

**SPECTRAL ANALYSIS OF CASPIAN LEVEL VARIATIONS**

Alexey A. Lyubushin

Institute of the Physics of the  
Earth, Russian Academy of  
Sciences, 123995, Moscow,  
Russia, Bolshaya Gruzinskaya,  
10;

fax : +007-095-2556040;  
e-mail: [lubushin@mtu-net.ru](mailto:lubushin@mtu-net.ru)

Pieter H.A.J.M. van Gelder

Faculty of Civil Engineering and  
Geosciences, Delft,  
The Netherlands  
(corresponding author);

fax: +31-15-2785124;  
e-mail: [p.vangelder@ct.tudelft.nl](mailto:p.vangelder@ct.tudelft.nl)

Mikhail V. Bolgov

Institute of Water Problems,  
Russian Academy of Sciences,  
119991, Moscow, Russia,  
Gubkina, 3;

e-mail: [caspian-sea@mtu-net.ru](mailto:caspian-sea@mtu-net.ru)

**ABSTRACT**

The problem of extracting the most common spectral variations within multidimensional time series of Caspian Sea level variation on the shore stations is considered. Accurate data of 15 stations with a sampling time interval of 6 hours over a time period of observations since the beginning of 1977 till the end of 1991 are available for this purpose. The Fourier-aggregated signal method allows extract stationary common harmonics which form 2 groups of tidal sea level variation at the vicinity of 12 and 24 days periods. Besides that a long-periodic common variation was extracted with the period approximately 12.85 years. The non-stationary collective effects within period range from 4 up to 8 days were detected by estimating canonical coherences of multiple time series within moving time window. Non-stationary common effects are associated with influence of wind.

**INTRODUCTION**

Using modern time series analysis it is possible to discover characteristic time scales in datasets of environmental variables. For instance, the annual record of hurricane activity in the North Atlantic basin for the period 1886-1996 has been examined in [1] from the perspective of time series analysis. Singular spectrum analysis combined with the maximum entropy method was used on the time series of annual hurricane occurrences over the entire basin to extract the dominant modes of oscillation. Their near-decadal oscillation was a

new finding. Speculations as to the cause of the near-decadal oscillation of BE hurricanes center on changes in Atlantic SSTs possibly through changes in evaporation rates. Specifically, cross-correlation analysis pointed to solar activity as a possible explanation is given in [1]. Not only in coastal studies, but also in hydrology spectral analysis can lead to interesting observations. In [2] a yet unexplainable periodic component of 4.2 years in the annual average discharges (of different West-European rivers) on basis of daily discharge data over 100 years was detected.

This paper is devoted to application of two methods for processing multidimensional time series which were elaborated initially for purposes of earthquake prediction [3-4]. The main purpose of these methods is extracting common or collective components within variations of different scalar time series. These collective effects could be stationary, acting all the period of observation or non-stationary, occurring time from time. Thus, two methods are necessary. The 1<sup>st</sup> method (for stationary collective effects extracting) is called the Fourier-aggregated signal [3] and it processes all samples available simultaneously. The 2<sup>nd</sup> method could be called the method of by-component canonical coherences and it processes samples within moving time window of the certain length [4]. Note that the 2<sup>nd</sup> method has been applied for processing multidimensional river's runoff time series for detecting common climatic components within their variations [5].

## INITIAL DATA

Fig.1 presents all graphics of initial time series (the first 15 curves) which are Caspian Sea level variations measured at the network of 15 onshore observational stations (covering more or less uniformly all perimeter of the Caspian Sea) with sampling time interval 6 hours from the beginning of 1977 till the end of 1991. It could be noticed that time series are highly correlated for 1 year period and that all of them have a rather strong linear trend (well-known period of Caspian Sea level increasing). All further results were obtained after preliminary de-trending of these time series. Besides that it should be underlined the series of rather strong outliers within time series (especially for time series 30, 48, 59, corresponding to shallow part of the Caspian Sea) – this is the influence of strong winds. The last curve on the Fig.1 is the graphic of aggregated signal – see below.

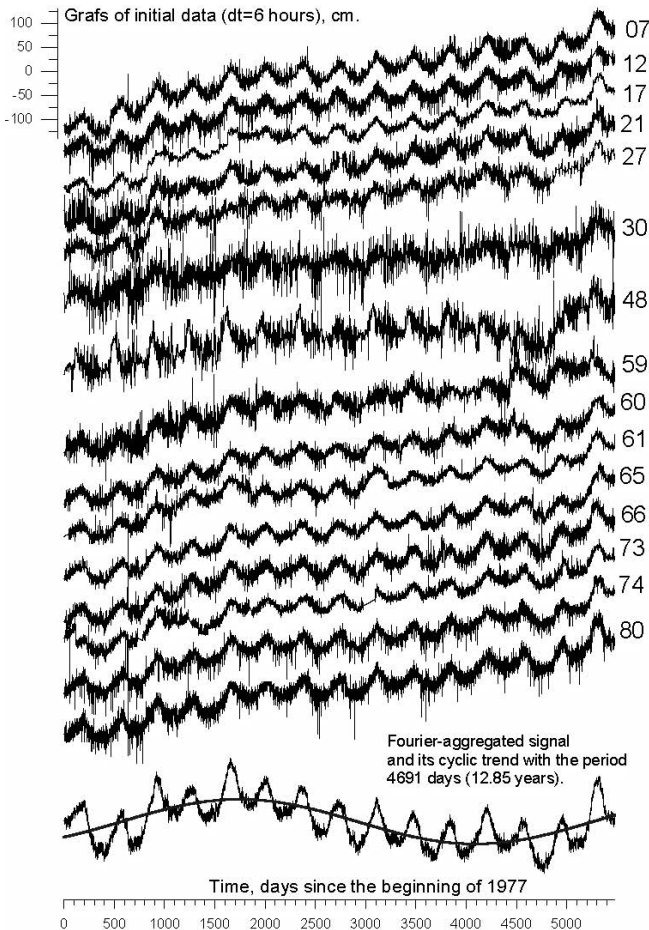


Fig.1. Graphics of all 15 initial time series (with correspondent indexes of observational stations) and their Fourier-aggregated signal with best-fit cyclic trend with the period 12.85 years (the last curve).

## METHODS

Let  $\vec{Z}(t)$  be an  $\mathbf{l}$ -dimensional time series of measurements from a monitoring system ( $t$  - discrete time index),  $S_{zz}(\omega)$  - its spectral matrix for the frequency value  $\omega$  (this complex matrix is nonnegative and hermitian, and so its eigenvalues are real and nonnegative),  $\lambda_1(\omega)$  - maximal eigenvalue of spectral matrix. According to the spectral method of principal components [6],  $\lambda_1(\omega)$  is a power spectrum of some hypothetical scalar time series  $W_1(t)$  (a first principal component time series), constructed by multi-channel filtering of the initial time series  $\vec{Z}(t)$ , by using as a frequency filter an eigenvector of spectral matrix  $S_{zz}(\omega)$ , corresponding to the maximal eigenvalue  $\lambda_1(\omega)$ . A first principal component time series carries maximum information about joint behavior of the scalar components of the vector time series  $\vec{Z}(t)$  (for Gaussian time series). Thus, if for some frequency bands the value of  $\lambda_1(\omega)$  increases considerably relative to neighbor background fluctuations, this means that for such values of frequency  $\omega$  the collective behavior is increasing also.

Let now the  $\mathbf{l}$ -dimensional vector  $\vec{Z}(t)$  be split into two vectors: an  $\mathbf{m}$ -dimensional vector  $\vec{X}(t)$  and an  $\mathbf{n}$ -dimensional vector  $\vec{Y}(t)$ , where  $\mathbf{l} = \mathbf{n} + \mathbf{m}$ . Without restriction of generality let  $\mathbf{m} \leq \mathbf{n}$ . This splitting could have the following physical meaning:  $\vec{X}(t)$  is composed of the results of measurements of some geophysical field at  $\mathbf{m}$  points and  $\vec{Y}(t)$  is composed of observations of another field at  $\mathbf{n}$  points. Now we want to know for which frequency values  $\omega$  the interaction between these two fields is maximal or minimal. Another situation, that could be reflected in such decomposition, is two points of observations: at one point there are records of variations of  $\mathbf{m}$  different geophysical parameters and at another point - of  $\mathbf{n}$  parameters. Let us now ask a question: how can we describe the interaction between the two geophysical regions, represented by these two points, in various frequency bands, using all available information?

To answer us such a question, a notion of *maximal canonical coherence* is useful [6]. Let us consider the matrix  $U(\omega)$ , which is a following product of inverted spectral and cross-spectral matrices:

$$U(\omega) = S_{xx}^{-1}(\omega) \cdot S_{xy}(\omega) \cdot S_{yy}^{-1}(\omega) \cdot S_{yx}(\omega) \quad (1)$$

It can be seen that when both series are not vectors but scalars, then  $\mathbf{U}(\omega)$  becomes the usual squared spectrum of coherence. It can be shown that the eigenvalues of  $\mathbf{U}(\omega)$  are real, nonnegative and  $\leq 1$ . These eigenvalues can be interpreted as the squared spectrum of coherence between some scalar time series, which are called *canonical components* of the initial time series  $\bar{\mathbf{X}}(\mathbf{t})$  and  $\bar{\mathbf{Y}}(\mathbf{t})$ . Let  $\mu_1^2(\omega)$  be the maximum eigenvalue of the  $\mathbf{U}(\omega)$  (i.e. the maximal canonical coherence). Then if for some values of frequency the maximum canonical coherence increases considerably and approaches the value of 1, it means that for this frequency the statistical relation between the two vector time series is strong.

Let now introduce a notion of *by-component canonical coherence*  $v_i^2(\omega)$  [3-4] as a maximum canonical coherence in a situation, when the time series  $\bar{\mathbf{Y}}(\mathbf{t})$  is composed only of the  $\mathbf{i}$ -th scalar component of the time series  $\bar{\mathbf{Z}}(\mathbf{t})$  ( $\mathbf{n}=1$ ) and  $\bar{\mathbf{X}}(\mathbf{t})$  - of all the *other* components of  $\bar{\mathbf{Z}}(\mathbf{t})$  ( $\mathbf{m}=\mathbf{1}-\mathbf{1}$ ). The value of  $v_i^2(\omega)$  describes the “strength” of the relation between variations with frequency  $\omega$  of the  $\mathbf{i}$ -th component and *the set of all other* components. Computing the product value of all by-component canonical coherences gives a spectral statistic that describes the “strength” of *joint relations* between all components of  $\bar{\mathbf{Z}}(\mathbf{t})$  at a given frequency  $\omega$  :

$$\kappa(\omega) = \prod_{i=1}^{\mathbf{l}} v_i^2(\omega) \quad (2)$$

It is evident that  $\mathbf{0} \leq \kappa(\omega) \leq \mathbf{1}$  and, hence, the closer the value of  $\kappa(\omega)$  is to 1, the stronger are the effects of collective behavior of the scalar components of  $\bar{\mathbf{Z}}(\mathbf{t})$  at a given frequency  $\omega$  .

Let  $\tau$  be a time coordinate of the moving time-window, for example, its center,  $\mathbf{L}$  be the number of samples in a time-window, and  $\delta t$  be the sampling time interval. Computing statistics (2) not over the whole interval of observation, but in a moving time-window, we will obtain a two-parameters function  $\kappa(\tau, \omega)$  . The time-window and the sampling time interval define a frequency band, which could be investigated with the help of this statistics:  $2\pi / ((\mathbf{L} - \mathbf{1})\delta t) \leq \omega \leq \pi / \delta t$

Suppose that for some time intervals and frequency bands  $(\tau, \omega)$  the value of  $\kappa(\tau, \omega)$  considerably exceeds the level of its statistical background fluctuations. Then we shall say that *a synchronous signal is observed for this*  $(\tau, \omega)$ .

Finally, let us introduce the notion of *aggregated signal* [3]. Qualitatively an aggregated signal could be defined as a scalar signal, which accumulates in its own variations *only those* spectral components, which are present *simultaneously in each* scalar time series of the multidimensional signal to be analyzed. Moreover, an algorithm of aggregation suppress those spectral components, which are presented in any one of the scalar components but absent in the others (these components could be called *local disturbance* signals). The main purpose of constructing an aggregated signal is to bring out the common trends in low-frequency geophysical network data series, which indicate an increase of collective behavior.

To formalize the notion of aggregated signal let us exclude the  $\mathbf{i}$ -th scalar component  $\mathbf{Z}_i(\mathbf{t})$  from the multidimensional time series  $\bar{\mathbf{Z}}(\mathbf{t})$  and try to filter the  $(\mathbf{1}-\mathbf{1})$ -dimensional series  $\bar{\mathbf{X}}^{(\mathbf{i})}(\mathbf{t})$ , composed of the other components, so that the filtered scalar signal  $\mathbf{C}_i^{(\mathbf{Z})}(\mathbf{t})$  has a maximum canonical coherence with  $\mathbf{Z}_i(\mathbf{t})$  for each frequency value. For this purpose we must use, as a frequency filter for  $\bar{\mathbf{X}}^{(\mathbf{i})}(\mathbf{t})$ , an eigenvector, corresponding to the maximal eigenvalue of the matrix  $\mathbf{U}(\omega)$ , where the series  $\mathbf{Z}_i(\mathbf{t})$  is taken as  $\bar{\mathbf{Y}}(\mathbf{t})$ , and  $\bar{\mathbf{X}}^{(\mathbf{i})}(\mathbf{t})$  as  $\bar{\mathbf{X}}(\mathbf{t})$ . It is clear that this eigenvalue equals  $v_i^2(\omega)$ . If  $\mathbf{Z}_i(\mathbf{t})$  contains some noise, which is present only in this component and absent from the other components of  $\bar{\mathbf{Z}}(\mathbf{t})$ , then the noise will be absent from  $\mathbf{C}_i^{(\mathbf{Z})}(\mathbf{t})$  as a consequence of its construction. At the same time  $\mathbf{C}_i^{(\mathbf{Z})}(\mathbf{t})$  retains all spectral components of  $\mathbf{Z}_i(\mathbf{t})$  which are common to the other scalar components of  $\bar{\mathbf{Z}}(\mathbf{t})$ , i.e. to the  $(\mathbf{1}-\mathbf{1})$ -dimensional signal  $\bar{\mathbf{X}}^{(\mathbf{i})}(\mathbf{t})$ . Let us call  $\mathbf{C}_i^{(\mathbf{Z})}(\mathbf{t})$  a *canonical component* of the scalar time series  $\mathbf{Z}_i(\mathbf{t})$ .

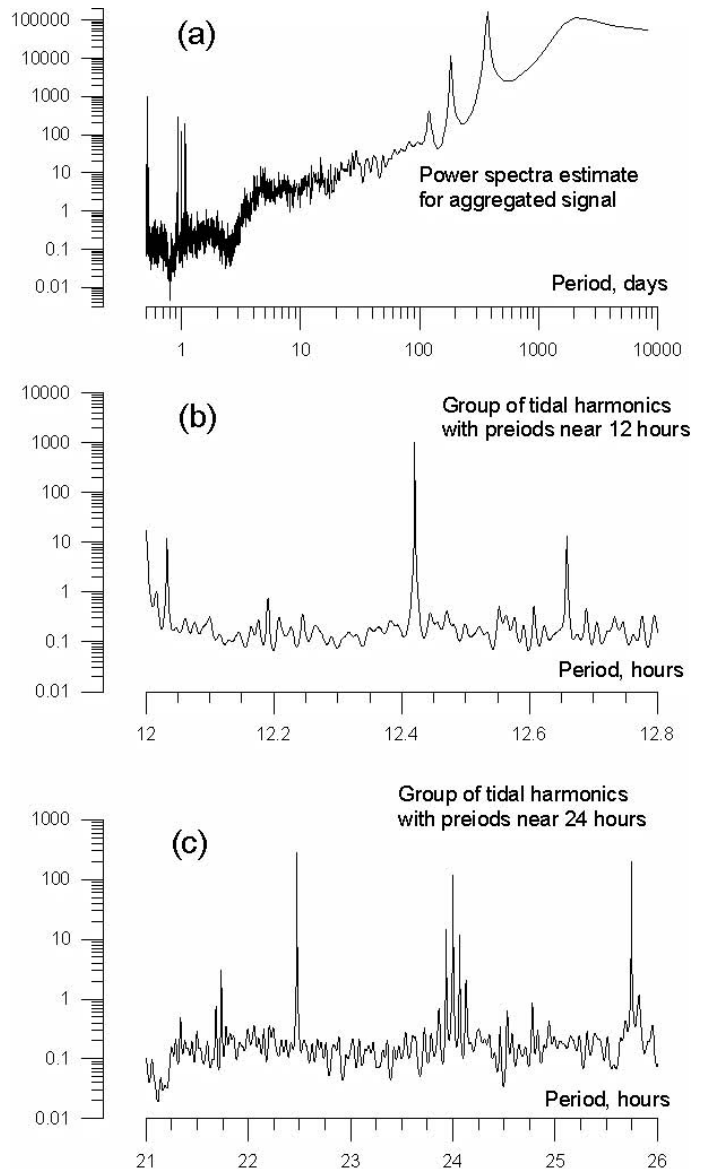
Now let us define an aggregated signal  $\mathbf{A}_z(\mathbf{t})$  of the multidimensional time series  $\bar{\mathbf{Z}}(\mathbf{t})$  as the first principal component of the multidimensional series  $\bar{\mathbf{C}}^{(\mathbf{Z})}(\mathbf{t})$ , composed of the canonical components  $\mathbf{C}_i^{(\mathbf{Z})}(\mathbf{t})$  of each of the scalar time series from the initial series  $\bar{\mathbf{Z}}(\mathbf{t})$ . The

difference between  $\mathbf{A}_z(\mathbf{t})$  and a “simple” first principal component  $\mathbf{W}_1(\mathbf{t})$  must be emphasized. In both cases the signals are constructed by multi-channel frequency filtering, using spectral matrix eigenvectors corresponding to maximal eigenvalues, as a filter. But for  $\mathbf{W}_1(\mathbf{t})$  the spectral matrix is that of the initial series  $\bar{\mathbf{Z}}(\mathbf{t})$ , whereas for  $\mathbf{A}_z(\mathbf{t})$  the spectral matrix is that of  $\bar{\mathbf{C}}^{(z)}(\mathbf{t})$ . Although in both cases detection of common spectral components takes place, the aggregated signal  $\mathbf{A}_z(\mathbf{t})$  has advantages in comparison with  $\mathbf{W}_1(\mathbf{t})$ , because the algorithm of aggregation eliminates individual noise completely, whereas they could “penetrate” into  $\mathbf{W}_1(\mathbf{t})$ , especially when the noise has the character of intense monochromatic signals.

For practical realization of these methods we must estimate spectral matrices. Below we use 2 methods for this: non-parametric Fourier transform based – for calculating aggregated signal and using multi-dimensional AR-model (5-th AR-order) – for estimating statistics (2) within moving time window [6,7].

## RESULTS

The last curve on the Fig.1 presents aggregated signal. It is evident that this signal has intensive seasonal component and that a considerable part of un-correlated noises, which are present in initial signals, was suppressed by the aggregating procedure. Besides that, this signal contains an explicit long-periodic cyclic trend. Let us find its period by fitting cyclic trends with some probe values of periods and chose those period which provides the minimum residual variance. This period turns to be equal to 12.85 years. The first idea is that this period is rather close to well-known period of solar activity and, thus, has a climatic origin.



**Fig.2.** Aggregated signal power spectra estimate within different frequency bands (vertical axes normalised).

Fig.2 presents graphics of aggregated signal power spectra estimate within 3 frequency bands: for the general frequency range (Fig.2(a)); at the vicinity of 12-hours variations (Fig.2(b)) and at the vicinity of 24-days variations (Fig.2(c)). The following period values, corresponding to the spectral peaks, could be identified:

- 1 year and its half and its third part period values (because of asymmetric form of seasonal sea level variation);
- 12.00, 12.03, 12.42 and 12.66 hours within half-day group of tidal variations;
- 21.74, 22.48, 23.94, 24.00, 24.07, 24.13 and 25.74 hours within 1-day group of variations.

Besides that there is one more interesting spectral peculiarity – a rather smoothed peak with central period near 6 days. As we will see below this peak is a cumulative result of action of non-stationary collective behavior within period range from 4 up to 8 days.

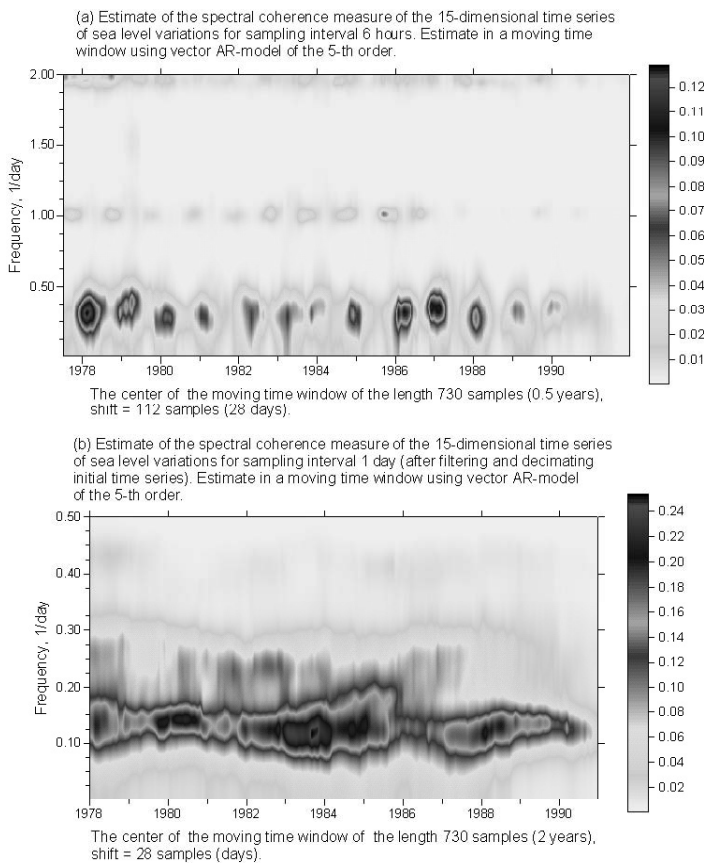


Fig.3. Maps of evolution of the by-components canonical coherences product for 2 variants of moving time windows: length = 0.5 year (a) and length = 2 years (b).

Fig.3 presents 2 maps of the statistics  $\kappa(\tau, \omega)$  for the length of time window = 0.5 years (Fig.3(a)) and for the length = 2 years (Fig.3(b)). It should be noticed that for the 2<sup>nd</sup> case we made transform from sampling time interval 6 hours to 1 day (by decimation with suppressing removed high-frequency bands). This transform was performed in order to make low-frequency common variations more explicit. The Fig.3(a) presents “islands” of collective behavior increasing for periods near 12 and 24 hours which occur with 1-year periodicity. This result is quite natural and arises due to tidal variations modulated by seasonal Volga river floods. The Fig.3(b) is more interesting. The 1-years modulated bursts of coherent behavior disappear because of rather long moving time window – 2 years (they are averaged within such window). But a new coherent signal arises – strong coherent non-stationary behavior within periods range from 4 till 8 days. We propose that these periods are some kind of Caspian Sea own periods which are excited by external force – strong winds.

## CONCLUSIONS

Results such as presented in this paper are useful and necessary for coastal studies. Design of harbor basins, offshore structures, beach nourishment programs, and so on, need accurate descriptions of the sea level variations, in particular its spectral densities.

## REFERENCES

[1] Elsner J.B., Kara A.B. and Owens M.A., 1999, “Fluctuations in North Atlantic hurricane frequency”, *Journal of Climate*, **12**, pp. 427-437.

[2] Pieter van Gelder, Vadim Kuzmin, Paul Visser, 2000 “Analysis and statistical forecasting of trends of hydrological processes in climate changes”, *Proceedings of the International Symposium on Flood Defence*, Eds. Frank Toensmann and Mafred Koch, **1**, pp. D13-D22, September 20-23, 2000, Kassel, Germany.

[3] Lyubushin A.A., 1998b, “An Aggregated Signal of Low-Frequency Geophysical Monitoring Systems”, *Izvestiya, Physics of the Solid Earth*, **34**, 1998, pp. 238-243.

[4] Lyubushin A.A., 1998a, "Analysis of Canonical Coherences in the Problems of Geophysical Monitoring", *Izvestiya, Physics of the Solid Earth*, **34**, pp. 52-58.

[5] Lyubushin A.A., Pisarenko V.F., Bolgov M.V. and Rukavishnikova T.A., 2003, "Collective effects investigation within river's runoff time series", *Russian Meteorology and Hydrology*, No.7, (pp.76-88 – for Russian issue; pages for English translation will be known at the middle of 2004).

[6] Brillinger D.R., 1975, "*Time series. Data analysis and theory*", Holt, Rinehart and Winston, Inc., N.Y., Chicago, San Francisco.

[7] Marple S.L.(Jr.), 1987, "*Digital spectral analysis with applications*", Prentice-Hall, Inc., Englewood Cliffs, New Jersey.