



Applied Mathematics and Nonlinear Sciences 5(2) (2020) 531-542



Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

An Bayesian Learning and Nonlinear Regression Model for Photovoltaic Power Output Forecasting

Wengen Gao[†], Qigong Chen

Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Ministry of Education,

AnHui Polytechnic University China

Submission Info

Communicated by Juan Luis García Guirao Received February 24th 2020 Accepted May 26th 2020 Available online September 15th 2020

Abstract

Photovoltaic power system is taking a significant percentage of power system and the demands for accurate forecasting of the power outputs is surging. In prior works, the forecasting problem was formulated as a regression problem, however, which most cannot guarantee that the forecasted outputs is nonnegative. To solve this problem, we proposed a novel probabilistic model by using nonlinear regression and Bayesian learning method. In the paper, we present the detailed theoretical derivations and interpretations. The simulation results show the validity and feasibility of the proposed algorithm by comparing with the traditional SVM algorithm.

Keywords: Bayesian Learning, Power Output Forecasting, Data-based Regression, Probabilistic Model **AMS 2010 codes:** 60H30,93C41.

1 Background

Renewable energy sources are of more and more significant importance in the current and future power supply systems [1, 2], especially, the photovoltaic(PV) power techniques has achieved tremendous progress in the industry and research fields. In the past years, the total cumulative solar PV power capacity has reached 178GW [3, 4]. Moreover, photovoltaic(PV) power takes a percentage of 8 in the gross power consumption in Italy and 7.1 in Germany in the year of 2015 [5,6]. The large-scale deployments of PV system brings the surging demands for the management and scheduling operations on the PV power system, which greatly depends on the forecasting of PV system power outputs in [7–9]. Genially, PV power outputs are determined by the randomness of solar irradiance in the area of interest, which indicates that the power outputs are variable. Therefore, amount of models and methods have been proposed to approximate the PV power outputs under different conditions.

[†]Corresponding author. Email address: ahpuchina@ahpu.edu.cn In [10], the power regression is modeled based on analysis of the images of the weather in the UC San Diego, which provides a god tested for the solar energy. Similarly, the analysis of cloud images or weather is employed in [11–13]. However, these methods requires expensive equipments to obtain the cloud or weather images, which is not favorable to lower the price of PV power systems. In [14], the forecasting of PV power output is implemented based on the real-time collection of solar irradiance through the irradiance sensor network. In [15], the images of cloud motions are obtained through geostationary satellite to predict the medium and short-term solar radiation. The above approaches can provide good performance in forecasting PV power outputs, which need additional hardware or complex operations.

Besides applying equipments and complex operations, various forecasting algorithms are also proposed. Common approaches are based on employing the machine learning classification and regression methods. In [16], aerosol index, which has evident linear correlation with solar radiation attenuation, is used to train the artificial neural network (ANN) and forecast the power outputs in the next 24 hours. Similarly, the ANN method is also employed to implement the forecasting of the PV power outputs in [17, 18]. Support vector machine(SVM) was also employed to learn and model the relationship and relevance between the input data such as solar radiation and the output of PV power in [19–21]. In [22, 23], multiple linear regression (MLR) modeled the power outputs of PV system based on the features of solar radiation and the weather data. In [24], K-nearest neighbour(K-NN) was employed to build the forecast model based on the non-common data. In [25], ANN, SVM, KNN and MLR are analyzed and the effect of selecting input data for the learning algorithms are analyzed.

Another approach is based on the probabilistic model to forecast probability density function associated with PV power outputs based on the features of input data in [26]. In [27], a versatile probability method based on pair copula construction to model the PV power system. Similarly, a chronological probability is employed to model the output of PV power system based on conditional probability and nonparametric kernel density estimation in [28]. Moreover, the conditional probability associated with the outputs of PV power is also utilized to predict the outputs in the future. In [29], the Bayesian sparse learning that incorporates the features of input data to learn the likelihood function of the outputs of PV power. In the above probabilistic model, the prediction of PV power outputs is inevitably negative, which is resulted from the models and do not follow the positivity of the outputs. Therefore, a sparse Bayesian learning algorithm that guarantees the positivity of the outputs and approximates the relevance between the input features and power outputs is proposed in this paper.

The rest of the paper is organized as follows. In Section II, the forecasting problem is molded as a Poisson regression problem and the regression problem is implemented on the basis of sparse Bayesian learning. The simulation results of proposed algorithm for the forecasting performance are presented in Section III. The conclusion and acknowledgement are given in Section IV and Section V, respectively.

2 System Model

Generally, the basic principle for photovoltaics is the photovoltaic effect, which transform the solar energy to the electrical energy in the semi-conductors. The output power of the photovoltaics is modeled as follows,

$$P_{out\,put} = P_{std} \varepsilon_T R_t,\tag{1}$$

where P_{output}^{t} is the output power at the time *t* and P_{std} is the power output at the standard condition, R_t is the strength of solar radiation at time *t* and ε_T is the parameters that are associated with the are PV cell temperature in current time step and standard test conditions, which is given by

$$\varepsilon_T = \frac{\left[1 + \tau_T \left(T_t - T_0\right)\right]}{R_0},\tag{2}$$

where τ_t is the temperature co-efficient of power and T_t and T_0 are PV cell temperature in current time step and temperature in standard test conditions respectively.

The above model reveals the fact the output power P_{output}^t of a PV cell is mainly affected by the solar radiations R_t and the environmental temperature T_t which both directly depend on different weather types, such as sunny days, cloudy days and rainy days. However, it is known that the outputs of the PV system are not identical even the solar radiations and the environmental temperatures are the same, which are resulted from the ignorance of weather types information the equation (1) and (2). Hence, an more reasonable way is to build a model that incorporates the weather types and the strength of solar radiations to model and predict the PV cell power outputs.

Based on previous discussions, the outputs of a PV system is nonnegative and can be regarded as integers(with low resolutions in large scale systems), and the outputs cannot be modeled by simple support vector machine, Gaussian process or relevance vector machine, which will leads to negative outputs predictions. In order to alleviate the problem, we employ a generalized linear model, poisson regression was built based on the hierarchical Bayesian learning.

2.1 Poisson Regression Model

In the regression of PV power outputs, a training set $\left\{ \{x_{it}^{v}, P_{it}^{v}\}_{i=1}^{N} \right\}_{t=1}^{M}$ are given and x_{it}^{v} represents the input data x_{it}^{v} and v means solar radiations and the weather types. The variable v is defined and quantized as

$$\mathbf{v} = \begin{cases} 0 \quad P_{it} \text{ in rainy weather} \\ 1 \quad P_{it} \text{ in cloudy weather} \\ 2 \quad P_{it} \text{ in sunny weather} \end{cases}$$
(3)

The index *i* represents the day-time step and *t* means the time slots in a day and P_{it} is the corresponding PV power output. In the Poisson Regression, the power outputs is assumed to follow Poisson distribution as

$$f(P_{it}^{\nu}) = \frac{\rho^{P_{it}^{\nu}} e^{-\rho}}{P_{it}^{\nu}!},$$
(4)

where ρ is the natural parameter of the Poisson distribution.

Assuming that the power outputs is linear combinations of the inputs vector, which is given by

$$P_{it}^{\nu} = \sum_{k=1}^{K} \omega_k \phi\left(x_{it}^{\nu}, \boldsymbol{x}^{\nu}\right) + \varepsilon_{it}^{\nu},$$
(5)

where ε_{it}^{v} is a Gaussian noise with zero mean and variance δ^{2} . $\boldsymbol{\omega}^{v} = [\boldsymbol{\omega}_{1}^{v}, ..., \boldsymbol{\omega}_{K}^{v}]$ is the feature weights and $\phi(x_{it}^{v}, \boldsymbol{x}^{v})$ is the kernel function, which is defined as follows,

$$\boldsymbol{\phi}\left(\boldsymbol{x}_{it}^{\boldsymbol{\nu}},\boldsymbol{x}^{\boldsymbol{\nu}}\right) = \left[\boldsymbol{\varphi}^{T}\left(\boldsymbol{x}_{it}^{\boldsymbol{\nu}}\right)\boldsymbol{\varphi}\left(\boldsymbol{x}_{11}^{\boldsymbol{\nu}}\right), \cdots, \boldsymbol{\varphi}^{T}\left(\boldsymbol{x}_{it}^{\boldsymbol{\nu}}\right)\boldsymbol{\varphi}\left(\boldsymbol{x}_{K}^{\boldsymbol{\nu}}\right)\right]^{T},\tag{6}$$

where \mathbf{x}^{v} is the set of all inputs vectors in the weather type v, which is stacked in the time sequence and given by $\mathbf{x}^{v} = [x_{1}^{v}, \dots, x_{K}^{v}]^{T}$. $\boldsymbol{\varphi}(\bullet)$ is the function that projects the input features into dimensional spaces. For example, a universal mapping function [31, 32] is defined as

$$\varphi(\bullet) = \exp\left(-\frac{(\boldsymbol{x}^{\nu} - \bullet)^2}{h}\right),\tag{7}$$

and projects the input features into infinite dimension spaces, which is widely used in the high dimensional regressions and classifications.

Based on (5), the likelihood function can be formulated as follows,

$$f(P_{it}^{\nu}|x_{it}^{\nu}) = \mathscr{N}\left(P_{it}^{\nu}|\sum_{k=1}^{K}\omega_{k}^{\nu}\phi(x_{it}^{\nu},\boldsymbol{x}^{\nu}),\delta^{2}\right).$$
(8)

In Bayesian learning, the weight parameters $\boldsymbol{\omega}$ is assumed to be a vector of random variables that subjects to independent Gaussian prior distributions, which is given by

$$f\left(\boldsymbol{\omega}^{\nu}|\boldsymbol{\lambda}^{\nu}\right) = \prod_{i=1}^{K} \mathcal{N}\left(\boldsymbol{\omega}_{i}^{\nu}|0,\boldsymbol{\lambda}_{i}^{-1}\right),\tag{9}$$

where λ_i^{-1} is the inverse variance of the Gaussian prior distribution and $\boldsymbol{\lambda}^{-1} = [\lambda_1^{-1}, \dots, \lambda_{\text{NM}}^{-1}]^T$. By combining the equation (8) and (9), the posterior of $\boldsymbol{\omega}^{\nu}$ can be formulated as follows,

$$f(\boldsymbol{\omega}^{\nu}|\mathbf{P}^{\nu},\boldsymbol{x}^{\nu},\boldsymbol{\lambda}) \propto f(\mathbf{P}|\boldsymbol{x}^{\nu},\boldsymbol{\lambda},\boldsymbol{\omega}^{\nu})f(\boldsymbol{\omega}^{\nu}|\boldsymbol{\lambda}), \qquad (10)$$

where $\mathbf{P}^{\mathbf{v}} = [P_1^{\mathbf{v}}, \cdots, P_K^{\mathbf{v}}]^T$.

By subsisting the details, the posterior distribution can be given by

$$f(\boldsymbol{\omega}^{\nu}|\mathbf{P}^{\nu},\boldsymbol{x}^{\nu},\boldsymbol{\lambda}) = (2\pi)^{\frac{K}{2}} \boldsymbol{\Lambda}^{\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\omega}^{\nu}-\boldsymbol{\mu})^{T} \boldsymbol{\Lambda}(\boldsymbol{\omega}^{\nu}-\boldsymbol{\mu})\right),$$
(11)

where $\boldsymbol{\mu}$ is the mean, which is formulated as

$$\boldsymbol{\mu} = \frac{\boldsymbol{\Lambda}^{-1} \mathbf{M}^T \mathbf{P}^{\nu}}{\delta^2},\tag{12}$$

where Λ is the inverse covariance matrix of $\boldsymbol{\omega}$ and is given by

$$\mathbf{\Lambda} = \frac{\mathbf{M}^T \mathbf{M}}{\delta^2} + \mathbf{\Lambda}_{\boldsymbol{\lambda}}^{-1},\tag{13}$$

where $\mathbf{\Lambda}_{\boldsymbol{\lambda}}^{-1} = diag\left(\lambda_1^{-1}, \cdots, \lambda_K^{-1}\right)$ is a inverse diagonal covariance matrix of $\boldsymbol{\omega}^{\nu}$. $\mathbf{M} \in \mathbb{R}^{K \times K}$ is a matrix with (i, j)th element $\phi\left(x_i^{\nu}, x_j^{\nu}\right)$.

Similarly, the posterior distribution $f(\boldsymbol{\lambda}, \delta^2 | \mathbf{P}^{\nu}, \mathbf{x}^{\nu})$ can be formulated as

$$f(\boldsymbol{\lambda}, \delta^2 | \mathbf{P}^{\nu}, \mathbf{x}^{\nu}) \propto f(\mathbf{P} | \mathbf{x}^{\nu}, \boldsymbol{\lambda}, \delta^2) f(\boldsymbol{\lambda}) f(\delta^2), \qquad (14)$$

By assuming that λ and δ^2 follows the uniform prior distribution, then $f(\lambda, \delta^2 | \mathbf{P}^v, \mathbf{x}^v)$ can reformulated as

$$f(\boldsymbol{\lambda}, \delta^2 | \mathbf{P}, \boldsymbol{x}) \propto f(\mathbf{P} | \boldsymbol{x}, \boldsymbol{\lambda}, \delta^2).$$
 (15)

By following the results in [33], $\boldsymbol{\omega}^{\nu}$ and δ^2 can be updated as follows,

$$\lambda_k^{(\eta+1)} = \frac{\theta_i}{\mu_k},\tag{16}$$

$$\delta^{(\eta+1)} = \sqrt{\frac{\|\mathbf{P}^{\nu} - \mathbf{M}\boldsymbol{\mu}\|}{K - \theta_i \Lambda_i}}.$$
(17)

where $\theta_i = 1 - \lambda_k^{(\eta)} \Lambda_i$, Λ_i is the *i*th diagonal element of inverse covariance matrix Λ and μ_k is the *i*th element of the posterior mean of μ .

\$ sciendo

534

2.2 Poisson Regression of PV System Power based on SBL: Bayesian Estimation

In this subsection, the Sparse Bayesian Learning method is proposed for the Poisson regression of PV system power outputs.

Given the equation (4) and (5), we redefine the natural parameter ρ as follows,

$$\rho_{it} = \exp\left(\boldsymbol{\omega}^T \phi\left(x_{it}^{\nu}, \boldsymbol{x}^{\nu}\right)\right). \tag{18}$$

So the Poisson distribution can be reformulate as

$$f(\mathbf{P}^{\nu}|\mathbf{x}^{\nu}, \boldsymbol{\omega}^{\nu}) = \prod_{i=1}^{K} \frac{\rho_{ii}^{P_{ii}} e^{-\rho}}{P_{ii}^{\nu}!} = \prod_{i=1}^{K} \frac{\exp\left(e^{\phi(x_{ii}^{\nu}, \mathbf{x}^{\nu})} + P_{ii}^{\nu} \phi\left(x_{ii}^{\nu}, \mathbf{x}^{\nu}\right)\right)}{P_{ii}^{\nu}!},$$
(19)

which can be expressed as follows

$$f(\mathbf{P}^{\nu}|\mathbf{x}^{\nu}, \boldsymbol{\omega}^{\nu}) = \prod_{i=1}^{K} \frac{1}{P_{it}^{\nu}!} \prod_{i=1}^{K} \frac{1}{\Gamma(P_{it}^{\nu})} \frac{\exp\left(e^{\phi(x_{it}^{\nu}, \mathbf{x}^{\nu})} + P_{it}^{\nu}\phi\left(x_{it}^{\nu}, \mathbf{x}^{\nu}\right)\right)}{P_{it}^{\nu}!},$$
(20)

where $\Gamma(P_{it}^{v})$ is a Gamma function. By applying an approximation in, the above result can be approximated as follows,

$$f(\mathbf{P}^{\nu}|\mathbf{x}^{\nu}, \boldsymbol{\omega}^{\nu}) \approx \prod_{i=1}^{K} \frac{1}{P_{it}^{\nu}} N\left(\log P_{it}^{\nu}, P_{it}^{-1}\right)$$

$$= (2\pi)^{\frac{K}{2}} \Lambda_{P}^{-\frac{1}{2}}$$

$$\exp\left(-\frac{1}{2} \left(\sum_{i=1}^{K} \omega_{k}^{\nu} \phi\left(x_{it}^{\nu}, \mathbf{x}^{\nu}\right) - \log \mathbf{P}^{\nu}\right)^{T}\right)$$

$$\exp\left(-\frac{1}{2} \Lambda_{P}^{-1} \left(\sum_{i=1}^{K} \omega_{k}^{\nu} \phi\left(x_{it}^{\nu}, \mathbf{x}^{\nu}\right) - \log \mathbf{P}^{\nu}\right)\right),$$

(21)

where $\Lambda_{\mathbf{P}} = diag(P_1, \cdots, P_K)$ is the covariance matrix.

Taking logarithm to both sides of (21), the log-posterior of $\boldsymbol{\omega}$ can be rewritten as

$$\log f(\boldsymbol{\omega}^{\nu}|\mathbf{P}^{\nu},\boldsymbol{x}^{\nu},\boldsymbol{\lambda}) \propto \log p(\boldsymbol{\omega}^{\nu}|\boldsymbol{\lambda}) + \log f(\mathbf{P}^{\nu}|\boldsymbol{x}^{\nu},\boldsymbol{\omega})$$
$$\approx \sum_{i=1}^{K} \log \mathscr{N}\left(\boldsymbol{\omega}^{\nu}|0,\boldsymbol{\lambda}_{i}^{-1}\right) + \log \left[f(\mathbf{P}^{\nu}|\boldsymbol{x}^{\nu},\boldsymbol{\omega}^{\nu})\right].$$
(22)

By simple manipulations, the above complicated posterior formulation can be rewritten as

$$\log f(\boldsymbol{\omega}^{\nu} | \mathbf{P}^{\nu}, \boldsymbol{x}^{\nu}, \boldsymbol{\lambda}) \propto -\frac{1}{2} \left(\sum_{i=1}^{K} \omega_{k}^{\nu} \phi(\boldsymbol{x}_{it}^{\nu}, \boldsymbol{x}^{\nu}) - \log \mathbf{P}^{\nu} \right)^{T}$$

$$\Lambda_{\mathbf{P}}^{-1} \left(\sum_{i=1}^{K} \omega_{k}^{\nu} \phi(\boldsymbol{x}_{it}^{\nu}, \mathbf{x}^{\nu}) - \log \mathbf{P}^{\nu} \right)$$

$$-\frac{1}{2} [\boldsymbol{\omega}^{\nu}]^{T} \Lambda_{\mathbf{P}}^{-1} \boldsymbol{\omega}^{\nu}.$$
(23)

Hence, the posterior distribution can be formulated as

$$f(\boldsymbol{\omega}^{\nu}|\mathbf{P}^{\nu},\boldsymbol{x}^{\nu},\boldsymbol{\lambda}) = \mathscr{N}\left(\boldsymbol{\omega}^{\nu}|\boldsymbol{\tilde{\mu}},\boldsymbol{\tilde{\Lambda}}\right),\tag{24}$$

\$ sciendo

where $\tilde{\boldsymbol{\mu}}$ is given by

$$\tilde{\boldsymbol{\mu}} = \tilde{\boldsymbol{\Lambda}} \mathbf{M}^T \boldsymbol{\Lambda}_{\mathbf{P}} \log \mathbf{P}^{\mathbf{V}}, \tag{25}$$

where $\tilde{\Lambda}$ is the inverse variance matrix, which given by

$$\tilde{\mathbf{A}} = \mathbf{M}^T \mathbf{A}_{\mathbf{P}} \mathbf{M} + \mathbf{A}_{\boldsymbol{\lambda}}.$$
(26)

So the Bayesian estimation of $\boldsymbol{\omega}^{\nu}$ can be obtained from the mean $\tilde{\boldsymbol{\Lambda}}$.

2.3 Bayesian Learning: Bayesian Estimation

Based on Bayesian rule, the posterior distribution of λ can be given by

$$f(\boldsymbol{\lambda}|\boldsymbol{x}^{\boldsymbol{\nu}},\boldsymbol{P}^{\boldsymbol{\nu}}) \propto f(\boldsymbol{P}^{\boldsymbol{\nu}}|\boldsymbol{x}^{\boldsymbol{\nu}},\boldsymbol{\lambda})f(\boldsymbol{\lambda}).$$
(27)

Without any extra information of λ , it is assume that the prior distribution of λ is uniform distribution. Hence,

$$f(\boldsymbol{\lambda}|\boldsymbol{x}^{\nu},\boldsymbol{P}^{\nu}) \propto f(\boldsymbol{P}^{\nu}|\boldsymbol{x}^{\nu},\boldsymbol{\lambda}).$$
(28)

Then, the above likelihood can be given by

$$f(\mathbf{P}^{\nu}|\mathbf{x}^{\nu},\boldsymbol{\lambda}) = \int_{\boldsymbol{\omega}^{\nu}} f(\mathbf{P}^{\nu}|\mathbf{x}^{\nu},\boldsymbol{\omega}^{\nu}) f(\boldsymbol{\omega}^{\nu}|\boldsymbol{\lambda}) d\boldsymbol{\omega}^{\nu}.$$
(29)

By using the approximation results in (21), it yields

$$f(\mathbf{P}^{\nu}|\mathbf{x}^{\nu},\boldsymbol{\lambda}) \approx (2\pi)^{-\frac{K}{2}} \prod_{i=1}^{K} \sqrt{\boldsymbol{\lambda}}_{i} \boldsymbol{\lambda}_{\mathbf{P}}^{-\frac{1}{2}} \exp\left([\log \mathbf{P}^{\nu}]^{T} \boldsymbol{\Lambda}_{\mathbf{P}} [\log \mathbf{P}^{\nu}] - \tilde{\boldsymbol{\mu}}^{T} \tilde{\boldsymbol{\lambda}} \tilde{\boldsymbol{\mu}}\right).$$
(30)

By taking log operation with respect to both sides of (30), taking derivatives with respect to λ_i and setting it to zero, it leads to

$$\boldsymbol{\lambda}_{i}^{(\eta+1)} = \frac{1 - \boldsymbol{\lambda}_{i}^{(\eta)} \boldsymbol{\Lambda}_{i}^{-1}}{\boldsymbol{\mu}_{i}^{2}},$$
(31)

where Λ_i is the *i*th element of the Λ and μ_i is the *i*th element of μ .

2.4 Prediction of New Inputs Based on Poisson Regression Model

Given the estimation results of $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$, the prediction of new input \boldsymbol{x}^* can be formulated as follows

$$f(P^*|\mathbf{x}^*, \mathbf{x}, \mathbf{P}) = \int f(P^*|\mathbf{x}^*, \boldsymbol{\omega}^v) f(\boldsymbol{\omega}^v | \mathbf{x}^v, \mathbf{P}^v, \boldsymbol{\lambda}) d\boldsymbol{\omega}^v$$

= $\int f(P^*|\mathbf{x}^*, \boldsymbol{\vartheta}^*) f(\boldsymbol{\vartheta}^* | \mathbf{x}^v, \mathbf{P}^v, \boldsymbol{\lambda}) d\boldsymbol{\vartheta}^*$
= $\int f(P^*|\mathbf{x}^*, \boldsymbol{\theta}^*) f(\boldsymbol{\theta}^* | \mathbf{x}^v, \mathbf{P}^v, \boldsymbol{\lambda}) d\boldsymbol{\theta}^*,$ (32)

where $\boldsymbol{\theta}^* = \exp(\boldsymbol{\vartheta}^*) = \exp\left(\boldsymbol{\phi}^T(x^*, \boldsymbol{x}) \, \tilde{\boldsymbol{\mu}}\right)$ and due to the linear transformation $\boldsymbol{\vartheta}^* = \boldsymbol{\phi}^T(x^*, \boldsymbol{x}) \, \tilde{\boldsymbol{\mu}}$, it can be obtained

$$f(\boldsymbol{\vartheta}^*|\boldsymbol{x}^{\boldsymbol{\nu}}, \mathbf{P}^{\boldsymbol{\nu}}, \boldsymbol{\lambda}) = \mathcal{N}\left(\tilde{\boldsymbol{\mu}}_{\boldsymbol{\vartheta}}, \tilde{\boldsymbol{\delta}}_{\boldsymbol{\vartheta}}^2\right),\tag{33}$$

where

$$\tilde{\boldsymbol{\mu}}_{\boldsymbol{\vartheta}} = \boldsymbol{\phi}^{T} \left(\boldsymbol{x}^{*}, \boldsymbol{x} \right) \left(\mathbf{M}^{T} \boldsymbol{\Lambda}_{\mathbf{P}} \mathbf{M} + \boldsymbol{\Lambda}_{\boldsymbol{\lambda}} \right) \mathbf{M}^{T} \boldsymbol{\Lambda}_{\mathbf{P}} \log \mathbf{P}^{\boldsymbol{\nu}}, \tag{34}$$

$$\tilde{\delta}_{\vartheta}^{2} = \phi^{T}(x^{*}, \boldsymbol{x}) \left(\mathbf{M}^{T} \mathbf{\Lambda}_{\mathbf{P}} \mathbf{M} + \mathbf{\Lambda}_{\boldsymbol{\lambda}} \right) \phi(x^{*}, \boldsymbol{x}).$$
(35)

\$ sciendo

536

Due to $\boldsymbol{\theta}^* = \exp(\boldsymbol{\vartheta}^*)$, it is obtained that

$$f(\boldsymbol{\theta}^* | \boldsymbol{x}^{\boldsymbol{v}}, \mathbf{P}^{\boldsymbol{v}}, \boldsymbol{\lambda}) \approx Gamma(m, n),$$
(36)

where

$$m = \frac{1}{\tilde{\delta}_{th}^2},\tag{37}$$

$$n = \tilde{\delta}^2_{\vartheta} e^{\tilde{\mu}_{\vartheta}}.$$
(38)

Following the results in [34], the likelihood function of prediction can be approximated as

$$f\left(P^*|\mathbf{x}^*, \mathbf{x}^{\nu}, \mathbf{P}^{\nu}\right) = \frac{\Gamma(m+P^*)}{\Gamma(1+P^*)\Gamma(m)} \left(\frac{1}{1+n}\right)^m \left(\frac{n}{1+n}\right)^{P^*}.$$
(39)

So the closed-form prediction can be obtained by maximizing the likelihood function.

Based on the above derivations, the detailed algorithm can be formulated as Algorithm 1 as

Algorithm 1 Poisson Kernel Regression Based Sparse Bayesian Learning

- 1: Input the training set $\left\{ \left\{ x_{it}^{v}, P_{it}^{v} \right\}_{i=1}^{N} \right\}_{t=1}^{M}$; 2: Set the convergence criterion for $\boldsymbol{\omega}$ by using the difference between the current estimation and the next estimation;
- 3: Set $\eta = 1$ and the maximum iteration number to be $\eta_{max} = 50$;
- 4: Initialize the parameter $\boldsymbol{\omega}$;
- 5: Initialize the threshold value $\boldsymbol{\omega}_{th}$;
- 6: Initialize the RVs matrix by setting $\mathbf{P}_{RV} = \mathbf{P}$;
- while Maximum iteration or convergence criteria is reached do 7:
- Creating the kernel matrix according to (6); 8:
- Calculate the inverse covariance matrix of $\boldsymbol{\omega}$ according to (26); 9:
- Calculate the mean vector according to (25); 10:
- Updating the hyper-paramter as $\boldsymbol{\lambda}_{i}^{(\eta+1)} = \frac{1-\boldsymbol{\lambda}_{i}^{(\eta)}\boldsymbol{\Lambda}_{i}^{-1}}{\boldsymbol{\mu}_{i}^{2}};$ 11:
- Eliminate the $\boldsymbol{\omega}_i$ and the samples \mathbf{P}_i with $\boldsymbol{\omega}_i > \boldsymbol{\omega}_{th}$; 12:
- Updating kernel matrix by using the eliminated samples; 13:
- 14: end while
- 15: Output the estimation of $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$

2.5 **Summary and Analysis**

Based on combination of Poisson regression and SBL, the power output of PV system can be formulated as a regression problem. Based on the strength of solar radiations in different type weathers, the regression problem can can be classified into three sub-problems.

In each regression problem, the weights of input vectors are dominated by independent zero-mean Gaussian distribution, which is different form the Bayesian prior with identical Gaussian distributions. Meanwhile, the sparsity of the weights are guaranteed by the zero mean and the variance parameter λ , which will alleviate training complexity and time.

On one hand, the complexity of the proposed algorithm is dominated by the step 9 in Algorithm 1 which requires a matrix inverse operation. In the step 9, the complexity is scale with the order of $N(\mathbf{P}_{RV})$. Furthermore,

537

Situations	Sunny	Sunny/Cloudy	Rainy/Cloudy
RMSE of PR-SBL	1.145	11.861	8.343
RMSE of SVM	22.290	22.281	18.715

 Table 1
 The detailed RMSE of three different situations in two regression methods

according to Algorithm 1, the number of P_{RV} will decrease with the iteration process, which means that the complexity decreases in each iteration.

On the other hand, the complexity of prediction is proportional to the number of RV samples. By comparing the analysis results of complexity to other algorithms, it is concluded that the proposed algorithm is less than the related kernel count data regression model including Kernel Probabilistic Regression and Probabilistic Regression [35].

3 Numerical Results

In the section, the data collected from the real PV power system in Anhui Polytechnic University PV power platform will be applied. The installed capacity of the platform reaches 100 kWh, which is deployed on the roof of main administration building in the campus. PV power data and corresponding weather data are collected in a season. Fig.1 shows the collected PV power data in seven different days.

In order to make the weather data clear, the data is shifted by one unit in vertical orientation, which is shown in Fig.2. The RMSE results are obtained through 1000 times Monte Carlo independent experiments, and is defined as RMSE = $\frac{1}{N} \sum_{i=1}^{N} (\hat{P}_{it} - P_{it})^2$, where \hat{P}_{it} is the forecasted PV power output and P_{it} is the true PV power output. *N* is the number of the true data.

In Fig.3, the data is collected in the sunny data and the forecasting values based on Poisson regression is closed to true data and has no negative outputs while SVM regression poses negative outputs and has larger error in table I.

Fig.4 and Fig.5 show the simulation results in hybrid weather. In both situations, the Poisson regression based on SBL algorithm can achieve better performance in forecasting and nonnegativity.

In simulation results, the PV power regression is more complicated under the hybrid weather conditions. In super short-term regression, the other factors, such as environmental temperature and wind speed, can be regarded as stable and unchanged in a sole weather type. Thus just the time sequence correlation is considered. The proposed Poisson regression based on SBL can also incorporate the environmental temperature and wind speed to the input data, then the input data forms a vector and the SBL algorithm can still provide good performance, which can be found in [35].

By combining all the simulation results, the proposed PR-SBL algorithm can provide accurate and nonnegative forecasting values of PV power, which outperforms the SVM algorithm in both aspects. The superiorities is resulted from the Poisson distribution assumption and statistical learning mechanism. Specifically, SBL is a data-driven iterative algorithm and updates the hyper-parameters in a hierarchical way, which prevails over the SVM. Moreover, the assumption that the outputs of PV power subjects to Poisson distribution guarantees the nonnegativity of the predicted data. Furthermore, the assumption can be used by adopting the maximum entropy principle according to the physical situations.

4 Conclusion

The forecasting problem is of vital significance for the management and schedules in the renewable energy sources, such as the PV power system. The traditional nonparametric regression methods cannot guarantee the nonnegativity of the output. In this paper, a regression model based on Poisson distribution and sparse



Fig. 1 The collected PV power outputs data



Fig. 2 The collected Quantized weather data



Fig. 3 The forecasting PV power outputs in sunny days



Fig. 4 The forecasting PV power outputs in sunny/cloudy days



Fig. 5 The forecasting PV power outputs in rainy/cloudy days

Bayesian learning algorithm is proposed to solve the nonnegative PV power forecasting problem. The detailed principles of PR-SBL algorithm and the simulation results are illustrated. The simulation results demonstrate the superiorities and accuracies of the proposed algorithm. Moreover, the proposed algorithm is feasible to other exponential family distribution other than Poisson distribution, which deserves more investigations in the future.

Acknowledgment

The work is supported by the National Natural Science Foundation of China 61572032, Natural Science Foundation of Anhui Province 1908085MF215 and Key Natural Science Research Project of Anhui Province KJ2017A107.

References

- A. Tascikaraoglu, B. M. Sanandaji, G. Chicco, V. Cocina, F. Spertino, O. Erdinc, N. G. Paterakis, and J. P. S. Catalato (2016), Compressive spatio-temporal forecasting of meteorological quantities and photovoltaic power, *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 1295–1305.
- [2] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, and S. Wang (2018), Short-term solar power forecasting based on weighted gaussian process regression, *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 300–308.
- [3] S. Chai, M. Niu, Z. Xu, L. L. Lai, and K. P. Wong (2016), Nonparametric conditional interval forecasts for pv power generation considering the temporal dependence, in 2016 IEEE Power and Energy Society General Meeting (PESGM), pp. 1–5.
- [4] D. Pepe, G. Bianchini, and A. Vicino (2016), Model estimation for pv generation forecasting using cloud cover information, 2016 IEEE International Energy Conference (ENERGYCON), pp. 1–6.
- [5] E. Nuao, M. Koivisto, N. Cutululis, and P. Sarensen (2017), Simulation of regional day-ahead pv power forecast scenarios, 2017 IEEE Manchester PowerTech, pp. 1–6.
- [6] Y. Zhang, M. Beaudin, R. Taheri, H. Zareipour, and D. Wood(2015), Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators, *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2253–2262.
- [7] J. Vasilj, P. Sarajcev, and D. Jakus, Pv power forecast error simulation model(2015), in 2015 12th International Conference on the European Energy Market (EEM), pp. 1–5.
- [8] L. Oneto, F. Laureri, M. Robba, F. Delfino, and D. Anguita(2018), Data-driven photovoltaic power production nowcasting and forecasting for polygeneration microgrids, *IEEE Systems Journal*, vol. 12, no. 3, pp. 2842–2853.
- [9] X. G. Agoua, R. Girard, and G. Kariniotakis(2018), Short-term spatio-temporal forecasting of photovoltaic power production, *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 538–546,.
- [10] B. Urquhart (2011), Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed, ASES.

- [11] R. Marquez and C. F. Coimbra(2013), Intra-hour dni forecasting based on cloud tracking image analysis, *Solar Energy*, vol. 91, pp. 327–336.
- [12] H. Yang, B. Kurtz, D. Nguyen, B. Urquhart, C. W. Chow, M. Ghonima, and J. Kleissl(2014), Solar irradiance forecasting using a ground-based sky imager developed at uc san diego, *Solar Energy*, vol. 103, pp. 502–524.
- [13] S. Quesada-Ruiz, Y. Chu, J. Tovar-Pescador, H. Pedro, and C. Coimbra(2014), Cloud-tracking methodology for intrahour dni forecasting, *Solar Energy*, vol. 102, pp. 267–275.
- [14] A. T. Lorenzo, W. F. Holmgren, M. Leuthold, C. K. Kim, A. D. Cronin, and E. A. Betterton (2014), Short-term pv power forecasts based on a real-time irradiance monitoring network, 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), pp. 0075–0079.
- [15] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker Jr, D. Renné, and T. E. Hoff (2010), Validation of short and medium term operational solar radiation forecasts in the us, *Solar Energy*, vol. 84, no. 12, pp. 2161–2172.
- [16] J. Liu, W. Fang, X. Zhang, and C. Yang (2015), An improved photovoltaic power forecasting model with the assistance of aerosol index data, *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 434–442.
- [17] W. Fei, M. Zengqiang, S. Shi, and Z. Chengcheng (2011), A practical model for single-step power prediction of gridconnected pv plant using artificial neural network,' 2011 IEEE PES Innovative Smart Grid Technologies, pp. 1–4.
- [18] T.-C. Yu and H.-T. Chang (2012), The forecast of the electrical energy generated by photovoltaic systems using neural network method, *Electric Information and Control Engineering (ICEICE)*, 2011 International Conference on. IEEE, pp. 2758–2761.
- [19] J. Shi, W. J. Lee, Y. Liu, Y. Yang, and P. Wang (2012), Forecasting power output of photovoltaic systems based on weather classification and support vector machines, *IEEE Transactions on Industry Applications*, vol. 48, no. 3, pp. 1064–1069.
- [20] S. Qijun, L. Fen, Q. Jialin, Z. Jinbin, and C. Zhenghong (2016), Photovoltaic power prediction based on principal component analysis and support vector machine, 2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia), pp. 815–820.
- [21] Y. Liu, J. Zhao, M. Zhang, F. Liu, H. Ouyang, H. Fang, Q. Hao, and Y. Lu (2016), A novel photovoltaic power output forecasting method based on weather type clustering and wavelet support vector machines regression,2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 29–34.
- [22] S. Liu and M. Dong (2016), Quantitative research on impact of ambient temperature and module temperature on shortterm photovoltaic power forecasting, 2016 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE), pp. 262–266.
- [23] H. K. Elminir, Y. A. Azzam, and F. I. Younes (2007), Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models, *Energy*, vol. 32, no. 8, pp. 1513–1523.
- [24] S. I. Sulaiman, T. K. A. Rahman, I. Musirin, and S. Shaari (2015), Artificial neural network versus linear regression for predicting grid-connected photovoltaic system output, *Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2012 IEEE International Conference on.* IEEE, pp. 170–174.
- [25] B. Saghafian, S. Anvari, and S. Morid (2013), Effect of southern oscillation index and spatially distributed climate data on improving the accuracy of artificial neural network, adaptive neuro-fuzzy inference system and k-nearest neighbour streamflow forecasting models, *Expert Systems*, vol. 30, no. 4, pp. 367–380.
- [26] H. Long, Z. Zhang, and Y. Su(2014), Analysis of daily solar power prediction with data-driven approaches, *Applied Energy*, vol. 126, pp. 29–37.
- [27] M. J. Sanjari and H. B. Gooi (2017), Probabilistic forecast of pv power generation based on higher order markov chain,*IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2942–2952.
- [28] W. Wu, K. Wang, B. Han, G. Li, X. Jiang, and M. L. Crow (2015), A versatile probability model of photovoltaic generation using pair copula construction, *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1337–1345.
- [29] Z. Ren, W. Yan, X. Zhao, W. Li, and J. Yu (2014), Chronological probability model of photovoltaic generation, *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1077–1088.
- [30] M. Yang, S. Fan, and W. J. Lee (2012), Probabilistic short-term wind power forecast using componential sparse bayesian learning, in 48th IEEE Industrial Commercial Power Systems Conference, pp. 1–8.
- [31] Y. Sun, Y. Yuan, and G. Wang (2014), Extreme learning machine for classification over uncertain data, *Neurocomputing*, vol. 128, pp. 500–506.
- [32] J. Moon, J. Park, E. Hwang, and S. Jun(2017), Forecasting power consumption for higher educational institutions based on machine learning, *The Journal of Supercomputing*, pp. 1–23.
- [33] M. E. Tipping(2001), Sparse bayesian learning and the relevance vector machine, *Journal of machine learning research*, vol. 1, no. Jun, pp. 211–244.
- [34] A. B. Chan and N. Vasconcelos(2012), Counting people with low-level features and bayesian regression,(2017) IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 2160–2177, April 2012.
- [35] Y. Jia, S. Kwong, W. Wu, R. Wang, and W. Gao(2017), Sparse bayesian learning-based kernel poisson regression, *IEEE Transactions on Cybernetics*, pp.1–13.