Ricardo Marquez
Mechanical Engineering and Applied Mechanics Program,
School of Engineering,
University of California,
Merced, CA 95343

Carlos F. M. Coimbra[1]
Mem. ASME
Department of Mechanical and Aerospace Engineering,
Jacobs School of Engineering,
Center of Excellence in Renewable Resource Integration,
University of California,
San Diego, La Jolla, CA 92093
e-mail: ccoimbra@ucsd.edu

Proposed Metric for Evaluation of Solar Forecasting Models

This work presents an alternative metric for evaluating the quality of solar forecasting models. Some conventional approaches use quantities such as the root-mean-square-error (RMSE) and/or correlation coefficients to evaluate model quality. The direct use of statistical quantities to assign forecasting quality can be misleading because these metrics do not convey a measure of the variability of the time-series for the solar irradiance data. In contrast, the quality metric proposed here, which is defined as the ratio of solar uncertainty to solar variability, compares the forecasting error with the solar variability directly. By making the forecasting error to variability comparisons for different time windows, we show that this ratio is essentially a statistical invariant for each forecast model employed, i.e., the ratio is preserved for widely different time horizons when the same time averaging periods are used, and therefore provides a robust way to compare solar forecasting skills. We employ the proposed metric to evaluate two new forecasting models proposed here, and compare their performances with a persistence model. [DOI: 10.1115/1.4007496]

1 Introduction

One of the leading impediments for achieving higher market penetration and power grid connectivity of solar and wind technologies is the variable nature of such resources. Recent studies on higher penetration impact of renewables have emphasized the need for accurate forecasts if large variable capacities are to be achieved [1–3]. Historically, most of the power grid variability has been on the load (demand) side because fossil and nuclear generation are designed to operate in stable, dispatchable and controllable fashion. In order to increase renewable power capacity penetration, independent system operators must now cope with generation variability. Therefore load forecasting errors have economic consequences on electricity markets, as well as other operational impacts [2].

Current efforts in developing forecast models include the development of artificial neural networks (ANNs) for time series predictions, semi-empirical models based on satellite images and predictions based on national or regional weather predictions (see, e.g., Refs. [4–6]). Many of the existing forecast models have been shown to predict the 1-h integrated solar irradiance 1- or 2-h ahead to good accuracy [7–9] and forecast models for same-day [4,6], 1–5 days ahead have also been developed with good accuracy and reliability. Effective forecasting methodologies and relevant inputs considered can vary widely depending on the time horizons considered.

However, consideration of the relative advantages of different forecasting methodologies is not straightforward because different authors use different evaluation criteria, and also because the solar radiation data sets are dependent on geographic location, time of year, and climate. The second issue is particularly significant as it is relatively easier to forecast the solar irradiance during clear day periods, and therefore an oversimplistic forecast model can yield very good conventional statistical metrics for those conditions when solar irradiation is highly predictable. The same oversimplistic forecast model would certainly fail under different geoclimatic conditions. The purpose of this work is to present an alternative approach to evaluate the quality of forecast models by defining and quantitatively comparing the solar resource variability (V) and the forecast uncertainty (U). The observation that leads to the relationship between V and U is that forecast model errors are typically higher during wet (cloudy) days than during dry (clear) days, as clearly demonstrated in a few recent studies [4,6]. In this paper, we analyze forecast models and quantify the relationship between "variability" and "uncertainty" in order to produce a consistent metric that is independent of the time horizon under consideration.

Since the forecast quality measures that is introduced depends on using clear sky solar irradiation model as a normalization factor to compute V and U and to define a persistence model, we consider two clear sky models. We also test a definition of V and U and a persistence model based on the extraterrestrial solar irradiance so that, unlike the clear sky models, location dependent parameters can be avoided. The clear sky and persistence models are described in Secs. 2 and 3, respectively. In Sec. 4, the solar irradiance variability is defined and discussed following the work from Refs. [10,11]. After these preliminaries, the goal here is to evaluate the forecasting using a more robust and meaningful metrics which is based on determining the average forecasting skill ($\langle S \rangle$). In Sec. 5, we describe the forecasting model evaluation procedure. In Sec. 6, two forecasting models are developed and, along with the persistence models, are evaluated in terms of the conventional and proposed evaluation metrics. The rest of the paper discusses further forecasting evaluations for various time-horizons and other possible applications of understanding forecast uncertainty and variability.

2 Clear-Sky Models

This section describes two clear sky models which are used for persistence forecasts and for normalization in the evaluation metric described later. The parameters of the clear sky models are site specific to the location of the study (Merced, CA.) and for the period of the study (Jan. 1–Oct. 31, 2010) and therefore may not be accurate for other locations. The interested reader can find other well studied clear sky models from [12]. Here we only choose two clear sky models, one which is a based on a polynomial regression and the other based on the European Solar Radiation Atlas (ESRA) clear sky model, so that the sensitivity of the methods for forecasting of clear sky days can be examined. The sensitivity should be small since short-term forecasting involves the use of previous irradiance values that can be used to adjust the clear sky predictions. This particular feature of short-term forecasting is shown using a persistence model in Sec. 3.

[1]Corresponding author.

Contributed by the Solar Energy Division of ASME for publication in the JOURNAL OF SOLAR ENERGY ENGINEERING. Manuscript received April 25, 2011; final manuscript received July 2, 2012; published online October 23, 2012. Assoc. Editor: Carsten Hoyer-Klick.

Table 1 Comparing clear sky and extraterrestrial models with clear sky and clearness persistence models

	RMSE (W/m^2)		MBE (W/m^2)	
	I_{clr}	$\hat{I}_p$	I_{clr}	$\hat{I}_p$
Polynomial	28.1	12.9	−12.8	−0.0906
ESRA	26.6	20.7	7.44	−0.186
Extraterrestrial	257	31.5	−252	7.28

2.1 CS Model 1: Polynomial Fit. The first clear sky model that we are considering is a third degree polynomial function of cosine of the solar zenith angle $\cos(\theta)$

$$I_{clr,poly} = c_3(\cos\theta)^3 + c_2(\cos\theta)^2 + c_1(\cos\theta) + c_0 \quad (1)$$

where the coefficients where determined as: $c_3 = 688$, $c_2 = 1103$, $c_1 = 518$, $c_0 = -3$, and $I_{clr,poly}$ is in W/m^2. This model was generated using selected clear sky days in the data set collected in the year 2010 at the University of California Merced solar observatory station. The accuracy for a few selected days is shown in Table 1.

2.2 CS Model 2: ESRA. As a second clear sky model, $I_{clr,ESRA}$ is computed using the ESRA model adopted from Refs. [13,14]. The algorithms used are developed by Ref. [13] and were obtained from the accompanying c.d. to Ref. [14]. The ESRA model only depends on a site dependent Linke-Turbidity factor (T_L). Here, T_L is set to $T_L = 4$ so that the bias for all of the clear sky days of the entire data set is minimized.

The polynomial and ESRA clear sky models are compared in Fig. 1 which shows that models do not exactly match. Although their correlation is high, there is a spread of RMSE = $14.7 W/m^2$. It is intentional to have a difference between the polynomial and the ESRA clear sky models so that we can test the resulting sensitivity to the evaluations of the developed forecasting models.

3 Persistence Models

3.1 Clear Sky Persistence Models. The clear sky persistence models are defined as having the clear sky conditions (ratio between the measured irradiance to the clear sky irradiance) persist for the next time-step

$$\hat{k}(t+\Delta t) = k^*(t) = \frac{I(t)}{I(t)_{clr}} \quad (2)$$

From $\hat{k}(t+\Delta t)$, the prediction for the next value of solar irradiance value $I(t+\Delta t)$ is computed as

$$\hat{I}(t+\Delta t)_p = \hat{k}(t+\Delta t) I(t+\Delta t)_{clr} \quad (3)$$

This simple persistence model only relies on I_{clr} which is time and location dependent. Since the persistence model only relies on I_{clr}, there can also be several persistence models depending on how I_{clr} is estimated. Here we consider two persistence models based on the clear sky models already presented

$$\hat{I}(t+\Delta t)_{poly} = k_{poly}(t) I_{clr,poly}(t+\Delta t) \quad (4)$$

and

$$\hat{I}(t+\Delta t)_{ESRA} = k_{ESRA}(t) I_{clr,ESRA}(t+\Delta t) \quad (5)$$

An illustration of the persistence model, $\hat{I}(t+\Delta t)_{poly}$, is shown in Fig. 2 which shows two prominent characteristics of persistence model in general. The first characteristic, is that large forecasting errors mainly occur during abrupt changes in $I(t)$ which are a result of passing opaque clouds. The second characteristic is the obvious lag of the abrupt change as can be clearly seen in the second day in Fig. 2. During cloudless periods, the persistence model is extremely effective as a forecast model as there is very little error resulting from the persistence of the clear sky values.

3.2 Clearness Persistence Model. As a slight variation of the persistence models just described, we also apply a *clearness* persistence model which is based on the extraterrestrial solar radiation (I_0) rather than I_{clr}, that is computed in W/m^2 as

$$I_0 = G_0 \cos(\theta_z) \quad (6)$$

where $G_0 = 1367\ W/m^2$ is the solar constant. Moreover, this persistence does not involve any location-dependent fitting parameters as the clear sky models, however, when compared to irradiance the I_0 has a large positive bias I_0. The resulting clearness-based persistence model is similarly defined as

$$\hat{I}(t+\Delta t)_0 = k_0(t) I_0(t+\Delta t) \quad (7)$$

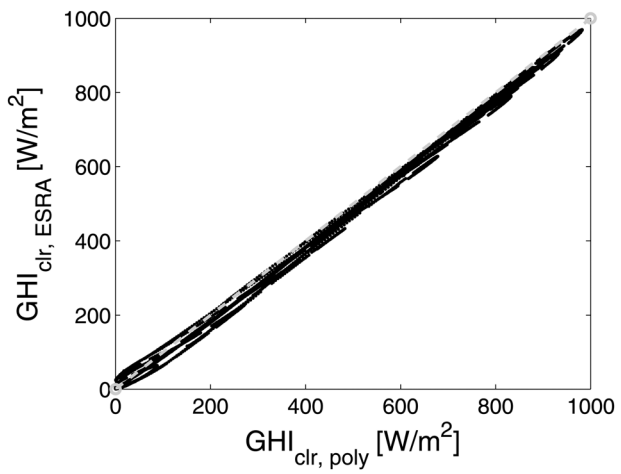

Fig. 1 Comparison of ESRA and polynomial-fit clear-sky models. The coefficient of determination between the two models is (R^2) = 0.998 and the RMSE is RMSE = 14.7 W/m^2.

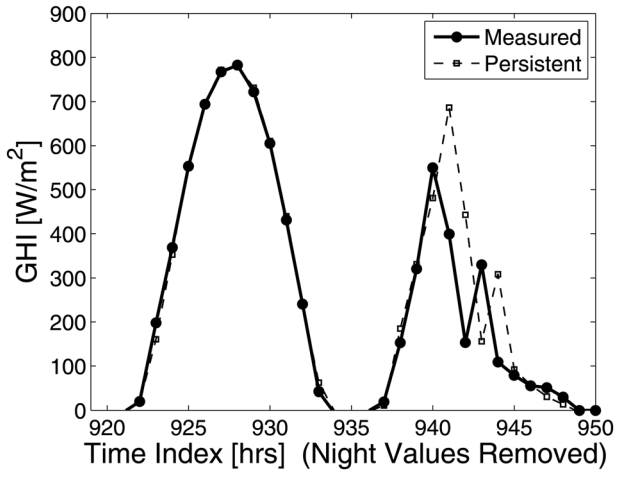

Fig. 2 Example of persistent model performance for a clear and a partially cloudy day (Mar. 20–21, 2010). The clear day is approximated very well by persistent model, whereas a "time delay" is observed for the partially cloudy day.

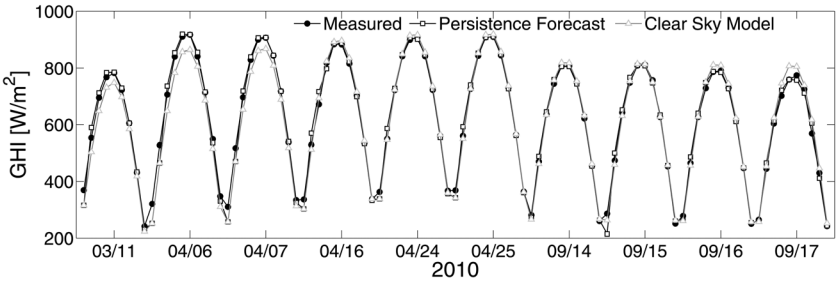

Fig. 3 Measured, modeled, and forecasted clear sky days arbitrarily selected for 2010. This figure illustrates the improved accuracy of a clear sky persistence forecast model over original clear sky model. The RMSEs are, respectively, 20.7 W/m^2 and 26.6 W/m^2 for the clear sky persistence forecast model and the original model.

Applying this persistence model gives us an extreme case, where the normalization factor has a large positive bias compared to measured clear sky irradiance values and will be useful to evaluate the sensitivity of how accurate clear sky models need to be for short-term forecasting.

3.3 Clear Sky Modeling and Forecasting Accuracy. This section is intended to clarify the difference between the accuracy of clear sky models (I_{clr}) and persistence forecast models ($\hat{I}_p$) during clear days. The main difference is that the persistence models utilize the clear sky (or clearness) index at time t as an indication of the irradiance value at $t + \Delta t$, whereas the clear sky model does not use this information. Clear sky models are generally used with satellite-based modeling to predict irradiance for location where no other measurements exist, however, in forecasting applications, on-site solar irradiance measurements are generally available. As a result, the persistence models are more accurate than the clear sky models.

To demonstrate, consider the arbitrarily selected clear days shown in Fig. 3 which shows the measured hourly average values of the solar irradiance $I(t)$, the polynomial based clear sky model $I_{\text{clr,poly}}$, and the polynomial based persistence forecast model $\hat{I}_{\text{poly}}$. As this graph shows, the clear sky model fails to overlap well the irradiance values for a few of the days while on other days there is close overlap. The persistence, on the other hand, is fairly accurate for each of the days shown. The same analysis can be done with other clear sky (or extraterrestrial) and persistence models. Applying the RMSE (defined in Sec. 5) and mean-bias-error (MBE) as levels of accuracy for the days shown in Fig. 3, we get the results shown in Table 1. According to the obtained values, the persistence-based forecasts are much more accurate than the accuracy of a clear sky model indicate. Similar results can be obtained for other forecasting horizons besides 1-h. Based on our experience, clear sky modeling accuracy is not critical for short-term forecasting applications.

4 Solar Resource Variability

The variability of solar irradiance at the ground level is due to several factors such as the presence of participating gases in the atmosphere (CO_2, H_2O, etc.), aerosols, cloud cover, and solar position [14]. Most of the solar variability, however, is due to the latter two factors. The variability due to solar position is completely deterministic while the variability due to clouds is considered mostly stochastic because precise models for cloud dynamics have been proven elusive. The portion of solar variability that is of most concern to forecast models is the cloud-induced (or stochastic) component [3,10,15], thus we refer to solar variability as the standard deviation of the step-changes of the ratio of the measured solar irradiance to that of a clear sky solar irradiance so that the diurnal variability is neglected

$$V = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{I(t)}{I_{\text{clr}}(t)} - \frac{I(t-1)}{I_{\text{clr}}(t-1)}\right)^2} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\Delta k(t))^2} \quad (8)$$

In the above definition for V, k is either the clear sky or clearness index depending on whether we use the $I_{\text{clr,poly}}$, $I_{\text{clr,ESRA}}$ or I_0. This formulation of variability is essentially the same as in Refs. [11,16], except for the modification to include the deterministic changes ($\Delta k(t)$) as is done in Refs. [10,15]. For very small time intervals of less than 5 min this modification is not too important because the deterministic (solar position dependent) variations are small. Figure 4 shows values of $\Delta k(t)$ for a sequence of clear and cloudy days. For clear days, the fluctuations $\Delta k(t)$ are much smaller than for the cloudy days where large ramps are apparent in the $\Delta k(t)$ time-series signal.

5 Uncertainty in Solar Resource: Forecasting Errors

5.1 Conventional Statistical Metrics to Characterize Model Quality. The coefficient of determination is a comparison of the variance of the errors to the variance of the data which is to be modeled

$$R^2 = 1 - \frac{\text{Var}(\hat{I} - I)}{\text{Var}(I)} \quad (9)$$

where the denominator in the second term of the right hand side of the equation is the calculated variance $\text{Var}(I)$ of the data set, not to be mistaken for the variability ($\approx \sqrt{\text{Var}(\Delta k)}$).

Another common evaluation metric is the RMSE which yields a measure of the average spread of the forecasting errors. The RMSE is calculated as,

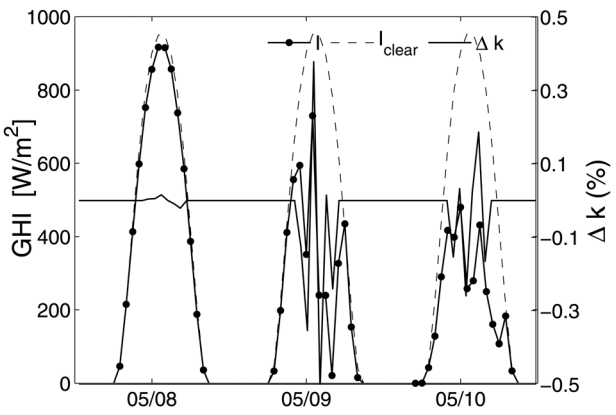

Fig. 4 Time series of global horizontal irradiance (*I*) values, estimated clear-sky *I* and calculated values of stochastic step changes, Δk (Data for May 8–10, 2010 in Merced, CA)

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{I}(t) - I(t)\right)^2} \tag{10}$$

where the summation is carried over the entire data set. Typically night values are removed in the above calculations of R^2 and RMSE and only 1 value is given to summarize the entire data set. Both the RMSE and the R^2 do not help to quantify the amount of variability actually in the irradiance data. Other solar forecasting metrics include normalization of RMSE (relative RMSE), the MBE, the mean absolute error, the mean absolute percentage error, and the correlation coefficient (ρ). Again, none of these metrics embed a sense of variability of the irradiance time-series data.

5.2 Proposed Metric. Here we define the uncertainty as the standard deviation of a model forecast error divided by the estimated clear sky value of the solar irradiance over a subset time window of N_w data points

$$U = \sqrt{\frac{1}{N_w}\sum_{t=1}^{N_w}\left(\frac{\hat{I}(t) - I(t)}{I_{\text{clr}}(t)}\right)^2} \tag{11}$$

This definition is related to the commonly used RMSE [4,6,8,9] where sometimes relative or normalized RMSE is used [4,6,9] with respect to the average irradiance. The present definition of solar resource uncertainty is closely related to those except that in our definition, the normalization is made with respect to I_{clr} for which we will use $I_{\text{clr,poly}}$, $I_{\text{clr,ESRA}}$ or I_0 in our evaluations.

The following metric directly evaluates the variability that is effectively reduced by the forecasting models by taking the difference between U and V and normalizing it with respect to V

$$s = \frac{V - U}{V} \tag{12}$$

where U and V are calculated over the same data set. The metric for evaluating the quality of forecast models is more simply computed by considering the ratio between uncertainty, U, and variability, V, such that

$$s = 1 - \frac{U}{V} \tag{13}$$

The forecast quality measure is defined above is such that when $s = 1$ it means that the solar irradiance is perfectly forecasted, and when $s = 0$ the solar irradiance variability dominates the forecast. By definition, the persistence model should have a forecast quality measure of $s = 0$, as can be shown by comparing Eqs. (3), (8), and (11). The ratio U/V would be then defined as a measure of the benefit of a forecast method with respect to the persistency forecast based on the clear sky index. If s is negative for a developed forecast model, then that model performs worse than a persistence forecast. A typical forecast model should be characterized with values between 0 and 1, with higher values indicating better forecasting.

Since U and V are random variables, it follows that s is also a random variable. To obtain a representative value of s, we take the average value $\langle s \rangle$ as the indication of forecast skill. The average is obtained by calculating U and V for various time-window subsets. If a time window contains a large number of clear days, then both U (forecasting error) and V (the solar irradiance variability) will be small for that time window, thereby, the relative amount of error to variability is preserved. The time windows are selected by fixing N_w (the window size) and computing U_j and V_j over each jth window in the time series. The window partitioning is illustrated in Fig. 5, showing solar irradiance and calculated

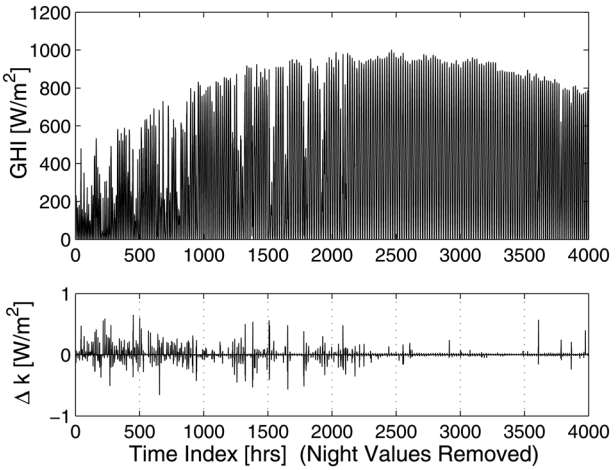

Fig. 5 Time series of solar irradiance and Δk. The figure illustrates the partition of the time series into window sizes of $N_w = 500$ h. Each dashed vertical line represents the boundaries of the 500-h time windows.

values of Δk_{poly} for the data set spanning Jan. 1–Oct. 31. Low confidence experimental values due electrical power issues that occurred in May and July were removed from the calculations. In the case shown in Fig. 5, the window size is $N_w = 500$ and are separated by the dashed vertical bars. As mentioned above, night values are not included in Fig. 5, nor are they used in the calculations below.

6 Application of Proposed Metric to Forecast Model Evaluation

The metric proposed is now applied to solar forecasting models based on ANNs. We employ feedforward ANNs to approximate future hourly values of the $I(t)$ using lagged values of the time series. In the first case, we use only the time-series of hourly averaged $I(t)$, and in the second case, we use more information of $I(t)$ by including 30-min and 6-min moving averages and standard deviations computed from 30-s interval data. In this work, we do not consider applying an input selection scheme to determine the most relevant inputs (see e.g., Ref. [4]), but rather limit the scope to determining whether or not including several more inputs than just the hourly averaged values of $I(t)$ time series alone will improve the forecasting performance of the models. The forecasting performances are evaluated based on the conventional and the proposed metric (s) in order to compare and contrast the quality of the models.

6.1 NAR and NARX Forecasting Models. The forecast model including only the hourly averaged $I(t)$ time-series as an input is referred to as the nonlinear autoregressive (NAR) forecasting model, and the model including additional inputs is referred to as the nonlinear autoregressive with exogenous inputs (NARX) forecasting model. The NAR model for 1-h ahead predictions can be mathematically expressed as

$$\hat{I}(t + \Delta t) = f(I(t), I(t - \Delta t), \ldots, I(t - n\Delta t)) \tag{14}$$

where $n + 1$ is the number of time delays of the time series $I(t)$ which are included as inputs to predict $I(t + \Delta t)$. For this application, we set the number of time delays to 2 (e.g., $I(t)$, $I(t - \Delta t)$ are used to predict $I(t + \Delta t)$). The function (f) is based on a feedforward ANN structure where the number of hidden neurons is set to 10. The values of the network weights are determined by the "early-stopping" method for ANN training where the data is split into three sets—a training set for computing directional derivatives of the errors in weight space, a testing set for determining

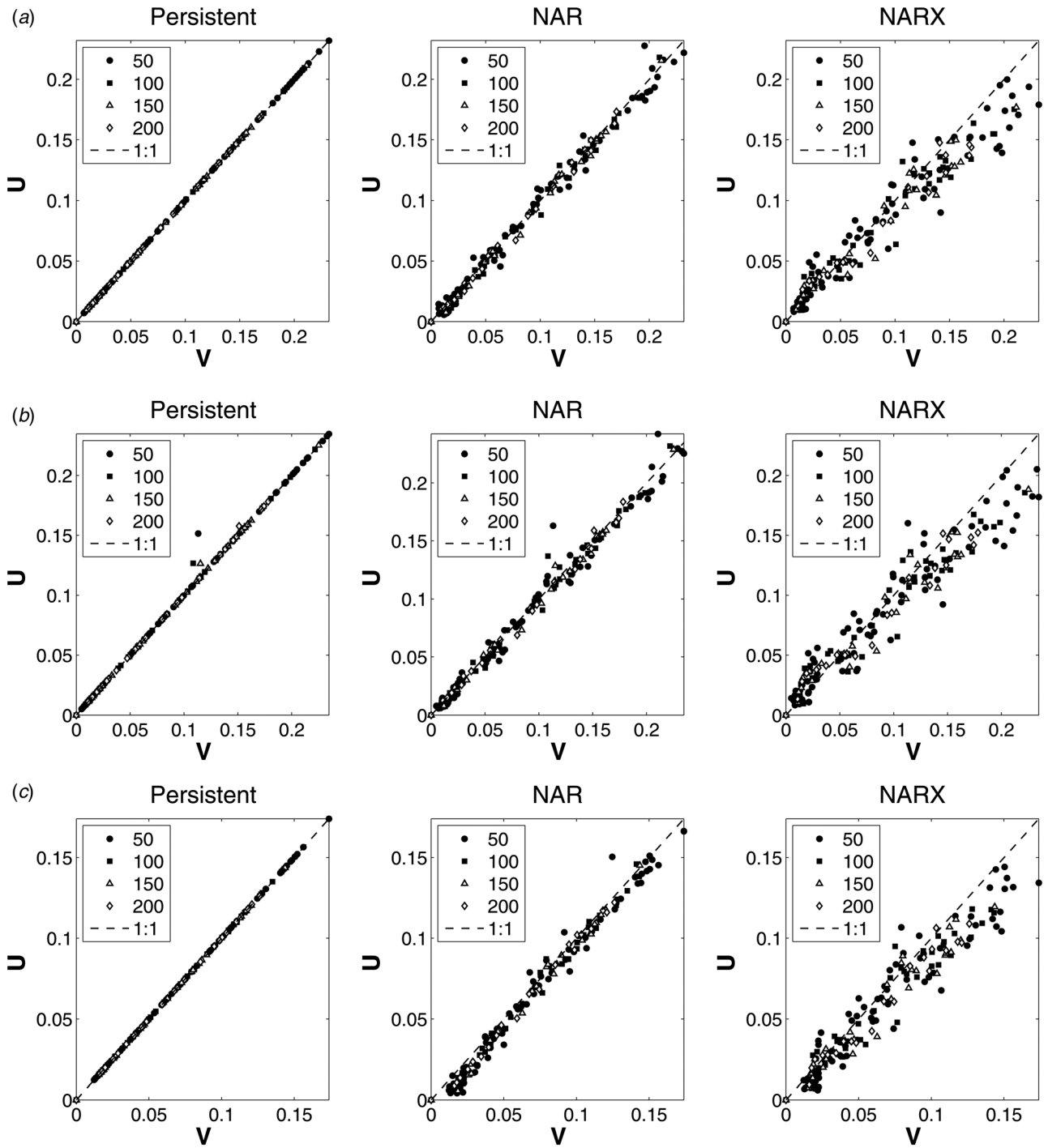

Fig. 6 Scatter plot of *U* and *V* using various clear sky models including a polynomial-based, the ESRA-based, and the clearness index model which uses extraterrestrial irradiance for normalization

when to stop training, and a validation set which is not used at all during the ANN training [14,17]. Data from Oct. 15–31, 2010 is used for validation and the rest of the data from Jan. 1, 2010–Oct. 14, 2010 is split randomly into 80% for the training set and 20% for the testing sets. The ANNs are implemented using the MATLAB Neural Network Toolbox Version 7.0.

The NARX model is similar to the NAR model except that more time-series signals are utilized in the forecast scheme

$$\hat{I}(t+\Delta t) = f(I(t), I(t-\Delta t), \ldots, I(t-n\Delta t)) \\ u_1(t), u_1(t-\Delta t), \ldots, u_2(t), \ldots, u_m(t-n\Delta t)) \quad (15)$$

where m is the number of exogenous inputs. In this case, the u's are 30-min and 6-min backwards moving averages (MA) and standard deviations (SD) of clearness index values which were calculated from 30-s interval data, denoted as k' to distinguish from k which is the clear sky index for hourly averages of I. For example, the 30-min MA and SD is calculated as

$$u_{\mathrm{MA}_{30\mathrm{min}}}(t) = \frac{1}{N} \sum_{\tau=t-30\,\mathrm{min}}^{t} k'(\tau) \quad (16)$$

and

$$u_{\mathrm{SD}_{30\mathrm{min}}}(t) = \sqrt{\frac{1}{N}\sum_{\tau=t-30\mathrm{min}}^{t}\left(k'(\tau)-\overline{k'}(t)\right)^2} \tag{17}$$

where $\bar{k}'(t) = u_{\mathrm{MA30min}}(t)$. The 6-min MA and SD are similarly defined. These inputs are an attempt to use the trends at the last moments of the current hour to forecast the next 1-hour time step. Again, f in Eq. (15) is also a feedforward ANN which contains more input neurons than the NAR model and the number of time delays are set to 2 for each signal. The number of hidden neurons are set to 10, and the early-stopping method is used for adjusting the weights.

6.2 Comparison Between Persistent and Forecasting Models. The forecast quality evaluations were performed for the data set collected from Jan. 1–Oct. 31 of 2010. Figure 6, shows scatter plots of U_j versus V_j computed for each jth time-window partitioning of the data set for $N_w = 50, 100, 150$, and 200. These plots allow us to visualize the forecasting performance over various time window subsets with different variability values. The difference between each of the evaluations is the normalization factor in the calculations of U and V, where we used on separate occasions $I_{\mathrm{clr,poly}}$, $I_{\mathrm{clr,ESRA}}$, and I_0. These plots show that the persistence models (also distinctive in terms of the normalization variable) all result in $s = 0$, since $U_j = V_j$ for any window j. Note that, for the extraterrestrial case, U_j and V_j have a smaller numerical range than the cases for using $I_{\mathrm{clr,poly}}$ and $I_{\mathrm{clr,ESRA}}$. This is because $I_0 > I_{\mathrm{clr,poly}}$ and $I_{\mathrm{clr,ESRA}}$. In all cases, the general trends and conclusions are the same which can be visualized by the scatter points in Fig. 6. The NAR model forecast quality seems to be not much better than persistence, while the NARX model does show some significant forecasting benefit since many of the scatter points fall below the 1:1 line.

The results are quantitatively evaluated by computing an approximation of $\langle s \rangle$. As mentioned earlier, s is a random variable which depends on the ratio U/V. The statistical average of this the slope of the scatter points as shown in Fig. 5, since the scatter plots from each of the models form almost a linear relationship. The slopes calculations are repeated using $N_w = 10, 11, \ldots, 200$ and the results are plotted in Fig. 7, where we can see that, values of $\langle s \rangle$ converge to a certain value as N_w increases. For the persistence models $\langle s \rangle = 0$, for the NAR model $2\% < \langle s \rangle < 5\%$, and for the NARX model $10\% < \langle s \rangle < 15\%$. Again, there is very little difference in the resulting approximation of $\langle s \rangle$ amongst which normalization factor is used in the evaluation procedures.

Numerical values of $\langle s \rangle$ obtained using $N_w = 200$ are given in Table 2, along with the more common forecast quality metrics, the R^2 and the RMSE. Considering first R^2, which range 0.964–0.977 on the validation data set, this forecasting quality measure gives the impression that the forecasts are very accurate even for the persistence models. This performance measure, however, is misleading because, by definition, the persistence model does not capture any of the solar irradiance variability. Similarly, the RMSEs, which range from 48.8–59.4 $\mathrm{W/m^2}$, give misleading conclusions if compared to other RMSEs in the literature (e.g., Refs. [4,6–9,18]) without first knowing about the solar irradiance variability conditions in those studies. By using the R^2 or RMSEs, it is not clear either how much those difference actually translate to differences in forecast quality. In terms of the $\langle s \rangle$ metric, it is clear that each persistence model has no forecasting quality since $\langle s \rangle = 0$ which means that $U = V$ (all the uncertainty is due to the variability).

Comparing the $\langle s \rangle$ values for the NAR and NARX models the $\langle s \rangle$ values are slightly different depending on the normalization factor. When the $I_{\mathrm{clr,poly}}$ and $I_{\mathrm{clr,ESRA}}$ are used, the difference of the computation of $\langle s \rangle$ leads us within $\pm 1\%$, there is some, although small, sensitivity on forecasting quality when using different normalization factors.

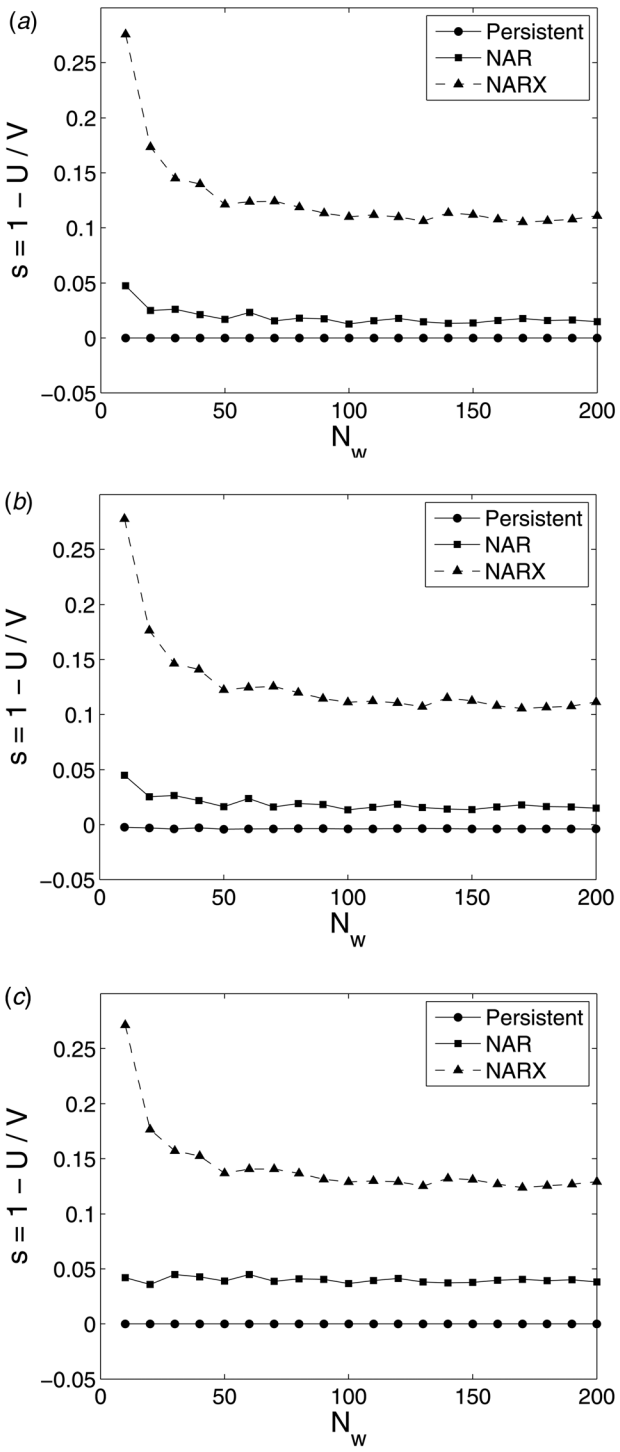

Fig. 7 Evaluation of $\langle s \rangle = 1 - U/V$ versus N_w (time-window sizes) after modifying algorithm with different clear sky and persistence models

The time series of the developed forecasting models, the persistent model and the measured $I(t)$ values are shown in Fig. 8. The lines of the measured values and NARX model predictions are in bold in order to reduce clutter and also to emphasize the performance of the NARX model which seems to perform the best. This figure illustrates that the models perform similarly well over the clear day. In general, each model performs very well on clear days as opposed to the highly variable days where most of the larger errors occur.

Table 2 Forecasting quality metrics for the persistent, NAR, and NARX models on validation and training data sets. For comparisons, the metric $s = 1 - U/V$ is based on three normalization factors.

Training set Model	R^2	RMSE (W/m^2)	s_{poly}	s_{ESRA}	s_0
$I_{p,0}$	0.964	59.4	0	–	–
$I_{p,ESRA}$	0.969	55.5	–	0	–
$I_{p,poly}$	0.973	52.4	–	–	0
NAR	0.972	53.2	1.74%	2.27%	3.71%
NARX	0.977	48.8	11.56%	12.02%	13.1%
Validation set Model	R^2	RMSE (W/m^2)	s_{poly}	s_{ESRA}	s_0
$I_{p,0}$	0.918	62.4	0	–	–
$I_{p,ESRA}$	0.926	59.5	–	0	–
$I_{p,poly}$	0.934	56.5	–	–	0
NAR	0.924	60.2	−2.56%	−2.53%	2.66%
NARX	0.949	49.4	16.12%	16.25%	20.1%

6.3 Comparison With a Previously Validated Forecast Model. In this section, we show how the NAR and NARX models can be compared to a previously validated forecast model by Perez et al. [6]. The Perez et al. model is based on cloud motion forecasts (CMF) and was used to validate solar irradiance forecasts of 1–5 h ahead for several climatically distinct sites for a period spanning from Aug. 23, 2008–Aug. 31, 2009. In the CMF technique, satellite derived images are used to extract pixel values of clearness index ($k(t)$) at time t. The motions of the clouds are then predicted and are used to determine future images from which values of $k(t+1)$ are inferred. From the $k(t+1)$ predictions, solar irradiance forecasts are obtained. The study in Ref. [6] is relevant for comparisons here because the persistent model is defined equivalently to the present work. Specifically, both persistent models make use of the current clearness index value to predict future values of solar irradiance.

The models are compared by observing the improvements over persistency and, furthermore, the improvement over the persistent models are approximate to the proposed metrics, that is, $U/V \approx \mathrm{RMSE}/\mathrm{RMSE}_p$, where RMSE_p is the RMSE of the persistence model. To empirically show this, the RMSEs of the persistent, NAR, and NARX models were calculated using $N_w = 200$ hours then we plot the RMSEs of the NAR and NARX forecasts versus the RMSEs for the persistent model as shown in Fig. 9(*a*). The slopes obtained by the regression fits are equivalent to the slopes describing the average of the ratio U/V. Taking the slope quantities to estimate $\langle s \rangle$ we get values: $\langle s \rangle = 1 - 1.004 = -0.4\%$ and $\langle s \rangle = 1 - 0.863 = 13.7\%$, which are approximately equally to those in Table 2 for the NAR and NARX models, respectively. The approximate equivalence between U/V and RMSE/ RMSEp can also be established from the definitions of U, V, and RMSE and realizing that the normalization factors effectively cancel out when taking the ratios. Estimating U/V with RMSE/ RMSEp is much easier than the procedure used to produce the graphs in Fig. 9, so this approach is recommended. However, we emphasize here the rationale for proposing the metric, which is that the metric gives a rigorous measure of the effective probing into the random variability (Eq. (13)).

In Table 2 of Ref. [6], RMSEs for 1-h ahead forecasts of the CMF model and the persistent model are given. These values are used to produce Fig. 9(*b*), where we calculate a regression line after setting the intercept value at zero. The point inside the red circle was considered an outlier and neglected from the fit (for this datum, the persistent model performed much better than the CMF model in terms of the RMSE). Just as the value of $\langle s \rangle$ for the NAR and NARX models from Fig. 9(*a*), the $\langle s \rangle$ value of the CMF model is estimated to be $1 - 0.923 \approx 8\%$. This value is close to the value obtained by NARX model value reported here, therefore, a NARX-type approach seems to produce comparable forecasting performance to the CMF model approach.

7 Testing With a Numerical Weather Prediction Forecast Study

Here we consider an interesting case to test our proposition that forecasting errors are mainly due to solar variability and to affirm using performance measures that directly compare solar irradiance forecast errors with solar irradiance variability. In Ref. [19], solar irradiance forecasts were developed a spatially distributed array of sensors over Germany. An important result highlighted in this study, is the reduction of overall forecasting errors, as evaluated using the RMSE, when averaged over all the sensors. Here we explain their observations of the error reductions using the framework recently proposed by Ref. [15] for studying variability for an ensemble of PV systems. According to Refs. [15,10], the variability for a fleet of PV systems can be determined from knowledge of the individual standard deviations $\sigma^i_{\Delta t}$ of the fluctuations and the correlation coefficients $\rho^{i,j}_{\Delta t}$ of the fluctuations for each pair of stations from

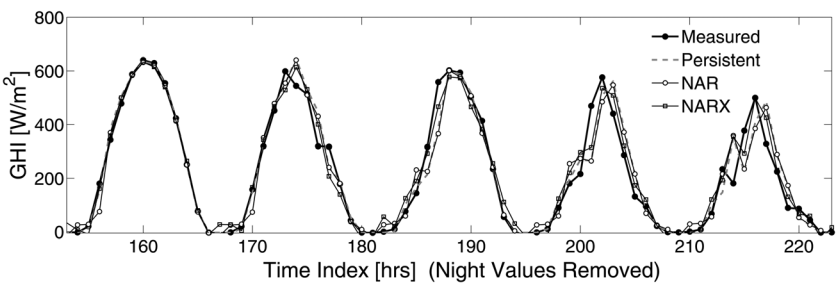

Fig. 8 Hourly forecasting comparisons for five consecutive days (Oct. 27–31, 2010) in the validation data set with night values removed

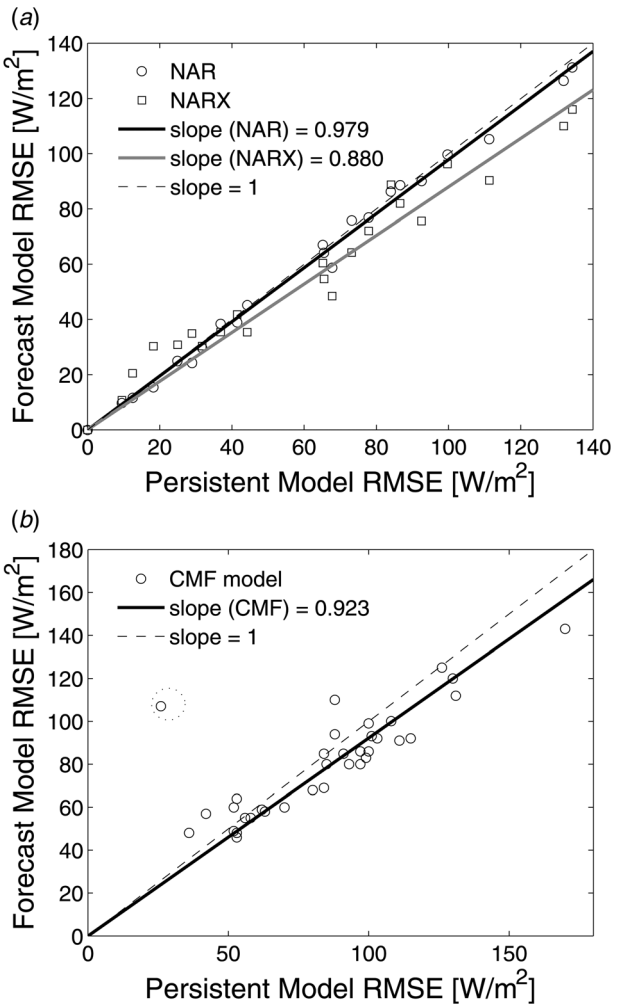

Fig. 9 Root mean square errors (RMSEs) for different forecast models versus RMSE of persistent model: (*a*) NAR and NARX model and (*b*) CMF model [6]. The outlier point within the dashed circle was ignored for calculating the regression line in (*b*).

$$\sigma_{\Delta t}^{\text{Fleet}} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{\Delta t}^{i}\sigma_{\Delta t}^{j}\rho_{\Delta t}^{i,j}} \tag{18}$$

Based on the results obtained in Secs. 4, 5, and 6 of this paper, we anticipate that $U \propto V$, thus we make the assumption that for the forecasts in Ref. [19], $\Delta k \propto \varepsilon$, where ε is a forecast error. Using this assumption, we proceed by adapting Hoff and Perez's variability model to the ensemble average solar irradiance forecasting error and also assume that, for all i, $\sigma_{\Delta t}^{i} = \sigma_{\Delta t}^{1}$ where now we let $\sigma_{\Delta t}^{i}$ and $\rho_{\Delta t}^{i,N}$ represent the forecast RMSE and the cross correlation of the forecast errors, respectively. After expanding the first summation we get

$$\frac{\sigma_{\Delta t}^{\text{Fleet}}}{\sigma_{\Delta t}^{1}} = \frac{1}{N}\sqrt{\left(\sum_{i=1}^{N}\rho_{\Delta t}^{i,1} + \sum_{i=1}^{N}\rho_{\Delta t}^{i,2} + \ldots + \sum_{i=1}^{N}\rho_{\Delta t}^{i,N}\right)} \tag{19}$$

Each summation term can be represented as an average N times the average ($N\langle\ldots\rangle = \sum\ldots$) and, after assuming $\langle\rho_{\Delta t}^{i,1}\rangle = \langle\rho_{\Delta t}^{i,2}\rangle = \ldots = \langle\rho_{\Delta t}^{i,N}\rangle = \langle\rho\rangle$, we reexpress Eq. (19) as

$$f_R = \frac{\sigma_{\Delta t}^{\text{Fleet}}}{\sigma_{\Delta t}^{1}} = \sqrt{\langle\rho\rangle}. \tag{20}$$

This equation, which only depends on the cross correlations ($\langle\rho\rangle$), gives us a very simple expression for estimating the reduced combined uncertainty for a network of spatially distributed solar irradiance measurements sites. As it turns out, $\langle\rho\rangle$ (the cross correlations of Δk or ε) can very accurately be modeled as a function of distance x between stations. In fact, when plotting $\langle\rho\rangle$ of Δk (as in Ref. [15]) or ε (as in Ref. [19]) versus x the plots appear to follow a similar decreasing trend. In Ref. [19], $\rho(x)$ was modeled with an exponential parametric function, $\rho = \exp((a_1 x)^{a_2})$, however, the parameters were not given. The fitting function was re-estimated here by selecting points on the model fit of their graph, but this time using the parametric exponential function $\exp(a_3x^3 + a_2x^2 + a_1x + a_0)$. The results are shown in Fig. 10 where the selected points are indicated by the square markers and the dashed line is the fit. Using the fit, we approximate the reduced error factor f_R given by Eq. (20) as

$$f_R(x) = \sqrt{1/x\int_0^x \rho(x')dx'} \tag{21}$$

Applying this equation gives the solid line in Fig. 10 which matches closely the observed values obtained from Ref. [19].

It is also interesting to point a possible lower bound for f_R which occurs when all the forecast errors are uncorrelated. In this situation, $f_R = 1/\sqrt{N}$ as can be verified from Eq. (20). This is the same result for reduced solar irradiance variability predicted in Refs. [11,10] and observed approximately from experiments in Refs. [20,21].

8 Applications of the Forecast Metric

As compared to a single site, the solar variability can be significantly reduced by distributing solar power generation sites over geographically diverse locations [11,16]. Building distributed solar power sites can take some time however due to a number of reasons (building permits, capital costs, etc.). One immediate way to avert negative impacts of variability of the solar resource is through reliable forecasting. We demonstrated that a well trained forecasting model has the effect of reducing stochastic variability which is the real concern for solar power. In one sense, the purpose of forecasting is to reduce the amount of stochastic variability because this type of variability is difficult to manage. By applying the proposed metric, we can provide an estimate on the amount of reduced stochastic solar variability and therefore provide some reassurance that some of the solar variability could be handled. If the solar variability is reduced

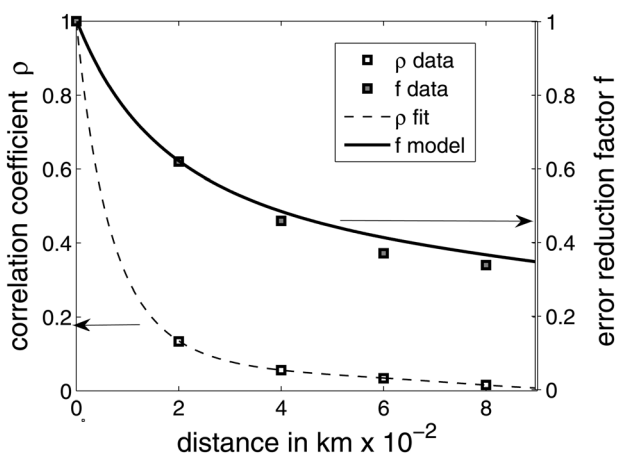

Fig. 10 Empirical data compared with modeling predictions of uncertainty (forecast errors) reduction

by geographically distributed generation, then the gains are compounded by improved forecasting, as shown in Sec. 7. In Ref. [19], the "relative" RMSE (rRMSE) for one location are found to be on the order 36%, where as for an ensemble of sites the rRMSE is dramatically reduced to 13% for one-day ahead forecasts.

As recent renewable energy integration studies have reported [1–3], the variability and intermittency of renewables will likely have significant impacts on reserve and storage requirements. Assuming that the amount of reserve and storage capacity needed to back-up solar energy is proportional to the solar variability, then we can potentially reduce these requirements with reliable forecasting. The forecast metric allows one to estimate by how much a forecasting model will save in storage and reserve requirements. Because the variability of solar irradiance is an important quantity to characterize (e.g., probability of large ramp rates, frequency of fluctuations, etc.) for the purposes of estimating future impacts of including solar energy, it is beneficial to examine historical solar irradiance data sets in order to characterize solar irradiance variability on a temporal and spatial basis. Based on the observations provided here, it can be reasonably expected that a forecasting model, if it is trained well enough over data sets that thoroughly represent the climatology for a given location, should retain their performance in terms of the forecast metric when applied to locations with similar climates. This kind of analysis would be valuable information for simulating impacts due to the uncertainty and variability of solar availability over certain geographical regions of interest.

9 Conclusions

A new forecasting metric is proposed and compared to conventional forecasting performance metrics such as the RMSE and the coefficient of determination (R^2). We showed that the proposed metric is a good indicator of the extent to which a forecast model effectively predicts the stochastic variability of solar irradiance. The purpose of forecasting models is essentially to prescribe the nontrivial component of the solar irradiance, and hence improve our understanding of the solar resource variability. The concept is illustrated by comparing three forecast models. A persistent clearness index model is used as baseline, so that improvement over the baseline model would only occur when the forecast models demonstrate predictability for the stochastic variability of solar irradiance. As expected, persistent models are characterized by low forecast quality when the solar irradiance exhibits cloud-induced fluctuations, but perform very well for clear sky periods. Two other models based on artificial neural networks were also applied, one of which included short-range moving averages as inputs which did produce effective forecasts by reducing the stochastic variability by 12–20%. In terms of the conventional statistical metrics, each model performs well with R^2 above 90% and RMSEs below 65 W/m^2 over the entire data sets. The advantage of the the proposed metric over existing metrics is that the forecast quality is robust since the ratio U/V is preserved over all time window subsets.

Acknowledgment

CFMC and RM gratefully acknowledge the financial support given by the California Energy Commission (CEC) under the PIER RESCO Project PIR-07-036, by the National Science Foundation (NSF) CNS division Grant No. 0923586, and also by the Eugene Cotta-Robles (ECR) Fellowship program of the University of California and the Southern California Edison Fellowship program. Seed and continued support by the Center for Information Technology Research in the Interest of Society (CITRIS) is gratefully appreciated.

References

[1] Lew, D., and Piwko, R., 2010, "Western Wind and Solar Integration Study," National Renewable Energy Laboratories, Technical Report No. NREL/SR-550-47781.

[2] California Independent System Operator (CAISO), 2010, "Integration of Renewable Resources: Operational Requirements and Generation Fleet Capability at 20 Percent RPS," available online at http://www.caiso.com/2804/2804d036401f0.pdf.

[3] Rodriguez, G. D., 2010, "A Utility Perspective of the Role of Energy Storage in the Smart Grid," Power and Energy Society General Meeting, IEEE, Minneapolis, MN, July 25–29, pp. 1–2.

[4] Marquez, R., and Coimbra, C. F. M., 2011, "Forecasting of Global and Direct Solar Irradiance Using Stochastic Learning Methods, Ground Experiments and the NWS Database," Sol. Energy, **85**(5), pp. 746–756.

[5] Pedro, H. T. C., and Coimbra, C. F. M., 2012, "Assessment of Forecasting Techniques for Solar Power Output With No Exogenous Inputs," Sol. Energy, **86**, pp. 2017–2028.

[6] Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renne, D., and Hoff, T. E., 2010, "Validation of Short and Medium Term Operational Solar Radiation Forecasts in the US," Sol. Energy, **84**(5), pp. 2161–2172.

[7] Cao, J., and Lin, X., 2008, "Study of Hourly and Daily Solar Irradiation Forecast Using Diagonal Recurrent Wavelet Neural Networks," Energy Convers. Manag., **49**(6), pp. 1396–1406.

[8] Mellit, A., 2008, "Artificial Intelligence Technique for Modelling and Forecasting of Solar Radiation Data: A Review," Int. J. Artif. Intell. Soft Comput., **1**, pp. 52–76.

[9] Martin, L., Zarzalejo, L. F., Polo, J., Navarro, A., Marchante, R., and Cony, M., 2010, "Prediction of Global Solar Irradiance Based on Time Series Analysis: Application to Solar Thermal Power Plants Energy Production Planning," Sol. Energy, **84**(10), pp. 1772–1781.

[10] Mills, A., and Wiser, R., 2010, "Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power," Lawrence Berkeley National Laboratory, Technical Report No. LBNL-3884E.

[11] Hoff, T. E., and Perez, R., 2010, "Quantifying PV Power Output Variability," Sol. Energy, **84**(10), pp. 1782–1793.

[12] Ineichen, P., 2006, "Comparison of Eight Clear Sky Broadband Models Against 16 Independent Data Banks," Sol. Energy, **80**, pp. 468–478.

[13] Zarzalejo, L. F., Polo, J., and Ramirez, L., 2004, "Gc_model5_irradiance," (Matlab computer program) CD-ROM accompanying Ref. [14].

[14] Badescu, V., 2008, *Modeling Solar Radiation at the Earth Surface*, Springer-Verlag, Berlin/Heidelberg.

[15] Hoff, T. E., and Perez, R., 2011, "Modeling PV Fleet Output Variability," Sol. Energy, **86**(8), pp. 2177–2189.

[16] Lave, M., and Kleissl, J., 2010, "Solar Variability of Four Sites Across the State of Colorado," Renewable Energy, **35**(12), pp. 2867–2873.

[17] Bishop, C., 1995, *Neural Networks for Pattern Recognition*, Oxford University, Great Clarendon Street, Oxford, UK.

[18] Marquez, R., Gueorguiev, V. G., and Coimbra, C. F. M., 2012, "Forecasting Solar Irradiance Using Sky Cover Indices," ASME J. Sol. Energy Eng. (in press).

[19] Lorenz, E., Hurka, J., Heinemann, D., and Beyer, H. G., 2009, "Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., **2**(1), pp. 2–10.

[20] Lave, M., Kleissl, J., and Arias-Castro, E., 2012, "High-Frequency Irradiance Fluctuations and Geographic Smoothing," Sol. Energy, **86**(8), pp. 2190–2199.

[21] Marcos, J., Marroyo, L., Lorenzo, E., Alvira, D., and Izco, E., 2011, "Power Output Fluctuations in Large Scale PV Plants: One Year Observations With One Second Resolution and a Derived Analytic Model," Prog. Photovoltaics, **19**(2), pp. 218–227.