Energy 98 (2016) 40–49

Contents lists available at ScienceDirect

Energy

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

Improving short term load forecast accuracy via combining sister forecasts

Jakub Nowotarski [a, b], Bidong Liu [b], Rafał Weron [a, *], Tao Hong [b]

[a] Department of Operations Research, Wrocław University of Technology, Wrocław, Poland
[b] Energy Production and Infrastructure Center, University of North Carolina at Charlotte, USA

ARTICLE INFO

Article history:
Received 19 July 2015
Received in revised form
28 December 2015
Accepted 31 December 2015
Available online 3 February 2016

Keywords:
Electric load forecasting
Forecast combination
Expert aggregation
Sister forecasts

ABSTRACT

Although combining forecasts is well-known to be an effective approach to improving forecast accuracy, the literature and case studies on combining electric load forecasts are relatively limited. In this paper, we investigate the performance of combining so-called sister load forecasts, i.e. predictions generated from a family of models which share similar model structure but are built based on different variable selection processes. We consider 11 combination algorithms (three variants of arithmetic averaging, four regression based, one performance based method and three forecasting techniques used in the machine learning literature) and two selection schemes. Through comprehensive analysis of two case studies developed from public data (Global Energy Forecasting Competition 2014 and ISO New England), we demonstrate that combing sister forecasts outperforms the benchmark methods significantly in terms of forecasting accuracy measured by Mean Absolute Percentage Error. With the power to improve accuracy of individual forecasts and the advantage of easy generation, combining sister load forecasts has a high academic and practical value for researchers and practitioners alike.

1. Introduction

Short term load forecasting is a critical function for power system operations and energy trading. The increased penetration of renewables and the introduction of various demand response programs in today's energy markets has contributed to higher load volatility, making forecasting more difficult than ever before [26,33,34,38]. Over the past few decades, many techniques have been tried for load forecasting, of which the popular ones are artificial neural networks, regression analysis and time series analysis (for reviews see e.g. Refs. [9,24,46]). The deployment of smart grid technologies has brought large amount of data with increasing resolution both temporally and spatially, which motivates the development of hierarchical load forecasting methodologies. The GEFCom2012 (Global Energy Forecasting Competition 2012) stimulated many novel ideas in this context (the techniques and methodologies from the winning entries are summarized in Ref. [25]).

In the forecasting community, combination is a well-known approach to improving accuracy of individual forecasts [3]. Many combination methods have been proposed over the past five decades, including simple average, OLS (Ordinary Least Squares) averaging, Bayesian methods, and so forth (for reviews see Refs. [19,43]). Simultaneously, approaches known as expert aggregation, committee machines or ensemble averaging that typically involve boosting, bagging or random forests have been developed in the machine learning community (for a review see Ref. [40]). However, researchers from the two communities seem to be unaware of the parallel developments [47]. With this paper we try to bridge the gap between these two groups, at least in the context of load forecasting.

Although forecast combination has recently received considerable interest in the electricity price forecasting literature [5,35,39,47] and despite the early applications in load forecasting [6,41], load forecast combination is still an under-developed area. Since weather is a major driving factor of electricity demand, some research efforts were devoted to combining weather forecasts [15,16] and combining load forecasts from different weather

* Corresponding author.
E-mail addresses: jakub.nowotarski@pwr.edu.pl (J. Nowotarski), bdliu0824@gmail.com (B. Liu), rafal.weron@pwr.edu.pl (R. Weron), hongtao01@gmail.com (T. Hong).

http://dx.doi.org/10.1016/j.energy.2015.12.142
0360-5442/

forecasts [10,17]. There are also some studies on combining forecasts from wavelet decomposed series [2,18,29] or independent experts within exponential smoothing [42], nonlinear regression and generalized additive model [12,18], neural network [29], random forest [28], particle swarm optimization [44], and meta-learning [32] frameworks. However, to our best knowledge, there is no comprehensive study on the use of different combination schemes in load forecasting. In particular, only Devaine et al. [12] evaluate different combining schemes, the remaining papers focus on the individual forecasting methods (i.e. experts).

The fundamental idea of using forecast combination is to take advantage of the information that is underlying the individual forecasts and is often unobservable to forecasters. The general advice is to combine forecasts from diverse and independent sources [3,4], which has been followed by most of the aforementioned load forecasting papers. In practice, however, combining independent forecasts has its own challenge. If the independent forecasts were produced by different experts, the cost of implementing forecast combination is often unaffordable to utilities. On the other hand, if the independent forecasts were produced by the same forecaster using different techniques, the individual forecasts often present varying degrees of accuracy, which may eventually affect the quality of forecast combination (for discussions see Refs. [31,47]).

This paper examines a novel approach to load forecast combination: combining sister load forecasts. The sister forecasts are predictions generated from a family of models, or sister models, which share similar model structure but are built based on different variable selection processes, such as different lengths of calibration window and different group analysis settings. The idea of sister forecasts was first proposed and used by Liu et al. [31]; where the authors combined sister load forecasts to generate probabilistic (interval) load forecasts rather than point forecasts as done in this paper. In the forecast combination literature, a similar but less general idea was proposed by Pesaran and Pick [37]; where the authors combined forecasts from the same model calibrated from different lengths of the calibration window. In the machine learning literature, Devaine et al. [12], proposed another related approach. To yield multiple experts, the authors arbitrarily changed some of the parameters responsible for the long-term trend or the short-term effects of a (semi-) parametric model – either nonlinear regression or a generalized additive model.

The contribution of this paper is threefold:

1. This is the first empirical study of combining sister forecasts in the point load forecasting literature. Given that the proposed method is easy to implement compared to computing independent expert forecasts, our approach has far reaching consequences for practitioners.
2. To our best knowledge, this is the most extensive study so far on combining point load forecasts, considering 11 combination and two selection schemes, representatives from both the forecasting and machine learning literature.
3. The two presented case studies are based on publicly available data (GEFCom2014 and ISO New England), which enhances the reproducibility of our work by other researchers.

The rest of this paper is organized as follows. Section 2 introduces the sister load forecasts, 11 combination methods to be tested, and two benchmark methods to be compared with. Section 3 describes the setup of the two case studies. Section 4 discusses the forecasting results, while Section 5 wraps up the results and concludes the paper.

2. Combining sister load forecasts

2.1. Sister models and sister forecasts

When developing a model for load forecasting, a crucial step is variable selection. Given a large number of candidate variables and their different functional forms, we have to select a subset of them to construct the model. The variable selection process may include several components, in particular data partitioning, the selection of error measures and the choice of the threshold to stop the estimation process. Applying the same variable selection process to the same dataset, we should get the same subset of variables. On the other hand, different variable selection processes may lead to different subsets of variables being selected. Following Liu et al. [31]; we call the models constructed by different (but overlapping) subsets of variables *sister models* and forecasts generated from these models – *sister forecasts*.

In this study we use a relatively rich family of regression models to yield the sister forecasts. The rationale behind this choice is twofold. Firstly, regression analysis is a load forecasting technique widely used in the industry [10,20,23,24,27,46]. Secondly, in the load forecasting track of the GEFCom2012 competition attracting hundreds of participants worldwide, the top four winning entries used regression-type models [25]. Nevertheless, other techniques – such as neural networks, support vector machines or fuzzy logic – can also fit in the proposed framework to generate sister forecasts.

We start from a generic regression model that served as the benchmark in the GEFCom2012 competition:

$$\hat{y}_t = \beta_0 + \beta_1 M_t + \beta_2 W_t + \beta_3 H_t + \beta_4 W_t H_t + f(T_t), \tag{1}$$

where $\hat{y}_t$ is the load forecast for time (hour) t, β_i are the coefficients, M_t, W_t and H_t are the month-of-the-year, day-of-the-week, and hour-of-the-day classification variables corresponding to time t, respectively, T_t is the temperature at time t, and

$$f(T_t) = \beta_5 T_t + \beta_6 T_t^2 + \beta_7 T_t^3 + \beta_8 T_t M_t + \beta_9 T_t^2 M_t + \beta_{10} T_t^3 M_t + \beta_{11} T_t H_t + \beta_{12} T_t^2 H_t + \beta_{13} T_t^3 H_t. \tag{2}$$

Note that to improve the load forecasts we could apply further refinements, such as processing holiday effects and weekday grouping (see e.g. Ref. [23]). However, the focus of this paper is not on finding the optimal forecasting models for the datasets at hand. Rather on presenting a general framework that lets the forecaster improve prediction accuracy via combining sister forecasts, starting from a basic model, be it regression, an ARIMA process or a neural network.

Like in Liu et al. [31]; the differences between the sister models built on the generic regression defined by Eqs. (1) and (2) are the amount of lagged temperature variables $\sum_{lag=1}^{N_{lag}} f(T_{t-lag})$, $N_{lag} = 0,1,2,\ldots$, and lagged daily moving average temperature variables $\sum_{d=1}^{N_d} f(\tilde{T}_{t,d})$, $N_d = 0,1,2,\ldots$, where $\tilde{T}_{t,d} = \frac{1}{24}\sum_{k=24d-23}^{24d} T_{t-k}$ is the daily moving average temperature of day d, added to Eq. (1). Hence the whole family of models used here can be written as:

$$\hat{y}_t = \beta_0 + \beta_1 M_t + \beta_2 W_t + \beta_3 H_t + \beta_4 W_t H_t + f(T_t) + \sum_{d=1}^{N_d} f\left(\tilde{T}_{t,d}\right) + \sum_{lag=1}^{N_{lag}} f\left(T_{t-lag}\right). \tag{3}$$