Neurocomputing 182 (2016) 118–132

Contents lists available at ScienceDirect

Neurocomputing

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

Forecasting electricity load with advanced wavelet neural networks

Mashud Rana, Irena Koprinska *

School of Information Technologies, University of Sydney, Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 28 September 2014
Received in revised form
16 November 2015
Accepted 8 December 2015
Communicated by Dr. P. Zhang
Available online 14 December 2015

Keywords:
Electricity load forecasting
Neural networks
Wavelet transform with best basis selection
Mutual information

ABSTRACT

Electricity load forecasting is a key task in the planning and operation of power systems and electricity markets, and its importance increases with the advent of smart grids. In this paper, we present AWNN, a new approach for very short-term load forecasting. AWNN decomposes the complex electricity load data into components with different frequencies that are predicted separately. It uses an advanced wavelet transform with entropy cost function to select the best wavelet basis for data decomposition, mutual information for feature selection and neural networks for prediction. The performance of AWNN is comprehensively evaluated using Australian and Spanish electricity load data for one-step and multi-step ahead predictions, and compared with a number of benchmark algorithms and baselines.

1. Introduction

Forecasting electricity load is required for the efficient and reliable operation of modern electricity networks. It is used for numerous decisions, including the optimal scheduling of generators, setting the reserve and planning maintenance. Very Short-Term Load Forecasting (VSTLF) – predicting the load minutes to hours ahead – is particularly important in deregulated and competitive electricity markets. The market participants (electricity generation and transmission companies and independent market operators) use continuous forecasting to plan and support their bids, decisions and operations.

VSTLF is becoming even more important with the emerging smart grid technology [1]. A large amount of data from smart meters is being collected at very short intervals (e.g. from 5 min to 1 h), at various levels of granularity (e.g. household, block, suburb, region, etc.). This data can be used to better understand the usage patterns of the customers, and plan changes and incentives in order to optimize the electricity usage, ensuring reliable supply at a minimum cost. The demand-respond mechanism of the smart grid will depend critically on VSTLF as it introduces interactions between the electricity price and demand, and also high penetration of renewable energies [1].

The accuracy of the electricity load forecasting is very important. An over-prediction will lead to employing too many generators and unnecessary increase in the spinning reserve and operating costs. An under-prediction may lead to endangering the system reliability due to insufficient resources to meet the security requirements or higher cost due to purchasing expensive spot market electricity to meet the demand.

In this paper, we consider the following VSTLF task. Given a time series of electricity loads $X_1, X_2, \ldots, X_t$ recorded every 5 min or every 60 min up to time t, our task is to predict the future load X_{t+h} where h is the forecasting horizon. We consider one-step ahead prediction, i.e. $h=1$, and multi-step ahead prediction, i.e. $h=1,\ldots,12$. Weather variables are not included since the weather changes are already captured in the load series for forecasting horizons smaller than a few hours [2].

Forecasting the electricity load with high accuracy is a challenging task. The electricity time series is complex, with non-linear dependencies, and contains both periodic and random components. The periodic components are due to the weekly and daily nested cycles. The random components are due to the inherent fluctuations in the household electricity usage, changes in industry usage, e.g. big consumers with unknown hours of operation, and also variations due to weather changes, special events, calendar and economic factors, and malfunction of measuring devices.

In this paper we study the potential of advanced wavelet decomposition and nonlinear algorithms for feature selection and prediction, to build accurate prediction models for VSTLF. We apply a wavelet decomposition to find a set of optimal frequency components to represent the data and then predict each of these components separately. The motivation is that the load at a given time can be seen as a combination of components with different frequencies, e.g. low frequency components representing the more regular patterns and high frequency components representing the irregular fluctuations. By identifying these components and predicting them separately, we might be able to build more accurate prediction models.

* Corresponding author.
E-mail addresses: mashud.rana@sydney.edu.au (M. Rana), irena.koprinska@sydney.edu.au (I. Koprinska).

http://dx.doi.org/10.1016/j.neucom.2015.12.004
0925-2312/

Our contribution can be summarized as follows:

1. We present an approach for time series forecasting called Advanced Wavelet Neural Network (AWNN), and its application for VSTLF. It uses an advanced wavelet algorithm for load decomposition (wavelet packet with best basis selection), Mutual Information (MI) for feature selection and Neural Networks (NNs) as prediction algorithm.
   The wavelet packet transform with best basis selection decomposes the load into a subset of different frequency components, that are optimal with respect to a given cost function, appropriate for the task. In contrast, the standard wavelet and wavelet packet decompositions do not optimize the wavelet representation. We compare the performance of the best-basis wavelet with these two standard wavelet decompositions, and also without wavelets, and show that the best-basis wavelet improves the performance. We also highlight the importance of using a shift-invariant wavelet transform and the need to address border distortion during wavelet decomposition. We propose and evaluate a prediction-based approach for minimizing border distortion. Wavelet transform with best basis selection has not been applied for load forecasting and the problems of shift invariance and border distortion have not received enough attention in previous work.
2. We propose an efficient feature selection method based on MI. Most of the methods for feature selection in electricity load forecasting are based on autocorrelation and can capture only linear relationships, while MI is also able to capture non-linear dependencies. Our method is efficient (i) as it is based on MI estimation using k nearest neighbor distances and (ii) as we apply a two-step process that firstly identifies a small set of candidate features and then computes the MI for them. For each wavelet component, we apply a small (1-week) sliding window to form a set of candidate features. The MI between each candidate and the target variable is computed and the most informative candidates are selected to form the final feature set. As a prediction algorithm we employ a multi-layer perceptron NN, trained with the Levenberg–Marquardt algorithm which offers faster convergence than the standard backpropagation gradient descent algorithm.
3. We conduct a comprehensive evaluation of the performance of AWNN using two different datasets of electricity load data for two years: Australian 5-min data and Spanish 60-min data. The different geographic location and time resolution are chosen to better assess the robustness of AWNN. We evaluate the impact of the wavelet decomposition and analyze the performance of AWNN for different hours of the day, different months of the year and different days of the week (working, weekend and holidays). We also compare AWNN with state-of-the-art versions of exponential smoothing and Auto-Regressive Integrated Moving Average (ARIMA) methods, a typical industry model, a number of baselines and also linear regression and model tree rules prediction algorithms instead of NN, with and without wavelet transforms. In addition to one-step ahead prediction, we test the performance of AWNN for multi-step ahead prediction (up to 12 steps ahead) using an iterative methodology.

The rest of this paper is organized as follows: Section 2 reviews the related work and Section 3 presents our proposed approach; Section 4 describes the experimental setup, Section 5 presents and discusses the results, and Section 6 concludes the paper.

2. Related work

VSTLF, i.e. predicting the load minutes and hours ahead, is a relatively new area, which has become important with the arrival of competitive electricity markets and the smart grid. In contrast, short-term load forecasting, i.e. predicting the load one to several days ahead, has been a very active area of research, for a survey see [3,4]. Our review focuses on load forecasting and wavelet based methods for forecasting. It is organized into four parts: approaches for VSTLF, wavelet-based approaches for load forecasting (both very short-term and short-term), wavelet-based approaches for other energy time series forecasting and wavelet packet based approaches.

2.1. Approaches for VSTLF

There are two main groups of approaches for VSTLF: 1) statistical, e.g. exponential smoothing, moving average, ARIMA and linear regression, and 2) computational intelligence, e.g. NNs, fuzzy rules, support vector machines. The statistical approaches are often criticized as they are model-based [4] – they fit a model based on prior knowledge about the relationship between the inputs and the output, and also assume linear relationship between the input variables and the output variable. In contrast, NNs, the most popular representatives of the second group, can model non-linear relationships, and learn them from a set of training examples. NNs, however, have a number of drawbacks: they are very sensitive to parameter tuning, may not converge to a satisfactory solution or take too long to converge and also do not have an in-built mechanism for feature selection.

Taylor's work [2,5,6] is one of the most prominent and successful examples of statistical approaches for VSTLF. In [5] he considered 30-min British data and showed that for one-step ahead prediction the best method was double seasonal ARIMA, followed by Holt–Winters exponential smoothing with double, weekly and daily seasonality. In [2] 1-min British data was used to predict the load up to 30 min ahead. The best methods for 5-min ahead prediction were double seasonal Holt–Winters exponential smoothing, restricted intraday exponential smoothing and double seasonal ARIMA, achieving MAPE of about 0.25%. In [6] Taylor et al. considered the task of predicting the hourly electricity load for Rio de Janeiro, for forecasting horizon from 1 to 24 h. They compared four methods: ARIMA, double seasonal Holt–Winters exponential smoothing, backpropagation NN and a PCA-based linear regression. They found again that the exponential smoothing was the most accurate method, in addition to being the simplest and fastest method.

Various NN-based approaches have been proposed. Charytoniuk and Chen [7] used NN for 10-min ahead load forecasting from the previous 20–90 min load data achieving MAPE=0.4–1%. Shamsollahi et al. [8] applied NN for 5-min ahead load forecasting obtaining MAPE=0.12%. Both approaches forecasted load differences instead of actual load. In [9] we considered one-step ahead prediction from 5-min Australian data and built seasonal and yearly models. Applying autocorrelation feature selection, we found that LR was more accurate than NN, achieving MAPE=0.294%. The results also showed that the seasonal models did not improve accuracy in comparison to a single yearly model.

Troncoso et al. [10,11] considered hourly data and proposed WNN, a weighted nearest neighbor approach for simultaneously predicting the 24 hourly values for the next day. WNN first finds k similar days (nearest neighbors) and then computes a 24 dimensional vector (the prediction), where each element is the weighed linear combination of the load for the days following the neighbors. A generalization of this method, called PSF, that combines clustering with matching of sequences of cluster labels, was proposed in [12]. PSF was tested on three electricity load datasets (Spanish, Australian and American) and achieved MAPE=4.02–5.97%, outperforming NN and GARCH autoregressive models. A