Physics – Uspekhi **39** (11) 1085 – 1108 (1996)

REVIEWS OF TOPICAL PROBLEMS PACS numbers: **02.30.–f**, **02.90. + p**, **92.60.–e**, 92.60.Ry

Wavelet analysis: basic theory and some applications

N M Astaf'eva

Contents

Abstract. The basic theory of the wavelet transform, an effective investigation tool for inhomogeneous processes involving widely different scales of interacting perturbations, is presented. In contrast to the Fourier transform, with the analysing function extending over the entire axis of time, the two-parametric analysing function of the one-dimensional wavelet transform is well localised in both time and frequency. The potential of the method is illustrated by analysing familiar model series (such as harmonic, fractal, and those with various types of singularities) and the long-term variation of some meteorological characteristics (Southern Oscillation index and global and hemispheric temperatures). The analysis of a number of El Niño events and of the temporal behaviour of the Southern Oscillation index reveals periodic components, local periodicity features and time scales on which self-similarity structures are seen. On the whole, both stochastic and regular components seem to be present. The global and hemispheric temperatures are qualitatively similar in structure, the main difference — presumably due to the greater amount of land and stronger anthropogenic factor — being that the warming trend in the Northern Hemisphere is slightly stronger and goes first in time.

1. Introduction

The notion of a wavelet (a small wave if taken literally) has evolved relatively recently — it was introduced by Grossmann and Morlet in the middle of the 80s as applied to the analysis of properties of seismic and acoustic signals [1]. At present, the family of analysing functions dubbed wavelets is being increasingly used in problems of pattern recognition; in processing and synthesising various signals, speech for instance; in analysis of images of any kind (these may be iris images, X-ray picture of a kidney, satellite images of clouds or a planet surface, an image of mineral, etc.); for study of turbulent fields, for contraction (compression) of large volumes of information, and in many other cases.

The wavelet transform of a one-dimensional signal involves its decomposition over a basis obtained from a soliton-like function (wavelet), possessing some specific properties, by dilations and translations. Each of the functions of this basis emphasises both a specific spatial (temporal) frequency and its localisation in physical space (time).

Thus, unlike the Fourier transform traditionally used in signal analysis the wavelet transform offers a two-dimensional expansion of a given one-dimensional signal, with the frequency and the coordinate treated as independent variables. As a result, the possibility emerges of analysing the signal simultaneously in physical (time, coordinate) and frequency spaces. What was said can readily be generalised to multidimensional signals or functions.

In foreign literature, it is already a conventional practice to term the Fourier spectrum a single spectrum, in distinction to the spectrum obtained from the coefficients of wavelet transform, which is referred to as a time-scale spectrum, or wavelet spectrum.

The area where the wavelets find use is not reduced to the analysis alone of properties of signals and fields of arbitrary nature obtained either numerically, experimentally or observationally. Wavelets are coming into use in direct numerical simulations — as some hierarchical basis well suited to describe the dynamics of complex nonlinear processes characterised by interaction of perturbations in wide ranges of spatial and temporal scales.

N M Astaf'eva Space Research Institute, Russian Academy of Sciences, ul. Profsoyuznaya 84/32, 117810 Moscow, Russia
Tel. (7-095) 333 21 45. E-mail: ast@iki.rssi.ru

Received 23 May 1996, revised 18 July 1996
Uspekhi Fizicheskikh Nauk **166** (11) 1145 – 1170 (1996)
Translated by S D Danilov

Results of numerous numerical experiments suggest that for large Reynolds numbers the major part of volume occupied by a turbulent fluid remains inactive with respect to energy dissipation and, correspondingly, the inverse energy cascade. This phenomenon, called intermittency, can be conveniently explored by wavelet analysis. The latter permits one to recover spatially distributed features of the object studied, determine if the intermittency is present and how the regions of dissipation are distributed, obtain both local high-frequency and global low-frequency information related to the object, and a great deal of other information, with reasonable accuracy and without redundancies.

One faces known difficulties when processing short high-frequency signals, or signals with localised frequencies. The wavelet transform proves to be an extremely efficient tool in adequately decoding such signals since elements of its basis are well-localised and possess a moving time-frequency window.

It is not a coincidence that many researchers refer to wavelet analysis as a 'mathematical microscope', as this term accurately conveys the remarkable capability of the method to offer a good resolution at different scales. The capability of this microscope to reveal the internal structure of an essentially inhomogeneous process (or field) and expose its local scaling behaviour has been demonstrated through many classical examples such as the fractal Weierstrass functions and probability measures of the Cantor series, to mention but a few. Application of wavelet analysis to a turbulent velocity field in a wind tunnel under large Reynolds numbers for the first time offered a vivid confirmation of the Richardson cascade. The analogy between the energy cascade and the structure of the multifractal inhomogeneous Cantor series was explicitly settled. Even more efficient was the application of wavelet analysis to the multifractal invariant measures of several well-known dynamical systems that model transitions to chaos observed in dissipative systems.

Thus, wavelets can be successfully applied to solve various problems. They are, however, not widely known to researchers dealing with analysis of experimental and observational data. In this work, an attempt is made to outline, clearly and simply when possible, the basics of the theory needed in practical applications of the wavelet transform to processing signals of various nature.

In Section 2 the analogy between the Fourier series and wavelet series expansions is shown and basic definitions of the wavelet transform are introduced. Section 3 describes the features and properties of functions that could form the basis of a wavelet transform and gives some examples of the most widely-used wavelets. In Section 4, the properties of wavelet transforms are listed, a number of important physical characteristics are introduced and some potentialities of wavelet analysis are illustrated. This material is largely based on monographs and collections of works [2–4] and on excellent works by Ingrid Daubechies [5] and Marie Farge [6]. Examples of wavelet transform applications to modelling signals of various nature are considered in Section 5; Section 6 is dedicated to results of wavelet analysis of observational meteorological data series.

2. From the Fourier transform to the wavelet transform

The integral Fourier transform and Fourier series underlie harmonic analysis. Fourier coefficients resulting from the transform are amenable to a quite simple physical interpretation, with the simplicity by no means reducing the importance of the ensuing inferences on the character of the signals studied. Making use of the integral Fourier transform and Fourier series (in calculations, analytical transformations) is straightforward; all necessary properties and formulae can be written with the help of only two real-valued functions $\sin t$ and $\cos t$ (or with a single complex-valued, the sinusoidal wave $\exp(\mathrm{i}t) = \cos t + \mathrm{i}\sin t, \mathrm{i} = \sqrt{-1}$), and proved relatively easily.

The wavelet transform is not so readily and widely recognised because it came into use not such a long time ago and its mathematical principles are still at the stage of active development. For that reason, in order to achieve a clear exposition we will introduce below [4] the necessary concepts of wavelet method by analogy with the Fourier analysis the significance and potentialities of which for a wide research community raise no doubts and are verified by long and successful practice.

The definitions, properties and corollaries will be presented for one-dimensional functions or data series. When necessary, all that can be readily generalised to multidimensional cases. Specifically, we will speak of time-dependent functions, series, and accordingly, of frequencies. Without any loss of generality, an independent coordinate can be a spatial one (with corresponding wave numbers), or any other.

2.1 Fourier series

Let us recall a few concepts which will be needed below. Let $L^2(0, 2\pi)$ be the space of square-integrable functions with a finite energy (norm)

$$\int_0^{2\pi} |f(t)|^2 \,\mathrm{d}t < \infty, \qquad t \in (0, 2\pi). \tag{1}$$

This is the definition of a piecewise-continuous function $f(t)$. It can be periodically continued and defined on the entire axis $R(-\infty, \infty)$:

$$f(t) = f(t - 2\pi), \qquad t \in R.$$

Any function $f(t)$ belonging to the space of 2π-periodic square-integrable functions can be expanded into a Fourier series

$$f(t) = \sum_{-\infty}^{\infty} c_n \exp(\mathrm{i}nt). \tag{2}$$

The coefficients c_n in Eqn (2) have the form

$$c_n = (2\pi)^{-1} \int_0^{2\pi} f(t) \exp(-\mathrm{i}nt) \,\mathrm{d}t, \tag{3}$$

and the series (2) converges uniformly to $f(t)$:

$$\lim_{M, N \to \infty} \int_0^{2\pi} \left| f(t) - \sum_{-M}^{N} c_n \exp(\mathrm{i}nt) \right|^2 \mathrm{d}t = 0.$$

Notice that

$$w_n(t) = \exp(\mathrm{i}nt), \qquad n = \ldots, -1, 0, 1, \ldots \tag{4}$$

is the orthonormalised basis in $L^2(0, 2\pi)$ constructed with the help of dilations of a single function $w(t) = \exp(\mathrm{i}t)$ so that $w_n(t) = w(nt)$.

To summarise, each 2π-periodic square-integrable function can be obtained as a superposition of scale transformations of the basis function $w(t) = \exp(\mathrm{i}t) = \cos t + \mathrm{i} \sin t$, i.e. it is a composition of sinusoidal waves with different frequencies (and with coefficients dependent on the wave number).

Recall that by virtue of Parseval's theorem coefficients of Fourier series are constrained by the identity

$$(2\pi)^{-1} \int_0^{2\pi} |f(t)|^2 \,\mathrm{d}t = \sum_{-\infty}^{\infty} |c_n|^2 . \tag{5}$$

2.2 Wavelet expansion

Consider the space $L^2(R)$ of functions $f(t)$ defined on the entire real axis $R(-\infty, \infty)$ and possessing a finite energy (norm)

$$E_f = \int_{-\infty}^{\infty} |f(t)|^2 \,\mathrm{d}t < \infty . \tag{6}$$

The functional spaces $L^2(0, 2\pi)$ and $L^2(R)$ are essentially different. In particular, the local mean of every function from $L^2(R)$ should tend to zero at $\pm\infty$. However a sinusoidal wave does not belong to $L^2(R)$ and consequently, the family of sinusoidal waves w_n cannot be the basis of functional space $L^2(R)$. Let us try to find reasonably simple functions to construct the basis of space $L^2(R)$.

'Waves' forming the space $L^2(R)$ must tend to zero at $\pm\infty$, and, for practical reasons, the quicker they decay, the better. Therefore, for the basis functions, let us consider wavelets — well-localised soliton-like 'small waves'.

As in the case of space $L^2(0, 2\pi)$ which can be completely formed with the help of a single basis function $w(t)$, we construct the functional space $L^2(R)$ relying on a single wavelet $\psi(t)$. Note that the latter can be either a single-frequency wavelet, or a wavelet with a frequency band. We begin with discrete transformations.

How can the entire axis be covered with the help of a single localised function which tends rapidly to zero? The most simple choice is to adopt a system of shifts (translations) along the axis. For the sake of simplicity let them be integers, i.e. $\psi(t - k)$.

We introduce an analogue of a sinusoidal frequency. Again, for the sake of simplicity and definiteness we take it as a power of two: $\psi(2^j t - k)$. Here j and k are integers $(j, k \in I)$.

In this manner, departing from discrete scale transformations $(1/2^j)$ and translations $(k/2^j)$, we may describe all frequencies and cover the entire axis with the help of a single basis wavelet $\psi(t)$.

Let us recall the norm definition:

$$\|p\|_2 = \langle p, p\rangle^{1/2} ,$$

$$\langle p, q\rangle = \int_{-\infty}^{\infty} p(t) q^*(t) \,\mathrm{d}t .$$

The asterisk denotes complex conjugation. Consequently,

$$\|\psi(2^j t - k)\|_2 = 2^{-j/2} \|\psi(t)\|_2 ,$$

i.e., if a wavelet $\psi(t) \in L^2(R)$ is of unit norm, then all wavelets of the family $\{\psi_{jk}\}$, of the form

$$\psi_{jk}(t) = 2^{j/2} \psi(2^j t - k) , \quad j, k \in I, \tag{7}$$

are also normalised to unity, $\|\psi_{jk}\|_2 = \|\psi\|_2 = 1$.

A wavelet $\psi \in L^2(R)$ is termed orthogonal if the family $\{\psi_{jk}\}$ defined by Eqn (7) is an orthonormalised basis in the functional space $L^2(R)$, i.e.

$$\langle \psi_{jk}, \psi_{lm}\rangle = \delta_{jl} \delta_{km} ,$$

and each function $f \in L^2(R)$ can be represented as the series

$$f(t) = \sum_{j,k=-\infty}^{\infty} c_{jk} \psi_{jk}(t) , \tag{8}$$

whose uniform convergence in $L^2(R)$ implies that

$$\lim_{M_1, N_1, M_2, N_2 \to \infty} \left\| f - \sum_{-M_2}^{N_2} \sum_{-M_1}^{N_1} c_{jk} \psi_{jk} \right\|_2 = 0 .$$

The simplest example of an orthogonal wavelet offers the HAAR wavelet named after Haar, who suggested it. It is defined by the relationship

$$\psi^{\mathrm{H}}(t) = \begin{cases} 1 , & 0 \leqslant t < 1/2 , \\ -1 , & 1/2 \leqslant t < 1 , \\ 0 , & t < 0 , \quad t \geqslant 1 . \end{cases} \tag{9}$$

One may readily see that any two functions ψ^{H}_{jk} and ψ^{H}_{lm} obtained from this wavelet by formula (7) with the help of dilations $1/2^j$ and $1/2^l$ and translation $k/2^j$ and $m/2^l$ are orthogonal and have a unit norm.

Let us construct a basis of functional space $L^2(R)$ with the help of continuous scale transformations and translations of a wavelet $\psi(t)$ with arbitrary values of basis parameters — the scaling coefficient a and translation parameter b:

$$\psi_{ab}(t) = |a|^{-1/2} \psi\left(\frac{t - b}{a}\right) , \quad a, b \in R , \quad \psi \in L^2(R) . \tag{10}$$

Based on this, we write the integral wavelet transform:

$$\begin{aligned}[W_\psi f](a, b) &= |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \psi^*\left(\frac{t - b}{a}\right) \mathrm{d}t \\ &= \int_{-\infty}^{\infty} f(t) \psi^*_{ab}(t) \,\mathrm{d}t .\end{aligned} \tag{11}$$

Adhering to further analogy with the Fourier transform, the coefficients $c_{jk} = \langle f, \psi_{jk}\rangle$ of the expansion (8) of function f in wavelet series can be defined through the integral wavelet transform:

$$c_{jk} = [W_\psi f]\left(\frac{1}{2^j}, \frac{k}{2^j}\right) . \tag{12}$$

Below, instead of $[W_\psi f](a, b)$, the designations $W(a, b)$, $W_\psi f$ or $W[f]$ will sometimes be used for the coefficients of wavelet transform (also called the wavelet amplitude).

Thus every function from $L^2(R)$ can be represented by a superposition of scale transforms and translations of basis wavelet, i.e. it is a composition of 'wavelet waves' [with coefficients being functions of wave number (frequency, scale) and translation parameter (time)].

Using a discrete wavelet transform (discrete frequency-time space formed by integer translations and dilations in powers of two) enables one to prove many aspects of wavelet theory [2–5] related to the completeness and orthogonality of basis, convergence of series, etc. Such proofs are needed, say in information compressing or in numerical modelling problems, i.e. in cases where it is necessary to accomplish the expansion constrained to a minimum number of independent coefficients of the wavelet transform and to have an exact formula for the inverse transform. As applied to signal analysis a continuous wavelet transform (11) is more convenient; although it possesses some ambiguity related to a continuous variation of the scaling coefficient a and the translation parameter b, this is rather an advantage here allowing for more complete and clearer presentation and analysis of the data.

2.3 Inverse wavelet transform

Sinusoidal wave forms an orthonormalised basis in the functional space $L^2(0, 2\pi)$, so the Fourier transform is reversible and there exists the inverse Fourier transform. The orthonormality of a wavelet-derived basis in the space $L^2(R)$ depends both on the choice of the basis wavelet and on the way the basis is constructed (values of basis parameters a, b).

Undoubtedly, a wavelet can be considered as a basis function in $L^2(R)$ only in the case when the basis derived from it is orthonormal and the inverse transform exists. It should, however, be noted that rigorous proofs of completeness and orthogonality are complicated and very involved; corresponding examples can be found in Refs [2–5] where the theory of wavelets is elaborated. In addition, for practical purposes it is frequently enough to have stability and approximate orthogonality of the system of functions used for the decomposition, i.e. it suffices for the system to be a quasi-basis. As a rule, such quasi-basis wavelets are in fact employed in signal analysis.

For a detailed theory and proofs the reader is referred to the above-mentioned references. Here we restrict ourselves to writing inverse transforms in two particular cases described above: for the basis (7) admitting dilations and translations $(1/2^j, k/2^j)$, $j, k \in I$, and the basis (10) constructed with arbitrary parameters (a, b), $a, b \in R$.

With basis parameters (a, b), $a, b \in R$ the inverse wavelet transform is expressed in the same basis (10) as the direct one:

$$f(t) = C_\psi^{-1} \iint \left[W_\psi f\right](a, b)\, \psi_{ab}(t)\, \frac{\mathrm{d}a\, \mathrm{d}b}{a^2}\,, \qquad (13)$$

C_ψ is the normalising coefficient (analogous to $(2\pi)^{1/2}$ that normalises the Fourier transform):

$$C_\psi = \int_{-\infty}^{\infty} \left|\hat{\psi}(\omega)\right|^2 |\omega|^{-1}\, \mathrm{d}\omega < \infty$$

(the hat designates the Fourier transform).

The requirement for the coefficient C_ψ to be finite bounds the class of functions $\psi(t) \in L^2(R)$ that may be used as basis wavelets. In particular, it is apparent that the Fourier transform $\hat{\psi}$ must be equal to zero at the coordinate origin $\omega = 0$, and, consequently, at least one moment should be zero:

$$\int_{-\infty}^{\infty} \psi(t)\, \mathrm{d}t = 0\,.$$

Most frequently, only positive frequencies are needed in applications, i. e. $a > 0$; accordingly, the wavelet should satisfy the condition

$$C_\psi = 2\int_0^{\infty} \left|\hat{\psi}(\omega)\right|^2 \omega^{-1}\, \mathrm{d}\omega = 2\int_0^{\infty} \left|\hat{\psi}(-\omega)\right|^2 \omega^{-1}\, \mathrm{d}\omega < \infty\,.$$

A stable basis in the case of discrete wavelet transform is defined in the following way.

A function $\psi \in L^2(R)$ is termed an R-function if the basis $\{\psi_{jk}\}$ defined by (7) is the Riesz basis in the sense that there exist two constants A and B, $0 < A \leqslant B < \infty$, for which the relation

$$A\left\|\{c_{jk}\}\right\|_2^2 \leqslant \left\| \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{jk}\, \psi_{jk} \right\|_2^2 \leqslant B\left\|\{c_{jk}\}\right\|_2^2$$

holds with any (bounded, double square-summed) set $\{c_{jk}\}$:

$$\left\|\{c_{jk}\}\right\|_2^2 = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |c_{jk}|^2 < \infty\,.$$

For any R-function there exists a basis $\{\psi^{jk}\}$ dual to the basis $\{\psi_{jk}\}$ (in the sense that $\langle \psi_{jk},\, \psi^{lm} \rangle = \delta_{jl}\delta_{km}$) with whose help the reconstruction formula emerges

$$f(t) = \sum_{j,\,k=-\infty}^{\infty} \langle f, \psi_{jk} \rangle\, \psi^{jk}(t)\,. \qquad (14)$$

If ψ is an orthogonal wavelet and $\{\psi_{jk}\}$ is an orthonormal basis, $\{\psi^{jk}\}$ and $\{\psi_{jk}\}$ coincide. Then (14) is the inverse transform formula. If ψ is not an orthogonal wavelet, but a dyadic R-wavelet, then it possesses a double $\overset{*}{\psi}$ with whose help the dual basis to the family $\{\psi_{jk}\}$ is constructed in much the same way as the basis (7):

$$\psi^{jk}(t) = \overset{*}{\psi}_{jk}(t) = 2^{j/2}\overset{*}{\psi}(2^j t - k)\,, \qquad j, k \in I\,. \qquad (15)$$

In a general case, the reconstruction formula (14) may not even be a wavelet series in the sense that $\overset{*}{\psi}$ is not a wavelet, and $\{\psi^{jk}\}$ may possess no dual basis constructed in the manner of (10).

2.4 Time-frequency localisation

The Fourier transform and Fourier series are an excellent mathematical tool for the physical interpretation of processes by analysing signals that characterise them. However there are cases where their efficiency is insufficient.

A real signal always (or as a rule) belongs to the space $L^2(R)$. The Fourier transform of a signal $f(t)$ with a finite energy defined by the norm $\|f\|_2$ represents the spectrum of this signal:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-\mathrm{i}\omega t)\, \mathrm{d}t\,.$$

Under some circumstances, the physical interpretation based on this formula appears to be difficult. For example, to obtain spectral information at a given frequency one needs to have both the preceding and future information; moreover, the formula does not take into account the fact that frequency may evolve with time. The Fourier transform, for example, is not capable of distinguishing between the signal composed of two sine functions with different frequencies and a signal composed by the same functions switched on in an alternating way one after another (an example is considered in Section 5).

Besides, it is known that signal frequency is inversely proportional to signal duration. Therefore, to extract high-frequency information at a fair resolution one should retrieve it from relatively short time intervals, but not the entire signal; conversely, low-frequency information should be extracted from relatively long temporal intervals.

Some of the difficulties described can be obviated by making use of windowed Fourier transform. Still, an infinitely stretched basis function (sine wave) is not fitted to discriminate actually localised information. In contrast, the element of basis of the wavelet transform is a well-localised function rapidly going to zero out of some finite interval thus allowing for a 'localised spectral analysis'. The sense of this somewhat strange word combination will soon become clear. In other words, the wavelet transform automatically offers a variable time-frequency window, narrow for small scales and wide for large ones.

Which are the parameters of the time-frequency window associated with the wavelet transform? Since either the wavelet ψ itself or its Fourier transform $\hat{\psi}$ decay quite rapidly, both can be used as window functions with 'centre' and 'width' defined as follows.

For a nontrivial window function $z(t) \in L^2(R)$ [the function $tz(t)$ should also belong to $L^2(R)$] its centre $\langle t \rangle$ and radius Δ_z are expressed as

$$\langle t \rangle = \frac{1}{\|z\|_2^2} \int_{-\infty}^{\infty} t |z(t)|^2 \, \mathrm{d}t \,,$$

$$\Delta_z = \frac{1}{\|z\|_2} \left[\int_{-\infty}^{\infty} \big(t - \langle t \rangle\big)^2 |z(t)|^2 \, \mathrm{d}t \right]^{1/2} ,$$

with the window width being $2\Delta_z$.

Let $\langle t \rangle$, Δ_ψ, $\langle \omega \rangle$, and $\Delta_{\hat{\psi}}$ be the centres and radii of wavelet ψ and its Fourier transform $\hat{\psi}$, respectively, defined by these formulae. Then the integral wavelet transform (11) implicitly uses the 'time window'

$$\left[\mathrm{win}_t\right] = \left[b + a\langle t \rangle - 2a\Delta_\psi, \, b + a\langle t \rangle + 2a\Delta_\psi\right] , \tag{16}$$

i.e. there is a time localisation with the window centre at $b + a\langle t \rangle$ and window width $4a\Delta_\psi$.

Introduce the function $\eta(\omega) = \hat{\psi}\big(\omega + \langle \omega \rangle\big)$ which is also a window function with its centre at zero and radius $\Delta_{\hat{\psi}}$. Recognising that $\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle / 2\pi$, one may write the integral wavelet transform (11) for the Fourier transform $\hat{f}$ in the form

$$W(a,b) = |a|^{1/2} \int_{-\infty}^{\infty} \hat{f}(\omega) \exp(\mathrm{i}b\omega) \eta^* \left(a \left[\omega - \frac{\langle \omega \rangle}{a} \right] \right) \mathrm{d}\omega \,. \tag{17}$$

If we neglect the phase shift for the current moment, it becomes obvious that the transform (17) gives a localised information on the spectrum $\hat{f}(\omega)$ of the signal $f(t)$ with the 'frequency window'

$$\left[\mathrm{win}_\omega\right] = \left[\frac{\langle \omega \rangle}{a} - \frac{1}{a} \Delta_{\hat{\psi}}, \, \frac{\langle \omega \rangle}{a} + \frac{1}{a} \Delta_{\hat{\psi}} \right] . \tag{18}$$

Frequency localisation occurs in the window with centre at $\langle \omega \rangle / a$ and width $2\Delta_{\hat{\psi}}/a$.

Notice that the ratio of the central frequency to the window width,

$$\frac{\langle \omega \rangle}{a} \left(\frac{2\Delta_{\hat{\psi}}}{a} \right)^{-1} = \frac{\langle \omega \rangle}{2\Delta_{\hat{\psi}}}$$

is independent of the exact position of central frequency, and the time-frequency window $[\mathrm{win}_t] \cdot [\mathrm{win}_\omega]$ of area $4\Delta_\psi \Delta_{\hat{\psi}}$ becomes narrower with higher central frequency $\langle \omega \rangle / a$ and expands at low frequency (see Fig. 1a).

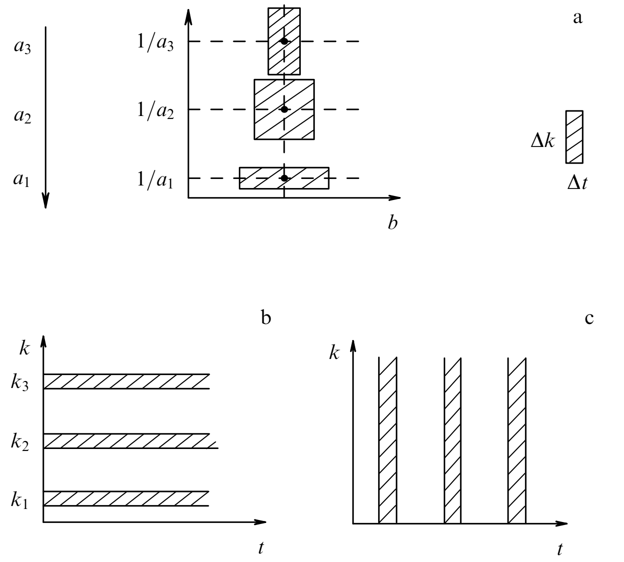

Figure 1. Time-frequency localisation of transforms with different analysing functions: (a) wavelets, (b) Fourier harmonics, (c) Shannon functions.

By way of comparison, Fig. 1 shows the localisation in time-frequency space of transforms with other analysing functions: the Fourier transform (Fig. 1b) and the Shannon transform in which the analysing function is Dirac's function (Fig. 1c).

It is readily seen that the Fourier transform is well suited to localise the frequency (but at the expense of time resolution), while the Shannon transform is not capable of frequency localisation. Contrastingly, the wavelet transform possesses a variable window localised near a time instant chosen and expanding with growing scale, a feature most desirable for retrieving spectral information. All this hinges on the fact that basis functions of the transforms listed above are, respectively, a sinusoidal wave localising only the frequency, Dirac's function localising only the time instant, and soliton-like wavelet localising fairly well both the temporal scale and the time instant.

In order to demonstrate the advantages of the wavelet transform as a method of localised spectral analysis, we compare the wavelet transform (11) with the windowed, or short-time, Fourier transform frequently used in signal analysis,

$$F(\omega, b) = \int f(t) z(t - b) \exp(\mathrm{i}\omega t) \, \mathrm{d}t$$

— the transform of the signal preliminarily multiplied by the window function z. Therefore $F(\omega, b)$ is the signal expansion in the family of functions $z(t-b) \times \exp(\mathrm{i}\omega t)$ generated from a single function $z(t)$ with the help of translations b in time and translations ω in frequency. At the same time, the result of the wavelet transform $W(a,b)$ is the signal expansion in the family $\psi\big((t-b)/a\big)$ derived from a single function $\psi(t)$ with translation b in time and dilation a also in time. The wavelet transform thus looks like a continuous set of windowed Fourier transforms with different windows for different frequencies.

Hence the basis functions of a windowed Fourier transform have the same resolution in time and frequency $\big(z(t), \hat{z}(\omega)\big)$ within the entire transform plane, whereas the basis functions of a wavelet transform possess time resolution

$(\psi(t/a))$ decreasing with scale a and frequency resolution $(\hat{\psi}(a\omega))$ increasing with that scale. This property of the wavelet transform is indeed extremely advantageous for signal analysis: rapid variations of signals are well localised (high-frequency characteristics), while to reveal slowly varying characteristics a reasonably low-frequency resolution will do. The wavelet transform owing to its variable time-frequency window is equally well suited to recovering either the high- or low-frequency characteristics of signals.

Recall that the Fourier transform of a data sequence sampled through equal time intervals Δt cannot provide the frequency discretisation better than $\Delta\omega = \Delta t/2$ (the Nyquist frequency; this is a particular demonstration of the indeterminacy principle with respect to time and frequency localisations). Analogous constraint on the wavelet transform may be expressed through the relation $\Delta t\Delta\omega \geqslant (4\pi)^{-1}$.

One may easily infer that the advantages of wavelets may be extremely helpful not only in analysing complex signals but also in solving equations that describe the processes of interaction between perturbations of different scales.

3. Basis functions of wavelet transform

So far we have used the term 'wavelet' to designate some soliton-like function; concepts associated with it were introduced and pertinent properties were described. In existing literature we have not encountered any successful conventional definition of a wavelet. To be specific, we present here the most simple, in our opinion, definition [3] which relies on the notions introduced above.

3.1 Wavelet definition

Any localised R-function $\psi \in L^2(R)$ is called an R-wavelet (or simply wavelet) if there exists a function $\overset{*}{\psi} \in L^2(R)$ (its conjugate, double) such that families $\{\psi_{jk}\}$ and $\{\overset{*}{\psi}_{jk}\}$ constructed in compliance with Eqns (7) and (15) are paired bases in the functional space $L^2(R)$.

Each wavelet ψ defined in this way, whether it is orthogonal or not, allows for an arbitrary function $f \in L^2(R)$ to be represented as series (8) with coefficients being determined by the integral wavelet transform of f relative to $\overset{*}{\psi}$.

Wavelet double $\overset{*}{\psi}$ is unique and is itself also an R-wavelet. The pair $(\psi, \overset{*}{\psi})$ is symmetric in the sense that ψ, in turn, is the double of $\overset{*}{\psi}$.

For an R-wavelet ψ exhibiting property of orthogonality $\overset{*}{\psi} \equiv \psi$, and $\{\psi_{jk}\} \equiv \{\overset{*}{\psi}_{jk}\}$ is the orthogonal basis.

For many practical applications it suffices for wavelet ψ to be semiorthogonal, i. e. to have the Riesz basis $\{\psi_{jk}\}$ satisfying the condition $\langle\psi_{jk}, \psi_{lm}\rangle = 0$ at $j = l, j, k, l, m \in I$.

An R-wavelet is termed non-orthogonal if it is not a semiorthogonal wavelet. However, as an R-wavelet, it has a double and the pair $(\psi, \overset{*}{\psi})$ offers the possibility of forming the families $\{\psi_{jk}\}$ and $\{\overset{*}{\psi}_{jk}\}$ that satisfy the biorthogonality condition $\langle\psi_{jk}, \overset{*}{\psi}_{lm}\rangle = \delta_{jl}\delta_{km}$, $j, k, l, m \in I$ and allow for construction of valid wavelet series and the reconstruction formula.

The need for an inverse wavelet transform (or reconstruction formula) entails most of the constraints a wavelet is subject to.

3.2 Wavelet criteria

For practical applications it is important to know the criteria which a function should satisfy in order to be a wavelet; we present them here and also consider, by way of examples, certain well-known functions, inquiring whether they observe these criteria.

Localisation. Unlike the Fourier transform, the wavelet transform is based on a localised basis function. A wavelet must be localised both in time and frequency.

Zero mean:

$$\int_{-\infty}^{\infty} \psi(t)\,\mathrm{d}t = 0\,. \tag{19}$$

In applications it is frequently necessary to have additionally first m zero moments:

$$\int_{-\infty}^{\infty} t^m\psi(t)\,\mathrm{d}t = 0\,. \tag{20}$$

Such a wavelet is referred to as an mth-order wavelet. Wavelets possessing many zero moments enable one to ignore most regular polynomial components of signals and analyse their small-scale fluctuations and high-order features.

Boundedness:

$$\int \left|\psi(t)\right|^2 \mathrm{d}t < \infty\,. \tag{21}$$

The estimate of good localisation and boundedness could be written as

$$\left|\psi(t)\right| < \left(1 + |t|^n\right)^{-1} \quad \text{or} \quad \left|\hat{\psi}(\omega)\right| < \left(1 + |k - \omega_0|^n\right)^{-1},$$

where ω_0 is the dominant frequency of a wavelet; the integer n should be as large as possible.

Basis self-similarity. A characteristic feature of a wavelet transform basis is its self-similarity. All wavelets of a given family $\psi_{ab}(t)$ have the same number of oscillations as the basis wavelet $\psi(t)$ does — they are derived from it by scale transforms and translations. It is for this reason that the wavelet transform is successfully used in fractal analysis (see, for example, Ref. [7]).

This can be illustrated by examples from Ref. [6] which presents a number of functions and their Fourier transforms. For comparison, among them there are both wavelets and functions which cannot be wavelets for one reason or another.

For instance, the δ-function and sine function do not satisfy the necessary condition of being localised in both time and frequency domains: the δ-function well-localised in the t-space does not possess this property in the k-domain; conversely, a sine function well-localised in the k-space does not show this property in the t-space.

The Gabor function

$$G(t) = \exp\left[\mathrm{i}\Omega(t - t_0) - \mathrm{i}\vartheta\right] \exp\left[-\frac{(t - t_0)^2}{2\sigma^2}\right] \frac{1}{\sigma(2\pi)^{1/2}}$$

is defined as a modulated Gaussian function with four parameters: the shift t_0, standard (root mean square) deviation σ, modulation frequency Ω and phase shift ϑ. Expansion in the Gabor functions is that in modulated fragments of sine functions. The duration of fragments is the same at different frequencies which yields different number of oscillations for different harmonics. Hence it follows that the Gabor function, being localised in both t and k spaces, cannot be chosen as a basis function for the wavelet transform since the basis created with it will not exhibit self-similarity.

The HAAR wavelet [see Ref. (9)] exemplifies an orthogonal discrete wavelet giving birth to an orthonormalised basis. Its drawback is non-smoothness — the presence of sharp boundaries in the t-space manifested as infinite (decaying as k^{-1}) 'tails' in the k-space and form asymmetry. There are applications for which these drawbacks are immaterial, and sometimes the one-sidedness of wavelet serves as an advantage. One may often encounter another very similar, also discrete but symmetric FHAT wavelet, better known as the 'French hat' (it looks like a top hat):

$$\psi(t) = \begin{cases} 1\,, & |t| \leqslant 1/3\,, \\ -1/2\,, & 1/3 < |t| \leqslant 1\,, \\ 0\,, & |t| > 1\,, \end{cases}$$

$$\hat{\psi}(k) = 3\Theta(k)\left(\frac{\sin k}{k} - \frac{\sin 3k}{3k}\right),$$

here $\Theta(k)$ is Heaviside's function [$\Theta(k) = 1$ at $k > 0$ and $\Theta(k) = 0$ at $k \leqslant 0$].

The FHAT wavelet, behaving irregularly in the time domain and decaying too slowly in the frequency domain, as well as the LP wavelet (named after Littlewood and Paley, see Ref. [6]), conversely, with sharp boundaries in the k-space and improperly decaying in the t-space, could be attributed to limiting cases; practically all wavelets fill in between them.

3.3 Examples of analysing wavelets

Since the wavelet transform is a scalar product of analysing wavelet with a given scale and a signal explored, coefficients $W(a, b)$ contain combined information on both the analysing wavelet and the signal (similar to coefficients of the Fourier transform which bear imprints of both the signal and sinusoidal wave).

The choice of analysing wavelet is as a rule dictated by the character of information to be derived from the signal. Every wavelet has specific features in time and frequency domains and sometimes by applying different wavelets one may reveal more fully or emphasise one or other signal characteristic.

If the analogy mentioned with the 'mathematical microscope' is pursued further, then the translation parameter b fixes the focusing point of the microscope, the scale factor a, the magnification, and finally, the choice of the basis wavelet ψ determines the optical quality of the microscope.

Real-valued bases are frequently constructed from derivatives of Gaussian functions

$$\psi_m(t) = (-1)^m\,\partial_t^m\left[\exp\left(-\frac{t^2}{2}\right)\right],$$

$$\hat{\psi}_m(k) = m(\mathrm{i}k)^m\exp\left(-\frac{k^2}{2}\right)$$

(here $\partial_t^m = \partial^m[\ldots]/\partial t^m$, $m \geqslant 1$). Higher derivatives have more zero moments and permit one to retrieve information on features of higher orders contained in signals.

Figures 2a, b show wavelets obtained for $m = 1$ and $m = 2$, respectively. Due to their shape they came to be known, respectively, as the WAVE wavelet and the MHAT wavelet, or the 'Mexican hat' (looks like a sombrero).

The MHAT wavelet, with its narrow energy spectrum and two zero moments (zeroth and first) is well suited to analyse complex signals. Generalised to the two-dimensional case, the MHAT wavelet is frequently taken to analyse isotropic fields. If the derivative is taken in one direction, an anisotropic basis can be obtained, with a good angular resolution [6]. To create it, rotations should be added to scale transforms and translations. In this case mathematical microscope gains

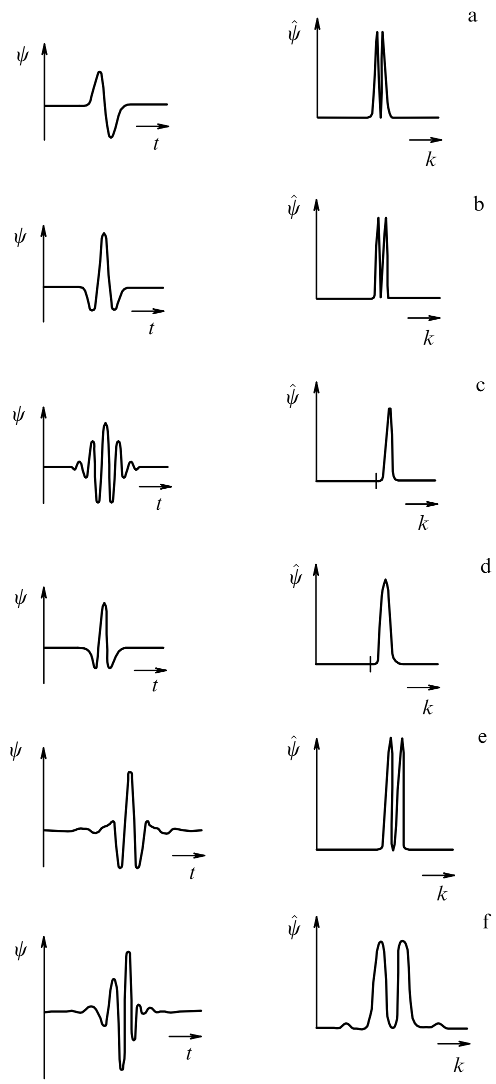

Figure 2. Examples of widely used wavelets: (a) WAVE, (b) MHAT, (c) Morlet, (d) Paul, (e) LMB, (f) Daubechies. Left and right columns show wavelet in time and Fourier representations, respectively.

also the potentialities of a polariser with the polarisation angle being proportional to the angle of wavelet turn.

Based on the Gaussian function, there is the well-known DOG wavelet (difference of Gaussians):

$$\psi(t) = \exp\left(-\frac{|t|^2}{2}\right) - 0{,}5\exp\left(-\frac{|t|^2}{8}\right),$$

$$\hat{\psi}(k) = \frac{1}{(2\pi)^{1/2}}\left[\exp\left(-\frac{|k|^2}{2}\right) - \exp\left(-2|k|^2\right)\right].$$

Examples of complex-valued wavelets are presented in Fig. 2c, d (only their real components are shown). Of them a complex-valued basis derived from the Morlet wavelet well-localised in both k- and r-spaces is most frequently used:

$$\psi(r) = \exp(\mathrm{i}k_0 r)\exp\left(-\frac{r^2}{2}\right),$$

$$\hat{\psi}(k) = \Theta(k)\exp\left[-\frac{(k-k_0)^2}{2}\right].$$

It is a plane wave modulated by a Gaussian of unit width. Figure 2c shows the Morlet wavelet for $k_0 = 6$. With an increase of k_0, the angular sensitivity of the basis also increases, but at the expense of the spatial one.

The Paul wavelet [8], frequently employed in quantum mechanics,

$$\psi(t) = \Gamma(m+1)\,\frac{\mathrm{i}^m}{(1-\mathrm{i}t)^{m+1}}\,,$$

$$\hat{\psi}_m(k) = \Theta(k)(k)^m \exp(-k)$$

is shown in Fig. 2d for $m = 4$ (the greater m, the more zero moments the wavelet has).

The complex-valued wavelets presented above are progressive. In this manner one calls the wavelets which have zero Fourier coefficients with negative wave numbers. They are well fitted to analyse signals for which the causality principle is important: such wavelets preserve the time direction and do not introduce a spurious interference between the past and future.

Notice that in analysing a complex-valued one-dimensional signal, or making use of a complex-valued analysing wavelet, the wavelet transform results in two-dimensional arrays of the module and phase of coefficients

$$W(a,b) = \left|W(a,b)\right| \exp\left[\mathrm{i}\Phi(a,b)\right].$$

Figures 2e, f show examples of wavelets which are often employed to create orthogonal discrete bases [of type (7)] with the help of the Mallat procedure [9]: the LMB wavelet, suggested by Lemarie, Meyer, and Battle [10, 11] and one of the Daubechies wavelets [5]. These are biorthogonal wavelets possessing a conjugate (double) required to obtain the reconstruction formula. In the works cited, the reader may find additional examples of such wavelets and corresponding procedures of their construction.

4. Properties and potentialities of wavelet transform

The one-dimensional Fourier transform offers one-dimensional information on relative contributions (amplitudes) from different time scales (frequencies). The wavelet transform of one-dimensional series leads to a two-dimensional array of wavelet amplitude — the magnitude of coefficients $W(a,b)$. Its distribution in the space (a,b) = (time scale, time localisation) supplies information on the evolution of components with different scales in time and is referred to as the wavelet transform coefficient spectrum, time-scale (-frequency) spectrum, or wavelet spectrum, to distinguish it from the single spectrum of the Fourier transform.

4.1 Result representation methods

The spectrum $W(a,b)$ of a one-dimensional signal represent a surface in three-dimensional space. There are different methods for visualising such surfaces. Instead of presenting these surfaces themselves, one often resorts to presenting their projections on the plane ab by drawing isolines, or isopleths, which enable the changes in the wavelet amplitude intensity to be followed as functions of scale and time. Another approach is to draw patterns of local extremum lines (so-called 'skeleton') which distinctly shape the structure of the process under analysis. The term 'skeleton' reflects properly the structure of local extremum line patterns (see examples), and we shall adopt it for brevity.

In those cases where one needs to display a very broad scale range, it can be more preferably done by using logarithmic coordinates, for example $(\log a, b)$, rather than linear ones.

We illustrate these methods by taking a specific signal as an example. A corresponding physical interpretation will not be touched on. Results shown in Fig. 3 were obtained for the time series of solar wind ion flux (data were collected with MONITOR detector [12] installed on Prognoz-8 satellite; results of their analysis are partly published in Ref. [12]). In calculations, the MHAT wavelet was employed.

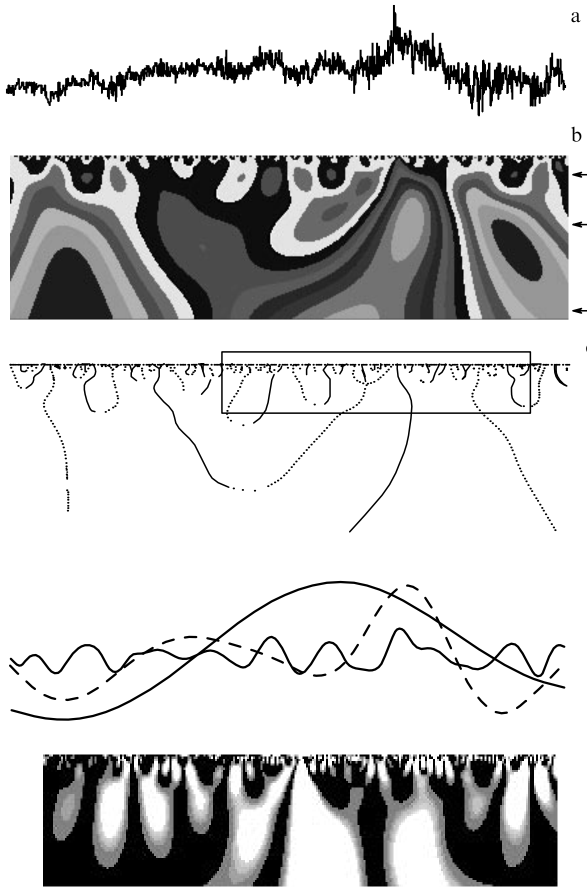

Figure 3. An example of wavelet transform applied to a real signal — solar wind ion flux: (a) signal subject to analysis, (b) pattern of coefficients $W(a,b)$, (c) pattern of local extremum lines, (d) time dependence of coefficients $W(a,b)$ for scales marked by arrows in (b), (e) energy density distribution $E_W(a,b)$ for a fragment within the frame in plate (c).

Figures 3a – c demonstrate, respectively, the series analysed, the pattern of wavelet transform coefficients and the skeleton, two latter projected on the plane ab (time scale, time); the abscissa shows the time (or translation parameter), and the ordinate corresponds to the time scale.

In Figure 3b, dark areas correspond to positive, and light, to negative values of $W(a,b)$; the gray scale is used to outline ranges of values taken by $W(a,b)$ within respective areas. Apparently, the magnitude of wavelet amplitude at a point (a_0, b_0) is the greater (by absolute value) the stronger is the correlation between a wavelet of a given scale and the signal behaviour in the vicinity of $t = b_0$. The pattern of coefficients reveals that the process is composed of components of different scales: the $W(a, b)$ extrema are seen at different scales, their intensity varies both with time and scale.

Figure 3c shows a corresponding picture of local extremum lines — those that join, from scale to scale, the extrema of each feature of the surface (the crest or valley) separately. Many authors believe that the skeleton not only visualises the structure of the process explicitly and without unnecessary details, but *de facto* contains complete information concerning it. In Fig. 3c, the solid lines denote the positions of maxima of $W(a, b)$; the minima are shown by the dotted lines.

We have already mentioned that the wavelet transform decomposes a process analysed into waves that make it up, the components of different scales, and additionally, offers time-localised information about the process. A horizontal section of the pattern given in Fig. 3b, with scale a being specified, shows the temporal course of the scale chosen; Fig. 3d displays the behaviour of three components in the vicinity of values of scale a indicated in Fig. 3b. A vertical section of the coefficient pattern at a particular moment t_0 demonstrates how the process behaves in the vicinity of that moment (one may determine whether an irregularity is present, its order, and the set of scales involved, see the analysis of local regularity in Section 4.3).

It is noteworthy that the value taken by a function analysed at a point t_0 influences the magnitudes of transform coefficients in a temporal domain that grows with the scale, within a so-called angle of influence (Fig. 4a). It seems clear that the influence angle looks different if the scale varies according to some other law (logarithmic, algebraic) which does not coincide with the linear one.

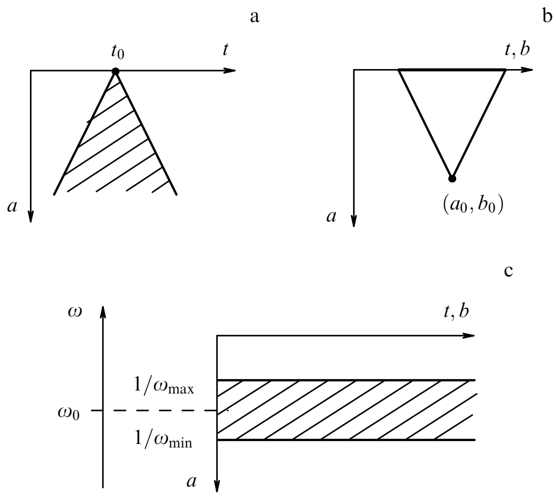

Figure 4. The angle of influence (a, b) and the band of influence (c).

In turn, the coefficient $W(a, b)$ at a point (a_0, b_0) depends on the magnitudes of series terms from the temporal domain (or integration domain) near b_0 falling within the same angle of influence (Fig. 4b). The domain becomes broader as the scale a_0 grows, i.e., high-frequency (or small-scale) information is computed based on small portions of the series, while low-frequency information, is based on the large portions.

The maximum angle of influence (reliability angle) singles out the reliability domain out of which the coefficients $W(a, b)$ are calculated over intervals extending beyond the series boundaries (by complemented series). Since any series analysed is always finite and one always needs maximum information, the approximate data obtained outside the influence angle (with some uncertainty) are often retained. In order to diminish the error the series is supplemented in one way or another with account taken for its behaviour (by mean value, known temporal course, etc.).

The wavelet transform can also be expressed through the Fourier transforms of signal $\hat{f}(\omega)$ and wavelet $\hat{\psi}(\omega)$. It is an easy matter to show that the influence of Fourier component $\hat{f}(\omega_0)$ is experienced by coefficients $W(a, b)$ belonging to a horizontal strip $\omega_{\min} < a\omega_0 < \omega_{\max}$ (see Fig. 4c); in turn coefficient $W(a, b)$ at point (a_0, b_0) is influenced by all Fourier components of the signal $\hat{f}(\omega)$ for which $\omega_{\min} < a_0\omega < \omega_{\max}$.

The span of the influence angle and the width of the influence band are dependent on the basis wavelet. For instance, the conventional MHAT wavelet (see Fig. 2b) is well-localised in time and possesses a narrow power spectrum. This favourable property specifically implies that the coefficients $W(a, b)$ depend on a fraction of wavelet frequency range, i.e. the width of influence band $(\omega_{\min}, \omega_{\max})$ is not large.

Questions related to numerical algorithms remain outside the scope of this study. We shall restrict ourselves to giving merely several brief and useful, in our opinion, practical recommendations.

The discrete wavelet transform is well-fitted to a fast numerical algorithm (see, for example Ref. [13]) based on widely disseminated fast Fourier transform procedure.

The continuous transform is usually implemented through direct numerical integration. The most simple (and quick) verification of the numerical algorithm is to compute the wavelet transform of Dirac's function (as a result, the analysing wavelet should be reproduced at every scale), or a Gaussian function (the result to be found numerically can be readily derived analytically). That the discretisation has a sufficient density for a particular scale can be tested by computing the wavelet transform of the analysing wavelet itself and checking that there are no spurious marginal effects.

In subsequent analysis we will need the properties of the wavelet transform. They all follow from the material already presented in Section 3.

4.2 Wavelet transform properties

We have already mentioned that coefficients of wavelet transform contain combined information on analysing wavelet and a signal analysed. Nonetheless, the wavelet transform allows one to recover unspoiled information on the signal, since some of its properties are independent of a specific choice of the analysing wavelet. This independence makes these simple properties very important.

We write out basic elementary properties of the wavelet transform of function $f(t)$. The designation $W[f] = W(a, b)$ will be adopted for brevity.

Linearity:

$$\begin{aligned} W[\alpha f_1(t) + \beta f_2(t)] &= \alpha W[f_1] + \beta W[f_2] \\ &= \alpha W_1(a, b) + \beta W_2(a, b) . \end{aligned} \tag{22}$$

It follows therefore that the wavelet transform of a vector function is a vector with components which are wavelet transforms of the respective components of the vector analysed.

Translational invariance:

$$W\big[f(t-b_0)\big] = W(a, b-b_0)\,. \tag{23}$$

This property leads to the commutativity of differentiation, in particular, $\partial_t W[f] = W[\partial_t f]$ (here $\partial_t = \partial/\partial t$). Together with the first property it implies the permutability for vector analysis derivatives.

Invariance to dilations (contractions):

$$W\left[f\left(\frac{t}{a_0}\right)\right] = \frac{1}{a_0}\, W\left(\frac{a}{a_0}\,,\,\frac{b}{a_0}\right). \tag{24}$$

This property makes it possible to determine whether a function analysed has singularities and to investigate their character (see Section 4.3.1).

Aside from these three elementary properties independent of the choice of analysing wavelet, the wavelet transform has a few others. In our opinion, the most important and helpful among them are the following.

Time-frequency localisation and existence of time-frequency window and influence angle (it would be more correct to speak of time-scale localisation). The parameters of time-frequency window are given in Section 2.4.

Differentiation:

$$W\big[\partial_t^m f\big] = (-1)^m \int_{-\infty}^{\infty} f(t)\,\partial_t^m\big[\big(\psi_{ab}^*(t)\big)\big]\,\mathrm{d}t\,. \tag{25}$$

Thus to ignore, for example, large-scale polynomial components and analyse singularities of higher order or small-scale variations of function f one may perform differentiation of either the analysing wavelet or the function itself. This is a highly useful property, especially if one remembers that the function f is often defined through a sequence of values while the analysing wavelet is given by a formula.

For the wavelet transform there exists **an analogue of Parseval's theorem** and the identity

$$\int f_1(t) f_2^*(t)\,\mathrm{d}t = C_\psi^{-1} \iint W_1(a,b)\, W_2^*(a,b)\, \frac{\mathrm{d}a\,\mathrm{d}b}{a^2} \tag{26}$$

holds. Hence it follows that the signal energy can be calculated in terms of amplitudes (coefficients) of wavelet transform, just in the manner it is computed through the components of the Fourier transform

$$E_f = \int f^2(t)\,\mathrm{d}t = \int \big|A(\omega) - \mathrm{i}B(\omega)\big|^2\,\mathrm{d}\omega\,.$$

The definitions and properties of one-dimensional continuous wavelet transform can be generalised to multidimensional and discrete cases. Each of them has its own peculiarities. We shall not address them here because further we will only exploit a continuous wavelet transform of one-dimensional functions.

4.3 Some applications of wavelet analysis

Having wavelet spectra available one may compute useful characteristics of the process under study and analyse many of its intrinsic properties. We shall describe in more detail what information can be obtained on signal singularities and power characteristics.

4.3.1 Analysis of local regularity [6, 13]. Consider some consequences of the scale invariance property (24).

If $f \in C^m(t_0)$, i.e. the function analysed has continuous derivatives up to m-th order at a point t_0, its wavelet transform coefficients at $b = t_0$ comply with the inequality

$$W(a, t_0) \leqslant a^{m+3/2}$$

as $a \to 0$.

If $f \in \Lambda^\alpha(t_0)$, i.e. the function analysed belongs to the space of Holder functions with the exponent α (it will be recalled that this implies the function f to be continuous, not necessarily differentiable at t_0, but satisfying the condition $\big|f(t+t_0) - f(t)\big| = c|t_0|^\alpha,\ \alpha < 1,\ c = \mathrm{const} > 0$), the coefficients of its wavelet transform at $b = t_0$ should obey the relation

$$W(a, t_0) \simeq c a^{\alpha+1/2}$$

as $a \to 0$.

The wavelet transform has an inherent property that $W(a,t)$ is a regular function even with $f(t)$ being irregular. Total information on a singularity $f(t)$ may possess (localisation t_0, intensity c, exponent α) is contained in asymptotic behaviour of coefficients $W(a, t_0)$ at small a. If coefficients diverge at small scales then f has a singularity at t_0 and the exponent of the singularity α is wholly determined by the inclination of the dependence $\log\big|W(a,t_0)\big|$ with respect to $\log a$. Conversely, if they are close to zero at small scales in the vicinity of t_0 then this implies f to be regular at t_0.

The property outlined is often successfully used when analysing fractal and multifractal signals [14, 15]. A typical property of fractal sets is their asymptotic self-similarity. When looking at f near point t_0 at different magnifications one sees practically the same function:

$$f(\lambda t + \lambda t_0) - f(\lambda t) \simeq \lambda^{\alpha(t_0)}\big[f(t+t_0) - f(t)\big]\,.$$

Basis of the transform is self-similar; one may easily deduce that the coefficients of the transform also scale with the same exponent as does the function analysed:

$$W(\lambda a, t_0 + \lambda b) \simeq \lambda^{\alpha(t_0)}\, W(a, t_0)\,.$$

Hence the scaling exponent $\alpha(t_0)$ can be derived. As is known, it is closely related to the fractal dimension of the set. In this fashion the analysis of multifractal set may provide the spectra of exponents and dimensions.

It should be emphasised that the analysis of local regularity is in a certain sense universal — it does not depend on the choice of analysing wavelet.

4.3.2 Energy characteristics [6, 16]. Consider some consequences of identity (26). The existence of the analogue of Parseval's theorem for the wavelet transform implies that in the space of real-valued functions the total energy of the signal f can be expressed in terms of the amplitude of wavelet transform as

$$E_f = \int f^2(t)\,\mathrm{d}t = C_\psi^{-1} \iint W^2(a,b)\, \frac{\mathrm{d}a\,\mathrm{d}b}{a^2}\,. \tag{27}$$

The signal energy density $E_W(a,b) = W^2(a,b)$ characterises energy levels (the excitation levels) of the signal $f(t)$ explored in the space $(a,b) = $ (scale, time).

Figure 3e shows the pattern of energy density distribution obtained for a portion of the solar wind ion flux series. In that pattern, the lightest areas correspond to largest values of $E_W(a,b)$, denser colour, up to black, corresponds to decrease of $E_W(a,b)$ down to zero. To better illustrate the details, the distribution of the energy density is shown for a portion of the series and only for the upper third of the scale range (this fragment is singled out by the frame in Fig. 3c).

The fragment shown illustrates that the energy is distributed inhomogeneously over scales — there are some particular scales. Both patterns display the nonstationary structure of the process analysed with elements of quasi-periodicity, evolving frequencies and local periodicity ranges at different scales.

Local energy spectrum. One of the main features of the wavelet transform is its capability to reveal localised characteristics and local properties of processes. Paradoxical as the words 'local power spectrum' may sound, the nature of wavelet transform is nevertheless such that this term has valid grounds for existence. Let us clarify this point.

Given the energy density $E_W(a,b)$ we may, with the help of window, determine the local energy density at a point b_0 (or t_0)

$$E_\xi(a,t_0) = \int E_W(a,b)\xi\left(\frac{b-t_0}{a}\right) \mathrm{d}b \, .$$

Window function ξ 'shapes' the range near t_0 and satisfies the equality

$$\int \xi(b)\, \mathrm{d}b = 1 \, .$$

Choosing Dirac's delta for ξ we arrive at the following expression for the local energy spectrum

$$E_\delta(a,t_0) = W^2(a,t_0) \, .$$

This quantity makes it possible to analyse time behaviour of the energy cascade over scales — the exchange in energy between components of the process possessing different scales at any given point in time.

Global energy spectrum. The total energy is distributed over scales in compliance with the global energy spectrum of the wavelet transform coefficients:

$$E_W(a) = \int W^2(a,b)\, \mathrm{d}b = \int E_W(a,b)\, \mathrm{d}b \, . \tag{28}$$

This quantity is also called the scalogram, or wavelet variance.

Figure 5a shows the power spectrum $E_F(\omega)$ of the solar wind ion flux (see Fig. 3a) and the scalogram E_W — the global energy spectrum of the wavelet transform coefficients obtained for the same signal. The spectra match each other quite satisfactorily, however the spectrum computed by wavelet transform coefficients is considerably smoother. The reason is that the wavelet spectrum of signal energy E_W corresponds to a smoothed power spectrum E_F. One may see that by expressing the energy spectrum $E_W(a)$ through the signal power spectrum in the Fourier space $E_F(\omega) = \left|\hat{f}(\omega)\right|^2$:

$$E_W(a) = \int E_F(\omega)\left|\hat{\psi}(a\omega)\right|^2 \mathrm{d}\omega \, .$$

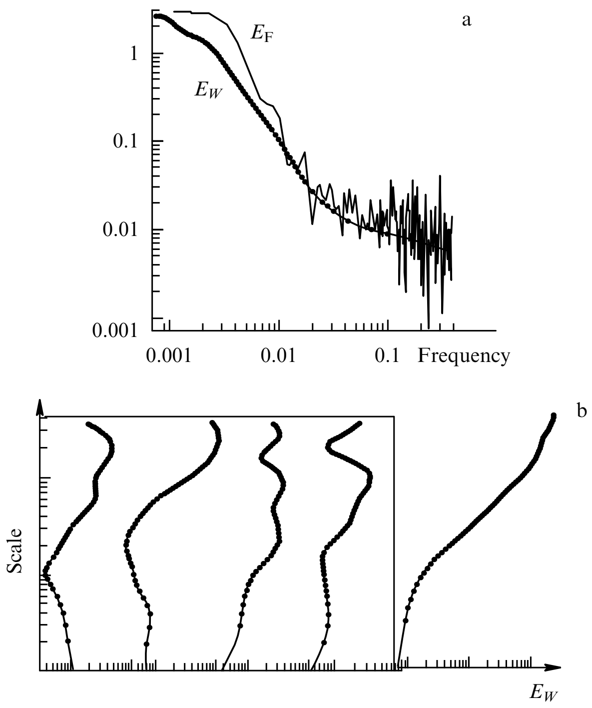

Figure 5. Power spectra of solar wind ion flux: (a) the power spectrum E_F and the scalogram E_W vs. frequency, (b) scalograms for four time intervals (in the frame) and for the whole process vs. scale (increases to the top). Plots are in log-log axes.

Now it becomes apparent that the scalogram E_W corresponds to the power spectrum E_F smoothed at each scale by the Fourier spectrum of analysing wavelet.

The wavelet transform, by supplying us with the time-scale spectrum, allows for more localised information to be obtained. Figure 5b displays four scalograms $E_W(a)$, each found by a convolution with a corresponding quarter of the series, instead of the entire series (see Fig. 3a). For comparison, the same figure shows the global spectrum (in Fig. 5a it is plotted as a function of frequency; in Fig. 5b scalograms are plotted against the scale which increases to the top). The scalograms obtained for particular ranges make it possible to follow the evolution of energy distribution by scales.

The signal energy is expressed through the energy spectrum as

$$E_f = C_\psi^{-1} \int E_W(a)\, \frac{\mathrm{d}a}{a^2} \, . \tag{29}$$

Thus the quantity E_f is proportional to the area under the curve $E_W(a)/a^2$, while the scalogram reflects relative contributions from different scales to the total energy and reveals the energy distribution over scales.

The function under analysis is of finite energy and the analysing wavelet is of zero mean. Therefore the energy spectrum $E_W(a)$ should tend to zero at both extremes of the scale domain and consequently exhibit at least a single maximum. The positions of maxima (peaks) of that kind in the Fourier spectrum $E_F(\omega)$ are conventionally associated with frequencies and corresponding intrinsic modes of the signal analysed, in which the main fraction of signal energy is contained. Maxima of energy spectrum $E_W(a)$ are treated in a similar way — they define those scales of the process which contribute maximally to the total energy E_f.

Using a simple example, let us illustrate the links between the scale emerging as a result of wavelet transform and that recovered from the Fourier spectrum. Let

$$f(t) = \sin(\omega_0 t) = \sin\frac{2\pi t}{T_0};$$

its wavelet transform expressed through the Fourier transforms [see, for instance, Eqn (17)] is

$$W(a,b) = \frac{i}{2}\left[\exp(i\omega_0 b)\hat{\psi}(a\omega_0) + \exp(-i\omega_0 b)\hat{\psi}(-a\omega_0)\right],$$

while $E_W(a) = \left|\hat{\psi}(a\omega_0)\right|^2$ is its spectrum. The necessary and sufficient condition for the presence of peak at a scale $a = a_0$ is that of derivative $\mathrm{d}\hat{\psi}(a\omega_0)/\mathrm{d}a$ being zero at $a = a_0$. This condition is satisfied at $a_0\omega_0 = \omega_\psi$, where ω_ψ is a constant, dependent on the wavelet ψ, with dimension of frequency. For many wavelets the constant ω_ψ yields to analytical calculations: given HAAR and MHAT wavelets, it is equal 1.484π and $\sqrt{2}$, respectively.

In practice, if scalogram $E_W(a)$ shows a peak at $a = a_0$, a characteristic scale is defined as $d = T_0/2 = a_0\pi/\omega_\psi$. A factor of $1/2$ appearing here is responsible for the fact that one is not looking for a period, but for a scale of elementary event or detail. In this sense a sine function shows two elementary details over a period.

The constant ω_ψ is found for a simple function. Expanding the result on arbitrary, even non-harmonic signals we assume that the location of the maximum of the spectrum $E_W(a)$ (i.e. the scale revealed) could be interpreted as the mean duration of an elementary event (events) contributing most efficiently into the energy of the process under study. This fact is verified for numerous known signals with different wavelets and is recognised to give a very good approximation (see, for example, Ref. [16]).

We present two more characteristics that are expressible in terms of energy density — the measure of local intermittency and the contrast measure of the signal under analysis.

Measure of local intermittency

$$I_W(a,t) = \frac{E_W(a,t)}{\left\langle E_W(a,t)\right\rangle_t}$$

is the measure of local deviations of the spectrum at each scale from the mean field; it enables one to assess the degree of nonuniformity of energy distribution over scales (the angular brackets denote averaging).

That $I_W(a,t) = 1$ for all a and t implies that energy is distributed uniformly and all local energy spectra are similar; $I_W(a,t_0) = \alpha$ implies that the contribution coming from a component with scale a at a point t_0 exceeds α times the average over t.

Time-scale contrast measure

$$C_W(a,t) = \frac{E_W(a,t)}{E'_W(a,t)}, \qquad E'_W(a,t) = \int_{a'=0}^{a'=a} E_W(a',t)\,\mathrm{d}a'$$

serves to detect even the slightest variations in a signal when one needs, for example, to reveal the structure of a weak signal or weak variations against the background of extended structure (build-in structures).

5. Applications of the wavelet transform to model signals

In this section we show the utility of the wavelet transform as regards visualising various typical features contained in signals. The wavelet transform is applied to model signals composed of functions with well-known properties (see also Ref. [17]).

For each example we present a plot of the series analysed. The coefficients $W(a,b)$ are shown as projections on the plane ab (time scale, time). Time is plotted against the abscissa, and time scale (it grows linearly to bottom), against the ordinate. Similar to Fig. 3b, light and dark areas correspond, respectively, to positive and negative values of $W(a,b)$, with gray scale used to show the ranges of $W(a,b)$ magnitudes. Patterns of local extremum (or local maximum) lines are presented in the same coordinates.

The results presented are obtained with the MHAT wavelet. Calculations were carried out for a rectangular area in the plane of parameters a, b. Data series were continued to achieve that. Particular ways of continuation are shown in plots of functions analysed. In patterns of local maxima lines, the reliability triangle, or the angle of influence are shown.

5.1 Harmonic function

The wavelet transform was applied to series of sine functions

$$f(t) = \sin\frac{2\pi t}{T_1} + \alpha\sin\frac{2\pi t}{T_2}.$$

The results of wavelet transform of such a function can be easily compared with those given by the conventional Fourier transform. Magnitudes of periods T_1 and T_2 as well as constant α for the signals analysed are given in Table 1. It also displays the numbers of the respective figures and signals. Coefficients of wavelet transform are expected to provide scales $D_1 = T_1/2$ and $D_2 = T_2/2$.

Signal 1. It is shown in Fig. 6a and is composed by sine functions with noticeably different frequencies (as one may see from the plot, the series is continued by the mean value).

The pattern of coefficient magnitudes, Fig. 6b, reveals multiple periodically recurrent details in its upper part (at small values of scale a) which evolve from a resonance between high-frequency component of the signal and small-scale wavelets, as well as three dark and two light areas at large scales (positive and negative values of $W(a,b)$, respectively), which are the result of a strong correlation between large-scale wavelet and low-frequency component of the signal containing only two and a half of period.

Figure 6c displays lines of local extrema; the solid lines denote local maxima and the dotted ones correspond to local minima. The pattern of local maxima (Fig. 6d) shows only the

Table 1. Parameters of harmonic functions.

T_1	T_2	α	Number of figure	Number of signal
200	10	0.4	6	1
200	10	0.4	–	2
50	–	–	–	3
25	–	–	8	4
25	50	1	9a	5
25	50	1	9a	6
25	23	1	10a	7
25	23	1	10b	8

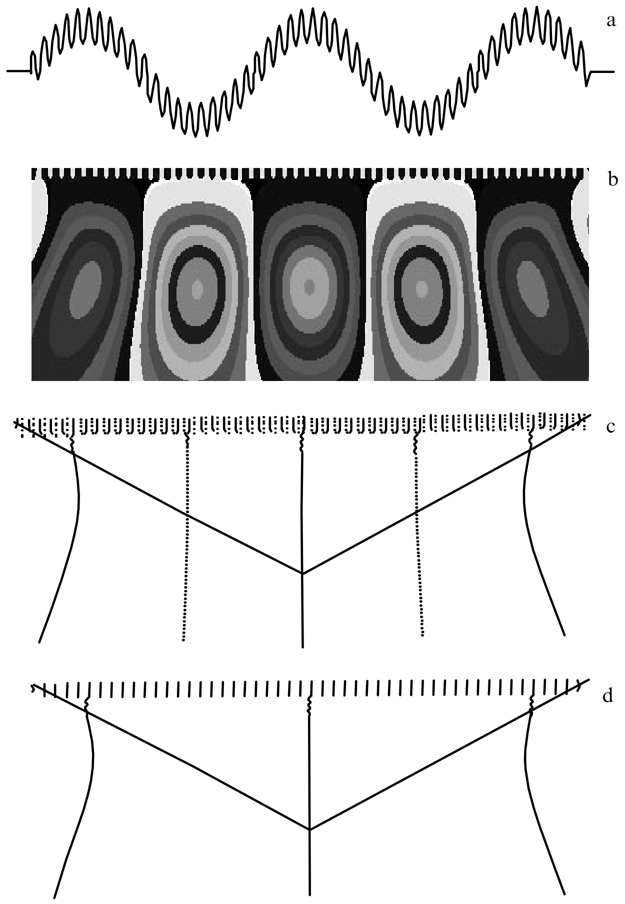

Figure 6. Signal 1 (a) and the results of its wavelet transform: patterns of coefficients (b), local extremum lines (c) and local maximum lines (d).

lines related to positive extrema. Both patterns reveal a periodic structure of the signal analysed and represent the time-scale skeleton of the process described by the signal. Lines marking off the positions of local extrema of wavelet-transform coefficients are related to the extrema of function studied — points where its derivative changes sign.

In what follows we will demonstrate only one of the skeletons; if lines of local minima bear no particular information and only complicate the pattern, as they do for signal 1, they will be excluded and only the pattern of local maxima lines will be illustrated.

Signal 2 represents a part of signal 1. The pattern of wavelet transform coefficients permits one to speak of the large-scale component in the signal despite its entire extent is covered by only a single period of the large-scale component of signal 1.

Energy spectrum $E_W(a)$ in this simple case allows for determining even the extension of the large-scale component in the signal, though it is represented by a single period. Figure 7 shows the energy spectrum $E_W(d)$; it will be recalled that for the MHAT wavelet the typical time-scale d is linked with the scale a of wavelet transform as $d = a\pi/\sqrt{2}$. The solid line in this figure plots the energy spectrum of signal 2, dashed line plots the same for signal 1. In both cases the peaks corresponding to scales of 5 and 100 are easily discernible; the small-scale part of the spectrum is given in more detail in the insert in the upper right corner of the figure. Given the Fourier spectra, one is able to recover only the high-frequency constituent in these signals.

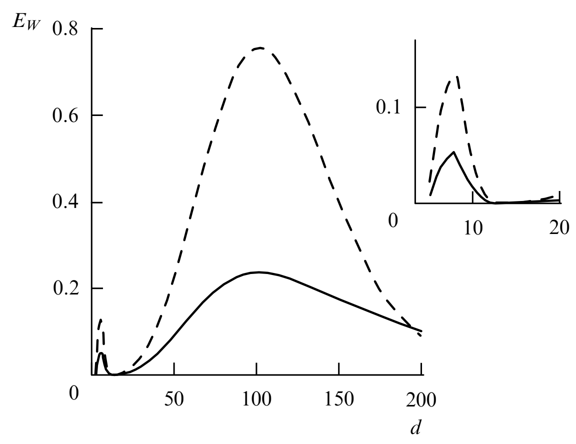

Figure 7. Power spectra $E_W(d)$ of signals 1 (dashed line) and 2 (solid line); in the upper right corner the small-scale part is displayed at magnification.

Signals 3 and 4 are sine functions with periods of 50 and 25, respectively. Figure 8 presents signal 4 and patterns of wavelet transform coefficients and local maxima lines obtained for it. Upper parts of the patterns demonstrate a periodic behaviour of signals. Dark and light large-scale details in the lower part of coefficient pattern are due to boundary effects and are of very low intensity; in patterns this may be recognised by noting that these large-scale areas have substantially fewer colour levels than basic periodic details (here — only a single level), and wave lines of local extrema.

Both the same length of local maxima lines (within the reliability triangle) and the periodicity with which they appear point to a single typical frequency of the signal and a fixed period. This is also supported by a peculiar kind of a 'hatched' structure shown by the skeleton in the bottom part of the reliability triangle; if there were a few frequencies present (see results for signal 6 and 8) it would change to an 'interference' structure.

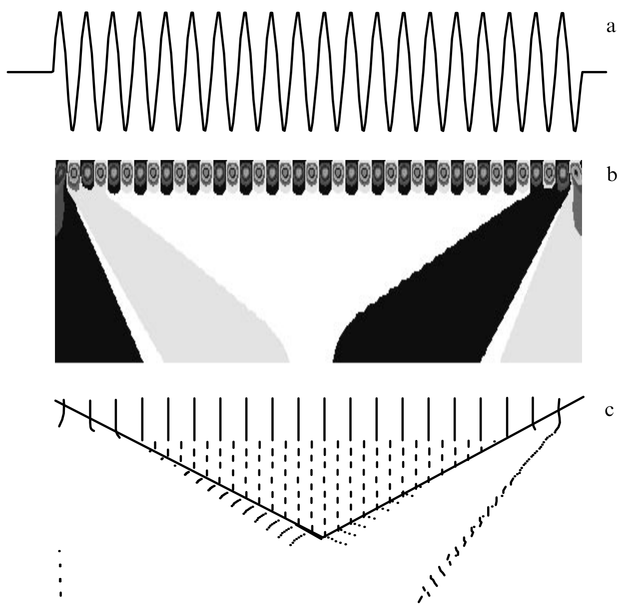

Figure 8. Signal 4 and patterns of coefficients and local maximum lines.

That distinctive interference (or hatched, as for signal 4) structure of skeleton appears in cases when the length of series is sufficient for all inherent scales to be realised within the reliability triangle (this holds for all signals except for the first one) and owes its existence to the fact that after realisation of all scales the coefficients $W(a, b)$ decrease abruptly, become very small and approach zero in an oscillating manner (hence wave lines of local maxima).

The spectra $E_W(d)$ computed by magnitudes of wavelet transform coefficients for signals 3 and 4 exhibit peaks at scales 25 and 12.5, respectively.

Signal 5 and 6 are in fact different combinations of 3 and 4. Signal 5 is composed of 3 and 4 switched on in turn while signal 6 is their sum. Signals 5 and 6 are of interest since they are indistinguishable for the Fourier transform: one may easily verify that their Fourier spectra $E_F(\omega)$ practically coincide. For the wavelet transform they are quite different, as may be concluded from the patterns of coefficients and local maxima lines computed for them and given in Fig. 9a and 9b.

In contrast, wavelet energy spectra of these signals as well as their Fourier counterparts are very similar because they are obtained by convolution over the entire series length. They possess broad maxima covering both scales. We note that nonstationary properties of signals, say, evolving frequencies (scales) could be well resolved with the help of localised spectral analysis when, in calculating spectra $E_W(d)$, the convolution of wavelet transform coefficients is performed over a fraction of series instead of its total length.

Signal 7. Figure 10a displays results found with a signal that differs from 4 only in a phase shift by π in the middle of the series. Clearly, the Fourier transform of this signal differs from that for the signal without phase shift by showing additional peaks. The explanation is that any detail of the signal influences all Fourier coefficients and frequencies.

The presence of such details in a signal, with account for a direct link between frequencies in spectrum and characteristic scales of the process, may noticeably distort the interpretation. Commonly, the presence of additional peaks in spectra is attributed to the existence of several scales which are completely absent in the example considered: there is only a single scale.

The wavelet transform copes with a singularity of that kind by localising it (see patterns of coefficients and local maxima lines in Fig. 10a).

Signal 8. Figure 10b shows results obtained for a signal composed of two sine functions with very close frequencies; that the second frequency is present is only seen in tilted local maxima lines. Similar tilts may also appear due to low-frequency modulation, however in that case there will be no distinctive skeleton interference pattern which indicates that all scales present in the signal are realised.

5.2 Signal with a singularity

Signals frequently contain isolated singularities in the form of pulse, jump, power-like singularity, etc. These may be either the details intrinsic to the process in hand, or spurious details caused, say, by instrumental failures. The Fourier transform of a signal that is regular everywhere except for a single singular point bears information about that point in all its coefficients. Isolated singularities are practically not amenable to filtration and thus distort both the spectrum and the signal recovered.

The wavelet transform was applied to signals with singularities of some types mentioned above. Figure 11, along with signals, presents local extremum line patterns (they are more informative) and several isopleths of coefficients $W(a, b)$ drawn schematically (positive and negative values are represented by solid and dashed lines, respectively).

All point singularities contained in signals are accompanied by local maximum lines emanating from these points. Their number depends on the type of the singularity and analysing wavelet.

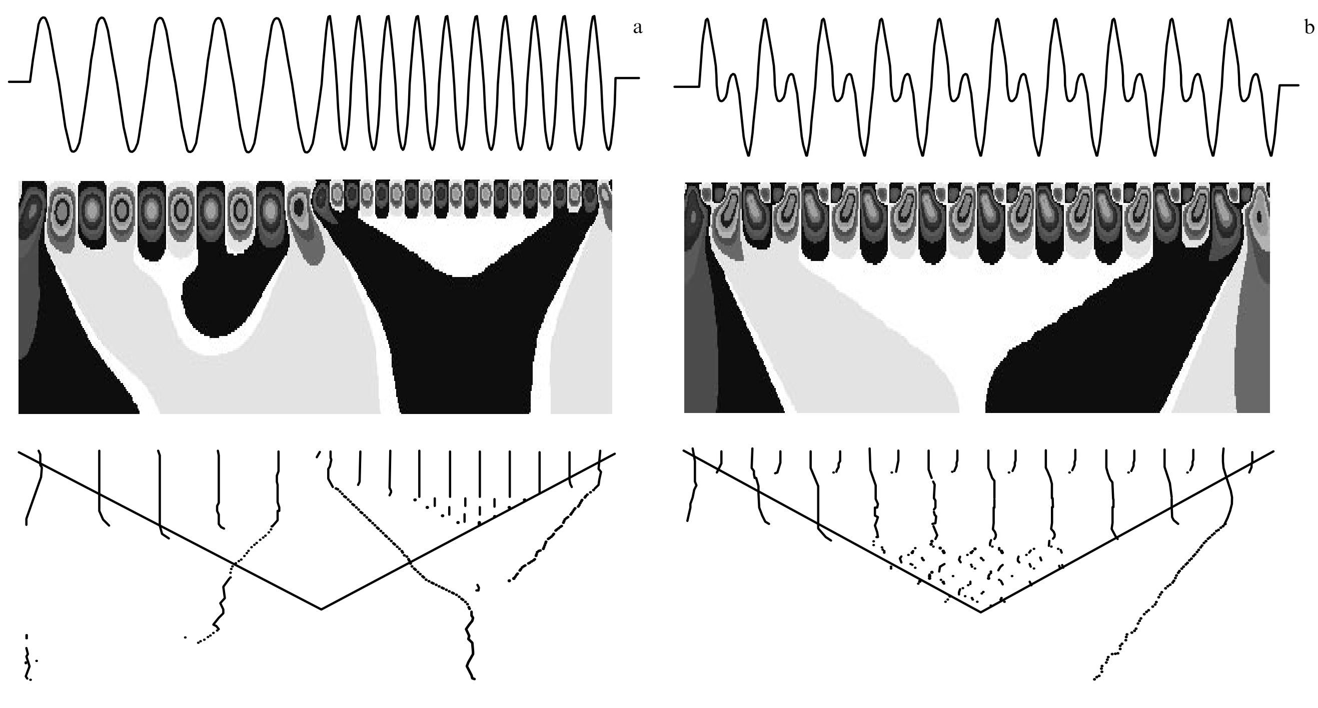

Figure 9. Signals 5 (a) and 6 (b) — the same as in Fig. 8.

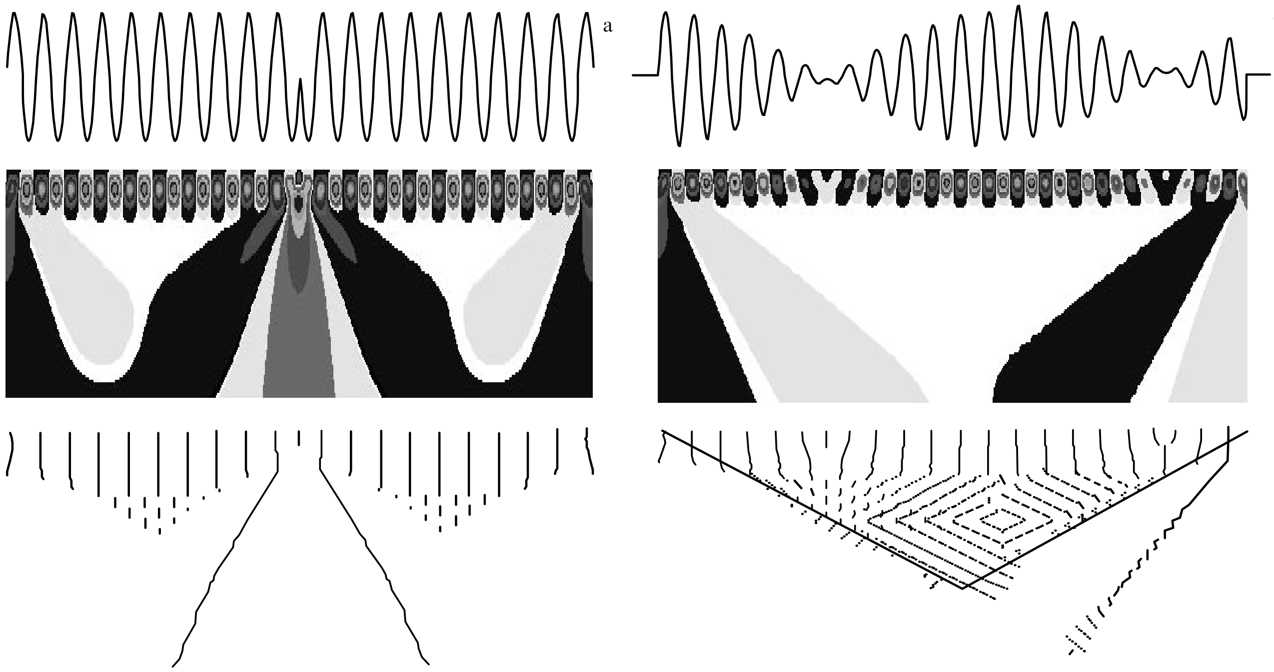

Figure 10. Signals 7 (a) and 8 (b) — the same as in Fig. 8.

Examples from Section 5.1 have indicated that if the derivative of function analysed changes its sign, this is seen as a local extremum line of distribution $W(a, b)$.

Figure 11a shows the transform of a signal with δ-like singularity. Two next examples (Fig. 11b, c) contain singularities in which the change of derivative sign is accompanied by the function discontinuity (algebraic singularities of $|t^5|$ and $t^{2/3}$ forms, respectively). In this case the point of singularity localisation is associated with three lines of skeleton. The central line is that of minimum or maximum depending on whether the signal attains its maximum or minimum at this point. Results of wavelet transform for the δ-function look similarly.

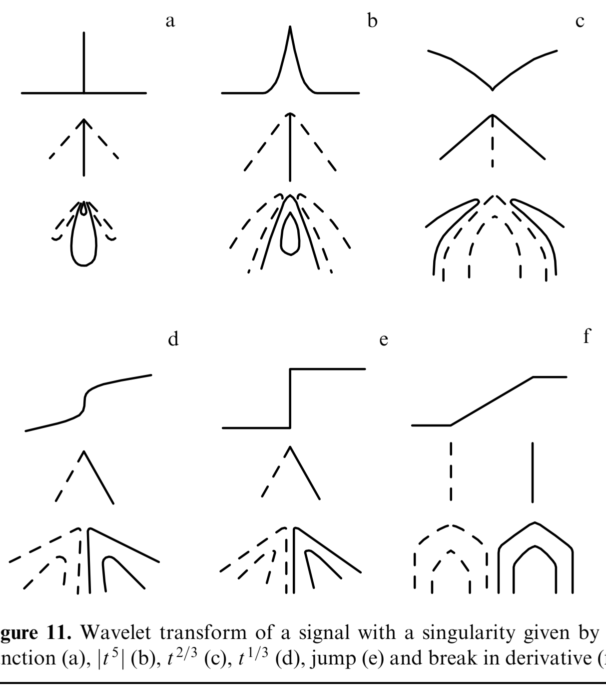

Figure 11. Wavelet transform of a signal with a singularity given by δ-function (a), $|t^5|$ (b), $t^{2/3}$ (c), $t^{1/3}$ (d), jump (e) and break in derivative (f).

Singularities of lower order, such as inflection of function $t^{1/3}$ (Fig 11d) or jump (Fig. 11e) bring about two local extremum lines, and a break of derivative (Fig. 11f) leads to a single line.

Thus the wavelet transform is capable of revealing the location of singularity — wavelet transform coefficients of a smooth function are small and increase abruptly as a singularity is encountered, visualising it by local extremum lines. The character of singularity at the point can be determined from the asymptotic behaviour of wavelet transform coefficients as the scale tends to zero. For example, the coefficients of the wavelet transform of δ-function are maximal at small scales and decay abruptly with scale growing, tracing the behaviour of the singularity itself. An isolated singularity influences wavelet transform coefficients locally and can be easily eliminated from the signal or corrected.

The results given here were obtained with the help of MHAT wavelet possessing two zero moments (zeroth and first). The wavelets with only one zero moment are not capable of distinguishing singularities of derivatives. The wavelet transform coefficient distribution obtained with such wavelets exhibits less local extremum lines. For example, for the first three signals there will be no central lines; the singularity shown by the derivative of a signal with inflection (Fig. 11d) will not be seen at all, and the jump in signal (Fig. 11c) will lead to a single extremum line.

The higher the order of analysing wavelet and the greater the number of zero moments, the better the wavelet transform distinguishes between singularities.

Note that specific features of wavelet transform with different order wavelets can be advantageously used to reveal the presence and behaviour of trends, the largest-scale components in signals. On applying the reconstruction formula to coefficients and subtracting the resultant signal from the original one, such large-scale component can be separated out: since wavelets have zero moments they convert to zero a constant contribution, linear or polynomial trends, etc. In this way, subjectivity can be avoided which is almost always present if trends are determined by other means.

5.3 Fractal set

The wavelet transform owing to its hierarchical basis is well-suited for analysing cascade processes, fractal and multi-fractal sets which have a hierarchical nature.

We present an example concerning analysis of a fractal set formed on the basis of a homogeneous triadic Cantor set. As known, when constructing the first generation of this set, an interval is divided into three parts and the middle of them is excluded; for the second generation the same procedure is applied to two remaining intervals, and so at each subsequent stage, to infinity. Figure 12a displays the first stages of this construction.

Based on the set constructed, the Cantor dust, a numeric set, is created from zeros and ones (zeros correspond to excluded parts of the interval).

Figure 12b presents the patterns of coefficients and local maximum lines. They are reasonably detailed, however linear scale sweep does not allow for presenting a broad scale range. To demonstrate general character of the process, Fig. 12c shows the skeleton in logarithmic axes.

The pattern of coefficients displays the hierarchical structure of the set presented. Even more clearly it is seen in the patterns of local maximum lines. The skeleton not only reveals the hierarchical structure, but also shows how the fractal measure was constructed on which the set is formed.

Every stage of the cascade process, every scale subdivision is marked off on the local maximum pattern by branching, the appearance of a peculiar 'forge': the line marking the local maximum position bifurcates into two independent local maximum lines. This is the invariably recurrent feature since the measure is self-similar and monofractal.

It is known that the fractal dimension, or self-similarity dimension of the homogeneous Cantor set $D_{\mathrm{f}} = \ln m / \ln s$, where m is the branching rate and s is the scale factor. In the case of triadic set $D_{\mathrm{f}} = \ln 2 / \ln 3$. The dimension can also be assessed by using wavelet transform coefficients as the limit (with scale going to zero) of ratio $\ln N(a) / \ln a$; here $N(a)$ is the number of local maxima. The higher the order of the Cantor set generation used, the more accurately its dimension can be determined; for 10th–11th generations the value computed by wavelet transform coefficients coincides practically with that found analytically.

For comparison, Fig. 12d pictures local maximum lines of a random process. One may see how different, even qualitatively, are the 'tree-like' structure of the skeleton of the cascade process and the 'grass-like' skeleton of the random process (they could be compared with periodic skeletons of harmonic functions and 'bushes' of lines marking off the singularities of signals).

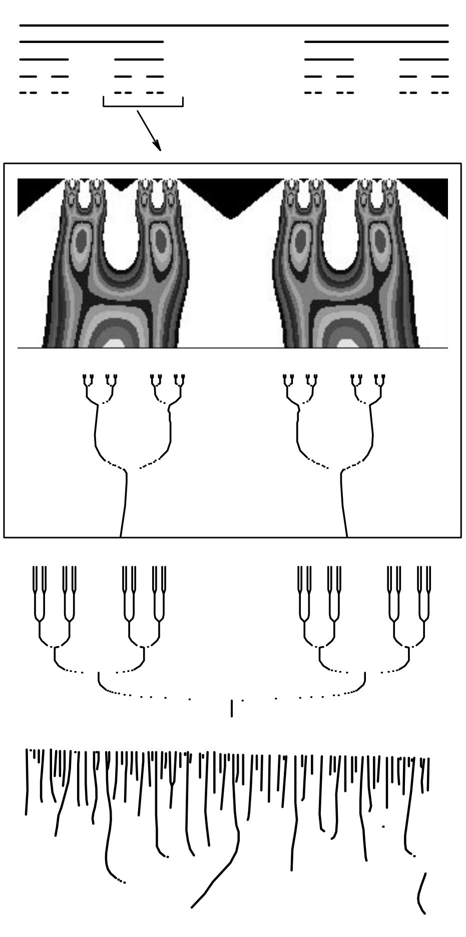

Figure 12. First generations of a homogeneous triadic Cantor series (a), fragments of patterns of coefficients and local maximum lines (b), patterns of local maximum lines for a triadic homogeneous Cantor series (c) and a random process (d) in logarithmic scale.

6. Analysis of meteorological time series

In this section we present results that ensue from the analysis of real data — long-term observations of variations in certain meteorological parameters.

6.1 El Niño and Southern Oscillation

The Southern Oscillation is a large-scale process that develops over the extent of the Pacific ocean. It is closely related to the El Niño phenomenon, the sudden warming of oceanic waters along the Pacific coast of Central America. The global process of planetary scale in the atmosphere–ocean system, El Niño and the Southern Oscillation (ENSO), has a pronounced impact on the dynamics of the entire planet's climatic system, by influencing the Hadley and Walker circulations and the locations of regions of active convection in the tropics.

Processes in the tropics, supposedly, have an important bearing on the climate dynamics on scales of decades or more. We shall describe briefly this extremely interesting and important phenomenon, the ENSO, a peculiar kind of dialogue between the wind and the sea.

El Niño (Spanish for the Christ Child) is the name the fishmen from Ecuador and Peru use to call a warm stream in the ocean that appears most frequently close to Christmas and persists during a certain period. There is no fishing during that time since near-coast upwelling is suppressed. The warm stream (and vacancies) may appear in May, or even in June, but the name El Niño for it and associated phenomena has become conventional.

In 1920s a known British scientist Gilbert Walker worked in India on the problem of forecasting monsoons that sometimes bring devastating rainfalls. In particular, while studying barometric data he discovered a dependence between data recorded at stations in the western and eastern Pacific. The term 'Southern Oscillation' was introduced by him to designate anomalies in atmospheric surface pressure along the tropical belt. The alteration of anomaly signature resembles a gigantic pendulum pumping air masses between the Eastern and Western Hemispheres.

Near the action centres of the Southern Oscillation where its signature has opposite signs there are stations located on Tahiti island (17° S, 150° W) and in Darwin (12° S, 150° E). It is agreed that the time series of difference in normalised pressure anomalies at these stations [the Southern Oscillation index $C(t)$] is the most pertinent characteristic of the Southern Oscillation time behaviour.

That there exist coupling between these two phenomena, El Niño (in the ocean) and the Southern Oscillation (in the atmosphere) was perceived much later, after a strong El Niño of 1957. For long, this coupling was interpreted using the Wyrtki hypothesis [18], well illustrated in Ref. [19]. Its essence is as follows.

Under normal conditions, $C \cong 0$, north-east and south-east trade winds drive warm water into the western Pacific; the sea level there is 40 cm above that in the eastern part. The circulation is accompanied by upwelling — the rise of deep water, which is cool and rich in nutrients, near the Pacific coast of South America. If $C > 0$, the phenomena outlined are more pronounced.

When the index C decreases and becomes negative the pressure gradient between the eastern and western zones of the tropical Pacific weakens noticeably. Without opposing wind drag, warm waters flow to the east, reach South America and split into northward and southward currents and a reflected wave propagating westward. The region of warm water quickly broadens.

The increase in temperature in the eastern and central Pacific changes locations of convective regions in the atmosphere. Convection is usually bounded to Indonesia and the western Pacific. As the index of the Southern Oscillation diminishes, a period of very dry weather begins in the area of Australian-Indonesian centre of activity, while in the central and eastern Pacific, normally poor in rains, heavy rainfalls begin. The population of coastal South America suffers from floods and squalls; due to the cessation of the near-coast upwelling and transport of water rich in nutrients, the fishes, birds, and animals migrate or perish. These periods happen to be persistent and then bring about a real ecological disaster.

The change in location of atmospheric convective regions involves not only the Pacific, but also the entire tropics. Dry weather comes to the western coast of Africa and South America where the precipitation level is usually normal. The displacement of tropical cyclone tracks is observed. During El Niño the number of days with tropical cyclones considerably decreases in the Atlantic and increases in the French Polynesia.

A nice illustration to the coupling between the Southern Oscillation and El Niño is suggested by results of Ref. [20] which shows the time-latitude section of temperature anomalies at the equator and a synchronous course of the Southern Oscillation index for 1979–1989. Isotherms there are plotted using the data from the Washington Center for Climate Analyses. The time span considered includes three warm and three cool episodes.

Comparison of latitude-time sections of surface temperature anomalies at the equator with the synchronous course of the Southern Oscillation index $C(t)$ reveals certain regularities. There are at least three points worth mentioning. There exists a negative correlation between C and the oceanic surface temperature: the greater C the lower is the temperature, and the larger the absolute value of negative C the greater is the temperature. Positive temperature anomalies are more persistent than the negative phases of index C: during El Niño the decrease in C begins simultaneously with the appearance of positive temperature anomaly in the eastern Pacific, whereas the increase of C after passing a minimum begins 3 – 4 months prior to the temperature drop. The temperature in the equatorial zone of the Atlantic varies independently of anomalies in the Pacific, while temperature oscillations in the west and equatorial Indian ocean and in the eastern Pacific correlate.

Anomalous warming of waters in the Pacific and Indian oceans entails the warming of the equatorial atmosphere and sharpens the pole–equator temperature contrast. In turn, this intensifies zonal circulation.

The unusual localisation of regions of augmented convection perturbs the circulation of the atmosphere not only within the equatorial belt, but over the globe. Weather anomalies caused by that are observed also in middle latitudes. Thus, a strong El Niño of 1982–1983 resulted in extraordinary strong cyclones passing during winter of 1982/83 from the northern Atlantic across Scandinavia to the east. It will be recalled that storms associated with them smeared the Kursh spit. Many natural phenomena occurring during El Niños led to very severe consequences both ecologically and economically.

It is therefore not surprising that this fascinating process, the Southern Oscillation – El Niño, and its outcomes are for many years the subject of constant scientific interest. The last decade was dedicated to its thorough study in the framework of the TOGA program. New data have been obtained and the interpretation of coupling between the Southern Oscillation and El Niño, based on the Wyrtki hypothesis, did suffer certain modifications. It was established that the periodicity and scenarios of the ENSO alter with time noticeably. For example, the surface temperature anomaly can migrate not only westward, but also eastward, the pool of warm water may not reach the Ecuador and Peruvian coasts and remain bounded within the central Pacific, and so on.

There is a large volume of publications dealing with El Niño and the Southern Oscillation, however, many details still have not received rigorous treatment; especially this concerns the scales exceeding decades and less than a few months.

The following time series are the object of our study : the data that bear witness to the dynamics of El Niño events for the past 500 years from Ref. [21] (Section 6.2), monthly mean magnitudes of the Southern Oscillation index $C(t)$ for a period from 1882 to 1992 from Ref. [20] (Section 6.3) and daily magnitudes of index $C(t)$ for a period from 1981 to 1991 (Section 6.4) collected by collaborates of D M Sonechkin (Gidrometeorological Centre of Russia); partly the results of analysis of the Southern Oscillation index dynamics are published in Ref. [22].

To reveal the coupling between the global warming and El Niño we analysed the Jones series [23] — those of anomalies in semiannual surface air temperatures (both global and hemispheric) for a period from 1854 to 1990 (see Section 6.5).

6.2 500 years of El Niño history

Data on pressure anomaly and surface temperature embrace a period slightly exceeding a century since regular observations began in the recent past. There have been numerous

Table 2. El Niño events from 1470 to 1989 [21].

Years	Events	Sum	Years	Events	Sum
1470s	73-74, 79	2	1730s	36	1
1480s	84, 87-88	2	1740s	40, 42, 44, 47	4
1490s	90,93, 96-97, 99	4	1750s	54, 56	2
1500s	09-10	1	1760s	61, 63, 65-66, 69	4
1510s	17-18	1	1770s	71, 73, 75, 78-79	4
1520s	20-21, 25-26, 29-30	3	1780s	82-83, 86-87	2
1530s	32-33, 39	2	1790s	91, 98	2
1540s	41, 45, 47	3	1800s	03-04, 06-07	2
1550s	52-53, 59-60	2	1810s	12-13, 15, 17	3
1560s	65, 67-68	2	1820s	21, 24-25, 28	3
1570s	78	1	1830s	31-32, 36-37	2
1580s	85-86	1	1840s	44-45	1
1590s	90-91, 95-96	2	1850s	50, 54-55, 57-58	3
1600s	00, 04, 07	3	1860s	64, 66, 68	3
1610s	10, 14-15, 18-19	3	1870s	77-78	1
1620s	24-25	1	1880s	80-81, 84, 88	3
1630s	34-35, 37-38	2	1890s	91, 96, 99-00	3
1640s	40-41, 47	2	1900s	02, 04-05	2
1650s	50	1	1910s	11, 13-14, 18-19	3
1660s	60-61	1	1920s	25-26	1
1670s	71, 74	2	1930s	30	1
1680s	80-81, 84, 87-88	3	1940s	40-41, 44-45	2
1690s	92, 96	2	1950s	51, 53, 57-58	3
1700s	01, 04, 07	3	1960s	63, 65, 68-69	3
1710s	15-16	1	1970s	72, 76	2
1720s	20-21, 23, 25	3	1980s	82-83, 86-87	2

Table 3. Occurrence of El Niño events per decade — the demonstration of 70-year cycle [21].

Years, events	1500s	1510s	1520s	1530s	1540s	1550s	1560s
	1	1	3	2	3	2	2
	1570s	1580s	1590s	1600s	1610s	1620s	1630s
	1	1	2	3	3	1	2
	1640s	1650s	1660s	1670s	1680s	1690s	1700s
	2	1	1	2	3	2	3
	1710s	1720s	1730s	1740s	1750s	1760s	1770s
	1	3	1	4	2	4	4
	1780s	1790s	1800s	1810s	1820s	1830s	1840s
	2	2	2	3	3	2	1
	1850s	1860s	1870s	1880s	1890s	1900s	1910s
	3	3	1	3	3	2	3
	1920s	1930s	1940s	1950s	1960s	1970s	1980s
	1	1	2	3	3	2	2
Sum	11	12	12	20	20	15	17

attempts to reconstruct more than a thousand-year-long history of the ENSO. In these attempts one resorts to indirect evidence which may shed light on El Niño events (fossil remnants, chronicles, etc.). For example, this may be accounts about droughts and floods, snow and ice layer states on mountain tops, data on fossil microflora, the growth rate of corral skeletons, rings of tree growth, etc.

In our opinion, the data in Ref. [21] on frequency of El Niño events for the last 500 years seem to be the most reliable. They are presented in Table 2 of Ref. [21] and depart from the evidence on the number of typhoons in South China, cold winters in Western Asia, droughts in Australia, floods on Nile and in Peru which might be caused by El Niño events.

The series analysed consists of zeros (for years when, as the author of Ref. [21] suggests, all indirect data point to the absence of El Niño) and ones (in El Niño years). For a period lasting from 1470 to 1987, 114 El Niño events were discovered, with a mean interval between them of about 4.5 years (Table 2 lists these particular years). The author of Ref. [21] plausibly suggests that there may exist a 70-year cycle and as an illustration gives data averaged by decades and arranged in order presented in Table 3.

Figure 13a plots the dependence of El Niño events against time. For the sake of convenience, numbers presented are averaged over decades — it is seen that for a decade there occurs from one to four El Niño events. The pattern of wavelet transform of the basic series, consisting of zeros and ones, is shown in Fig. 13b. In this figure the time axis is aligned with the abscissa of the plot, and the scale increases to the bottom, up to values of 300 years (to make explicit all large-scale details that can be recovered with the finite series available).

In Figure 13b, one's attention is called to the fact that a scale of about a century subdivides the coefficient pattern into two noticeably different regions. In the lower part of the pattern, one may discern only two and a few more of large-scale details (with a scale surpassing 150 years). A completely different structure is seen in the upper part of the pattern: the dynamics is almost entirely bounded to scales below 100 years.

The distribution of energy density $E_W(a,b)$ shown in Fig. 13c for the upper third of the scale range (up to 100 years) reveals in more detail how the process behaves with time. Despite the apparent nonstationarity, a structure resembling a periodic one can be discerned on certain scales. Both at the beginning and the end of the time interval analysed there are three details with scales of 40 years (they are less intense at the end of the series). In the middle of the pattern this 40-year periodicity is interrupted — here only several details with scales of 25 – 30 years are seen which form two large-scale details (approximately of 85 – 90 years). Additionally, we may distinguish a few ranges of local periodicity with scales ranging from 8 to 11 years (or even less).

A scale of 9 – 10 years is conventionally attributed to an 18.6-year moon cycle which is present in the dynamics of droughts and floods, and is seen in the temperature course in North America [24]. A scale of nearly 40 years may be connected with 70 – 80 year cycle of El Niño events mentioned in Refs [20, 25]; its origin is still unclear. Attempts to relate it to volcanic activity do not seem to be well substantiated. The third typical scale of about 150 years (three details of largest scale in Fig. 13b) may be of limited accuracy due to the finite extent of series analysed and hence considered doubtful.

Thus the wavelet transform, one of whose most potent features is the ability to analyse the structure of inhomogeneous processes, shows the following. There exist local periodicities of El Niño events with scales in between 8 and 11 years. A special study is needed to clarify whether they are dominated by an 18.6-year moon cycle or the 22-year cycle of Sun's activity. As for scales less than 7 – 8 years, there is no sense to discuss them based on yearly data.

The wavelet analysis does not reveal a stable 70 – 80-year El Niño cycle: there are several long epochs of similar duration (about 40 years) whose periodicity breaks in the middle of series. The reasons for this are as yet obscure. They may be either real physical reasons or those due to the data compiling procedure: the data analysed have been reconstructed by the author of Ref. [21] from a huge volume of indirect evidence, and the absence of some data (for the

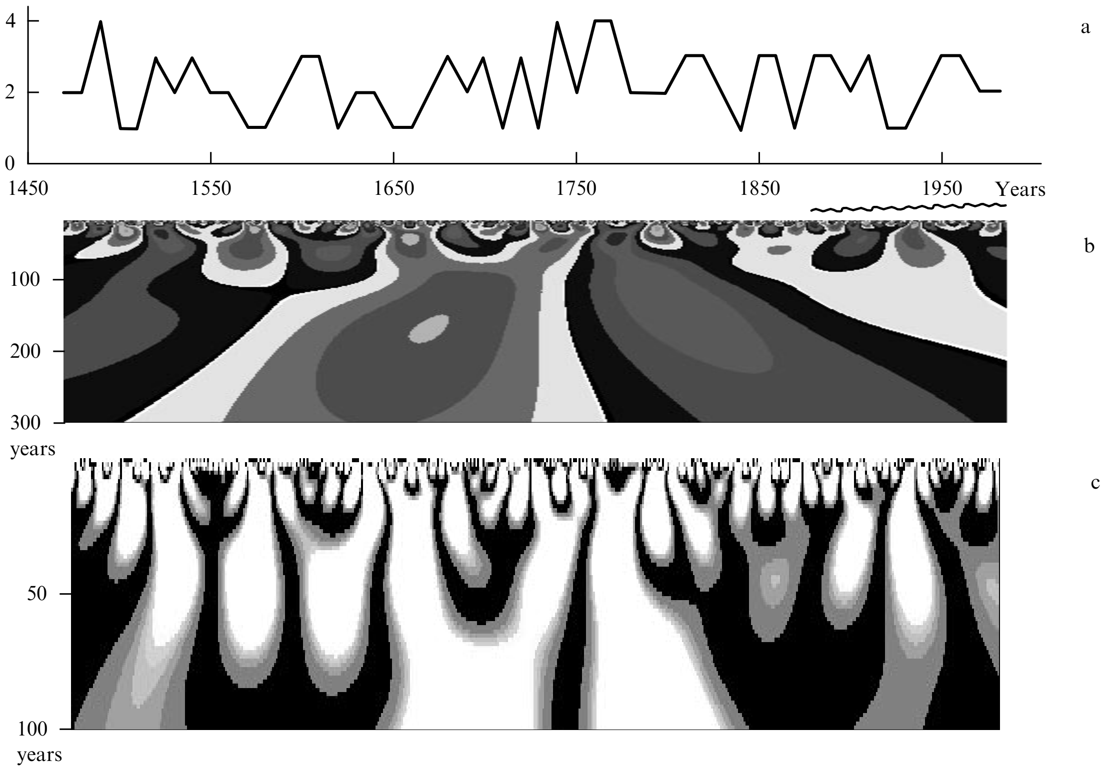

Figure 13. Dynamics of El Niño events for 500 years (a), patterns of coefficients (b) and energy density distribution (c).

period that corresponds to the middle of series) could lead to pattern distortion. There are some indications (for example, the preserved phase of oscillations) pointing to the existence of a stable 70–80-year cycle of El Niño events. More extended periods cannot be addressed because of finiteness of the series.

When interpreting the results one should remember that the data analysed are reconstructed by indirect evidence and aside from that, do not contain information on the duration and intensity of the process (they only 'resolve' whether an El Niño event occurred during a particular year) as well as on anti-El Niño and La Niña which are inseparable parts of the ENSO process. Much more information of incomparable higher reliability is contained in observational data on the Southern Oscillation index variations collected for the last century.

6.3 Monthly mean magnitudes of the Southern Oscillation index

Recall that the Southern Oscillation index $C(t)$ is the difference in pressures, normalised in some particular way, measured at sea level at stations located near the Southern Oscillation action centres. Conventionally it is computed in the following way. At input there are pressure series at Tahiti and Darwin $p_{\mathrm{y,m}}$ (the indices y, m denote year and month). They serve to compute series of normalised pressure at stations $P_{\mathrm{y,m}} = (p_{\mathrm{y,m}} - \langle p_{\mathrm{m}} \rangle)/\varepsilon$, here $\langle p_{\mathrm{m}} \rangle$ is the long-term mean (norm) calculated by monthly means for a period of 1951–1980; ε is the standard mean deviation calculated by all pressure anomaly data for the same period. Next, one computes normalised pressure anomalies between Tahiti (index T) and Darwin (index D) stations $\delta_{\mathrm{y,m}} = P^{\mathrm{T}}_{\mathrm{y,m}} - P^{\mathrm{D}}_{\mathrm{y,m}}$. Finally, the index of the Southern Oscillation is calculated: $C_{\mathrm{y,m}} = \delta_{\mathrm{y,m}}/\delta$ (with δ being the standard deviation of all differences $\delta_{\mathrm{y,m}}$ for the same reference period from 1951 to 1980). Negative values of this index are closely related to El Niño, and positive, to La Niña.

Figure 14 shows variations in monthly mean magnitudes of $C(t)$ for the last 108 years (Fig. 14a) and the pattern of wavelet transform coefficients (Fig. 14b) in a scale range chosen to enclose, where possible, all large-scale details of the process described by this finite series (the scale grows linearly to bottom, up to 97 years). Figure 14c shows the distribution of energy density for monthly mean magnitudes of $C(t)$ in a smaller-scale region, the scale varies up to 10.5 years.

Of notice are two large-scale minima almost at series boundaries and the branching 'tree' of positive extrema between them. The tree trunk and two of its branches, emerging just near the base, single out two practically the similar periods — between El Niños at tree margins and the trunk — of 39.8 years (which complies with a scale of 40 years found with 500-year series and testifies in favour of existence of 75–80 year cycle).

Energy density distribution (Fig. 14c) suggests that small-scale constituent of the process contains ranges of local periodicity and, besides, that the process exhibits qualitative differences in time intervals belonging and not belonging to the tree. The inner part of the tree largely contains details with characteristic scales of 30 and 12 months, whereas the outer parts are composed of details with a scale of 18 months.

The presence of ages with different temporal structure could be a plausible explanation to the change in character of El Niño scenarios after the middle of the 70s, as noted by many authors. In this respect it should be mentioned that when studying the ENSO process (and perhaps not only it) one should avoid drawing very stringent conclusions based on observations of the last few decades. An analysis of

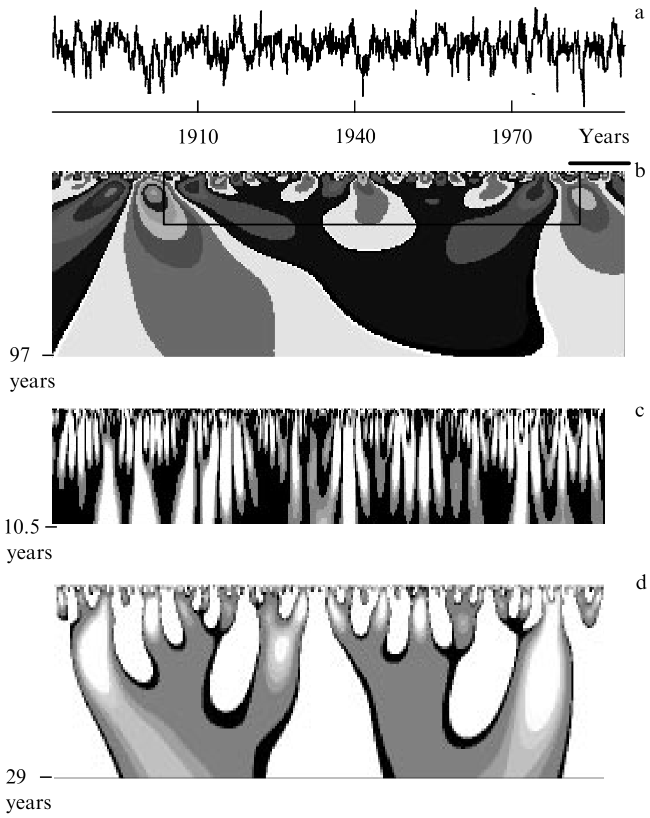

Figure 14. Monthly mean magnitudes of the Southern Oscillation index (a), pattern of coefficients (b), pattern of energy density distribution (c), and pattern of coefficients (d) for a fragment within the frame in plate (b).

considerably more representative data is needed (in our case, for example, a series for the last 30 years would be too short).

We however return to the structure of the coefficient pattern. Two large-scale minima at the tree boundaries and one between them correspond to intense and persistent El Niño events of 1899 – 1902, 1940 – 1941, and 1982 – 1983. By making comparison with 500-year data it can be easily inferred that these strongest El Niños of the current century are quite ordinary events. Despite the fact that 500-year data are averaged over 12 months and do not bear information on process intensity, the details of patterns match reasonably well. One may consider the El Niño chronology reconstructed for the past 500 years as being able to describe the dynamics of the Southern Oscillation fairly well for large time spans (of the order of several decades).

The branches of the tree are shown in Fig. 14d in more detail; here the scale grows up to 29 years. The patterns of coefficients (and local maximum lines) exhibit an intricate hierarchical structure. If, near the tree base, we see only the growth of scale in the form of period doubling, the two main branches of the tree bifurcate each in its own manner: the left one into three branches and the right one into two branches. At finer scales one may observe that alterations in doubling and tripling, seemingly irregular, persist. One may also encounter branching of a watershed type when several secondary local maximum lines progressively join the main one as scale increases.

Branching that looks similar can be modelled by the Cantor set when from a unit interval (in our case this corresponds to approximately 80 years) a middle part is eliminated, next the operation being applied to the parts remained, and so on (model of doubling the local maximum lines); or two fifths of the unit interval are eliminated, with the same procedure being applied in every generation (the model of tripling).

In our case the process is more intricate in two respects. First there are branchings of two (or even more) types alternating in an irregular manner — this is indicative of an irregular or multifractal set. Second, unlike the standard construction of the Cantor set in which that part of the interval which is eliminated never appears again in any generation, here we encounter the case when branching of both signs is seen, as if in an interval excluded some of its parts are recovered in subsequent generations. Whether the result of such a cascade process will be the Cantor dust, or any other fractal set, depends on the entire set of branching rules which can hardly be deduced with a finite data set available.

Thus the wavelet analysis of monthly mean values of the Southern Oscillation index reveals the self-similar structure of the data and the presence of a process resembling a cascade one on scales from a month to several decades (up to 70 – 80 years). Whether the resultant 'tree-like' structure of local maximum lines is indicative of the cascade of period doubling or tripling, quasiperiodic or any other behaviour exhibited by the system, and whether the branching of both signs in local extremum lines observed at small temporal scales implies that there is a cascade at scales on the orders of month or less, can be elucidated only by further investigations. This in turn will require the analysis of series with a refined resolution.

6.4 Daily magnitudes of the Southern Oscillation index

The series under analysis of daily mean magnitudes of the index $C(t)$ embraces a decade (unfortunately out of the tree described in Section 6.3). Figures 15a, b show the series $C(t)$ itself and the pattern of its wavelet transform coefficients (the scale grows up to 3 years).

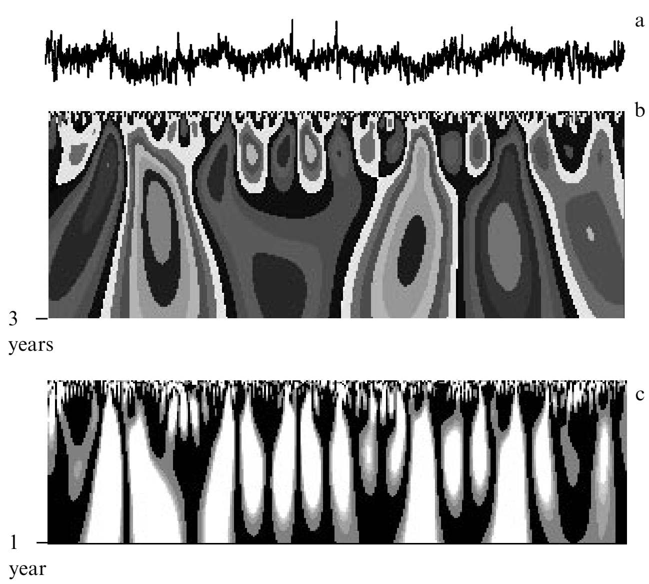

Figure 15. Daily magnitudes of the Southern Oscillation index for a period marked out with line in Fig. 14b (a), the pattern of coefficients (b) and energy density distribution (c).

The comparison shows that monthly mean magnitudes describe interannual variability very accurately, and the annual course, fairly well. The interannual variability is much better described by the detailed series, however most of typical features are seen in patterns of wavelet transform coefficients for both series.

The pattern of energy distribution calculated for the Southern Oscillation index (see Fig. 15c, the scale varies up to 1 year there) exhibits a clearly defined yearly course and an essentially nonstationary structure of the process at smaller scales. Rigorous analysis reveals ranges of local periodicity in this nonstationary structure with scales measured by a week (most maxima occur at scales ≈ 5 and ≈ 9 days), a month (≈ 25 days), 8 – 9 months, and about 2 years (≈ 22 months).

These scales as well as those described in Sections 6.2 and 6.3 can be discerned in power spectra shown in Fig.16. Here we present power spectra computed both by Fourier transform coefficients (Fig. 16a), E_{F}, and by wavelet transform coefficients (Fig. 16b), E_W, of monthly mean (dashed) and daily mean magnitudes of $C(t)$. The power spectrum E_W [see (28)] corresponds to the Fourier spectrum smoothed at each scale. Both spectra contain relatively long intervals of power-like behaviour across the scales from several days to a month and from two – three months to a year.

Thus the process described by the time series $C(t)$, as most processes encountered in the nature, evolves in a very broad range of time scales. Fourier spectra of the series investigated exhibit the presence of noise with fairly high amplitudes, however the peaks at particular typical periods (which coincide practically with those in power spectra obtained by wavelet transform coefficients) stand out clearly. The patterns of energy distribution $E_W(a, b)$ presented above show intricate nonstationary behaviour of the process, the presence of periodic and aperiodic constituents at different scales.

Such a process could be formed by superimposing a stochastic component on several regular ones. We take advantage of Takens' procedure (embedding theory for nonlinear dynamical systems [27]) to construct the phase space corresponding to a given time realisation and to create a feasible attractor.

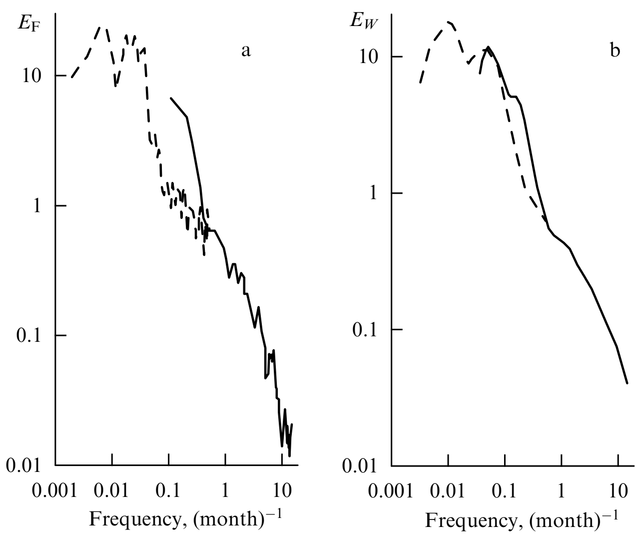

Figure 16. Power spectra (a) and scalograms (b) for mean monthly (dashed lines) and daily (solid lines) magnitudes of $C(t)$.

In compliance with Takens, we construct an m-component state vector from $C(t)$ in the following way:

$$X_i = \{x_1(t_i), x_2(t_i), \ldots, x_m(t_i)\}.$$

Here $x_k(t_i) = x(t_i + (k-1)\tau)$ and τ is the time lag. The distribution of state vectors composes the reconstructed phase space of dimensionality m.

Two-dimensional projections of the trajectories are presented in Fig. 17 for different values of the lag parameter τ. The first three portraits are based on the daily series, and the last two, on the series of monthly mean magnitudes. All trajectories are finite, but do not show any explicit periodic structure. Peculiar loops are seen in the first two portraits, yet they disappear with a lag of one month and reappear for $\tau = 3$ months. Except for that the portraits with lags of one and three months practically duplicate each other.

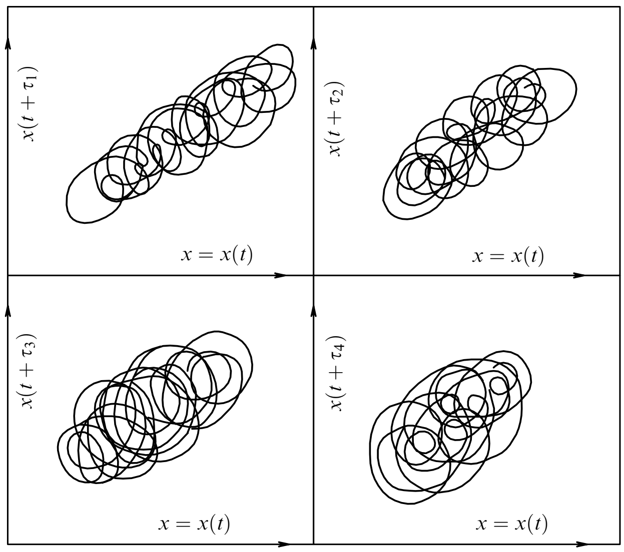

Figure 17. Projections of phase space obtained on the basis of series of daily and mean monthly magnitudes of $C(t)$: $\tau_1 = 1$, $\tau_2 = 3$, $\tau_3 = 30$ days; $\tau_4 = 3$ months.

A portrait with the time lag $\tau = 12$ months, i.e. constructed for a marginally long or even too long lag for the realisation available, differs from the preceding ones and is of interest in that its trajectory possesses three axes (one could imagine a dynamical system with three stationary points).

The boundedness of trajectories and not too large dimensionality of the system suggest that a model of the process could be created and its behaviour could be predicted, at least in principle, over not very long intervals. It is however apparent that the duration of the realisation available does not allow for rigorous conclusions to be made, while the form of the last projections hints at least at the need for analysing more representative data.

6.5 Global temperatures and the ENSO

Figure 18 presents the plot of annual mean values of air surface temperature anomaly for a period from 1854 to 1990 and the pattern of corresponding wavelet transform coefficients in a scale span that covers practically all large-scale details of the process described by the series in question (the scale grows linearly to 85 years).

It is noteworthy that a scale of 25 – 30 years separates the coefficients pattern into two noticeably different domains. In

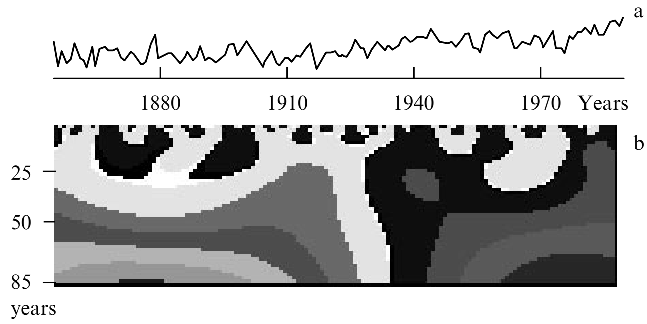

Figure 18. Anomalies of yearly mean global temperature and the pattern of wavelet transform coefficients.

the lower part of the pattern, only two large-scale details are seen. An entirely different structure is seen in the upper part where practically the overall temporal dynamics of the data analysed is comprised.

The lower part of the coefficient pattern reflects the fact that the data analysed contain a large-scale component resembling a positive trend — the light area to the left corresponds to negative values of coefficients which increase with time (to the right), and in the right lower part of the pattern almost a similar region (dark) is formed where coefficients are positive.

A much more involved and, at first sight, exhibiting no definite order structure is observed across scales of 25 – 30 years. A closer examination permits one to discriminate typical recurrent details: there exists a range of temporal scales within which the pattern has almost a quasiperiodic structure. One may argue that while the largest-scale component of the process analysed looks like a linear positive trend, at mesoscales the process resembles one composed of several harmonic oscillations.

Figure 19 with the same scales, and for the same time interval, presents the plot of annual mean values of air surface temperature anomaly for the Northern (Fig. 19a) and Southern (Fig. 19b) Hemispheres separately, and wavelet transform coefficient patterns for hemispheric data: Fig. 19c and 19d refer, respectively, to the Northern and Southern Hemispheres.

Qualitatively, the coefficient patterns of hemispheric data display a similar structure. The most apparent difference between the Hemispheres is in the time lag of the trend in the Southern Hemisphere: global warming begins earlier and is more pronounced in the Northern Hemisphere. The basic reason for that can be the inhomogeneous distribution of land which is mostly contained in the Northern Hemisphere. An anthropogenic factor, whose influence is most pronounced there, should also not be ruled out. Account should also be taken for the fact that meteorological stations are sparser over the Southern Hemisphere and the data are probably less representative there.

We note that the trend obtained by subtracting the series reconstructed by the inverse wavelet transform from the original data does not exhibit, at the beginning of the current century, a discontinuity of derivative usually attributed to anthropogenic impact.

Like the global temperature anomalies, the temporal dynamics of analysed hemispheric data is entirely contained across scales approximately up to 25 – 30 years.

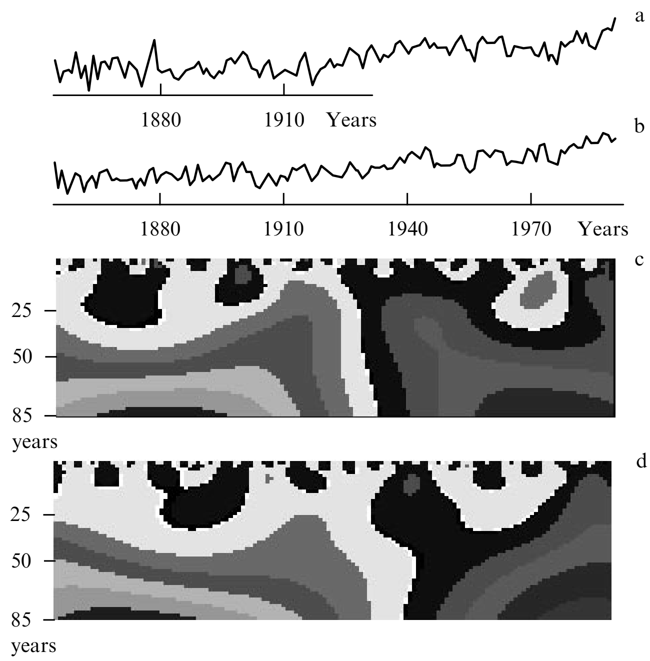

Figure 19. Anomalies of yearly mean global temperature for the Northern and Southern Hemispheres and respective patterns of wavelet transform coefficients.

Figure 20 shows fragments of hemispheric patterns with a better temporal resolution: the scale grows up to 30 years in Fig. 20a and to 8 years in Fig. 20b (time axis remains the same as in Figs. 18, 19). To facilitate the comparison of hemispheric data the patterns are arranged to match by side of small scales — in lower parts of both plates there are patterns for the Southern Hemisphere, while in the upper one, the mirror images for the Northern Hemisphere (correspondingly, the scale grows to the top of the pattern there).

Well pronounced cyclicity in periods of warming and cooling on scales of nearly 10 – 11 years (Fig. 20b) forms more extended epochs of warm and cold climate. The recurrent, relatively large-scale details in Fig. 20a bear witness to more or less persistent epochs (with duration between 10 and 30 years), with enhanced or reduced mean annual temperature, replacing each other during the whole period analysed without exhibiting any apparent regularity. The warming epochs can be seen in the 60s and 80s of the last

Figure 20. Fragments of the coefficient patterns for the Northern and Southern Hemispheres.

century, at the turn of the century, in the middle of this century, and at the end of the series.

Here, as in the results of analysis of ENSO characteristics, local periodicities stand out clearly. A relatively stable 10 – 11 year cycle is seen at the beginning of the series. Then its structure is subject to some changes, yet one may observe as a relatively stable 5 – 6 year cycle settles (this is especially seen in the pattern for the Northern Hemisphere). This period of time corresponds to the tree branches in the wavelet transform coefficient pattern for the Southern Oscillation index (Section 6.3). The larger-scale dynamics of mean annual temperature anomalies does not exhibit any stable periodicity.

A wide interest in the ENSO history is due to the fact that sometimes the occurrence rate of El Niño is associated with global climate warming [28]. The author of Ref. [21] however mentions that he has not discovered any close links between these processes.

In actual fact whether there is a direct link between the occurrence of El Niño events and global warming is difficult to answer, since, as one may suggest, not only the occurrence alone, but also the duration of events and their strength are of importance. As mentioned previously, changes in these characteristics of El Niño are well reflected in the behaviour of the Southern Oscillation index.

Figure 21 presents, in the same scales for handy comparison, patterns of wavelet transform coefficients for series of temperature anomaly in the Southern Hemisphere (it is in the upper part of the figure, the scale increases to the top) and the coefficient pattern for the Southern Oscillation index (it is in the lower part, scale increases to the bottom). The patterns are black-white; black areas correspond to positive values of coefficients (and are related to warm epochs in the pattern for temperature anomalies), white ones refer to negative values (negative values of $C(t)$ are associated with El Niño events).

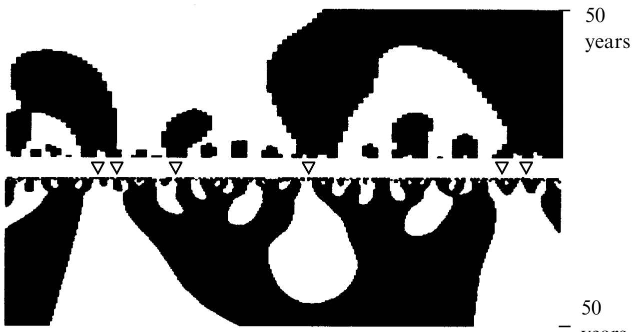

Figure 21. Fragments of wavelet transform coefficient patterns for monthly mean magnitudes of the Southern Oscillation index and yearly mean magnitudes of the Southern Hemispheric temperature anomalies.

Triangles mark only the most significant El Niños for that period which occurred in 1899 – 1900, 1902, 1913 – 1914, 1940 – 1941, 1970, and 1882–1983. It is easy to see that all more or less persistent warm periods are tied with the most intense El Niños (even the structure of black details in the upper pattern repeats that of white details in the lower one). However seemingly the occurrence of El Niño does not influence decisively the global warming.

If spoken generally, El Niño can influence the climate either directly or indirectly. A direct influence (through the atmosphere) occurs at time scales of 1 – 2 years — during El Niño itself. Indirect influence (through the ocean) as it turned out [29] can occur on scales of decades or even more.

Satellite observations indicate that waves of planetary scale, reflected from the coast of America during the extremely strong El Niño of 1982 – 1983, crossed the Pacific and ten years later perturbed the Kuroshio. As a result, a large volume of warm water penetrated from the southern coast of Japan into the middle latitudes of the Pacific thereby increasing noticeably the surface temperature at high latitudes of the north-west Pacific. Hence, the consequences of El Niño events assimilated by the ocean could appear to be long-lived.

7. Conclusions

Using simple examples of the utility of the wavelet transform in analysing one-dimensional functions with well-defined properties, we have shown the potentialities of this relatively new mathematical tool that enables one to reveal and explicitly expose the structure (quasiperiodic, self-similar, etc.) of the process described by a function under analysis, giving simultaneously information on typical scales of the process.

That the model examples were one-dimensional does not impose any restrictions on the applicability of the wavelet transform: the definitions and properties can be easily generalised to multidimensional cases, and dimensionless variable (time in our case) can be any scalar or vector quantity.

Time-scale sweep resulting from the wavelet transform of a signal reveals not only oscillations with well-expressed period, but also nonstationary oscillations, localised periodicities, etc.

The energy (or variance) of the wavelet transform coefficients $E_W(a)$ is proportional to the variance of data subject to the analysis and gives the distribution of process energy over scales. The availability of this characteristic at a local level allows one, for instance, when dealing with turbulent processes not only to retrieve a set of typical scales, but also to rigorously determine the scales related to coherent structures and explore the intermittency of the process.

Based on wavelet transform coefficients, or the behaviour of local extrema, one can calculate the dimension of the set analysed, or the spectrum of dimensions if it is multifractal.

Filtering and reconstruction properties of the transform allow information processing (smoothing, decomposition into components, convolution) without losing significant details. Breaks in continuity, jumps or other irregularities due to variations in characteristics measured, and faults or noise introduced by measuring instruments, can be easily detected, localised, and, if necessary, cut out or corrected.

The potentialities of wavelet transform have also been shown on examples of several observational time series.

The wavelet transform seems to be a very promising mathematical tool not only in tasks that involve signals of different nature but also for solving equations describing complex nonlinear processes with interactions in broad ranges of scales.

It should be emphasised that the wavelet transform can by no means replace the existing harmonic analysis or diminish its advantages. It is simply another tool which enables looking at a process from a somewhat different viewpoint — that of other analysing function (or the family of functions).

The author is indebted to the Russian Foundation for Basic Research for a partial support (grant No. 93-01-17342) and to Yu A Kravtsov for his interest in this work and valuable comments.

References

1. Grossmann A, Morlet J *SIAM J. Math. Anal.* **15** 723 (1984)
2. *Wavelets* (Eds J M Combes, A Grossmann, P Tchamitchian) (Berlin: Springer-Verlag, 1989)
3. *Wavelets and Their Applications* (Ed. R Coifman) (Boston: Jones and Barlett Publ., 1992)
4. *Wavelet Analysis and Its Applications* (Vol. 1: *An Introduction to Wavelets* Vol. 2: *Wavelets*: *A Tutorial in Theory and Applications*) (San Diego: Academ. Press Inc., 1992)
5. Daubechies I *Comm. Pure Appl. Math.* **41** 906 (1988); *IEEE Trans. Inform. Theory* **36** 961 (1990); *Ten Lectures on Wavelets* (CBMS Lecture Notes Series) (Philadelphia: SIAM, 1991)
6. Farge M *Ann. Rev. Fluid Mech.* **24** 395 (1992)
7. Pouligny B et al. *J. Appl. Cryst.* **24** 526 (1991); Muzy J F, Bacry E, Arneodo A *Phys. Rev. Lett.* **67** 3515 (1991)
8. Paul T *J. Math. Phys.* **25** (11) 3252 (1984)
9. Mallat S G *Trans. Am. Math. Soc.* **315** 69 (1989); *IEEE Trans. Patt. Anal. Mach. Int.* **31** 674 (1989)
10. Lemarie P G, Meyer Y *Rev. Math. Ibero-Americana* **2** 1 (1986)
11. Battle G *Commun. Math. Phys.* **110** 601 (1987)
12. Zastenker G N et al. *Kosmicheskie Issledovaniya* **20** 900 (1982) [*Cosmic Res.* **20** 640 (1982)]; Astaf'eva N M, Zastenker G N, Éĭges P E *Kosmicheskie Issledovaniya* **34** 407 (1996)
13. Beylkin G, Coifman R, Rokhlin V *Comm. Pure Appl. Math.* **44** 141 (1991)
14. Holschneider M *J. Stat. Phys.* **50** 963 (1988)
15. Arneodo A, Grasseau G, Holschneider M *Phys. Rev. Lett.* **61** 2281 (1988)
16. Collineau S, Brunet Y *Boundary-Layer Meteorology* **65** 357 (1993)
17. Astaf'eva N M, Preprint No. 1891 (Moscow: IKI RAN, 1994); Preprint No. 1946 (Moscow: IKI RAN, 1996)
18. Wyrtki K *J. Phys. Oceanogr.* **5** 572 (1975)
19. MacKenzie D *New Scientist* (16 April) 16 (1987)
20. Sidorenkov N S *Trudy Gidrometeotsentra SSSR* (316) 31 (1991)
21. Wang Shaowu *Acta Meteorologica Sinica* **6** (1) 47 (1992)
22. Astaf'eva N M, Sonechkin D M *Dokl. Russian Akad. Nauk* **344** (4) 539 (1995)
23. *JPCC, 1990*: *Climate Change*. *The JPCC Scientific Assessment* (Eds J T Houghton, F J Jenkins, J J Ephraums) (Cambridge: Cambridge University Press, 1990)
24. Currie R G *J. Geophys. Res.* **89** 7213 (1984)
25. Quinn W H, Neal V T, Antunez de Maylo S E *J. Geophys. Res.* **92** 14449 (1987)
26. Haudler R, Andsager K *Inter. J. Climato.* **10** 413 (1990)
27. Takens F *Lecture Notes in Mathematics* (Eds D Rand, L-S Young) (New York: Springer-Verlag, 1981) p. 366
28. Enfield D B, Cids L *TOGA Notes* (1) 1 (1990)
29. Jacobs G A et al. *Nature* (London) **370** 360 (1994)