



# КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

УДК 519.65

## МЕТОД МОДЕЛИРОВАНИЯ ДАННЫХ КРИТИЧЕСКОЙ ЧАСТОТЫ НА ОСНОВЕ СОВМЕЩЕНИЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И МОДЕЛЕЙ АВТОРЕГРЕССИИ ПРОИНТЕГРИРОВАННОГО СКОЛЬЗЯЩЕГО СРЕДНЕГО

**О. В. МАНДРИКОВА<sup>1</sup>**  
**Н. В. ГЛУШКОВА<sup>2</sup>**

*<sup>1)</sup> Институт космических исследований и распространения радиоволн ДВО РАН*

*<sup>2)</sup> Камчатский государственный технический университет*

*e-mail:*  
*oksanam1@mail.kamchatka.ru*

В статье предложен метод моделирования и анализа данных критической частоты ионосферы, основанный на совместном применении методов вейвлет-преобразования и моделей авторегрессии – проинтегрированного скользящего среднего. Метод позволяет проанализировать характерные особенности ионосферных данных и автоматически обнаружить аномальные изменения, возникающие в периоды повышенной солнечной или сейсмической активности. Апробация метода выполнена на данных критической частоты ионосферы foF<sub>2</sub>, полученных на обсерватории «Паратунка» (с. Паратунка, Камчатский край).

Ключевые слова: вейвлет-преобразование, модель авторегрессии-проинтегрированного скользящего среднего, параметры ионосферы, аномалии.

На фоне регулярных изменений, обусловленных суточным и сезонным ходом, в регистрируемых ионосферных параметрах наблюдаются аномальные особенности длительностью от несколько десятков минут до нескольких часов [1-3]. Эти аномалии имеют форму пиков различной амплитуды и возникают на фоне мощных ионосферных возмущений, обусловленных активностью Солнца, в сейсмоактивных областях возникающих накануне сильных сейсмических явлений [2, 3]. Сложная структура данных и наличие негладких локальных особенностей делает неэффективным для их выделения и анализа традиционные методы [2-4], которые описывают достаточно узкий класс процессов и позволяют выделить некоторые наиболее устойчивые характеристики структуры данных.

В настоящее время для анализа данных сложной структуры используются методы адаптивной фильтрации, основанные на применении экстраполирующих вейвлет-фильтров [2, 6, 7]. По сравнению с Фурье-анализом теория вейвлетов дает возможность рассмотреть многие явления на общем уровне и создать эффективный теоретический и технический аппарат в области обработки и анализа данных сложной структуры. Вейвлеты позволяют на основе небольшого числа параметров получить аппроксимации требуемой точности [7], и, что самое главное, ввиду большого разнообразия базисных функций, могут быть использованы для широкого спектра регистрируемых природных данных. Авторами данной работы на их основе совместно с моделями авторегрессии – про-



интегрированного скользящего среднего (АРПСС) разработан метод многокомпонентного моделирования и анализа ионосферных параметров, позволяющий выделить и проанализировать характерные особенности и аномалии, возникающие в периоды повышенной солнечной или сейсмической активности.

Для идентификации модели выполняется отображение данных в пространство вейвлетов на основе конструкции кратномасштабного анализа и их представление в виде сглаженной составляющей и детализирующих компонент [3, 6]. Сглаженная составляющая аппроксимируется моделью АРПСС, процедура оценки параметров этой составляющей модели подробно описана в работе [3]. Настоящая работа фактически является ее продолжением, где описан процесс идентификации модели для данных разных временных сезонов. Апробация модели выполнена на данных критической частоты ионосферы, регистрируемых на Камчатке (станция «Паратунка», с. Паратунка, Камчатский край). В процессе анализа статистики в данных выявлены аномальные особенности, возникающие в периоды повышенной солнечной активности или обусловленные процессами в литосфере, формирующиеся накануне сильных сейсмических явлений (анализировались события энергетического класса с  $k \geq 12,5$  в радиусе  $R \sim 300$  км от П.-Камчатского).

### Многокомпонентное моделирование временного ряда

Рассмотрим в качестве базового пространства регистрируемых дискретных данных  $f_0(t)$  замкнутое пространство с разрешением  $j = 0: V_0 = \text{clos}_{L^2(R)}(2^0 \phi(2^0 t - k)) : k \in Z$ , порожденное скэйлинг-функцией  $\phi \in L^2(R)$  [7]. На основе кратномасштабных разложений до уровня  $m$  получим следующее представление данных [3]

$$f_0(t) = \sum_{j=-1}^{-m} g[2^j t] + f[2^{-m} t], \text{ где } g[2^j t] \in W_j, f[2^{-m} t] \in V_{-m}, \quad (2)$$

$W_j$  - пространство с разрешением  $j$ , порожденное вейвлет-базисом  $\Psi_{j,n}(t) = 2^{j/2} \Psi(2^j t - n)$ ; компоненты  $g[2^j t] = \sum_n d_{j,n} \Psi_{j,n}(t)$ , где  $d_{j,n} = \langle f, \Psi_{j,n} \rangle$ , являются детализирующими компонентами, характеризуют локальные свойства данных; компонента  $f[2^{-m} t] = \sum_k c_{-m,k} \phi_{-m,k}(t)$ ,  $c_{-m,k} = \langle f, \phi_{-m,k} \rangle$  является аппроксимирующей составляющей.

Для идентификации модели используется итеративный подход, состоящий из следующих операций:

Отображение данных в вейвлет-пространство до уровня  $m$  на основе операции (2);

Анализ стохастических свойств полученной аппроксимирующей составляющей  $f[2^{-m} t]$  уровня  $m$ . В случае выполнения для нее условия строгой стационарности, выбор модели из класса моделей АРПСС [5] для ее аппроксимации;

Получение предварительных оценок параметров выбранной модели АРПСС;

Выполнение диагностических проверок полученной модели АРПСС. Если погрешность удовлетворяет требованию, модель готова к использованию и переход на шаг 5, в противном случае изменение уровня  $m$  вейвлет-разложения и возврат на шаг 1;

Анализ детализирующих компонент разложения  $g[2^j t]$ ,  $j = \overline{-1, m}$  [3, 6] и выделение в данных локальных особенностей.

Объединение полученных представлений в общую многокомпонентную конструкцию.

Полученная многокомпонентная модель будет иметь следующий вид:



$$f(t) = f_{1,j}(t) + f_2(t) = \sum_k s_k^j(t) \phi_{j,k}(t) + \sum_j g[2^j t], \quad (3)$$

где  $f_{1,j}(t) = \sum_k s_k^j(t) \phi_{j,k}(t)$  – параметрическая составляющая модели разрешения

$j$ ,  $s_k^j(t) = \sum_{l=1}^{p_j} \gamma_l^j \omega_{k-l}^j(t) - \sum_{n=1}^{h_j} \theta_n^j a_{k-n}^j(t)$ ,  $\gamma_l^j$  – параметры авторегрессии компоненты разложения разрешения  $j$ ,  $\omega_k^j(t) = \nabla^d c_{j,k}(t)$ ,  $p_j$  – порядок авторегрессионной модели компоненты разложения разрешения  $j$ ,  $\nabla^d$  – оператор взятия разности порядка  $d$ ,  $\theta_k^j$  – параметры скользящего среднего компоненты разрешения  $j$ ,  $h_j$  – порядок модели скользящего среднего компоненты разрешения  $j$ ,  $a_k^j$  – ошибки модели компоненты разрешения  $j$ ,  $f_2(t) = \sum_j g[2^j t]$  – непараметрическая составляющая модели, несущая информацию о локальных особенностях и представленная комбинацией вейвлет-коэффициентов.

В процессе идентификации модели мы изучаем структуру данных и аппроксимируем ее. Если данные содержат аномалию, то предполагаем, что произойдет изменение их структуры. Выделение аномалий в этом случае может быть построено на обработке остаточных ошибок модели  $a_k^j$ , полученных на основе прогнозирования: если сумма квадратов ошибок прогноза, рассчитанных в наперед заданном скользящем окне, существенно возрастает, то это говорит об изменении структуры входных данных вследствие наличия в них аномалии. Процедура обнаружения аномалии в компоненте модели  $f_{1,j}(t) = \sum_k s_k^j(t) \phi_{j,k}(t)$  может быть построена следующим образом:

1. Прогнозирование значения  $s_{k+q}^j$ ,  $q \geq 1$  определяет прогноз  $s_k^j$  в момент  $t = k$  с упреждением  $q$ . Значение  $s_{k+q}^j$  на основе полученной модели (3) определяется следующим образом:

$$s_{k+q}^j(t) = \sum_{l=1}^{p_j} \gamma_l^j \omega_{k+q-l}^j(t) - \sum_{n=1}^{h_j} \theta_n^j a_{k+q-n}^j(t).$$

2. Остаточная ошибка модели на уровне разложения  $j$  определяется как разность между прогнозным и фактическим значениями данных в момент времени  $t = k + q$ :  $a_{k+q}^j(t) = s_{k+q, прогноз}^j(t) - s_{k+q, фактич}^j(t)$ .

3. Факт обнаружения аномалии можно определить на основе соотношения:  $A_N = \frac{1}{N} \sum_{q=1}^N (a_{k+q}^j(t))^2$ , где  $N$  – длина окна наблюдения.

В случае выполнения условия:

$$A_N > P, \quad (4)$$

где  $P$  – некоторое наперед заданное пороговое значение, будем считать, что пределах окна наблюдения имеем аномалию.

### Идентификация многокомпонентной модели ионосферных данных

В процессе исследований были обработаны данные критической частоты ионосферного слоя F2 за период 1969 – 2006 гг., регистрацию данных ведет Институт космических исследований и распространения радиоволн ДВО РАН (Камчатский край, с. Паратунка). Вариации критической частоты содержат пропуски, что существенно затрудняет процесс построения моделей. С целью уменьшения погрешности получаемых результатов были выбраны временные периоды с наименьшим количеством пропусков.

Учитывая сезонный характер ионосферного процесса, данные предварительно были разделены на сезоны. В качестве базисных функций использовался класс ортогональных вейвлетов Добеши: db2, db3, db4.

На основе операции (2) было выполнено отображение данных в вейвлет-пространство и получены компоненты-разложения до 3-го уровня включительно. Попытка идентификации моделей АРПСС была выполнена для исходных данных и аппроксимирующих компонент 1-го, 2-го и 3-го уровней разложения. Диагностические проверки полученных моделей АРПСС за разные временные периоды показали, что структура исходных временных рядов и компонент 1-го, а также, для некоторых лет, 2-го уровней разложения является сложной, и они не могут быть аппроксимированы моделью АРПСС (наблюдалась существенная автокорреляция остатков). Наилучшие результаты при диагностике моделей АРПСС были получены для аппроксимирующих компонент 3-го уровня разложения. Анализ остаточных ошибок моделей показал, что наилучшую аппроксимацию данных позволяет получить базисная функция db3.

Результаты моделирования данных зимнего периода времени показали, что все модели АРПСС компонент 3-го уровня разложения имеют второй порядок и близкие значения параметров, что позволило определить общую модель, имеющую вид:

$$f(t) = \sum_k s_k^{j=3}(t) \phi_{j=3,k}(t) + \sum_j g[2^j t], s_k^{j=3}(t) = (1 + 0.9B)^2 (1 - B) \omega_k^3(t) + a_k(t), \quad (4)$$

где  $B^l \omega_k(t) = \omega_{k-l}(t)$ ,  $a_k(t)$  - остаточные ошибки модели.

На рис.1, в качестве примера, показаны результаты моделирования аппроксимирующей компоненты 3-го уровня разложения за период 18.01.1970-02.02.1970 при использовании общих (на основе модели (4)) (рис. 1а) и фактических (полученных при идентификации) (рис. 1б) параметров (пунктир – исходные данные, сплошная линия – значения прогноза).

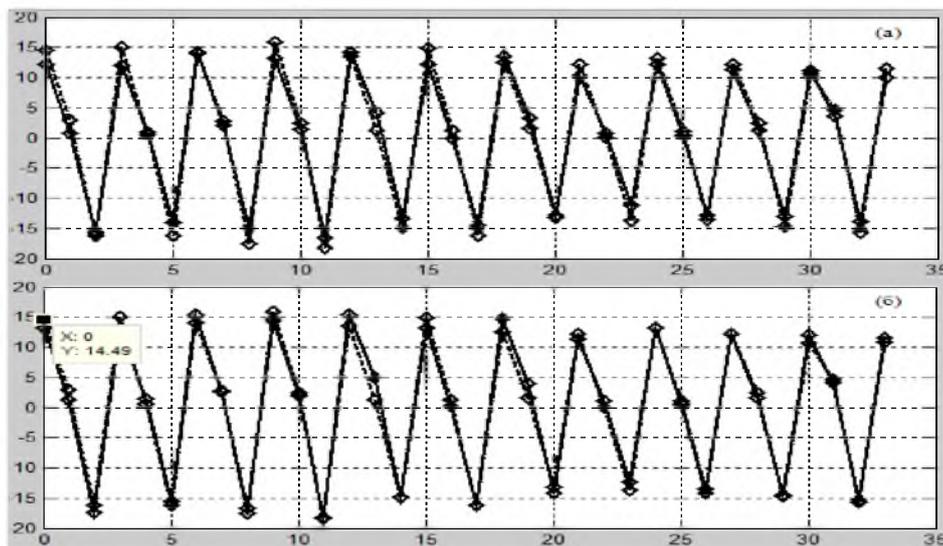


Рис. 1. Результаты моделирования ионосферных данных за период 18.01.1970-02.02.1970: (а) – при использовании общих параметров; (б) – при использовании фактических параметров (пунктир – исходные данные, сплошная линия – значения прогноза)

В процессе анализа остаточных ошибок результатов моделирования данных зимних периодов времени были выделены аномалии. Их сопоставление с данными индекса геомагнитной активности К, характеризующего солнечную активность, и с данными каталога землетрясений показало, что они возникают в периоды повышенной солнечной

активности (рис.2), а также могут наблюдаться в периоды сильных землетрясений на Камчатке (рис.3).

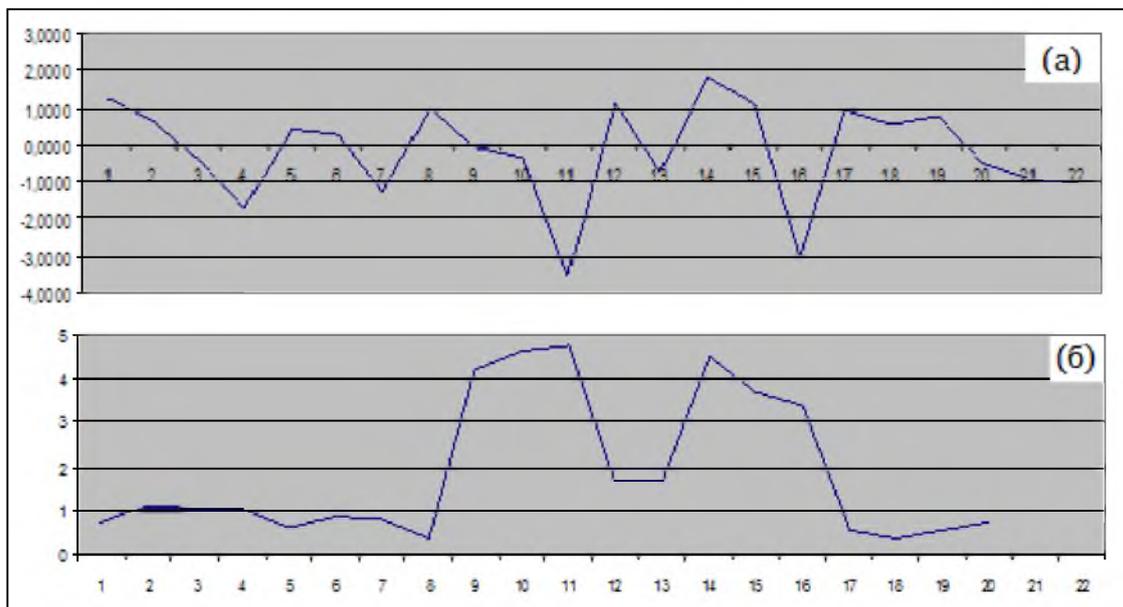


Рис. 2. Результаты обработки данных за период 15.02.1987-26.02.1987. а) ряд остаточных ошибок модели  $a_{k+q}^j(t), q = 1, j = 3$ ; б) результаты расчета величины  $A_N, N = 3$

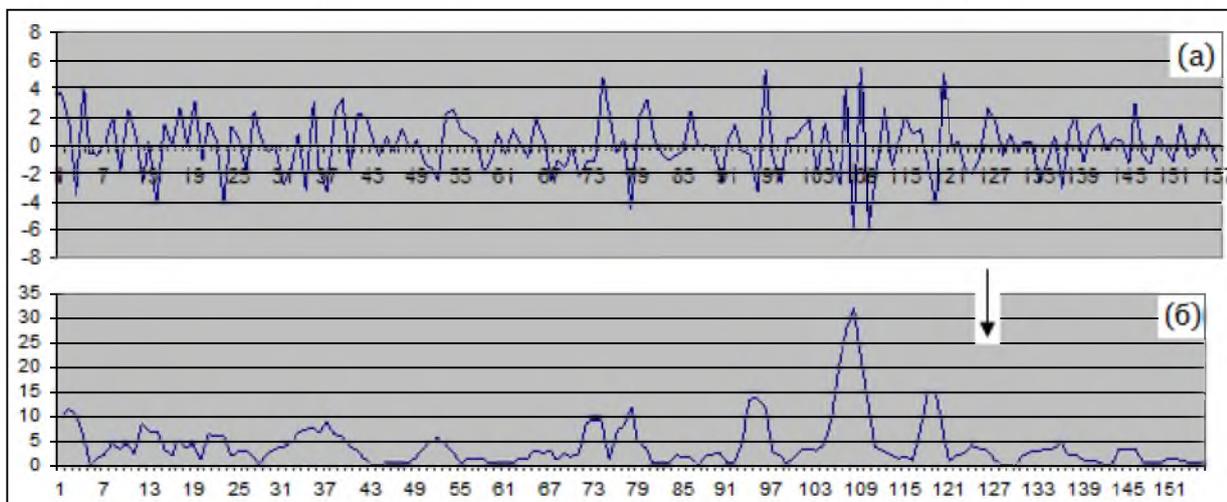


Рис. 3. Результаты обработки данных за период 01.01.2006-28.02.2006. а) ряд остаточных ошибок модели  $a_{k+q}^j(t), q = 1, j = 3$ ; б) результаты расчета величины  $A_N, N = 3$ . Стрелкой показан момент возникновения сильного землетрясения (дата, энергетический класс  $k = 13,8$ )

Результаты моделирования данных летних периодов времени показали, что все АРПСС модели компонент 3-го уровня разложения имеют третий порядок. В таблице 1 представлены параметры моделей этих компонент. Особо следует отметить момент яркого проявления в результатах моделирования солнечного одиннадцатилетнего цикла (1970, 1982 гг., отмечены в таблице 1 серым цветом). Модели этих лет имеют наиболее близкие значения параметров.

Анализ результатов обработки данных летних периодов времени показал, что в периоды повышенной сейсмической активности (1970, 1972, 1982 гг.) параметры моделей отличаются от параметров моделей в сейсмически спокойные периоды времени (1969,

1979 гг.). На рис.4, в качестве примера, показан исходный временной ряд за период 08.08.1970-31.08.1970 и результаты его моделирования, стрелками отмечены моменты возникновения сильных землетрясений, произошедших за анализируемый период на Камчатке (22 августа и 30 августа, 1970 г). Анализ результатов расчета величины  $A_N$ , характеризующей наличие в данных аномалии, показывает, что за несколько суток до землетрясения в данных наблюдается аномальная особенность длительностью 32 часа (один отсчет соответствует периоду времени 8 часов).

Таблица 1

**Параметры моделей данных fo F2-летнего периода времени**

Анализируемый временной интервал	Значение первого параметра	Значение второго параметра	Значение третьего параметра
06.05.1969-25.05.1969	-0.4095	-0.4951	0.2449
08.08.1970-31.08.1970	-0.3975	-0.3852	0.5787
26.06.1972-22.07.1972	-0.2267	-0.3485	0.2228
04.05.1979-23.05.1979	-0.652	-0.692	0.1547
01.05.1982-30.05.1982	-0.3909	-0.3449	0.4816

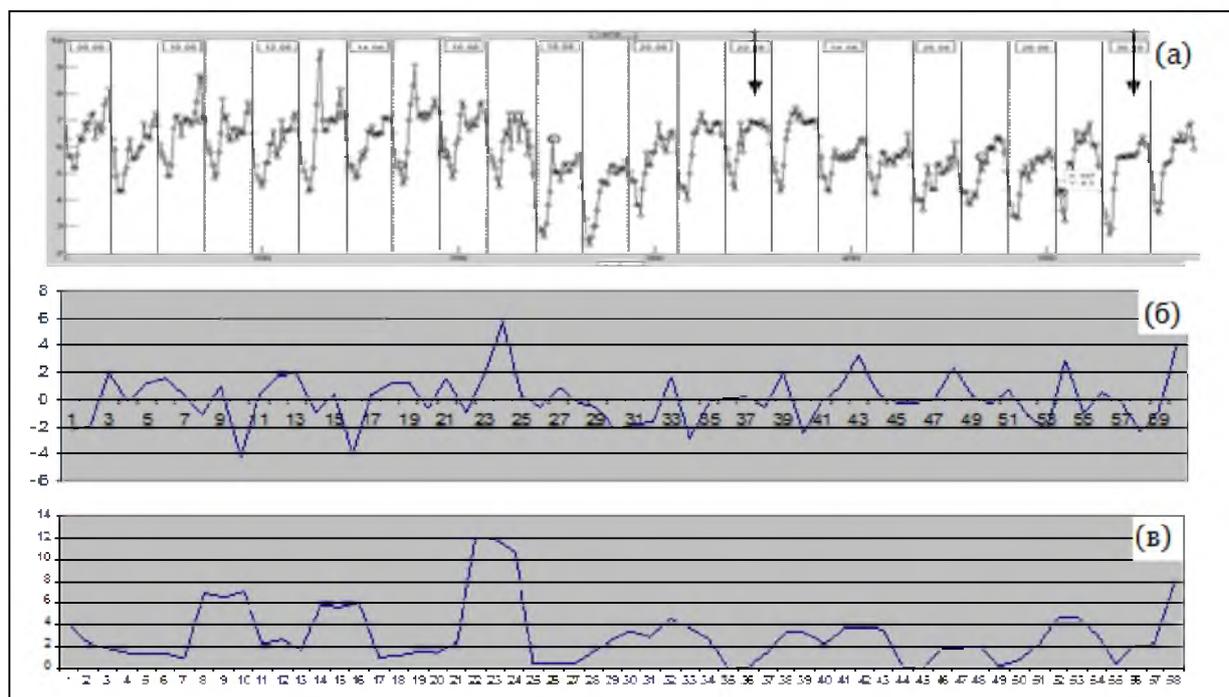


Рис. 4. (а) – исходный временной ряд, период регистрации 08.08.1970-31.08.1970; (б) – ряд остаточных ошибок модели  $a_{k+q}^j(t), q = 1, j = 3$ ; (в) – результаты расчета величины  $A_N, N = 3$ .  
Стрелкой показаны моменты возникновения сильных землетрясений на Камчатке (дата, энергетический класс  $k = 13,8$ )

**Выводы**

В работе предложен метод моделирования данных критической частоты ионосферы, в основе которого лежит совмещение методов авторегрессии- проинтегрированного скользящего среднего и вейвлет-преобразования. Метод позволяет проанализировать



характерный ход процесса и выделить аномальные особенности, связанные с повышенной солнечной или сейсмической активностью.

Построены модели данных критической частоты для зимних и летних периодов времени. В процессе моделирования обнаружены аномалии, возникающие в периоды повышенной солнечной активности и накануне сильных землетрясений на Камчатке.

Работа поддержана грантом Президента Российской Федерации МД-2199.2011.9, грантом РФФИ – ДВО РАН №11-07-98514-р\_восток\_a и грантом «У.М.Н.И.К.»-2011 г.

### Литература

1. Дёмин М.Г. Ионосфера Земли. Плазменная гелиогеофизика. – М.: Физматлит, 2008. Т. II. – С.92-163.
2. Богданов В.В., Геппенер В.В., Мандрикова О.В. Моделирование нестационарных временных рядов геофизических параметров со сложной структурой. – С.-Петербург: ЛЭТИ. 2006. – 107 с.
3. Мандрикова О.В., Глушкова Н.В. Многокомпонентное моделирование и анализ аппроксимирующих компонент критической частоты foF2 на основе вейвлет-преобразования и моделей авторегрессии // Междунар. конф. по мягким вычислениям и измерениям (SCM'2011): Санкт-Петербург: Издательство СПбГЭТУ "ЛЭТИ", СПб, 2011, т. 2. – С. 139-143.
4. Siegmund Brandt. Data Analysis. Statistical and Computational Methods for Scientists and Engineers. Пер. с англ. – М.: Мир, ООО «Издательство АСТ», 2003. – 668 с.
5. Марпл.-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ. – М.: Мир, 1990. – 265 с.
6. Мандрикова О.В., Горева Т.С. Метод идентификации структурных компонентов сложного природного сигнала на основе вейвлет-пакетов // Научно-технический журнал "Цифровая обработка сигналов", 2010. № 1. – С.45-50.
7. Stephane Mallat, A Wavelet tour of signal processing. Пер. с английского, – М.: Мир, 2005.

## A METHOD OF MODELING DATA OF THE CRITICAL FREQUENCY BASED ON COMBINATION OF THE WAVELET TRANSFORM AND AUTOREGRESSIVE MODELS

**O.V. MANDRIKOVA<sup>1</sup>**  
**N.V. GLUSHKOVA<sup>2</sup>**

<sup>1</sup>*Institute of Cosmophysical  
Researches and Radio Wave  
Propagation*

<sup>2</sup>*Kamchatka State Technical  
University*

*e-mail:*  
*oksanam1@mail.kamchatka.ru*

This article proposes a method of modeling and analysis data of the critical frequency of the ionosphere, based on the joint application of the wavelet transform and autoregressive models – integrated over a moving average. The method allows to analyze the characteristics of ionospheric data and automatically detect abnormal changes that occur during periods of increased solar and seismic activity. Testing of the method is executed on the data of the ionosphere critical frequency foF2, obtained at the Observatory "Paratunka" (p. Paratunka, Kamchatka).

Key words: wavelet transform, autoregressive model, parameters of the ionosphere, anomalies.