Renewable Energy 87 (2016) 15–30

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

A semiparametric spatio-temporal model for solar irradiance data

Joshua D. Patrick [a, *], Jane L. Harvill [b], Clifford W. Hansen [c]

[a] Department of Statistics, One Shields Avenue, University of California, Davis, CA 95616-5270, USA
[b] Department of Statistical Science, P.O. Box 97140, Baylor University, Waco, TX 76798-7140, USA
[c] P.O. Box 5800, Sandia National Laboratories, Albuquerque, NM 87185-1033, USA

ARTICLE INFO

Article history:
Received 5 May 2015
Received in revised form
26 August 2015
Accepted 1 October 2015
Available online xxx

Keywords:
Irradiance
Spatio-temporal model
Nonseparability
Lattice data
Semiparametric time series

ABSTRACT

We evaluate semiparametric spatio-temporal models for global horizontal irradiance at high spatial and temporal resolution. These models represent the spatial domain as a lattice and are capable of predicting irradiance at lattice points, given data measured at other lattice points. Using data from a 1.2 MW PV plant located in Lanai, Hawaii, we show that a semiparametric model can be more accurate than simple interpolation between sensor locations. We investigate spatio-temporal models with separable and nonseparable covariance structures and find no evidence to support assuming a separable covariance structure. Our results indicate a promising approach for modeling irradiance at high spatial resolution consistent with available ground-based measurements. Such modeling may find application in design, valuation, and operation of fleets of utility-scale photovoltaic power systems.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Accurate modeling of power output from photovoltaic (PV) power plants requires accurate estimates of the plane-of-array (POA) irradiance over each plant's footprint. For single, spatially contiguous plants, power correlates reasonably well with the aggregate irradiance over the plant [1]. Currently plant aggregate irradiance is estimated by either 1) averaging measurements from many irradiance sensors [2] or 2) transforming data from a single measurement point [3]. Various transformations of single measurements to plant-aggregate irradiance are proposed [4–6] all of which attempt to produce a time series of values representative of a spatial average over an area of interest by smoothing, in some fashion, a time series of measurements at a single point. However, these smoothing transformations are themselves calibrated by a set of spatially diverse measurements and are not inherently capable of estimating the spatial variability of irradiance within the area of interest at each point in time. Areal average irradiance can also be estimated from satellite data (e.g., [7]) albeit with somewhat coarse spatial and temporal resolution on the order of 1 km^2 and 30 min.

Past research efforts [summarized in Ref. [2] have focused on accurate representation of the temporal variability in spatially-averaged irradiance. Increasing concerns about the effects of large numbers of PV systems connected to an electrical grid have led to significant interest in better understanding of the spatial variability of irradiance as well as the temporal variability of the spatial average. For example, spatial variability of irradiance, and hence PV power, over a utility's service area, or over the service area of a distribution feeder circuit, causes variable levels of power to be injected at different locations which may require mitigation efforts to maintain circuit voltages within acceptable limits. Contracts for purchase of power from large PV plants often require after-the-fact estimation of the potential energy production (in addition to measurement of actual production) over short time periods to accurately compensate for utility-directed limits on power and for equipment outages; these estimates would benefit from irradiance data with high spatial and temporal resolution. To better understand spatial variability and to more accurately determine potential production from large plants, utilities have deployed networks of irradiance sensors across service territories [e.g. Ref. [8] or are requiring deployment of many sensors at large plants [e.g. Ref. [9]. Suppliers of satellite-based irradiance estimates are upscaling in time by simulating cloud movements [10] or running numerical weather simulations [11], although spatial coarseness remains. Statistical simulations have been employed to create representative irradiance data at high time resolution over large areas [12] but with coarse spatial resolution. Here, we explore statistical modeling

* Corresponding author.
E-mail addresses: jdpatrick@ucdavis.edu (J.D. Patrick), jane_harvill@baylor.edu (J.L. Harvill), cwhanse@sandia.gov (C.W. Hansen).

http://dx.doi.org/10.1016/j.renene.2015.10.001
0960-1481/© 2015 Elsevier Ltd. All rights reserved.

of global horizontal irradiance (GHI) at high resolution spatial and temporal scales, i.e., on the order of 100 m^2 and 1 min.

A challenge in modeling irradiance data is incorporating the interaction between time and space. Especially in the presence of advecting clouds, the irradiance observed at one location is likely to also be observed at other locations but with a time shift. Thus we anticipate that irradiance will exhibit a spatial autocorrelation that varies with time. Spatio-temporal models explicitly account for this autocorrelation and thus may predict aggregate irradiance more accurately than does a simple spatial average.

In spatio-temporal modeling, the time and spatial covariances structures are classified as either separable or nonseparable. Assuming separability allows for simpler methods for estimating the model but this assumption is not realistic for the complex interactions between the space and time components of irradiance data. We propose a spatio-temporal model that does not assume a separable covariance structure to the data. Our model incorporates a semiparametric method for modeling the time series component and a lattice structure for the spatial component.

The remainder of this paper is organized as follows. In Section 2, we present a survey of the literature on time series and spatio-temporal approaches for irradiance data. In Section 3, we discuss how the time series structure is modeled via a semiparametric model fitted with a method known as spline-backfitted kernel (SBK) estimation. In Section 3.3, we introduce the spatio-temporal model, and compare the model's performance assuming either a separable or a nonseparable covariance structure to evaluate whether separability can be assumed. In Section 4, we apply the model to irradiance data from the La Ola photovoltaic plant in Lanai, HI. Finally, we provide discussion and conclusions in Section 5.

2. Literature survey

Literature reports several efforts at modeling time-series of irradiance at individual locations, and substantially fewer attempts to construct spatio-temporal models of irradiance considering several proximal locations. The literature on individual time-series modeling includes approaches based on autoregressive integrated moving average (ARIMA) analysis [e.g. [13], non-linear autoregressive analysis [14], regression analysis [15], artificial neural networks analysis [16], *k*-Nearest Neighbors algorithm [16] and Bayesian inference [16]. These approaches focus on forecasting irradiance considering only the measurements at a selected location separately from measurements at other locations. Paoli et al. [16] considers a type of artificial neural network known as Multi-Layer Perceptron (MLP) network and finds their method performs as well or better than other methods such as ARIMA analysis, Bayesian inference, and *k*-Nearest Neighbors. Yang et al. [13] introduce an ARIMA model that incorporates low-resolution, ground-based cloud cover data to obtain next hour solar irradiance. The authors state that their ARIMA model outperforms all other time series forecasting methods in four of the six stations they tested. In both the MLP and ARIMA methods, the model does not incorporate a spatial component but only models irradiance in time.

The literature on spatio-temporal modeling of solar irradiance and/or power from PV systems comprise three general types: 1) lattice and geo-statistical models, which may predict irradiance at individual spatial locations; 2) variability models which estimate statistics for the time series of power from a large PV system or from a fleet of PV system; and 3) models to estimate the aggregate output of a collection of PV systems from measurements at a single location. We discuss each category in turn.

Among lattice models [17], model irradiance data from ten sensors roughly at 5 km spacing using a spatio-temporal autoregressive moving average (STARMA) model. The STARMA model incorporates the Euclidean distances between two points in order to model the spatial structure of the data. However, the STARMA model used in Ref. [17] assumes a separable covariance structure, an assumption which we find to be questionable at the scale of a single PV plant. Glasbey et al. [18] model irradiance with a nonseparable spatio-temporal covariance structure for data taken at 22 sites in the city of Edinburgh, Scotland, with distances between sites ranging from 0.1 to 20 Km. They found the nonseparable model performed marginally better with a mean square difference between estimated and fitted covariances of 0.0032 as compared to 0.0033 for the separable model. Spatio-temporal covariance structures were explored in Refs. [19,20]. In both papers, a geographical plane transformation by multidimensional scaling is used to make the process isotropic. After the transformation, a separable or nonseparable covariance model is used for time-forward kriging in Ref. [19] and an empirical covariance structure is used to describe the correlation between sites based on distance in Ref. [20]. The methods used are then applied to data taken over Singapore (an area of about 700 km^2).

Literature reports many studies intended to quantify and/or model the variability in the output of a large PV plant or the aggregate output of a fleet of PV plants. Typically, variability is quantified by computing statistics for the time series of changes in irradiance over a fixed time interval. Generally, variability models either estimate the variance of the time series of aggregate output from a fleet of PV systems, or the correlation between time series of changes in irradiance, or PV output power as a function of distance between the sites. Both types of results are of great importance for managing and balancing energy on an electrical grid with connected PV systems. However, most of these variability models do not allow for prediction of the time series of irradiance at a location of interest from measurements at nearby locations.

A model for maximum output variability was presented in Ref. [10] with changes assumed to be uncorrelated. Hoff and Perez [21] extended this model by incorporating correlation coefficients between the variability of individual plants of a PV fleet based on site distance and implied cloud speed. Perez et al. [22] used time-shifted data from a single location to examine the potential variability of clearness index by examining site pair correlation for distances ranging from 100 m to 100 km and for time intervals of 1 min, 5 min, and 15 min. Smaller site pair distances are examined in Ref. [23] in which coherence spectra and wavelet analysis are used for quantifying geographic smoothing. Using the coherence spectra approach they found that strong correlations exist for station pairs within 3 km at time intervals as short as 10 min. For longer time intervals, they found the correlation decreases with distance, although not uniformly in all directions. After applying the wavelet transform to one station and comparing to the average of six stations, the authors found that at time intervals shorter than about 5 min, the six stations are independent, resulting in reduced variability. For longer time intervals the reduction in variability lessens with almost no reduction at about 68 min.

Moving beyond isotropic models for correlation, i.e., based only on time interval and distance between locations [24], presented an anisotropic model which describes correlation for power production and for changes in power production as a function of spatial location in an X,Y plane and time interval. They consider forecasting in time for a fleet of PV systems spread over a 50 km $\times$ 50 km area. A mathematical model for the spatio-temporal correlations between changes in irradiance is presented in Ref. [25]. This model is based on a spatial Poisson process that is used to simulate the probability that advecting clouds cover two sites. This correlation model is inherently anisotropic since it takes into account the cross-wind

distance, the along-wind distance, and the length of the motion path of the clouds.

Finally, aggregate power from a large system or fleet of systems over a defined area is estimated from measurements at a single point by transforming the time series of measurements [6]: describes a moving average technique [5]; proposes a low-pass filter; and [4] uses wavelet decomposition. The wavelet method uses a variability model for correlation among locations as described above to moderate each wavelet model's amplitude to estimate the time series of aggregate power. The authors validate their simulation method at two plants, one of which is a 1 km × 1 km distributed generation plant and the other is a 3 km × 2 km utility scale PV plant. The wavelet method permits estimation of the aggregate output but not of the contribution to the aggregate from individual systems.

The spatio-temporal model that we propose improves upon the works of [13] and [14] because we do not assume a parametric form for the time component of the model, and improves on [17] through the nonseparable covariance structure. While the works of [23] and [4] examined correlation as a function of distance between sites, we incorporate a lattice structure and utilize the simultaneous autoregressive spatial (SAR) model in which the sites are connected through a "neighborhood." Our model explains the correlation through this neighborhood structure and does not take as inputs the wind distance and cloud motion paths as in Ref. [25]. The use of the SAR model is based on viewing the sites within a utility scale PV plant as fixed locations instead of geostatistical points such as in Refs. [19,20].

3. Modeling irradiance

Let $Q_{s,t}$ represent an observable process, e.g., measured GHI, at time t and location s for $t = 1,2,\ldots,T$ and $s = 1,2,\ldots,S$. If there is no interaction in time and space, the covariance function of $Q_{s,t}$ can be written as a product of two functions where one function is dependent on time only, and the other on location alone. Such a covariance function is called "separable." However, when interactions in space and time are present, the covariance function is "nonseparable; " i.e., it cannot be factored into two separate functions. Spatio-temporal models with separable covariance are much easier to implement. But in the presence of space-time interaction, separable models do not perform well, and can lead to misleading or incorrect conclusions.

For modeling $Q_{s,t}$, consider

$$Q_{s,t} = R_{s,t} + Z_{s,t}, \quad t = 1,\ldots,T, \quad s = 1,\ldots,S, \tag{1}$$

where, at time t and location s,$R_{s,t}$ represents the true irradiance signal and $Z_{s,t}$ is a noise process. Furthermore, decompose the noise process into a sum of three terms,

$$Z_{s,t} = X_{s,t} + Y_{s,t} + \varepsilon_{s,t}, \tag{2}$$

where $X_{s,t}$ is a time series process at location s, $Y_{s,t}$ is a spatial process at time t, and $\varepsilon_{s,t}$ is a multivariate error process with mean zero and $TS \times TS$ covariance matrix $\mathbf{\Sigma}(s,t)$. If the process is separable, then the covariance matrix can be written as $\mathbf{\Sigma}(s,t) = \mathbf{\Lambda}(t) \otimes \mathbf{\Gamma}(s)$, where $\mathbf{\Lambda}(t)$ is a $T \times T$ temporal covariance matrix, $\mathbf{\Gamma}(s)$ is an $S \times S$ spatial covariance matrix, and $\otimes$ is the Kronecker product [26].

There are a variety of methods for fitting separable spatio-temporal models to space-time data. A review of space-time analysis methods and their computational counterparts can be found in Ref. [27] or [28]. We consider three approaches to fitting model (2). The first approach fits a spatial model at each time. Then spatial residuals are computed, and for each location a time series model is fitted to the spatial residuals at each location. The second approach models the time series at each location, computes time residuals, and then fits a spatial model at each time to the residuals. Both of these approaches carry the assumption that the covariance structure is separable in space and time. The third approach removes the separability assumption, jointly modeling time and space using the spatio-temporal model introduced in Section 3.4.

3.1. Modeling the time series component

For modeling time series data arising from a dynamic process, such as solar irradiance, nonlinear models often out-perform linear models [29]. Although the class of nonlinear time series models is infinitely large, there are many popular parametric nonlinear models including the bilinear model [30], the exponential autoregressive model [31], and a variety of threshold autoregressive models [32,33]. When one of these parametric models is known to be appropriate for analyzing the time series, it should be used for analyzing the series. However in the analysis of solar irradiance, no specific class of parametric nonlinear model has been shown to be generally applicable, and therefore we pursue a semiparametric approach. In this section, we examine only the time component of the model. So to ease notation, for the remainder of this section, we consider the location fixed, and suppress the s subscript; that is $X_{s,t} = X_t$ for a fixed value of s.

A highly versatile semiparametric model is the functional coefficient autoregressive model of order p (FCAR(p)), first introduced by Ref. [34]. The FCAR(p) model has an additive autoregressive structure, but with coefficients that vary as a function of some variable, u say, which can be exogenous to the series X_t. In the pure time series context, u is a lagged value of the series, and we write $u_t = X_{t-d}$. In this paper, we restrict the FCAR(p) models to those with $u_t = X_{t-d}$, and so define the FCAR(p) model as

$$X_t = m_0(u_t) + \sum_{j=1}^{p} m_j(u_t)X_{t-j} + \omega_t, \quad t = p+1,\ldots,T \tag{3}$$

where $u_t = X_{t-d}$, $d \le p$, $m_j(\cdot)$, $j = 0,1,2,\ldots,p$ are measurable functions of u, and $\{\omega_t\}$ is a sequence of independent and identically distributed (IID) random variables with mean zero and constant variance.

Reasonable use of the FCAR(p) model requires only that the model is additive, and places few restrictions on the functional coefficients. To illustrate the versatility of the FCAR(p) model, note that if $m_0(u_t) = 0$, and $m_j(u_t) = \alpha_j$, $j = 1,2,\ldots,p$ are constants, then the FCAR(p) reduces to a linear autoregressive model of order p, $X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \omega_t$. Another example is, for each $j = 1,\ldots,p$, the coefficients are of the form $m_j(X_{t-d}) = \alpha_j + \beta_j \exp\{-\delta X_{t-d}^2\}$. Then the FCAR($p$) model reduces to the exponential autoregressive model of [31]. Moreover, the FCAR(p) formulation allows for a mixture of models; for example, $m_1(X_{t-d}) = \alpha_1$ and $m_2(X_{t-d}) = \alpha_2 + \beta_2 \exp\{-\delta X_{t-d}^2\}$. For applications, the versatility of the FCAR(p) model is shown in Refs. [34–36]. In Ref. [34], an FCAR(2) model is fit to the monthly records of cases of chickenpox in New York City and an FCAR(8) model is applied to Wolf's annual sunspot numbers data set. An ecological application can be found in Ref. [35] in which an FCAR(2) model is fit to the annual Canadian lynx fur auction sales. The estimated coefficient functions in this application have a nice interpretation of predator–prey interaction. An economic application can be found in Ref. [36] in which a bivariate FCAR(p) model is fit to U.S. GNP and unemployment rate.

Fan and Yao [37] contains a review of methods for fitting the FCAR(p) model, and related inferential procedures. In the following

section, we propose a more recent, improved method for fitting the FCAR(p) model.

3.2. Spline-backfitting kernel estimation

With no presupposed form for the functional coefficients, we propose a semiparametric method for finding pointwise estimates of the functions $m_j(u)$, $j = 0,1,2,\ldots,p$. A number of methods are proposed in the statistics literature. Chen and Liu [38] and [35] propose a kernel regression approach to fitting the model. Harvill and Ray [39] extend the procedure to the case when the series is a vector process. More recently, spline-backfitted kernel (SBK) estimation has been proposed as a means for fitting semiparametric models like the FCAR(p) model. SBK estimation is an adaptation of the backfitting algorithm of [40]. By using backfitting, we take advantage of the additive structure of the model to reduce the dimensionality. Instead of estimating the p coefficient functions simultaneously, we pre-estimate the functions, backfit, and then estimate the functions separately.

The SBK method uses an under-smoothed centered standard spline procedure to pre-estimate the $m_j(u), j = 0,1,2,\ldots,p$. These pre-estimates, also called "oracle" estimates, are used to find pseudo-responses. Then the pseudo-responses are used to estimate the $m_j(u)$ through a kernel estimator; e.g., the Nadaraya–Watson estimator. The SBK method was first proposed by Ref. [41] for estimating nonlinear additive autoregressive models. Wang and Yang [42] adapt the SBK method for IID data [43], adapt it to generalized additive models, and [44] to partially linear additive models. Liu and Yang [45] propose the SBK method for additive coefficient models.

The ability to estimate $m_j(u)$, $j = 0,1,2,\ldots,p$ relies on the good approximation properties of spline estimators. For any $j = 0,1,2,\ldots,p$, assume $m_j(\cdot)$ is sufficiently smooth. Without loss of generality, u can be defined on the compact interval [0,1]. Define the integer $N \approx T^{2/5}logT$, and let $H = (N+1)^{-1}$ [see Ref. [41], Appendix]. Let $0 = \xi_0 < \xi_1 < \cdots < \xi_N < \xi_{N+1} = 1$ denote a sequence of equally spaced knots. There is a set of basis functions $b_0(u), b_1(u),\ldots b_{N+1}(u)$ and a set of constants $\tilde{\lambda}_{0,j}, \tilde{\lambda}_{1,j}, \ldots, \tilde{\lambda}_{N+1,j}$ such that the spline estimator of the j-th coefficient is

$$m_j(u) \approx \tilde{m}_j(u) = \sum_{k=0}^{N+1} \tilde{\lambda}_{k,j} b_k(u). \tag{4}$$

For the basis functions, we choose the linear B-spline basis, defined by

$$b_k(u) = \left(1 - \frac{|u - \xi_k|}{H}\right)_+ = \begin{cases} (N+1)u - k + 1, & \xi_{k-1} \le u < \xi_k, \\ k + 1 - (N+1)u, & \xi_k < u \le \xi_{k+1}, \\ 0, & \text{otherwise}. \end{cases}$$

The coefficients $\tilde{\lambda}_{0,j}, \tilde{\lambda}_{1,j}, \ldots, \tilde{\lambda}_{N+1,j}$ are estimated via least squares; that is, the $\tilde{\lambda}_{k,j}, k = 0, 1, \ldots, N+1, j = 0, 1, 2, \ldots, p$ are the values of $\lambda_{k,j}$ that minimize the sum of squares

$$\sum_{t=p+1}^{T} \left[X_t - \sum_{k=0}^{N+1} \lambda_{k,0} b_k(u) - \sum_{j=1}^{p} \left\{ \sum_{k=0}^{N+1} \lambda_{k,j} b_k(u) \right\} X_{t-j} \right]^2. \tag{5}$$

The spline-estimated functional coefficients are then used to compute "pseudo-responses." Specifically, for each j'=0,1,2,…,p, $j' \neq j$, the pseudo-responses are defined by

$$\tilde{W}_{t,j'} = X_t - \sum_{j=1, j \neq j'}^{p} \tilde{m}_j(u) X_{t-j}, \quad t = p+1, p+2, \ldots, T.$$

For each j'=0,1,2,…,p, let $\tilde{\mathbf{W}}_{j'} = (\tilde{W}_{p+1,j'}, \cdots, \tilde{W}_{T,j'})'$ represent the vector of pseudo-responses, and define the matrix

$$\mathbf{M} = \text{diag}\left\{ K_h\left(X_{p+1-d} - u\right), \ldots, K_h(X_{T-d} - u) \right\},$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$, $K(\cdot)$ is a kernel function, and $h > 0$ is a bandwidth. Then the SBK estimator of $m_{j'}(u)$ is

$$\widehat{m}_{j'}(u) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \left(\frac{1}{T} \mathbf{C}'\mathbf{M}\mathbf{C} \right)^{-1} \frac{1}{T} \mathbf{C}'\mathbf{M}\tilde{\mathbf{W}}_{j'}, \tag{6}$$

where

$$\mathbf{C}' = \begin{bmatrix} X_{p+1} & X_{p+2} & \cdots & X_T \\ X_{p+1}\left(X_{p+1-d} - u\right) & X_{p+2}\left(X_{p+2-d} - u\right) & \cdots & X_T(X_{T-d} - u) \end{bmatrix}$$

The idea behind SBK estimation is to under-smooth in the pre-estimates in order to reduce the bias. This under-smoothing leads to a larger variance which is reduced in the kernel estimation step. The use of splines for the pre-estimates is computationally fast while using kernel smoothing provides convenient asymptotic results [45].

To illustrate, consider a series of $T = 500$ observations from the exponential autoregressive model of order $p = 2$ (EXPAR(2)) given by

$$X_t = \left\{0.5 - 1.1e^{-50X_{t-1}^2}\right\} X_{t-1} + \left\{0.3 - 0.5e^{-50X_{t-1}^2}\right\} X_{t-2} + 0.2\omega_t, \tag{7}$$

where the ω_t are standard normal errors. A time plot of a mean-centered realization of length 500 of such a series is given in Fig. 1. Since X_{t-1} is the functional variable, and is one of the autoregressive lags, the model in (7) must be rewritten and treated as

$$X_t = m_1(X_{t-1}) + m_2(X_{t-1})X_{t-2} + 0.2\omega_t.$$

Consequently the functional coefficients of the autoregressive terms are

$$m_1(u_t) = 0.5u_t - 1.1u_t e^{-50u_t^2} \quad \text{and} \quad m_2(u_t) = 0.3 - 0.5e^{-50u_t^2},$$

where $u_t = X_{t-1}$.

To estimate the functional coefficients, begin by accounting for the variability in the response due to the term $m_1(u_t)$. Remove that variability, and use the pseudo-responses to estimate $m_2(u_t)$. Noting that the maximum lag is 2, we have.

1. For $j = 1,2$ in Eq. (4), fit a spline to the mean-centered data. Note that there is no $m_0(u_t)$ in this example. The result is an estimate $\tilde{m}_1(u_t)$ of $m_1(u_t)$ and $\tilde{m}_2(u_t)$ of $m_2(u_t)$.
2. Compute pseudo-responses $\tilde{W}_{t,2}, t = 3, \ldots, T$ using

$$\tilde{W}_{t,2} = X_t - \tilde{m}_1(u_t), \quad t = 3, 4, \ldots T.$$

These pseudo-responses are a proxy for the original realization, but with the effect of the $m_1(u_t)$ removed.

3. Fit a kernel regression to the pseudo-responses to get the SBK estimate $\widehat{m}_2(u_t)$ of $m_2(u_t)$.

Repeat steps 2 and 3, reversing the roles of m_1 and m_2.

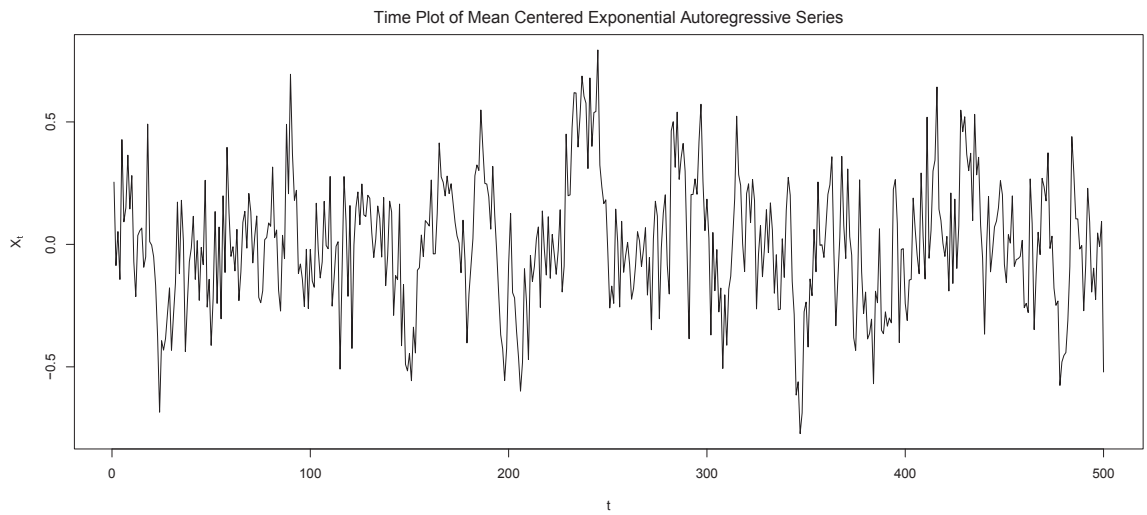


Fig. 1. Mean-centered realization of length $T = 500$ from an EXPAR(2) model given in Eq. (7).

Fig. 2 shows the estimation results of a simulated series from the exponential autoregressive model in Eq. (7) with IID standard normal ω_t and 500 samples. The dark curves of dots are the estimated functions, and the solid (thin) lines are the true functions. The dashed lines are the 95% pointwise confidence bands.

3.3. Spatial modeling for lattice data

For a fixed time t, consider a lattice process Y_s, $s = 1,2,\ldots S$. In this section, to ease notation, the time index t is suppressed. Let $\mathcal{N}_s$ represent a neighborhood around location s. The simultaneous autoregressive (SAR) model is defined as

$$Y_s = \sum_{j' \in \mathcal{N}_s} \beta_{s,j'} Y_{j'} + \delta_s,$$

where $\beta_{s,j'}$ is a set of coefficients that induces the spatial autocorrelation between locations j' and s in $\mathcal{N}_s$, and δ_s are independent, zero-mean, constant variance errors. The SAR model was first introduced by Ref. [46]. The adjective "simultaneous" describes the S autoregressions that occur simultaneously at each data location in the formulation. To fit this model in Section 4, we will employ a two nearest neighbor structure to define $\mathcal{N}_s$. The model is fitted using maximum likelihood estimators which are obtained using the R package spdep [47].

3.4. Spatio-temporal modeling

We now introduce two spatio-temporal models both of the form in (1). The first model we consider uses the noise process defined in (2) and assumes a separable covariance structure for this process. The time series structure, $X_{s,t}$, is modeled as an FCAR model using the SBK method. Values of p and d in (3) are allowed to vary

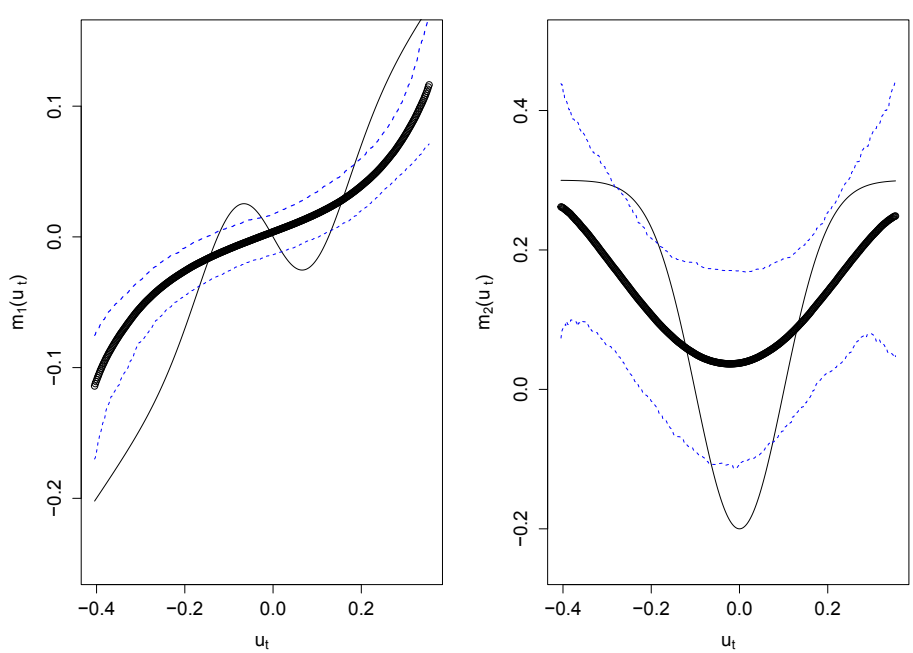


Fig. 2. Spline-backfitted kernel estimates of the coefficients $m_0(u) = 0.5u - 1.1u exp\{-50u^2\}$ (left panel) and $m_1(u) = 0.3 - 0.5exp\{-50u^2\}$ (right panel). Heavy curves are the estimates $\widehat{m}_0(u)$ and $\widehat{m}_1(u)$; thinner lines are the true functions $m_0(u)$ and $m_1(u)$; dashed lines are 95% confidence bands.

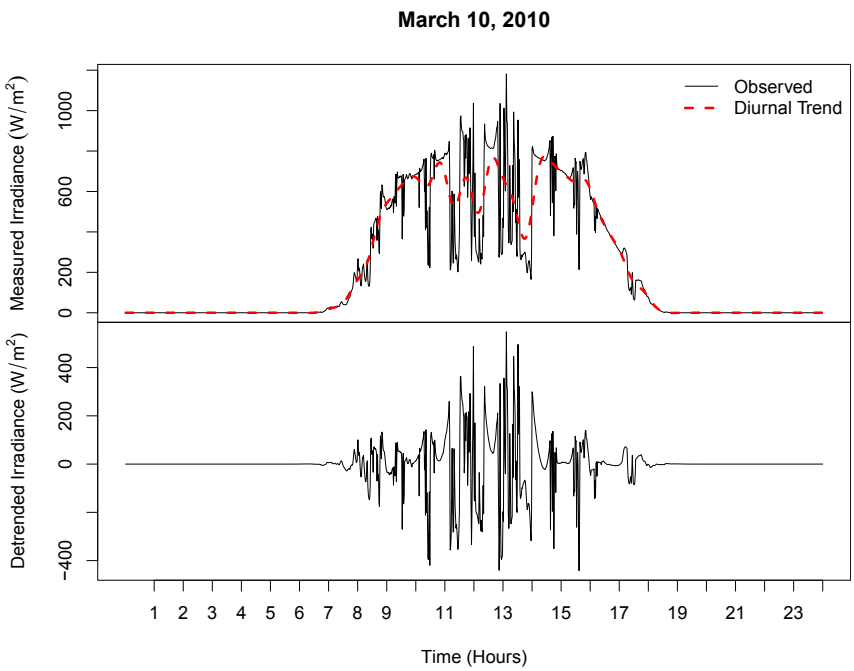


Fig. 3. Top graph is the time plot of 1-min averages (solid black) of irradiance measurements for March 10 with the local polynomial kernel estimate (dashed red) superimposed. The bottom plot is transformed irradiance (residuals after using local polynomial kernel regression to remove the diurnal trend). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between locations, and the spatial structure, $Y_{s,t}$, is modeled separately using a SAR model at each time. By modeling the time and space components separately, we are implicitly assuming separability. If this assumption is appropriate, then the order in which the two models are fit (time-then-space, or space-then-time) should not matter.

The second spatio-temporal model does not assume separability. Combining the FCAR(p) model with a generalized version of the SAR model, we define the space-time functional coefficient simultaneous autoregressive (FCSAR) model as

$$Z_{s,t} = \sum_{w=1}^{b} \sum_{\ell \in \mathcal{N}_s} \beta_{s,\ell,w} Z_{\ell,t-w} + \sum_{k=1}^{p_s} m_{k,s}(Z_{s,t-d_s}) Z_{s,t-k} + \varepsilon_{s,t}, \qquad (8)$$

where b is the spatial time order for the spatial component in the model, $\beta_{s,\ell,w}$ is the spatial autocorrelation between locations s and ℓ at a time lag of w, $Z_{s,t-ds}$ is a delay variable, and $\varepsilon_{s,t}$ are IID with mean zero and constant variance σ_ε^2. We allow the values of p_s and d_s to vary among locations. The FCSAR model is based on the space-time simultaneously specified autoregressive model of [26].

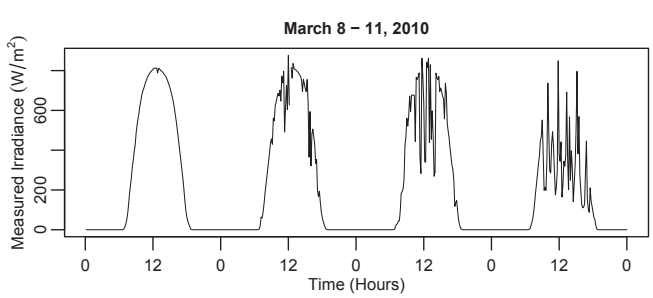


Fig. 4. Time plot of 10-min time averaged GHI for March 8–11, 2010. These days illustrate the three cloud cover categories we used for the Lanai data. March 8 is clear, March 10 is partly cloudy, and March 11 is mostly cloudy.

In Ref. [26], the model in (8) has a linear autoregressive structure to the time series term for which a Bayesian approach is used for estimation. For our model, we estimate the $\beta_{s,\ell,w}$'s using maximum likelihood. Define $I_S \in \mathbb{R}^{S\times S}$ as an identity matrix, $\mathbf{D} \in \mathbb{R}^{S\times S}$ as a matrix of weights with elements $\delta_{ij} = 1/(2S)$ if the jth sensor is one of the two nearest neighbors of sensor i and $\delta_{ij} = 0$ otherwise, and ρ as a spatial autocorrelation parameter. Let $\mathbf{Z}_t = (Z_{1,t}, Z_{2,t}, \ldots, Z_{S,t})'$ and $\mathbf{Z} = \mathbf{D}\,(\mathbf{Z}_{t-1}, \ldots, \mathbf{Z}_{t-b})$. The log-likelihood function is

Table 1
Root mean squared error (RMSE, in W/m^2) for the four spatio-temporal models of the days with clear, partly cloudy, and mostly cloudy conditions. Columns S-T and T-S contain the RMSE for the separable spatio-temporal models. Column S-T contains RMSE for data with the spatial component fit first, then time; Column T-S contains RMSE with the time component fit first, then space. The last two columns contain RMSE for the nonseparable FCSAR model with spatial time orders $b = 1$ and $b = 2$, respectively. The values in parentheses are the percentages of the mean GHI for that day that the RMSE represents.

	Date	Separable		FCSAR	
		S-T	T-S	$b = 1$	$b = 2$
Clear	Feb. 3	0.64 (0.1)	2.32 (0.5)	0.38 (0.1)	0.16 (0.1)
	Feb. 16	3.58 (0.8)	15.03 (3.2)	2.28 (0.5)	1.36 (0.3)
	Mar. 8	0.62 (0.1)	1.99 (0.4)	0.37 (0.1)	0.22 (0.1)
	Mar. 19	0.60 (0.1)	6.33 (1.1)	0.36 (0.1)	0.22 (0.1)
	Oct. 21	0.94 (0.2)	4.01 (0.9)	0.65 (0.1)	0.42 (0.1)
	Dec. 16	1.13 (0.2)	6.17 (1.2)	0.87 (0.2)	0.44 (0.1)
Partly cloudy	Mar. 7	8.26 (2.6)	63.68 (19.8)	6.79 (2.1)	3.35 (1.0)
	Apr. 1	7.88 (2.2)	99.53 (27.3)	6.96 (1.9)	4.51 (1.2)
	May 10	6.18 (2.1)	56.02 (19.1)	5.84 (2.0)	3.72 (1.3)
	Jun. 4	9.69 (2.8)	50.91 (14.7)	6.34 (1.8)	3.36 (1.0)
	Jun. 28	4.07 (1.4)	55.00 (18.8)	2.78 (1.0)	1.76 (0.6)
	Nov. 15	10.70 (3.0)	58.14 (16.3)	9.07 (2.5)	6.00 (1.7)
Mostly cloudy	Feb. 1	8.61 (2.4)	62.60 (17.7)	7.80 (2.2)	4.88 (1.4)
	Mar. 15	13.26 (4.0)	116.71 (35.0)	10.65 (3.2)	5.50 (1.6)
	Apr. 6	6.00 (1.9)	48.16 (15.0)	3.93 (1.2)	2.43 (0.8)
	May 31	11.23 (2.2)	91.11 (17.8)	8.53 (1.7)	5.08 (1.0)
	Aug. 3	6.05 (2.0)	64.36 (21.0)	5.10 (1.7)	2.82 (1.0)
	Oct. 27	4.54 (1.2)	93.69 (25.4)	3.23 (0.9)	1.89 (0.5)

$$\ln L = C + \sum_{t=1}^{T} \ln\left|I_S - \rho\mathbf{D}\right| - \sum_{t=1}^{T}\left(\frac{S}{2}\right)\ln(\mathbf{e}'\mathbf{e}) \tag{9}$$

where C represents a constant not involving the parameters,

$$\mathbf{e} = \mathbf{e}_\beta - \rho\mathbf{D}\mathbf{Z}_t + \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{D}\mathbf{Z}_t,$$

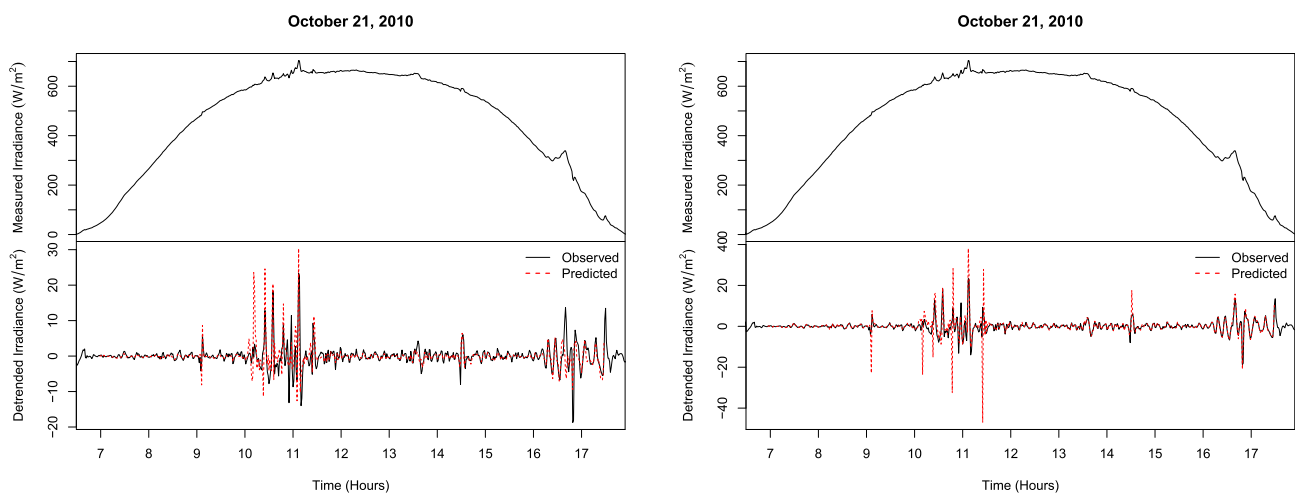
$$\mathbf{e}_\beta = \mathbf{Z}_t - \mathbf{Z}\boldsymbol{\beta}_s,$$

$$\boldsymbol{\beta}_s = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Z}_t,$$

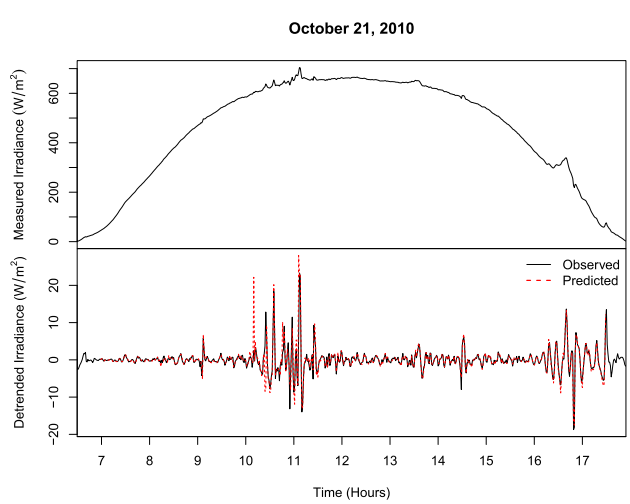
and $\boldsymbol{\beta}_s = (\beta_{s,\ell,1},\ldots,\beta_{s,\ell,b})'$. Note that the weight matrix we use forces each neighboring parameter to be the same, i.e., $\beta_{s,1,w} = \beta_{s,2,w}$ in the two nearest neighbors structure. We maximize (9) with respect to $\boldsymbol{\beta}_s$ for each location s. For more detail on estimating a SAR model using maximum likelihood see Ref. [48]. The functional coefficients $m_{k,s}(Z_{s,t-ds})$ are estimated using the SBK method defined in Section 3.2. The values of p_s and d_s are chosen for each site s by minimizing the mean square error of the fit to the observed time series at s. The value of the bandwidth h for the kernel smoothing in the SBK method is chosen via a modified, multifold cross-validation procedure at each site s [see Ref. 39]

4. Application and discussion

To illustrate the utility and compare the performance of the proposed models, we model GHI data at the 1.2 MV La Ola PV plant on the island of Lanai, Hawaii. The La Ola PV plant contains a grid of 12 single-axis tracked arrays arranged in three columns and four rows covering a total area of approximately 250 m by 250 m. The La Ola plant is located at 20.7669N 156.9229W, near the southern end of the island of Lanai, atop the plateau in a subtropical dry forest climate [49]. Lanai lies in the rainfall shadow of Maui and thus is relatively dry, yet measured irradiance shows that variable irradiance due to cloud movement are the dominant pattern [50]. We chose to model GHI rather than POA irradiance to illustrate a more



(a) October 21, Time-Space. (b) October 21, Space-Time.



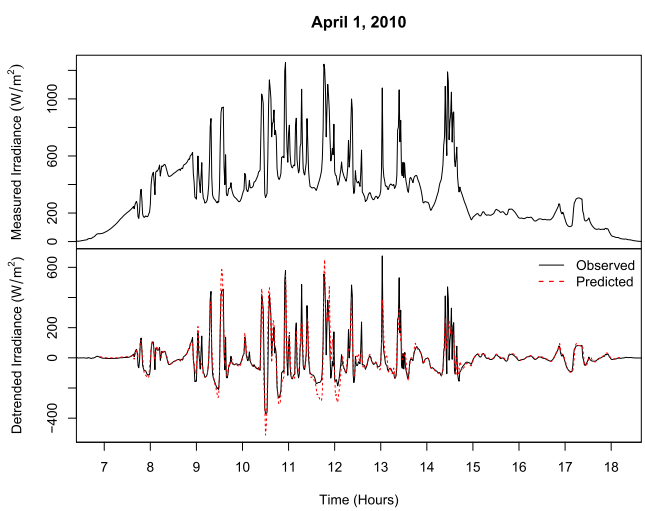
(c) October 21, Nonseparable.

Fig. 5. Time plots of 1-min time averaged irradiance and of transformed irradiance with predicted values superimposed in red for a clear day, October 21. Forecasting was conducted using the (a) time-then-space fitting (b) space-then-time fitting (c) nonseparable model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

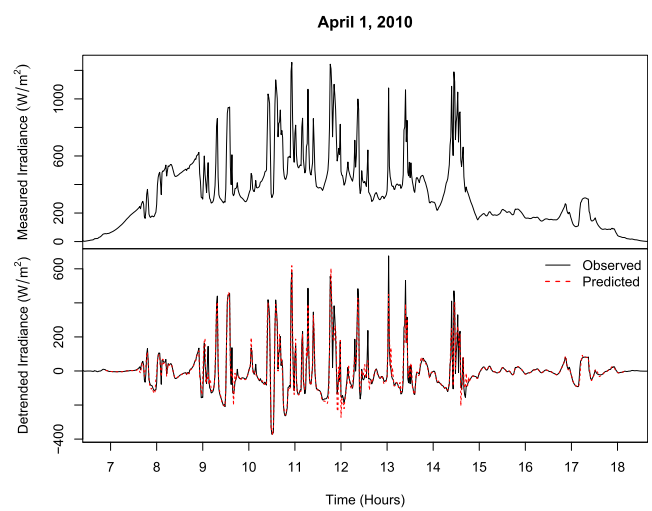
general application of our method. However, the La Ola data are POA irradiance rather than GHI. Sandia National Laboratories and SunPower Corporation designed an irradiance measurement system in part to study the effects of the movement of cloud shadows across the PV arrays on the power output of the plant [1]. Plane-of-array (POA) irradiance (in W/m^2) is measured at the midpoint of each tracking array using LiCor-200 pyranometers.

Before fitting the models it is necessary to remove the diurnal trend, a step which we found somewhat difficult. Clear sky models are available for removing trends from measured GHI data; a review of some of these models can be found in Ref. [51]. We set out to use clear sky models to remove the diurnal trend, which would present no great difficulty for measured GHI. We know of no equivalent "clear-sky" model for POA irradiance (although, if the tracking algorithm is known with sufficient precision, such a model could be assembled by applying a GHI-to-POA translation model, e.g., the DISC model of [52] to the output of a clear-sky model). We translated POA irradiance to GHI by assuming the isotropic sky model for the sky diffuse irradiance and using concurrent measurements of diffuse horizontal irradiance (DHI) and direct normal irradiance (DNI) from a nearby rotating shadowband radiometer (RSR) operated by the National Renewable Energy Laboratory. Because tracker rotations are not measured we estimated the angle of incidence on the modules using a generic algorithm for single-axis tracking [53]. Even with the use of measured DHI and DNI, the estimated GHI profiles were not well-matched with the output of available clear-sky models, and the clear sky models performed poorly in removing the trend. Consequently, we removed the diurnal trend in the estimated GHI by using a local polynomial kernel regression implemented in the KernSmooth package [54] in the R programming software.

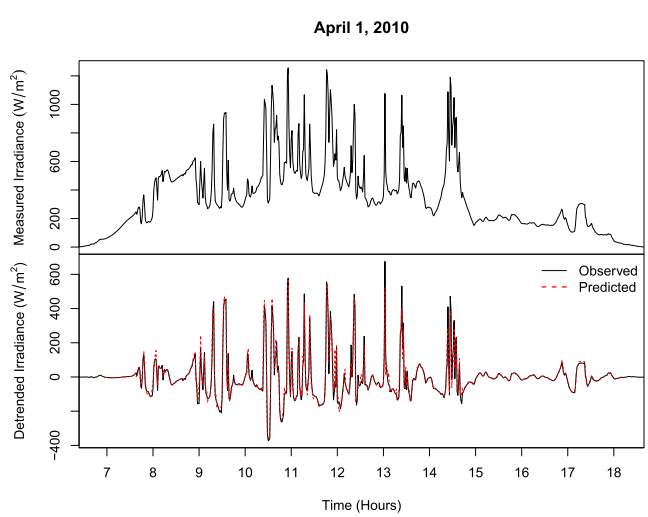
We selected one year (i.e., January 1, 2010 to December 31, 2010) of POA irradiance measurements, which are recorded every second. We observed little to no variability from one irradiance



(a) April 1, Time-Space.



(b) April 1, Space-Time.



(c) April 1, Nonseparable.

Fig. 6. Time plots of 1-min time averaged irradiance and of transformed irradiance with predicted values superimposed in red for a partly cloudy day, April 1. Forecasting was conducted using the (a) time-then-space fitting (b) space-then-time fitting (c) nonseparable model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measurement to the next at 1 s intervals and consequently reduced the data by time averaging. We investigated time-averages of lengths of 30 s, 1 min, 5 min, and 10 min. Longer time averages (e.g., 15 and 20 min) were also considered but did not appear to be significantly different from the 10 min averages. We chose to use 1-min average data for our exploratory work.

The top time plot in Fig. 3 contains the 1-min time averages of estimated GHI in solid black superimposed with the local polynomial kernel regression estimate in dashed red for March 10. The bottom time plot contains the residuals, hereafter referred to as "transformed irradiance," obtained after removing the diurnal trend by subtracting the kernel fit.

We examined a large number of time plots of GHI to find days with different variability characteristics. For each day, the weather condition was classified visually as being in one of three categories: clear, partly cloudy, and mostly cloudy, by the variability and magnitude of GHI. Fig. 4 shows an example of how the different categories looked visually. We show the measured GHI for March 8–11 for which March 8 is considered a clear day due to its smoothly changing irradiance similar in magnitude to that predicted a clear sky model. March 9 and 10 are considered as partly cloudy. These days show variation in GHI due to moving clouds with a general trend that is similar to a clear sky day (e.g., March 8). March 11 is considered a mostly cloudy day because the general trend is much lower than the magnitude of irradiance during a clear sky day. Days with irradiance that changed among these categories were not considered in this paper. In Section 5 we discuss future work in which a weather covariate can be used to allow the model to be fitted to days with varying weather conditions.

For 2010, in Lanai, HI, only six days could be classified as "clear" throughout the entire day. For partly cloudy and mostly cloudy conditions, we found many days. For both of these weather conditions, six days in 2010 were randomly selected.

For each selected day, we explored whether assuming space-time covariance separability in (1) would be justified. Using the

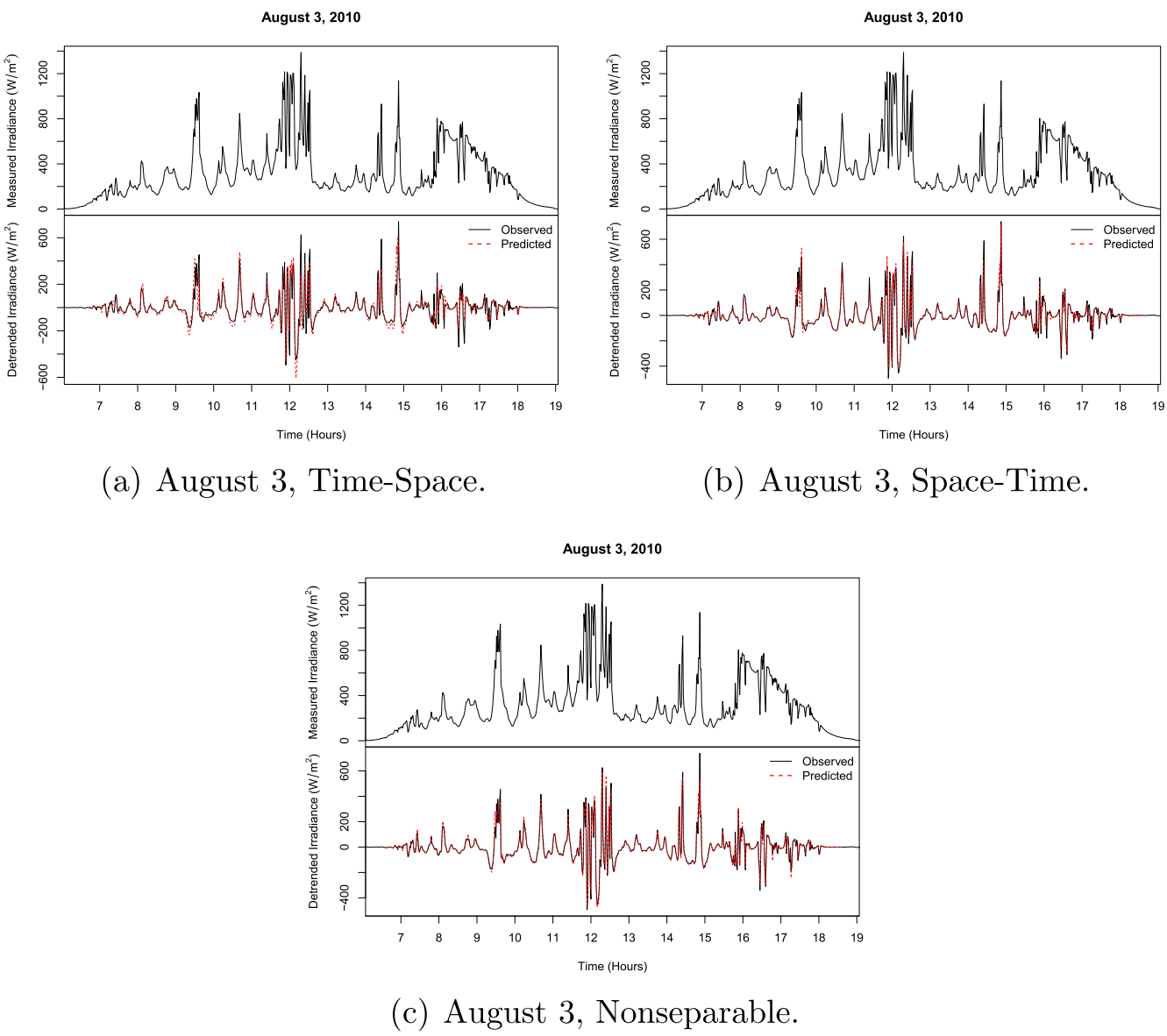
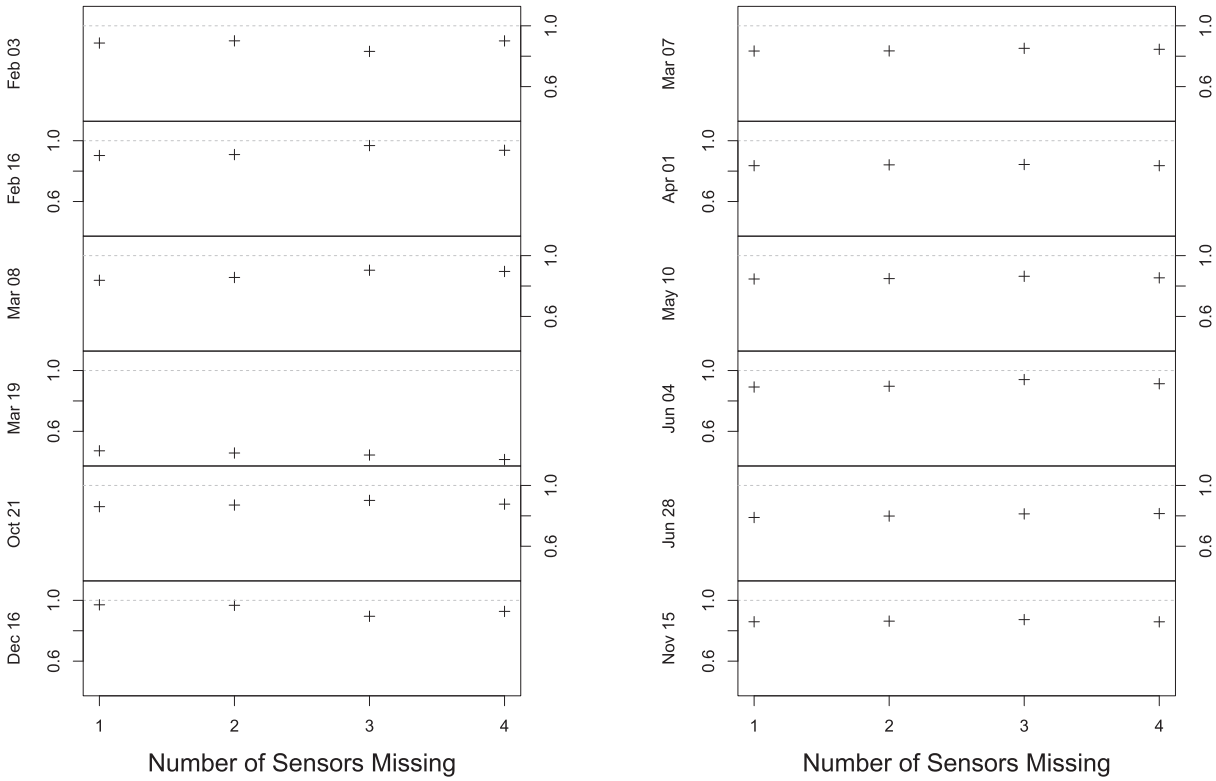
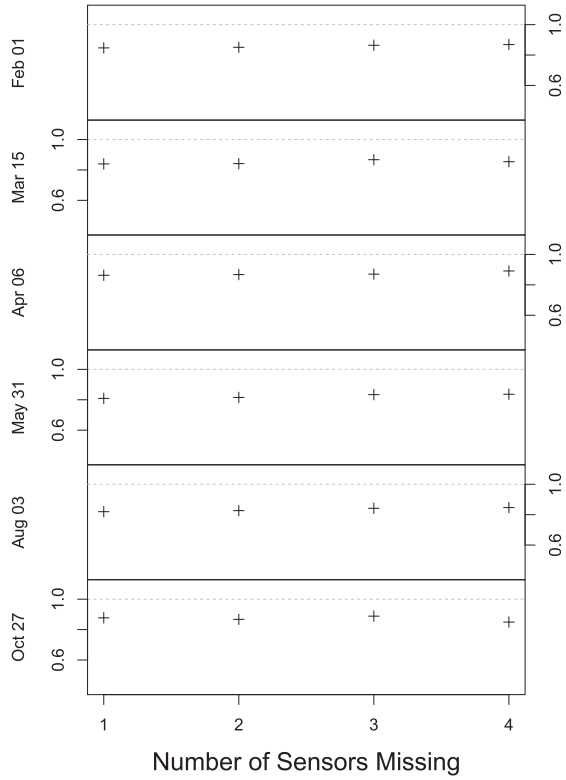


Fig. 7. Time plots of 1-min time averaged irradiance and of transformed irradiance with predicted values superimposed in red for a mostly cloudy day, August 3. Forecasting was conducted using the (a) time-then-space fitting (b) space-then-time fitting (c) nonseparable model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a) Clear days.

(b) Partly cloudy days.



(c) Mostly cloudy days.

Fig. 8. Plots of the ratios of the RMPE for (a) the clear days, (b) the partly cloudy days, and (c) the mostly cloudy days. The ratios are calculated as the RMPE of the FCSAR model divided by the RMPE of linear interpolation.

Table 2
Root mean squared error (RMSE, W/m^2) for the nonseparable spatio-temporal models with $b = 2$ of the days with clear, partly cloudy, and mostly cloudy conditions. The columns are the RMSE for data at 30-s, 1-min, 5-min, and 10-min averages. The values in parenthesis are the R_a^2.

Condition	Date	30-sec	1-min	5-min	10-min
Clear	Feb. 3	0.38 (0.990)	0.30 (0.993)	0.19 (0.999)	0.16 (0.999)
	Feb. 16	4.98 (0.952)	3.68 (0.970)	1.82 (0.988)	1.36 (0.963)
	Mar. 18	1.79 (0.970)	0.99 (0.989)	0.42 (0.999)	0.22 (0.992)
	Mar. 19	0.81 (0.999)	0.63 (0.998)	0.32 (0.952)	0.22 (0.993)
	Oct. 21	3.94 (0.846)	2.21 (0.926)	0.67 (0.991)	0.42 (0.983)
	Dec. 16	0.67 (0.998)	0.61 (0.998)	0.49 (0.952)	0.44 (0.993)
Partly cloudy	Mar. 7	25.66 (0.920)	17.68 (0.960)	5.42 (0.991)	3.35 (0.997)
	Apr. 1	31.75 (0.932)	24.65 (0.960)	6.99 (0.996)	4.51 (0.998)
	May 10	25.91 (0.927)	19.68 (0.954)	7.70 (0.987)	3.72 (0.999)
	June 4	22.05 (0.907)	20.60 (0.920)	7.21 (0.986)	3.36 (0.995)
	June 28	13.85 (0.935)	10.12 (0.966)	3.88 (0.994)	1.76 (0.999)
	Nov. 15	27.69 (0.908)	22.56 (0.935)	9.24 (0.979)	6.00 (0.987)
Mostly cloudy	Feb. 1	26.46 (0.899)	24.12 (0.917)	9.27 (0.982)	4.88 (0.993)
	Mar. 15	27.85 (0.938)	20.29 (0.968)	6.05 (0.996)	5.50 (0.997)
	Apr. 6	24.22 (0.883)	19.44 (0.930)	4.01 (0.997)	2.43 (0.997)
	May 31	42.26 (0.899)	30.26 (0.943)	9.60 (0.988)	5.08 (0.997)
	Aug. 3	25.92 (0.948)	20.59 (0.966)	6.15 (0.995)	2.82 (0.998)
	Oct. 27	7.29 (0.980)	5.89 (0.985)	2.77 (0.996)	1.89 (0.999)

separable models we fit the data in two ways: space-then-time and time-then-space. If the separability assumption is appropriate, then the two models are equivalent and should yield similar results. For the space-then-time approach, we first fit the SAR model to the 16 sensors for each time, t. We obtain the residuals from the fitted SAR model, and then apply the FCAR model to each sensor separately. For the time-then-space model, we first fit the FCAR model to the detrended irradiance for each of the 16 sensors, and then the residuals from the fitted FCAR model are fit with the SAR model at each time point. For each approach the root mean square errors

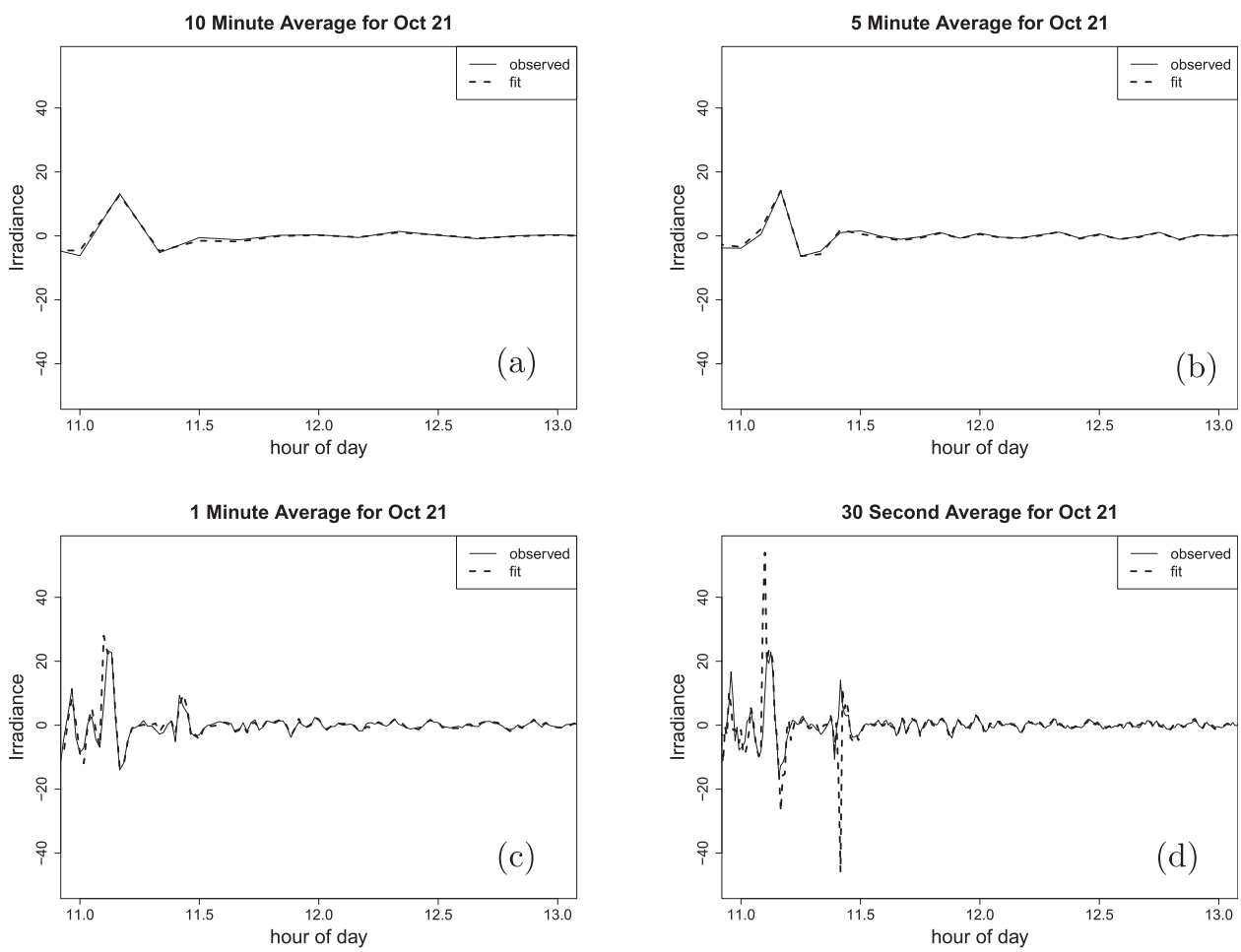


Fig. 9. Plots of the detrended observed irradiance and the fit of the FCSAR model for sensor 1 from 11:00 to 13:00 on October 21 (a clear day). The different plots are for different averages: (a) 10 min, (b) 5 min, (c) 1 min, and (d) 30 s.

(RMSEs) (over all sensor locations and times) for eighteen days with three different weather conditions are found in the first two columns of Table 1. For all days considered in this study, the RMSE for the model that fit space first is considerably smaller than when time was fit first. This is a strong indication that the assumption of separable covariance structures is not supported and that nonseparable models should be employed.

For a fixed time t, because the sensors are at fixed locations, the spatial structure can properly be considered a lattice. Consequently, for the nonseparable model we fit the FCSAR model in (8) for spatial time orders $b = 1,2$. The last two columns in Table 1 contain values of RMSE for these two fits. The FCSAR model with $b = 2$ has the smallest RMSE for all 18 days, indicating the best fit among the models considered. For cloudy and partly cloudy conditions RMSE decreases substantially from $b = 1$ to $b = 2$ indicating that a lagged model is needed for greater prediction accuracy. For comparing accuracy between days with different irradiance levels, we provide the percentage of GHI that the RMSE represents. These are shown in the parentheses in Table 1.

Figs. 5–7 contain six plots, grouped in three pairs. Each figure displays one sensor location for one day: a clear day (October 21, Fig. 5); a cloudy day (April 1, Fig. 6); and a partly cloudy day (August 3, Fig. 7). For any one pair of plots, the top graph contains the GHI data represented by a solid black line, and the modeled GHI represented by a red dashed line. The bottom graph contains the detrended GHI data (solid black line) and the detrended modeled GHI (red dashed line). For all three days, the set of two plots labeled (a) were fit using a separable time-then-space approach; the two sets of plots labeled (b) were fit using a separable space-then-time approach; and the plots labeled (c) were fit using the nonseparable FCSAR model with $b = 2$. The collection of figures illustrates the nonseparable approach yields the best fit, regardless of the weather conditions, which is in agreement with minimum RMSE in Table 1. However, where RMSE is an aggregate measure of goodness-of-fit, the plots illustrate that at individual time points, the goodness-of-fit is uniformly better for the nonseparable model.

Forecasting the FCSAR model in time is largely dependent on using the SBK method for forecasting the FCAR term in (8). In Ref. [55], methodology is presented for forecasting a FCAR model using the SBK method. In this paper, we examine the performance of forecasting (8) in space for unobserved locations by using cross-validation. We simulated unobserved locations by omitting one or several sensors from our data set, and compared FSCAR model performance with a commonly used interpolation technique to judge the potential improvement offered by the FCSAR model.

Unobserved data are often estimated by interpolating between nearby sensors; one such technique is natural neighbor interpolation which comprises a weighted average with weights determined by a Voronoi partition [56]. A Voronoi partition divides the space that contains the sensors into regions. Each sensor will have a corresponding region consisting of all points closer to that sensor than to any other. We constructed a Voronoi partition on the set of

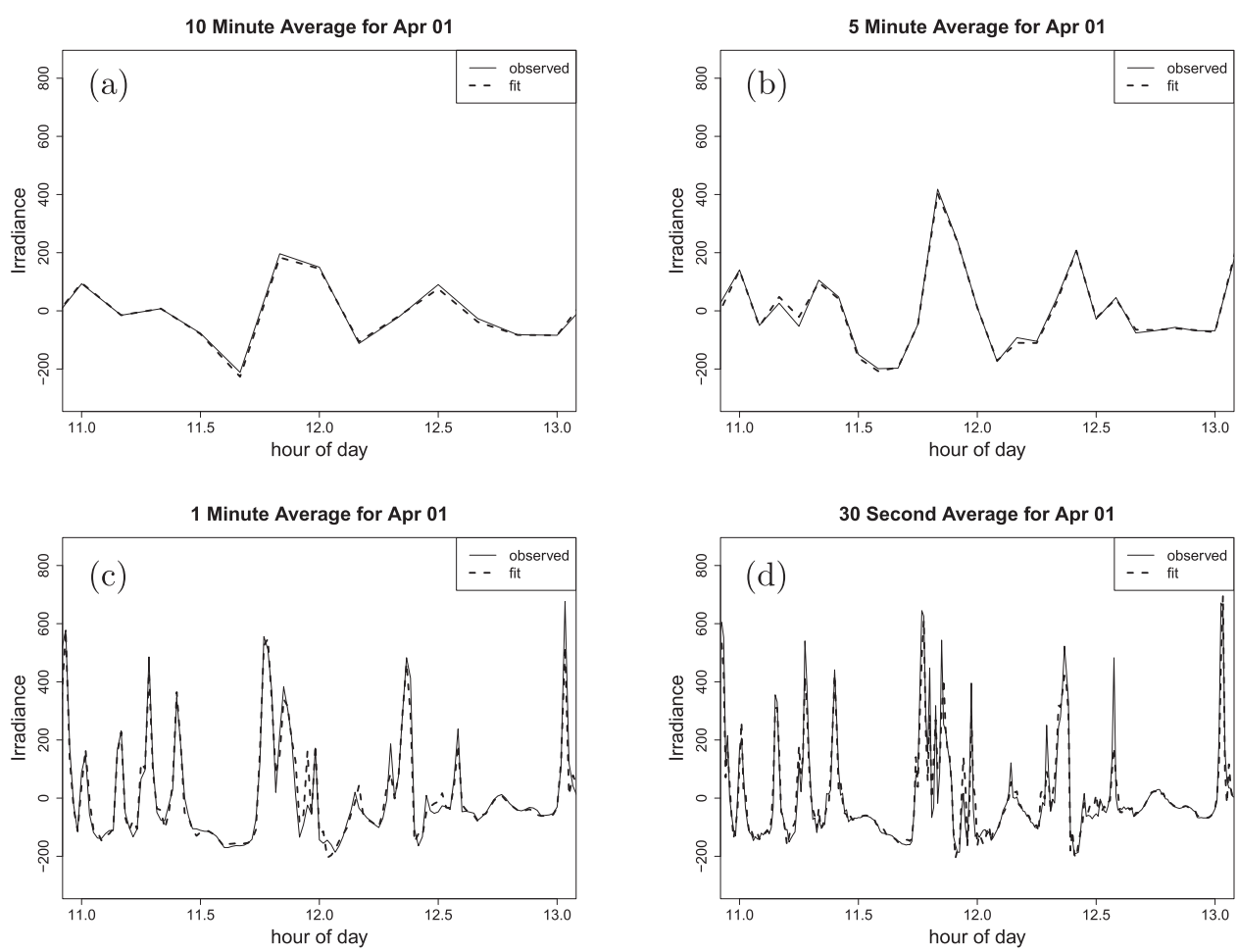


Fig. 10. Plots of the detrended observed irradiance and the fit of the FCSAR model for sensor 1 from 11:00 to 13:00 on April 1 (a partly cloudy day). The different plots are for different averages: (a) 10 min, (b) 5 min, (c) 1 min, and (d) 30 s.

training sensors along with the location of the missing sensor. For cross-validation, we took the weighted average of the training sensors where the weights are determined by the size of the regions. This weighted average is used for the prediction for the missing sensors.

We fit the FCSAR model to the training set of sensors with $b = 2$ and using a two nearest neighbor structure for $\mathcal{N}_s$. For each missing sensor, we determined the two nearest neighbors and predicted the irradiance by using the estimated β's for those neighbors.

For our set of 16 sensors, we calculated the predictions with $k = 1,2,3,4$ missing sensor locations. For $k > 1$, we predicted for each missing location one at a time. We calculated the root mean prediction error (RMPE) as

$$RMPE_{\Omega_i} = \frac{1}{Tk} \sum_{s \in \Omega_i} \left(\sum_{t=1}^{T} \left(\tilde{Z}_{s,t} - Z_{s,t} \right)^2 \right),$$

where Ω_i is the ith set of k missing sensors, $\tilde{Z}_{s,t}$ is the predicted irradiance for the sth sensor at time t, and $Z_{s,t}$ is the observed irradiance. The $RMPE_{\Omega_i}$ is calculated for all $K = \binom{16}{k}$ possible combinations of k missing sensors. The mean RMPE is calculated as

$$\overline{RMPE}_k = \frac{1}{K} \sum_{i=1}^{K} RMPE_{\Omega_i}.$$

To compare the FCSAR model to the interpolation method, we take the ratio

$$\frac{\overline{RMPE}_k \text{ for FCSAR}}{\overline{RMPE}_k \text{ for interpolation}}.$$

The ratios for all 18 days are plotted in Fig. 8. All ratios are less than one indicating that the FCSAR model performs better at predicting the missing sensors.

To examine the effect of different time averaging windows on the FSCAR model's performance we fit the model for a range of time averaging windows, from 10 min down to 30 s. For each day, we calculated the model's RMSE as well as the adjusted coefficient of determination R_a^2. The adjusted coefficient of determination R_a^2 quantifies the level of agreement between the data and a fitted model taking into account the number of variables used in the model:

$$R_a^2 = 1 - \frac{SS_{\text{Fit}}/(TS - \nu_{\text{Fit}})}{SS_{\text{Total}}/(TS)},$$

where

$$SS_{\text{Fit}} = \sum_{s=1}^{S} \sum_{t=1}^{T} \left(\tilde{Z}_{s,t} - Z_{s,t} \right)^2,$$

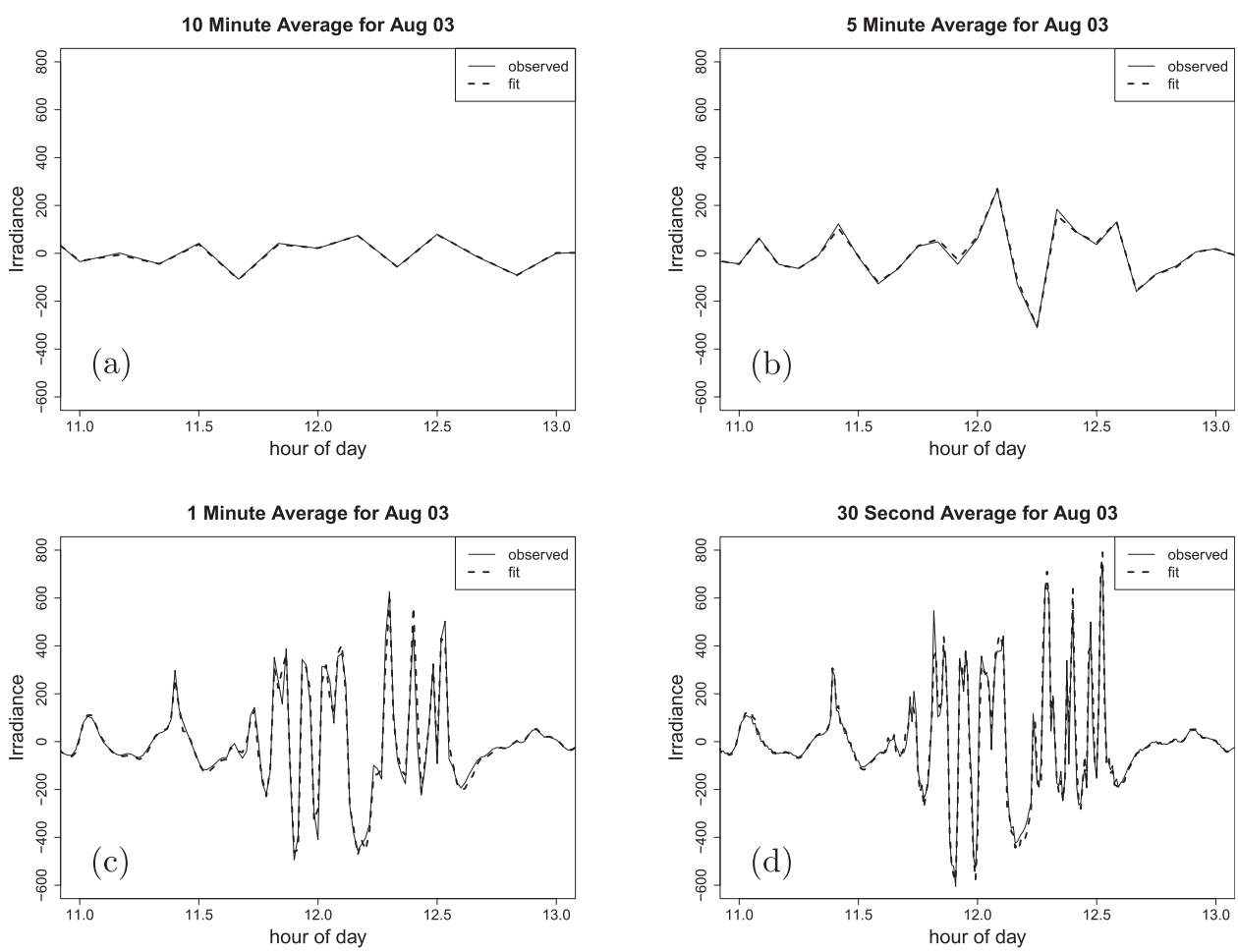


Fig. 11. Plots of the detrended observed irradiance and the fit of the FCSAR model for sensor 1 from 11:00 to 13:00 on August 3 (a mostly cloudy day). The different plots are for different averages: (a) 10 min, (b) 5 min, (c) 1 min, and (d) 30 s.

$$SS_{\text{Total}} = \sum_{s=1}^{S} \sum_{t=1}^{T} \left(Z_{s,t} - \overline{Z}\right)^2,$$

$$\overline{Z} = \frac{1}{TS} \sum_{s=1}^{S} \sum_{t=1}^{T} Z_{s,t},$$

$\tilde{Z}_{s,t}$ is the predicted irradiance for the sth sensor at time t, $Z_{s,t}$ is the observed irradiance, and ν_{Fit} is the number of parameters used in the fit. Since we are using kernel regression to fit the time series, we must estimate the number of parameters associated with that regression. For the SBK estimate of the kth coefficient function in (8), the effective number of parameters is the trace of the smoother matrix

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \left(\frac{1}{T}\mathbf{C}'\mathbf{M}\mathbf{C}\right)^{-1} \frac{1}{T}\mathbf{C}'\mathbf{M}$$

in (6) [see Refs. [40,57]. We calculate the total number of parameters as the sum of the parameters in the first double sum in (8) plus the sum of the effective number of parameters for the FCAR term. The values of R_a^2 and RMSE are shown in Table 2. We show the fits of three days for the different time averages in Figs. 9–11.

Table 2 shows that as the time averaging window decreases, RMSE increases and and R_a^2 decreases, both indicating increasing disagreement between data and model. However, as the time averaging window decreases, variance in time averaged data at any

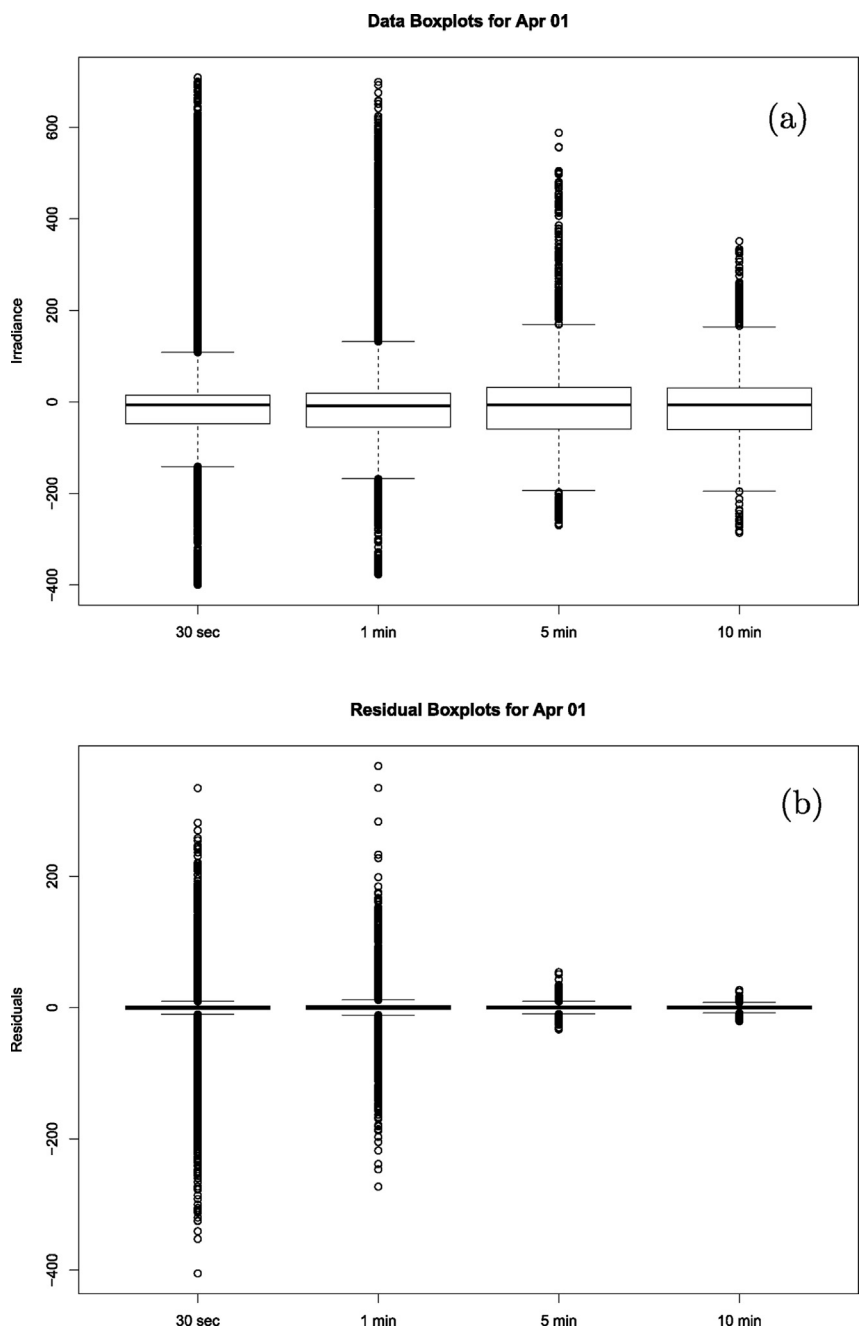


Fig. 12. Boxplots for (a) the observed detrended data and (b) the residuals of the fit of the FCSAR model with $b = 2$ for April 1 (a partly cloudy day).

individual location increases substantially (Figs. 9–11). As ramps in the data increase in both magnitude and frequency the largest residuals of the fitted model also increase. Similar patterns are evident in the spatially-averaged data. Fig. 12 compares distributions for the spatially averaged detrended irradiance data and corresponding distributions residuals for the fitted FSCAR model for a partly cloudy day. As the averaging window decreases., outliers increase in both the data and the model residuals also increase, leading to the increasing RMSE and decreasing R_a^2 evident in Table 2. However, the FCSAR model continues to fit the bulk of the data equally well across all time averaging windows, as is demonstrated by the relatively constant boxes and whiskers across the different time averages. Thus, the FCSAR model follows time averaged data equally well for various averaging windows.

5. Conclusion

We have presented a novel nonseparable spatio-temporal model for GHI data. This approach, termed the FCSAR model, outperforms a natural neighbor interpolation when predicting GHI at unobserved locations over the footprint of a PV system. We compared the nonseparable FCSAR model with simpler, separable models, and find little support for models that assume a separable covariance structure. The FSCAR model integrates an FCAR form for the time series component of the model and a SAR form for the spatial component. The FCAR(p) form of the time series component of our nonseparable model makes the FCSAR model flexible and reliable, and may be suitable for fitting irradiance data in general. Currently, the model is fit separately on each day. Further research will consider validating the fitted models by comparing predicted aggregate irradiance with generated power for a much larger solar power plant than La Ola. Future work may also explore adding a weather condition covariate that will allow the model to be fit over days with different weather conditions, by permitting the coefficient functions in the time series structure to vary based on weather condition. The use of the SBK method for estimating the time series structure has an assumption that the second derivative of the coefficient function is continuous [e.g. see 41, 55]. The use of a weather condition covariate may result in coefficient functions that do not satisfy this assumption. Patrick et al. [55] examine an example for which the assumption is relaxed and shows the SBK method still performs well in estimation and forecasting. Thus, the use of an weather condition covariate will be worth exploring.

Acknowledgments

The research was performed under contract (PO 1303122) with Sandia National Laboratories, a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Dr. Patrick's work was largely completed as a part of his dissertation work while he was at Baylor University. The authors thank Justin Sims for his help in creating the graphs provided in the figures throughout the manuscript.

References

[1] S. Kuszamaul, A. Ellis, J. Stein, L. Johnson, Lanai high-density irradiance sensor network for characterizing solar resource variability of MW-scale PV system, in: 35th Photovoltaic Specialists Conference, 2010. Honolulu, HI.
[2] A.R. Dyreson, E.R. Morgan, S.H. Monger, T.L. Acker, Modeling solar irradiance smoothing for large PV power plants using a 45-sensor network and the wavelet variability model, Sol. Energy 110 (2014) 482–495.
[3] M. Lave, A. Ellis, J.S. Stein, Simulating Solar Power Plant Variability: a Review of Current Methods, Tech. Rep., Sandia National Laboratories, Albuquerque, NM, 2013a. Sandia National Laboratories (SNL-CA), Livermore, CA (United States)
[4] M. Lave, J. Kleissl, J.S. Stein, A wavelet-based variability model (WVM) for solar PV power plants, Sustain. Energy IEEE Trans. 4 (2) (2013b) 501–509.
[5] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira, E. Izco, From irradiance to output power fluctuations: the PV plant as a low pass filter, Prog. Photovoltaics Res. Appl. 19 (5) (2011) 505–510.
[6] A. Longhetto, G. Elisei, C. Giraud, Effect of correlations in time and spatial extent on performance of very large solar conversion systems, Sol. Energy 43 (2) (1989) 77–84.
[7] SolarAnywhere, www.solaranywhere.com, 2015.
[8] J. Bing, P. Krishnani, O. Bartholomy, T. Hoff, R. Perez, T. Vargas, Solar monitoring, forecasting, and variability assessment at SMUD, Proc. World Renew. Energy Form (May 13–17, 2012) (Denver, CO).
[9] Southern California Edison, Pro-forma Renewable Power Purchase and Sale Agreement, 2, 2014. Exhibit P-2b, https://scerps.accionpower.com/_scerps_1401/home.asp.
[10] T.E. Hoff, R. Perez, Quantifying PV power output variability, Sol. Energy 84 (10) (2010) 1782–1793.
[11] J.D. Frank, P. Beaucage, M.C. Brower, High frequency irradiance and power output for grid integration and storage evaluations, in: Photovoltaic Specialist Conference (PVSC), 2014, pp. 1011–1013, 2014 IEEE 40th, IEEE.
[12] M. Hummon, E. Ibanez, G. Brinkman, D. Lew, Sub-hour solar data for power system modeling from static spatial variability analysis, Natl. Renew. Energy Lab. (November 12–13, 2012).
[13] D. Yang, P. Jirutitijaroen, W. M. Walsh, Hourly solar irradiance time series forecasting using cloud cover index, Solar Energy.
[14] C.A. Glasbey, Nonlinear autoregressive time series with multivariate Gaussian mixtures as marginal distribution, J. R. Stat. Soc. Ser. C Appl. Statistics 50 (2001) 143–154.
[15] G. Reikard, Predicting solar radiation at high resolutions: a comparison of time series forecasts, Sol. Energy 83 (3) (2009) 342–349.
[16] C. Paoli, C. Voyant, M. Muselli, M.-L. Nivet, Forecasting of preprocessed daily solar radiation time series using neural networks, Sol. Energy 84 (12) (2010) 2146–2160.
[17] C.A. Glasbey, D.J. Allcroft, A spatiotemporal autoregressive moving average model for solar radiation, J. R. Stat. Soc. Ser. C Appl. Statistics 57 (3) (2008) 343–355.
[18] C. Glasbey, R. Graham, A. Hunter, Spatio-temporal variability of solar energy across a region: a statistical modelling approach, Sol. Energy 70 (4) (2001) 373–381.
[19] D. Yang, C. Gu, Z. Dong, P. Jirutitijaroen, N. Chen, W.M. Walsh, Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging, Renew. Energy 60 (2013) 235–245.
[20] D. Yang, Z. Dong, T. Reindl, P. Jirutitijaroen, W.M. Walsh, Solar irradiance forecasting using spatio-temporal empirical kriging and vector autoregressive models with parameter shrinkage, Sol. Energy 103 (2014) 550–562.
[21] T.E. Hoff, R. Perez, Modeling PV fleet output variability, Sol. Energy 86 (8) (2012) 2177–2189.
[22] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker, T.E. Hoff, Short-term irradiance variability: Preliminary estimation of station pair correlation as a function of distance, Sol. Energy 86 (8) (2012) 2170–2176.
[23] M. Lave, J. Kleissl, E. Arias-Castro, High-frequency irradiance fluctuations and geographic smoothing, Sol. Energy 86 (8) (2012) 2190–2199.
[24] V.P. Lonij, A.E. Brooks, A.D. Cronin, M. Leuthold, K. Koch, Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors, Sol. Energy 97 (2013) 58–66.
[25] E. Arias-Castro, J. Kleissl, M. Lave, A Poisson model for anisotropic solar ramp rate correlations, Sol. Energy 101 (2014) 192–202.
[26] M.W. Woolrich, M. Jenkinson, J.M. Brady, S.M. Smith, Fully Bayesian spatio-temporal modeling of fMRI data, Med. Imaging IEEE Trans. 23 (2) (2004) 213–231.
[27] J.L. Harvill, Spatio-temporal Processes, in: Wiley Interdisciplinary Reviews: Computational Statistics, vol. 2, John Wiley & Sons, Inc., New York, 2010, pp. 375–382.
[28] N. Cressie, C. Wikle, Statistics for Spatio-temporal Data, John Wiley & Sons, Inc., Hoboken, New Jersey, 2011.
[29] H. Tong, Non-linear Time Series: a Dynamical System Approach, Oxford Statistical Science Series, Oxford University Press, United Kingdom, 1993.
[30] T. Subba Rao, M.M. Gabr, An Introduction to Bispectral Analysis and Bilinear Time Series Models, Springer, Berlin, 1984.
[31] V. Haggan, T. Ozaki, Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model, Biometrika 68 (1981) 189–196.
[32] H. Tong, Threshold Models in Nonlinear Time Series Analysis, Lecture Notes in Statistics, Springer, New York, 1983.
[33] D. van Dijk, Smooth Transition Models: Extension and Outlier Robust Inference, Ph.D. thesis, Erasmus University, Rötterdam, Netherlands, 1999.
[34] R. Chen, R. Tsay, Functional coefficient autoregressive models, J. Am. Stat. Assoc. (1993) 298–308.
[35] Z. Cai, J. Fan, Q. Yao, Functional coefficient regression models for nonlinear time series, J. Am. Stat. Assoc. 95 (451) (2000) 941–956.
[36] J.L. Harvill, B.K. Ray, A note on multi-step forecasting with functional coefficient autoregressive models, Int. J. Forecast. 21 (4) (2005) 717–727.
[37] J. Fan, Q. Yao, Nonlinear Time Series, Springer, New York, 2003.
[38] R. Chen, L. Liu, Functional coefficient autoregressive models: estimation and

tests of hypotheses, J. Time Ser. Analysis 22 (2) (2001) 151–173.
[39] J.L. Harvill, B.K. Ray, Functional coefficient autoregressive models for vector time series, Comput. Statistics Data Anal. 50 (2006) 3547–3566.
[40] T.J. Hastie, R.J. Tibshirani, Generalized Additive Models, 43, Chapman & Hall/CRC, Boca Raton, 1990.
[41] L. Wang, L. Yang, Spline-backfitted kernel smoothing of nonlinear additive autoregression model, Ann. Statistics 35 (2007) 2474–2503.
[42] J. Wang, L. Yang, Efficient and fast spline-backfitted kernel smoothing of additive models, Ann. Inst. Stat. Math. 61 (2009) 663–690.
[43] R. Liu, L. Yang, W.K. Härdle, Oracally Efficient Two-step Estimation of Generalized Additive Model, 649, Humboldt University, Collaborative Research Center, 2011. Tech. Rep.
[44] S. Ma, L. Yang, Spline-backfitted kernel smoothing of partially linear additive model, J. Stat. Plan. Inference 141 (1) (2011) 204–219.
[45] R. Liu, L. Yang, Spline-backfitted kernel smoothing of additive coefficient model, Econ. Theory 12 (1) (2010) 29–59.
[46] P. Whittle, On stationary processes in the plane, Biometrika 41 (1954) 434–449.
[47] R. Bivand, with contributions by Micah Altman, L. Anselin, R. Assuno, O. Berke, A. Bernat, G. Blanchet, E. Blankmeyer, M. Carvalho, B. Christensen, Y. Chun, C. Dormann, S. Dray, R. Halbersma, E. Krainski, P. Legendre, N. Lewin-Koh, H. Li, J. Ma, G. Millo, W. Mueller, H. Ono, P. Peres-Neto, G. Piras, M. Reder, M. Tiefelsdorf, D. Yu., Spdep: Spatial Dependence: Weighting Schemes, Statistics and Models, URL http://CRAN.R-project.org/package=spdep, R package version 0.5–56, 2013.
[48] O. Schabenberger, C.A. Gotway, Statistical Methods for Spatial Data Analysis, CRC Press, 2004.
[49] U.S. Forest Service, Life Zone Map of Lana'i, 2011, in: http://www.fs.fed.us/psw/topics/ecosystem_processes/tropical/restoration/lifezone/hawaii/Lanai.jpg.
[50] J. Johnson, B. Schenkman, A. Ellis, J. Quiroz, C. Lenox, Initial operating experience of the 1.2-MW La Ola photovoltaic system, in: Photovoltaic Specialists Conference (PVSC), vol. 2, 2012, pp. 1–6, 2012 IEEE 38th, IEEE.
[51] M.J. Reno, C.W. Hansen, J.S. Stein, Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis, Tech. Rep. SAND2012-2389, Sandia National Laboratories, Albuquerque, NM, 2012.
[52] E.L. Maxwell, A Quasi-physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation, Tech. Rep., Solar Energy Research Institute, Golden, CO, 1987.
[53] E. Lorenzo, L. Narvarte, J. Munoz, Tracking and back-tracking, Prog. Photovoltaics Res. Appl. 19 (6) (2011) 747–753.
[54] M. Wand, KernSmooth: Functions for Kernel Smoothing for Wand & Jones (1995), 2012. R package version 2.23–8, http://CRAN.R-project.org/package=KernSmooth.
[55] J. D. Patrick, J. L. Harvill, J. R. Sims, Spline-backfitted kernel forecasting for functional-coefficient autoregressive models, Submitted to International Journal of Forecasting. Manuscript available at http://arxiv.org/abs/1502.03486.
[56] R. Sibson, A brief description of natural neighbour interpolation, Interpreting multivariate data 21.
[57] Z. Cai, R.C. Tiwari, Application of a local linear autoregressive model to BOD time series, Environmetrics 11 (3) (2000) 341–350.