). New and improved methods to estimate day-ahead quantity and quality of solar irradiance. Applied Energy, 137, 240-249.

Lam, R. K., & Yeh, H. G. (2014, November). PV ramp limiting controls with adaptive smoothing filter through a battery energy storage system. In Green Energy and Systems Conference (IGESC), 2014 IEEE (pp. 55-60). IEEE.

Li, H., Ma, W., Wang, X., & Lian, Y. (2011). Estimating monthly average daily diffuse solar radiation with multiple predictors: a case study. Renewable energy, 36(7), 1944-1948.

Lin, C. H., Hsieh, W. L., Chen, C. S., Hsu, C. T., & Ku, T. T. (2012). Optimization of photovoltaic penetration in distribution systems considering annual duration curve of solar irradiation. Power Systems, IEEE Transactions on, 27(2), 1090-1097.

Mammoli, A., Menicucci, A., Caudell, T., Ellis, A., Willard, S., & Simmins, J. (2013, August). Low-cost solar micro-forecasts for PV smoothing. InTechnologies for Sustainability (SusTech), 2013 1st IEEE Conference on(pp. 238-243). IEEE.

Martín, L., Zarzalejo, L. F., Polo, J., Navarro, A., Marchante, R., & Cony, M. (2010). Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. Solar Energy, 84(10), 1772-1781.

Mellit, A., Benghanem, M., & Kalogirou, S. A. (2006). An adaptive wavelet-network model for forecasting daily total solar-radiation. Applied Energy,83(7), 705-722.

National Renewable Energy Laboratory. (July 2015). Renewable Portfolio Standards. Retrieved from: http://www.nrel.gov/tech_deployment/state_local_governments/basics_portfolio_standards.html

National Renewable Energy Laboratories. (December 2012). Impacts of Solar Power on Operating Reserve Requirements. Retrieved from: http://www.nrel.gov/docs/fy13osti/56596.pdf

Office of Energy Efficiency and Renewable Energy. (2015). Solar. Retrieved from: http://energy.gov/eere/renewables/solar

Outlook, A. E. (2010). Energy information administration. Department of Energy.

Perera, B. K., Ciufo, P., & Perera, S. (2013, November). Point of common coupling (PCC) voltage control of a grid-connected solar photovoltaic (PV) system. In Industrial Electronics Society, IECON 2013-39th Annual Conference of the IEEE (pp. 7475-7480). IEEE.

Perez, M. J., & Fthenakis, V. M. (2013, June). Long-distance interconnection as solar resource intermittency solution: Optimizing the use of energy storage and the geographic dispersion+ interconnection of solar generating facilities. In Photovoltaic Specialists Conference (PVSC), 2013 IEEE 39th (pp. 3367-3373). IEEE.

Perez, M. J., & Fthenakis, V. M. (2013, June). Long-distance interconnection as solar resource intermittency solution: Optimizing the use of energy storage and the geographic dispersion+ interconnection of solar generating facilities. In Photovoltaic Specialists Conference (PVSC), 2013 IEEE 39th(pp. 3367-3373). IEEE.

Reikard, G. (2009). Predicting solar radiation at high resolutions: A comparison of time series forecasts. Solar Energy, 83(3), 342-349.

Solar Energy Industries Association (SEIA). (2015). Solar Investment Tax Credit (ITC). Retrieved from: http://www.seia.org/policy/finance-tax/solar-investment-tax-credit

Solar Energy Industries Association (SEIA). (2015). Solar Market Insight Report 2015 Q1.

Solar Energy Industries Association (SEIA). (March 2016). U.S. Solar market Insight.

Solar Energy Industries Association (SEIA). (May 2015). National RES Bill Draws Strong Solar Industry Support. Retrieved from: http://www.seia.org/news/national-res-bill-draws-strong-solar-industry-support

Sözen, A., Arcaklıoğlu, E., Özalp, M., & Çağlar, N. (2005). Forecasting based on neural network approach of solar potential in Turkey. Renewable Energy,30(7), 1075-1090.

Thongpron, J., Sangpanich, U., Limsakul, C., Chenvidya, D., Kirtikara, K., & Jivacate, C. (2004). Study of a PV-grid connected system on its output harmonics and voltage variation. Asian J. Energy Environ, 5(1), 59-73.

U.S. Department of Energy. (May 2014). 2014: The Year of Concentrating Solar Power.

U.S. Department of Energy: Energy Efficiency & Renewable Energy (EERE). (2014). 2014 Renewable Energy Data Book.

U.S. Department of Energy: Energy Efficiency & Renewable Energy (EERE). (2009). High Penetration of Photovoltaic (PV) System into the Distribution Grid.

U.S. Department of Energy: Loan Programs Office. (February 2015). Investing in American Energy.

U.S. Department of Energy: Sun Shot. (September 2014). Photovoltaic System Pricing Trends: Historical, Recent and Near-Term Projections.

U.S. Department of Energy. (2015). Quadrennial Technology Report 2015.

U.S. Energy Information Administration. (December 2015). Short-Term Energy Outlook.

U.S. Energy Information Administration. (March 2015). Today in Energy.

U.S. Energy Information Administration. (March 2016). Short-Term Energy Outlook.

Voyant, C., Paoli, C., Muselli, M., & Nivet, M. L. (2013). Multi-horizon solar radiation forecasting for Mediterranean locations using time series models.Renewable and Sustainable Energy Reviews, 28, 44-52.

Wang, G., Ciobotaru, M., & Agelidis, V. G. (2014). Power smoothing of large solar PV plant using hybrid energy storage. Sustainable Energy, IEEE Transactions on, 5(3), 834-842.

Wang, F., Mi, Z., Su, S., & Zhao, H. (2012). Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters. Energies, 5(5), 1355-1370.

Wei, W. W. S. (1994). Time series analysis. Addison-Wesley Publishing.

Widén, J., Carpman, N., Castellucci, V., Lingfors, D., Olauson, J., Remouit, F. ... & Waters, R. (2015). Variability assessment and forecasting of renewables: A review for solar, wind, wave and tidal resources. Renewable and Sustainable Energy Reviews, 44, 356-375.

Wilcox, S.; Andreas, A.; (2013). Solar Resource and Meteorological Assessment Project (SOLRMAP). Southwest Solar Research Park, Arizona (Data). http://dx.doi.org/10.5439/1052563.

Wirth, H. (May 2015). Recent Facts about Photovoltaics in Germany. Fraunhofer Institute for Solar Energy Systems.

# APPENDIX A

## MATLAB PROGRAM - TIME SERIES FORECASTING OF GHI

```
clc;
clear all
hours = [1:1:168]; % hours in one week
pow = importdata('Phoenix_solar_hourly_Jan_0000.txt');
pow = pow';

measured = importdata('Phoenix_solar_hourly_Jan_measured.txt');
meas=zeros(24,1);
for i=1:1:24
   meas(i)=measured(i);
end


%------------------------------------------------------------------------

%% Plot original data

figure(1)
Subplot(3,1,1)
plot(hours,pow)
title('Phoenix hourly global irradiance data for one week_January_0000')

% examine time-series regression models
des = [ones(168,1)', hours]; % Design matrix
beta = des\hours;
u = pow - des*beta; % Residuals

subplot(3,1,2)
autocorr(u)
subplot(3,1,3)
parcorr(u)


%------------------------------------------------------------------------

%% Plot first difference

hoursDiff = 1:1:167; %remove one hour to do differencing

diffPow = diff(pow);
diffDes = [ones(167,1)', hoursDiff]; % Design matrix
beta = diffDes\hoursDiff;
u = diffPow - diffDes*beta; % Residuals

```

```
figure(2)
subplot(3,1,1)
plot(hoursDiff,diffPow)
title('Difference of per minute global irradiance data for one
week_January_Phoenix')

% examine time-series regression models
des = [ones(167,1)', hoursDiff]; % Design matrix
beta = des\hoursDiff;
u = diffPow - des*beta; % Residuals

subplot(3,1,2)
autocorr(u)
subplot(3,1,3)
parcorr(u)

%------------------------------------------------------------------------

%% Identify estimates with ARMA model (no seasonality)
% AR and MA models out to 3 lags with AR and 2 with MA, with no integrated
term

pMax = 3; % AR lag
qMax = 2; % MA lag
pSize = pMax+1; % for the results matrices
qSize = qMax+1;
LogL = zeros(pSize,qSize);
SumPQ = LogL;

for p = 0:pMax
   for q = 0:qMax
      Mdl = arima(p,0,q);
      [~,~,LogL(p+1,q+1)] = estimate(Mdl,pow','Display','params');
      SumPQ(p+1,q+1) = p+q;
   end
end

logL = reshape(LogL,pSize*qSize,1);
numParams = reshape(SumPQ,pSize*qSize,1) + 2;
aic = aicbic(logL,numParams);
AIC = reshape(aic,pSize,qSize)

a(1) = min(aic)
```

```
[bestP,bestQ] = find(AIC == a(1)); % this returns the matrix location with the
lowest AIC, but since we started with 0, we need to subtract 1
bestP = bestP - 1 % returns 2
bestQ = bestQ - 1 % returns 0

% re-output the best model
Mdl1 = arima(bestP,0,bestQ);
BestMdl = estimate(Mdl1,pow','Display','params');

%------------------------------------------------------------------------
%% Identify estimates with ARIMA model (no seasonality)

% AR and MA models out to 3 lags with AR and 2
% with MA, with 1 integrated term

pMax = 3; % AR lag
qMax = 2; % MA lag
pSize = pMax+1; % for the results matrices
qSize = qMax+1;
LogL = zeros(pSize,qSize);
SumPQ = LogL;

for p = 0:pMax
   for q = 0:qMax
      Mdl = arima(p,1,q);
      [~,~,LogL(p+1,q+1)] = estimate(Mdl,pow','Display','params');
      SumPQ(p+1,q+1) = p+q;
   end
end

logL = reshape(LogL,pSize*qSize,1);
numParams = reshape(SumPQ,pSize*qSize,1) + 2;
aic = aicbic(logL,numParams);
AIC = reshape(aic,pSize,qSize)

a(2) = min(aic)

[bestP,bestQ] = find(AIC == a(2)); %returns the matrix location with the lowest
                                   %AIC, but since we started with 0, we need to
                                   subtract 1
bestP = bestP - 1 % returns 2
bestQ = bestQ - 1 % returns 1
% re-output the best model
Mdl2= arima(bestP,1,bestQ);
```

```
BestMdl = estimate(Mdl2,pow','Display','params');

%% Identify estimates with ARIMA model (seasonality)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(1,0,0)(1,0,0)[24]

Mdl3 = arima('ARLags',1,'SARLags',24);
[EstMdl,~,LogL] = estimate(Mdl3,pow','Display','params');
EstMdl
numParams = 2 + 2;
a(3) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(2,0,1)(0,1,0)[24]

Mdl4= arima('ARLags',2,'MALags',1','Seasonality',24);
[EstMdl,~,LogL] = estimate(Mdl4,pow','Display','params');
EstMdl
numParams = 2 + 2;
a(4) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(2,1,1)(1,1,1)[24]

Mdl5
=arima('ARLags',2,'D',1,'MALags',1,'SARLags',24,'Seasonality',24,'SMALags',24);
[EstMdl,~,LogL] = estimate(Mdl5,pow','Display','params');
EstMdl
numParams = 4 + 2;
a(5) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(0,1,1)(1,1,1)[24]

Mdl6 = arima('D',1,'MALags',1,'SARLags',24,'Seasonality',24,'SMALags',24);
```

```
[EstMdl,~,LogL] = estimate(Mdl6,pow','Display','params');
EstMdl
numParams = 3 + 2;
a(6) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(0,0,1)(1,1,1)[24]

Mdl7 = arima('MALags',1,'SARLags',24,'Seasonality',24,'SMALags',24);
[EstMdl,~,LogL] = estimate(Mdl7,pow','Display','params');
EstMdl
numParams = 3 + 2;
a(7) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(0,1,1)(1,0,1)[24]

Mdl8 = arima('D',1,'MALags',1,'SARLags',24,'SMALags',24)
[EstMdl,~,LogL] = estimate(Mdl8,pow','Display','params');
EstMdl
numParams = 3 + 2;
a(8) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(0,0,1)(1,0,1)[24]

Mdl9 = arima('MALags',1,'SARLags',24,'SMALags',24);
[EstMdl,~,LogL] = estimate(Mdl9,pow','Display','params');
EstMdl
numParams = 3 + 2;
a(9) = aicbic(LogL,numParams)

%-------------------------------------------------------------------------------------------------
------

% ARIMA(0,0,2)(1,0,1)[1440]

Mdl10 = arima('MALags',2,'SARLags',24,'SMALags',24);
```

```
[EstMdl,~,LogL] = estimate(Mdl10,pow','Display','params');
EstMdl
numParams = 3 + 2;
a(10) = aicbic(LogL,numParams)

%-----------------------------------------------------------------------------------------------
------

% ARIMA(0,0,2)(1,1,1)[24]

Mdl11 = arima('MALags',2,'SARLags',24,'Seasonality',24,'SMALags',24);
[EstMdl,~,LogL] = estimate(Mdl11,pow','Display','params');
EstMdl
numParams = 3 + 2;
a(11) = aicbic(LogL,numParams)

%-----------------------------------------------------------------------------------------------
------



% ARIMA(0,0,2)(1,0,0)[24]

Mdl12 = arima('MALags',2,'ARLags',1,'SARLags',24);
[EstMdl,~,LogL] = estimate(Mdl12,pow','Display','params');
EstMdl
numParams = 2 + 2;
a(12) = aicbic(LogL,numParams)

%-----------------------------------------------------------------------------------------------
------


% ARIMA(0,0,2)(0,0,1)[24]

Mdl13 = arima('MALags',2,'SMALags',24);
[EstMdl,~,LogL] = estimate(Mdl13,pow','Display','params');
EstMdl
numParams = 2 + 2;
a(13) = aicbic(LogL,numParams)


%-------------------------------------------------------------------------
%   Simulates and forecasts timeseries data
```

```
%-------------------------------------------------------------------------

%a=reshape(a,[2,1])
bestaic=min(a)
for n=1:1:13
   if bestaic==a(n)                          % Find minimum AIC value
      newmdl=eval(['Mdl' num2str(n)]);       % Find model with minimum AIC
value
      break;
   else
      continue;
   end
end
%newmdl=Mdl2;
EstMdl = estimate(newmdl,pow');

%forecast
[Y,YMSE] = forecast(EstMdl,96,'Y0',pow');

result=[Y YMSE];

dlmwrite('Phoenix_results_hourly_January_0000.csv',result)

%-------------------------------------------------------------------------------------------------
------

[x,y,v]=find(meas);
meas=meas(x);
F=Y(x);                                  %Obtain non-zero values of forecasted result
diff=meas-F;
rmse= sqrt(sum(diff.^2)/length(diff)); % Calculate rmse
rmse=vpa(rmse,6)

%-------------------------------------------------------------------------------------------------
------

%Plot the forecast results

figure(3)
plot(pow,'Color',[.7,.7,.7]);
hold on

h1 = plot(169:264,Y,'k','LineWidth',2);
legend(h1,'Forecast','Location','NorthWest')
```

```
308     title('Phoenix Global irradiance forecast_January_0000')
309
310     hold off
311
312     %-----------------------------------------------------------------------------------------------
313     ------
```

# APPENDIX B

# MATLAB PROGRAM – CALCULATION OF INCIDENT SOLAR RADIATION FROM GHI

```
clc;
clear all;
GHI_meas=xlsread('Incident.xlsx','G2:G25');
GHI_forecast=xlsread('Incident.xlsx','H2:H25');
zenith=xlsread('Incident.xlsx','I2:I25');
ETR=xlsread('Incident.xlsx','J2:J25');
incidence=xlsread('Incident.xlsx','K2:K25');
beta=33;
ref=0.6;

%-----------------------------------------------------------------------------------------------
------
%Calculation of kt

for n=1:1:24
if GHI_meas(n)==0 || ETR(n)==0
   kt(n)=0;
else
   kt(n)=GHI_meas(n)/ETR(n);
end
end

%-----------------------------------------------------------------------------------------------
------
% Caluculation of Gd_measured

for t=1:1:24
   if kt(t)<=0.22
      Gd_meas(t)=GHI_meas(t)*(1-0.09*kt(t));
   elseif kt(t)>0.22 && kt(t)<=0.8
         Gd_meas(t)=GHI_meas(t)*(0.9511-0.1604*kt(t)+4.388*((kt(t))^2)-
16.638*((kt(t))^3)+12.336*((kt(t))^4));
   else
       Gd_meas(t)=GHI_meas(t)*0.165;
   end
end
Gd_meas=Gd_meas';

%-----------------------------------------------------------------------------------------------
------
% Calculation of Gb_measured

Gb_meas=(GHI_meas-Gd_meas)./cosd(zenith);
for t=1:1:24
```

```
    if Gb_meas(t)>=0
       Gb_meas(t)=Gb_meas(t);
    else
      Gb_meas(t)=0;
    end
end

%-------------------------------------------------------------------------------------------------
------
% Calculation of Incident solar radiation for measured GHI

Rb=cosd(incidence)./cosd(zenith);
Rd=(1+cosd(beta))/2;
Ra=(1-cosd(beta))/2;
Ai=Gb_meas./1367;
f=sqrt((Gb_meas)./(Gb_meas+Gd_meas));
for t=1:1:24
    if f>=0
       f;
    else
       f=0;
    end
end
G_incidence_meas=(Gb_meas+Gd_meas.*Ai).*Rb+Gd_meas.*Rd.*(1-
Ai).*[1+f.*(sind(beta/2))^3]+ref.*Ra.*(Gb_meas+Gd_meas)

%-------------------------------------------------------------------------
 %Calculation of Gd_forecast

for t=1:1:24
    if kt(t)<=0.22
       Gd_forecast(t)=GHI_forecast(t)*(1-0.09*kt(t));
    elseif kt(t)>0.22 && kt(t)<=0.8
          Gd_forecast(t)=GHI_forecast(t)*(0.9511-0.1604*kt(t)+4.388*((kt(t))^2)-
16.638*((kt(t))^3)+12.336*((kt(t))^4));
    else
        Gd_forecast(t)=GHI_forecast(t)*0.165;
    end
end
Gd_forecast=Gd_forecast';

%-------------------------------------------------------------------------------------------------
------
% Calculation of Gb_forecast
```

```
 89
 90    Gb_forecast=(GHI_forecast-Gd_forecast)./cosd(zenith);
 91    for t=1:1:24
 92       if Gb_forecast(t)>=0
 93          Gb_forecast(t)=Gb_forecast(t);
 94       else
 95         Gb_forecast(t)=0;
 96       end
 97    end
 98
 99    %-----------------------------------------------------------------------------------------------
100    ------
101    % Calculation of incident solar radiation for forecasted GHI
102
103    Rb=cosd(incidence)./cosd(zenith);
104    Rd=(1+cosd(beta))/2;
105    Ra=(1-cosd(beta))/2;
106    Ai=Gb_forecast./1367;
107    f=sqrt((Gb_forecast)./(Gb_forecast+Gd_forecast));
108    for t=1:1:24
109       if f>=0
110          f;
111       else
112          f=0;
113       end
114    end
115    G_incidence_forecast=(Gb_forecast+Gd_forecast.*Ai).*Rb+Gd_forecast.*Rd.*(1-
116    Ai).*[1+f.*(sind(beta/2))^3]+ref.*Ra.*(Gb_forecast+Gd_forecast)
117
118    %-----------------------------------------------------------------------------------------------
```