# MULTICHANNEL TIME-SERIES MODELLING AND PREDICTION BY WAVELET NETWORKS


*Aleš Procházka* [1], *Jonathan Smith* [2]

[1] Prague University of Chemical Technology, Department of Computing and Control Engineering
Technická 1905, 166 28 Prague 6, Czech Republic
Tel: +42 2 2435 4198; fax: +42 2 2431 1082
e-mail: prochaz@vscht.cz

[2] South Bank University, School of Engineering Systems and Design
103 Borough Road, London SE1 OAA, England,
Tel: +44 171 8157666; fax: +44 171 8157699
e-mail: smithjh@sbu.ac.uk



**ABSTRACT**

Multichannel time-series result from observations of a given engineering, biomedical, econometric or environmental variable taken at different locations. Processing this type of signal presents problems associated with its extrapolation in given space ranges and its possible prediction. This paper presents a comparison of seasonal AR modelling of such signals and the application of wavelet networks to the system identification and prediction of a particular signal. The choice of wavelet functions and the optimization of their coefficients is discussed as well. Each method suggested in the paper is verified for simulated signals at first and then used for the analysis of real signals, including the observation of air pollution. All algorithms are written in the MATLAB environment.


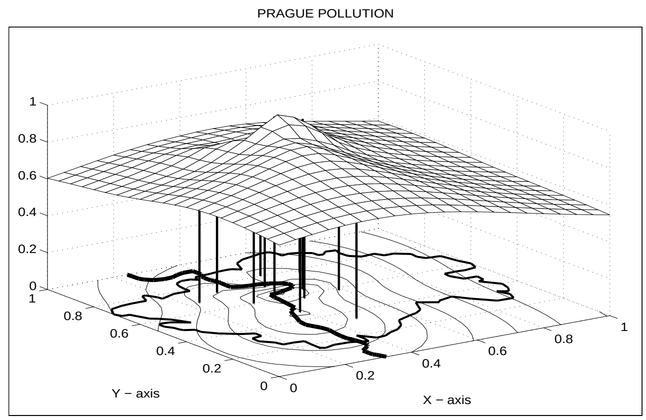


Figure 1: Interpolation of air pollution components over the given Prague region

## 1 INTRODUCTION

Multichannel observations are available in many problems, allowing subsequent system analysis and modelling. An example of such a system is presented in Fig. 1 derived from measurements of sulphur dioxide air pollution in 12 different locations in Prague and its extrapolation over the whole city region. Separate signal analysis is able to verify periodic signal components and the application of a STFT shows the time evolution of a given pollutant. An example of the results obtained from an analysis of data from the Santinka measuring station is given in Fig. 2.

Modelling of this type of system is studied in various papers based on different approaches. These include linear methods of signal analysis [14], nonlinear models often using artificial neural networks [7, 8, 13, 15, 18], wavelet functions [1, 9, 16, 20, 21] and wavelet networks [2, 17, 23, 24]. Many papers are also devoted to the use of radial basis functions [3, 10, 11, 22].

While model coefficient evaluation is relatively simple for linear systems, various optimization methods must be applied for nonlinear systems including gradient search and its modifications [6, 7] and genetic algorithms use [5, 19].

## 2 SYSTEM MODELLING

Modelling and prediction using both the classical and adaptive approaches [12] are usually based on linear models. To begin with, 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

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

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

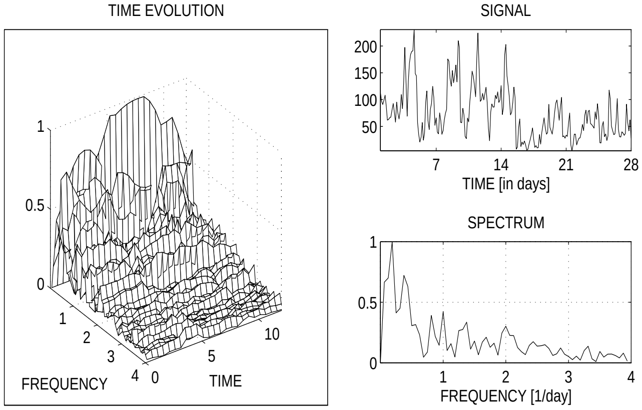


Figure 2: Analysis of air pollution components at the Santinka measuring station presenting **(a)** Time evolution resulting from the short-time Fourier transform, **(b)** Selected time-series plot and its spectral analysis

problem of optimal modelling [12] 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 second layer output $\mathbf{A2}$ $(= y(n))$ in the form

$$\mathbf{A1} = F1\,(\mathbf{W1} \star \mathbf{P} + \mathbf{B1}) \quad (2)$$

$$\mathbf{A2} = F2\,(\mathbf{W2} \star \mathbf{A1} + \mathbf{B2}) \quad (3)$$

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 is usually performed for a sigmoidal function $F1$ and a linear function $F2$. 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. 1.

The analysis provided in this paper is devoted to the use of wavelet functions in the first layer of neural networks using an algorithmic approach presented in Fig. 3. Appropriate signal preprocessing is assumed at first including delay-free filtering and normalization.

# 3 WAVELET NETWORKS

Neural networks are often used for nonlinear signal modelling and can provide better results compared to those achieved by linear methods, provided their structure and corresponding transfer functions are selected properly. Although sigmoidal functions are used in many cases, it is possible to apply other functions including wavelets, which have been studied in many papers recently [2, 24]

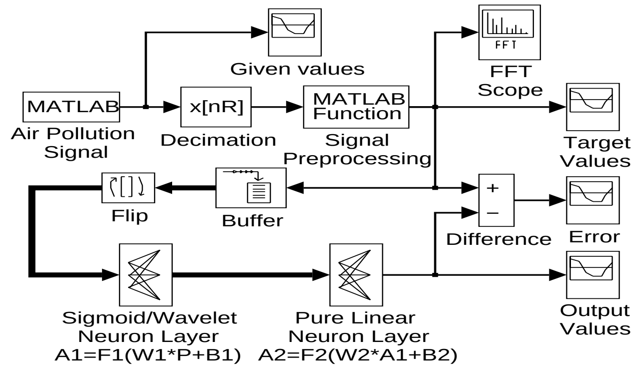


Figure 3: Air pollution time-series analysis and prediction by a neural network structure

in connection with signal analysis, compression and approximation.

The set of wavelet functions [1, 4, 9, 16] is usually derived from the initial (mother, basic) wavelet $h(t)$ dilated by value $a$, translated by $b$ and normalized defining (baby) functions

$$h_{a,b}(t) = \frac{1}{\sqrt{a}}\ h(\frac{t-b}{a}) \quad (4)$$

Wavelet coefficients are often defined in the discrete case by powers of two resulting in dilation $a = 2^m$ and translation $b = n\ 2^m$ implying

$$h_{m,n}(t) = \frac{1}{\sqrt{2^m}}\ h(2^{-m}\ t - n) \quad (5)$$

The basic wavelets can be either real [9] or complex [16]. It can be defined by an analytical expression or by a dilation equation, providing tools for wavelet transform evaluation based upon the pyramidal algorithm [20] which allows the decomposition and possible perfect reconstruction of the signal.

Some selected wavelets, defined in analytical form include

- Gaussian derivative

$$h(t) = -t\ e^{-t^2/2} \quad (6)$$

- Shannon wavelet function

$$h(t) = \frac{sin(\pi\ t/2)}{\pi\ t/2}\ cos(3\ \pi\ t/2) \quad (7)$$

- Modulated Gaussian (Morlet) function

$$h(t) = e^{j\omega_0\ t}\ e^{-t^2/2} \quad (8)$$

- Harmonic wavelet function

$$h(t) = \frac{1}{j\pi\ t}(e^{j\ 2\ \pi\ t} - e^{j\pi\ t}) \quad (9)$$

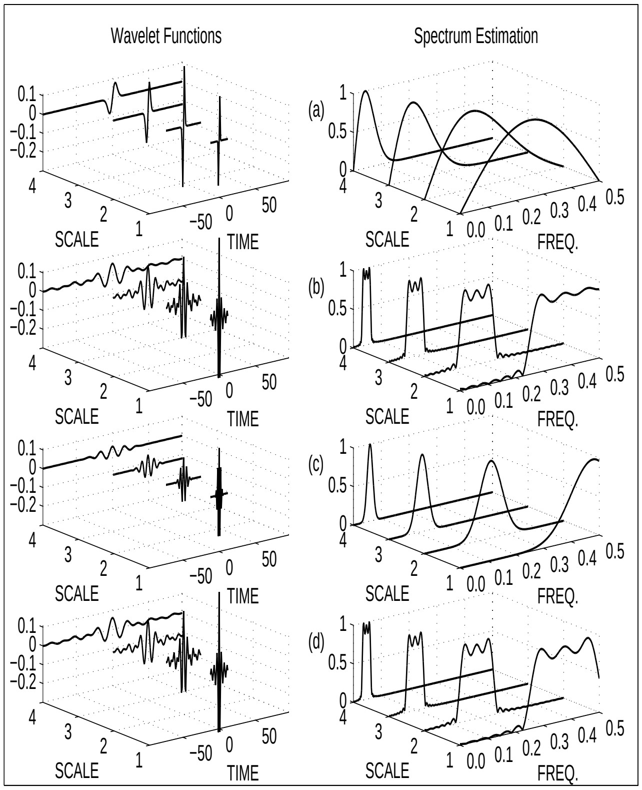


Figure 4: The set of original and dilated wavelet functions in the time and frequency domains for **(a)** Gaussian derivative, **(b)** Shannon wavelet, **(c)** Modulated Gaussian (Morlet) and **(d)** Harmonic wavelet function

From the signal processing point of view it is possible to consider the initial wavelet as a pass-band filter. Wavelet dilation by a factor $2^m$ is equivalent to a pass-band compression. This result is presented in Fig. 4 for various wavelet functions, including the harmonic wavelet function defined by Eq. (9) which was introduced by Prof. D. E. Newland [16]. The initial wavelet function covers the upper half of the normalized frequency range. This function is modified by a scaling index $m = 1, 2, \cdots$ using Eq. (5) causing it to dilate further and its corresponding spectrum to compress.

Wavelet functions can be used for signal analysis [9] using the wavelet transform. Because of its extremely desirable time and frequency localization properties, the wavelet transform can be used as an alternative to the short-time Fourier transform for feature extraction, parameter estimation and pattern recognition of non-stationary signals as well as their compression and coding. Another application of wavelet functions is in their use as transfer functions in neural networks for signal modelling and prediction.

Wavelet networks were introduced in [2, 24] to overcome the poor convergence of sigmoidal neural networks. The output of a $1 - S1 - 1$ network can be evaluated in the form

$$A2(n) = \sum_{j=1}^{S1} w2(1,j)\; h_{a_j,b_j}(p(n)) \qquad (10)$$

using the set of $S1$ wavelet functions $h_{a,b}(t)$. Initialization and optimization of such a network using genetic algorithms is discussed in [18] and in the case of radial basis functions, the results of [11] can be used for signal prediction. System modelling discussed below compares results obtained using different approaches to this problem.

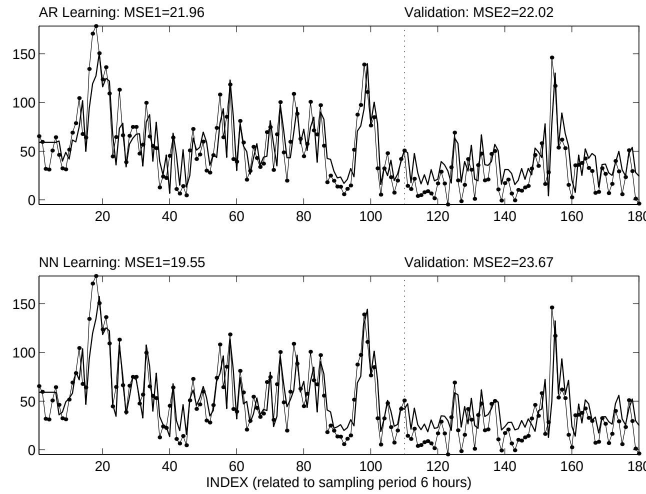


Figure 5: Prediction of one sample (6 hours) ahead based upon given (dotted) observations both in the learning and validation parts using **(a)** AR model of the 6th order and **(b)** Neural network 6-3-1 with Gaussian derivative wavelet transfer functions

Table 1: MEAN SQUARE ERROR OF ONE STEP AHEAD AIR POLLUTION PREDICTION AT THE SANTINKA STATION FOR VALUES DECIMATED TO 4 SAMPLES PER DAY COMPARING LINEAR AND NONLINEAR MODELLING BOTH IN THE LEARNING AND VALIDATION PARTS

| *Type* | *Ordinary Model* [1 2 3 4 5 6] | | *Seasonal Model* [1 2 3 7 8 9] | |
|---|---|---|---|---|
| | *Training* | *Valid.* | *Training* | *Valid.* |
| *AR Model* | 21.96 | 22.02 | 23.33 | 23.74 |
| *Tansig NN* | 20.79 | 24.08 | 19.27 | 27.63 |
| *Wavelet NN* | 19.55 | 23.67 | 19.53 | 28.77 |

# 4 RESULTS

The methods and algorithms discussed above have been applied to multichannel observations of air pollution in the Prague region, following their initial digital filtering, decimation and preprocessing. Results of a given time-series prediction both in the learning and validation parts are presented in Fig. 5, showing a comparison of the linear and nonlinear models.

Tab. 1 compares the results achieved by a linear model and nonlinear model after optimization of its coefficients, both for sigmoidal and Gaussian wavelet transfer functions and a selected network structure. All models produce errors of the same order in the given example for both ordinary and seasonal models. For a properly selected combination of several time-series in the given region to define pattern values it is possible to expect further better results in this case.

Problems which require further investigation include the selection of the optimal model structure and the design of efficient optimization methods.

# ACKNOWLEDGEMENT

This paper has been supported by the Department of Computing and Control Engineering and by grant No. 201/94/0130 of the Grant Agency of the Czech Republic. All real signals were kindly provided by the Meteorological Institute in Prague.

# References

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

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

[3] H. Demuth and M. Beale. *Neural Network Toolbox*. The MathWorks, Inc., Natick, Mass., Jan. 1994.

[4] N. J. Fliege. *Multirate Digital Signal Processing*. John Wiley & Sons, New York, 1994.

[5] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Company, Mass., 1989.

[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] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computing*. Addison-Wesley Publishing Company, California, 1991.

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

[10] S. Chen, C. F. N. Cowan, and P. M. Grant. Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks. *IEEE Trans. Neural Networks*, 2(2):302–309, 1991.

[11] 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.

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

[13] A. Lapedes and R. Farber. Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling, LA-UR-87-2602. Technical report, Los Alamos National Laboratory, USA, 1987.

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

[15] J. R. McDonnell and D. Waagen. Evolving recurent perceptrons for time-series modeling. *IEEE Trans. on Neural Networks*, 5(1):24–38, Jan. 1994.

[16] D. E. Newland. *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*. Longman Scientific & Technical, Essex, U.K., 3rd ed., 1994.

[17] A. Procházka, J. Jech, and J. Smith. Wavelet Transform Use in Signal Processing. In *31st International Conference in Acoustics*, pages 209–213. Czech Technical University, 1994.

[18] 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.

[19] A. Procházka and V. Sýs. Time Series Prediction Using Genetically Trained Wavelet Networks. In *Neural Networks for Signal Processing 3 - Proc. of the 1994 Workshop*, pages 195–203. IEEE, 1994.

[20] O. Rioul and M. Vetterli. Wavelets and Signal Processing. *IEEE SP Magazine*, Oct. 1991.

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

[22] B. A. Whitehead and T. D. Choate. Evolving space-filling curves to distribute radial basic functions over an input space. *IEEE Transactions on Neural Networks*, 5(1):15–23, January 1994.

[23] Q. Zhang. Regressor Selection and Wavelet Network Construction. Technical report, IRISA, France, 1993.

[24] Q. Zhang and A. Benveniste. Wavelet Networks. *IEEE Trans. on Neural Networks*, 3(7):889–899, 1992.