ARTICLE IN PRESS

International Journal of Forecasting ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond

Tao Hong[a,*], Pierre Pinson[b], Shu Fan[c], Hamidreza Zareipour[d], Alberto Troccoli[e], Rob J. Hyndman[c]

[a] University of North Carolina at Charlotte, USA
[b] Technical University of Denmark, Lyngby, Denmark
[c] Monash Business School, Monash University, Clayton, Australia
[d] University of Calgary, Alberta, Canada
[e] World Energy & Meteorology Council, UK

ARTICLE INFO

Keywords:
Electric load forecasting
Electricity price forecasting
Wind power forecasting
Solar power forecasting
Probabilistic forecasting
Forecasting competition

ABSTRACT

The energy industry has been going through a significant modernization process over the last decade. Its infrastructure is being upgraded rapidly. The supply, demand and prices are becoming more volatile and less predictable than ever before. Even its business model is being challenged fundamentally. In this competitive and dynamic environment, many decision-making processes rely on probabilistic forecasts to quantify the uncertain future. Although most of the papers in the energy forecasting literature focus on point or single-valued forecasts, the research interest in probabilistic energy forecasting research has taken off rapidly in recent years. In this paper, we summarize the recent research progress on probabilistic energy forecasting. A major portion of the paper is devoted to introducing the Global Energy Forecasting Competition 2014 (GEFCom2014), a probabilistic energy forecasting competition with four tracks on load, price, wind and solar forecasting, which attracted 581 participants from 61 countries. We conclude the paper with 12 predictions for the next decade of energy forecasting.

1. Introduction

In today's competitive and dynamic environment, the energy supply, demand and prices are becoming increasingly volatile and unpredictable. More and more decision-making processes in the energy industry require a comprehensive outlook of the uncertain future. Many decision makers are relying on probabilistic forecasts to quantify these uncertainties, rather than point forecasts.

Here, we use the term "energy forecasting" to refer to "forecasting in the energy industry", which includes but is not limited to the forecasting of the supply, demand and price of electricity, gas, water, and renewable energy resources. Probabilistic forecasts can take various forms, e.g., from quantile to full density forecasts, and probabilistic forecasts for multi-categorical variables or functional data. The business needs for probabilistic energy forecasts spread across the planning and operations of the entire energy value chain.

Thousands of papers on energy forecasting have been published over the past half-century. Hong (2014) provided an overview of energy forecasting, tracing the forecasting practices back to the inception of the electric power industry. A few recent literature review articles have offered more comprehensive views for various subdomains of energy forecasting, such as wind power forecasting (Pinson, 2013; Zhang, Wang, & Wang, 2014), electric load forecasting (Hong & Fan, in this issue), and electricity price forecasting (Weron, 2014). While there

* Corresponding editor.
E-mail address: hong@uncc.edu (T. Hong).

http://dx.doi.org/10.1016/j.ijforecast.2016.02.001
0169-2070/

Please cite this article in press as: Hong, T., et al., Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. International Journal of Forecasting (2016), http://dx.doi.org/10.1016/j.ijforecast.2016.02.001

are differences among the subdomains, there are some common challenges:

(1) Data cleansing. The real-world data for energy forecasting are not always clean. Data cleansing was recognized as one of the challenges of the Global Energy Forecasting Competition 2012 (GEFCom2012), which focused on point forecasting (Hong, Pinson, & Fan, 2014). Data cleansing remains a challenge for probabilistic energy forecasting.
(2) Probabilistic forecasting methodologies. Different subdomains in energy forecasting have different levels of maturity in their probabilistic forecasting. Which probabilistic forecasting methods, such as the ones reviewed by Gneiting and Katzfuss (2014), are applicable to energy forecasting? What specific methodologies are most suitable for a given subdomain? What is the best way to generate input scenarios? Which techniques are most effective at generating probabilistic forecasts? How to simulate residuals?
(3) Forecast combination. Combining a set of point forecasts usually results in more accurate and robust point forecasts. Can we adopt a similar concept for probabilistic energy forecasting? Should we combine point forecasts to give one new point forecast first, then generate a probabilistic forecast, or should we combine point forecasts to give a probabilistic forecast directly? How should probabilistic forecasts be combined to generate a better probabilistic forecast?
(4) Integration. A probabilistic forecasting process can be dissected into several components, such as scenario generation, modeling, and post-processing. An optimal outcome from one component may not be the optimal one for the entire process. The challenge is to integrate the various steps in order to obtain a high quality probabilistic forecast. A similar challenge of integration was also recognized in GEFCom2012 (Hong, Pinson, & Fan, 2014).

In order to maintain the momentum initiated by GEFCom2012, stimulate research activity, and tackle challenges in probabilistic energy forecasting, we decided to start two initiatives simultaneously: (1) organizing a special section for the *International Journal of Forecasting* on probabilistic energy forecasting; and (2) organizing GEFCom2014 with the plan to include the winning methodologies in the same special section.

A call for papers on probabilistic energy forecasting was released on October, 2013. We received 34 submissions, of which seven were accepted for publication in this special section. In addition, we also collected 13 papers from the top entries of GEFCom2014 and one paper from the winning entry of an in-class probabilistic load forecasting competition following a setup similar to that of the probabilistic load forecasting track of GEFCom2014. In total, this special section has collected 21 papers in addition to this hybrid editorial and review paper.

This paper serves four purposes: (1) to discuss the seven non-GEFCom2014 papers collected for this special issue; (2) to introduce the GEFCom2014 and the winning methodologies; (3) to provide an outlook for the field of probabilistic energy forecasting; and (4) to publish the data related to GEFCom2014 and the in-class competition. Section 2 summarizes the non-GEFCom2014 papers. Section 3 discusses the organization of GEFCom2014. Sections 4–7 introduce the four tracks of GEFCom2014, including the problem, the data, and the methods followed by the winning teams. The paper concludes in Section 8 with an outlook for the next decade of probabilistic energy forecasting.

2. Non-GEFCom2014 papers

The seven non-GEFCom2014 papers include three on demand forecasting, two on price forecasting and two on renewable generation forecasting. The subjects being forecasted include electricity and gas demand, electricity prices, wind speed and wave energy. Table 1 lists the authors and titles of the papers.

2.1. Electricity and gas demand forecasting

The majority of the load forecasting literature has focused on point forecasting. Probabilistic load forecasting (PLF) has become attractive to the load forecasting community only over the last decade. In this issue, Hong and Fan offer a tutorial review on PLF. Because most of the studies in the PLF literature have been developed from point load forecasting techniques and methodologies, this tutorial review begins by covering a selection of papers on point load forecasting. The authors then review the research progress on PLF made by two groups, the business consumers of load forecasts and the load forecasters. After reviewing the point and probabilistic load forecasting literature, the authors dissect the PLF problem into three elements, namely the input, model and output. They then introduce ways of producing probabilistic load forecasts from each element and evaluating the probabilistic load forecasts. Finally, the authors conclude the review with an in-depth discussion of future research needs. In addition to their review of the probabilistic load forecasting literature and tutorial about the production and evaluation of probabilistic load forecasts, the authors also offer their opinions about the myth of best techniques, the novelty and significance of load forecasting research, and the importance of integration in load forecasting.

Antoniadis et al. propose a flexible nonparametric function-valued forecasting model. The predictor can be viewed as a weighted average of the futures of past situations, where the weights increase with the similarity between the past situations and the actual one. This strategy can provide simultaneous point predictions at multiple horizons. In addition, these weights also induce a probability distribution that can be used to produce bootstrap pseudo-predictions. Prediction intervals can then be constructed after obtaining the corresponding bootstrap pseudo-prediction residuals. In their paper, Antoniadis et al. propose to obtain prediction intervals that are valid simultaneously for the multiple prediction horizons that correspond to the relevant path forecasts. The methodology is demonstrated using a dataset from the French grid.

Anomaly detection is an important step in data analysis, and has been regarded as one of the challenges of both