A Stochastic Model for Solar Photo-Voltaic Power for Short-Term Probabilistic Forecast

Article · June 2017

CITATIONS
4

READS
439

3 authors:

Raksha Ramakrishna
Arizona State University
16 PUBLICATIONS **32** CITATIONS
SEE PROFILE

Anna Scaglione
Arizona State University
459 PUBLICATIONS **12,791** CITATIONS
SEE PROFILE

V. Vittal
Arizona State University
373 PUBLICATIONS **19,231** CITATIONS
SEE PROFILE

A Stochastic Model for Solar Photo-Voltaic Power for Short-Term Probabilistic Forecast

Raksha Ramakrishna, *Student Member, IEEE*, Anna Scaglione, *Fellow, IEEE*, Vijay Vittal, *Fellow, IEEE*

***Abstract*—In this paper the model for solar Photo-Voltaic (PV) power previously proposed by the authors is extended to provide short term solar PV power forecasts. The proposed discrete model for solar PV power incorporates the stochasticity due to clouds by modeling the attenuation of power using three different parameters. These parameters are extracted from a dataset using dictionary learning techniques. The statistical behavior of parameters highlights how solar PV power switches between the three classes of *sunny*, *overcast* and *partly cloudy* during the day. To capture the same, different stochastic models are proposed for the three regimes. Then, an online solar power prediction algorithm is proposed by employing this switching phenomenon. This stochastic model is capable of generating probabilistic forecasts as well. Sample future scenarios of power generated with the proposed method could be used for stochastic optimization. To illustrate the technique developed, a set of solar PV power data from a single rooftop installation in California is analyzed and the effectiveness of the model in fitting the data and providing short term point forecasts is verified. The proposed prediction method outperforms two reference models considered, *persistence* and *diurnal persistence*.**

***Index Terms*—Solar PV power modeling, Short-term solar power prediction, Roof-top solar panels, Dictionary learning, Hidden Markov Models**

I. Introduction

Solar power generation, both from PV farms and roof-top solar panel installations is on the rise leading to their increasing penetration into traditional energy markets. Hence, consideration of solar PV power resource while analyzing electric grid operations is gaining great significance. Accurate models that can not only provide solar power forecasts, but also capture the uncertainty in the random process are necessary to address decision problems such as stochastic optimal power flow (SOPF) [1], probabilistic power flow studies [2], designing microgrids [3], solar power shaping [4] and reserve planning. The difficulty in modeling solar PV power output stems from the inherent stochasticity due to clouds and other atmospheric phenomena. A vast array of literature exists in the area of short-term forecasts for solar power. A majority of the approaches taken can be broadly classified as being physical, statistical or a hybrid of the two methods (see e.g. [5] for a review). Physical methods employ astronomical relationships [6], meteorological conditions and numerical weather prediction (NWPs) for an improved forecast [7], [8]. Such studies are based on modeling the clear sky radiation using earth-sun geometry, panel tilt and orientation, temperature and wind speed [9], [10]. Some also use irradiance information available from databases to determine the value of power for a geographical location considered. Other papers use static images of clouds in the sky recorded by a total sky imager (TSI) [11] or utilize a network of sensors recording cloud motion [12] to predict solar power. These models rely on a deterministic mapping given additional information to produce an estimate of the power generated by the panel.

The most prevalent statistical methods for solar power forecasting include stationary time-series modeling such as using autoregressive (AR) models [13]–[15]. One of the advantages of these methods is that they are driven by the actual electric power data, which means that they do not depend on having a lot of additional information like the previous literature cited, and are adaptive. However, these methods are born to model stationary normal processes and are generally not very well suited to fully capture the non-stationarity of solar power production. In the same class of statistical methods , there exist other works that capture variability in solar PV power [16]–[19] and black box models like using artificial neural networks (ANN) [20] and support vector machines (SVM) [21] to predict solar power.

In the proposed methodology the advantages offered by both the aforementioned approaches are exploited. The model proposed provides a statistical description of stochasticity of the electric power signal that is inspired by the physical behavior of solar PV power, while being completely adaptive.

Prior work in [22] by the authors briefly described in Section II involved the development of a parametric model that was proven to efficiently capture the effect of clouds on solar PV power while providing a compact representation.

In this paper, the prior modeling technique is utilized and extended to fit a switching process to solar PV power, using which a solar power prediction algorithm to provide short-term forecasts is designed. The key contributions of this paper are:

- The definition of an underlying switching process of solar PV power that consists of periods that can be classified as *sunny*, *overcast* and *partly cloudy* and introduction of stochastic models for the three regimes. This is detailed in section III.
- A change detection algorithm that identifies a switch in the process from the data.
- A hidden Markov model (HMM) for the *partly cloudy* regime utilizing the sparsity of parameters pertaining to attenuation of power described in subsection III-3.
- The design and analysis of an online algorithm for short term solar power prediction by employing the switching process and the relevant stochastic models for *sunny*, *overcast* and *partly cloudy* as outlined in section IV.

The authors are with the School of Electrical, Energy and Computer Engineering (ECEE), Arizona State University, Tempe, AZ 85281, USA. This work was funded in part by the Advanced Research Projects Agency- Energy (ARPA-E), U.S. Department of Energy, under Award Number de-ar0000696 and by National Science Foundation under grant number CPS-1549923 . The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

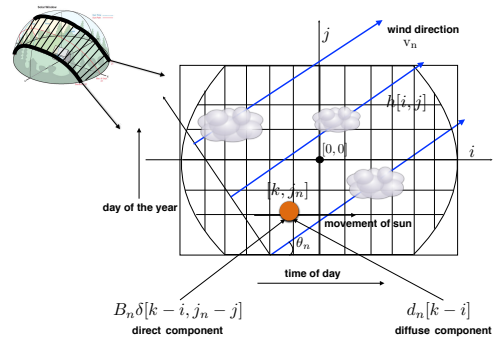


Fig. 1. Figure representing the sun's path across the sky over the days on a plane. The orange dot marks the position of the sun at time k on day n. The wind trajectory is described by blue lines.

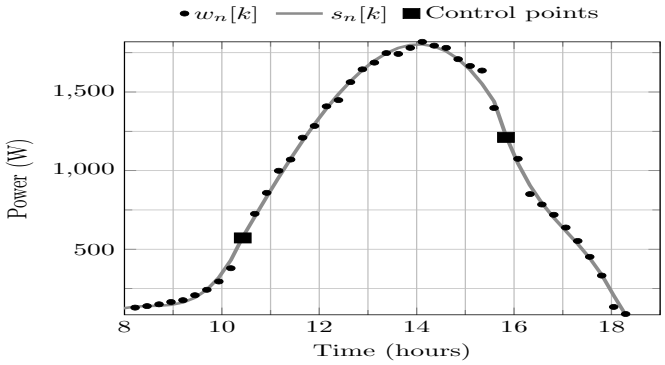


Fig. 2. Plot showing $s_n[k]$ and $w_n[k]$ for a sunny day on October 22, 2009

The prediction results using the proposed method, as seen in subsection V-B, indicate the validity of the approach. In fact, the proposed prediction also outperforms two reference models of *persistence* and *diurnal persistence* (c.f subsection V-B1). Also, one could complement the proposed method by using weather prediction and cloud imagery as additional information in order to improve the point forecasts. Since stochastic models are available, the method also has the ability to provide probabilistic forecasts which are very useful while making decisions under uncertainty. Section VI includes the conclusions and future work.

II. Discrete time Model

The following discrete time model for solar PV power output was derived in detail in the authors' prior work in [22]. It hypothesizes that the panel sums solar irradiation from the sky by weighting each contribution with a bi-dimensional gain function that handles the scaling factors to obtain total electrical power. The solar irradiation is attenuated by clouds modeled as a random mask that subtracts a percentage of the light coming from the patch of sky it covers at a certain time. The motion of the clouds over the panel can be approximated to be moving at a constant speed in a certain direction throughout the day. This assumption is reasonable considering the size of the panel relative to that of the displacement of the clouds. It is known that the solar irradiation has two major components [23], *direct beam component* and *diffuse beam component*. Each of these components is attenuated by the cloud coverage in different ways. Fig.1 summarizes the idea behind the discrete time model. If the received solar power is denoted by $w_n[k] \triangleq w_n(kT)$ on day n where T is the sampling period and $k \in (-N, N)$, the discrete time model for solar power $w_n[k]$ for day n is given by

$$w_n[k] = s_n[k] - (p_n^b[k] + p_n^d[k]) + \eta_n[k]$$

where $s_n[k] \triangleq s_n(kT)$ is the solar power if the nth day is sunny, $p_n^b[k]$ and $p_n^d[k]$ are the components pertaining to direct and diffuse beam component attenuation by the clouds respectively and $\eta_n[k]$ is Gaussian measurement noise. In the next two subsections a parametric model for the solar power output without and with cloud attenuation is provided. For the moment everything is deterministic. However, the extraction of the parameters will be used to construct and validate the stochastic model in Section III.

1) Sunny days parametrization

Physics based models give explicit expressions for $s_n[k]$ accounting for the geographical location, orientation and tilt of the panel and time [23]. Since each location can have possible variations with shading and a variety of panel orientations, these expressions are not employed. Instead, each cloudless day is modeled using a simple basis expansion model, whose expansion coefficients are periodically updated to reflect seasonal variations. Let $\mathcal{S}$ denote the set of sunny days. For the nth day $n \in \mathcal{S}$ the solar PV power samples are modeled as:

$$w_n[k] \equiv s_n[k] = \sum_{q=0}^{Q} \mathrm{s}_{nq} b_q(k) \tag{1}$$

where the choice of basis is three sets of non-overlapping cubic splines that cover three daylight periods delimited by two control points k_{n1}, k_{n2} . C^1 and C^2 continuity are not evident from the data which leads to the assumption of discontinuous parts. The control points k_{n1} and k_{n2} are chosen as time instants at which an abrupt change in the data is observed. The continuity of C^0 at the two control points is then forced. This can be done easily done by constructing the basis using $Q = 9$ i.e. 10 functions that are derived from Bernstein polynomials [24] of degree $\nu = 3$, defined as:

$$B_{j,\nu}(t) = \binom{\nu}{j} t^j (1-t)^{\nu-j} \mathrm{rect}(t), \;\; j = 0, 1, 2, 3 \tag{2}$$

where $\mathrm{rect}(t)$ denotes the rectangular function between $[0, 1)$. It is seen that the basis which would directly impose C^0 continuity at the control points is:

$$b_{\nu i+j}(k) = B_{j,\nu}(t_i), \;\; i = 0, 1, 2 \tag{3}$$

$$t_0 = (k+N)/(k_{n1}+N), \qquad -N \le k \le k_{n1} \tag{4}$$

$$t_1 = (k-k_{n1})/(k_{2n}-k_{1n}), \qquad k_{n1} \le k \le k_{n2} \tag{5}$$

$$t_2 = (k-k_{n2})/(N-k_{n2}), \qquad k_{n2} \le k \le N \tag{6}$$

Thus, for a sunny day at the most 10 parameters plus 2 control points are needed. The approximated $s_n[k]$ for one such sunny day is shown in Fig. 2. It highlights the very specific pattern obtained in October due to shading.

2) Cloudy days parametrization

By referring to the authors' previous work [22] where the expressions for $p_n^b[k]$ and $p_n^d[k]$ were derived as,

$$p_n^b[k] \approx a_n^b[k]s_n[k], \quad a_n^b[k] = \sum_{\ell \in \mathcal{B}} a_\ell \delta[k - r_\ell] \tag{7}$$

$$p_n^d[k] \approx \sum_q \tilde{h}[q] z_n[k-q] \tag{8}$$

and where $a_n^b[k]$ is the stochastic time series capturing the direct beam sudden power attenuations caused by clouds whose trajectories intersect that of the sun. The diffuse beam attenuation, instead, is modeled as the convolution of a one-dimensional filter $\tilde{h}[k]$ with a stochastic input $z_n[k]$ that represents the cloud attenuation. Furthermore, to explain the increase of power even beyond the expected sunny day power $s_n[k]$, along the lines of direct beam attenuation the following term is introduced to be present only when $w_n[k] > s_n[k]$:

$$p_n^e[k] \approx a_n^e[k]s_n[k], \quad a_n^e[k] = \sum_{\ell \in \mathcal{E}} a_\ell \delta[k - k_\ell] \tag{9}$$

This term captures the so called *edge of the cloud effect* that has been reported in literature [25], [26]. The edge of some clouds, $\ell \in \mathcal{E}$ act like a magnifying lens when their paths intersect with that of the sun thereby boosting the power. It is important to note that the edge of cloud effect cannot occur simultaneously with cloud related attenuation and in general this term will be far sparser. This observation directly ties to the formulation of the regression problem presented next.

3) Regression problem

From subsections II-1 and II-2 the complete model for power on a given day n is characterized by:

$$\mathbf{w}_n = \mathbf{s}_n - \mathbf{U}(\mathbf{S}_n^b \mathbf{a}_n^b + \mathcal{T}(\tilde{\mathbf{h}})\mathbf{z}_n) + \tilde{\mathbf{U}}\mathbf{S}_n^b \mathbf{a}_n^e \tag{10}$$

where $U(.)$ is the Heaviside step function operating element-wise, $\mathbf{U} = \text{diag}(U(\mathbf{s}_n - \mathbf{w}_n))$,$\tilde{\mathbf{U}} = \mathbb{I} - \mathbf{U}$ and $\mathbb{I}$ is the identity matrix of size $2N$. $\mathbf{w}_n(i) = w_n(i), \mathbf{s}_n(i) = s_n(i), \mathbf{a}_n^e(i) = a_n^e(i), \mathbf{a}_n^b(i) = a_n^b(i)$ are $2N$ dimensional positive real vectors, $\mathbf{S}_n^b = \text{diag}(\mathbf{s}_n) \in \mathbb{R}^{2N \times 2N}$, $\mathbf{z}_n(i) = z_n(i), \mathbf{z}_n \in \mathbb{R}_+^{(2N+M-1)\times 1}$, $\tilde{\mathbf{h}}(i) = \tilde{h}(i), \tilde{\mathbf{h}} \in \mathbb{R}_+^{M\times 1}$ and $\mathcal{T}(\tilde{\mathbf{h}}) \in \mathbb{R}_+^{2N\times(2N+M-1)}$ is the Toeplitz matrix with first column $\left[\tilde{h}[M-1], \quad \mathbf{0}^{1\times 2N-1}\right]^T$ and first row $\left[\tilde{h}[M-1], \ldots, \tilde{h}[0], \quad \mathbf{0}^{1\times 2N-1}\right]$. Here the estimation of the cloud coverage parameters is seen as a blind deconvolution problem that falls in the class of sparse dictionary learning problems [27], [28], usually solved by alternating between the estimation of the vectors $\mathbf{z}_n, \mathbf{a}_n^b, \mathbf{a}_n^e$ by sparse coding [29] and the estimation of filter $\tilde{\mathbf{h}}$ over multiple iterations. More specifically, as in a typical sparse coding problem formulation, estimates can be obtained by solving:

$$\begin{aligned} \min_{\tilde{\mathbf{h}}, \mathbf{z}_n, \mathbf{a}_n^b, \mathbf{a}_n^e} \quad & \sum_n \|\mathbf{U}\left(\mathbf{s}_n - \mathbf{w}_n - \mathbf{S}_n^b \mathbf{a}_n^b - \mathcal{T}(\tilde{\mathbf{h}})\mathbf{z}_n\right) \\ & + \tilde{\mathbf{U}}\left(\mathbf{s}_n - \mathbf{w}_n + \mathbf{S}_n^b \mathbf{a}_n^e\right)\|_2^2 \\ & + \sum_n \lambda_1(\mathbf{1}^T \mathbf{a}_n^e) + \lambda_2(\mathbf{1}^T \mathbf{a}_n^b) + \lambda_3(\mathbf{1}^T \mathbf{z}_n) \\ \text{subject to} \quad & \mathbf{a}_n^b \geq 0, \quad , \mathbf{a}_n^e \geq 0, \quad \mathbf{z}_n \geq 0 \quad \forall n, \quad \tilde{\mathbf{h}} \geq 0 \\ & \tilde{\mathbf{U}}\left(\mathbf{S}_n^b \mathbf{a}_n^b + \mathcal{T}(\tilde{\mathbf{h}})\mathbf{z}_n\right) = \mathbf{0}, \quad \mathbf{U}\mathbf{S}_n^b \mathbf{a}_n^e = \mathbf{0} \end{aligned} \tag{11}$$

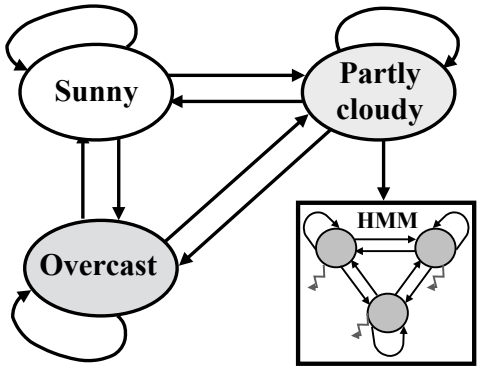


Fig. 3. Block diagram highlighting the proposed switching process between stochastic models

The algorithm is initialized with the filter being a scaled Hamming window of length M, $\tilde{h}[q] = g \times (0.54 - 0.46\cos(2\pi q/(M-1)))$. To address the scale ambiguity inherent in blind deconvolution problems, the scale g is chosen such that $\tilde{h}[q]$ and $s_n[k]$ have similar amplitudes. Regularization constants are chosen such that $\lambda_1 \geq \lambda_2 \gg \lambda_3$ in order to force $\mathbf{a}_n^b$ and $\mathbf{a}_n^e$ to be sparser than $\mathbf{z}_n$. This rationale is justified since only a subset of total number of clouds have the possibility of directly occluding the sun. It was shown in [22] that the proposed model led to an excellent fit with the data. Even though the regression problem is solved in a completely deterministic fashion, such a model allows the separation of the components and study a plausible stochastic model for them. This is explained in detail in the next section III.

III. Stochastic models for classification and forecast of solar power data

In spite of the fact that the switching nature is not intrinsically part of the model discussed above, as reported by the authors in [22], the results of the deterministic fit after solving (11) highlighted the switching nature of the solar irradiation phenomenon. Solar PV power produced in a period of time can be broadly classified as coming from *sunny, overcast* or *partly cloudy* models. The model switches between the three classes as shown in Fig. 3 due to weather changes. In this section, a stochastic model for each of the three classes is proposed. The first application of this model is for *change detection*, i.e. to identify the switch between classes to provide a forecast by assuming that the model persists. The second application is for probabilistic short-term forecast.

1) Stochastic model for sunny period

For sunny periods, it is hypothesized that the solar power is the deterministic solar power pattern i.e.,

$$w_n[k] = s_n[k] + \eta_n[k] \tag{12}$$

The modeling error is given by $\eta_n[k] \sim \mathcal{N}(0, \sigma^2)\ \forall k$.

2) Stochastic model for overcast period

During overcast periods, the attenuation of solar power is mostly from the diffuse beam component [23] which is why $z_n[k]$ accounts for the relevant attenuation. Also, there is an average component in the overcast days for $z_n[k]$ that mimics a scaled version of sunny day pattern $s_n[k]$. Since this

attenuation is smooth, the model for overcast period is:

$$w_n[k] \approx \alpha_n s_n[k] + \eta_n[k], \tag{13}$$

where α_n can be thought of as the attenuation of sunny day power. The parameter α_n is analogous to *clear sky index* defined as $w_n[k]/s_n[k]$ that many papers use to model solar PV power [30]. However, all the samples in the overcast period are used to estimate α_n unlike the determination of clear sky index. This leads to robustness with respect to noise.

3) Stochastic model for partly cloudy period

The model for the partly cloudy period is slightly more involved, due to the presence of all the three parameters. However, a hidden Markov model (HMM) is able to capture the underlying on-off process that characterizes the sparse parameters in periods with fast moving clouds that cause sharp fluctuations in solar PV power.

The observed solar PV power data $w_n[k]$ is modeled as coming from underlying hidden states that are Markovian in nature. Let the state/latent variable in this model, $\mathbf{q}_k$ be the support of the unknown sparse parameters, $(z_n[k], a_n^b[k], a_n^e[k])$. Their relationship is governed by the following equations:

$$w_n[k] = s_n[k] - \mathbf{P}\ \text{diag}\left(\mathbf{\Phi}\mathbf{q}_k\right)\mathbf{x}_k \tag{14}$$

$$\mathbf{q}_{k+1} = \mathbf{A}^T\mathbf{q}_k + \boldsymbol{\nu}_{k+1} \tag{15}$$

$$\text{where } \mathbf{P} = \begin{bmatrix}\tilde{h}[M-1] & \dots & \tilde{h}[0] & s_n[k] & -s_n[k]\end{bmatrix},$$

$$\mathbf{x}_k = \begin{bmatrix}z_n[k-M+1] & \dots & z_n[k] & a_n^b[k] & a_n^e[k]\end{bmatrix}^{\mathbf{T}}$$

Let the total number of states be represented by $\mathcal{N}_s$. Then, $\mathbf{A} \in \mathbb{R}^{(\mathcal{N}_s \times \mathcal{N}_s)}$ is the state transition matrix where $\mathbf{A}(i,j)$ is the probability of going from state i to state j and $\boldsymbol{\nu}_{k+1}$ is the noise. The state vector $\mathbf{q}_k \in \mathbb{R}^{(M+2)\times 1}$ is a binary vector taking values from the set of coordinate vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{\mathcal{N}_s}\}$ where $\mathbf{e}_i \in \mathcal{R}^{\mathcal{N}_s}$ has a 1 at position i and zero elsewhere. The matrix $\mathbf{\Phi} \in \mathbb{R}^{(M+2)\times \mathcal{N}_s}$ contains the possible combinations of presence and absence of the coefficients in $\mathbf{x}_k$ where each combination corresponds to one state. Certain assumptions are made to decrease the number of states. Firstly, $\begin{bmatrix}z_n[k-M+1] & \dots & z_n[k]\end{bmatrix}$ is restricted to have $\ell < M$ non-zero entries. Secondly, $a_n^e[k]$ cannot co-exist with the other parameters due to the fact that edge of cloud effect is indicative of the absence of attenuation. Furthermore, as a simplification, it is also assumed that direct beam and diffuse beam attenuations do not occur together which means that the total number of states is $\mathcal{N}_s = \binom{M}{\ell} + 3$ Notice the absence of noise term in the observation equation (14). This stems from the fact that measurement noise is not included since the 'noisy' nature of the solar power data is caused by the fast movement of clouds rather than by erroneous measurements.

The simplest case of choosing $\ell = 1$ and having $\mathcal{N}_s = M+3$ states is considered. All non-zero parameters in $\mathbf{x}_k$ are hypothesized to come from independent exponential distributions. While in state i a certain $w_n[k]$ is observed:

$$w_n[k] = \begin{cases} s_n[k],\ i = 1 \\ s_n[k] - \tilde{h}[i-2]z_n[k-i+2],\ i = 2, \dots, M+1, \\ s_n[k] - s_n[k]a_n^b[k],\ i = M+2 \\ s_n[k] + s_n[k]a_n^e[k],\ i = \mathcal{N}_s \end{cases}$$

The corresponding conditional probability distribution given the state i is denoted as $\tilde{f}_i(w_n[k]) \triangleq f_{\underline{w}_n[k]}(w_n[k]|\mathbf{q}_k = \mathbf{e}_i)$ and is equal to,

$$\tilde{f}_i(w_n[k]) = \begin{cases} \delta(s_n[k] - w_n[k]),\ i = 1 \\ \frac{C_i\lambda_z}{\tilde{h}[i-2]} \exp\left\{-\frac{\lambda_z(s_n[k]-w_n[k])}{\tilde{h}[i-2]}\right\} i = 2, \dots, M+1 \\ \frac{C_i\lambda_a^b}{s_n[k]} \exp\left\{-\frac{\lambda_a^b}{s_n[k]}\left(s_n[k] - w_n[k]\right)\right\}, i = M+2 \\ \frac{\lambda_a^e}{s_n[k]} \exp\left\{-\frac{\lambda_a^e}{s_n[k]}\left(w_n[k] - s_n[k]\right)\right\}, i = \mathcal{N}_s \end{cases}$$

where C_i is the normalizing constant for the probability distribution given by

$$C_i^{-1} = \begin{cases} 1 - \exp\left\{-\lambda_z s_n[k]/h[i-2]\right\},\ i = 2, 3, \dots M+1 \\ 1 - \exp\left(-\lambda_a^b\right),\ i = M+2 \end{cases}. \tag{16}$$

The normalization is done so that $w_n[k] \in [0, s_n[k]]$.

A. Learning the parameters of HMM for partly cloudy periods

The models for sunny and the overcast periods are such that the only thing that can be predicted is the mean of the process in both cases, but not the noise $\eta_n[k]$ which by construction is assumed to be i.i.d. during the corresponding period. Hence, the problem of learning the stochastic parameters of the model to perform predictions is non-trivial only during partly cloudy periods. To do so, it is assumed that the values of the parameters $\lambda_z, \lambda_a^b, \lambda_a^e$ of conditional probability distributions are known. It was seen that the algorithm is not very sensitive to the exact values of these parameters as long as they follow $\lambda_z \leq \lambda_{a^b} \leq \lambda_a^e$ which is consistent with the results of the regression problem. The probability of starting from a state i denoted by $\pi_i = 1/\mathcal{N}_s$ is also assumed to be known. In order to learn the the state transition matrix $\mathbf{A}$, Viterbi training [31] or segmental k-means [32] approach was adopted. Let $\xi = \mathbf{A}(i,j)$ be the set of unknown parameters to be estimated. Let $\tilde{N}$ be the number of samples in a certain block of solar PV power data, sequence $\mathbf{Q} = \mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{\tilde{N}}$ and $\mathbf{W} = w_n[1], w_n[2], \dots, w_n[\tilde{N}]$ denote the sequence of solar power observations. In the Viterbi training algorithm, instead of maximizing the likelihood over all possible state sequences $\bar{\mathcal{Q}}$, the likelihood is maximized only over the most probable state sequence to find the estimates of parameters in ξ. The algorithm starts with an initial estimate for all the unknown parameters $\xi_0 = \mathbf{A}_0$ and performs this maximization iteratively [32],

$$\hat{\xi}_m = \arg\max_{\xi}\left(\max_{\mathbf{Q}} f(\mathbf{W}, \mathbf{Q}|\hat{\xi}_{m-1})\right) \tag{17}$$

where m is the iteration number and

$$f(\mathbf{W}, \mathbf{Q}|\xi) = p(\mathbf{q}_1)\prod_{k=1}^{\tilde{N}} p(w_n[k]|\mathbf{q}_k, \xi)\prod_{k=1}^{\tilde{N}-1} p(\mathbf{q}_{k+1}|\mathbf{q}_k, \xi)$$

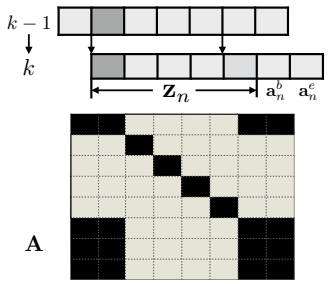


Fig. 4. The specific way in which state transition from time instant $k-1$ to k takes places determines the structure of the state transition matrix $\mathbf{A}$

The inner maximization is performed by using a dynamic programming algorithm known as Viterbi algorithm [33] which is a recursive method. As a result of this maximization,

$$\hat{\mathbf{Q}}_m = \arg\max_{\mathbf{Q}} f(\mathbf{W}, \mathbf{Q}|\hat{\xi}_{m-1}) = \hat{\mathbf{q}}_1^m, \hat{\mathbf{q}}_2^m, \ldots, \hat{\mathbf{q}}_{\tilde{N}}^m,$$

the most likely state sequence at iteration m which best describes the observed data. Later, maximum-likelihood (ML) estimates $\hat{\xi}_m$ are estimated,

$$\hat{\xi}_m = \arg\max_{\xi} \left(f(\mathbf{W}, \hat{\mathbf{Q}}_m|\hat{\xi}_{m-1}) \right) \tag{18}$$

Maximizing $\log f(\mathbf{W}, \hat{\mathbf{Q}}_m|\hat{\xi}_{m-1})$ with respect to $\mathbf{A}_m(i,j)$ under the constraint that $\mathbf{A}_m$ is stochastic since it is the state transition matrix i.e., $\sum_{j=1}^{\mathcal{N}_s} \mathbf{A}_m(i,j) = 1$, gives

$$\hat{\mathbf{A}}_m(i,j) = \frac{N_{ij}}{\sum_{j=1}^{\mathcal{N}_s} N_{ij}} \tag{19}$$

where N_{ij} is the number of times the transition from state i to state j occurs within the state sequence $\hat{\mathbf{Q}}_m$. Following from III-3 wherein the number of active coefficients at time k in $\mathbf{x}_k$ is restricted to 1, only a limited number of transitions from state i are possible and not to all $\mathcal{N}_s$ states. Also, since z_n is the input to a filter with memory M, it means that $M-1$ components need to be retained and shifted while a new one comes in. All of the above reasons give the state transition matrix $\mathbf{A}$ a sparse and specific structure as shown in Fig.4 which is forced on $\mathbf{A}_0$ during the initialization . As a result, only 16 entries of the matrix need to be estimated instead of $(\mathcal{N}_s)^2$ irrespective of the size of the filter M.

IV. Change Detection and Solar power prediction

The premise for prediction is the persistence in the weather condition for the time horizon over which a prediction of solar power is provided. Therefore, the proposed prediction algorithm has two steps:

- Classification of the solar power from a given period as coming from one of the three classes of models: *sunny, overcast, partly cloudy*
- Assuming that this weather condition persists for the duration of the prediction horizon and provide with a point forecast corresponding to the class decided in the classification step.

Such a scheme captures the inherent switching behavior that solar power exhibits i.e. that of going from one model to another while persisting for a certain duration in each of these. Note that the classification step can be skipped if prior knowledge in the form of weather prediction is available. The prediction algorithm utilizes a rolling horizon wherein prediction is improved as more data comes in.

A. Classification algorithm for solar power

The classification algorithm uses the stochastic models for the solar power data as detailed in III. Let $w_n[k],\ k \in (\kappa_1, \kappa_2)$ be the solar power samples that have to be classified. It is easy to decide in favor of *sunny* model by computing the error, $\sum(w_n[k] - s_n[k])^2$. If it is less than some power threshold $\tilde{p}$, then it is classified as a sunny period. If that is not the case, the hypotheses *overcast* ($\mathcal{H}_0$) or *partly cloudy* ($\mathcal{H}_1$) are tested.

$$\mathcal{H}_0 : w_n[k] = \alpha_n^\kappa s_n[k] + \eta_n[k],\ k \in (\kappa_1, \kappa_2)$$
$$\mathcal{H}_1 : w_n[k] = s_n[k] - \mathbf{P}\ \text{diag}\left(\mathbf{\Phi}\mathbf{q}_k\right)\mathbf{x}_k$$
$$\mathbf{q}_{k+1} = \hat{\mathbf{A}}^T \mathbf{q}_k + \boldsymbol{\nu}_k, \quad k \in (\kappa_1, \kappa_2)$$

Let $\mathbf{W}_\kappa = w_n[\kappa_1], \ldots, w_n[\kappa_2],$ and $\mathbf{Q}_\kappa = \mathbf{q}_{\kappa_1}, \ldots, \mathbf{q}_{\kappa_2}$. It is a composite hypothesis testing problem since α_n^κ is unknown. The maximum likelihood estimate of $\hat{\alpha}_n^\kappa$ is,

$$\hat{\alpha}_n^\kappa = \frac{\sum_{k=\kappa_1}^{\kappa_2} \left(w_n[k] s_n[k]\right)^2}{\sum_{k=\kappa_1}^{\kappa_2} \left(s_n[k]\right)^2}. \tag{20}$$

Generalized likelihood ratio is not computed. Instead the error,

$$\sum_{k=\kappa_1}^{\kappa_2} (\eta_n[k])^2 = \sum_{k=\kappa_1}^{\kappa_2} (w_n[k] - \hat{\alpha}_n^\kappa s_n[k])^2 \tag{21}$$

is compared with a predefined threshold and also the value of α_n is compared with a heuristically set threshold. These rules decide the classification of data as *overcast* model or *partly cloudy*. If the decision is in favor of *partly cloudy*, the most likely state sequence $\mathbf{Q}_\kappa$ that generated the power observations $\mathbf{W}_\kappa$ is determined using the Viterbi algorithm with state transition matrix $\hat{\mathbf{A}}$.

B. Prediction for each class of model

Based on the classification results on $w_n[k],\ k \in (\kappa_1, \kappa_2)$, a solar power forecast, $\hat{w}_n[k], k \in (\kappa_2+1, \kappa_2+\chi)$ is provided. Here, χ is the length of the prediction horizon.

1) Prediction using sunny model

When the detection algorithm chooses the hypothesis that the current solar power data is from a *sunny* model, then:

$$\hat{w}_n[k] = s_n[k],\ \forall k \in \{\kappa_2 + 1, \kappa_2 + 2, \ldots, \kappa_2 + \chi\} \tag{22}$$

Note that the deterministic sequence of the sunny day solar power pattern is known beforehand, and it is updated at a very slow pace on days that are classified as being sunny, to adjust for seasonal variations.

2) Prediction using overcast model

When the test on $\sum_{k=\kappa_1}^{\kappa_2}(\eta_n[k])^2$ and $\hat{\alpha}_n$ decides that hypothesis $\mathcal{H}_0$ is true in the duration $k \in (\kappa_2 \kappa_1)$, then:

$$\hat{w}_n[k] = \hat{\alpha}_n^\kappa s_n[k],\ \forall k \in \{\kappa_2 + 1, \kappa_2 + 2, \ldots, \kappa_2 + \chi\}$$

where $\hat{\alpha}_n$ is estimated using (20).

3) Prediction using partly cloudy model

Since solar PV power on a partly cloudy day has an underlying Markov Model, the estimated state transition matrix $\hat{\mathbf{A}}$ is used to determine the most likely future state sequence:

$\mathbf{Q}_{pred} \triangleq \hat{\mathbf{q}}_{\kappa_2+1}, \ldots \hat{\mathbf{q}}_{\kappa_2+\chi}$ as

$$\mathbf{Q}_{pred} = \max_{\mathbf{Q}} \left(p(\mathbf{q}_1) \prod_{k=\kappa_2+1}^{\kappa_2+\chi-1} p(\mathbf{q}_{k+1}|\mathbf{q}_k, \xi) \right) \tag{23}$$

by using a modified Viterbi algorithm: Define

$$\zeta_k(i) = \max_{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{k-1}} (p(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k = \mathbf{e}_i))$$

Then,

$$\zeta_{k+1}(j) = \max_i \zeta_k(i) a_{ij}$$

Let $\tilde{j}$ is the last seen state before prediction started i.e. $\mathbf{q}_{\kappa_2} = \mathbf{e}_{\tilde{j}}$. The recursion is:

$$\zeta_k(j) = \max_{1 \le i \le \mathcal{N}_s} \zeta_{k-1}(i) a_{ij}, \quad \psi_k(j) = \arg\max_{1 \le i \le \mathcal{N}_s} \zeta_{k-1}(i) a_{ij}$$

with the initialization:

$$\zeta_1(i) = 1 \ \forall i = 1, 2, \ldots, \mathcal{N}_s, \quad \psi_1(i) = \tilde{j}$$

and termination at:

$$j^*_{\kappa_2+\chi} = \arg\max_{1 \le i \le \mathcal{N}_s} \zeta_{\kappa_2+\chi}(i).$$

At this point the state sequence backtracking is:

$$j^*_k = \psi_{k+1}(j^*_{k+1}), \ k = \kappa_2 + 1, \kappa_2 + 2, \ldots, \kappa_2 + \chi - 1$$

After $\mathbf{Q}_{pred}$ is determined, vector $\hat{\mathbf{x}}_k$ is created to generate a point prediction. An estimate of vector $\hat{\mathbf{x}}_k$ is given by

$$\hat{\mathbf{x}}_k = \begin{bmatrix} \hat{z}_n & \ldots & \hat{z}_n & \hat{a}^b_n & \hat{a}^e_n \end{bmatrix}^{\mathbf{T}} \text{ where} \tag{24}$$

$$\hat{z}_n = \arg\min_{z_n} \sum_{k \in \mathcal{L}, i \in \mathcal{B}} \left(w_n[k] - s_n[k] + \tilde{h}[i-1] z_n \right)^2,$$

$$\text{subject to } z_n \ge 0$$

$$\mathcal{L} = \{k \mid \mathbf{q}_k = \mathbf{e}_{i=2,\ldots,M+1}, k \in (\kappa_1, \kappa_2)\},$$

$$\mathcal{B} = \{i \mid \mathbf{q}_k(i+1) = 1, k \in (\kappa_1, \kappa_2), 1 < i \le M+1\}$$

$$\hat{a}^b_n = \begin{cases} 1 - \hat{\alpha}^\kappa_n, & \hat{\alpha}^\kappa_n < 1 \\ 1/\lambda^b_a, & \text{otherwise} \end{cases}, \quad \hat{a}^e_n = \begin{cases} \hat{\alpha}^\kappa_n - 1, & \hat{\alpha}^\kappa_n > 1 \\ 1/\lambda^e_a, & \text{otherwise} \end{cases}$$

Then, the prediction of solar power is given by,

$$\hat{w}_n[k] = s_n[k] - \mathbf{P} \ \text{diag}\left(\mathbf{\Phi} \hat{\mathbf{q}}_k\right) \hat{\mathbf{x}}_k, \quad k \in (\kappa_2 + 1, \kappa_2 + \chi).$$

The estimate $\hat{z}_n$ is obtained from present power measurements which are emissions of the hidden states $i = 2, \ldots, M+1$. This is akin to estimating the size and intensity of one single cloud that is responsible for the diffuse beam attenuation in the time frame considered. This is retained in the prediction as it is equivalent to the propagation of this single cloud which is responsible for the future attenuation in power.

The estimates of $\hat{a}^b_n$ and $\hat{a}^e_n$ are more heuristic however. This is due to the fact that these parameters are responsible for the sudden and sharp transition in the value of power and it is very difficult to predict them. Therefore, the values of $\hat{a}^b_n$ and $\hat{a}^e_n$ are adjusted in a way so that, $\hat{w}_n[k] = \hat{\alpha}^\kappa_n s_n[k]$ when $\mathbf{q}_k = \mathbf{e}_{i=M+1,M+2}$. However, whenever $\hat{\alpha}^\kappa_n < 1$ when state $i = M+2$, the parameters $\hat{a}^b_n$ and $\hat{a}^e_n$ are replaced with their mean values.

The prediction algorithm is summarized in Fig.5.

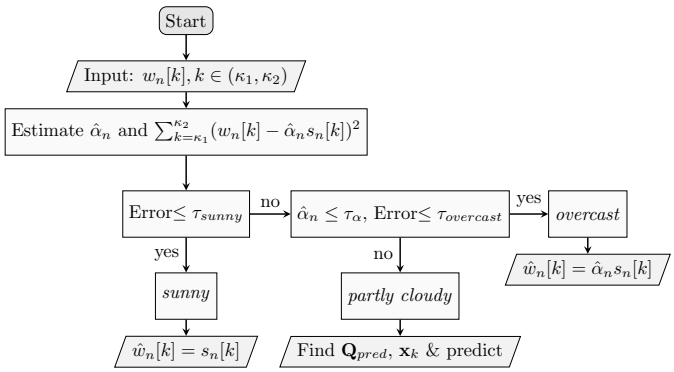


Fig. 5. Flowchart of the solar power prediction algorithm

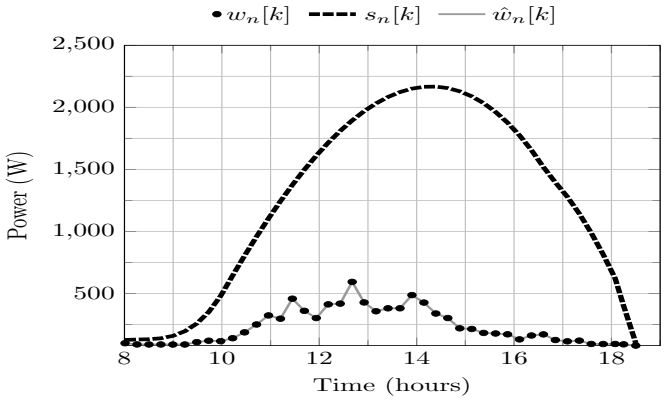


Fig. 6. Fit of the model to an *overcast day*

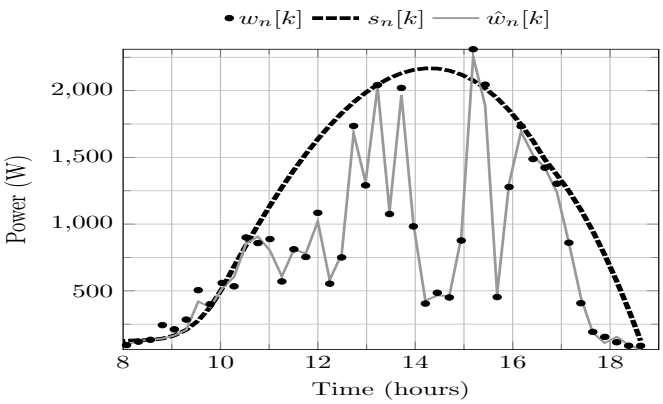
V. Numerical Results

A. Description of the dataset

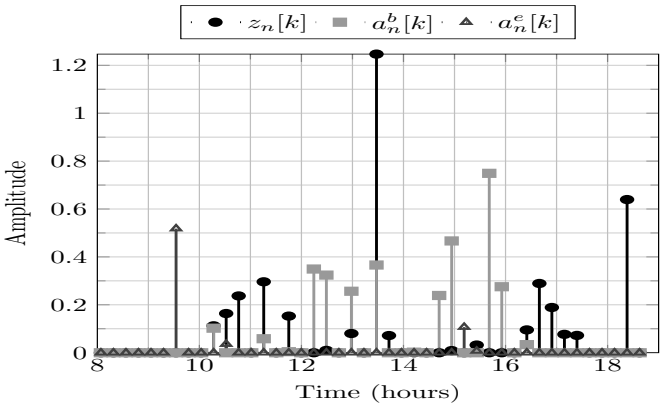
The dataset used for this work was from a rooftop panel installation in California and was provided by SolarCity. This dataset was also used in authors' prior work in [22]. The format of this solar power data consisted of current (in A), voltage measurements (in V) and timestamps (in Hours) at the inverter approximately every 15 minutes recorded for a duration of two years. Each panel had a rating of 170 W and there were a total of 22 panels. Therefore, the nameplate rating of all panels combined was $170 \times 22 = 3740$ W. The length of the filter $\tilde{h}[q]$ used for fitting was $M = 8$ which translates to having 2 hours of memory. Normalized mean square error (NMSE) is used as the error metric where,

$$\text{NMSE}_n = \frac{\sum_k (w_n[k] - \hat{w}_n[k])^2}{\sum_k (w_n[k])^2}$$

As reported in [22], the maximum NMSE is approximately 0.05 which proves the good fit the model provides. As also seen in [22], the results for fitting highlight the switching of the solar irradiation between the three classes of *sunny, overcast and partly cloudy*. Fig.6 reflects the fitting for an overcast day with potentially persistent clouds that cause the power to look like a scaled version of the sunny day pattern. This is incorporated in the proposed model for the *overcast* period. Fig.7a highlights the fit of the model to a *partly cloudy* day. The parameters of $\mathbf{a}^b_n$ and $\mathbf{a}^e_n$ are larger and less sparse on



(a) Fit to a *partly cloudy* day with sharp power fluctuations.



(b) Parameters for a *partly cloudy* day with sharp power fluctuations

Fig. 7. Fitting a *partly cloudy* day

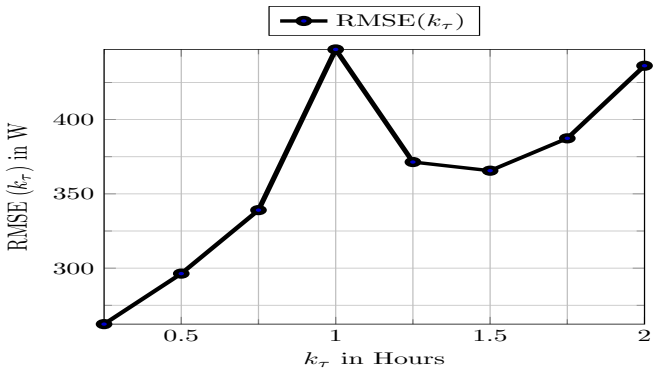


Fig. 8. RMSE for k_τ-step prediction using proposed model

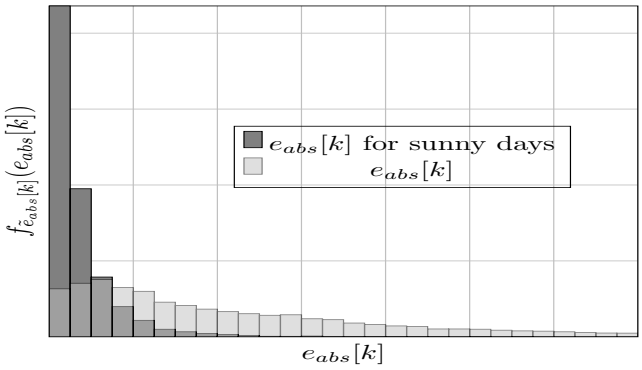


Fig. 9. Distribution of average prediction error

such days, as shown in Fig.7b. A natural result of inducing sparsity and non-negativity forced exponential distribution on the three parameters. Therefore, the *partly cloudy* period was appropriately modeled as a HMM to capture the on-off process that characterizes these sparse components.

B. Solar power prediction

In this subsection, the results of the prediction algorithm are provided. Prediction horizon was $\chi = 8$ i.e. 2 hours and filter length, $M = 5$ was used during *partly cloudy* conditions. The algorithm started with $\chi_{window} = 4$ samples (1 hour) for each day and predicted for the next χ samples. Then, the window was moved by one sample. Results are provided using the metrics of average absolute error and root mean squared error (RMSE) for the k_τ-step prediction ,

$$e_{abs}[k] = \frac{1}{\chi}\sum_{k_\tau=1}^{\chi} |w_n[k] - \hat{w}_n^{k_\tau}[k]|$$

$$\text{RMSE}(k_\tau) = \sqrt{\frac{\sum_k (w_n[k] - \hat{w}_n^{k_\tau}[k])}{\sum_k 1}}$$

where $\hat{w}_n^{k_\tau}[k]$ refers to the prediction at time k given $w_n[k - k_\tau]$ and values before it. Fig. 9 shows the distribution of average absolute error, $e_{abs}[k]$ for sunny days and other days separately. The mean absolute error is 172.7066 W which is approximately 4.6% of the nameplate rating of the panels. Fig. 8 plots the RMSE for different prediction horizons.

Fig. 10, 11 and 12 show the actual and predicted power for different days with a multitude of weather conditions. This predicted power is one-step prediction, $k_\tau = 1$. For days that are entirely *overcast* or *sunny*, predictions have little error as can be seen in fig. 10 and 11. These results highlight that the stochasticity of power in both these regimes is minimal leading to better predictions if the weather condition persists. However, there is higher error whenever there is a change in regime, for example going from *partly cloudy* condition to *overcast* around 12 PM as seen in Fig. 12. It can be attributed to the delay in detecting the change in model. This uncertainty cannot be avoided in days with sudden change in weather unless there is some additional information in the form of cloud motion information or accurate weather forecasts.

1) Comparison with reference models

In [13], the paper used *reference* models to evaluate the performance of the online short term prediction method without utilizing external inputs like numerical weather predictions (NWP), temperature or wind speed. Since the proposed method also does not use external inputs, the comparison is made with respect to two reference models that are naive predictors, *persistence* and *diurnal persistence*. In *persistence*, the prediction for the next k steps is the power value at this moment,

$$\hat{w}_n^{per}[k] = w_n[\kappa_2],\ k = \kappa_2 + 1, \ldots, \kappa_2 + \chi \qquad (25)$$

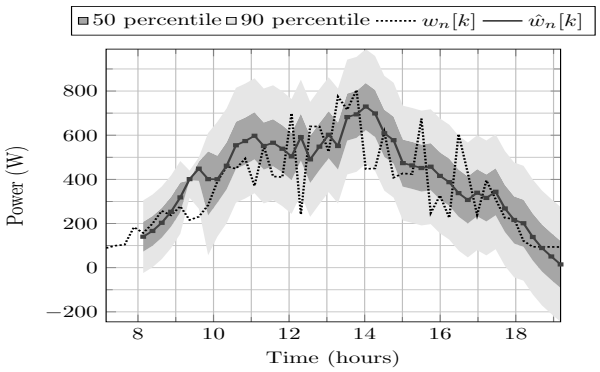


Fig. 10. Plot of actual and predicted value with one-step prediction for day that is *overcast*

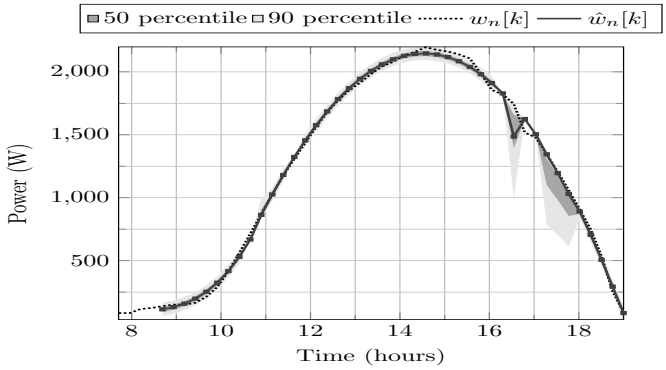


Fig. 11. Plot of actual and predicted value with one-step prediction for day that is *sunny* day

In *diurnal persistence*, the prediction for a time k is the power value at the same time on the previous day if available,

$$\hat{w}_n^{diurnal}[k] = w_{n-1}[k],\ k = \kappa_2 + 1, \ldots, \kappa_2 + \chi \qquad (26)$$

To evaluate the efficacy of the prediction algorithm, the improvement with respect to the naive reference defined was computed using the following metric,

$$\mathrm{I_{EC}}(k_\tau) = \frac{100 \times (\mathrm{EC}_{Ref} - \mathrm{EC})}{\mathrm{EC}_{Ref}} \qquad (27)$$

where EC is the evaluation criterion which is RMSE(k_τ) for different methods.

The improvement I_{EC} with respect to two naive predictors is plotted in Fig.13. As seen, the proposed method outperforms *persistence* by a large margin. However, the improvement with respect to *diurnal persistence* is not as pronounced due to the fact that there exist contiguous days with similar weather in a large proportion in the dataset making *diurnal persistence* a reasonable prediction model. It is important to note that the proposed method is better suited for shorter horizons i.e. less than 2 hours since persistence in weather condition is assumed. In the situation that no weather forecasts or other additional information is used, the performance of the proposed prediction algorithm is good. More importantly, since the proposed model is stochastic, it is possible to provide probabilistic forecasts of power over the desired horizon. Sample future power scenarios can be produced by considering all three stochastic models to be probable in the future. These scenarios are quite useful while solving stochastic optimization

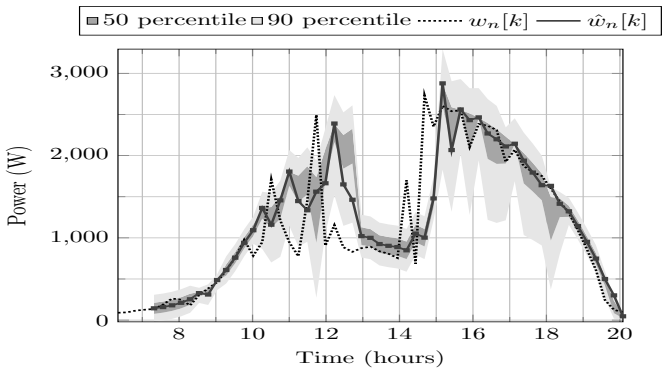


Fig. 12. Plot of actual and predicted value with one-step prediction for day with variety of weather conditions

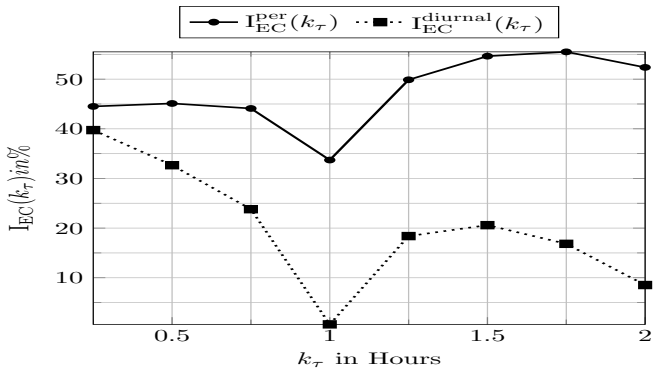


Fig. 13. Prediction improvement using the proposed model versus various *Reference* models for k_τ-step prediction

problems such as designing a battery storage policy [34]. This is future work.

VI. Conclusions

Stochastic models for different periods of *sunny, overcast* and *partly cloudy* were proposed in this paper along with online algorithm for short term probabilistic forecasts. The prediction algorithm was shown to compare favorably with two reference models (*persistence* and *diurnal persistence*). Future work includes accounting for the spatial correlation in solar power at multiple locations through low order models.

References

[1] M. Bazrafshan and N. Gatsis, "Decentralized Stochastic Optimal Power Flow in Radial Networks With Distributed Generation," *IEEE Transcations on Smart Grid*, 2016.

[2] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic Power Flow Studies for Transmission Systems With Photovoltaic Generation Using Cumulants," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2251–2261, November 2012.

[3] T. Schittekatte, M. Stadler, G. Cardoso, S. Mashayekh, and N. Sankar, "The impact of short-term stochastic variability in solar irradiance on optimal microgrid design," *IEEE Transcations on Smart Grid*, 2016.

[4] Y. Ghiassi-Farrokhfal, S. Keshav, C. Rosenberg, and F. Ciucu, "Solar power shaping: An analytical approach," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 1, pp. 162–170, January 2015.

[5] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin, and Z. Hu, "Photovoltaic and Solar Power Forecasting for Smart Grid Energy Management," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 4, pp. 38–46, December 2015.

[6] Hoyt. C. Hottel, " A Simple Model for Estimating the Transmittance of Direct Solar Radiation Through Clear Atmospheres," *Solar Energy*, vol. 18, pp. 129–134, 1976.

[7] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, August 2016.

[8] E. Lorenz, J. Hurka, D. Heinemann, and H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 2–10, March 2009.

[9] D. Larson, L. Nonnenmacher, and C. F. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest," *Renewable Energy*, vol. 91, pp. 11–20, January 2016.

[10] T. E. Hoff and R. Perez, "Quantifying PV power Output Variability," *Solar Energy*, 2010.

[11] C. W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, and B. Washom, "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed," *Solar Energy*, vol. 85, 2011.

[12] J. Bosch and J. Kleissl, "Cloud motion vectors from a network of ground sensors in a solar power plant," *Solar Energy*, vol. 95, pp. 13–20, May 2013.

[13] P. Bacher, H. Madsen, and H. A. Nielsen, "Online short-term solar power forecasting," *Solar Energy*, vol. 83, March 2009.

[14] E. B. Iversena, J. M. Moralesa, J. K. Møllera, and H. Madsen, "Probabilistic Forecasts of Solar Irradiance by Stochastic Differential Equations," *Environmetrics*, 2014.

[15] J. Boland, M. Korolkiewicz, M. Agrawal, and J. Huang, "Forecasting solar radiation on short time scales using a coupled autoregressive and dynamical system (CARDS) model," in *50th Annual Conference Australian Solar Energy Society*, December 2012.

[16] M. D. Tabone and D. S. Callaway, "Modeling Variability and Uncertainty of Photovoltaic Generation: A Hidden State Spatial Statistical Approach," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2965–2973, Nov 2015.

[17] M. Lave, J. Kleissl, and J. S. Stein, "A Wavelet-Based Variability Model (WVM) for Solar PV Power Plants," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 2, April 2013.

[18] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Preprocessing Uncertain Photovoltaic Data," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 1, pp. 351–352, January 2014.

[19] P. Shamsi, M. Marsousi, H. Xie, and W. Fries, "Dictionary learning for short-term prediction of solar pv production," *IEEE Power and Energy Society General Meeting*, July 2015.

[20] S. Bhardwaj and et.al, "Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model," *Solar Energy*, pp. 43–54, 2013.

[21] H. S. Jang, K. Y. Bae, H.-S. Park, and D. K. Sung, "Solar Power Prediction Based on Satellite Images and Support Vector Machine," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, July 2016.

[22] R. Ramakrishna and A. Scaglione, "A Compressive Sensing Framework for the analysis of Solar Photo-Voltaic Power," in *Conference Record of the Fiftieth Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 308–312.

[23] Gilbert. M. Masters, *Renewable and Efficient Electric Power Systems*. Wiley, 2004.

[24] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. Dover Publications, 1965.

[25] A. Kankiewicz, M. Sengupta, and D. Moon, "Observed impacts of transient clouds on utility-scale PV fields," in *ASES National Solar Conference*, 2010.

[26] A. E. Curtright and J. Apt, "The Character of Power Output from Utility-Scale Photovoltaic Systems," *Progress in Photovoltaics: Research and Applications*, vol. 16, pp. 241–247, September 2007.

[27] Ivana Tosic and Pascal Frossard, "Dictionary learning," *IEEE Signal Processing Magazine*, pp. 27–38, March 2011.

[28] B. Mailhé, S. Lesage, R. Gribonval, F. Bimbot, and P. Vandergheynst, "Shift- invariant dictionary learning for sparse representations: Extending K-SVD," in *European Signal Processing Conference*, vol. 4, 2008.

[29] Joel A. Tropp and Stephen J.Wright, "Computational Methods for Sparse Solution of Linear Inverse Problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.

[30] R. H. Inman, H. T. Pedro, and C. F. Coimbra, "Solar forecasting methods for renewable energy integration," *Progress in Energy and Combustion Science*, 2013.

[31] F. Jelinek, "Continuous Speech Recognition by Statistical Methods," *Proceedings of the IEEE*, vol. 64, no. 4, pp. 532–556, April 1976.

[32] B. H. Juang and L. Rabiner, "The Segmental K-Means Algorithm for Estimating Parameters of Hidden Markov Models," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 9, pp. 1639–1641, September 1990.

[33] G.D.Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, March 1973.

[34] W. B. Powell and S. Meisel, "Tutorial on Stochastic Optimization in Energy—Part I: Modeling and Policies," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1459–1467, March 2016.