NEURAL NETWORKS AND SEASONAL TIME-SERIES PREDICTION

Aleš Procházka

Prague Institute 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: prochaz@vscht.cz

ABSTRACT

Seasonal time series with dominant frequency components are very common in various information and control systems, biomedicine and environmental engineering. Methods for their analysis, modelling and prediction include both linear and nonlinear structures using autoregressive models, adaptive systems and neural networks. The paper presents basic algorithms for preprocessing of such signals including their filtering and subsampling at first to enable the following efficient signal approximation, modelling and prediction. The main part of the paper presents the comparison of linear autoregressive models and neural networks with different transfer functions including wavelets as well. Resulting methods presented in the paper are applied for processing of real values representing the river Elbe flow at a selected measuring station to study their modelling and prediction. The paper compares results obtained both by classical and adaptive methods using system history and appropriate learning methods for model optimization.

INTRODUCTION

Time-series prediction is a very important problem in many applications including control systems, engineering processes, bioengineering, environmental systems and econometrics. Mathematical tools used to predict signal behaviour cover linear models [11, 12, 18] and nonlinear systems based in many cases upon artificial neural networks [5, 8, 15] including methods of adaptive signal processing [5, 17] and system optimization [7, 19].

The paper presents comparison of various methods of signal prediction. The main part of the paper is devoted to the study of wavelet functions for signal analysis [2, 3, 9, 16, 21, 22] in connection with applications of filter banks [23]. Both these functions and radial basis functions can be used for signal analysis [4, 13], classification [20] and prediction [10]. The contribution presents the use of harmonic wavelet function [14] for signal analysis at first and the use of wavelets as the first layer transfer functions of artificial neural networks according to the structure given in Fig. 1. Results achieved by other methods are then compared.

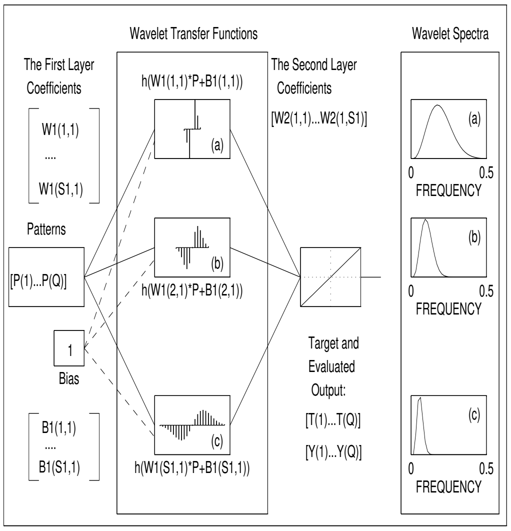

Figure 1: The basic structure of artificial neural network with wavelet transfer functions.

All algorithms presented in the paper were verified for simulated signal at first and then used for real signal processing. The problem of river flow forecast [1] has been chosen for signal analysis and prediction as a very instructive example of a seasonal time-series with a very clear periodicity of one year. Both classical and adaptive methods have been studied in this case.

TIME - FREQUENCY AND TIME - SCALE SIGNAL ANALYSIS

The processing of real signals include in most cases their 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 components for its modelling at first.

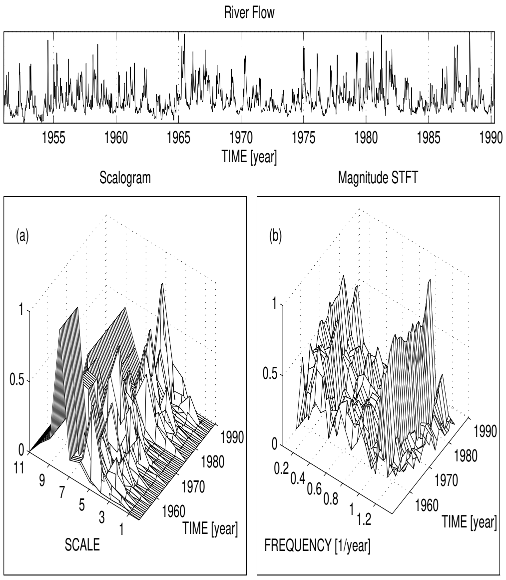

Figure 2: Analysis of a given signal representing river Elbe flow in the time period 1951 - 1990 measured with the sampling period of 1 day by **(a)** The Wavelet analysis based upon the harmonic wavelet functions pointing to the periodicity of 1 year (scale 5) and 7 years (scale 9), **(b)** The Short time Fourier transform with its window 14 years long providing similar results.

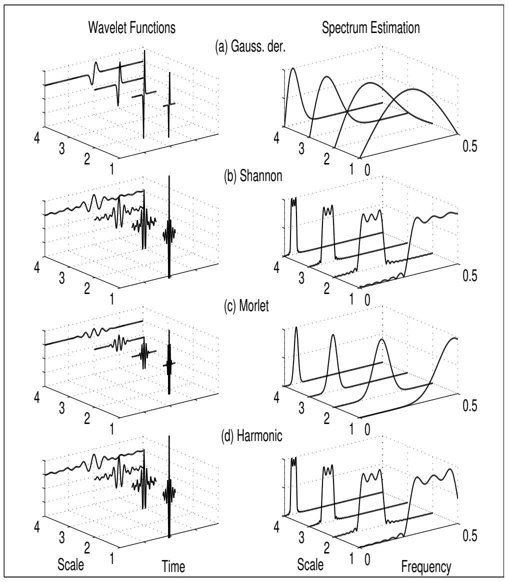

Figure 3: Initial and dilated wavelet functions in the time domain and their spectra: **(a)** Derivative of Gaussian function, **(b)** Shannon function, **(c)** Morlet function, **(d)** Harmonic function.

Signal analysis of non-stationary signals is in many cases evaluated by the Short time Fourier transform using the moving window of a chosen length and shape to estimate signal spectral components within its range. Results of such a preliminary analysis for the real time-series representing the river Elbe flow measured with the sampling period of one day in the period 1951-1990 is presented in Fig. 2. This analysis proves a dominant periodic component one year long and further components about 7 years long.

As the Short time Fourier transform (STFT) applies the window of the same length for signal segments choice it has the constant frequency resolution. To allow different resolution over frequency components it is often useful to apply Wavelet transform (WT). Results of signal analysis achieved for the same real signal are presented in Fig. 2 as well.

Basic properties of wavelet functions both in time and frequency domains are presented in Fig. 3 for different initial (mother, basis) wavelets $h(t)$ including the following selected functions:

- Gaussian derivative

$$h(t) = -t\ e^{-t^2/2} \tag{1}$$

- Shannon wavelet function

$$h(t) = \frac{sin(\pi\ t/2)}{\pi\ t/2}\ cos(3\ \pi\ t/2) \tag{2}$$

- Modulated Gaussian (Morlet) function

$$h(t) = e^{j\omega_0\ t}\ e^{-t^2/2} \tag{3}$$

- Harmonic wavelet function

$$h(t) = \frac{1}{j\pi\ t}(e^{j\pi\ t} - e^{j\pi/2\ t}) \tag{4}$$

Using the signal processing point of view it is possible to consider the initial wavelet as a pass-band filter.

In all cases the corresponding associated (baby) wavelet functions are derived using dilation coefficient usually defined on the dyadic grid by value $a = 2^m$ and translated by constant $b = k\ 2^m$ resulting in relation

$$h_{m,k}(t) = \frac{1}{\sqrt{a}}\ h(\frac{t-b}{a}) = \frac{1}{\sqrt{2^m}}\ h(2^{-m}\ t - k) \tag{5}$$

where m and k are integers. The wavelet dilation by factor 2^m implies its pass-band compression by the same factor according to Fig. 3.

The system of wavelet functions defines a special filter bank which can be used for signal components analysis. Application of this analysis for a real signal described above enables to detect periodic signal components useful for the selection of the given system modelling and its prediction.

SIGNAL PREDICTION

Many systems of signal modelling and prediction using both the classical and adaptive approaches [11] are based on linear models. Consider a selected time series $\{y(n)\}$ containing measurements with a given sampling period T. Assuming that the time-series is seasonal with its periodic component being N samples long, it is possible to apply a linear model for signal prediction m samples ahead in the form

$$y(n) = \sum_{j=1}^{R} a(j)\, y(n-k_j) + e(n) \qquad (6)$$

having R unknown coefficients with $m = min(k(j))$ for $j = 1, 2, \cdots, R$. It can be expected that the time-series is both serially related to previous values and seasonally related to values N samples delayed. The main problem of optimal modelling [11] in this case is both in the selection of the number of variables R and in their specification by values of vector $\mathbf{k}$. In this case $\mathbf{k} = [1, 2, \cdots, N, N-1, \cdots]$ for a given seasonal model structure.

A nonlinear neural network can be used in a similar way to evaluate the output $\mathbf{A1}$ of its first layer and the second layer output $\mathbf{A2}$ $(= y(n))$ in the form

$$\mathbf{A1} \;=\; F1\,(\mathbf{W1} \star \mathbf{P} + \mathbf{B1}) \qquad (7)$$

$$\mathbf{A2} \;=\; F2\,(\mathbf{W2} \star \mathbf{A1} + \mathbf{B2}) \qquad (8)$$

for selected transfer functions $F1$ and $F2$ respectively, given pattern vector $\mathbf{P}_{R,1}$ of previously observed R values and optimized matrices $\mathbf{W1}_{S1,R}$, $\mathbf{B1}_{S1,1}$, $\mathbf{W2}_{S2,S1}$ and $\mathbf{B2}_{S2,1}$ with their selected sizes. Signal approximation can be performed for a sigmoidal or wavelet function $F1$ and a linear function $F2$ using the basic structure presented in Fig. 1. For a one layer network, a linear transfer function and properly chosen pattern values, this model is equivalent to the seasonal model described by Eq. (6).

Even though 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. 4 assumes its two layers of sizes $R - S1 - 1$ with a nonlinear first layer transfer function $F1$ including either sigmoidal or wavelet functions and linear second layer transfer function. The buffer block is used in this case for the neural network input pattern definition. Output of such a model for any input vector $\mathbf{P}_{R,1}$ can be evaluated by relation

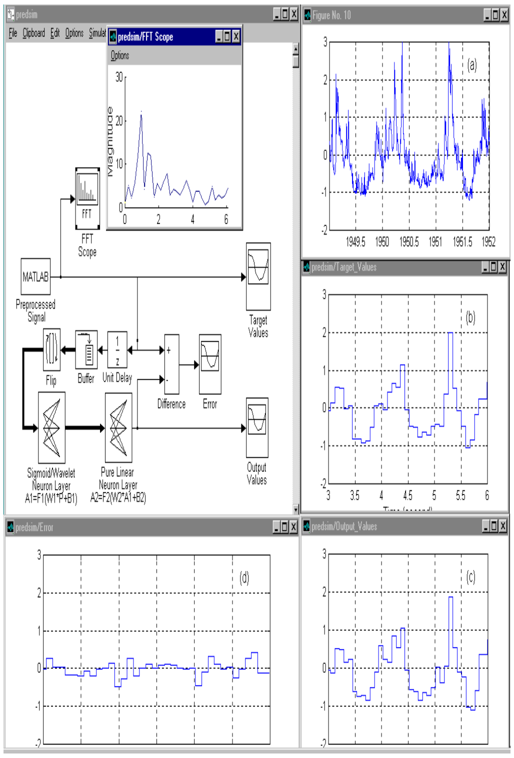

Figure 4: The two layer neural network use for one step ahead signal prediction with the observed sequence spectral analysis presenting **(a)** The original sequence standing for the river Elbe flow for a three years time segment and sampling period of one day, **(b)** The time series of average monthly flow and sampling period of 30 days evaluated from the original sequence, **(c)** Result of one step ahead signal prediction using the optimized neural network with sigmoidal transfer function, **(d)** Prediction error evolution.

$$\mathbf{Y} = \mathbf{W2} * F1\,(\mathbf{W1} * \mathbf{P} + \mathbf{B1}) + \mathbf{B2} \qquad (9)$$

To find all these coefficients a properly chosen optimization method must be used during the learning process.

While linear and nonlinear models described above assume the selection of the learning and validation parts it is often useful to apply the adaptive Widrow methods [5] to update system coefficients continuously. The 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

$$S = \sum_{n=1}^{P} (d_n - \sum_{j=1}^{R} (w_j g(n - k_j)))^2 \qquad (10)$$

using the selected vector $\mathbf{k}$ of time delays according to the seasonal time-series properties. The following gradient method results in the final iterative algorithm in the form

$$\mathbf{W} = \mathbf{W} - \alpha \frac{\partial S}{\partial \mathbf{W}} = \mathbf{W} + \alpha \mathbf{P}' \, (\mathbf{D} - \mathbf{P} * \mathbf{W}) \qquad (11)$$

for a chosen learning value α and pattern matrix $\mathbf{P}$ defined by a reference signal $\mathbf{G}$.

RESULTS

Methods presented above have been applied for the real signal processing representing water flow measured at the selected measuring station on the river Elbe in Czech republic 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 the flow control to enable stable transport conditions 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 have proved one year long seasonal component enabling the design of both seasonal linear model and nonlinear model as well. In all cases signal decimation has been applied at first to achieve signal sampling period one day, one week and one month long and 5 pattern values were formed by 2 values preceeding the predicted value together with 3 values one year old. Fig. 5 provides results for a seasonal autoregressive model evaluated during the period of 7 years from 1986 till 1992 for real values decimated to the final sampling period of one month and one sample ahead prediction. Results of achived by an associated model for the following signal processing during the validation period of 3 years from 1992 till 1995 are presented in the same figure as well.

Similar results are presented in Fig. 6 for a model using the two layer neural network 5-5-1 with a sigmoidal transfer function showing both the learning and validation parts. Fig. 7 presents results of the application of adaptive system for one month ahead prediction in the same time period.

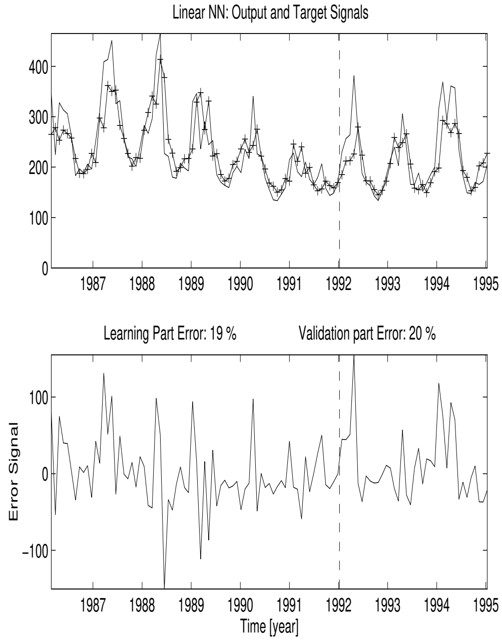

Figure 5: The learning and validation part for a real signal one step ahead prediction using seasonal AR model of the 5th order using sampling period of one month and taking into account the main signal periodic component one year long.

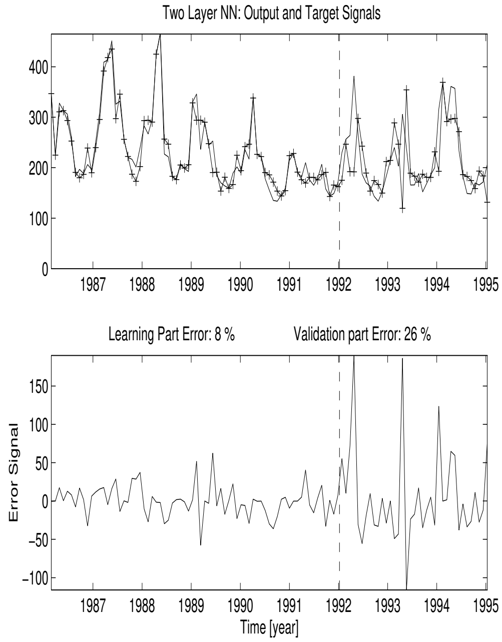

Figure 6: The learning and validation part for a real signal one step ahead prediction by the two layer 5-5-1 neural network with sigmoidal and linear transfer function respectively using in each pattern two previous signal values and three values one year old taking into account the main signal periodic component one year long.

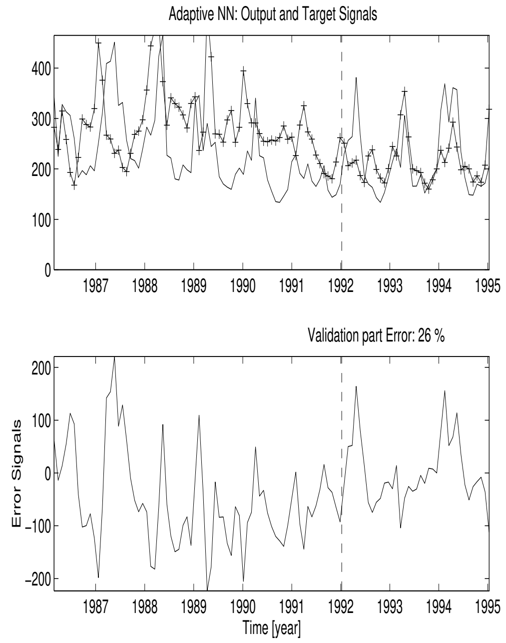

Figure 7: The real signal segment one step ahead prediction evaluated by the adaptive linear processor and Widrow learning algorithm taking into account the main signal periodic component one year long.

Typical results of one day, one week and one month ahead prediction are summarized in Tab. 1. Both in the learning and validation parts the mean square errors are compared. It is obvious that in this case the linear autoregressive model provides results with its mean square error of the same order as nonlinear structures.

CONCLUSION

The paper compares basic methods of signal prediction using both linear and nonlinear structures based upon neural networks with sigmoidal and wavelet transfer functions. All algorithms were verified for simulated signals at first and the used for real signals processing representing water flow measurements.

Further problems include the study of signal preprocessing, signal missing values restoration [6], signal noise removal using various methods including wavelet decomposition [13, 23] and problems of delay free filtering [11].

Table 1: Comparison of typical mean square errors of a real signal prediction using 5 selected historical values for one day, one week and one month ahead prediction using autoregressive seasonal model, two layer neural network with sigmoidal and wavelet transfer function and adaptive linear element.

Sampling Period [Days]	*Error (Initial/Validation) [%]*			
	AR	*Sigmoidal* Neural Network	*Wavelet* Neural Network	*Adaptive* Neural Network
1	7/7	7/7	7/7	17/7
7	18/16	14/15	14/16	29/16
30	19/20	14/25	12/20	31/19

Acknowledgement

The paper has been supported by the grant No. 201/94/0130 of the Grant Agency of Czech Republic. All real signals were kindly provided by the Czechoslovak Elbe River Transport company.

REFERENCES

[1] 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.

[2] A. Bruce, D. Donoho, and Hong-Ye Gao. Wavelet Analysis. *IEEE Spectrum*, 33(10):26–35, October 1996.

[3] I. Daubechies. The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Trans. Inform. Theory*, 36:961–1005, September 1990.

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

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

[6] 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.

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

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

[9] Y. T. Chan. *Wavelet Basics*. Kluwer Academic Publishers, Boston, 1995.

[10] 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.

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

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

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

[14] D. E. Newland. *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*. Longman Scientific & Technical, Essex, U.K., third edition, 1994.

[15] 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.

[16] J. Pesquet, H. Krim, and H. Carfantan. Time-Invariant Orthonormal Wavelet Representation. *IEEE Trans. Signal Processing*, 44(8):1964–1970, August 1996.

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

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

[19] A. Procházka and V. Sýs. Application of Genetic Algorithms for Wavelet Networks Signal Modelling. In *VIIth European Signal Processing Conference EUSIPCO-94*, pages II/1078–II/1081. European Association for Signal Processing, 1994.

[20] A. Procházka and M. Štorek. Wavelet Transform Use for Signal Classification by Self-Organizing Neural Networks. In *Fourth International Conference on Artificila Neural Networks ANN-97*. IEE, Cambridge, England, 1995.

[21] O. Rioul and M. Vetterli. Wavelets and Signal Processing. *IEEE SP Magazine*, pages 14–38, October 1991.

[22] G. Strang. Wavelets and Dilation Equations: A brief introduction. *SIAM Review*, 31(4):614–627, December 1989.

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