



The comparison of Holt–Winters method and Multiple regression method: A case study



Liljana Ferbar Tratar ^{a,*}, Ervin Strmčnik ^b

^a University of Ljubljana, Faculty of Economics, Kardeljeva ploščad 17, 1000 Ljubljana, Slovenia

^b University of Ljubljana, Faculty of Mechanical Engineering, Aškerčeva 6, 1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:

Received 28 December 2015

Received in revised form

17 April 2016

Accepted 25 April 2016

Keywords:

District heating

Heat load forecasting

Multiple regression

Holt–Winters methods

ABSTRACT

The European Union approach towards a low-carbon society in EU provides many measures. Appropriate heat load forecasting techniques offer opportunity for more effective schedule operations and cost minimization. The Company *Energetika Ljubljana* claims the largest district heating network in the Republic of Slovenia. Although the company has a 150-year tradition, the company has not implemented any of the advanced heat load forecasting methods. Especially long-term heat load forecasting methods offer many opportunities for the strategic planning and the optimal scheduling of heating resources, whereas short-term forecasting approach would help to reach the optimal daily operations and the maximum utilization of the company's resources. This paper presents forecasting approach for short- and long-term heat load forecasting on the three levels: monthly, weekly and daily forecasting bases. The comparison of the forecasting performances of Multiple regression and Exponential smoothing methods has been analysed. Based on chosen accuracy measures, Multiple regression was recognized as the best forecasting method for daily and weekly short-term heat load forecasting, whereas Holt–Winters methods ensured the best forecasting values in purpose of long-term heat load forecasting and monthly short-term heat load forecasting.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Current challenges facing the European Union are sustainable energy supply, reducing dependence on energy resources and energy efficiency [1]. In last years, The European Union accepted many documents regarding efficient energy production and sustainability. Growth strategy *Europe 2020* [2] represents a direction towards a smart, sustainable and inclusive growth of the European Union economy. Climate/energy is one of the fifth ambitious objectives to be reached by 2020. Challenges and solutions to a fully decarbonised power sector have been analysed in a practical guide to a prosperous low-carbon Europe (*Power Perspectives 2030* [3]), which is a contribution to the European Climate Foundation's *Roadmap 2050*. The Republic of Slovenia, as member of the European Union, tried to follow them and accepted Energy law in March 2014 and Action plan for energy efficiency 2014–2020 in May 2015. Forecasting is one of the possible measures for the efficient energy production, and wise management of energy resources. The

Slovenian company *Energetika Ljubljana* uses a very simple, one day ahead heat load forecasting model, which is mainly based on one external factor. Besides opportunities for short-term forecasting, the company has a huge potential for an implementation of long-term heat load forecasting approach. The proposed forecasting approach may lead to lower costs of energy production, a lower environment impact and higher security of supply.

The aim of this article is the identification of the best forecasting methods for long- and short-term daily, weekly and monthly heat load forecasting. We defined the following hypotheses: 1) »Multiple regression method is the most suitable forecasting method for short-term heat load forecasting«, 2) »Holt–Winters methods are the most appropriate methods for the long-term heat load forecasting«. Both hypotheses represent an important contribution to science. According to MSE (Mean Squared Error), MAPE (Mean Absolute Percentage Error), MARNE (Mean Absolute Range Normalized Error) forecasting measure and Theil's U-statistics, the first hypothesis confirms the results [4], which are used in decision making process for optimal daily operations and optimal utilization of resources and capabilities. The second hypothesis represents a novelty, which is very much related to long-term planning and optimal replenishment.

* Corresponding author.

E-mail address: liljana.ferbar@ef.uni-lj.si (L. Ferbar Tratar).

The Naive method has been chosen as a reference method, as suggested in Ref. [5]. Forecasting values have been compared to forecasting values of other methods. The Multiple regression method belongs to the group of more advanced methods. It's based on the consideration of the external factors. Dotzauer [6] figured out that the outdoor temperature affects the heat load the most. Other potential relevant external factors are solar radiation, wind speed and humidity [7]. The basic exponential smoothing methods have been developed by Holt [8] and Winters [9]. Extended exponential smoothing methods have been presented by the American professor Gardner [10]. One possible way of forecasting improvement represents relaxation of space parameter restrictions [11]. Hyndman, Akram in Archibald [12] mentioned, recommended values of space parameters are dependent on time series characteristic such as error, trend and seasonality. Multiple regression has been recognized as one of the most robust and reliable method for short-term heat load forecasting [4], although slightly better forecasting performance has been obtained by Neural network based model. As the Neural network based models are not very stable, robust and reliable [13], we didn't include them into our research. There has been increasing evidence that the accuracy of Neural network methods is comparable to existing methods, although in standard time series comparisons their performance still remains disappointing [14]. Our purpose was to investigate robust and reliable forecasting methods, which don't required special knowledge and special software such as neural network approach.

The remainder of this paper is structured as follows. Section 2 presents explanation of the methodology. Forecasting methods, included in our research activities, are described in the Section 3. Section 4 presents a case study and the data applied in our research. Research findings are presented in the end of the section. Section 5 consists of conclusion and ideas for the further research.

2. Methodology

Data was divided into two sets: training data set (WS1-Winter Season 1, WS2, WS3, WS4) and testing data set (WS5). In testing data set we had data for 123 days, it meant 18 weeks or 4 months. We calculated forecasting values for WS5 and then later compare calculated forecasting values to independent-real data in WS5. In case of short-term forecasting approach (Eq. (1)), we calculated forecasting value for one day ($m = 1$) ahead (daily short-term forecasting), one week ($m = 1$) ahead (weekly short-term forecasting) and one month ($m = 1$) ahead (monthly short-term forecasting):

$$F_{t+m}, m = 1 \quad (1)$$

where t represents an arbitrary time point in winter season WS5.

In the case of long-term forecasting approach (Eq. (2)), we calculated forecasting value for one hundred and twenty-three days ($m = 123$) ahead (daily long-term forecasting), eighteen weeks ($m = 18$) ahead (weekly long-term forecasting) and four ($m = 4$) months ahead (monthly long-term forecasting).

$$\begin{aligned} F_{t+m}, m = 1, 2, 3 \dots 123 & \text{(daily forecasting)} \\ F_{t+m}, m = 1, 2, 3, \dots 18 & \text{(weekly forecasting)} \\ F_{t+m}, m = 1, 2, 3, 4 & \text{(monthly forecasting)} \end{aligned} \quad (2)$$

where t represents last day/week/month in the winter season WS4.

For the evaluation of the forecasting methods we applied the following forecasting accuracy measures: MSE (Mean Squared Error), MAPE (Mean Absolute Percentage Error), MARNE (Mean Absolute Range Normalized Error) and Theil's U-statistics.

The most common forecasting measure is MSE:

$$MSE = \frac{1}{N} \sum_{t=1}^N (Y_t - F_t)^2, \quad t = 1, 2, 3, \dots N \quad (3)$$

where Y_t represents actual value, F_t forecasted value and N number of samples.

MAPE is a relative forecasting accuracy measure and it is scale independent measures. It's defined as:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| 100 \frac{Y_t - F_t}{Y_t} \right| [\%], \quad t = 1, 2, 3, \dots N \quad (4)$$

MARNE error was calculated as the average of the absolute differences of forecasted heat consumption F_t and actual heat consumption Y_t , normalized by the maximum transmission capacity of the district heating network Y_{max} [4], as it is shown in the following Eq. (5):

$$MARNE = 100 \frac{\frac{1}{N} \sum_{t=1}^N |F_t - Y_t|}{Y_{max}} [\%], \quad t = 1, 2, 3, \dots N \quad (5)$$

where N is the number of heat load samples in a monthly/weekly/daily resolution. MARNE error allowed us the comparison between actual and forecasted values among all samples of time series. If we compare error MARNE and error MAPE, MARNE hasn't problems with stability and divergence, if values are around zero. From the consumer perspective, there is no difference between lower and higher outdoor temperatures. Therefore we try to keep the same importance of error through the range Y . We have to be aware, that error MAPE penalized absolute error depends on value Y in specific time. Let's assume that we have the same absolute error E_1 and E_2 in both cases. If value Y_1 is lower, error $MAPE_1$ is bigger and if value Y_2 is higher, error $MAPE_2$ is smaller. Error MARNE is a relative measure depending on the size of the district heating system, and can be easily interpreted in technical or economical terms.

Theil's U-statistics has been used as the additional accuracy measure. It has been calculated through actual heat load values Y_t , Y_{t+1} and forecasted heat load values F_{t+1} as it is shown in the following Eq. (6):

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{F_{t+1} - Y_{t+1}}{Y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{Y_{t+1} - Y_t}{Y_t} \right)^2}} \quad (6)$$

If the value of Theil's U-statistics is higher than 1, it means that the forecasting technique is worse than the naïve method approach. If the value of Theil's U-statistics is lower than 1, it means that our technique is better than naïve method technique.

At this point, we would like to emphasize that lower values of the accuracy measures MSE, MAPE, MARNE, and Theil's U-statistics, represent a better forecasting performance. Testing accuracy measures was accepted as the main criterion for the forecasting performance of methods.

3. Forecasting methods

The Naive method was chosen as the reference method. A detailed forecasting performance analysis has been conducted for Multiple regression methods and exponential smoothing methods.

3.1. Naive method

A Naive method represents a benchmark method, and it is one of the most simple forecasting methods, it was applied as a

Download English Version:

<https://daneshyari.com/en/article/8073454>

Download Persian Version:

<https://daneshyari.com/article/8073454>

[Daneshyari.com](https://daneshyari.com)