View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

International Conference
"The European Electricity Market EEM-04"
September 20-22, 2004, Lodz, Poland

Proceedings Volume, pp. 135-142

MODELING AND FORECASTING ELECTRICITY LOADS: A COMPARISON

Rafał WERON
Hugo Steinhaus Center
for Stochastic Methods
Wrocław University of Technology
Wrocław (Poland)

Adam MISIOREK[1]
Institute of Power Systems Automation
Wroclaw (Poland)

Abstract

In this paper we study two statistical approaches to load forecasting. Both of them model electricity load as a sum of two components – a deterministic (representing seasonalities) and a stochastic (representing noise). They differ in the choice of the seasonality reduction method. Model A utilizes differencing, while Model B uses a recently developed seasonal volatility technique. In both models the stochastic component is described by an ARMA time series. Models are tested on a time series of system-wide loads from the California power market and compared with the official forecast of the California System Operator (CAISO).

1. INTRODUCTION

The forecasting of energy demand has become one of the major fields of research in electrical engineering. The power industry requires forecasts with lead times that range from the short term (a few minutes, hours or days ahead) to the long term (up to 20 years ahead). Short-term forecasts, in particular, have become increasingly important since the rise of the competitive energy markets.

Load forecasting is vital to the whole power industry, however, it is a difficult task. Firstly, because the load time series exhibit seasonality – at the daily, weekly and annual timescales. Secondly, because there are many exogenous variables that may be considered, with weather conditions being the most influential. It is relatively easy to get short-term forecasts with a few percent error, however, the financial costs of the error are so high that research is aimed at reducing it even by a fraction of a percentage point.

Most forecasting models and methods have already been tried out on load forecasting, with varying degrees of success. They may be classified into two broad categories: artificial intelligence based techniques and classical (or statistical) approaches. The former include expert systems, fuzzy inference, fuzzy neural models, and – in particular – artificial neural networks (ANN). In the 1990s much research has been carried out in this area. Nevertheless, the reports on the performance of ANNs in forecasting have not entirely convinced the researchers and the practitioners alike and the skepticism may be partly justified. Recent reviews and textbooks on forecasting argue that there is little evidence as yet that ANNs might outperform standard forecasting methods [7,9]. Reviews of ANN based forecasting systems have concluded that much work still needs to be done before they are accepted as established forecasting techniques and that they are promising but that "significant portion of the ANN research in forecasting and prediction lacks validity". Two major shortcomings were found to detract from the credibility of the results: the proposed ANN architectures were too large for the data at hand (the ANNs apparently overfitted the data) and the models were not systematically tested [7].

The statistical methods differ from the previous approach in that they forecast the current value of a variable by using an explicit mathematical combination of the previous values of that variable and, possibly, previous values of exogenous factors (specially weather and social variables). Models that have been applied recently include autoregressive (AR) models, linear regression models, dynamic linear or nonlinear models, ARMAX models, threshold AR models, methods based on Kalman filtering, optimization techniques, and curve fitting procedures. The statistical models are attractive because some physical interpretation may be attached to their components, allowing engineers and system operators to understand their behavior. At the same time they offer relatively good performance [5,11,13,14].

In this paper we evaluate two statistical approaches. Both of them model electricity load as a sum of two components – a deterministic (representing seasonalities) and a stochastic (representing noise). They differ in the choice of the seasonality reduction method. Model A utilizes differencing, while Model B uses a recently developed seasonal volatility technique [16]. In both models the stochastic component is described by an ARMA time series.

2. ANALYZED DATA

The analyzed time series of system-wide loads was constructed using data obtained from the California's Independent System Operator (CAISO, http://oasis.caiso.com). The models are calibrated to data from the period January 1, 1999 – December 31st, 2000, i.e. two full years. The following year is used for out of sample testing of the models. Recall, that it includes the period of soaring prices and San Francisco blackouts of January 2001. This rough period was selected because we wanted to "stress test" the models.

Due to a very strong daily cycle that we did not want to address in this paper, we created a 1096 days long sequence of daily loads, see Fig. 1. Apart from the daily cycle, the time series displays weekly and annual seasonality, which has to be removed before time series models can be fitted to the stochastic part.

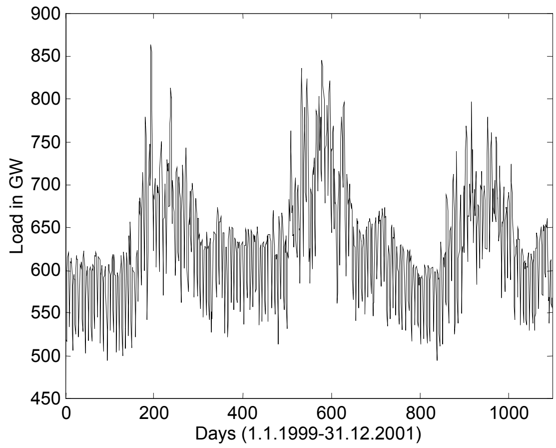

FIGURE 1. Daily system-wide loads in California (1999-2001)

3. SEASONALITY REDUCTION

3.1. Model A

As we have mentioned we model the electricity load Z_t as a sum of two components – a deterministic (or seasonal) X_t and a stochastic Y_t. The former can be treated in several ways. Probably the simplest one is to use differencing, which consists of subtracting selected previous load values from the current load [2]. In Model A we utilize the following formula:

$$Y_t = Z_t - X_t$$

where

$$X_t = \frac{1}{N}\sum_{i=1}^{N} Z_{t-iT} + \frac{1}{M}\sum_{j=1}^{M} Z_{t-jD} - \\ - \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} Z_{t-iT-jD}$$

T = 168 hours (i.e. one week), D = 24 hour (i.e. one day), N = 4 or 5 is the number of weeks used for calibration, and M = 7 is the number of days in a week. This procedure yields an approximately stationary sequence Y_t; the KPSS test [10] does not reject the hypothesis of stationarity at the 5% level. This allows us to model the residuals (i.e. the stochastic part Y_t) by an ARMA process.

3.2. Model B

The differencing technique applied in Model A has the disadvantage of being very sensitive to the load observed in the proceeding days or weeks. An alternative approach – which does not possess this deficiency – consists of fitting, typically via a non-linear least-squares routine, a sum of sine (or cosine) waves having different amplitudes, frequencies and/or phase angles [12]. However, in our case the daily data spans only a few years and no significant change of amplitude can be observed. Furthermore, as can be seen in Fig. 1, it is highly non-sinusoidal – it is rather flat throughout the year with a substantial hump in late summer and autumn. Because common trend and seasonality removal techniques (like the moving average algorithm, see below) do not work well when the time series is only a few cycles long, in Model B we applied a new seasonality reduction technique [16] to the annual cycle.

To remove the weekly cycle we used the moving average technique [4, p.30]. For the vector of daily loads $\{Z_1,\ldots Z_{731}\}$ the trend was first estimated by applying a moving average filter specially chosen to eliminate the weekly component and to dampen the noise:

$$\hat{m}_t = \tfrac{1}{7}(Z_{t-3} + \cdots + Z_{t+3})$$

where t = 4, …, 728. Next, we estimated the seasonal component. For each k = 1,…, 7, the average w_k of the deviations $\{(Z_{k+7j} - \hat{m}_{k+7j}), 4 \le k+7j \le 728\}$ was computed. Since these average deviations do not necessarily sum to zero, we estimated the seasonal component $\hat{s}_k$ as

$$\hat{s}_k = w_k - \tfrac{1}{7}\sum_{i=1}^{7} w_i$$

where k = 1, …, 7 and $\hat{s}_k = \hat{s}_{k-7}$ for k>7. The deseasonalized (with respect to the 7-day cycle) data was then defined as $d_t = Z_t - \hat{s}_t$ for t = 1, …, 731. Finally, we removed the trend from the deseasonalized data $\{d_t\}$ by taking logarithmic returns $r_t = \log(d_{t+1}/d_t)$, t = 1, … , 730.

After removing the weekly seasonality we were left with the annual cycle. Unfortunately, because of the short length of the time series (only 2 years), the method applied to the 7-day cycle could not be used to remove the annual seasonality. To overcome this we applied a new method which consists of the following [16]:

(i) calculate a 25-dayrolling volatility [8]

$$v_t = \sqrt{\tfrac{1}{24}\sum_{i=0}^{24}(R_{t+i} - \overline{R}_t)^2}$$

for t = 1, … , 730 and a vector of returns $\{R_t\}$ such that $R_1 = R_2 = \ldots = R_{12} = r_1$, $R_{12+t} = r_t$ for t = 1, … , 730, and $R_{743} = R_{744} = \ldots = R_{754} = r_{730}$;

(ii) calculate the average volatility for 1 year, i.e. in our case $\overline{v}_t = \dfrac{v_t^{1999} + v_t^{2000}}{2}$;

(iii) smooth the volatility by taking a 25-day moving average of $\overline{v}_t$;

(iv) finally, rescale the returns by dividing them by the smoothed annual volatility.

The obtained time series showed no apparent trend and seasonality (for details see [11]). Therefore, we treated it as a realization of a stationary process. Moreover, the dependence structure exhibited only short-range correlations. Both, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) rapidly tended to zero, which suggested that the deseasonalized load returns could be modeled by an ARMA-type process.

4. MODELING THE STOCHASTIC COMPONENT

The mean-corrected (i.e. after removing the sample mean) deseasonalized load returns were modeled by ARMA processes, i.e. processes of the form:

$$Y_t - \phi_1 Y_{t-1} - \ldots - \phi_p Y_{t-p} = \\ = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$$

where (p,q) denote the order of the model and $\{\varepsilon_t\}$ is a sequence of independent, identically distributed variables with mean 0 and finite variance σ^2 (denoted by $\mathrm{iid}(0;\sigma^2)$ in the text). In both models the maximum likelihood (ML) estimators were used to obtain estimates of $(\phi_1,\ldots,\phi_p,\theta_1,\ldots,\theta_q,\sigma^2)$. The ML estimators used here are based on the assumption of Gaussian noise $\{\varepsilon_t\}$. However, this does not exclude models with non-Gaussian noise since the large sample distribution of the estimators is the same for $\{\varepsilon_t\} \sim \mathrm{iid}(0;\sigma^2)$ regardless of whether or not $\{\varepsilon_t\}$ is Gaussian (see [3, Section 10.8]).

In Model A the model size $(\mathrm{p},\mathrm{q}) = (3,2)$ was selected based on the SAS implementation of the Extended Sample Autocorrelation Function (ESACF) and the Smallest CANonical (SCAN) correlation method that can tentatively identify the orders of a stationary or nonstationary ARMA process [6,15]. In Model B the parameter estimates and the model size $(\mathrm{p},\mathrm{q}) = (1,6)$ were selected to be those that minimize the bias-corrected version of the Akaike criterion, i.e. the AICC statistics, see [4, Section 5.5]. This time the ITSM and Matlab software was used, however, analogous results were also obtained with SAS.

After calibrating ARMA processes we had to test their residuals. As it turned out, in both cases, there was not sufficient evidence to reject the i.i.d. hypothesis of the residuals at the common 5% level, however, the distribution was not Gaussian but heavy tailed. The hyperbolic distribution [1] gave a very good fit to Model B residuals (for details see [11]).

5. FORECASTING PERFORMANCE

It is not so surprising that a model will perform well when evaluated by its fit to the data set to which it was adjusted. The real test is whether it will be capable of also describing new data sets coming from the same process. A suggestive and attractive way of comparing different models is to evaluate their performance when applied to a data set to which none of them was adjusted. The standard measure of goodness of fit is the difference between actual and forecasted outputs. The disadvantage of this method is that we have to save part of the data set (in our analysis – one year) for the comparisons and therefore cannot use all available information to build the model.

In the previous section we fitted ARMA models to the stochastic (or deseasonalized) components of the system-wide load from the period January 1, 1999 to December 31, 2000. Now, we test the performance of the model on data from the following year, i.e. from the period January 1 to December 31, 2001. For every day in the test period we run a day-ahead prediction. We apply an adaptive scheme, i.e. instead of using a single model for the whole sample, for every day in the test period we calibrate the best ARMA(3,2) and ARMA(1,6) model (for Model A and Model B, respectively) to the previous 730 values of the stochastic component and obtain a forecasted value for that day. The results are then "inverted" (the seasonality is added) and compared with the actual system-wide loads and the CAISO official day-ahead forecasts. The performance of the models can be observed in Figs. 2-7 and is summarized in Table 1 and Fig. 8.

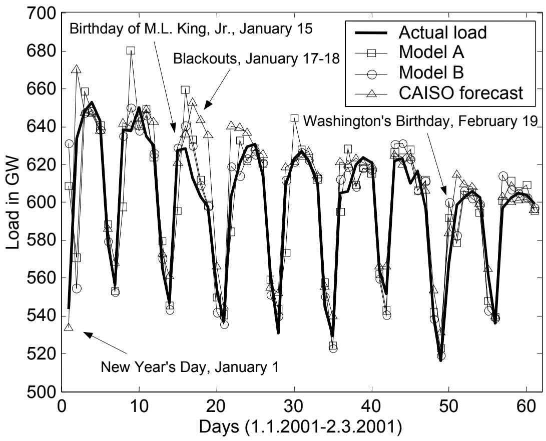

FIGURE 2. Daily system-wide loads in California (Jan.-Feb. 2001) compared with Model A, Model B and CAISO day-ahead forecasts.

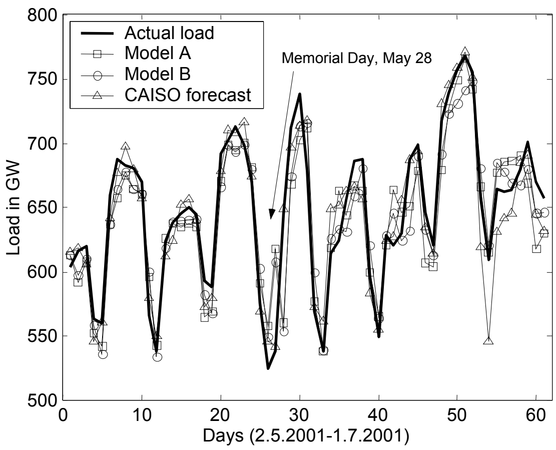

FIGURE 3. Daily system-wide loads in California (May-Jun. 2001) compared with Model A, Model B and CAISO day-ahead forecasts.

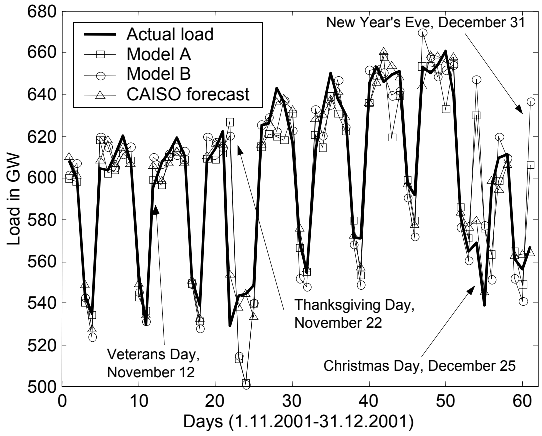

FIGURE 4. Daily system-wide loads in California (Nov.-Dec. 2001) compared with Model A, Model B and CAISO day-ahead forecasts.

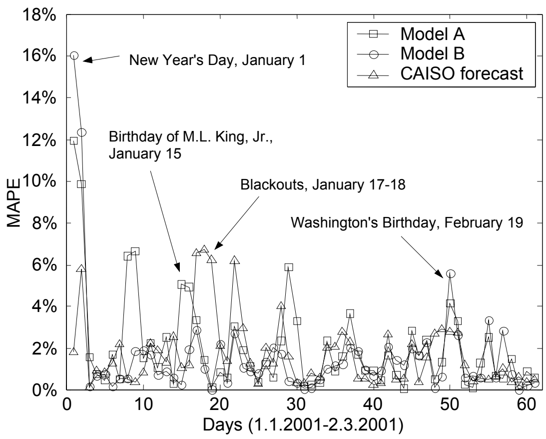

FIGURE 5. Mean Absolute Percentage Error (MAPE) of the day-ahead Model A, Model B and CAISO forecasts (Jan.-Feb. 2001).

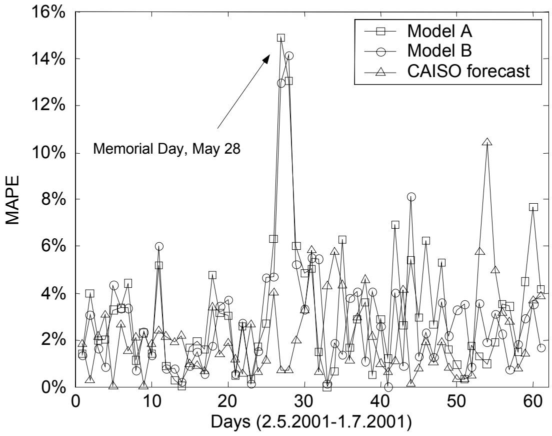

FIGURE 6. Mean Absolute Percentage Error (MAPE) of the day-ahead Model A, Model B and CAISO forecasts (May-Jun. 2001).

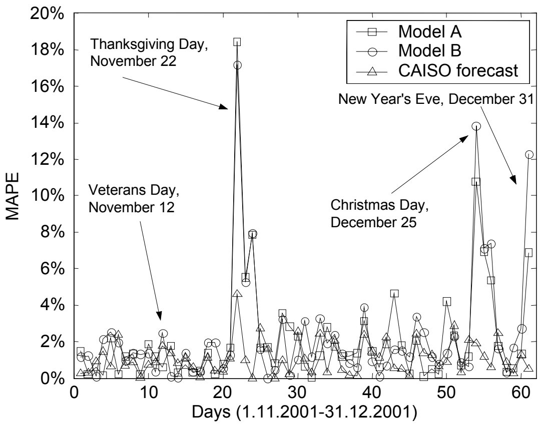

FIGURE 7. Mean Absolute Percentage Error (MAPE) of the day-ahead Model A, Model B and CAISO forecasts (Nov.-Dec. 2001).

TABLE 1. Mean Absolute Percentage Errors (MAPE) of the day-ahead Model A, Model B and CAISO forecasts for the whole year 2001. Best results are emphasized in bold.

MAPE	CAISO	Model A	Model B
With holidays	**1.84%**	2.30%	2.08%
Without holidays	1.77%	1.95%	**1.71%**

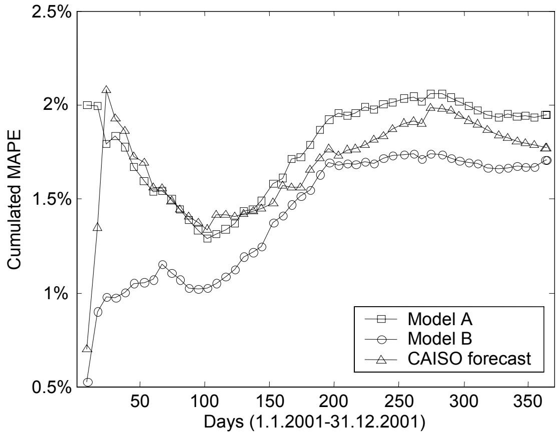

FIGURE 8. Cumulated Mean Absolute Percentage Error (MAPE) of the day-ahead Model A, Model B and CAISO forecasts for data with US national holidays excluded.

Looking at the Mean Absolute Percentage Error ($\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{x}_i - x_i\right| \times 100\%$) values for the whole test period we can observe that the CAISO forecast outperforms our models. Note, however, that the extreme differences between the actual load and Models A and B correspond to the US national holidays – New Year's Day (lag 1 in Fig. 5), Washington's Birthday (lag 50 in Fig. 5), etc. Obviously, our models cannot capture the holiday structure. Fortunately, this can be quite easily incorporated into them by simply subtracting a certain amount of GW for these holidays based on previous years' experience. When we compare the forecasting results for the same period but with US national holidays excluded our models perform much better. The results improve even more if we eliminate some of the days directly proceeding or following the holidays – in the bottom row of Table 1 and in Fig. 8 we additionally exclude four days with an abnormal consumption pattern (Jan. 2, May 29, Sept. 4, and Dec. 31). After the exclusions Model A performs significantly better (but still worse than the CAISO forecast) and Model B outperforms both Model A and the CAISO forecast. It is worth noting that Model B yields the smallest forecasting error throughout the whole year, see Fig. 8.

It is also quite surprising that both our models produce relatively small errors during the blackout days in San Francisco (see Fig. 5), while the CAISO forecast is at least twice worse around this time. In fact, as reported in [11], the MAPE of Model B is only 1.23% during the first two months of year 2001, compared to 1.71% for the CAISO forecast during this time period.

As we have already mentioned the differencing technique applied in Model A has the disadvantage of being very sensitive to the load observed in the proceeding days or weeks. This can be observed in Fig. 5 where the forecasting error of January 1 (lag 1) negatively influences the forecasts for January 8 (lag 8) and 15 (lag 15), i.e. a week and two weeks later.

6. CONCLUSIONS

Short-term load forecasting plays an important role in power system operation and planning. Accurate load prediction saves costs by improving economic load dispatching, unit commitment, etc. At the same time it enhances the function of security control. In this paper, we have evaluated two statistical approaches. Both of them model electricity load as a sum of two components – a deterministic (representing seasonalities) and a stochastic (representing noise). They differ in the choice of the seasonality reduction method – Model A uses differencing, while Model B a new seasonality removal technique.

The models were successfully applied to real data. A comparison was made between both models and the official forecasts of the California Independent System Operator (CAISO). The effectiveness of the approaches was demonstrated through a comparison of the real load data with short-term forecasted values. In terms of MAPE Model B yielded a smaller error than the CAISO day-ahead forecast during the whole year 2001. While being simpler Model A returned a higher error.

We strongly believe that the approach implemented in Model B is a universal one and can be applied not only to the California power market system-wide load but also to other power market data sets displaying seasonalities. Moreover, even for the more burdensome Model B the computational times are negligible and the method can be used in real time forecasting.

7. REFERENCES

1. O.E. Barndorff-Nielsen (1977) Exponentially decreasing distributions for the logarithm of particle size, Proc. Roy. Soc. London A 353, 401–419.
2. M. Borgosz-Koczwara, M. Kozłowski, A. Misiorek, T. Piesiewicz (2001) Analiza metod prognozowania procesu zapotrzebowania na energię elektryczną, Energetyka 12/2001, 759-764.
3. P.J. Brockwell, R.A. Davis (1991) Time Series: Theory and Methods, 2nd Edition, Springer, New York.
4. P.J. Brockwell, R.A. Davis (1996) Introduction to Time Series and Forecasting, Springer, New York.
5. D. Bunn (2000) Forecasting loads and prices in competitive power markets, Proc. IEEE 88 (2), 163-169.
6. Choi, ByoungSeon (1992), *ARMA Model Identification*, New York: Springer-Verlag, 129-132.
7. H.S. Hippert, C.E. Pedreira, R.C. Souza (2001) Neural networks for short-term load forecasting: a review and evaluation IEEE Trans. Power Systems 16 (1), 44–55.
8. V. Kaminski (1997) The challenge of pricing and risk managing electricity derivatives, in: P. Barber, ed., The US Power Market, Risk Books, London.
9. S. Makridakis, S.C. Wheelwright, R.J. Hyndman (1998) Forecasting – methods and applications, 3rd ed., Wiley, New York.
10. Misiorek, T. Piesiewicz (2002) Dekompozycja jako istotny etap prognozowania zapotrzebowania na energię elektryczną, Materiały IX Konf. Nauk.-Techn. RE-E'2002, Kazimierz Dolny, 13-15 maja 2002, tom II, 161-167.

11. J. Nowicka-Zagrajek, R. Weron (2002) Modeling electricity loads in California: ARMA models with hyperbolic noise, Signal Processing 82 (12), 1903-1915.
12. S.M. Pandit, S.M. Wu (1983) Time Series and System Analysis with Applications, Wiley, New York.
13. R. Sadownik, E.P. Barbosa (1999) Short-term forecasting of industrial electricity consumption in Brasil, J. Forecast. 18, 215–224.
14. M. Smith (2000) Modeling and short-term forecasting of New South Wales electricity system load, J. Bus. Econom. Statist. 18, 465–478.
15. R.S. Tsay, G.C. Tiao (1984) Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models, JASA 79 (385), 84-96.
16. R. Weron, B. Kozlowska, J. Nowicka-Zagrajek (2001) Modeling electricity loads in California: a continuous-time approach, Physica A 299, 344–350.

Adam Misiorek, M.Sc.
was born in 1975 in Kielce, Poland. He received the M.Sc. degree in applied mathematics from the Wroclaw University of Technology (WUT). Currently he is statistical analyst/programmer at Institute of Power Systems Automation (IASE) and a PhD student at the Institute of Mathematics at WUT. His areas of interest include forecasting and mathematical methods of risk management.

Mailing address:
Adam Misiorek
Institute of Power Systems Automation
1 Wystawowa Str., 51-618 Wroclaw
POLAND
Phone: (+48) (71) 348-42-21.
Fax: (+48) (71) 348-21-83.
e-mail: misiorek@iase.wroc.pl

Dr. Rafał Weron
received his M.Sc. (1995) and Ph.D. (1999) degrees in applied mathematics from the Wroclaw University of Technology (WUT), Poland. Currently he is Assistant Professor of Mathematics at WUT. His research focuses on risk management in the power markets and computational statistics as applied to finance and insurance. His professional experience includes development of insurance strategies for the Polish Power Grid Co. (PSE S.A.) and Hydro-storage Power Plants Co. (ESP S.A.) as well as implementation of option pricing software for LUKAS Bank S.A. (Credit Agricole Group). He has also been a consultant or executive teacher to a large number of banks and corporations.

Mailing address:
Rafal Weron
Hugo Steinhaus Center
Wroclaw University of Technology
Wyb. Wyspianskiego 27, 50-370 Wroclaw
POLAND
Phone: (+48) (71) 320-3530.
Fax: (+48) (71) 320-2654.
e-mail: rweron@im.pwr.wroc.pl

---

[1] Research partially supported by KBN grant no. 4 T10B 030 25.