Renewable Energy 43 (2012) 1–8

Contents lists available at SciVerse ScienceDirect

Renewable Energy

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

Review

Applications of Bayesian methods in wind energy conversion systems

Gong Li, Jing Shi*

Department of Industrial and Manufacturing Engineering, 1410 14th Avenue North, North Dakota State University, Dept 2485, PO Box 6050, Fargo, ND 58108, USA

ARTICLE INFO

Article history:
Received 20 June 2010
Accepted 23 December 2011
Available online 17 January 2012

Keywords:
Bayesian methods
Wind energy conversion systems
Wind resource assessment
Wind forecasting
Performance and reliability evaluation

ABSTRACT

The fast growth of wind power is in urgent need of more accurate, reliable, and adaptive modeling and data analysis methods for the characterization and prediction of wind resource and wind power, as well as reliability evaluation of wind energy conversion systems. Bayesian methods have shown unique advantages in statistical modeling and data analysis for the quantity of interest with uncertainty and variability. The adoption of Bayesian methods carries great potentials for various aspects in wind energy conversion systems such as improving the accuracy and reliability of wind resource estimation and short-term forecasts. This paper summarizes the basic theories of several Bayesian methods, and extensively reviews the literature addressing the applications of Bayesian methods in wind energy conversion systems. Based on the state-of-the-art review, the prospects of Bayesian methods in wind energy conversion systems are discussed on how to develop new applications and enhance the methods for existing applications. It is believed that Bayesian methods will be gaining more momentum in wind energy applications in the near future.

1. Introduction

According to World Wind Energy Association [1], the world wind power capacity reached 196,630 MW in 2011 after a nearly exponential growth for 5 years. The top two countries on the chart are China and the U.S. The installed wind power capacity in the U.S. increased from 11,575 MW in 2006 to 40,180 MW in 2010. More astonishing growth took place in China, in which the installed wind power capacity increase from 2599 MW in 2006 to 44,733 MW in 2011. Meanwhile, deep wind energy penetration into electricity market has been witnessed in many countries. For instance, it consists of approximately 21% of electricity use in Denmark in 2010 [1], and it is expected to contribute 20% of the total U.S. electricity supply and 23% of European electricity needs, respectively, by 2030 [2,3].

Modern wind power industry has witnessed huge progress in the past 30 years due to the R&D efforts, as reviewed in literature [4,5]. However, due to the uncertainty and intermittence of the wind resource, it is necessary and important to further improve the accuracy and reliability of characterization and assessment of the wind resource. Although numerous studies have been performed on modeling the wind speed frequency distributions [6] and predicting the short-term wind speed and wind power [7,8], they still cannot guarantee satisfactory results in general. Therefore, the large scale development of wind power is in urgent need of more accurate, reliable, and adaptive modeling and data analysis methods for the characterization and assessment of wind resource and wind power.

Bayesian methods have many unique advantages in statistical modeling and data analysis. Bayesian modeling techniques, such as hierarchical Bayesian modeling and Bayesian networks, provide a natural way to handle missing data, allow combination of data with domain knowledge, facilitate learning about causal relationships between variables, provide a method for avoiding the overfitting of data, predict with good accuracy even with rather small sample sizes, and can be easily combined with decision analytic tools [9]. Bayesian estimation and inference on the confidence intervals of parameters and probability values on hypotheses are more in line with commonsense interpretations [10]. Bayesian methods provide a way to formalize the process of learning from data to update beliefs in accord with recent notions of knowledge synthesis. It is readily adapted to complex random effects models that are more difficult to fit using classical methods than using modern sampling methods. Recently, Bayesian model selection and averaging [11] have gained increasing popularity in various fields, demonstrating their advantages over classical methods [12–14]. Therefore, it is appealing to apply Bayesian methods to wind energy systems.

Bayesian methods are relatively widely adopted in many other fields, but their applications on wind energy are still in early stage.

* Corresponding author. Tel.: +1 701 231 7119; fax: +1 701 231 7195.
E-mail address: jing.shi@ndsu.edu (J. Shi).

0960-1481/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.renene.2011.12.006

It is meaningful to review the start-of-the-art progress about the applications of Bayesian methods in wind energy field, and more importantly, to provide discussion and insights on how to extend their applications in the near future. The remainder of the paper is organized as follows. Typical Bayesian methods are briefly introduced in Section 2. In Section 3, the current applications of the Bayesian methods in the wind energy field are reviewed. In Section 4, a discussion is given on current application efforts and the potential applications. Finally, conclusive remarks are provided in Section 5.

2. Bayesian methods

2.1. Bayes' theorem

In Bayesian methods, the prior, the likelihood, and the posterior information are represented by probability distributions. A prior is a probability distribution representing one's knowledge or belief about an unknown quantity of interest before any corresponding data have been observed. A likelihood is a function of the parameters of a statistical model, reflecting how likely it is to observe current data if the parameters of interest would have current value. The posterior is the probability conditional on the collected data by combing the prior and the likelihood together via Bayes' theorem [10].

As expressed in Equation (1), Bayes' theorem relates two variables (or events), A and B, based on their prior (or marginal) probabilities and posterior (or conditional) probabilities,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

where $P(A|B)$ is the posterior probability of A conditional on B, $P(B|A)$ the prior of B conditional on A, and $P(B)$ the non-zero prior probability of event B, which functions as a normalizing constant. Thus, the relation is often simplified as $P(A|B) \propto P(B|A)P(A)$.

Bayes' theorem intuitively describes the way in which the belief about observing event A is updated by having observed event B. In Bayesian inference, the prior and the likelihood are mathematically combined to produce the posterior. The posterior is then used to determine the probabilities of specified effects. For Bayesian estimation, this theorem is often expressed as

$$p(\theta|x) = \frac{p(\theta)p(x|\theta)}{\int_{-\infty}^{\infty} p(\theta)p(x|\theta)\mathrm{d}\theta} \tag{2}$$

where $p(\theta)$ is the probability density of the prior distribution for the parameter(s) of interest, $p(x|\theta)$ is the probability density distribution of the posterior for data x given θ, and $p(\theta|x)$ is the posterior density of the distribution of θ given data x. The normalization of coefficients ensures that the integral of posterior is always equal to 1.

2.2. Hierarchical modeling and Bayesian network

The hierarchical modeling idea has recently become one of the most important topics in modern Bayesian analysis. It allows us to entertain a much richer class of models that can better capture our statistical understanding of the problem than a simpler model could. Especially, with the advance of Markov Chain Monte Carlo (MCMC) sampling methods, it has become possible to do the calculations on more complex models, thus rendering the hierarchical Bayesian approach more practical.

As mentioned above, given data x and parameter vector θ, a simple Bayesian analysis is to compute a posterior probability based on a prior $p(\theta)$ and likelihood $p(x|\theta)$ according to Equation (2). Often, however, the prior $p(\theta)$ is dependent on other parameters that are not mentioned in the likelihood (e.g., V, W). Thus, the prior $p(\theta)$ must be replaced by a prior $p(\theta|V,W)$, and the prior $p(V)$ and $p(W)$ on the newly introduced parameters are required, respectively, resulting in a posterior as follows,

$$p(\theta,V,W|x) \propto p(x|\theta)p(\theta|V,W)P(V)p(W) \tag{3}$$

The process may be repeated. For instance, the parameters V, W may depend in turn on additional parameter Z which will require its own prior. Eventually the process must terminate, with priors that do not depend on any other unmentioned parameters.

This algorithm can also be represented by a directed acyclic graph (DAG), a Bayesian network, as illustrated in Fig. 1(a). In a Bayesian network, the nodes represent the quantities of interest (variables) and missing edges encode conditional independencies between these variables. A Bayesian network is well known for its ability of providing a compact and simple representation of probabilistic information (uncertainty), allowing the creation of models associating a large number of variables [15].

A general hierarchical Bayesian network usually consists of two parts: the structural part and the probabilistic part. The structural part contains the variables of the network and describes the part-of relationships and the probabilistic dependencies between them. The part-of relationships in a structural part may be illustrated either as nested nodes (see Fig. 1(b)) or as a tree hierarchy (see Fig. 1(c)). The probabilistic part contains the conditional probability tables that quantify the links introduced at the structural part [15].

Hierarchical Bayesian Networks are a generalization of standard Bayesian Networks, where a node in the network may be an aggregate data type. This allows the random variables of the network to represent arbitrary structure types. Within a single node, there may also be links between components, representing probabilistic dependencies among parts of the structure. Hierarchical Bayesian Networks encode conditional probability dependencies in the same way as standard Bayesian Networks. They can express further knowledge about variable structures and use that knowledge to build more realistic probabilistic models.

2.3. Bayesian neural network

Bayesian neural network means to adopt Bayesian learning method for the implementation of a neural network model. As mentioned previously, Bayes' theorem reflects the dynamics of learning and accumulation of the knowledge. The prior distribution encapsulates the state of our current knowledge before we see any

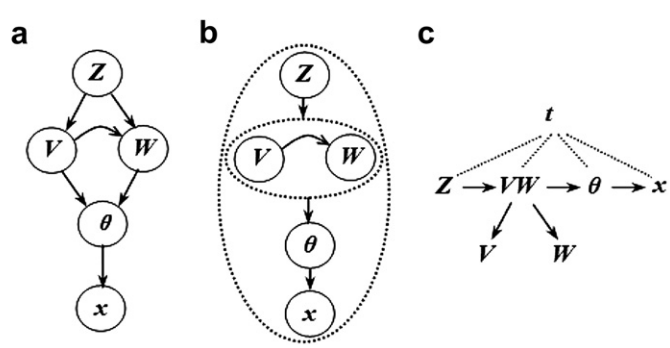

Fig. 1. A simple hierarchical Bayesian network structure. (a) Standard Bayesian network; (b) Nested representation; (c) Tree representation.