Annales Mathematicae et Informaticae
39 *(2012) pp. 159–171*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

The using of wavelet analysis in climatic challenges

Maxim Khaziakhmetov, Tatiana Zakharova

Moscow State University
e-mail: lsa@cs.msu.su

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

Nowadays there are many methods for processing of digital signals. A classic example is Fourier analysis. Using this transformation we build a decomposition of a signal by frequencies, so the result is easy for interpretation. But this method works well only with stationary signals, when we can find periodic components. Also using Fourier transform we lose information about coordinates of events in the original signal, because the result of transformation is in terms of frequencies.

Of course, there are alternative ways of signal processing. Wavelet analysis is one of them. Wavelet transform works in a very simple manner. It divides the original signal into two parts – approximation and details. This dichotomization can be repeated many times and we'll have decomposition with multilevel detailization. There are 2 ways for further work: to analyze the result of transformation interpreting it as something or to execute some operations with the result and then use inverse Wavelet transform.

This article is about the using of wavelet analysis in climatic challenges. The work of the authors of this article was to analyze water vapor field of the Earth searching for numerical patterns.

1. Basic information about Wavelet transform

Information about wavelets required for further discussion is placed in this section of the article.

1.1. 1D wavelet transform

1.1.1. 1D discrete wavelet transform

Definition 1.1. The function $\varphi(x) \in L^2(\mathbb{R})$ is scaling function if it can be represented as:

$$\varphi(x) = \sqrt{2} \sum_{n\in\mathbb{Z}} h_n \varphi(2x - n), \tag{1.1}$$

where $h_n, n \in \mathbb{Z}$ satisfy the condition

$$\sum_{n\in\mathbb{Z}} |h_n|^2 < \infty.$$

The equation (1.1) is scaling equation, the set of coefficients $\{h_n\}_{n\in\mathbb{Z}}$ is scaling filter.

Widely known Haar function:

$$F(x) = \begin{cases} 1, x \in [0, 1), \\ 0, \ \text{otherwise} \end{cases}$$

is scaling function, but centered Haar function

$$F(x) = \begin{cases} 1, x \in [-\frac{1}{2}, \frac{1}{2}), \\ 0, \ \text{otherwise} \end{cases}$$

cannot be classified to scaling functions.

Definition 1.2. Orthogonal multiresolution decomposition (or multiresolution analysis, or MRA) of $L^2(\mathbb{R})$-space is a sequence of confined spaces:

$$\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots,$$

with following properties:

1. $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$,
2. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,
3. $f(x) \in V_j \iff f(2x) \in V_{j+1}$,
4. $\exists \varphi(x) \in V_0 : \ \{\varphi_{0,n}(x) = \varphi(x - n)\}_{n\in\mathbb{Z}}$ – an orthonormal basis of V_0-space.

Using properties 3 and 4 of Definition 1.2 we can conclude that

$$\{\varphi_{j,n}(x) = \sqrt{2^j}\varphi(2^j x - n)\}_{n\in\mathbb{Z}}$$

is an orthonormal basis of V_j-space.

$\forall j$ we have $V_j \subset V_{j+1}$. Let W_j be an orthogonal complement of V_j to V_{j+1}, i.e. $V_{j+1} = V_j \oplus W_j$. Then $V_{j+1} = V_{j-1} \oplus W_{j-1} \oplus W_j$, because $V_j = V_{j-1} \oplus W_{j-1}$. Repeating this procedure we'll have

$$V_{j+1} = \overline{\bigoplus_{k=-\infty}^{j} W_k}.$$

According to property 1 of Definition 1.2 we can conclude that $L^2(\mathbb{R})$-space has an orthogonal decomposition:

$$L^2(\mathbb{R}) = \overline{\bigoplus_{k=-\infty}^{+\infty} W_k}.$$

Definition 1.3. If a sequence of subspaces $\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots$ is multiresolution analysis and $\forall j \in \mathbb{Z}$ W_j is the orthogonal complement of V_j to V_{j+1} then each such W_j is wavelet space and its elements are wavelets.

$\exists \psi(x) \in W_0$, called "mother wavelet", with properties

1. $\{\psi(x-n)\}_{n\in\mathbb{Z}}$ – an orthonormal basis of W_0-space,
2. $\{\psi_{j,n}(x) = \sqrt{2^j}\psi(2^j x - n)\}_{n\in\mathbb{Z}}$ – an orthonormal basis of W_j-space for every j.

Let $(f(x), g(x))$ be the scalar product in $L^2(\mathbb{R})$, i.e.

$$\langle f(x), g(x)\rangle = \int_{-\infty}^{+\infty} f(x)g(x)dx. \tag{1.2}$$

Let $f(x) \in V_{j+1}$ then we have the decomposition of $f(x)$:

$$f(x) = \sum_{k=-\infty}^{j} \sum_{n\in\mathbb{Z}} \langle f(x), \psi_{k,n}(x)\rangle \psi_{k,n}(x).$$

In practice the number of detailization levels is finite, so for $f(x)$ we have the following decomposition formula:

$$f(x) = \sum_{k=0}^{j} \sum_{n\in\mathbb{Z}} \langle f(x), \psi_{k,n}(x)\rangle \psi_{k,n}(x) + \sum_{n\in\mathbb{Z}} \langle f(x), \varphi(x-n)\rangle \varphi(x-n),$$

i.e. the signal from V_{j+1}-space is represented in terms of elements of spaces $\{V_0, W_0, \ldots, W_j\}$.

In case of discrete signal the formula (1.2) can be rewritten as the sum. If $x = \{x_n\}_{n\in\mathbb{Z}}$ is the digitization of the signal $x(t), t \in \mathbb{R}$, then wavelet coefficient

$a_{j,n}$ can be represented in terms of discrete convolution with the filter $h = \{h_n\}_{n\in\mathbb{Z}}$ corresponded to function $\psi_{j,n}$:

$$a_{j,n} = \sum_{k=-\infty}^{\infty} h_k x_{n-k}.$$

Usually the process of signal analysis starts from its representation in terms of a basis of V_{j+1}-space. Then we build a decomposition in bases of V_j č W_j constructing approximation and details. We can repeat decomposition or stop the process. So we have the following decomposition scheme:

$$\begin{array}{ccccccc}
V_{j+1} & \longrightarrow & V_j & \longrightarrow \; ... \; \longrightarrow & V_1 & \longrightarrow & V_0 \\
\downarrow & & \downarrow & & \downarrow & & \\
W_j & & W_{j-1} & & W_0 & &
\end{array}$$

We stopped at 0-index, so V_0-based approximation component consists of the most general trends of the original signal and W_0-based detailization includes the most spatially extended deviations from these trends.

As we noted in the abstract we can use decomposition coefficients as independent data and create a conclusion based on it or perform operations on them and then reconstruct the signal.

For the reconstruction process we have the following scheme:

$$\begin{array}{ccccccc}
V_{j+1} & \longleftarrow & V_j & \longleftarrow \; ... \; \longleftarrow & V_1 & \longleftarrow & V_0 \\
\uparrow & & \uparrow & & \uparrow & & \\
W_j & & W_{j-1} & & W_0 & &
\end{array}$$

1.1.2. 1D continuous wavelet transform

As we know, for 1D discrete wavelet transform the following formula is valid:

$$\psi_{j,n}(x) = \sqrt{2^j}\psi(2^j x - n), \text{ where } j,\ n \in \mathbb{Z}.$$

Continuous wavelet transform is constructed by allowing arbitrary real values of the parameters of scaling and shift (in discrete variant we should use powers of 2 for the scale parameter and integers for shift parameter). This generalization allows to select details of a signal with arbitrary size of their support.

Let $\psi(x)$ be wavelet. $\psi(x) \in L^2(\mathbb{R})$ and also

$$C_\psi = 2\pi \int_{-\infty}^{+\infty} |\omega|^{-1} |\hat{\psi}(\omega)|^2 d\omega < \infty. \tag{1.3}$$

(The relation (1.3) guarantees the existence of inverse continuous transform).

Let a be the scaling parameter and b is the shift parameter, then we have 2-parameter family of wavelets:

$$\psi_{ab}(x) = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right), \text{ where } a,\, b \in \mathbb{R}.$$

1D continuous wavelet transform is defined by the formula:

$$W_\psi\,[f]\,(a,b) = (f, \psi_{a,b}) = |a|^{-1/2}\int_{-\infty}^{+\infty} f(x)\overline{\psi\left(\frac{x-b}{a}\right)}dx. \tag{1.4}$$

It's obvious that the coefficients of 1D discrete wavelet transform can be computed as

$$c_{jk} = W_\psi[f]\left(\frac{1}{2^j}, \frac{k}{2^j}\right).$$

And as in discrete case we have inverse transform:

Theorem 1.4. *If $f(x) \in L^2(\mathbb{R})$ and (1.3) is satisfied, then we have inverse 1D continuous wavelet transform formula:*

$$f(x) = C_\psi^{-1}\iint W_\psi[f](a,b)\psi_{ab}(x)\frac{dadb}{a^2}.$$

1.2. 2D discrete wavelet transform

In this subsection we'll consider the case of functions from $L^2(\mathbb{R}^2)$-space.

The simplest way of generalization 1D wavelet transform to 2D wavelet transform is based on tensor product. We have the following representation for $L^2(\mathbb{R}^2)$-space:

$$L^2(\mathbb{R}^2) = L^2(\mathbb{R}) \otimes L^2(\mathbb{R}).$$

I.e. linear combinations of $f(x)g(y)$ construct dense set in $L^2(\mathbb{R}^2)$.

We'll define V_0^2 as tensor product of V_0:

$$V_0^2 = V_0 \otimes V_0.$$

According to this fact

$$\{\varphi_{0,k,n}(x,y) = \varphi(x-k)\varphi(y-n)\}_{k,n\in\mathbb{Z}}$$

is an orthonormal basis of V_0^2-space.

$V_j^2 = V_j \otimes V_j$ are scaled versions of V_0^2-space, for them we have the following relation

$$f(x,y) \in V_0^2 \iff f(2^jx, 2^jy) \in V_j^2.$$

So, as in 1D-case, there is a sequence of spaces $\ldots \subset V_{-1}^2 \subset V_0^2 \subset V_1^2 \subset \ldots$. Using equation $V_1 = V_0 \oplus W_0$ for $L^2(\mathbb{R}^2)$ we conclude that

$$\begin{aligned} V_1^2 = V_1 \otimes V_1 = (V_0 \oplus W_0) \otimes (V_0 \oplus W_0) = \\ = (V_0 \otimes V_0) \oplus (V_0 \otimes W_0) \oplus (W_0 \otimes V_0) \oplus (W_0 \otimes W_0) = \\ = V_0^2 \oplus (V_0 \otimes W_0) \oplus (W_0 \otimes V_0) \oplus (W_0 \otimes W_0). \end{aligned}$$

$V_0 \otimes W_0$, $W_0 \otimes V_0$, $W_0 \otimes W_0$ forms 2D-wavelet space W_0^2. The following facts exist:

space $V_0 \otimes W_0$ is constructed by shifts of function $\psi^H(x,y) = \varphi(x)\psi(y)$, we'll designate it as W_0^H – this is the space of horizontal details (ox-oriented homogeneous areas can be selected);

space $W_0 \otimes V_0$ is constructed by shifts of function $\psi^V(x,y) = \psi(x)\varphi(y)$, we'll designate it as W_0^V – this is the space of vertical details (oy-oriented homogeneous areas can be selected);

space $W_0 \otimes W_0$ is constructed by shifts of function $\psi^D(x,y) = \psi(x)\psi(y)$, we'll designate it as W_0^D – this is the space of diagonal details(diagonal inhomogeneous areas can be selected).

So we have the following decomposition:

$$V_{j+1}^2 = V_j^2 \oplus W_j^H \oplus W_j^V \oplus W_j^D \quad \forall j.$$

The following sets of functions are orthonormal bases of listed spaces:

$$\begin{aligned} &\{\varphi_{j,k,n}(x,y) = 2^j \varphi(2^j x - k)\varphi(2^j y - n)\}_{k,n\in\mathbb{Z}}; \\ &\{\psi_{j,k,n}^H(x,y) = 2^j \varphi(2^j x - k)\psi(2^j y - n)\}_{k,n\in\mathbb{Z}}; \\ &\{\psi_{j,k,n}^V(x,y) = 2^j \psi(2^j x - k)\varphi(2^j y - n)\}_{k,n\in\mathbb{Z}}; \\ &\{\psi_{j,k,n}^D(x,y) = 2^j \psi(2^j x - k)\psi(2^j y - n)\}_{k,n\in\mathbb{Z}}. \end{aligned}$$

I.e. there are 4 types of decomposition coefficients:

$$\langle f(x,y), \varphi_{j,k,n}(x,y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x,y)\varphi(2^j x - k)\varphi(2^j y - n)dxdy;$$

$$\langle f(x,y), \psi_{j,k,n}^H(x,y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x,y)\varphi(2^j x - k)\psi(2^j y - n)dxdy;$$

$$\langle f(x,y), \psi_{j,k,n}^V(x,y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x,y)\psi(2^j x - k)\varphi(2^j y - n)dxdy;$$

$$\langle f(x,y), \psi_{j,k,n}^D(x,y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x,y)\psi(2^j x - k)\psi(2^j y - n)dxdy;$$

Schemes of decomposition and reconstruction processes are

$$
\begin{array}{ccccccc}
V^2_{j+1} & \longrightarrow & V^2_j & \longrightarrow \dots \longrightarrow & V^2_1 & \longrightarrow & V^2_0 \\
\downarrow & & \downarrow & & \downarrow & & \\
W^H_j, W^V_j, W^D_j & & W^H_{j-1}, W^V_{j-1}, W^D_{j-1} & & W^H_0, W^V_0, W^D_0 & &
\end{array}
$$

$$
\begin{array}{ccccccc}
V^2_{j+1} & \longleftarrow & V^2_j & \longleftarrow \dots \longleftarrow & V^2_1 & \longleftarrow & V^2_0 \\
\uparrow & & \uparrow & & \uparrow & & \\
W^H_j, W^V_j, W^D_j & & W^H_{j-1}, W^V_{j-1}, W^D_{j-1} & & W^H_0, W^V_0, W^D_0 & &
\end{array}
$$

2. The task of analysis of water vapor field of the Earth

2.1. Introduction

The Earth atmosphere is very complex and unpredictable. Energy of the atmosphere is contained mostly in water vapor, because of its heat capacity. Study of processes in water vapor field can help us to explain and predict atmospheric phenomena, for example, cyclones and hurricanes' formation – they are the most interesting because of their consequences.

Many atmospheric phenomena have periodic nature, for example, it is easy to understand that water vapor field has an annual cycle of movements. But the most interesting phenomena for scientists are nonstationary. They should be localized in space and time and their parameters should be found so appropriate methods of research are required.

In our research of water vapor field we used wavelet analysis.

2.2. Essence of the task

For every day from 01.01.1999 to 31.12.2009 we had a digital map of the earth with the size 360×720 (grid is 0,5°). A value of every pixel on map is an average density (kg/m^2) of water vapor in spheric segment of the Earth with appropriate geographic coordinates and date (rounded to the nearest whole). So we had 3D-array of data from satellite images.

There are some problems with this array. The algorithm of value construction can be used only for water vapor field above the surface of the oceans, so by this reason the land surface pixels are filled with zeros. Also there are spaces on maps where in some days satellites didn't make images, these pixels are marked with "-20"-value.

It's obvious that the most convenient way to display this information is graphic – for every day we can draw a map of water vapor field as indexed image. The picture (1) demonstrates an example image (the matching of colors to values of density is on the right side): we can see all the problems listed above.

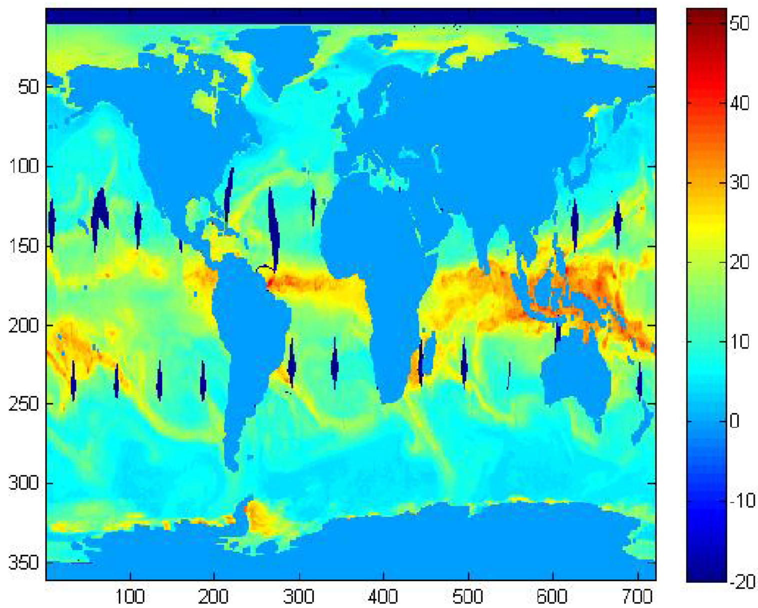

Figure 1: Water vapor field. 01.01.1999

So the essence of the task was to search for numeric patterns in this density array.

The preliminary analysis of data had given a very important fact. The variance of values for the first 7 years differs from the variance of values for the last 4 years (the algorithm of value construction had been changed by data provider), so we decided to use data from 01.01.1999 to 31.12.2005.

2.3. Research ways

We decided to divide research into 2 parts:

1. The research of time series in every discretization point;
2. Meridional analysis.

The main idea was to examine the results of wavelet decomposition on numeric patterns. All the algorithms were realized in Matlab using Wavelet Toolbox.

2.3.1. The research of time series

We had 3D-array where every 2D-layer is a density map. Fixing values of geographic coordinates we can extract time series for every point of discretization of the Earth's surface, so it had been done for data appropriate to surface of the oceans. An example of such time series is on the Figure 2.

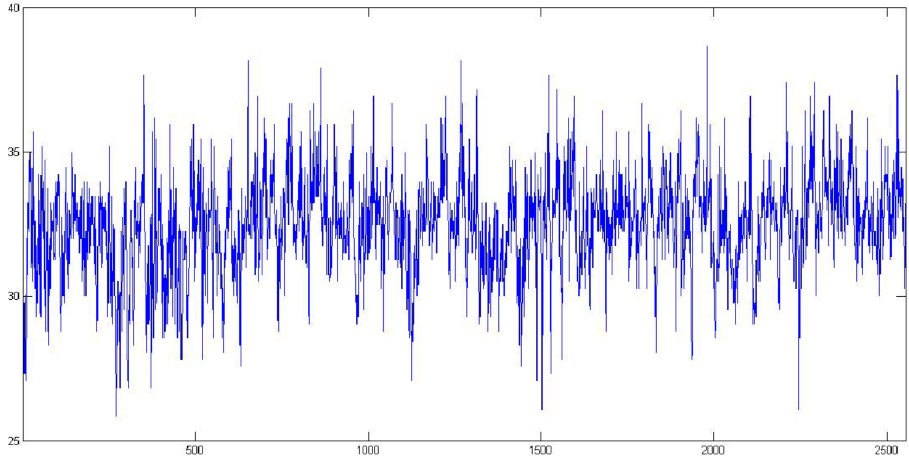

Figure 2: An example of time series (ox - time, oy - density)

In Matlab's Wavelet Toolbox we can analyze 1D arrays using both discrete and continuous wavelet transforms. But the second variant gives much more information about signal, because we can observe details of time series with arbitrary length instead of only the powers of 2. We had tested many wavelet families and for some of them found interesting patterns.

Results obtained with the use of wavelet Morlet (in terms of Matlab - "cmor1-1.5") we consider the most important. Wavelet Morlet is a complex function which can be written as

$$\psi(t) = e^{2\pi i t - \frac{t^2}{2}}.$$

Decomposition of every time series should be executed according to formula (1.4). On Figure 3 the result of 1D continuous complex wavelet transform for time series of point from Oceania is presented. This figure is built using Matlab's Wavelet Toolbox, so we can see specific Matlab notation for wavelet coefficients. For example, Ca, b means wavelet coefficient of continuous wavelet decomposition with scale parameter a and shift parameter b:

$$W_\psi[f](a, b) = Ca, b.$$

So we had such decomposition arrays for all available values of geographic coordinates. We had computed main frequencies for every pair (time series, values of scale parameter) using Fourier transform and for every value of scale parameter of decomposition generated a "frequency map". An example of such map is on the Figure 4. The area of 45° sl. — 45° nl. is presented because cyclonic activity is strong there.

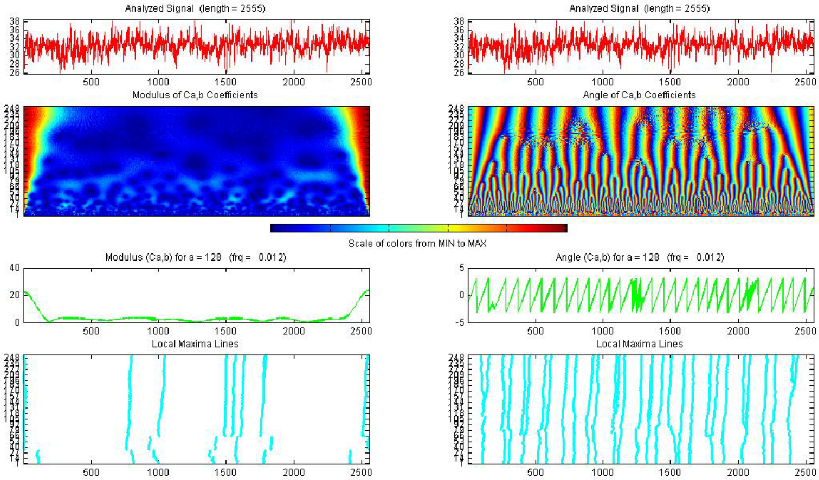

Figure 3: An example of continuous complex wavelet decomposition of time series which is presented on Figure 2

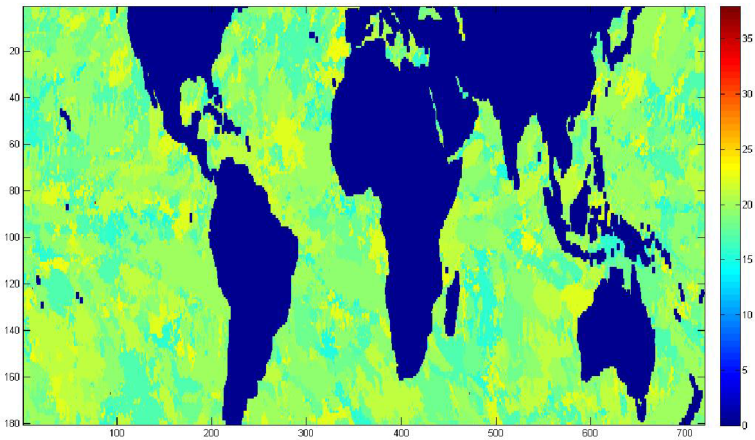

Figure 4: An example of “frequency map” (30-days activity)

Figure 4 demonstrates distribution of 30-days activity because the wavelet filter works with interval of coefficients of this length.

The staff of Institute of Space Researches had worked with this algorithm searching for physical interpretation of results represented on “frequency maps”. They had found some new numeric patterns in the Earth’s water vapor field such as subzones of variability of density of water vapor, known subzones had been localized better. Also all season effects and high day to day activity had been confirmed.

2.3.2. Meridional analysis

This part of researches is based on idea of joint analysis of data for every meridian. From original 3D array we extracted 2D array fixing the value of geographic longitude. So we had indexed images such as presented on Figure 5.

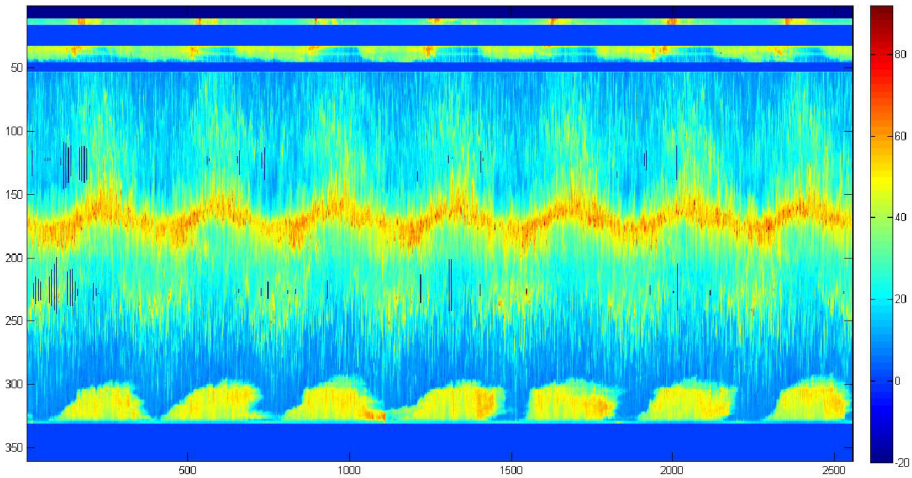

Figure 5: An example of distribution of density of water vapor during 1999-2005 for 21.5° el

According to strong variance of data from day to day we cannot use continuous wavelets. We will not reconstruct signal after decomposition, so we don't need any specific properties of wavelet basis in this context. Also it's comfortable when wavelet function is symmetric because transform results are unbiased. So we were using biorthogonal wavelets.

The physical interpretation of wavelet coefficients is rather simple. 2D discrete wavelet transform on every step of decomposition doubles the size of details it constructs. In our case it doubles them both by time and space. According to nature of selected basis we can conclude that the first level of decomposition consists of (day to day, 0,5° of latitude) details (for example the 5th level has details represented fluctuations during a month (approximately)). But we must take into account that there are 3 sorts of details: horizontal, vertical and diagonal. Horizontal details represent constant time nature of signal and difference by latitude. For vertical details we have an opposite situation. Diagonal details include space-and-time differences in signal.

There are many interesting facts we had found on decomposition images. In this article we are presented only some of them.

The first conclusion is based on the results of analysis of the first level of decompositions.We can say that day to day variability is stronger then interlatitudinal. Day to day activity is stronger in equatorial zone and mid-latitudes then in subtropics and circumpolar regions. Also it's easy to see that activity movings has seasonal component.

Another interesting fact we want to present in this article had been found on the 5th level of decomposition ("monthly details"). There is a strong difference between equatorial and subtropical monthly climatic activities. Similar result we see in circumpolar regions but we cannot be so sure in wavelet coefficients computed there because of edge effects.

2.4. Summary of current results

Distancing from facts listed above we can say that some relatively common results have already been achieved:

New mathematical methods for data processing has been proposed;

Using this methods the new concept "frequency map" has been introduced in the subject area;

Some numeric patterns for water vapor field have been discovered;

The hypothesis that "frequency maps" can help to predict atmospheric phenomena has been proposed.

2.5. Further plans

We have got many ideas about further work on this task and using of wavelet analysis in other research areas. We are working on the hypothesis of the relationship between water vapor field fluctuations and cyclonic activity. Many facts say that "frequency maps" can help us to make necessary mathematical tools for prediction of some atmospheric phenomena such as hurricanes. Some simple visual patterns were detected during comparison of sets of images such as Figures 6 and 7.

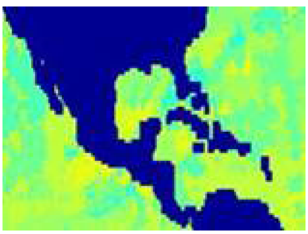

Figure 6: Element of frequency map

Figure 7: Trajectory of Catrina hurricane

So further researches are needed.

References

[1] Bogges A., Narkovich F. A. A First Course in Wavelets with Fourier Analysis. – Prentice Hall, 2001.

[2] Zakharova T. V., Shestakov O. V. Wavelet analysis. Study guide. – Max-Press, 2009 (in Russian).

[3] Mallat S. A wavelet tour of signal processing. – Academic Press, 1999.

[4] Smolentsev N. K. Basics of wavelet theory. Wavelets in Matlab. – DMK Press, 2005 (in Russian).