16

Wavelet Use for Noise Rejection and Signal Modelling

Aleš Procházka[1]
Martina Mudrová[1]
Martin Štorek[1]

ABSTRACT
Wavelet functions form an essential tool of modern signal analysis and modelling in various engineering, biomedical and environmental systems. The paper is devoted to selected aspects of construction of wavelet functions and their properties enabling their use for time-scale signal analysis and noise rejection. The main part of the paper presents the basic principles of signal decomposition in connection with fundamental signal components detection and approximation. The methods are verified for simulated signals at first and then used for real signal analysis representing air pollution inside a given region of North Bohemia covered by several measuring stations. The following part of the chapter is devoted to comparison of linear and non-linear signal modelling and prediction for original and de-noised signals. The paper emphasizes the algorithmic approach to methods of signal analysis and prediction and it presents the use of wavelet functions both for signal analysis and modelling.

16.1 Introduction

Signal analysis is one of the essential parts of digital signal processing in various information and control systems and is often based upon different discrete transforms. The wavelet transform forms in many cases a very efficient alternative to the short-time Fourier transform enabling signal analysis with varying time-frequency resolution. Its application further includes signal components detection and signal decomposition, data compression and noise rejection both for one-dimensional and two-dimensional signals [14, 15, 24]. The use of wavelet and radial basis functions as transfer functions of neural networks for signal modelling and prediction is presented in various papers as well [3, 9, 11, 21, 27]. Problems closely related to this topic include signal segmentation [6, 19], restoration, statistical signal processing [13] and efficient optimization methods [7, 20].

The theory of wavelet transforms is studied in many papers and books published recently [2, 15, 22, 23, 25] describing also the relations between the wavelet trans-

[1]Institute of Chemical Technology, Department of Computing and Control Engineering, Technická 1905, 166 28 Prague 6, Czech Republic,
E-mails: {Ales.Prochazka, Martina.Mudrova, Martin.Storek}@vscht.cz

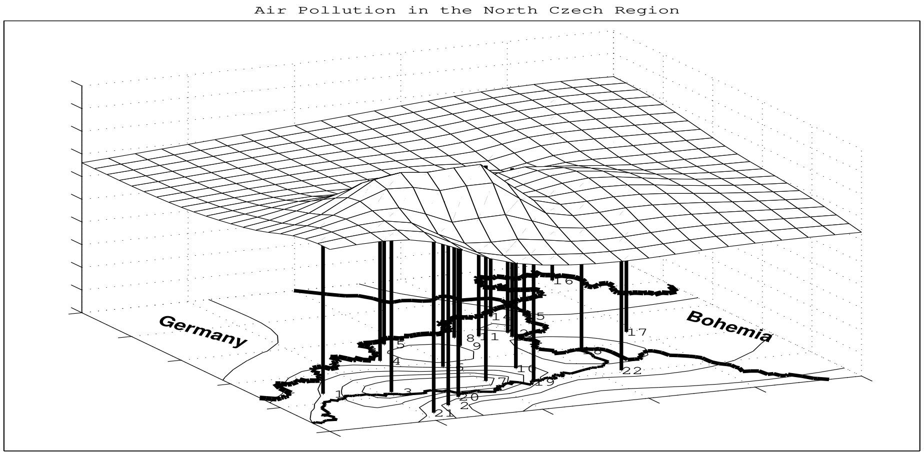

FIGURE 16.1. Air pollution in North Bohemia extrapolated from simultaneous measurements of 22 stations at a selected time

form and the short-time Fourier transform [10]. The analysis and the construction of wavelet functions both in the analytical and in the recurrent forms with description of their properties are important results of these studies.

The chapter is devoted to basic principles of signal analysis and decomposition by wavelet transform in connection with rejection of noise components forming an additive part of a given sequence. Results of these studies are applied to the processing of a given multichannel signal representing air pollution in North Bohemia. An example of such a signal is given in Fig. 16.1 using measurements at a selected time from 22 measuring stations in this region for signal interpolation and extrapolation. The goal of such a signal processing is to analyse noisy data to provide information about air pollution in the whole region and to predict the situation several sampling periods ahead. General methods of signal de-noising and prediction studied in the chapter are closely related to some other principles described in [1, 12, 18].

16.2 Signal Analysis and Wavelet Decomposition

Wavelet transforms (WT) provides an alternative to the short-time Fourier transform (STFT) for non-stationary signal analysis and both STFT and WT result in signal decomposition into two-dimensional function. Its independent variables stand for time and frequency in case of the STFT and they represent time and scale (closely related to frequency) in case of the WT. The basic difference between these two transforms is in the construction of the window function which has a constant length in case of the STFT (including rectangular, Blackman and other window functions) while in case of the WT wide windows are applied for low frequencies and short windows for high frequencies to ensure constant time-frequency resolution. Local and global signal analysis can be combined in this way.

In case of a continuous (non-stationary) signal $x(t)$ the continuous WT coefficients are given by relation

$$C(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\ W(\frac{t-b}{a})\ dt \tag{16.1}$$

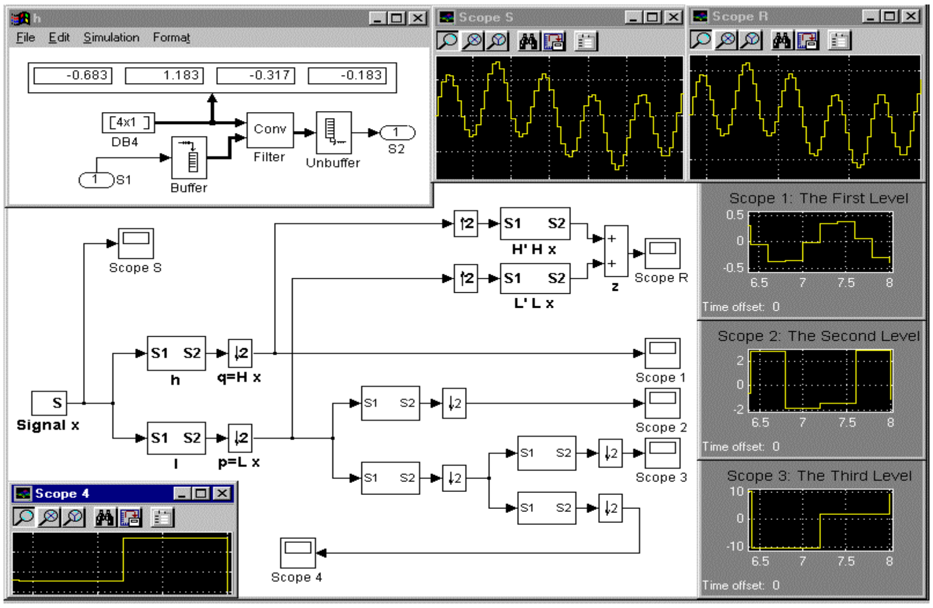

FIGURE 16.2. The Simulink model of the pyramidal filter bank structure used **(a)** to evaluate and to present wavelet transform coefficients for a given sequence $\{x(n)\}_{n=0}^{N-1}$ and complementary low-pass and high-pass filters at selected levels **(b)** to reconstruct the given sequence of values

where $W(t)$ is the basic (mother) wavelet function and values a and b stand for its translation and dilation respectively. There are strict criteria for the choice of such a basic wavelet function given by a requirement of the transforms reversibility [10, 14]. The most frequently used wavelets include Daubechies, Shannon, Morlet, Gaussian and harmonic functions [10, 15].

The discrete WT assumes the evaluation of wavelet coefficients for discrete parameters of dilation $a = 2^m$ and translation $b = k\ 2^m$ using only the initial (mother) wavelet function $W(t)$ localized at a particular position. The dilated and translated version in the form

$$W_{m,k}(t) = \frac{1}{\sqrt{a}}\, W\,(\frac{1}{a}\,(t-b)) = \frac{1}{\sqrt{2^m}}\, W\,(2^{-m}t - k)) \qquad (16.2)$$

is then used for signal analysis and description. Two basic approaches include definition of wavelet functions either in the analytical form or by recurrent equations. A very commonly used Daubechies wavelet function of order N is defined through the scaling function $\Phi(t)$ resulting from equation

$$\Phi(t) = l(0)\ \Phi(2t) + l(1)\ \Phi(2t-1) + l(2)\ \Phi(2t-2) + \cdots + l(L-1)\ \Phi(2t-L+1) \quad (16.3)$$

for coefficients $\{l(n)\}_{n=0}^{L-1}$. The (iterative) solution of this equation is based upon the relation

$$\Phi_j(t) = l(0)\ \Phi_{j-1}(2t) + l(1)\ \Phi_{j-1}(2t-1) + \cdots + l(L-1)\ \Phi_{j-1}(2t-L+1)$$

for $j = 1, 2, \cdots$ and the given starting Haar function

$$\Phi_0(t) = \begin{cases} 1 & \text{for } t \in \langle 0, 0.5) \\ 0 & \text{in other cases} \end{cases} \qquad (16.4)$$

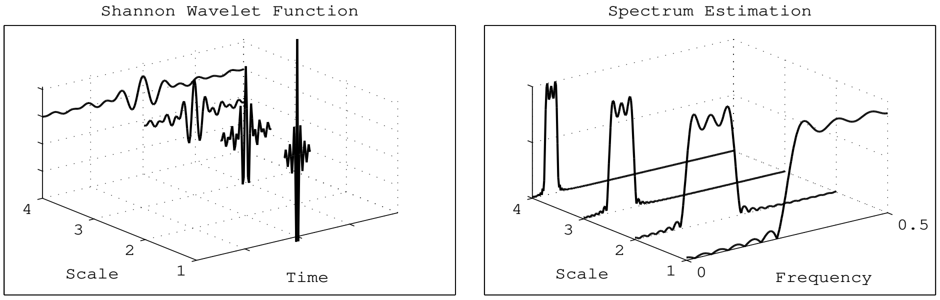

FIGURE 16.3. Analysis of the set of Shannon wavelet functions with its initial function in the form $W(t) = \sin(\pi * t/2)/(\pi * t/2) * \cos(3 * \pi * t/2)$ presenting its dilation and the corresponding spectrum compression

The initial wavelet function is then defined by relation

$$W(t) = h(0)\ \Phi(2t) + h(1)\ \Phi(2t-1) + h(2)\ \Phi(2t-2) + \cdots + h(L-1)\ \Phi(2t-L+1)$$

where $\{h(n)\}_{n=0}^{L-1} = \{(-1)^{n+1} l(L-1-n)\}_{n=0}^{L-1}$. The resulting set of wavelet functions can be then used for approximation of a given signal $x(t)$ in the form [15]

$$x(t) = c(0)\Phi(t) + c(1)W(t) + [c(2)\ c(3)] \begin{bmatrix} W(2t) \\ W(2t-1) \end{bmatrix} + [c(4)\ c(5)\ c(6)\ c(7)] \begin{bmatrix} W(4t) \\ W(4t-1) \\ W(4t-2) \\ W(4t-3) \end{bmatrix} + \ldots$$

for the set of wavelet transform coefficients $\{c(n)\}$.

The wavelet transform coefficients for a given column vector $\{x(n)\}_{n=0}^{N-1}$ may be evaluated efficiently using Mallat's pyramidal structure as shown in Fig. 16.2. This model enables the study of signal decomposition and reconstruction based upon the convolution between values of the given sequence and filter coefficients as shown in detail for block l in the upper-left corner of the Fig. 16.2 using Daubechies wavelet of order 4 in this case. From the signal processing point of view this algorithm assumes the use of the half band low-pass scaling sequence $\{l(n)\}_{n=0}^{L-1}$ to generate the corresponding wavelet sequence $\{h(n)\}_{n=0}^{L-1}$ which is convolved with the signal $\{x(n)\}$ and subsampled by two. Introducing decomposition matrices

$$\mathbf{L}_{N/2,N} = \begin{bmatrix} l(1) & l(0) & 0 & 0 & \cdots \\ l(3) & l(2) & l(1) & l(0) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & l(L-1) & \cdots & l(0) \end{bmatrix} \tag{16.5}$$

$$\mathbf{H}_{N/2,N} = \begin{bmatrix} h(1) & h(0) & 0 & 0 & \cdots \\ h(3) & h(2) & h(1) & h(0) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & h(L-1) & \cdots & h(0) \end{bmatrix} \tag{16.6}$$

it is possible to define the separation process for resulting signals of the first level decomposition $\mathbf{p} = \mathbf{L}\ \mathbf{x}$ and $\mathbf{q} = \mathbf{H}\ \mathbf{x}$. Signal reconstruction may be achieved by introducing the transposed matrices $\mathbf{L}_{N,N/2}^T$ and $\mathbf{H}_{N,N/2}^T$ as shown in Fig. 16.2 and the reconstructed signal $\mathbf{z}$ is given by

$$\mathbf{z} = (\mathbf{L}^T * \mathbf{L} + \mathbf{H}^T * \mathbf{H}) * \mathbf{x} = \mathbf{T}^T * \mathbf{T} * \mathbf{x} \tag{16.7}$$

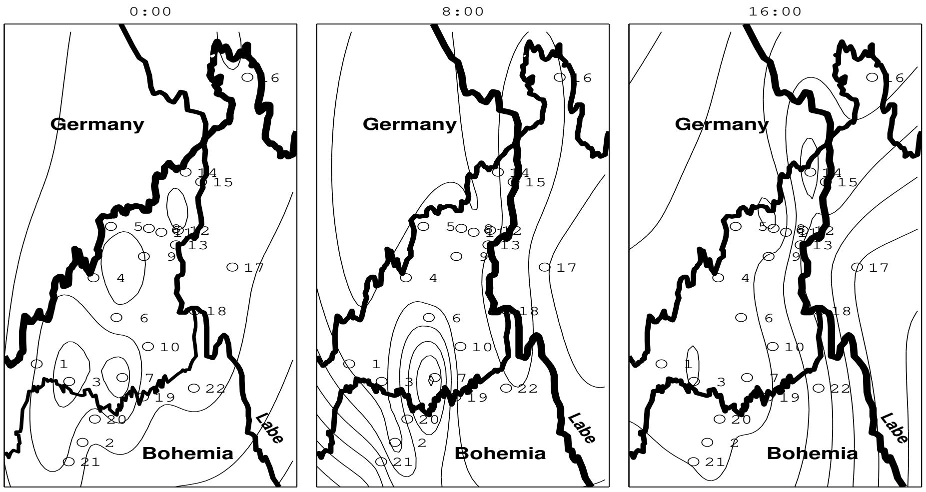

FIGURE 16.4. One day evolution of air pollution in North Bohemia extrapolated from 22 measuring stations in this region

The requirement for perfect signal reconstruction gives the necessary and sufficient condition for matrix $\mathbf{E} = \mathbf{T}^T\mathbf{T}$ to be the identity matrix. Matrix $\mathbf{T}$ defined by the original matrices $\mathbf{L}$ and $\mathbf{H}$ in the form

$$\mathbf{T} = \begin{bmatrix} \cdots & & & \cdots & \cdots & & & \cdots \\ l(L-1) & l(L-2) & l(L-3) & \cdots & l(0) & 0 & 0 & \cdots \\ h(L-1) & h(L-2) & h(L-3) & \cdots & h(0) & 0 & 0 & \cdots \\ 0 & 0 & l(L-1) & l(L-2) & \cdots & \cdots & l(0) & \cdots \\ 0 & 0 & h(L-1) & h(L-2) & \cdots & \cdots & h(0) & \cdots \\ \cdots & & & \cdots & \cdots & & & \cdots \end{bmatrix}$$

must be orthonormal in this case. The whole process of signal reconstruction from the first level of wavelet decomposition is illustrated in Fig. 16.2.

Various wavelet and scaling sequences can be used for signal analysis defined above. Their dilation correspond to spectrum compression according to Fig. 16.3. The most common choice include Daubechies wavelets even though their frequency characteristics give only an approximation to band-pass filters. On the other hand harmonic wavelets introduced in [15] can have broader application in many engineering problems owing to their very attractive spectral properties.

16.3 Real System Description

The region of North Bohemia represents a heavily industrial area with problems of air pollution which is monitored by several measuring stations to enable extrapolation and prediction of selected pollutants. As some measuring devices can be corrupted and resulting signal can be affected by the additive noise it is necessary to interpolate missing values first and then to reject signal noise components. The autoregressive (AR) method introduced in [6] can be applied very efficiently for restoration of missing values in the first stage of signal processing. The map with

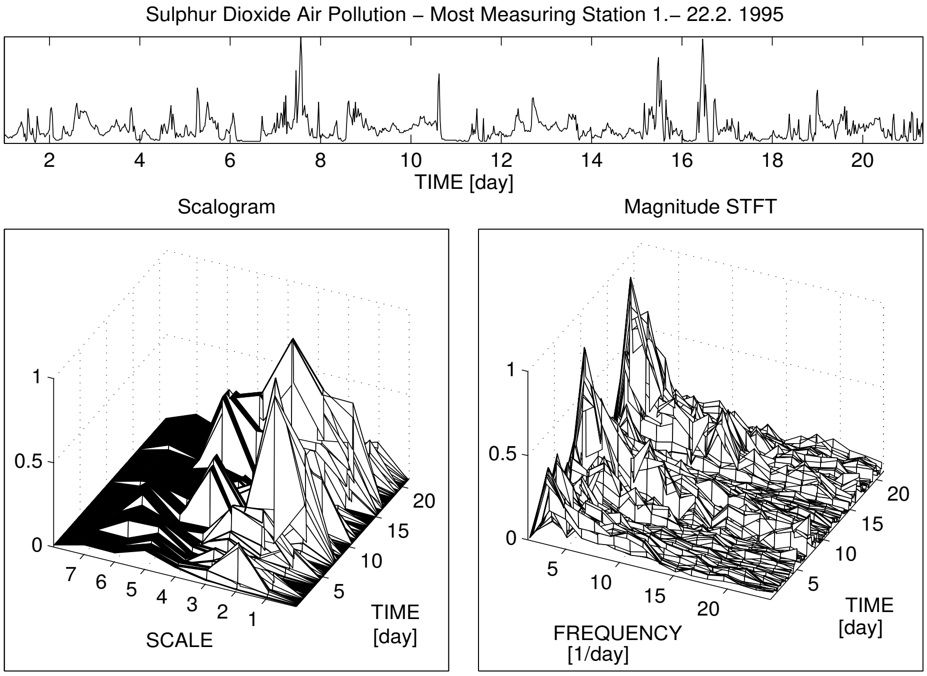

FIGURE 16.5. Spectrogram and scalogram of a given data segment representing air pollution measured in the city of Most during a three weeks period

the evolution of air pollution during a chosen day is presented in Fig. 16.4. The following preliminary analysis of a selected time series is given in Fig. 16.5 comparing results obtained by the STFT and WT. Owing to its definition the Wavelet transform provides better localization of signal peaks in this case thus enabling the use of an appropriate level of wavelet decomposition for the precise detection of extreme values of air pollution.

16.4 Signal Noise Rejection

Signal decomposition by a given set of wavelet functions into separate levels or scales results in the set of wavelet transform coefficients that can be used for signal compression, signal analysis, and in the case that the coefficients are not modified, they allow perfect signal reconstruction. In the case when only selected levels of signal decomposition are used, or wavelet transform coefficients are processed, it is possible to extract signal components or to reject its undesirable parts. Using the threshold method introduced by [24, 4] it is further possible to reject noise and to improve signal to noise ratio.

The de-noising algorithm assumes that the signal has low frequency components and that it is corrupted by additive Gaussian white noise with its power much lower than that of the analyzed signal. The whole method consists of the following steps:

- Signal decomposition using a chosen wavelet function up to the selected level and evaluation of wavelet transform coefficients

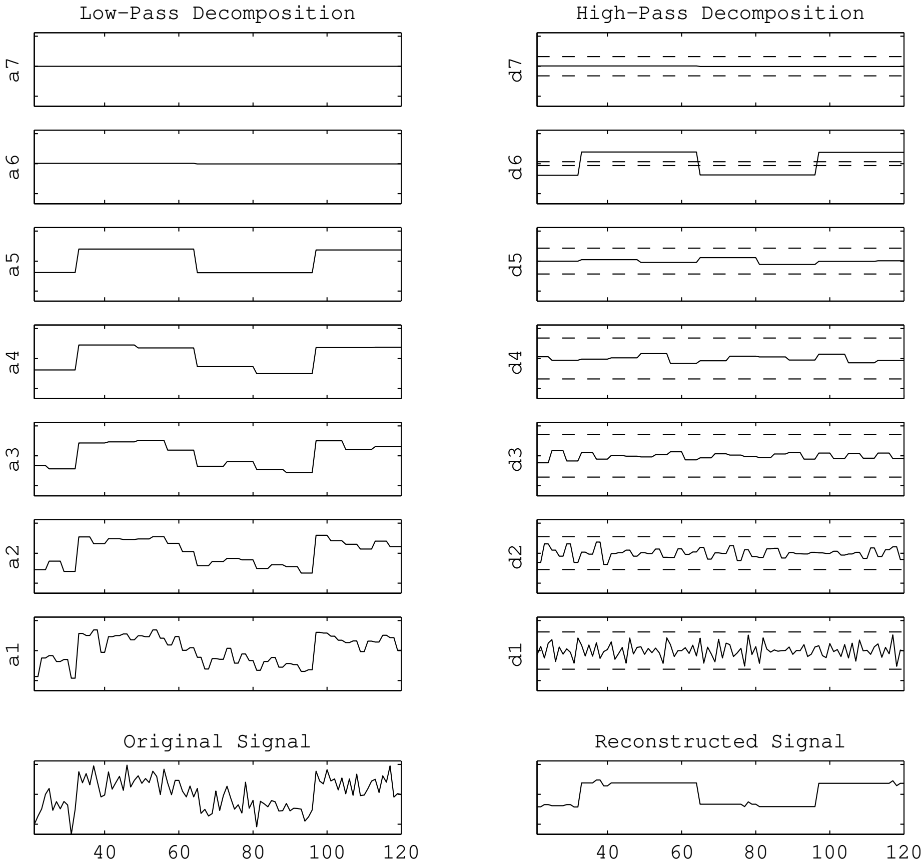

FIGURE 16.6. The use of the Haar wavelet function for decomposition of the given simulated signal $x(n) = 1$ for $n = N/4, \cdots, N/2-1, 3N/4, \cdots, N-1$ and $x(n) = 0$ for other values of n ($N = 128$) with an additive noise component and signal reconstruction for given threshold limits

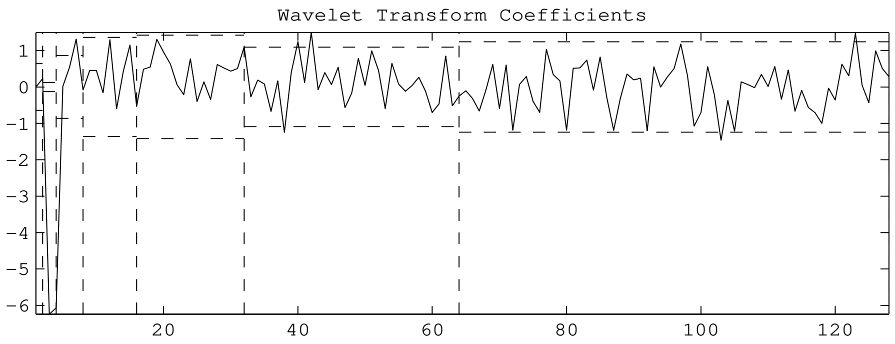

FIGURE 16.7. Wavelet transform coefficients $\{c(k)\}_{k=0}^{N-1}$ for the given signal $\{x(n)\}_{n=0}^{N-1}$ with its length $N = 128$ and selected threshold limits used for signal de-noising presenting segments of coefficients belonging to is decomposition level $l = D, D-1, \cdots, 1$ where $D = log_2(N)$

- The choice of threshold limits for each decomposition level and modification of its coefficients
- Signal reconstruction from modified wavelet transform coefficients

The success of this process depend upon the proper choice of wavelet functions, selection of threshold limits and their use.

The application of threshold limits to modify wavelet coefficients $\{c(k)\}_{k=0}^{N-1}$ include two basic approaches. The soft thresholding formula, for a chosen thresholding value δ, gives the new coefficients as

$$\overline{c}_{soft}(k) = \begin{cases} \text{sign } c(k) \ (\mid c(k) \mid -\delta) & \text{if} \mid c(k) \mid > \delta \\ 0 & \text{if} \mid c(k) \mid \leq \delta \end{cases} \tag{16.8}$$

The hard thresholding method gives

$$\overline{c}_{hard}(k) = \begin{cases} \text{sign } c(k) & \text{if} \mid c(k) \mid > \delta \\ 0 & \text{if} \mid c(k) \mid \leq \delta \end{cases} \tag{16.9}$$

Hard thresholding is the simplest method while soft thresholding has better mathematical properties and it is more appropriate for signal de-noising.

Selection of threshold limits has been studied extensively [4]. The choice is based either on statistical analysis of the given signal and its wavelet transform coefficients or on heuristic approaches. Basic adaptive threshold selection estimates the value of δ by the Stein's Unbiased Risk Estimate (SURE) [14]. Other methods include the fixed method, which assumes the value of $\delta = \sqrt{2\ \log{(N)}}$ and minimax method [4]. These methods result from the minimization of the mean-squared error to obtain a smooth reconstructed signal. Different thresholds can be further evaluated for separate levels by Birge-Massart strategy and by estimation of noise in the wavelet coefficients [5]. In the case that no thresholding is applied and all coefficients are used it is possible to obtain perfect reconstruction of the original signal $\{x(n)\}_{n=0}^{N-1}$. Properly chosen wavelet functions and threshold levels enable efficient extraction of low pass signal components and improvement of signal to noise ratio.

Fig. 16.6 together with Fig. 16.7 provides results of signal de-noising for a simulated sequence representing step function with an additive noise. In the case of step signal changes the Haar wavelet function can be used very efficiently with properly chosen threshold limits. Fig. 16.7 presents the sequence of these coefficients and limits used for signal processing. The same method can be applied for other signals and in many cases Daubechies wavelet functions can be used for signal de-noising in the similar way. Fig. 16.8 provides selected results of the real signal de-noising representing air pollution data. Such signal preprocessing is very useful for further signal modelling and prediction.

16.5 Signal Prediction

The problem of parameter estimation is very common in signal analysis, classification, prediction and system identification as well. In the simplest case it is possible to use the autoregressive (AR) model [8, 16, 17] of order N to describe a signal $\mathbf{x} = [x(0), x(1), \cdots, x(M{+}N{-}1)]$ by a vector of coefficients $\mathbf{a} = [a(0), \cdots, a(N{-}1)]'$.

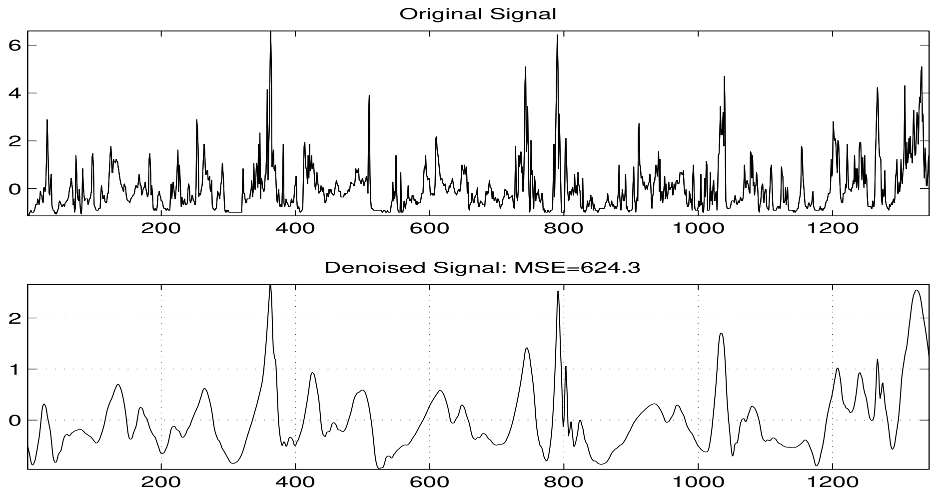

FIGURE 16.8. Wavelet de-noising applied for real signal representing a chosen segment of air pollution sequence with the sampling period of $T_s = 3$ hours processed by the Daubechies wavelet function of the 8th order

The problem of one step ahead prediction results in the solution of M equations of N variables $\mathbf{A}_{M,N}\ \mathbf{a}_{N,1} = \mathbf{y}_{M,1}$ in the form

$$\begin{bmatrix} x(N-1) & x(N-2) & \cdots & x(0) \\ x(N) & x(N-1) & \cdots & x(1) \\ & & \cdots & \\ x(N+M-2) & x(N-M-3) & \cdots & x(M-1) \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ \cdots \\ a(N-1) \end{bmatrix} = \begin{bmatrix} x(N) \\ x(N+1) \\ \cdots \\ x(N+M-1) \end{bmatrix} \quad (16.10)$$

The AR model order N can be estimated from the signal autocorrelation function as the number of its significant values [12]. Then it is possible to apply the singular value decomposition (SVD) for the linear regression model order selection and the following QR factorization for the selection of independent variables in the regression problem which make a fundamental contribution to the quality of the mathematical model. Such subset selection, described in [12, 26], is used in various applications of system identification and control and it allows the simplification of the autoregressive model especially in case of seasonal time series with significant periodic components.

Other methods of signal prediction are based on adaptive systems and non-linear models using artificial neural networks (ANN) in many cases. Let us assume the ANN with R inputs, $S1$ nodes of its first layer with non-linear (sigmoidal or wavelet) transfer function $F1$ and one output node with the linear transfer function. Applying the input column vector $\mathbf{P}_{R,1}$ to such ANN it is possible to evaluate its output

$$\mathbf{Y} = \mathbf{W2}{*}F1\ (\mathbf{W1} * \mathbf{P} + \mathbf{B1}) + \mathbf{B2} \quad (16.11)$$

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

To achieve better of signal prediction it is useful in many cases to apply signal de-noising in the first step before the chosen prediction method is applied. Selected results of AR modelling are presented in Fig. 16.9 comparing that achieved both

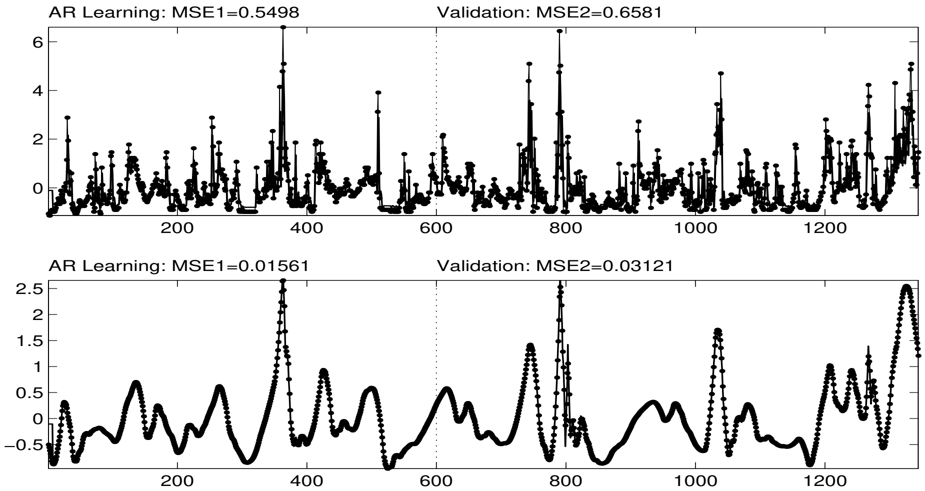

FIGURE 16.9. The prediction of a real signal with its sampling period $T_s = 3$ hours long for the original and de-noised signal using the autoregressive model of the 7th order

Type	*Original Signal*		*De-noised Signal*	
	Learning	*Validation*	*Learning*	*Validation*
AR Model	0.5498	0.6581	0.0156	0.0312
Tansig NN	0.4970	0.7728	0.0096	0.0606
Wavelet NN	0.5412	0.6598	0.0132	0.0328

TABLE 16.1. Mean square error of three hours air pollution prediction at the city of Most measuring station

for original signals and de-noised signals using wavelet signal decomposition described above. The mean square errors both in the learning and validation parts are summarized in Table 16.1 for various models and for original and de-noised signals. The experiments indicate that with signal de-noising in the preprocessing, similar prediction errors are achieved with AR models as are achieved with selected non-linear structures.

16.6 Conclusion

The contribution describes selected properties of the wavelet transform compared with that of the short-time Fourier transform. The main part of the paper presents

- the ability of the WT to decompose a given signal
- wavelet coefficient modification for noise rejection
- signal reconstruction

Methods and algorithm presented in the paper are applied for processing of a real signal representing air pollution to interpolate its missing values, to increase its signal to noise ratio and to predict its behaviour. Linear and non-linear methods are compared to find the most efficient method for given real signal modelling. Further problems include missing values restoration, analysis of wavelet de-noising methods, neural network transfer functions choice and their coefficients optimization for signal prediction.

Acknowledgments: 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 Czech Hydrometeorological Institute in Prague.

16.7 References

[1] P. Arena, S. Baglio, L. Fortuna, and G. Nunnari. Neural Networks to Predict Ozone Pollution in Industrial Areas. In *VIIIth European Signal Processing Conference EUSIPCO-96*. European Association for Signal Processing, 1996.

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

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

[4] D.L. Donoho. De-Noising by Soft-Thresholding. *IEEE Trans. on Information Theory*, 41(3):613–627, 1995.

[5] D.L. Donoho and I.M. Johnstone. Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrica*, pp. 425–485, 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, 1996.

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

[8] M.H. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley and Sons, Inc., New York, 1996.

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

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

[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] S.M. Kay. *Fundaments of Statistical Signal Processing*. Prentice Hall, 1993.

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

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

[16] B. Porat. *A Course in Digital Signal Processing.* John Wiley and Sons, Inc., New York, 1997.

[17] J.G. Proakis and D.G. Manolakis. *Digital Signal Processing.* Prentice Hall, 1996.

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

[19] A. Procházka, M. Sláma, and E. Pelikán. Bayesian Estimator Use in Signal Processing. *Neural Network World*, 6(2):209–213, 1996.

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

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

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

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

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

[25] M. Vetterli. Wavelets and Filter Banks: Theory and Design. *IEEE Trans. Signal Processing*, 40(9):2207–2232, 1992.

[26] M. Štorek and A. Procházka. Neural Networks in Signal Prediction. In *The First European Conference on Signal Analysis and Prediction ECSAP-97*, pp. 228–231. ICT Press, 1997.

[27] B.A. Whitehead and T.D. Choate. Cooperative-Competitive Genetic Evolution of Radial Basis Functions Centers and Widths for Time Series Prediction. *IEEE Trans. Neural Networks*, 7(4):869–880, 1996.