# Modelling the Belgian Gas Consumption using Neural Networks *

J. Suykens,† Ph. Lemmerling, W. Favoreel and B. De Moor
*Katholieke Universiteit Leuven, Dept. of Electrical Engineering, ESAT-SISTA*
*Kardinaal Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium*

M. Crepel and P. Briol
*Electrabel, Regentlaan 8, B-1000 Brussel, Belgium*

(Received ..... ; Accepted in final form .....)

**Abstract.** In this letter an accurate neural network model is proposed for the gas consumption in Belgium. It is a static nonlinear model, based on monthly data and contains the following inputs: temperature, difference between real and expected temperature, oil price, number of domestic clients and consumption by industry. Various interpretations are made on the identified models such as yearly error, normalized gas consumption, growth rate, uncertain linear model interpretation and sensitivity of the consumption with respect to the temperature. In contrast to traditional models, that depend only on the temperature, the present neural network models show excellent generalization ability, with small yearly errors on training and test set.

**Key words:** multilayer perceptron, local optimization, nonlinear modelling, time series

* This research work was carried out at the ESAT laboratory and the Interdisciplinary Center of Neural Networks ICNN of the Katholieke Universiteit Leuven, in the framework of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture (IUAP-17) and in the framework of a Concerted Action Project MIPS (Modelbased Information Processing Systems) of the Flemish Community and is sponsored by Electrabel.

† Author for correspondence: J. Suykens

## 1. Introduction

Gas is one of the important energy sources for industry and domestic clients in Belgium. A good knowledge of the evolution of the gas consumption is needed in order to make investments and for a better exploitation of the energy system. Because gas is used for heating of the houses by domestic clients, the gas consumption is depending on the temperature. Fig.1 (left) shows the relation between gas consumption and temperature for daily data[1] for the period of July 1 1982 till December 31 1994. Traditionally, the following model has been assumed by Electrabel:

$$\hat{G}_k = a + b\,\varphi(T_k), \tag{1}$$

where $\varphi(.)$ is defined as $\varphi = 16.5 - T_k$ if $T_k \leq 16.5$ and $\varphi = 0$ if $T_k > 16.5$. The motivation for this affine relation between $G_k$ and $\varphi(T_k)$ is clear from Fig.1 (right). $T_k$ denotes the daily or mean monthly temperature, depending whether the time index $k$ refers to daily or monthly data. $\hat{G}_k$ is the estimated gas consumption. In the model (1), the constant $a$ is related to the consumption by industry and $b$ to domestic clients. In order to obtain a better fitting on the training data, a time-varying model

$$\hat{G}_k = a_k + b_k\,\varphi(T_k) \tag{2}$$

has been investigated in the past. The time variation of $a_k, b_k$ is explained then with respect to other external signals, such as economic influences, number of clients, abnormality of temperature, i.e. the difference between the real and expected temperature, etc.

However, the time-varying model (2) is not very well suited in order to make extrapolations, because the evolution of $a, b$ in time is unknown. Therefore a neural network approach is studied here. All variables of interest are taken into account then in the input of the neural network, instead of the temperature only. Neural networks have been successfully applied already in literature to the electric load forecasting problem [1, 8]. We show that also for modelling the gas consumption, excellent results can be obtained by neural networks. The errors on the training set as well as on the test set are small, which is not the case for the traditional model (1). We also show how normalized gas consumption and growth rate can be calculated from the identified neural network models.

This letter is organized as follows. In Section 2 the structure of the neural network is explained. In Section 3 interpretations are made for

[1] Throughout this paper the data have been scaled (linearly) for confidentiality reasons, required by Electrabel.

the identified models such as yearly error, normalized gas consumption, growth rate, uncertain linear model interpretation and sensitivity of the consumption with respect to the temperature. In Section 4 simulation results of the neural network models are shown.

## 2. Neural network model for Belgian gas consumption

In order to model the gas consumption in Belgium, the following static nonlinear model is proposed here:

$$\hat{G}_k = f(T_k, T_k - T_k^N, P_k^*, N_k, I_k), \tag{3}$$

where $k$ denotes the monthly discrete time index, $\hat{G}_k$ the estimated monthly gas consumption, $T_k$ the mean monthly temperature, $T_k^N$ the normal mean monthly temperature, $P_k^*$ a transformed oil price, $N_k$ the number of domestic clients and $I_k$ the gas consumption by industry. The so-called normal temperature corresponds to the expected value of the temperature, measured over the last 30 years. Hence the input $T_k - T_k^N$ is a measure for abnormality of $T_k$. The transformed oil price $P_k^*$ is obtained by inverting and integrating the crude oil price $P_k$, such that $P_k^*$ is the sampled data version of $P^*(t) = \int_0^t [1/P(\tau)] d\tau$. The reason for this transformation is explained in Section 4. Data for $N_k$ and $I_k$, which are available on yearly basis, have been interpolated linearly. A scaled version of the input data is shown on Fig.2.

The nonlinear function $f(.)$ in (3) is parametrized then by a multilayer perceptron with one hidden layer

$$\hat{G}_k = w^T \tanh(V x_k + \beta), \tag{4}$$

where the input vector of the multilayer perceptron is equal to $x_k = [T_k; T_k - T_k^N; P_k^*; N_k; I_k]$, the weight vector of the output layer is $w \in \mathcal{R}^{n_h}$, the interconnection matrix of the hidden layer is $V \in \mathcal{R}^{n_h \times 5}$ and the bias vector is $\beta \in \mathcal{R}^{n_h}$, with $n_h$ the number of hidden neurons. The activation function tanh(.) is applied componentwise to the elements of a vector. It is well known that multilayer perceptrons with one or more hidden layers are universal approximators in the sense that they can approximate any continuous nonlinear function arbitrarily well on a compact interval [2, 5, 6].

In order to train the network (4), the following cost function is minimized:

$$\min_{w,V,\beta} \frac{1}{2n_1} \sum_{k=1}^{n_1} (G_k - \hat{G}_k)^2, \tag{5}$$

where $G_k$ is the measured monthly gas consumption and the available data are divided into a training set and test set, consisting of $n_1$ and $n_2$ data respectively. Many local optimization methods exist in order to solve the nonlinear least squares problem (5) [3, 4, 7, 9, 10, 12, 13], such as backpropagation (e.g. with adaptive learning rate and momentum term), Levenberg-Marquardt, quasi-Newton or conjugate gradient algorithms.

## 3. Interpretation of identified neural network models

The neural network models (4), identified according to (5), are interpreted then as follows:

- *Yearly error:*
  The yearly error $J_k$ of an identified model is defined as

$$J_k = 100 \, \frac{\sum_{l=k}^{k+12-1} \mid \hat{G}_l - G_l \mid}{\sum_{l=k}^{k+12-1} \; G_l} \qquad (\%) \tag{6}$$

  by taking a sliding window of one year with time index $k$.

- *Normalized gas consumption:*
  The normalized gas consumption $\hat{G}_k^N$ is obtained from the model (4), by $T_k := T_k^N$:

$$\hat{G}_k^N = f(T_k^N, 0, P_k^*, N_k, I_k). \tag{7}$$

  Hence, it expresses what the gas consumption would have been in case of a normal temperature, instead of the real temperature.

- *Growth rate:*
  The growth rate $\zeta_k$ of the gas consumption is defined on the normalized gas consumption, as follows:

$$\zeta_k = 100 \, \frac{\sum_{l=k+12}^{k+24-1} \; \hat{G}_l^N - \sum_{l=k}^{k+12-1} \; \hat{G}_l^N}{\sum_{l=k}^{k+12-1} \; \hat{G}_l^N} \qquad (\%). \tag{8}$$

  A sliding window of one year is taken with time index $k$.

- *Neural network models as uncertain linear models:*
  Neural network models are often regarded as black box models, constituting some nonlinear relationship between the inputs and outputs of the system. However, in [11] it is explained how neural network models can be interpreted as uncertain linear models. This

gives a better insight into the problem. For the model (4) this is done as follows:

$$\begin{aligned}\hat{G}_k &= w^T \tanh(V x_k + \beta) \\ &= w^T \Gamma(x_k) V x_k + w^T \Gamma(x_k) \beta \\ &= a(x_k) + b(x_k) T_k + c(x_k)(T_k - T_k^N) + d(x_k) P_k^* \\ &\quad + e(x_k) N_k + f(x_k) I_k\end{aligned} \tag{9}$$

with $\Gamma(x_k) = \mathrm{diag}\{\gamma_i\}$, $\gamma_i = \frac{\tanh(V(i,:)x_k+\beta_i)}{V(i,:)x_k+\beta_i} \in (0,1]$ and $a(x_k) = w^T\Gamma(x_k)\beta$, $b(x_k) = w^T\Gamma(x_k)V(:,1)$, $c(x_k) = w^T\Gamma(x_k)V(:,2)$, $d(x_k) = w^T\Gamma(x_k)V(:,3)$, $e(x_k) = w^T\Gamma(x_k)V(:,4)$, $f(x_k) = w^T\Gamma(x_k)V(:,5)$. Here $V(i,:)$ denotes the $i$th row of $V$ and $V(:,j)$ the $j$th column of $V$. The coefficients $\gamma_i$ indicate whether the neural network model is weakly or hardly nonlinear. If all $\gamma_i$ values are close to 1, the neural network model is close to linear.

– *Sensitivity:*
The sensitivity $s_k$ of $\hat{G}_k$ with respect to $T_k$ is based on the uncertain linear model interpretation (9)

$$s_k = \frac{E_{l=k}^{k+12-1}\,[b(x_l) + c(x_l)]}{E_{l=k}^{k+12-1}\,[G_l]}, \tag{10}$$

by taking a sliding window of one year with time index $k$. $E_i^j[.]$ denotes the expectation operator, applied to samples $i$ to $j$.

## 4. Simulation results

A comparison between the traditional model (1) and the neural network model (4) is made on Fig.3 and Fig.4. It shows the improvement obtained by introducing the four additional inputs in the nonlinear model (4). In both cases the models have been trained on the data before the vertical line ($n_1 = 100$) and tested on the data behind the vertical line ($n_2 = 27$). The first data point ($k = 1$) corresponds to July 1 1982. The model (1) has been identified in a linear least squares sense. For the neural network model, several numbers of hidden neurons were tested. Best results were obtained by taking $n_h = 5$, with respect to error on the training and test set. The training has been done using a scaled conjugate gradient algorithm [7], each time for 200 random starting points. The best result is shown on Fig.3-4. The corresponding evolution of the error on the training set and test set during optimization is shown on Fig.5. No overfitting occurs in this case.

For this identified neural network model, the normalized gas consumption has been calculated and is shown on Fig.6. From this normalized consumption the growth rate $\zeta_k$ has been obtained on Fig.7. An increase of about 4% was expected a priori, for the period around 1986. This is confirmed by the simulation results. Traditionally, it has been assumed that the sharp decrease of the oil price in 1986 is the main cause of this effect. Therefore, the transformed oil price $P_k^*$ has been taken as input of the neural network in (4), rather than $P_k$, such that the growth rate (which is essentially a time derivative of the gas consumption) is proportional to the inverse of the oil price (at least in case (3) is a linear mapping). Neural network models that consist only of the first input, i.e. the temperature, give wrong interpretations with respect to this curve of growth rate.

Finally, the identified neural network model has been interpreted as an uncertain linear model on Fig.8 (left). It shows the variation in time of the coefficients of $a$,$b$,$c$,$d$,$e$,$f$ of (9) with seasonal oscillations around some nominal value of certain parameters. The $\gamma_i$ coefficients are shown with respect to time on Fig.8 (right): 2 of 5 hidden neurons are permanently operating in their saturation region, which demonstrates that the identified model is indeed highly nonlinear. The sensitivity of the gas consumption with respect to the temperature (10) is shown on Fig.9. Traditionally, this has been considered as an invariant of the system. This is confirmed by the relatively small variation of $s_k$ on Fig.9.

## 5. Conclusion

In this letter it has been shown how the Belgian gas consumption can be accurately modelled by means of neural networks. Considerable improvements have been obtained compared to traditional models, especially for generalization on test data. This is done by not only considering the temperature as input to the neural network but all relevant variables of interest, including temperature, abnormality with respect to the expected temperature, oil price, number of domestic clients and the consumption by industry. The neural network models can be used in order to estimate normalized gas consumption and growth rates. This allows making better extrapolations toward the future, which is important for a good exploitation of the energy system.

## References


1. Asar A., McDonald J.R., "A specification of neural network applications in the load forecasting problem," *IEEE Transactions on Control Systems Technology*, Vol.2, No.2, pp.135-141, 1994.
2. Barron A.R., "Universal approximation bounds for superposition of a sigmoidal function," *IEEE Transactions on Information Theory*, Vol.39, No.3, pp.930-945, 1993.
3. Fletcher R., *Practical methods of optimization*, second edition, Chichester and New York: John Wiley and Sons, 1987.
4. Gill P.E., Murray W., Wright M.H., *Practical Optimization*, London: Academic Press, 1981.
5. Hornik K., Stinchcombe M., White H., "Multilayer feedforward networks are universal approximators," *Neural Networks*, Vol.2, pp.359-366, 1989.
6. Leshno M., Lin V.Y., Pinkus A., Schocken S., "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Networks*, Vol.6, pp.861-867, 1993.
7. Møller M.F., "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, Vol.6, pp.525-533, 1993.
8. Park D.C., El-Sharkawi M.A., Marks II R.J., Atlas L.E., Damborg M.J., "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, Vol.6, No.2, pp.442-449, 1991.
9. Rumelhart D.E., Hinton G.E., Williams R.J., "Learning representations by back-propagating errors," *Nature*, Vol.323, pp.533-536, 1986.
10. Sjöberg J., Zhang Q., Ljung L., Benveniste A., Delyon B., Glorennrc P., Hjalmarsson H., Juditsky A., "Nonlinear black-box modeling in system identification: a unified overview," *Automatica*, Vol.31, No.12, pp.1691-1724, 1995.
11. Suykens J.A.K., Vandewalle J.P.L., De Moor B.L.R., *Artificial Neural Networks for Modelling and Control of Non-Linear systems*, Kluwer Academic Publishers, Boston, Dec. 1995.
12. van der Smagt P.P., "Minimisation methods for training feedforward neural networks," *Neural Networks*, Vol.7, No.1, pp.1-11, 1994.
13. Zurada J.M., *Introduction to Artificial Neural Systems*, West Publishing Company, 1992.

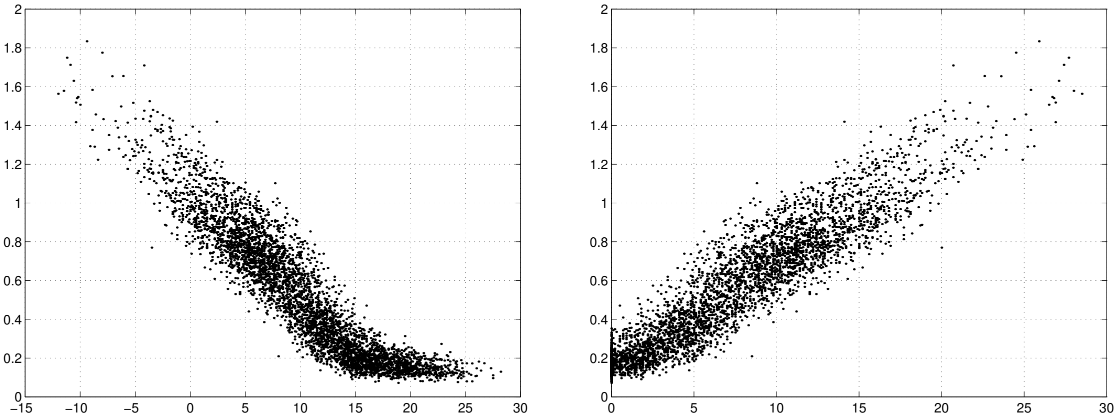


*Figure 1. (Left) Gas consumption (scaled) $G_k$ with respect to temperature $T_k$ for daily data from July 1 1982 till Dec. 31 1994. (Right) $G_k$ with respect to $\varphi(T_k)$. The function $\varphi(T_k)$ has been introduced traditionally in order to obtain an affine relationship between $G_k$ and the nonlinear function of $T_k$.*

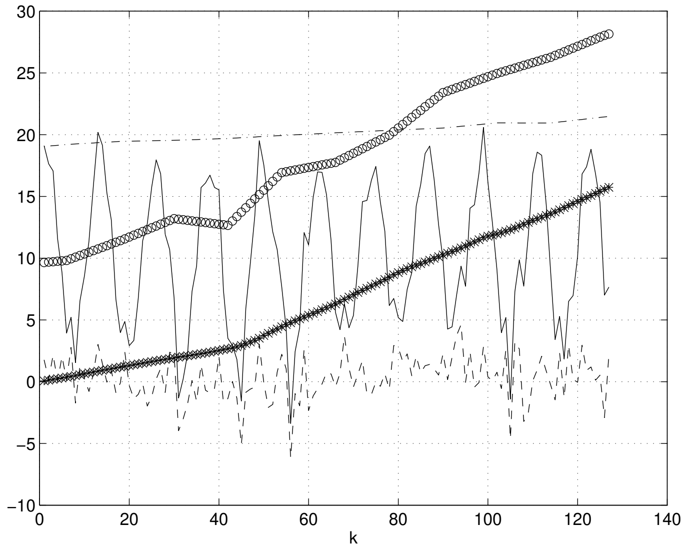


*Figure 2. Five inputs (scaled) for the neural neural network model. The data are monthly, starting from July 1 1982. (-) $T_k$: mean monthly temperature; (- -) $T_k - T_k^N$: difference between mean monthly temperature and normal mean monthly temperature; (***) $P_k^*$: transformed oil price; (-.) $N_k$: number of domestic clients; (ooo) $I_k$: gas consumption by industry.*

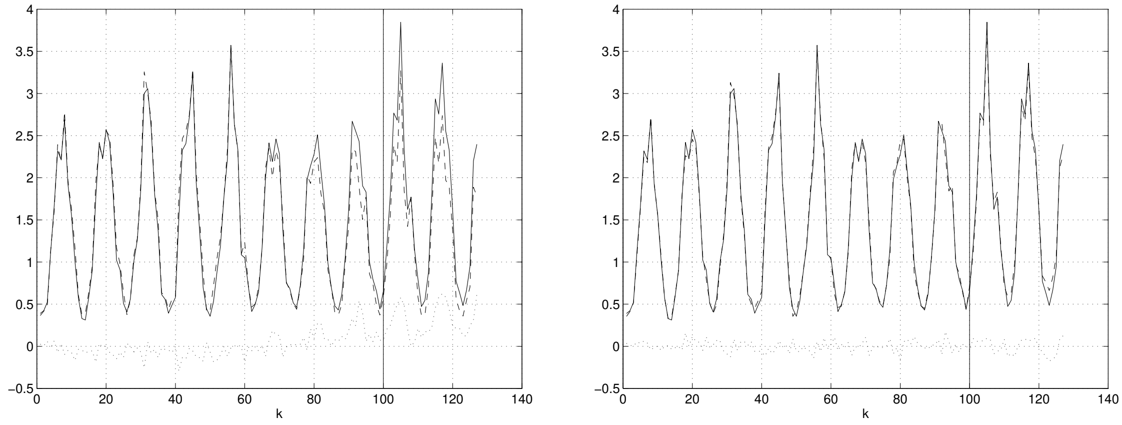


*Figure 3.* *Comparison between a traditional model (left), identified in linear least squares sense, and a neural network model (right), obtained after multistart local optimization. Training data are before the vertical line, test data behind. (-) Gas consumption $G_k$; (- -) estimated gas consumption $\hat{G}_k$ by the model; (...) error signal $G_k - \hat{G}_k$. The neural network model shows an excellent performance on the test data.*

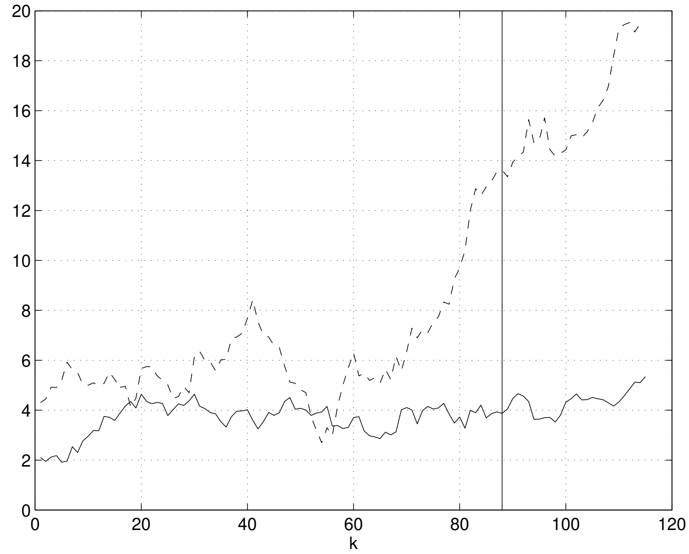


*Figure 4.* *Comparison between the traditional model (- -) and the neural network model (-) of Fig.3 for the yearly error in % with respect to time. It shows the considerable improvement by means of the neural network model, by taking into account the five variables of Fig.2 as input of the neural net.*

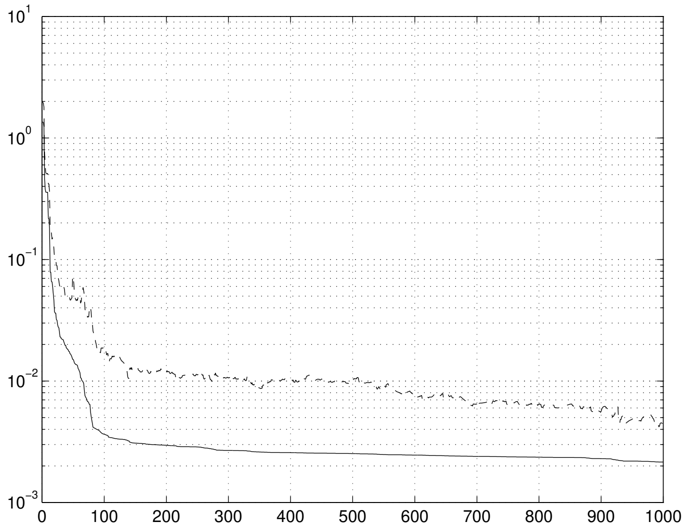


*Figure 5. Evolution of the error on the training set (-) and test set (- -), during optimization for a random starting point. The error is shown on logarithmic scale with respect to the iteration number. A scaled conjugate gradient algorithm has been used for training.*

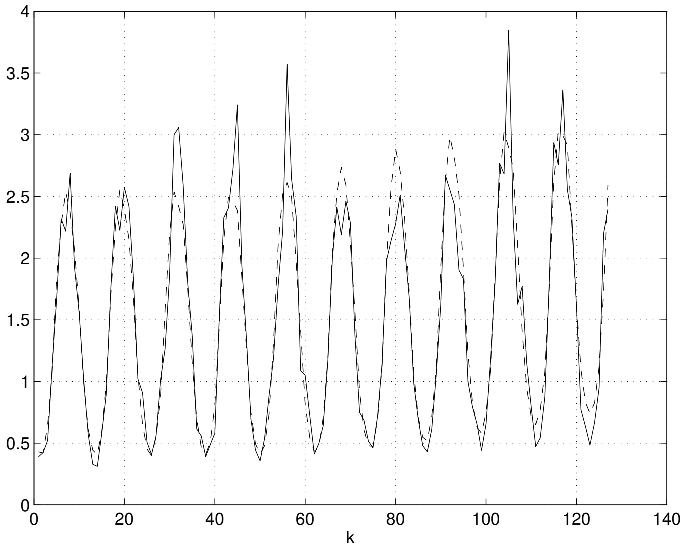


*Figure 6. Comparison between gas consumption $G_k$ (-) and normalized gas consumption $\hat{G}_k^N$ (- -) with respect to time. $\hat{G}_k^N$ has been obtained from the identified neural network model of Fig.3 (right).*

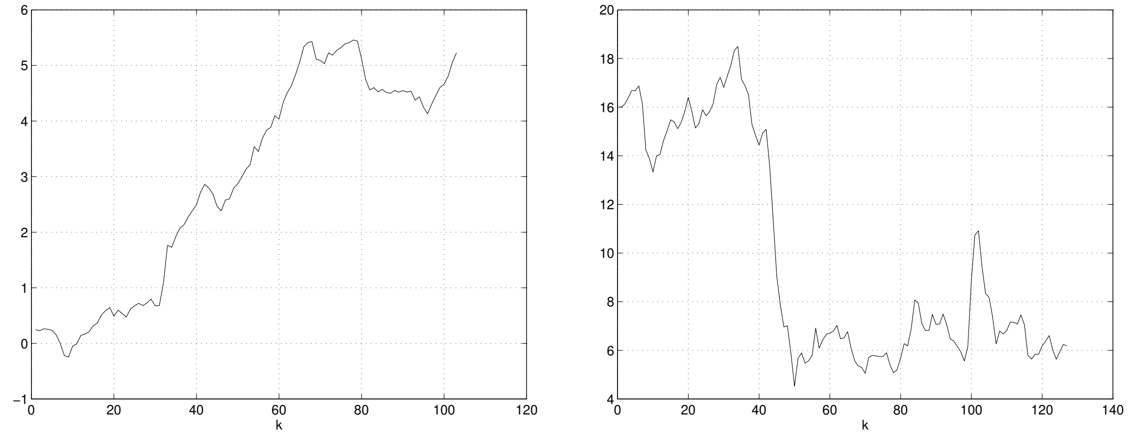


*Figure 7. (Left) Growth rate in % with respect to time, calculated from the normalized gas consumption of Fig.6; (Right) oil price (scaled) with respect to time. The sharp decrease of the price around 1986 is the main cause for the growth of consumption in this period.*

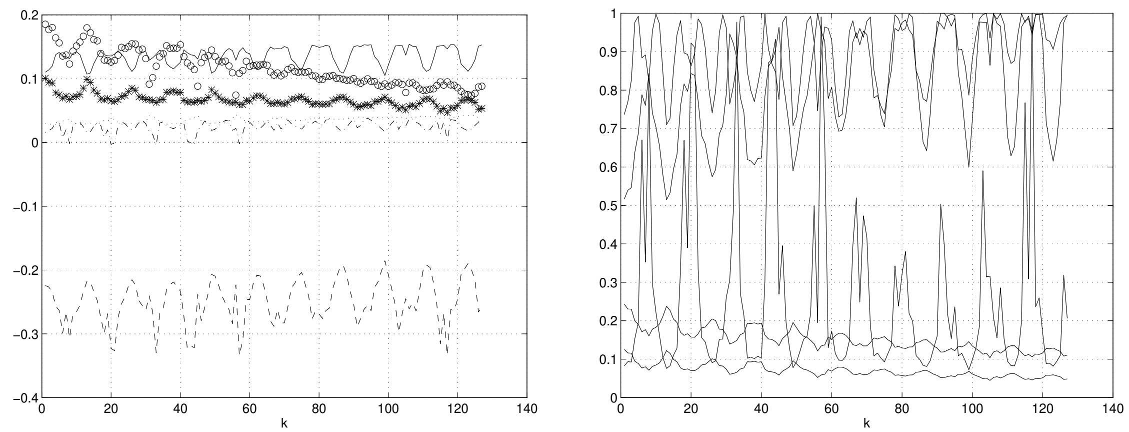


*Figure 8. (Left) Identified neural network model, interpreted as an uncertain linear model: (-) $a(x_k)$, (- -) $b(x_k)$, (-.) $c(x_k)$, (...) $d(x_k)$, (ooo) $e(x_k)$, (***) $f(x_k)$. Variations around nominal values of certain parameters, due to seasonal effects, are clearly visible; (Right) $\gamma_i$ coefficients with respect to time: 2 of 5 neurons are permanently in their saturation region, which indicates that the model is strongly nonlinear.*

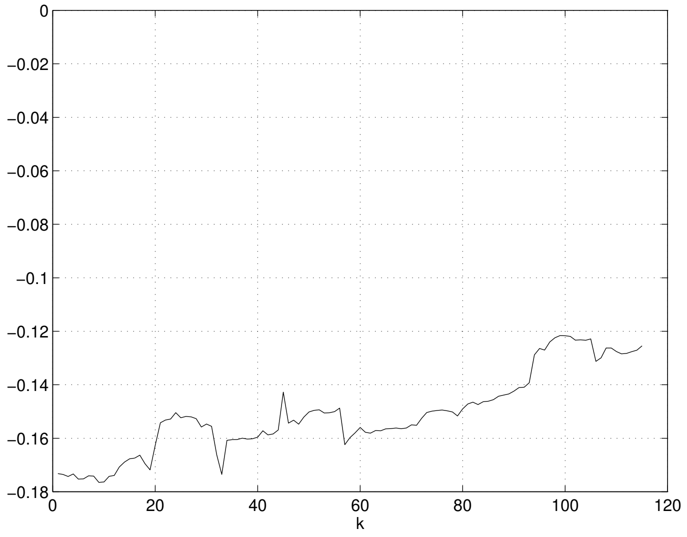


*Figure 9. Sensitivity of the estimated gas consumption with respect to time.*