Electric Power Systems Research 154 (2018) 401–412

Contents lists available at ScienceDirect

Electric Power Systems Research

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

New online load forecasting system for the Spanish Transport System Operator

Miguel López [a,*], Sergio Valero [a], Ana Rodriguez [b], Iago Veiras [c], Carolina Senabre [a]

[a] Universidad Miguel Hernández de Elche, Spain
[b] Red Eléctrica de España (REE), Spain[1]
[c] ALTEN, Spain[2]

ARTICLE INFO

Article history:
Received 12 December 2016
Received in revised form 16 August 2017
Accepted 4 September 2017

Keywords:
Load forecasting
Power demand
Neural Network application

ABSTRACT

This paper presents the implementation of a new online real-time hybrid load-forecasting model based on an autoregressive model and neural networks. This new system is currently running at the Spanish Transport System Operator (REE) and provides an hourly forecast for the current day and the next nine days timely every hour for the national system as well as 18 regions of Spain. These requirements impose a heavy computational burden that needs to be considered during the design phase. The system is developed to improve forecasting accuracy specifically on difficult days like hot, cold and special days. In order to achieve this goal, a deep analysis of the temperature series from 59 stations is made for each region and the relevant series are included individually in the model. Special days are also analyzed and a thorough classification of days is proposed for the Spanish national and regional system. The model is designed and tested with data from 2005 to 2015. The results provided for the period from December 2014 to October 2015 show how the addition of the proposed model to the TSO's ensemble causes a 5% RMSE overall error reduction and a 15% reduction on the 59 difficult days considered in the testing period.

1. Introduction

Short-term load forecasting (STLF) has been an active research topic for a long time. The changing characteristics of the consumers (air conditioning universalization or appearance of electric vehicle) and lately also of the producers (different renewables and distributed generation) keep forcing the industry to obtain more and more accurate forecasts every day. Short-term forecasting includes lead times from 1 hour to several days and it provides relevant information to system operators to ensure reliability of the system and to producers for determining schedules and utilization. Another application of STLF is the optimization of market bidding for both sides of the market. The deregulation of the Spanish markets in the past decade has put a lot of pressure on forecasts to improve trading profits.

As it was previously stated, STLF has received a lot of attention in the last decades [1–6]. Forecasting models have evolved from statistical models to more complex models based on different sorts of artificial intelligence. Statistical models include multiple linear regression models [7–9], exponential smoothing techniques [10] and time-series [11–14]. These methods offer accurate results and their research is currently active. Artificial intelligence in STLF comprises several techniques like Artificial Neural Networks (ANN) [15–19], fuzzy logic [16,18,20–23], Support Vector Machines (SVM) [24] or Evolutionary Algorithms [18,20,25–27]. The aforementioned categories refer to the mathematical entity that processes the data, the forecasting engine. Many of the referred techniques are combined in hybrid models that produce forecasts in several steps.

However, the forecasting engine is not the only key aspect of a forecasting models and other processes like data normalizing, filtering of outliers, clustering of data or decomposition by data transform [26–29] are also relevant. Specifically, this paper will focus on temperature and special day data treatment. The characteristics of the load (influence of meteorology, type of day or social events among others) need to be taken into account in order to develop an accurate model for the specific data base [3,4], therefore it is not possible to determine a single technique that outperforms the rest.

Moreover, those systems working under real conditions or tested as real world applications [30,31] are of special relevance [4]. The significance of created knowledge that is validated by con-

* Corresponding author.
E-mail address: m.lopezg@umh.es (M. López).
[1] www.ree.es.
[2] www.alten.es.

http://dx.doi.org/10.1016/j.epsr.2017.09.003

Nomenclature

Indices

d	Day
d^*	Similar day
h	Hour
m	Forecasting model
s	Weather station
t	Instant

Variables

$e(d, h)$	Forecasting error on day d at hour h
$f(m, d, h)$	Forecast from model m for day d at hour h
$L(T_{sd}, h)$	Load as a function of temperature at station s, on day d at hour h
$Lr(d, h)$	Actual load on day d at hour h
$p(d, h)$	Combined forecast for day d at hour h
T_{sd}	Temperature at station s, on day d
$y(t)$	Output of the model

Parameters and settings

$\alpha_{(1,3)}$	Slope of linearized relation between cold $^{(1)}$ and hot $^{(3)}$ temperatures and load
$\beta_{(1,3)}$	Intercept of linearized relation between cold $^{(1)}$ and hot $^{(3)}$ temperatures and load
p	Number of lags of the AR model
q	Number of lags of the MA model
ϕ	Coefficients for the AR part of the model
θ	Coefficients for the exogenous part of the model
n_y	Number of lags of the Neural Network model
n_u	Number of exogenous inputs of the Neural Network model
k	Intercept for the linear combination of forecasts
Λ	Weight for each model in the combination
n_s	Number of similar days in the optimization period
w	Weight of the similar days in the optimization function
THH_s,THH_c	Temperature thresholds for hot (s) and cold (c) days on station s

tinuous use by the industry instead of self-reported results under lab conditions is especially important for the advancement of the field.

Red Eléctrica de España (REE), the Spanish Transport System Operator (TSO) is a key agent for the efficiency of the electric system and it periodically seeks improvements on its STLF system. Specifically, their previous research work [32], states the relevance of understanding how the load reacts to special days and weather variables. Lack of accuracy in STLF for special days is the main loss derived from their forecasting system and, therefore, it is the focus of this research. Special days show a deviation from the expected load profile due to several reasons: national or regional holidays, Daylight Savings Time (DST), special periods like Christmas or Easter and even extreme temperatures. Even though the effects that these factors may have is generally known, in order to achieve the desired level of accuracy, it is required that each type of day identified for each electric system and even subsystems.

This paper presents a complete, functional forecasting system that is currently running at the REE headquarters providing timely forecasts every hour. The system's ability to perform accurately in a real environment is one of the key aspects of this research, but its main accomplishment is how the local weather variables and the calendar data is treated in order to build an information system in which a forecasting engine may be able to detect patterns and forecast future loads.

Considering all of the above, this paper presents a new hybrid model based on neural networks (NN) and autoregressive (AR) techniques that includes all the processing steps to be executed online (in a real-time environment) and provides a timely forecast not only for the national aggregate but also for each of the 18 regions that the Spanish TSO considers. The information treatment for both temperature data and calendar information conforms the key innovation presented in this paper:

- Section 2.2.2 describes a methodology which is valid to select the best locations from the available temperature data series as well as to obtain the key parameters for the treatment of the selected series.
- Section 2.2.3 describes the classification system for each type of day employing two levels of variables (exclusive and modifiers) to assign a proper classification to each day of the year. This system requires a deep understanding of the behavior of the system but its results prove that a more simplistic classification is insufficient to achieve the necessary accuracy.

The described procedures are original methods that outperforms the initial system described in Ref. [32].

The use of both NN and AR techniques have been widely described and it is not a key aspect of this model: Both methods have been employed as they produce low correlated error from similar input information. Each technique provides a separate forecast, and the final output is based on a linear combination of both techniques optimized over a period of time. The national output is obtained by a similar linear combination of both national forecasts and both aggregations of the regional forecasts. The actual forecast consists on a 24-h curve for each day from the current day to nine days ahead, considering the most relevant information for each horizon of prediction. Although, these forecasts are made hourly the results shown in this paper refer to the forecast made at 9 a.m. for the next day. The results of the proposed methods allow the TSO to improve the general accuracy of its whole system, especially on the difficult days that motivated this work.

This paper is organized as follows: Section 2 summarizes the characteristics of the databases used, the different techniques that have been evaluated and the definition of the definitive model used. On Section 3, the results of the model are shown along with a detailed analysis regarding each sub-model, meteorology, type of day and other variables affecting accuracy. Section 4 contains the conclusions and recommendations of use considering the results exposed on the previous section.

2. Materials and methods

The work described in this paper starts as a project to enhance the STLF system for the Spanish TSO (REE). The project started from scratch from the data acquisition system to the forecast data formatting and alarms system. The input of the system is the historic database along with a series of text files from REE and AEMET (National Weather Agency) with both new load and weather information. The output is a text file with a current-day and nine following days hourly forecast which is casted every hour. One of these files is created for each of the 18 zones of the peninsula plus one for the whole system. These files are then included in the TSO's forecasting ensemble.

During the development of the final models, several techniques were evaluated from which some were finally used and others were discarded. The criteria used to include each technique were based on forecasting results over testing periods of at least two years.