Middle-Term Electric Load Forecasting by Time Series Decomposition

Emil Pelikán

Institute of Computer Science, Pod Vodárenskou věží 2, 182 07 Prague 8
Czech Republic, email: emil@cs.cas.cz

Abstract. We present the results of our work in the frame of World-Wide Competition organized within the EUNITE network. For the middle-term electric load forecasting task with prediction horizon 1-31 days ahead we decided to develop a simple model based on decomposition of analyzed time series. The additive components were estimated using the training data collected by the Eastern Slovakian Electricity Corporation during the period 1997-1998. The model was validated by using of the real data from the first 20 days in December 1998. For final application the model was retrained and the requested forecasts were performed. The temperatures in January 1999, which form the inputs to the model, were estimated by moving average methods applied on historical data sets.

1 Introduction

Reliable and more accuracy forecast of time series in economy, energy, environment, transport and other sectors with forecasting horizon from hours up to several months needs to develop sophisticated prediction models. These models should process information coming from different sources, as are historical databases, expert knowledge, weather forecast reports, economical and sociological prognoses, internet etc. It is not a rule, that very complicated models, e.g. based on neural networks, can overcome simple models, based on easy understandable principles.

In this contribution we present our approach to the World-Wide Competition task organized within the EUNITE network. The task was to forecast the electric load consumption (daily maximum) in the Eastern Slovakia for the 31 days in January 1999. The historical load data (half-hourly) were available for the period 1997-1998 and daily averaged temperatures were available for the period 1995-1998. Other meteorological variables were not available.

We decided to use the model based on time series decomposition. The time series is decomposed to individual components which can extracted to the future. This approach was applied to electric load forecasting by several authors, for example, Bunn [1], Pelikán [6], Petrák [7] and Száthmary [8].

2 Time series decomposition (TSD) models

The general mathematical description of a decomposition approach is

$$L(t) = f(L_{trend}(t), L_{season}(t), L_{temp}(tmp), L_{base}(t), L_{irr}(t))$$

where

L(t) is a daily maximum of the load at time t,
L_{trend} *(t)* is a trend component at time t,
L_{season} *(t)* is a seasonal component at time t,
L_{temp} *(tmp) is a temperature component for the daily averaged temperature tmp,*
L_{base} *(t) is a baseload component at time t,*
L_{irr} *(t) is an irregular component at time t.*

An additive form of the previous equation is

$$L(t) = L_{trend}(t) + L_{season}(t) + L_{temp}(tmp) + L_{base}(t) + L_{irr}(t) .$$

In our model we suppose that the individual components have the following form:

Trend component:

$$L_{trend}(t) = \sum_{i=0}^{Nt} p_i t^i$$

Seasonal component:

$$L_{season}(t) = \sum_{k=1}^{Ns} [d_k \cos(\tfrac{2\pi}{365} kt) + e_k \sin(\tfrac{2\pi}{365} kt)]$$

Temperature component:

$$L_{temp}(tmp) = \sum_{i=0}^{Ntmp} b_i tmp^i$$

Baseload component:

$$L_{base}(t) = \sum_{i=1}^{Nb} a_i I_i(t),$$

where

$I_i(d)$ is a a binary 1/0 variable indicating the type of day (Monday-Saturday,
National holidays, Christmas etc.),
a_i, d_i, e_i, b_i, p_i are real parameters of the model,

Nt,Nb,Ntmp are integer parameters of the model.

3 Experiments

3.1 Data preprocessing

We tried to estimate the parameters of the time decomposition model (TSD) using the real load data collected by the Eastern Slovakian Electricity Corporation from the period January 1997-December 1998. Because the target forecasts are focused to the January 1999, we decided to use data from the "winter months" October-April only. The length of the training period is limited to two-years period only, therefore it would be difficult to estimate correctly the influence of the special days (holidays, Christmas period, etc.). Such cases should be processed with help of experts or manually by analyzing of historical sets. Therefore we excluded also these non-typical cases from our training set. Also some days with unusual load values, e.g. November 11, 1997 or February 1, 1998, were not taken into account. So, finally we decided to train our model using preprocessed data till the end of November 1998. We validate the model in the period December 1-20, 1998.

3.2 TSD model specification

With respect to the previous data preprocessing step we set the number of different types of days to *Nb=7* (Mo,Tu,We,Th,Fr,Sa,Su).

Although the relationship between the temperature and load consumption can be generally nonlinear, we supposed a linear form with *Ntemp=1* (we do not use the cases from the "summer period").

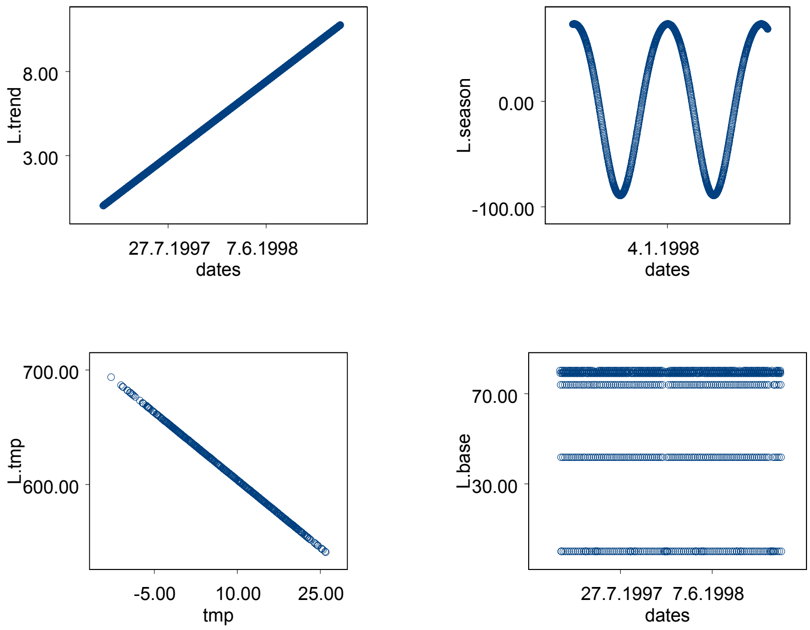
We know, that there is some inertia in the influence of the temperature to the load. Therefore we use the exponential smoothing (see e.g. [1],[2],[8]) applied to the daily averages temperatures *tmp* calculated by the formula

$$T_{smooth}(t) = \alpha * tmp(t) + (1-\alpha) * T_{smooth}(t-1)$$

We founded $\alpha = 0.8$ to be efficient value for daily maximum value forecasts. For seasonal component we use *Ns=3*.

3.3 Model results

The individual components estimated from the load time series form the period January97-April97,October97-April98 and October98-November98 are shown in Fig. 1. From this figure we can see a small positive trend (about 5MW per year), the behavior of periodical seasonal component, negative temperature dependency with temperature gradient about -4 MW/°C and baseload lines corresponding to the coeffi-



cients a_i for weekend days (Sunday -bottom line, Saturday -middle line) and working days (top lines). The irregular component part is shown in Fig.2.

Fig. 1. Trend, seasonal, temperature and baseload components estimated using the "winter data" from the period January97-November98.

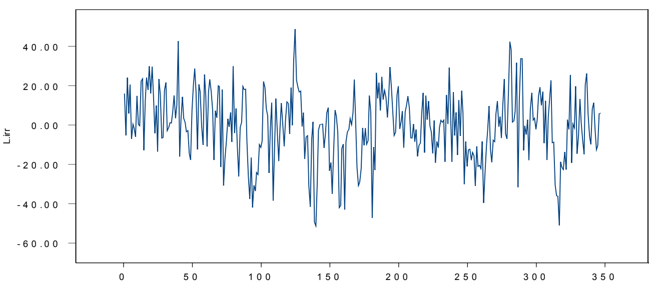


Fig. 2. Irregular component estimated using the "winter data" from the period January97-November98.

The real electric load (daily maximum) together with the developed DTS model forecasts during the first 20 days in December 1998 is shown in Fig.3. The MAPE (mean averaged percentage error) of the DSP model is shown in Table I.

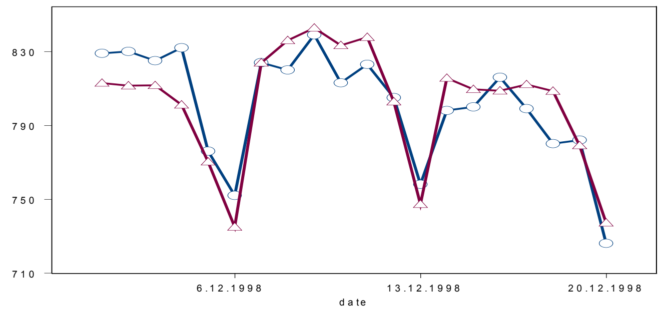


Fig. 3. Daily maximum electric load (circle) and its forecast (triangle) using the TSD model with real temperatures from the period December 1-December 20,1998.

3.4 Uncertainty analysis

The measure of uncertainty of TSD model can be estimated by analysis of the error process derived from the irregular component. The histogram of the errors is shown in Fig. 4. The standard deviation is approximately 17.5MW. This value can be used for setting the expected predicting intervals of our model. We can still improve our TSD model. The irregular part is not a white noise. In Fig. 5 the corresponding autocorrelation function is shown. From this figure we can detect "a short memory" of the error process up to the lag 5. It means, that the TSD model can be improved by adding the autoregressive ar ARMA filter for prediction of the irregular values. For example, if we estimate the AR(2) filter, we obtain two AR coefficients *ar1=0.35* and *ar2=0.15.* The standard deviation of the error process is reduced to 15.8MW. The MAPE of the improved model using this AR(2) filter is shown in Table I.

This short-term filtering cannot be used for the competition task, because the load from last two days is necessary for the prediction. This analysis is performed for demonstration purpose only.

Table 1. Mean percentage error (MAPE) of TSD model and TSD model with filtered irregular component

MAPE- TSD model	MAPE - TSD model with AR (2) filter
1.62%	1.58%

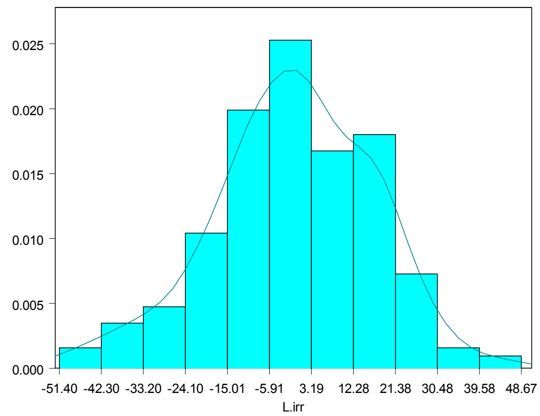


Fig. 4. Histogram of errors derived from the irregular component of the TSD model.

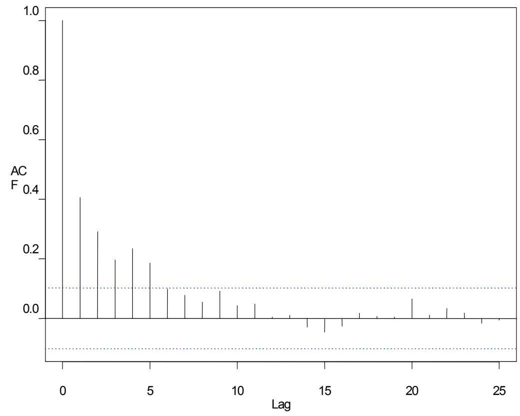


Fig. 5. Autocorrelation function of the irregular component.

4 Competition task

The developed TSD model needs the temperatures as an input. Because the temperatures for the January 1999 were not at disposal, we had to forecast them. The temperature forecast is a task for weather prediction companies and it is very difficult to have reliable values for several weeks ahead. We decided simply to use a moving averaged values calculated from the corresponding months in 1997-1998. The length of the window was selected to 14 days in the first scenario A and 6 days for the second scenario B. The forecasted temperature was located next to the center of the given window. For example, the forecasted temperature for January 15, 1999 was calculated as a mean of temperatures from January 8 to January 21in the case A and from January 12 to January 17 in the case B. There is no apriori reason to prefer 6 or 14 days, other number can be used as well. We use this two window lengths to demonstrate different scenarios, which can be used for simulation purposes by dispatchers. The graph of the forecasted temperatures are shown in Fig. 6 and the corresponding TSD load forecast for the scenario A is shown in Fig. 7. The numerical values of the forecasts are summarized in Table 2. The load for the first 7 days were manually corrected with respect to the national holidays in the middle of the first week in January 1999.

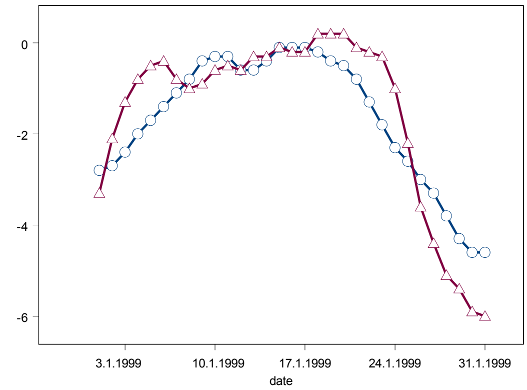


Fig. 6. Temperature forecasts for January 1999 using the moving average window with length 14 (circles) and 6 (triangles) .

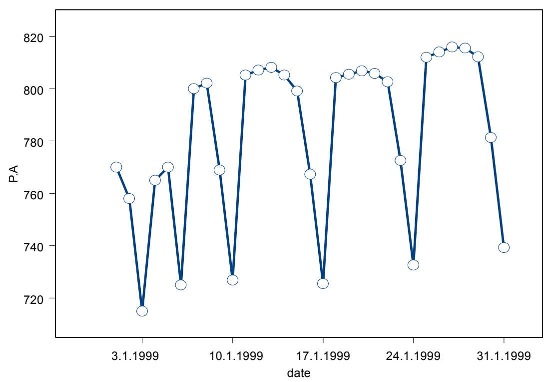


Fig. 7. The load predictions using the temperature forecasts (scenario A) for the period January 1999.

Table 2. Temperature and load forecasts for the 31 days in January 1999.

Date	Temperature (scenario A)	Load forecast	Temperature (scenario B)	Load forecast
1.1.1999	-2.8	770.0	-3.3	780.0
2.1.1999	-2.7	758.0	-2.1	755.0
3.1.1999	-2.4	715.0	-1.3	712.0
4.1.1999	-2.0	765.0	-0.8	760.0
5.1.1999	-1.7	770.0	-0.5	760.0
6.1.1999	-1.4	725.0	-0.4	722.0
7.1.1999	-1.1	800.0	-0.8	798.0
8.1.1999	-0.8	802.1	-1.0	802.9
9.1.1999	-0.4	768.9	-0.9	770.8
10.1.1999	-0.3	726.9	-0.6	728.0
11.1.1999	-0.3	805.2	-0.5	806.0
12.1.1999	-0.6	807.1	-0.6	807.1
13.1.1999	-0.6	808.1	-0.3	806.9
14.1.1999	-0.4	805.2	-0.3	804.9
15.1.1999	-0.1	799.1	-0.1	799.1
16.1.1999	-0.1	767.3	-0.2	767.6
17.1.1999	-0.1	725.5	-0.2	725.9
18.1.1999	-0.2	804.2	0.2	802.7
19.1.1999	-0.4	805.5	0.2	803.2
20.1.1999	-0.5	806.8	0.2	804.1
21.1.1999	-0.8	805.8	-0.1	803.1
22.1.1999	-1.3	802.6	-0.2	798.3
23.1.1999	-1.8	772.6	-0.3	766.8
24.1.1999	-2.3	732.6	-1.0	727.7
25.1.1999	-2.6	812.0	-2.2	810.4
26.1.1999	-3.0	814.0	-3.6	816.3
27.1.1999	-3.3	815.9	-4.4	820.1
28.1.1999	-3.8	815.5	-5.1	820.5
29.1.1999	-4.3	812.2	-5.4	816.4
30.1.1999	-4.6	781.3	-5.9	786.3
31.1.1999	-4.6	739.3	-6.0	744.7

5 Conclusion

The developed model based on time series decomposition belongs to a family of simple and easy understandable models, which can be used for middle-term electric load forecasting. The application of the model, which was presented in this contribution, was modified with respect to the length of load series (two years only) and to the available meteorological values (daily averaged temperatures only).

This restriction is reasonable for two reasons. Firstly, the economy in Slovakia is very changing (similarly to the economic situation in the Czech republic) and longer time series could be non-stationary. Secondly, the additional meteorological variables, which can help to forecasters (e.g. wind speed, cloud cover) are difficult to predict for longer horizon than for several days. But, if the short-term electric load forecast is asked, these meteorological variables would be very useful, especially for extreme values.

The application of the model is not limited to a linear form of the components (linear in parameters), but can be also extended to nonlinear ones. The quality of presented results will be strongly dependent on forecasted temperatures. Therefore we presented two scenarios, which can be extended using information from meteorological institutes. Such results, using different scenarios can help dispatchers in their decisions.

References

1. Bunn, D.W., Farmer E.D. : Comparative models for Electric Load Forecasting, New York, John Willey 1985
2. Makridakis S, Wheelwright S.C., Hyndman R.J. : Forecasting: Methods and Applications, Third edition, John Willey & Sons,1998
3. Pelikán E. a kol.: Interactive forecasting of the electric load using Kohonen self-organizing feature maps , Proc. of the World Congress of Neural Networks, San Diego, California, 1996, pp.443-446
4. Pelikán E., Petrák L. : Predikce spotřeby elektrické energie s ohledem na vývoj počasí, sborník semináře Aktuální otázky energetiky, EGU, Poděbrady 12-13.11.1998
5. Pelikán E., Matějka P., Sláma M., Novák M., Beran H., Vinkler K., Lukáš J.: Střednědobá predikce spotřeby elektrické energie pomocí Kohonenových map, Proceedings of the International Conference "Artificial Neural Networks and their Application Possibilities, ISBN 80-967593-1-0, Košice 18-19.11.1996, pp.45-49
6. Pelikán E.: Support Decision Forecasting Systems, (in Czech) Proceedings of the 2nd Prediction workshop NOSTRADAMUS´98, ISBN 80-214-1424-3, Zlin, October 7-8, 1999, pp. 76-80
7. Petrák L.: Dispečerské řízení odběru elektřiny v distribuční společnosti, doktorandská disertační práce, katedra elektroenergetiky FEL, Plzeň, 1998
8. Szathmáry,P., Kolcun M.: Predikcia denných diagramov zatiaženia v ES s využitím umelých neurónových sietí, edicia vedeckých spisov FEI TU Košice, ISBN 80-88964-80-6, 2001