WAVELET-BASED ROBUST COHERENCE MEASURE FOR DETECTING COLLECTIVE EFFECTS WITHIN MULTIPLE TIME SERIES

Alexey A. Lyubushin

Russian Academy of Sciences, Institute of the Physics of the Earth,
123995, Moscow, Russia, Bolshaya Gruzinskaya, 10;
e-mail: lubushin@mtu-net.ru

Abstract. The method for detecting collective effects within scalar components of multi-dimensional time series is presented. It is based on wavelet decomposition and was elaborated in for the problems of geophysical monitoring and earthquake prediction. The method is a wavelet modification of previously elaborated methods of collective effects extracting, which were realized for Fourier decomposition and using of classic multi-dimensional parametric models of multiple time series. The method constructs an estimate of scale-dependent measure of coherence behavior in a moving time window and is illustrated by processing multidimensional financial time series with the purpose to seek bursts of collective behavior within their variations.

INTRODUCTION

The method for detecting collective effects within scalar components of multi-dimensional time series is presented. The method is based on wavelet decomposition [Chui, 1992; Daubechies, 1992; Mallat, 1998] and was elaborated in [Lyubushin, 2000-2002; Lyubushin and Kopylova, 2004] for the problems of geophysical monitoring and earthquake prediction. The method is a wavelet modification of previously elaborated methods of collective effects extracting, which were realized for Fourier decomposition and using of classic multi-dimensional parametric models of multiple time series [Lyubushin, Lyubushin et al., 1998-1999]. The wavelet-based approach is more preferable because it is the most suitable for investigating transient effects within signals.

The method constructs an estimate of scale-dependent measure of coherence behavior in a moving time window. The scale-dependent coherence measure on the given detail level within given time window is the product of absolute values of canonical correlation coefficients [Hotelling, 1936; Rao, 1965] of wavelet coefficients of each scalar signal with respect to wavelet coefficients of all other signals. Thus, if we consider $\mathbf{q}, \mathbf{q} \geq 3$ time series then the wavelet-based coherence measure equals to the absolute value of the product of $\mathbf{q}$ canonical correlations. Absolute value of each canonical correlation describes “the strength” of connection of the considered scalar time series with the set of all other time series on the given detail level. It means that the product of $\mathbf{q}$ such values describes the strength of summary effect of collective behavior of the multiple time series. The main details of computing the used coherence measure are described in [Lyubushin, 2002], see also [Lyubushin, Kopylova, 2004]. It is essential to underline that the estimates of canonical correlations are constructed as a robust method [Huber, 1981], i.e. they are stable to the presence of outliers within data or within values of wavelet coefficients.

The method is illustrated by processing multidimensional financial time series with the purpose to seek bursts of collective behavior within their variations.

DESCRIPTION OF THE METHOD

Let $\mathbf{q} \geq 3$ be a general number of scalar time series to be analyzed simultaneously, τ be time index corresponding to the right-hand end of the moving time window of the length $\mathbf{N}$ samples. Let $\mathbf{M} = \min\{2^m : 2^m \geq \mathbf{N}\}$ be the minimum integer value of the type 2 in the integer power degree which is equal or greater than $\mathbf{N}$.

For each fragment of initial scalar time series corresponding to the current position of time window the following sequence of preprocessing operations are performed:

(i) the fragment is clarified from general linear trend;
(ii) optional: coming to time increments for the fragment;
(iii) the fragment is exposed to the procedure of *tapering* by cosine time window;
(iv) the fragment is renormalized to have a unit sample standard deviation, i.e. each value of the fragment is divided by the sample estimate of the standard deviation;
(v) the fragment is appended by zero values till the general length $\mathbf{M = 2^m}$.

Removing general linear trend for the fragment is a standard operation for suppressing those low-frequency components which could not be estimated statistically significant due to the finite length of the window. Optional operation (ii) of coming to time increments is an additional tool for making the sample within current time window to be more stationary. The tapering procedure [Jenkins, Watts, 1968] consists in multiplying the fragments of time series by the value which is equal to 1 for all inner samples of the time window and is going to zero when the time index is going to the left-hand or to the right-hand ends of the window. The law of going to zero of the tapering function is cosine. Afterwards (i.e. after general trends removing, tapering, renormalizing for unit standard deviation and appending by zero values) we perform discrete fast wavelet transforms [Press et al., 1996] of all these fragments using some orthogonal wavelet (Haar's wavelet for instance). Appending by zero values till the length $\mathbf{M = 2^m}$ is necessary for applying the fast wavelet transform.

As the result of the operations described above we will have a set of $\mathbf{q}$ wavelet coefficients for detail levels $\alpha = \mathbf{1,...,m}$.

$$\mathbf{c_j^{(\alpha,\tau)}(k), \quad j = 1,...,q; \quad \alpha = 1,...,m; \quad k = 1,...,M_\alpha = 2^{(m-\alpha)}} \tag{1}$$

The wavelet coefficients within the detail level, α describe variations of the signals with scales from $\mathbf{\Delta t \cdot 2^\alpha}$ till $\mathbf{\Delta t \cdot 2^{(\alpha+1)}}$, i.e. the 1st detail level is the most high-frequency. The number of wavelet coefficients on the detail level α is equal to $\mathbf{2^{(m-\alpha)}}$. But only the first part of $\mathbf{L_\alpha = 2^{(m-\alpha)} \cdot (N/M) = N \cdot 2^{-\alpha}}$ wavelet coefficients corresponds to the initial, non-appended by zero values parts of time series. Index $\mathbf{k}$ within formula (1) defines the position of scale-dependent vicinity inside the current moving time window which has the influence on the value of wavelet coefficient due to applying orthogonal finite support wavelet transform of the fragment.

Let $\mathbf{j_0}$ be a number of some time series, $\mathbf{1 \le j_0 \le q}$, and let us try to construct a measure describing the connection of the selected time series $\mathbf{j_0}$ with all other scalar time series within current time window. This measure should be scale-dependent of course. For this purpose let us consider a linear combination:

$$\sum_{\mathbf{j=1,\, j \ne j_0}}^{\mathbf{q}} \mathbf{c_j^{(\alpha,\tau)}(k)} \cdot \gamma_{\mathbf{j}} \tag{2}$$

with unknown coefficients $\gamma_{\mathbf{j}}$. It should be underlined that the index $\mathbf{j_0}$ is omit within the sum in the formula (2). The values of coefficients $\gamma_{\mathbf{j}}$ are found from solution the following minimization problem:

$$\sum_{\mathbf{k=1}}^{\mathbf{L_\alpha}} \left| \mathbf{c_{j_0}^{(\alpha,\tau)}(k)} - \sum_{\mathbf{j=1,\, j \ne j_0}}^{\mathbf{q}} \mathbf{c_j^{(\alpha,\tau)}(k)} \cdot \gamma_{\mathbf{j}} \right| \to \min_{\gamma_{\mathbf{j}}} \tag{3}$$

The robustness [Huber, 1981] of the procedure consists in minimizing the sum of absolute values in the formula (3) instead of the least squares approach, which follows to classical canonical correlation scheme of Hotelling [Hotelling, 1936; Rao, 1965]. For least squares approach the unknown coefficients $\gamma_{\mathbf{j}}$ in linear combination (2) are found by linear operation with sub-matrices of covariance matrix of wavelet coefficients whereas the robust minimization problem (3) must be solved numerically by the method of generalized gradient [Clarke, 1975; Shor, 1979] within each position of the moving time window.

Let $\gamma_{j,j_0}^{(\alpha,\tau)}$ be a solution of the problem (3). Let us define a sequence of scalar values:

$$\mathbf{d}_{j_0}^{(\alpha,\tau)}(\mathbf{k}) = \sum_{j=1,\, j\neq j_0}^{q} \mathbf{c}_j^{(\alpha,\tau)}(\mathbf{k}) \cdot \gamma_{j,j_0}^{(\alpha,\tau)} \quad (4)$$

which could be called a *robust canonical wavelet coefficients* for the selected time series $\mathbf{j_0}$. Now let us calculate correlation coefficient between the values of canonical coefficients $\mathbf{d}_{j_0}^{(\alpha,\tau)}(\mathbf{k})$ and initial wavelet coefficients $\mathbf{c}_{j_0}^{(\alpha,\tau)}(\mathbf{k})$ for $\mathbf{k = 1,...,L_\alpha}$. In order to provide robustness in all stages of final estimate we will use the robust formula for correlation coefficient [Huber, 1981] $\rho(\mathbf{x},\mathbf{y})$ for the pair of samples $\mathbf{x(k), y(k), k = 1,...,n}$:

$$\rho(\mathbf{x},\mathbf{y}) = \frac{\mathbf{S}(\hat{\mathbf{z}}^2) - \mathbf{S}(\breve{\mathbf{z}}^2)}{\mathbf{S}(\hat{\mathbf{z}}^2) + \mathbf{S}(\breve{\mathbf{z}}^2)} \quad (5)$$

where

$$\begin{aligned} &\hat{\mathbf{z}}(\mathbf{k}) = \mathbf{a}\cdot\mathbf{x}(\mathbf{k}) + \mathbf{b}\cdot\mathbf{y}(\mathbf{k}), \quad \breve{\mathbf{z}}(\mathbf{k}) = \mathbf{a}\cdot\mathbf{x}(\mathbf{k}) - \mathbf{b}\cdot\mathbf{y}(\mathbf{k}), \\ &\mathbf{a} = \mathbf{1}/\mathbf{S}(\mathbf{x}), \quad \mathbf{b} = \mathbf{1}/\mathbf{S}(\mathbf{y}), \quad \mathbf{S}(\mathbf{x}) = \mathbf{med}\,|\,\mathbf{x} - \mathbf{med}(\mathbf{x})\,| \end{aligned} \quad (6)$$

where $\mathbf{med(x)}$ means median value of the sample $\mathbf{x}$ and $\mathbf{S(x)}$ means absolute median deviation of the $\mathbf{x}$.

Substituting $\mathbf{x(k)}$ for $\mathbf{c}_{j_0}^{(\alpha,\tau)}(\mathbf{k})$, $\mathbf{y(k)}$ for $\mathbf{d}_{j_0}^{(\alpha,\tau)}(\mathbf{k})$ and $\mathbf{n}$ for $\mathbf{L_\alpha}$ within formulas (5) and (6) we will obtain the values of robust correlation coefficient $\nu_{j_0}(\alpha,\tau)$ describing "the strength" of connection of the selected time series $\mathbf{j_0}$ with all other time series. Let us call the value $\nu_{j_0}(\alpha,\tau)$ as *robust wavelet-based canonical correlation of the series* $\mathbf{j_0}$ (which is scale-dependent obviously). The necessity for using robust estimates is following from the strong instability of least squares calculations to the presence of outliers in initial time series or in the values of wavelet coefficients. Changing sequentially the number $\mathbf{j_0}$ of selected time series we will obtain the values of robust canonical correlations $\nu_k(\alpha,\tau)$ for all time series.

Statistical significance of the $\nu_k(\alpha,\tau)$-estimates depends on the number $\mathbf{L_\alpha}$ with the sum (3). But the number of wavelet coefficients $\mathbf{L_\alpha}$ rapidly decreases with increasing of α whereas for statistically significant estimates we must have some minimum value $\mathbf{L_{min}}$. The value of significance threshold $\mathbf{L_{min}}$ is the parameter of the method and it defines the maximum possible value of the detail level index α_{max} which could be analyzed by the method: $\alpha_{max} = \mathbf{max}\{\alpha : \mathbf{L_\alpha} \geq \mathbf{L_{min}}\}$.

After definition of the α_{max}-value the tapering preprocessing operation could be described in more details. The tapering operation is necessary for avoiding circular effects of discrete wavelet transform of the finite length sample [Press et al., 1996]. These effects are depending on the considered scales: the more is the scale, the longer are circular effects at the end of the transformed sample. Thus, the length of the end parts of tapering operations must be dependent on the maximum scale which is analyzed. That is why we have taken the length of end parts of the time window where the tapering operation is in action to be equal to $2^{(\alpha_{max}-1)}$.

The number of wavelet coefficients taking part in the estimate of the $\nu_k(\alpha,\tau)$-values rapidly decreases with increasing of the detail level number α. That is why standard deviations of statistical fluctuations of $\nu_k(\alpha,\tau)$-estimates are not uniform. In order to make statistical fluctuations of estimates dependent on the length of moving time window $\mathbf{N}$ only (and independent on the detail level number) let us introduce additional smoothing operation:

$$\bar{\nu}_k(\tau,\alpha) = \sum_{s=1}^{m_\alpha} \nu_k(\tau - s + 1, \alpha) / m_\alpha, \quad m_\alpha = 2^\alpha \tag{7}$$

The more is the number of detail level the deeper is averaging operation (7) using the values of estimates obtained by adjacent previous time windows. According to the formula (7) the efficient length of moving time window became scale-dependent and equals to $N_e^{(\alpha)} = N + 2^\alpha - 1$ samples.

A wavelet-based robust coherence measure (***WBRCM***) is defined by the formula:

$$\kappa(\tau,\alpha) = \prod_{k=1}^{q} |\bar{\nu}_k(\tau,\alpha)| \tag{8}$$

The measure (8) has possible values within boundaries from 0 up to 1. The more is the value of (8) the stronger is the collective effect of cooperation between analyzed signals on the time scales corresponding to the detail level number α. It should be noticed that the value (8) is the product of q values from the interval [0,1]. That is why the absolute values of $\kappa(\tau,\alpha)$ depend on the number q of analyzed time series and the comparison of $\kappa(\tau,\alpha)$-values could be possible for the same values of q. The most interest lays in relative peak values of the statistics (8) for different τ-values (i.e. for different positions of time window). For convenience of simultaneous comparison of $\kappa(\tau,\alpha)$-values for different detail levels let us start the index τ with initial value $N_e^{(\alpha_{max})} = N + 2^{\alpha_{max}} - 1$. The value of $\kappa(\tau,\alpha)$ is based on the information about the signals for time indexes $t: \quad \tau - N_e^{(\alpha)} \le t \le \tau$ strictly. Thus, the method has 3 free parameters: type of orthogonal wavelet, the length N of moving time window and the significance threshold L_{min}. The purpose of the analysis consists in obtaining and interpretation of extreme $\kappa(\tau,\alpha)$-values.

Dependence on the type of the used wavelet is rather essential. But this dependence could be removed if we consider the modification of the method which performs averaging coherence measure (8) within some vocabulary of used wavelets. Let $\gamma, \quad \gamma = 1,...,m_\gamma$ be the index denoting the used basis of orthogonal wavelets. We use 10 orthogonal compactly supported "usual" Daubechies wavelets with number of vanishing moments 1 (Haar's wavelet), 2, 3,…, m_γ =10. Let us introduce the wavelet type index γ into the formula for the statistics (8): $\nu_k(\tau,\alpha\,|\,\gamma), \quad k = 1,...,q$.

The multiple wavelet-based robust coherence measure (*multiple WBRCM*) is defined by the formula:

$$\tilde{\kappa}(\tau,\alpha) = \prod_{k=1}^{q} |\bar{\tilde{\nu}}_k(\tau,\alpha)| \tag{9}$$

where:

$$\begin{aligned} \bar{\tilde{\nu}}_k(\tau,\alpha) &= \sum_{s=1}^{m_\alpha} \tilde{\nu}_k(\tau - s + 1, \alpha) / m_\alpha, \quad m_\alpha = 2^\alpha \\ \tilde{\nu}_k(\tau,\alpha) &= \sum_{\gamma=1}^{m_\gamma} \nu_k(\tau,\alpha\,|\,\gamma) / m_\gamma \end{aligned} \tag{10}$$

Thus, free parameters of multiple WBRCM (9) are the length of moving time window N and significance threshold L_{min}.

CASE STUDY: FINANCIAL TIME SERIES.

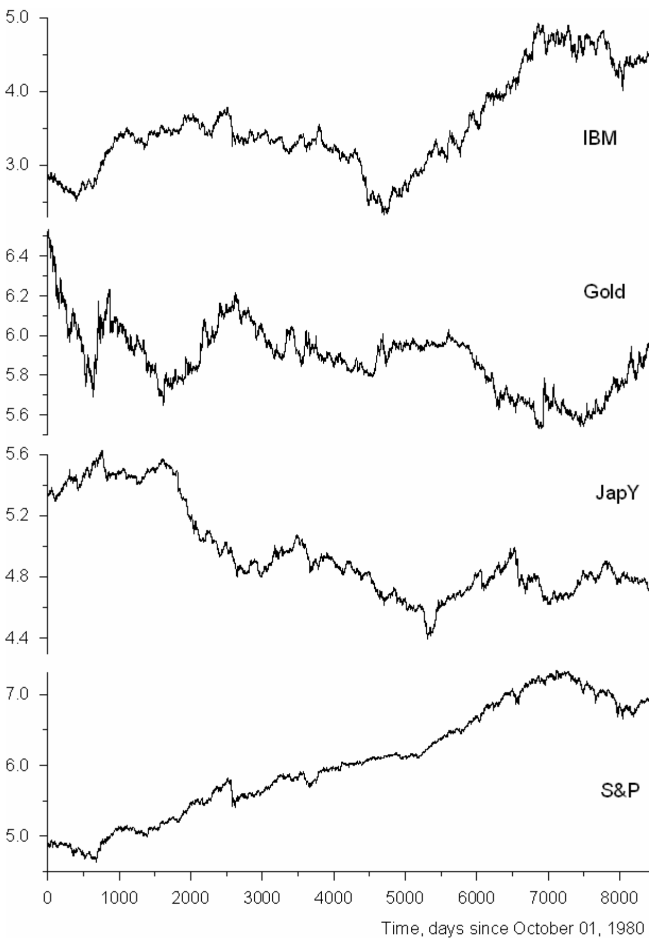

Fig.1. Graphics of initial data – 4 daily log-prices time series: IBM, Gold, JapY, S&P.

Weekends and holidays are filled by price values at the previous trading day (necessary to provide synchronous variations).

Time interval covers dates from October 01, 1980 till September 09, 2003, 8400 samples.

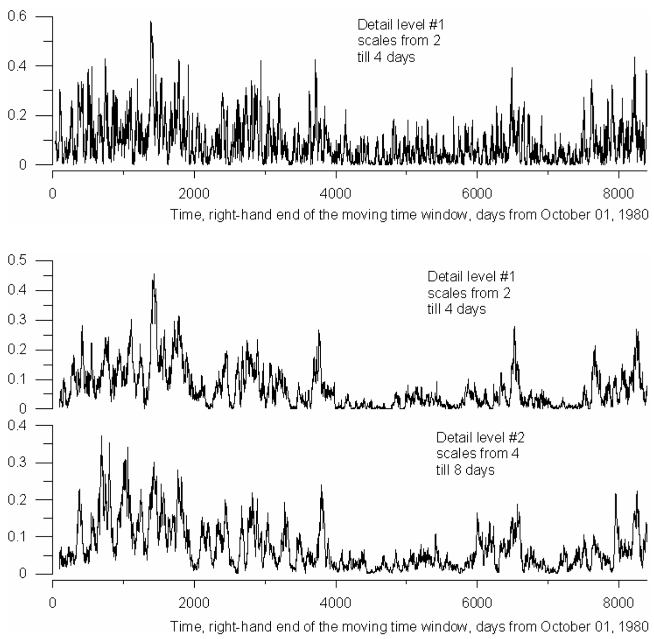

Fig.2. Multiple WBRCM Length of moving time window = 42 days (6 weeks) and 91 days (13 weeks, i.e. quarter of the year).

Coming to increments (i.e. to log-returns). Significance threshold $\mathbf{L}_{\mathbf{min}}$ =16.

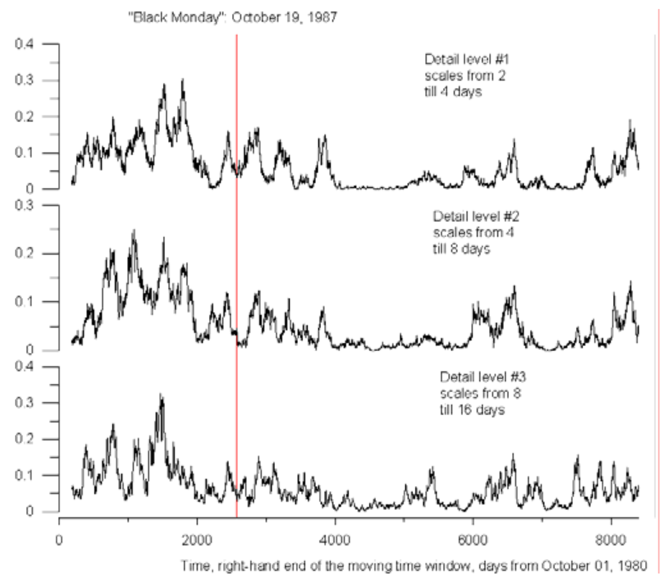

Fig.3. Multiple WBRCM Length of moving time window = 182 days (half of the year). Coming to increments (i.e. to log-returns). Significance threshold $\mathbf{L}_{\mathbf{min}}$ =16.

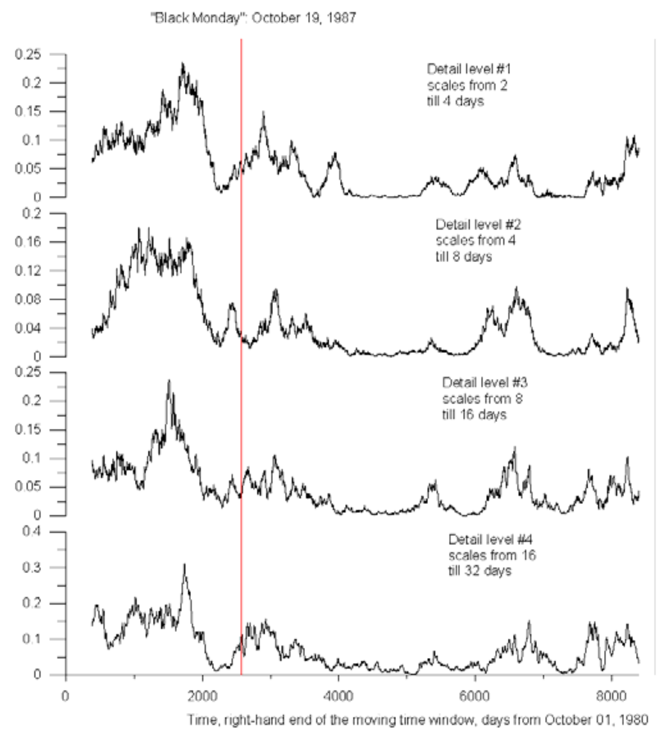

Fig.4. Multiple WBRCM Length of moving time window = 365 days (1 year). Coming to increments (i.e. to log-returns). Significance threshold $\mathbf{L}_{\mathbf{min}}$ =16.

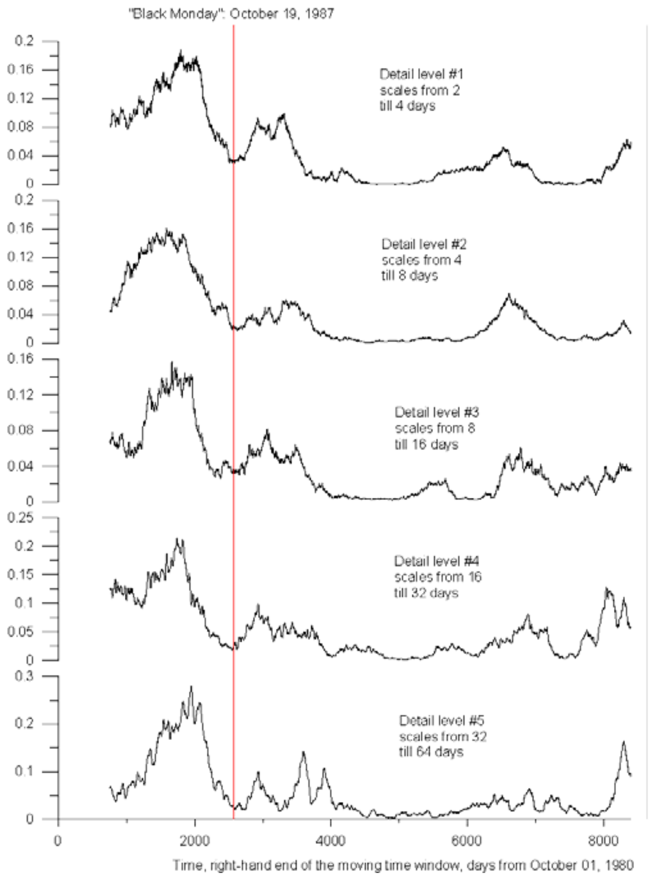

Fig.5. Multiple WBRCM Length of moving time window = 730 days (2 years). Coming to increments (i.e. to log-returns). Significance threshold $\mathbf{L}_{\mathbf{min}}$ =16.

REFERENCES

Chui, C.K. (1992) An Introduction to Wavelets, Academic Press, San Diego, CA.
Clarke E. (1975) Generalized gradients and applications //Trans. Amer. Math. Soc. 1975. Vol.205, No.2, pp. 247-262.
Daubechies I. (1992) Ten Lectures on Wavelets. No.61 in CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia.
Hotelling H. (1936). Relations between two sets of variates //Biometrika. Vol. 28. pp. 321-377.
Huber P.J. (1981) Robust statistics. John Wiley and Sons. New York, Chichester, Brisbane, Toronto
Lyubushin A.A. (1998(a)) Analysis of Canonical Coherences in the Problems of Geophysical Monitoring – Izvestiya, Physics of the Solid Earth, vol.34, pp. 52-58.
Lyubushin A.A. (1998(b)) An Aggregated Signal of Low-Frequency Geophysical Monitoring Systems. – Izvestiya, Physics of the Solid Earth, vol. 34, pp. 238-243.
Lyubushin A.A., V.A.Malugin and O.S.Kazantceva. (1999) Recognition of “Slow Events” in an Aseismic Region – Izvestiya, Physics of the Solid Earth, vol. 35, pp. 195-203.
Lyubushin A.A. (1999) Analysis of Multidimensional Geophysical Monitoring Time Series for Earthquake Prediction – Annali di Geofisica, Vol.42, No.5, pp.927-937.
Lyubushin A.A. (2000) Wavelet-Aggregated Signal and Synchronous Peaked Fluctuations in Problems of Geophysical Monitoring and Earthquake Prediction. – Izvestiya, Physics of the Solid Earth, vol.36, pp. 204-213.
Lyubushin A.A. (2001) Multidimensional Wavelet Analysis of Geophysical Monitoring Time Series – Izvestiya, Physics of the Solid Earth, vol. 37, pp. 474-483.
Lyubushin A.A. (2002) Robust wavelet-aggregated signal for geophysical monitoring problems – Izvestiya, Physics of the Solid Earth, vol.38, pp. 745-755.
Lyubushin A.A. and Kopylova G.N. (2004) Multidimensional Wavelet Analysis of Time Series of Electrotelluric Observations in Kamchatka – Izvestiya, Physics of the Solid Earth, vol.40, No.2, pp. 163-175.
Mallat S. (1998) A wavelet tour of signal processing. Academic Press. San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto. 577 p.
Press W.H., B.P.Flannery, S.A.Teukolsky, and W.T.Vettering. (1996) Numerical Recipes, 2-nd edition, Chapter 13, Wavelet Transforms, Cambridge Univ. Press, Cambridge.
Rao C.R. (1965) Linear statistical inference and its applications. John Wiley & Sons, Inc. N.Y., London, Sydney.
Shor N.Z. (1979) Methods of minimization non-differentiable functions and their applications (in Russian), Kiev, Naukova Dumka.