Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 97 (2013) 58–66

SOLAR ENERGY

www.elsevier.com/locate/solener

Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors

Vincent P.A. Lonij [a,*], Adria E. Brooks [a], Alexander D. Cronin [a], Michael Leuthold [b], Kevin Koch [c]

[a] University of Arizona, Department of Physics, 1118 E. 4th Street, Tucson, AZ 85716, United States
[b] University of Arizona, Department Atmospheric Sciences, 1118 E. 4th Street, Tucson, AZ 85716, United States
[c] Technicians For Sustainability, 612 N. 7th Ave., Tucson, AZ 85705, United States

Received 26 March 2013; received in revised form 5 July 2013; accepted 3 August 2013
Available online 31 August 2013

Communicated by: Associate Editor Jan Kleissl

Abstract

We report a new method to forecast power output from photovoltaic (PV) systems under cloudy skies that uses measurements from ground-based irradiance sensors as an input. This work describes an implementation of this forecasting method in the Tucson, AZ region where we use 80 residential rooftop PV systems distributed over a 50 km × 50 km area as irradiance sensors. We report RMS and mean bias errors for a one year period of operation and compare our results to the persistence model as well as forecasts from other authors. We also present a general framework to model station-pair correlations of intermittency due to clouds that reproduces the observations in this work as well as those of other authors. Our framework is able to describe the RMS errors of velocimetry based forecasting methods over three orders of magnitude in the forecast horizon (from 30 s to 6 h). Finally, we use this framework to recommend optimal locations of irradiance sensors in future implementations of our forecasting method.

Keywords: Irradiance; Variability; Cloud forecasting; Spatio-temporal correlation; Station-pair correlation

1. Introduction

Solar power generation at the utility-scale is a Grand Challenge (NAOE, 2008). A major problem is the intermittent output of solar power plants due to passing clouds and nighttime. Intermittency limits the adoption of solar power by utility companies and industry because of potentially unpredictable grid instabilities that may result. Additionally, unpredicted fluctuations in solar power production may cause the utility to either overproduce electricity or purchase additional electricity. Both scenarios are expensive and can negate the benefits of using solar power (Gowrisankaran et al., 2011). Fluctuations and intermittency can be mitigated with energy storage, spinning reserves, or demand response. However, optimal management of these three methods requires accurate forecasts of PV power output on several timescales. Such forecasts are the subject of the research presented in this paper.

Forecasts with multiple forecast horizons are valuable for utility operators and plant owners. For example, day-ahead forecasts are needed to determine an optimal energy trading strategy in the energy market. Hour-ahead and intra-hour forecasts are valuable for electric grid operators to better schedule spinning reserves and demand response.

Several methods exist to forecast PV power output, including numerical weather models and velocimetry of

* Corresponding author.
E-mail addresses: lonij@physics.arizona.edu (V.P.A. Lonij), cronin@physics.arizona.edu (A.D. Cronin).

0038-092X/$ - see front matter
http://dx.doi.org/10.1016/j.solener.2013.08.002

clouds using satellite images or ground based measurements of clouds (SolarAnywhere, 2012; Jayadevan et al., 2012; Chow et al., 2011; Perez et al., 2010; Hamill and Nerhkorn, 1993). However, these methods only outperform the persistence model for horizons either <10 min or >60 min.

In Section 2, we present a novel method to forecast solar power production using power measurements of a distributed network of PV systems (Lonij et al., 2012c). We will show how this method outperforms the persistence model for forecast horizons larger than 30 min. An advantage of using ground-based sensors is that PV power output can be inferred directly from the output of other PV systems without independent estimates of the height, density, reflectivity, or spectral properties of clouds. This is unlike satellite based or Numerical Weather Prediction (NWP) methods, which have to use radiative-transfer models to convert either satellite images or 3D cloud density data into an estimate of surface-level POA irradiance (Müller et al., 2004; Perez et al., 2004). Note that we use power measurements of PV systems as a measure of irradiance, this eliminates the need for a dedicated network of irradiance sensors and automatically accounts for the effect of temperature and angle of incidence on power forecasts.

In Section 3 we develop a framework to model station-pair correlations. We will show that our framework reproduces the station pair correlations measured in this work as well as correlations reported by several other authors. We then use our framework to predict the forecast accuracy of velocimetry based forecasts that span three orders of magnitude in the forecast horizon (from 30 s to 6 h). Finally, we discuss how to use this framework to devise a strategy to optimally place irradiance sensors for use in forecasting.

2. Forecasting of PV system output

The principal input to the forecasting algorithm described here consists of measurements of PV power output from 80 residential rooftop systems distributed over a 50 km by 50 km area are used to forecast PV power output. Details of the system specifications are given in Lonij et al. (2012b). Measurements are recorded at 15-min intervals, and each measurement represents the average AC power over the previous 15 min. Data presented here are obtained using existing infrastructure. Each of the PV systems in this study uses an inverter by SMA Solar Technology with a data communications card installed to record data. Data are transmitted over the Internet using an SMA Sunny WebBox. Although the forecasts we report here are based on historical data, a real-time forecasting system could be implemented with a software-only modification to our setup.

Fig. 1 shows power output for each of the 80 systems in the Tucson area at three different times. The dark points (indicating low output, due to a cloud) on this day shift

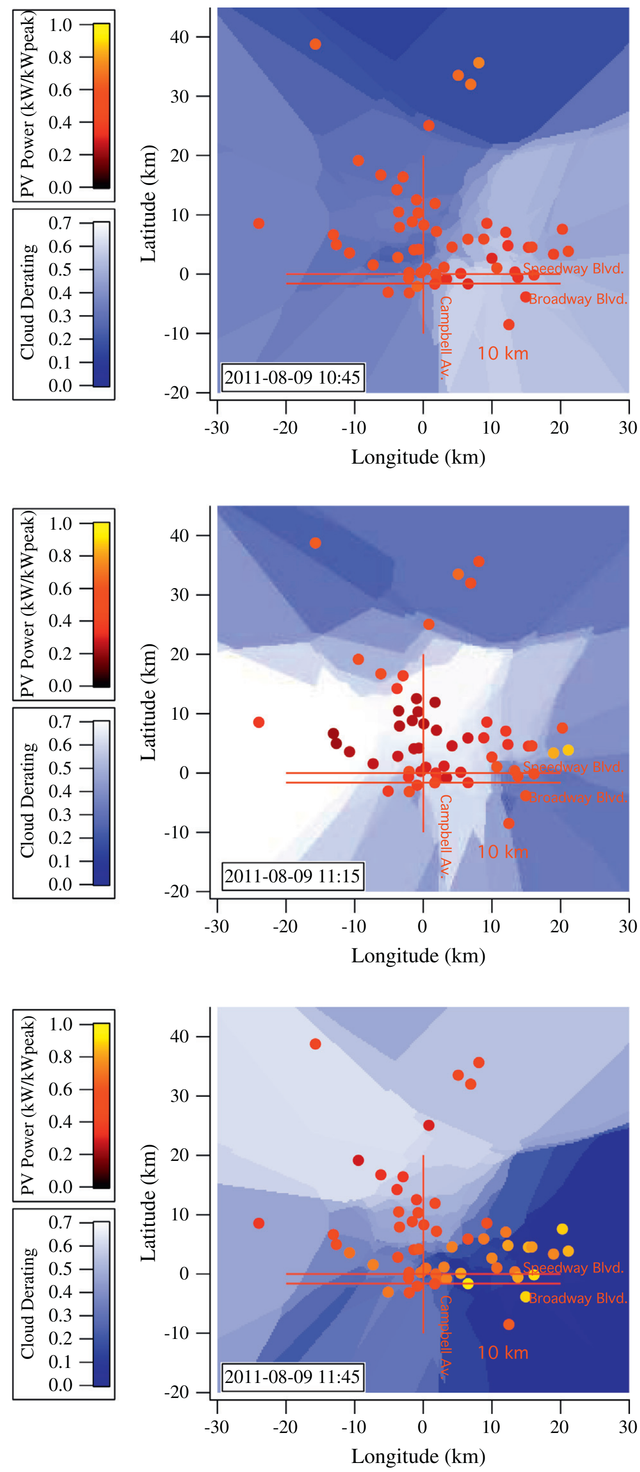

Fig. 1. Measured PV output at three times separated by 30-min. Dark (red) circles indicate stations with low power output, light (yellow) circles indicate stations with high power output. The background color is the interpolated clearness index (K), white areas are cloudy, blue areas are clear. During the period shown, a cloud drifts from the south–east tot the north–west. A movie of this is available at (Lonij et al., 2012a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from the southeast towards the northwest of the Tucson valley over the course of 1 h.

The ground sensor network presented here has an average nearest neighbor spacing of about 3 km and provides measurements every 15 min. This results in better spatial and temporal resolution than currently available operational forecasts based on GOES satellite images (SolarAnywhere, 2012) (whose forecasts have 10 km resolution and are updated approximately every hour).[1]

In our algorithm, once data are collected on a central server, PV output for each system is forecast as follows. First, a clear-sky expectation for the output of each system is obtained, as described by Lonij et al. (2012b). Subsequently, we correct for outages, system orientation, and partial shade due to permanent obstacles (not clouds) (Lonij et al., 2012b). Finally, we infer that deviations from the clear-sky operation of the system are the effect of clouds. Here we define the clearness index K at every location (x, y) and time (t) as

$$K(x,y,t) = \frac{POA(x,y,t)}{POA_{clear}(x,y,t)} \quad (1)$$

where, $POA(t)$ indicates the plane of array (POA) irradiance at time t and $POA_{Clear}(t)$ indicates the modeled POA irradiance in the absence of clouds. Since the output of any particular PV system normalized by peak power (kW/kW_{peak}) is approximately proportional to POA irradiance, we can write

$$K_i = \frac{p_i(t)}{p_{i,clear}(t)} \quad (2)$$

where $p_i(t)$ is the normalized power output (kW/kW_{peak}) for system i and $p_{i,clear}(t)$ is the normalized power that would be generated under a clear sky. Once K has been determined for every system at the present time t, then it is possible to forecast K at time $t + dt$, and at location (x, y) using

$$K(x,y,t+dt) = K(x - v_x dt, y - v_y dt, t) \quad (3)$$

where v_x and v_y are the x and y components of the cloud velocity.

Values of K at locations between the points where PV systems are located are determined by interpolation as follows. For each location (x, y), K is determined for the four closest PV systems $\{K_i\}$. We then compute $K(x, y) =$ MEDIAN $(\{K_i\})$. We use the median instead of, e.g., bilinear interpolation, because cloud edges are relatively sharp compared to the 3 km distance between measurement stations. Therefore picking a representative nearby system to estimate cloud cover is preferable over averaging.

A significant challenge in our forecasts is to determine the cloud edge velocity. Ground-based measurements of wind are not an accurate measure of the velocity of clouds. For example, in Fig. 1, where the clouds progress towards the northwest, ground-level wind measurements indicate wind from the northeast. Therefore we explored four ways of estimating cloud velocity:

1. Wind velocity from the National Oceanic and Atmospheric Administrations Rapid Update Cycle numerical weather model (NOAA RUC NWP) at the grid-point at latitude 32° and longitude −111° at an altitude of 700 mb.
2. A constant velocity for each hour that is numerically optimized (retrospectively) to minimize the RMS error of our power forecast. Note that this is not a true forecast since it uses future information. We use this method to explore how well our power forecasts perform if cloud velocity were perfectly known.
3. A Kalman Filter (Welch and Bishop, 2006; Maybeck, 1979) applied to the velocity determined by method 2. We use the Kalman Filter to extrapolate the data obtained with method 2, 2 h ahead. This method is therefore once again a true forecast. The Kalman Filter is applied separately to the x and y components of the wind vector. We model wind velocity as $v_K(t + \delta t) = v_K(t) + a_K \delta t$. The Kalman Filter returns maximum likelihood estimates for $v_K(t)$ and $a_K(t)$. For our forecast we then use $v(t) = v_K(t - \delta t) + a_K(t - \delta t)\ {}^{*}\delta t$, where $\delta t = 2$ h.
4. The persistence model, which assumes that the clearness index at a future time $t = t_0 + dt$ is the same as the clearness index at t_0. This is equivalent to assuming the cloud velocity is equal to zero.

Table 1 shows RMS error using these three different cloud velocity estimates for time horizons ranging from 15 min to 90 min. Table 2 shows mean bias error (MBE) defined as MEAN $(p_{forecast} - p_{measured})$ (Chow et al., 2011; Perez et al., 2010). Both Tables 1 and 2 show results for cloudy days in the period May 1, 2011 through April 30, 2012. Cloudy days are defined as days where K_i averaged over all systems and averaged over 24 h is less than 0.9. Fig. 2 shows the result of a forecast for one day in August of 2011.

For comparison, Tables 1 and 2 also list results from the persistence model. For time horizons ranging from 30 min

Table 1
RMS errors for power forecasts using forecast horizons ranging from 15 min to 90 min. RMS values are for cloudy days only and in units of normalized PV system output power (kW/kW_{peak}). The results of different methods to determine cloud velocity are shown. Note that method 2. is not a true forecast since it uses future information.

Horizon	NOAA (1.)	Kalman (3.)	Persistence (4.)	Optimized (2.)
15 min	0.065	0.067	0.062	0.061
30 min	0.082	0.082	0.084	0.055
45 min	0.090	0.092	0.097	0.069
60 min	0.098	0.100	0.105	0.079
75 min	0.106	0.107	0.113	0.092
90 min	0.112	0.114	0.120	0.102

[1] Solar Anywhere recently started to offer forecasts with forecast horizons less than one hour as well as 1 km × 1 km spatial resolutions, however, at the time of writing no peer-reviewed validation of these forecasts is available.

Table 2
Mean bias errors ($\times 1000$) for power forecasts using forecast horizons ranging from 15 min to 90 min. MBE values are for cloudy days only and in units of normalized PV system output power (kW/kW_{peak}). The results of different methods to determine cloud velocity are shown. Note that method 2. is not a true forecast since it uses future information.

Horizon	NOAA (1.)	Kalman (3.)	Persistence (4.)	Optimized (2.)
15 min	−2.80	−2.23	0.91	−0.27
30 min	−3.52	−3.14	2.08	−0.56
45 min	−2.58	−2.76	3.49	−0.10
60 min	−1.72	−1.91	5.09	1.07
75 min	0.16	−0.60	6.84	2.86
90 min	2.47	1.21	8.73	6.05

to 90 min, forecasts from our forecast provide smaller RMS errors and mean bias errors than the persistence model. This is significant because other forecasting methods based on NWP models, satellite images (Perez et al., 2010) or all-sky images (Jayadevan et al., 2012; Chow et al., 2011) are currently unable to beat the persistence model at these time-horizons for these (15 min) averaging intervals.

The best true forecast is obtained using cloud edge velocity obtained using method 1 (wind velocity form NOAA RUC model). The improvement over the persistence model is modest (about 8% for 45 min ahead forecasts), though this improvement is comparable to the improvement obtained with satellite based forecasting methods for a 1 h forecast horizon using 1 h averages (Perez et al., 2010).

Although method 2 is not a true forecast (because it uses a retrospectively optimized cloud velocity) it does give us an indication of how well the model can perform if cloud velocity were better known. Method 2 results in an RMS error that is 23% smaller than methods 1 and 3 at a forecast horizon of 45 min. This suggests cloud velocity vectors exist that will improve our forecast of PV system output by 23%. Improved estimates of wind velocity may be obtained in the future from a regionally optimized WRF

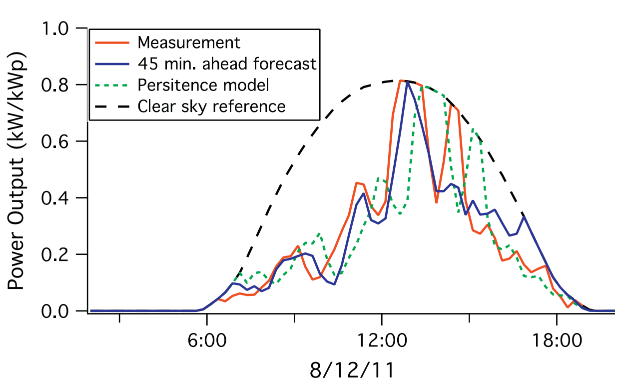

Fig. 2. Measurements and forecasts of PV performance using the network of 80 PV systems. This 45-min ahead forecast is made using NCDC forecast for wind velocity. For the day shown in the figure, the RMS error of the forecast is 0.07. The RMS error of the persistence model is 0.12.

model, ground based observations (e.g. using a camera), or by implementing a sensor network with higher temporal resolution (Bosch et al., 2013).

3. A model for spatio-temporal correlations of clouds

The success of velocimetry based forecasting methods is dependent on the spatio-temporal correlation properties of clouds in the geographical region of interest. In this section, we develop a model that can simultaneously describe correlations of power production as well as correlations for ramp-rates, at multiple time scales and averaging intervals. We use our model to describe our measurements as well as measurements from several other authors (Hoff and Perez, 2012; Perez et al., 2012; Perez et al., 2011; Mills and Wiser, 2010; Hoff and Perez, 2010; Murata et al., 2007; Glasbey et al., 2001). We will use our model to describe RMS forecasting errors and to recommend optimal locations for irradiance sensors to implement the method described in Section 2.

Several authors have studied changes in irradiance (i.e. ramp-rates), rather than cloud derating, and have suggested different empirical formulas to describe the spatio-temporal correlations of these ramp-rates (Perez et al., 2011; Mills and Wiser, 2010; Murata et al., 2007). Autocorrelation functions of ramp-rates in measured irradiance have been studied for individual measurement stations by Hoff and Perez (2012). In addition, correlation functions for ramp-rates have been studied as a function of station pair distance (Mills and Wiser, 2010; Hoff and Perez, 2010). Glasbey et al. (2001) proposed a formulation that included both spatial and temporal dependence of the correlation, however, this formulation did not allow for the possibility of forecasting because they did not include cloud drift. The effect of temporal and geographic smoothing of irradiance variability has been studied using a wavelet approach (Lave and Kleissl, 2013; Lave et al., 2012a,b).

The formulation presented here is able to capture the qualitative features previously reported by several authors including: monotonic decrease of the correlation function for ramp-rates as a function of station pair distance (Mills and Wiser, 2010; Murata et al., 2007), the zero-crossing of this correlation function when measured exclusively in the direction of clouds velocity (Hoff and Perez, 2012), and finally the forecasting accuracy of 3 different velocimetry based forecasting methods, including the results of Section 2 and work in Chow et al. (2011), Perez et al. (2010). Our model accurately describes forecast accuracies over a range of forecast horizons that spans nearly 3 orders of magnitude (from 30 s to 6 h).

3.1. The framework

We define a measurement of clearness index at a location $\vec{x} = \{x, y\}$ at time t as $K(\vec{x}, t)$. We assume the covariance function between two such measurements can be written as

$$Cov[K(\vec{x}_1, t_1), K(\vec{x}_2, t_2)] = f(\Delta\vec{x}, \Delta t, \{s_i\}) \tag{4}$$

where $\Delta t = t_1 - t_2$, $\Delta\vec{x} = \vec{x_1} - \vec{x_2}$, and $\{s_i\}$ is a set of parameters for the specific functional form of f. From now on, we will suppress $\{s_i\}$ from the notation for the sake of readability.

From this formula we will now derive the correlation function for ramp-rates (the derivative $\dot{K} = dK/dt$) as well as the effect of time-averaging of either K or $\dot{K}$. We will use the equation for the covariance between two stochastic variables, that are themselves sums of multiple stochastic variables:

$$Cov\left[\sum_i X_i, \sum_j Y_j\right] = \sum_i \sum_j Cov(X_i, Y_i) \tag{5}$$

where $\{X_i\}$ and $\{Y_i\}$ are sets of arbitrary stochastic variables. Using Eq. (5), we can derive the correlation function for measurements averaged over a time $\bar{t}$ defined as $\overline{K}(\vec{x}, t) = (\bar{t})^{-1} \int_0^{\bar{t}} K(\vec{x}, t + t')dt'$. We find

$$Cov[\overline{K}(\vec{x}_1, t_1), \overline{K}(\vec{x}_2, t_2)] = (\bar{t})^{-2} \int_0^{\bar{t}} \int_0^{\bar{t}} dt dt' f(\Delta\vec{x}, \Delta t + t - t') \tag{6}$$

Similarly we can derive the covariance function for $\dot{K}$ by inserting the definition of the derivative, $\dot{K}(\vec{x}, t) = \lim_{dt\to 0}[K(t + dt) - K(t)]/dt$, into Eq. (5); we find

$$Cov[\dot{K}(\vec{x}_1, t_1), \dot{K}(\vec{x}_2, t_2)] = -\frac{\partial^2 f(\Delta\vec{x}, \Delta t)}{\partial \Delta t^2} \tag{7}$$

For numerical derivatives, such as those studied by Hoff and Perez (2012), Mills and Wiser (2010), Hoff and Perez (2010) and Murata et al. (2007), the data is also averaged. Therefore, we first apply Eq. (6) and then Eq. (7) to find

$$Cov[\dot{\overline{K}}(\vec{x}_1, t_1), \dot{\overline{K}}(\vec{x}_2, t_2)] = -(\bar{t})^{-2} \int_0^{\bar{t}} \int_0^{\bar{t}} dt dt' \times \frac{\partial^2}{\partial \Delta t^2} f(\Delta\vec{x}, \Delta t + t - t') \tag{8}$$

This can be rewritten as

$$Cov[\dot{\overline{K}}(\vec{x}_1, t_1), \dot{\overline{K}}(\vec{x}_2, t_2)] = (\bar{t})^{-2}[f(\Delta\vec{x}, \Delta t + \bar{t}) + f(\Delta\vec{x}, \Delta t - \bar{t}) - 2f(\Delta\vec{x}, \Delta t)] \tag{9}$$

If instead of covariance, we want to study correlation between station pairs we may compute this using

$$\mathrm{Corr}[X, Y] = \frac{Cov[X, Y]}{\sqrt{Cov[X, X]Cov[Y, Y]}} \tag{10}$$

We will now validate this model with a specific expression for f.

3.2. Validation

We will validate the model of Section 3.1 using our measurements described in Section 2. Based on our measurement we will choose a specific functional form for $f(\Delta\vec{x}, \Delta t)$ in Eq. (4).

Correlations between system pairs for $\Delta t = 0$ as a function of north–south as well as east–west separation are shown in Fig. 3 (top). Fig. 3 (bottom) shows the station-pair correlation when Δt is chosen to maximize the correlation for each pair. Fig. 3 (bottom) shows an extended oval region of high correlations corresponding to $\Delta\vec{x} \approx \vec{v_c}\Delta t$, where $\vec{v_c}$ is the cloud velocity. This confirms that there is a drift component in the behavior of clouds. Fig. 4 shows a cross-section of Fig. 3 (top) along the east–west direction.

Based on Figs. 3 and 4 we conclude that for $\Delta t = 0$, the function f must be decreasing in $|\Delta\vec{x}|$. For $\Delta\vec{x} = 0$, the function must be decreasing in Δt. Finally, for a given Δt, the

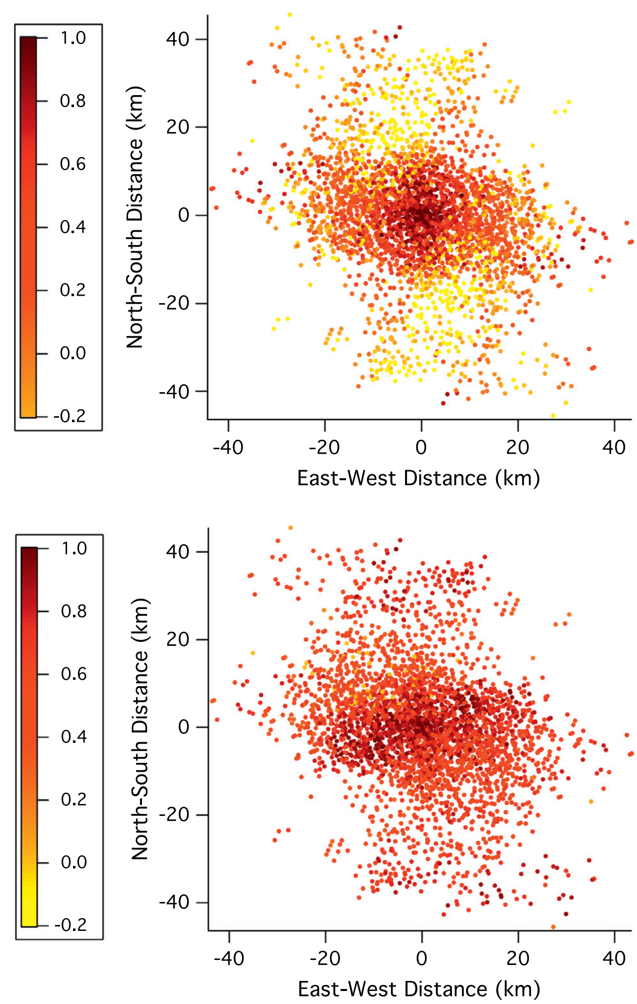

Fig. 3. Top: Spatial correlation between station pairs for one day as a function of separation in the east–west as well as north–south directions. Dark/red points in the center indicate high correlation, light/yellow points indicate low correlation. Bottom: Spatial correlation between station paris with the relative time-shift between stations adjusted to maximize correlation. Correlations are significantly increased for stations that are separated by a vector that is a particular multiple of the cloud velocity, i.e., the central circular region of high correlation in the top figure has become an extended oval. Correlations are calculated using 15 min averages of normalized power. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

covariance function reaches a maximum when $\Delta\vec{x} = \vec{v}_c\Delta t$, where $\vec{v}_c$ is the cloud velocity. The corresponding function for ramp-rates must have a zero-crossing as a function of distance parallel to the wind direction, but has no zero-crossing if averaged over all directions (Hoff and Perez, 2012), see Fig. 5.

There are two prevalent functional forms in the literature to describe covariances of clouds. An exponential form is used by Glasbey et al. (2001) and Mills and Wiser (2010). A function of the form $1/(1 + x/\bar{t})$ is used by Hoff and Perez (2012). The function proposed by Hoff and Perez (2012) does not yield the zero-crossing in Fig. 5, however, as we show below, an exponential function does produce this zero-crossing. We therefore choose and exponential function:

$$f(\Delta\vec{x}, \Delta t) = D_{00} = Ae^{-|\Delta\vec{x}-\vec{v}_c\Delta t|/\sigma_x}e^{-(|\Delta t|/\sigma_t)^q} \quad (11)$$

where σ_x, σ_t, q, and A are parameters to be determined from data. We define D_{01}, D_{10}, and D_{11} as the result of Eqs. (6), (7) and (9) respectively. The first exponential in Eq. (11) describes a drifting cloud pattern with a correlation function that decays as a function of distance. The second exponential describes the changes of the cloud pattern over time, i.e., the degree to which the cloud pattern at time t_1 is not simply a spatially shifted version of the pattern at time t_2. Although this simple functional form does not exactly match the average correlation shown in Fig. 4, considering the large day-to-day variation in the data, we opt for a simpler function in this work that captures all the qualitative features we require.

Fig. 5 shows the correlation function for ramp-rates (D_{11}) as a function of station-pair distance. There is a clear zero-crossing if $\Delta\vec{x}$ is parallel to the wind velocity, but not if $\Delta\vec{x}$ is perpendicular or if D_{11} is averaged over all directions. This behavior is consistent with results from Hoff and Perez (2012) and Mills and Wiser (2010); Murata et al., 2007. Furthermore, we can calculate the location of the zero-crossing as a function of averaging interval by neglecting the second exponential in Eq. (11), plugging this into Eq. (9) and setting the result equal to zero and solving for $\Delta x_{\parallel}$. This gives

$$\Delta x = \frac{\sigma_x}{2}\log(2\exp(\bar{t}|v_c|/\sigma_x) - 1) \approx \sigma_x\frac{\log(2)}{2} + \frac{\bar{t}|v_c|}{2} \quad (12)$$

The approximately linear dependence on $\bar{t}$ is consistent with measurements by Perez et al. (2012).

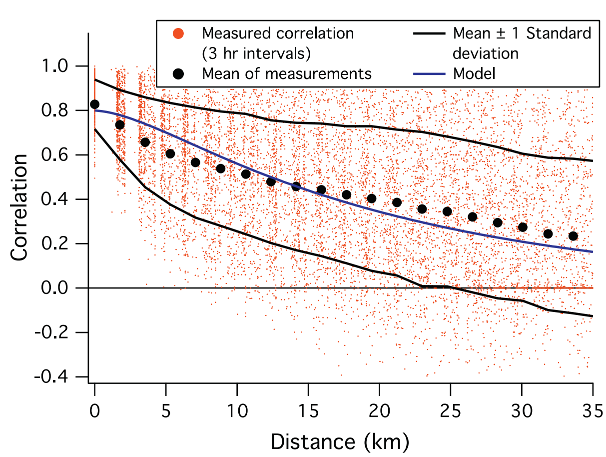

Fig. 4. A cross-section of Fig. 3 (top) in the east–west direction. The dots indicate measurement from 3 h periods. One year of data is shown. The black circles are averages for each distance. The figure shows that, on average, the correlation decreases for larger $\Delta\vec{x}$, but there is also a large spread. The model of Eq. (6) is shown as well. Correlations are calculated using 15 min averages of normalized power.

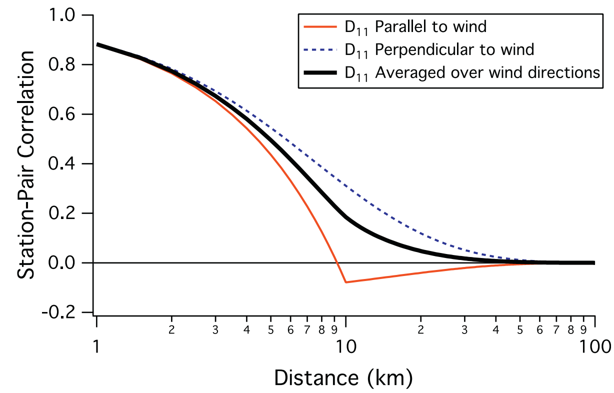

Fig. 5. Correlations between derivatives (ramp-rates) between different stations are modeled. The model reproduces both the monotonically decreasing correlation as a function of distance when averaged over all directions (Mills and Wiser, 2010; Murata et al., 2007), as well as the zero-crossing and subsequent minimum of the correlation if the station separation is parallel to the wind direction (Hoff and Perez, 2012).

3.3. Forecast errors

In this section we model irradiance forecast errors for velocimetry based forecasting methods using the framework developed in Section 3.1. To this end, an expression for the Mean Square Error (MSE) is first derived in terms of the covariance between a forecast m_f and a measurement m_s. By definition, MSE is given by

$$\mathrm{MSE} = \sum_{i=0}^{N}(m_{f,i} - m_{s,i})^2/N \quad (13)$$

where the subscript i refers to measurements at different times. Given that average and variance of derating due to clouds is similar for measurements in the same climatic region, it is reasonable to assume that the mean μ_f of the forecast and μ_s of the measurement are the same. We can therefore insert μ_f–μ_s inside the parentheses of Eq. (13) to obtain

$$\mathrm{MSE} \approx \sum_{i=0}^{N}(m_{f,i} - m_{s,i} - (\mu_f - \mu_s))^2/N$$
$$= Var(m_f - m_s) \quad (14)$$

We can rewrite the right hand side of this equation as

$$\mathrm{MSE} = Var(m_f) + Var(m_s) - 2Cov(m_f, m_s) \quad (15)$$

Using D_{01} to compute the variance and covariance in Eq. (15) we can plot the predicted RMSE along with the observed RMSE. Fig. 6 shows the result of the model as well as RMSE observed in different forecasting methods. The parameters σ_t, A, and q in Eq. (11) are adjusted to match the RMS errors for the method presented in Section 2 (the red circles in Fig. 6).

The RMS errors from other authors are scaled by a single factor to match with the model. This is justified because the other studies were conducted in different climate zones and at different times (e.g. time of day, or time of year). The dependence of RMS errors on forecast horizon for all measurements shown in Fig. 6 is consistent with the model.

Thus, we have demonstrated that we can predict RMS errors for a given forecast horizon. In the next Section we use this to determine the optimal locations for irradiance measurement stations.

3.4. Example: recommendation of sensor locations

In Section 2, the irradiance sensor (PV system) locations were predetermined by existing hardware installations. In future implementations of our forecasting method it is possible that no existing hardware is available, or so many sensors are already available that a subset must be chosen, in these situations it is useful to determine optimal sensor locations.

We now discuss optimal sensor placement in the case where forecasts are needed for a single large PV plant. A trade-off is possible between the number of sensors in the network (and therefore the cost), the accuracy of the forecasts, and the maximum forecasting horizon. We assume that wind can come from any direction with equal probability. Therefore, we will design a point-symmetric sensor pattern, i.e., the PV system is located at the origin and the sensors are located on concentric circles around the PV system.

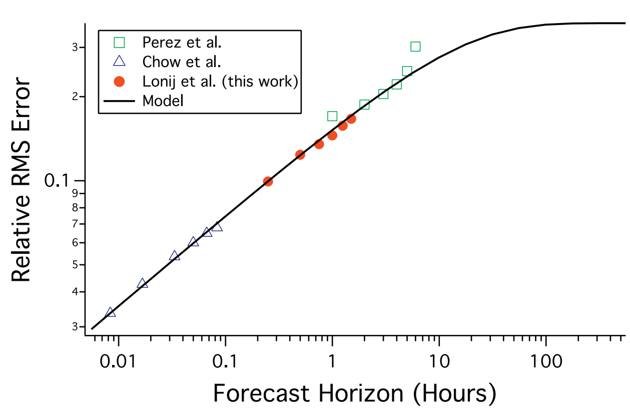

Fig. 6. RMS error as a function of forecast horizon for three different experiments. Triangles represent forecasts from Chow et al. (2011), squares represent forecasts from Perez et al. (2010), and circles represent the present study. Data from Chow et al. (2011) and Perez et al. (2010) are corrected for averaging interval using Eq. (6) and multiplied by 0.1 and 0.45 respectively. RMS error is also modeled using Eq. (15).

First, we first determine the spacing between sensors, both in the radial and axial directions. At one instant (i.e. for a constant cloud velocity vector) the radial spacing corresponds to spacing in the direction of cloud motion and axial spacing corresponds to spacing in the perpendicular direction.

To make forecasts, we use Eq. (3), which assumes that there is a station located at $\vec{x} = -\vec{v}_c dt$. If there is no station at that location, we can use the average of two nearby stations:

$$m_s(\vec{x}) = m_s(\vec{x}_i)/2 + m_s(\vec{x}_j)/2 \tag{16}$$

where $\vec{x}_i$ and $\vec{x}_j$ indicate the locations of the two nearest irradiance sensors. If we plug this expression for $m_s(\vec{x})$ into Eq. (15) and use Eq. (5) to expand the covariance and variance terms we find

$$\begin{aligned}\mathrm{MSE} = &\frac{3}{2} Var(m_s(\vec{x}_i)) + \frac{1}{2} Cov(m_s(\vec{x}_i), m_s(\vec{x}_j)) \\ &- 2[Cov(m_f, m_s(\vec{x}_i))/2 + Cov(m_f, m_s(\vec{x}_j))/2]\end{aligned} \tag{17}$$

We can use Eq. (17) to calculate the spacing between stations if we assume that x_i and x_j are equidistant from x (the worst case scenario) and set the RMSE to be no more than a given threshold (e.g. 3% higher than the minimum RMSE). Using the expression for covariance in Eq. (6) and the expression for f in Eq. (11), Eq. (17) can be solved numerically to yield the maximum distance between x_i and x_j. We can follow this procedure for station separations either perpendicular or parallel to the direction of cloud propagation.

Finally, we have to choose and averaging time $(\bar{t})$; we assume that for larger forecast horizons it is acceptable

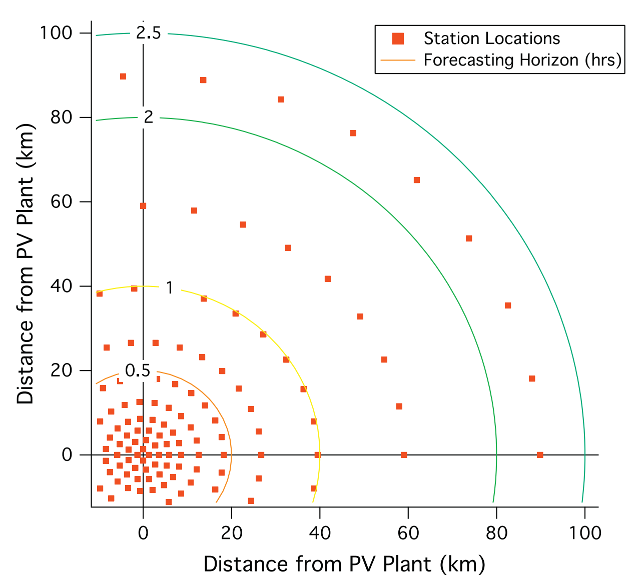

Fig. 7. Station locations calculated by numerically evaluating Eq. (17). The equivalent forecasting horizons, assuming a cloud velocity of 40 km/h are also shown.

Table 3
Parameters used in Eq. (11).

Parameter	Value
A	0.024
σ_x	20 km
σ_t	6 h
q	0.65

to use larger averaging intervals. This reflects the fact that small clouds are less significant when they are further away. Here we choose $\bar{t} = \Delta t/4$. Fig. 7 shows the results for the parameters shown in Table 3.

4. Conclusion

We presented results of an irradiance forecasting method that uses measurements from a network of residential PV systems. Forecasts using 15-min interval measurements from a network of distributed PV systems outperform the persistence model for forecast horizons ranging from 30 min up to 90 min. This is an improvement over forecasts based on satellite images, which are not available at 15 min intervals for forecasts horizons ranging from 30 min up to 90 min.

We observed that the RMS error of our forecasts for may be improved by 23% (for forecast horizon of 45 min) if we can find better ways to determine cloud velocity. This shows the overall capability of our model. Determining cloud velocity from measured PV data is challenging for our data set because the geographical area spanned by our dataset is small relative to the time resolution of our measurements. Simply stated, this is because it takes only a few time steps (1 h or so) for clouds to transit across our entire 50-km network.

Using wind velocities obtained from numerical weather models gives improved results. However, because cloud edge velocity is not always the same as wind velocity, there are still significant errors. In future work we will therefore explore other techniques to determine cloud velocity, including analysis of cloud images from satellites and from ground-based cameras. We will also explore if we can use a numerical weather models to predict on which days velocimetry based forecasts will perform well and on which days velocimetry based forecasts produce large errors.

Finally, we developed a framework to describe spatio-temporal correlations of cloud deratings. This framework is able to reproduce results from several different authors who studied different geographical regions. The model is able to describe both correlations in power output as well as correlations of ramp-rates.

We applied our framework to model forecasting errors and were able to predict RMS forecasting errors over three orders of magnitude in the forecasting horizon (from 30 s to 6 h). We also used our framework to determine optimal irradiance sensor locations for future implementations of our forecasting method.

Role of the funding sources

This work was funded by Tucson Electric Power (TEP), the University of Arizona, and the Arizona Research Institute for Solar Energy. TEP suggested that forecasts of PV power production could be useful if possible. The funding sources had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

References

Bosch, J.L., Zheng, Y., Kleissl, J., 2013. Deriving cloud velocity from an array of solar radiation measurements. Solar Energy 87, 196–203.

Chow, C.W., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., Washom, B., 2011. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. Solar Energy 85, 2881–2893. http://dx.doi.org/10.1016/j.solener.2011.08.025.

Glasbey, C., Graham, R., Hunter, A., 2001. Spatio-temporal variability of solar energy across a region: a statistical modelling approach. Solar Energy 70, 373–381. http://dx.doi.org/10.1016/S0038-092X(00)00152-3.

Gowrisankaran, G., Reynolds, S.S., Samano, M., 2011. Intermittency and the value of renewable energy. NBER Working Paper, No. 17086.

Hamill, T.M., Nerhkorn, T., 1993. A short-term cloud forecast scheme using cross correlations. Weather and Forecasting 8, 401–411.

Hoff, T.E., Perez, R., 2010. Quantifying PV power output variability. Solar Energy 84, 1782–1793. http://dx.doi.org/10.1016/j.solener.2010.07.003.

Hoff, T.E., Perez, R., 2012. Modeling PV fleet output variability. Solar Energy 86, 2177–2189. http://dx.doi.org/10.1016/j.solener.2011.11.005.

Jayadevan, V.T., Lonij, V.P.A., Rodriguez, J.J., Cronin, A.D., 2012. Forecasting solar power intermittency using ground-based cloud imaging, in: World Renewable Energy Forum, Denver, 2012.

Lave, M., Kleissl, J., 2013. Cloud speed impact on solar variability scaling – application to the wavelet variability model. Solar Energy 91, 11–21. http://dx.doi.org/10.1016/j.solener.2013.01.023.

Lave, M., Kleissl, J., Arias-Castro, E., 2012a. High-frequency irradiance fluctuations and geographic smoothing. Solar Energy 86, 2190–2199. http://dx.doi.org/10.1016/j.solener.2011.06.031.

Lave, M., Kleissl, J., Stein, J., 2012b. A wavelet-based variability model (wvm) for solar pv powerplants. IEEE Transactions on Sustainable Energy. http://dx.doi.org/10.1109/TSTE.2012.2205716.

Lonij, V.P.A., Brooks, A.E., Cronin, A.D., 2012a. Photovoltaics Group, University of Arizona. <http://www.uapv.org/>.

Lonij, V.P.A., Brooks, A.E., Koch, K., Cronin, A.D., 2012b. Analysis of 100 rooftop PV systems in the Tucson, AZ area. In: Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE.

Lonij, V.P.A., Jayadevan, V.T., Brooks, A.E., Rodriguez, J.J., Koch, K., Leuthold, M., Cronin, A.D., 2012c. Forecasts of PV power output using power measurements of 80 residential PV installs. In: Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE.

Maybeck, P.S., 1979. Stochastic models, estimation, and control, Mathematics in Science and Engineering, vol. 141.

Mills, A., Wiser, R., 2010. Implications of wide-area geographic diversity for short-term variability of solar power. DOE Report LBNL-3884E.

Müller, R., Dagestad, K., Ineichen, P., SchroedterHomscheidt, M., Cros, S., Dumortier, D., Kuhlemann, R., Olseth, J., Piernavieja, G., Reise, C., Wald, L., Heinemann, D., 2004. Rethinking satellite-based solar irradiance modelling. The SOLIS clear-sky module. Remote Sensing of Environment 91, 160–174.

Murata, A., Yamaguchi, H., Otani, K., 2007. A method to estimate the output fluctuation of many photovoltaic power generation systems dispersed in a wide area. IEEJ Transactions on Power and Energy 127, 645–652. http://dx.doi.org/10.1541/ieejpes.127.645.

NAOE, 2008. Grand challenges for engineering. National Academy of Engineering. <http://www.engineeringchallenges.org>.

Perez, R., Ineichen, P., Kmiecik, M., Moore, K., Renne, D., George, R., 2004. Producing satellite-derived irradiances in complex arid terrain. Solar Energy 77, 367–371. http://dx.doi.org/10.1016/j.solener.2003.12.016.

Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., K., Renne, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US. Solar Energy 84, 2161–2172. http://dx.doi.org/10.1016/j.solener.2010.08.014.

Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., K., Hoff, T., 2011. Parameterization of site-specific short-term irradiance variability. Solar Energy 85, 1343–1353. http://dx.doi.org/10.1016/j.solener.2011.03.016.

Perez, R., Kivalov, S., Schlemmer, J., Hemker Jr., K., Hoff, T.E., 2012. Short-term irradiance variability: preliminary estimation of station pair correlation as a function of distance. Solar Energy 86, 2170–2176. http://dx.doi.org/10.1016/j.solener.2012.02.027.

SolarAnywhere, Clean Power Research, 2012. <http://www.solaranywhere.com>.

Welch, G., Bishop, G., 2006. An Introduction to the Kalman Filter.