ARTICLE IN PRESS

International Journal of Forecasting ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Sequence of nonparametric models for GEFCom2014 probabilistic electric load forecasting

Ekaterina Mangalova *, Olesya Shesterneva

Siberian State Aerospace University named after academician M. F. Reshetnev, Krasnoyarsk, Russia

ARTICLE INFO

Keywords:
Probabilistic forecasting
Nadaraya–Watson estimator
Quantile estimation
Pinball loss function
Sequence of estimators

ABSTRACT

The probabilistic forecasting method proposed in this paper is based on the use of the sequence of Nadaraya–Watson estimators. It allows estimates of quantiles to be obtained without assumptions as to the probability distribution. The effectiveness of the approach is demonstrated during the Global Energy Forecasting Competition 2014 in the probabilistic electric load forecasting track.

1. Introduction

Load forecasting plays an important role in the management of the load, and has a great influence on the operations, control and planning of power systems. Accurate forecasts of the electricity consumption can provide significant economic benefits.

Most forecasting methods provide estimates of the future load as point forecasts. This means that we get only one "best" forecast, which is produced by a numerical model. Today, more and more decision making processes in the energy industry require more information than just a single value. The aim of the probabilistic electric load forecasting track is to forecast the probabilistic distribution (in quantiles) of the hourly loads for one utility on a rolling basis using hourly historical load and weather data from the utility (Hong et al., this issue).

There are two main approaches to probabilistic forecasting: the prediction error approach and the direct approach. The prediction error approach provides probabilistic forecasts of the errors of an existing deterministic forecasting model and adds uncertainty estimation to the existing spot forecasting system. Alternatively, the direct approach focuses on providing probabilistic predictions of the output variable directly (Juban, Siebert, & Kariniotakis, 2007).

In this paper, we propose a direct nonparametric approach based on the fitting of a sequence of Nadaraya–Watson estimators. The main idea of nonparametric estimators sequence building consists of the use of multiple forecasts, with each forecast computed in a different way. Then, postprocessing is used to convert the estimators in the sequence into probabilistic forecasts (quantiles of the predictive distribution). This technique avoids the need for prior assumptions as to the underlying probability distributions.

2. Algorithm description

The main idea of the proposed probabilistic forecasting algorithm consists of the use of multiple forecasts, with each forecast computed in a different way using the previous forecast. These forecasts are then transformed into quantiles of the predictive distribution. The scheme of the algorithm is presented in Fig. 1.

Training set T. During the learning process, the n-vector of outputs is changed. We start the learning process with the median estimation, and denote the training set used for it by $T^{\langle M\rangle} = T = (x_i, y_i)\ ,\ i = 1, 2, \ldots, n$, where M is the number of the median estimator in the sequence.

* Corresponding author.
E-mail address: e.s.mangalova@hotmail.com (E. Mangalova).

http://dx.doi.org/10.1016/j.ijforecast.2015.11.001
0169-2070/

Please cite this article in press as: Mangalova, E., & Shesterneva, O., Sequence of nonparametric models for GEFCom2014 probabilistic electric load forecasting. International Journal of Forecasting (2016), http://dx.doi.org/10.1016/j.ijforecast.2015.11.001

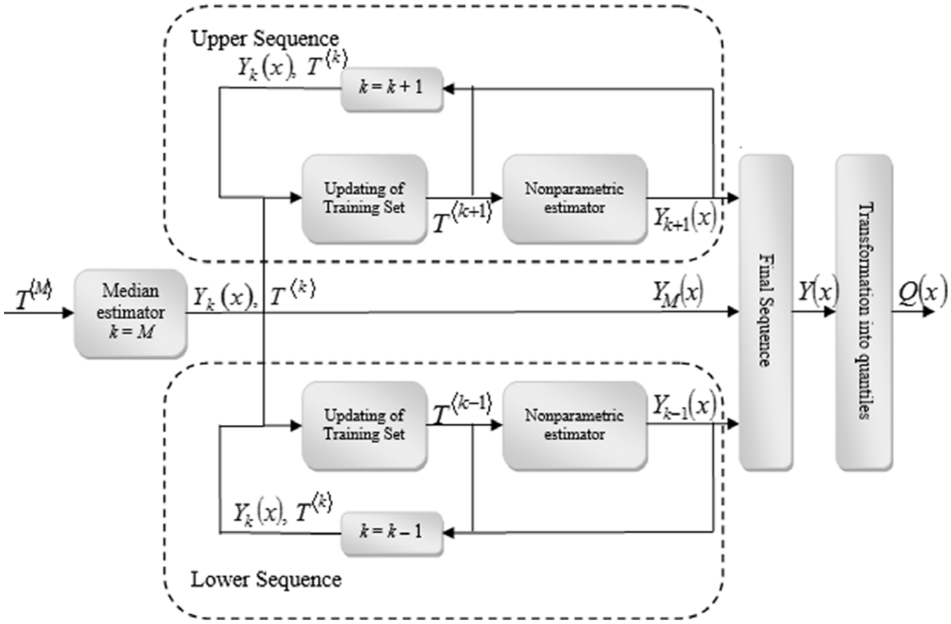

Fig. 1. Summary of the proposed algorithm.

Median estimator. Median estimation is the typical task of point forecasting according to the following criteria:

$$L\left(\hat{q}_{0.5}(x_i), y_i\right) = 0.5\left|\hat{q}_{0.5}(x_i) - y_i\right|$$
$$= \begin{cases} 0.5\left(\hat{q}_{0.5}(x_i) - y_i\right), & \hat{q}_{0.5}(x_i) > y_i, \\ 0.5\left(y_i - \hat{q}_{0.5}(x_i)\right), & \hat{q}_{0.5}(x_i) \le y_i. \end{cases} \quad (1)$$

The median (0.5-quantile) can be estimated as follows:

$$\hat{q}_{0.5}(x) = \text{median}\left\{ y_i,\ i = 1, \ldots, n : \prod_{j=1}^{m} K\left(x^{\langle j\rangle}, x_i^{\langle j\rangle}, c^{\langle j\rangle}\right) > 0 \right\}, \quad (2)$$

where K is the kernel function, $\bar{c} = \left\{c^{\langle j\rangle}, c^{\langle 2\rangle}, \ldots, c^{\langle m\rangle}\right\}$ is the vector of bandwidths. For example, the triangular kernel can be used

$$K\left(x^{\langle j\rangle}, x_i^{\langle j\rangle}, c^{\langle j\rangle}\right) = D\left(\frac{\left|x^{\langle j\rangle} - x_i^{\langle j\rangle}\right|}{c^{\langle j\rangle}}\right), \quad (3)$$

$$D(t) = \begin{cases} (1-t), & |t| \le 1, \\ 0, & |t| > 1. \end{cases} \quad (4)$$

At this stage of the algorithm, the following variations are possible:

- various kernel functions (different kernel functions can be used for different input variables);
- different bandwidths.

The process of constructing the nonparametric estimators sequence begins from the median estimation

$$Y_M(x) = \hat{q}_{0.5}(x), \quad (5)$$

where M is number of the middle estimator in the sequence, then splits into upper sequence building and lower sequence building.

Nonparametric estimators. The Nadaraya–Watson estimator (Hardle, 1992) is chosen as the base model in the proposed algorithm.

$$Y_k(x) = \frac{\sum_{i=1}^{n}\prod_{j=1}^{m} K\left(x^{\langle j\rangle}, x_i^{\langle j\rangle}, c^{\langle j\rangle}\right) y_i^{\langle k\rangle}}{\sum_{i=1}^{n}\prod_{j=1}^{m} K\left(x^{\langle j\rangle}, x_i^{\langle j\rangle}, c^{\langle j\rangle}\right)}, \quad (6)$$

where $y^{\langle k\rangle} = \left\{y_1^{\langle k\rangle}, y_2^{\langle k\rangle}, \ldots, y_n^{\langle k\rangle}\right\}$ is the updated output variable in the training set on the kth iteration. The peculiarities of the updating procedure for the training set are discussed in detail below.

The list of possible variations coincides with the variations that are allowed in median estimation. Generally speaking, other machine learning methods (such as neural networks, support vector machines, etc.) can also be used.

Updating procedure. The procedure of sequence building looks like the procedure of model generation in some ensemble learning methods. As with bagging algorithms, we want to obtain estimators that are as diverse as possible. However, our task is more complex, because we impose an additional requirement on the estimators:

$$Y_{k-1}(x) \le Y_k(x), \quad k = 2, \ldots, 2M-2. \quad (7)$$

This requirement means that, for each input x_i, the forecast of the kth estimator must be no less than the forecast of the $(k-1)$th estimator. In other words, we want to obtain a non-decreasing sequence of estimators.

The requirement in Eq. (7) means that standard diversity generation techniques (bootstrapping, using different parameters, etc.) are unsuitable. For this reason, we propose a regular procedure for updating the training set in order to generate a non-decreasing sequence of estimators.

Please cite this article in press as: Mangalova, E., & Shesterneva, O., Sequence of nonparametric models for GEFCom2014 probabilistic electric load forecasting. International Journal of Forecasting (2016), http://dx.doi.org/10.1016/j.ijforecast.2015.11.001