NEURAL NETWORKS IN SIGNAL PREDICTION

Martin ŠTOREK, Aleš PROCHÁZKA

Prague University of Chemical Technology, Department of Computing and Control Engineering
Technická 1905, 166 28 Prague 6, Czech Republic
Phone: +42-2-2435 4198 ⋆ Fax: +42-2-2431 1082 ⋆ E-mail: storekm@vscht.cz, prochaz@vscht.cz

ABSTRACT

Methods of signal modelling and prediction provide a very important tool of various information and control systems in technology, biomedicine and environmental engineering. The paper presents the analysis of signal preprocessing including its filtering and subsampling at first as a basic tool for the following efficient signal modelling and prediction. The main part of the paper is devoted to the comparison of linear autoregressive serially and seasonally related models and nonlinear systems based upon neural networks. The singular value decomposition and QR factorization is then studied for model structure selection. Resulting methods are verified for simulated signals at first and then used for real signal processing. General algorithms presented in the paper are used for the analysis of sampled values of the river flow to allow their preprocessing and prediction both by classical and adaptive methods using system history.

1 INTRODUCTION

Signal prediction is a very important tool of information engineering with its application in technology, control systems, bioengineering, environmental systems and econometrics. Its theoretical background covers autoregressive models [11, 16] and nonlinear systems based upon artificial neural networks [7, 4, 3, 2, 14] including methods of adaptive signal processing and system optimization [6, 8, 15].

The paper presents possibilities of signal preprocessing and subsampling to enable the following efficient signal prediction and it compares results of signal prediction achieved both by classical and adaptive methods. Basic general algorithms presented are verified for simulated signals at first and then used for real signal processing.

2 SIGNAL PREPROCESSING

Methods of signal preprocessing [9] include in most cases its spectral analysis, filtering, noise rejection, restoration of missing values and subsampling. In case of signal prediction discussed further it is essential to find basic signal features for its modelling. Results of such a preliminary analysis resulting from the short time Fourier transform (STFT) for the real time series representing the river Elbe flow measured with the sampling period of one day in the period 1941-1995 is presented in Fig. 1. This analysis proves a dominant periodic component one year long and further components about 7 years long.

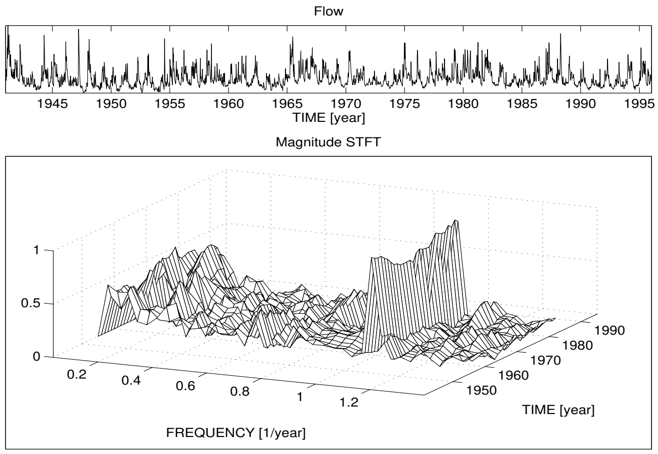

Fig. 1: Analysis of a given signal representing river Elbe flow in the time period 1941 - 1995 measured with the sampling period of 1 day by the short time Fourier transform with its window 28 years long.

Further problems of signal preprocessing include the compensation of delay in signal filtering, the choice of a proper cutoff frequency and the subsampling rate for its decimation in connection with the desired goal of the whole information or control system. The basic structure used in further discussions is presented in Fig. 2.

3 NONLINEAR TIME-SERIES MODELLING AND PREDICTION

Methods of signal modelling and prediction include linear and nonlinear systems. While linear autoregressive models are sufficient in many applications including both serially and seasonally related time series it is necessary to use nonlinear systems in special cases. The basic neural network structure presented in Fig. 2 assumes its two layers of sizes $R - S1 - 1$ with a transfer function $F1$ including both sigmoidal and wavelet functions. Output of such a model for any input vector $\mathbf{P}_{R,1}$ is defined by relation

$$\mathbf{Y} = \mathbf{W2} * F1\left(\mathbf{W1} * \mathbf{P} + \mathbf{B1}\right) + \mathbf{B2} \qquad (1)$$

where matrices $\mathbf{W1}_{S1,R}, \mathbf{W2}_{1,S1}$ define neural network coefficients and vectors $\mathbf{B1}_{S1,1}, \mathbf{B2}_{1,1}$ represent biases.

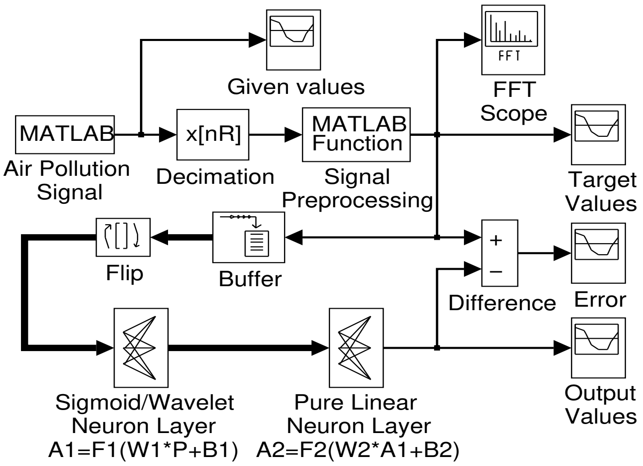

Fig. 2: System for signal preprocessing and prediction

4 ADAPTIVE SIGNAL PROCESSING

The basic structure of the adaptive system for signal prediction using the Widrow method [4] is presented in Fig. 3. It represents the simplification of a more general system presented in Fig. 2 to one layer only. The method assumes a proper selection of training values and network coefficients optimization based upon the gradient method. Each input vector is presented in a sequence to the network causing the adjustment of network weights and biases according to the resulting error. As the linear network is continuously updated it can respond to changes in the relationship between past and future values of a signal. This property is very useful in slowly changing systems using its historical values and model evolution.

Principle of the learning process of an adaptive element can be explained for a time series $\mathbf{D} = [d_1, d_2, \cdots, d_P]'$ and a reference signal $\mathbf{G} = [g_1, g_2, \cdots, g_P]'$. Assuming an adaptive processor with its initial coefficients $\mathbf{W} = [w_1, w_2, \cdots w_Q]'$ it is possible to apply the mean square method to minimize function

$$S = \sum_{k=1}^{P} (d_k - (w_1 g_k + w_2 g_{k-1} + \cdots + w_N g_{k-N+1}))^2 \quad (2)$$

The following gradient method results in the final iterative algorithm in the form

$$\mathbf{W} = \mathbf{W} - \alpha \frac{\partial S}{\partial \mathbf{W}} = \mathbf{W} + \alpha \mathbf{X}' \, (\mathbf{D} - \mathbf{P} * \mathbf{W}) \qquad (3)$$

for a chosen learning value α and pattern matrix $\mathbf{P}$ defined by a reference signal $\mathbf{G}$. An example of results of such a process is presented in Fig. 3 together with corresponding neural network structure.

The whole algorithm presented in Fig. 3 has been verified for a simulated signal at first and then applied for real signal processing as well. The basic problem connected with real signal processing is in a proper selection of pattern values. In a given application results of spectral analysis has been used and 10 pattern values were formed by 5 values preceeding the predicted value together with 5 values one year old.

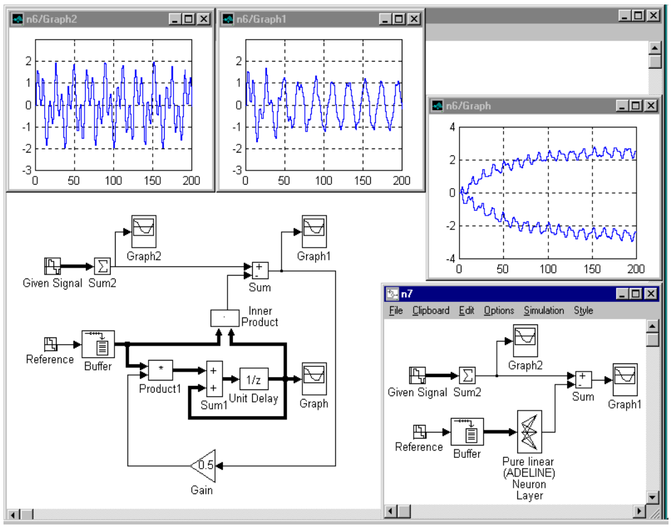

Fig. 3: Principle of adaptive signal prediction for simulated time series processing

5 PARAMETER ESTIMATION

The problem of parameter estimation is very common in system identification and it includes model structure selection, evaluation of its parameters and model validation. The following discussion is devoted to

- the application of the singular value decomposition (SVD) for the linear regression model order selection
- the use of QR factorization for the selection of independent variables in a regression problem having the fundamental contribution to the quality of the mathematical model

Such a subset selection is based upon basic vector and matrix operations [1, 9, 10, 12, 17] and it is used in various applications of system identification, control and signal prediction.

Let us study a signal $\mathbf{y} = [y(1), y(2), \cdots, y(M+N)]$. Its simple autoregressive (AR) model of order N is defined by a vector $\mathbf{a} = [a(1), a(2), \cdots, a(N)]'$ providing the estimate of its one sample ahead prediction $\mathbf{A}_{M,N}\ \mathbf{a}_{N,1} = \mathbf{y}_{M,1}$,

$$\begin{bmatrix} y(N) & Y(N-1) & \cdots & y(1) \\ y(N+1) & Y(N) & \cdots & y(2) \\ & & \cdots & \\ y(N+M-1) & Y(N-M-2) & \cdots & y(M) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \cdots \\ a(N) \end{bmatrix} = \begin{bmatrix} y(N+1) \\ y(N+2) \\ \cdots \\ y(N+M) \end{bmatrix} \quad (4)$$

The full-set AR model maximum order N can be estimated from the signal autocorrelation function as the number of its significant values. An example of such a function together with the estimate of the power spectrum is given in in Fig. 4 for a seasonal time series representing the river Elbe flow with the sampling period one month long. A one year and about 7 years long period results from this analysis. Owing to this signal properties it is possible to find the subset g values long of the full model of order N with. To find this subset the fundamental variables can be selected using the singular value decomposition of matrix

$$\mathbf{A}_{M,N} = \mathbf{U}_{M,M}\ \mathbf{S}_{M,N}\ \mathbf{V'}_{N,N} \tag{5}$$

at first producing the diagonal matrix $\mathbf{S}$ with nonnegative diagonal elements in decreasing order and orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$. The number g of dominant diagonal values of matrix $\mathbf{S}$ define the size of the subset model even though in real cases (Fig. 4) the selection of the tolerance value can not be precisely stated. The following QR factorization with column pivoting

$$\mathbf{V'}_{N,N}\ \mathbf{P}_{N,N} = \mathbf{Q}_{N,N}\ \mathbf{R}_{N,N} \tag{6}$$

can be applied for matrix $\mathbf{V'}$ decomposition into matrix $\mathbf{Q}$ with orthonormal columns and upper triangular matrix $\mathbf{R}$. The permutation matrix $\mathbf{P}$ enables to rearrange columns of the original matrix in order of their Euclidian norm and their importance.

To solve the system of equation (3) it is possible to use decomposition (5) at first to form equation

$$\mathbf{U}_{M,M}\ \mathbf{S}_{M,N}\ \mathbf{V'}_{N,N}\ \mathbf{a}_{N,1} = \mathbf{y}_{M,1} \tag{7}$$

Introducing a new variable $\mathbf{z} = \mathbf{V'}\ \mathbf{a}$ and taking into account properties of decomposition matrices it is possible to find the solution of this equation using MATLAB notation in the form

$$\mathbf{z}_{N,1} = \mathbf{U'}_{N,M}\ \mathbf{y}_{M,1}\ ./\ diag(\mathbf{S}_{M,N}) \tag{8}$$

Owing to properties of decomposition matrices only first N columns of matrix $\mathbf{U}$ are used here. Using the permutation matrix $\mathbf{P}$ for selection of the most important columns of matrix $\mathbf{V'}$ the model coefficients can be evaluated by relation

$$[\mathbf{a}_{g,1};\ \mathbf{a}_{N-g,1}] = \mathbf{P'}_{N,N}\ \mathbf{V}_{N,N}\ \mathbf{z}_{N,1} \tag{9}$$

Model of the pseudorank g is defined by its coefficients resulting form Eq. 9 and corresponding values of the permutation matrix $\mathbf{P}$. Model validation resulting from various pseudoranks are given in Fig. 4 as well allowing its selection using the mean square error or taking into account different information criteria.

6 RESULTS

Methods presented above have been verified for simulated signals at first and then used for the real signal processing representing water flow measured at the selected measuring station each day for last 55 years. The main goal of this analysis is in the long term prediction useful from the ecological point of view and providing information for flow control to enable stable transport conditions as well.

Spectral analysis verified one year long seasonal component enabling the design of both seasonal linear model and nonlinear model as well. In all cases the model structure selection has been evaluated through the SVD and QR algorithm. A simple model resulting from this approach is based upon the following difference equation

$$y(N+1) = a(3)y(N-2) + a(7)y(N-6) + a(11)y(N-10)$$

but the more complicated structures has been tested, too.

Results of one step ahead prediction by a neural network are summarized in Tab. 1. In all cases signal decimation has been applied at first to achieve signal sampling period one day, one week and one month long respectively. The adaptive system has been then trained for one step ahead prediction. In the selected validation part the maximum and the mean errors have been compared. Fig. 5 represents results of one day ahead prediction presenting both results of the adaptive learning process and system

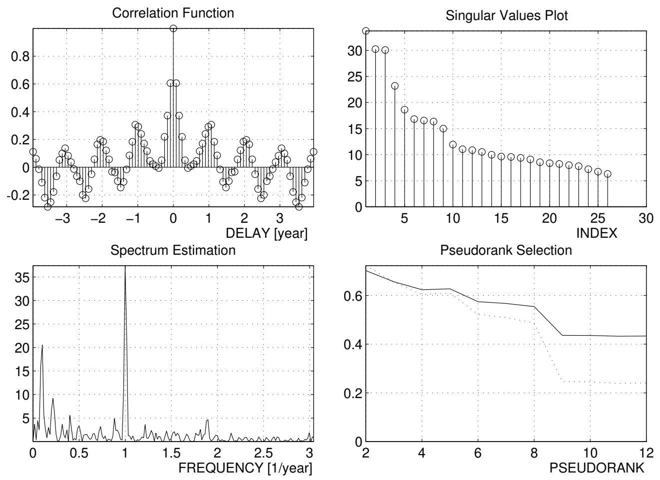

Fig. 4: Real time series analysis representing river Elbe flow sampled with one month period in years 1951 - 1980 providing **(a)** Autocorrelation function, **(b)** Spectral analysis **(c)** Singular values of the matrix formed for AR model coefficients evaluation for the full set model order $N = 26$ and the pseudorank in the range of $k = 2, 3, \cdots, 12$, **(d)** Model mean square error for the model subset selection together with results of the Akaike information criterium

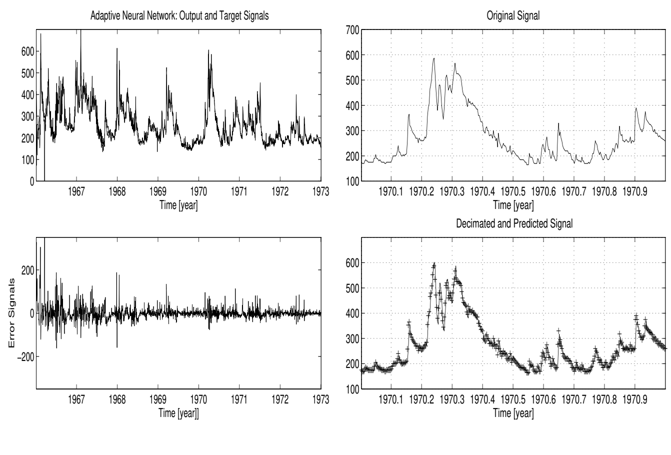

Fig. 5: Results of system adaptation during the learning process for one day ahead prediction system validation for a selected time period

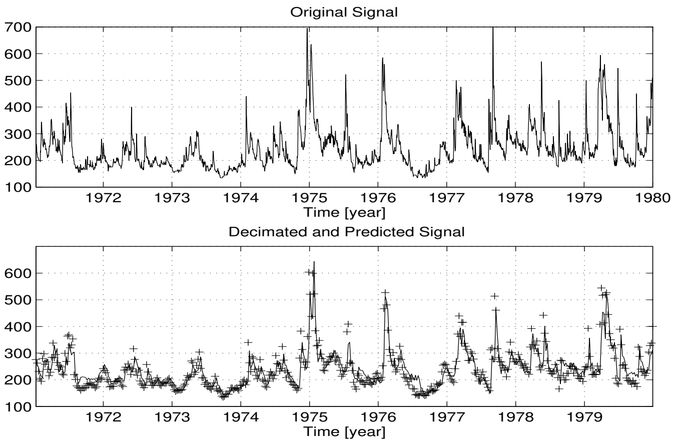

Fig. 6: Results of one week ahead prediction during a selected time period

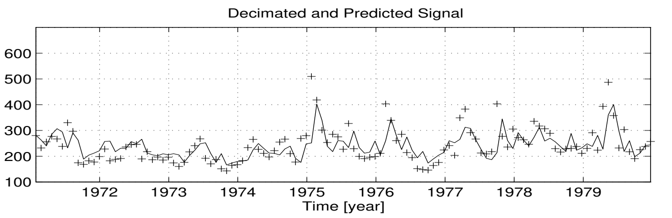

Fig. 7: Results of one month ahead prediction during a selected time period

validation. Similar results of one step ahead prediction in case of one week and one month sampling period are provided in Fig. 6 and Fig. 7 respectively.

7 CONCLUSION

Submitted paper describes possibilities of signal prediction both for simulated and real signals representing water flow measurements. In all cases the structure of the mathematical model has been estimated using the singular value decomposition principle.

Further problems include signal missing values restoration [5], signal noise removal using various methods including wavelet decomposition [18, 13], delay free filtering [9] and basic features extraction for signal prediction.

TABLE 1: TYPICAL ERRORS OF THE ADAPTIVE SIGNAL PREDICTION USING 10 SELECTED HISTORICAL VALUES OF THE GIVEN TIME SERIES

Prediction Extend	*Maximum Error [%]*	*Mean Error [%]*
One Day	33	7
One Week	53	18
One Month	55	23

Acknowledgement

The paper is supported by the grant No. 201/94/0130 of the Grant Agency of Czech Republic. All real signals were kindly provided by the Ústí measuring station.

Reference

[1] S. D. Conte ane C. Boor. *Elementary Numerical Analysis.* McGraw-Hill Kogakusha, Ltd., Japan, 1972.

[2] A. Atyia, S. El-Shoura, S. Shaheen, and M. El-Sherif. River Flow Forecasting Using Neural Networks. In *World Congress on Neural Networks*, pages 461–464, San Diego, California, 1996. Lawrence Erlbaum Associates.

[3] B. Delyon, A. Juditsky, and A. Benveniste. Accuracy Analysis for Wavelet Approximation. *IEEE Trans. on Neural Networks*, 6(2):332–348, 1995.

[4] H. Demuth and M. Beale. *Neural Network Toolbox.* The MathWorks, Inc., Natick, Massachusetts 01760, January 1994.

[5] W. Etter. Restoration of Discrete-Time Signal Segments by Interpolation Based on the Left-Sided and Right-Sided Autoregressive Parameters. *IEEE Trans. Signal Processing*, 44(5):1124–1135, May 1996.

[6] M. T. Hagan and M. B. Menhaj. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Trans. Neural Networks*, 5(6):989–993, November 1994.

[7] S. Haykin. *Neural Networks.* IEEE Press, New York, 1994.

[8] E. S. Chng, S. Chen, and B. Mulgrew. Gradient Radial Basis Function Networks for Nonlinear and Nonstationary Time Series Prediction. *IEEE Trans. on Neural Networks*, 7(1):190–194, 1996.

[9] P. P. Kanjilal. *Adaptive Prediction and Predictive Control.* Peter Peregrinus Ltd., IEE, U.K., 1995.

[10] G. Linfield and J. Penny. *Numerical Methods Using Matlab.* Ellis Horwood, Hertfordshire, U.K., third edition, 1995.

[11] L. Ljung. *System Identification.* Prentice Hall, Englewood Cliffs, N.J., 1987.

[12] S. Mika. *Numerické metody algebry.* SNTL, Praha, 1985.

[13] M. Misiti, Y. Misiti, G. Oppenheim, and J. M. Poggi. *Wavelet Toolbox.* The MathWorks, Inc., Natick, Massachusetts 01760, March 1996.

[14] E. Pelikan, Matejka P., Slama M., and Vinkler K. Interactive Forecasting of the Electric Load Using Kohonen Self-Organizing Feature Maps. In *World Congress on Neural Networks*, pages 435–442, San Diego, California, 1996. Lawrence Erlbaum Associates.

[15] H. Phillips. Adaptive Forecasting. In *World Congress on Neural Networks*, pages 504–507, San Diego, California, 1996. Lawrence Erlbaum Associates.

[16] A. Procházka, M. Mudrová, and J. Fiala. Non-Linear Time Series Modelling and Prediction. *Neural Network World*, 6(2):215–221, 1996.

[17] A. Ralston. *A First Course in Numerical Analysis.* McGraw-Hill, Inc., New York, 1965.

[18] G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, 1996.