

ОБНАРУЖЕНИЕ И АНАЛИЗ АНОМАЛИЙ В ДАННЫХ КРИТИЧЕСКОЙ ЧАСТОТЫ ИОНОСФЕРЫ НА ОСНОВЕ СОВМЕЩЕНИЯ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И АВТОРЕГРЕССИОННЫХ МОДЕЛЕЙ

Н.В. ГЛУШКОВА¹
О.В. МАНДРИКОВА²

¹Институт космических исследований и распространения радиоволн Дальневосточное отделение РАН, г. Петропавловск-Камчатский
²Камчатский государственный технический университет, г. Петропавловск-Камчатский

e-mail:
nv.glushkova@yandex.ru,
oksanam1@mail.kamchatka.ru

Описан метод многокомпонентного моделирования ионосферных данных, основанный на совмещении вейвлет-преобразования с авторегрессионными моделями, позволяющий выполнить анализ данных и выделение аномальных особенностей. Апробация метода и основанных на нем алгоритмов моделирования проводилась на данных критической частоты foF2, записанных на станции «Паратунка» (п-ов Камчатка). При моделировании данных выделены особенности, возникающие в периоды повышенной сейсмической активности на Камчатке.

Ключевые слова: вейвлет-преобразование, модель авторегрессии-проинтегрированного скользящего среднего, критическая частота, аномалии.

Введение. Предметом исследований являются данные критической частоты F2-слоя ионосферы. Одной из важных задач обработки и анализа ионосферных данных является задача контроля состояния ионосферы и автоматическое выделение и интерпретация аномалий, возникающих в периоды ионосферных возмущений [1]. Сложная структура регистрируемых параметров делает поставленную задачу весьма сложной. Данные критической частоты ионосферы содержат суточный, сезонный ход и локальные особенности, возникающие в периоды ионосферных возмущений и несущие полезную информацию о процессе, что делает неэффективным для их анализа традиционные методы [2]. На основе совместного применения вейвлет-преобразования с авторегрессионными моделями в работе предложен метод многокомпонентного моделирования ионосферных данных и обнаружения аномальных особенностей. Построены модели критической частоты ионосферы F2-слоя по данным Камчатки. В работе выполнено моделирование, анализ данных и выделены аномалии, возникающие в периоды ионосферных возмущений. Анализ аномалий показал, что они наблюдаются в периоды повышенной сейсмической активности на Камчатке.

Описание метода. Рассмотрим замкнутое пространство $V_j = \text{clos}_{L^2(R)}(2^j \phi(2^j t - k)) : k \in Z$ с разрешением $j = 0$, порожденное скэйлинг-функцией $\phi \in L^2(R)$ [3], где $L^2(R)$ – пространство Лебега. На основе кратномасштабного анализа (КМА) до уровня m получим представление данных в виде [4]:

$$f_0(t) = \sum_{j=-1}^{-m} (g[2^j t] + e[2^j t]) + f[2^{-m} t], \quad (1)$$

где $\left[\begin{matrix} \cdot \\ \cdot \\ \cdot \end{matrix} \right]_{-m} \left[\begin{matrix} \cdot \\ \cdot \\ \cdot \end{matrix} \right]_j$ – пространство с разрешением j , порожденное $f[2^{-m} t] \in V$, $g[2^j t] \in W$

вейвлет-базисом $\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k)$; $g[2^j t] = \sum_k d_{j,k} \Psi_{j,k}(t)$ – детализирующие компонен-

ты, где коэффициенты разложения $d_{j,k} = \langle f, \Psi_{j,k} \rangle$; $\left[\begin{matrix} \cdot \\ \cdot \\ \cdot \end{matrix} \right]_{-m,k}$ – аппроксимиру-

$$f[2^{-m} t] = \sum_k c_{-m,k} \phi_{-m,k}(t)$$

ющая компонента, где коэффициенты разложения $c_{-m,k} = \langle f, \phi_{-m,k} \rangle$; $e[2^j t]$ – шумовые составляющие (предполагается, что шум белый), разрешение j (соответствует масштабу $(-j)$).

Следуя результатам работы [5] подавление шума выполним на основе применения пороговой функции

$$P_T(x) = \begin{cases} x, & \text{если } |x| \geq T \\ 0, & \text{если } |x| < T \end{cases} \quad (2)$$

для каждой детализирующей компоненты $g[2^j t] \in (1)$. Порог $T = \sigma^2$, где σ^2 – дисперсия шума. Дисперсию шума σ^2 можно оценить на основе соотношения [3]: $\sigma^2 \approx \text{Med} \left\langle \left\langle f, \Psi_{j,k} \right\rangle \right\rangle_{0 \leq k < N}$, где Med – медиана, $j = -1$, N – длина компоненты.

Для идентификации модели применим следующие операции:

На основе КМА получим представление данных в виде (1) и для каждой детализирующей компоненты выполним подавление шума.

Восстановим каждую из полученных компонент $f[2^{-m} t]$ и $g[2^j t]$, $j = \overline{-1, m}$.

Используя традиционный подход [3], определим модели из класса моделей АРПСС для аппроксимации каждой из полученных восстановленных компонент;

Выполним диагностические проверки полученных моделей. Если погрешность модели компоненты удовлетворяет требованию, то будем считать, что данная компонента описывает характерные особенности структуры данных;

Объединим полученные представления в общую многокомпонентную конструкцию:

$$f(t) = \sum_{\mu=1, M} \sum_{k=1, N_j^\mu} s_{j,k}^\mu(t) b_{j,k}^\mu(t), \quad (3)$$

где $s_{j,k}^\mu(t) = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k-l}^\mu(t) - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k-n}^\mu(t)$, $\gamma_{j,l}^\mu$ – параметры авторегрессии компоненты с номером μ , $\omega_{j,k}^\mu(t) = \nabla^{v_j} \beta_{j,k}^\mu$, $\beta_{j,k}^\mu$ – коэффициенты разложения компоненты с номером μ , p_j^μ – порядок авторегрессионной модели компоненты с номером μ , ∇^{v_j} – оператор взятия разности порядка v_j , $\theta_{j,n}^\mu$ – параметры скользящего среднего модели компоненты с номером μ , h_j^μ – порядок модели скользящего среднего компоненты с номером μ , $a_{j,k}^\mu$ – ошибки модели компоненты с номером μ , M – количество моделируемых компонент, N_j^μ – длина компоненты с номером μ , $b_{j,k}^\mu$ – базис компоненты с номером μ , j – разрешение.

Процедура выделения аномалий может быть построена на обработке остаточных ошибок полученных моделей компонент при выполнении операции прогнозирования:

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

$$s_{j,k+q}^\mu(t) = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k+q-l}^\mu(t) - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k+q-n}^\mu(t)$$

2. Остаточные ошибки компоненты модели с номером μ масштаба $(-j)$ определяются как разность между прогнозными и фактическими значениями данных в момент времени $t = k + q$: $a_{j,k+q}^\mu(t) = s_{j,k+q, \text{прогноз}}^\mu(t) - s_{j,k+q, \text{фактич}}^\mu(t)$.

3. Обнаружение аномалии в компоненте с номером μ масштаба $(-j)$ можно выполнить на основе проверки условия:

$$D_{U_j} = \frac{1}{U_j} \sum_{q=1}^{U_j} (a_{j,k+q}^\mu(t))^2 > T_{A_j}, \quad (4)$$

где T_{A_j} – некоторое наперед заданное пороговое значение, определяющее наличие в данных аномалии масштаба $(-j)$, U_j – длина окна наблюдения на масштабе $(-j)$.

Моделирование данных критической частоты ионосферы F2-слоя. В экспериментах использовались часовые данные fOF2 за период 1968-2011гг, регистрируемые на Камчатке. В качестве базисных функций использовались ортогональные вейвлеты Добеши порядка 3. На основе КМА было получено представление данных в виде (1). Уровень разложения $m = 3$ определялся статистически и основывался на результатах работы [4]. Далее, на основе операции (2), были подавлены шумовые составляющие $e^{[2^j t]}$ (рис. 1). Результаты оценки дисперсии шума представлены в табл. 1. Анализ результатов табл. 1 показывает, что уровень шума носит случайный характер.

Таблица 1

Результаты оценки дисперсии шума в данных fOF2

Анализируемый период	23.01.1981 06.02.1981	16.01.2002 25.02.2002	15.01.2006 04.02.2006	10.12.2010 31.12.2010	08.02.2011 27.02.2011
Дисперсия шума	0.3152	0.2068	0.1768	0.1716	0.1654

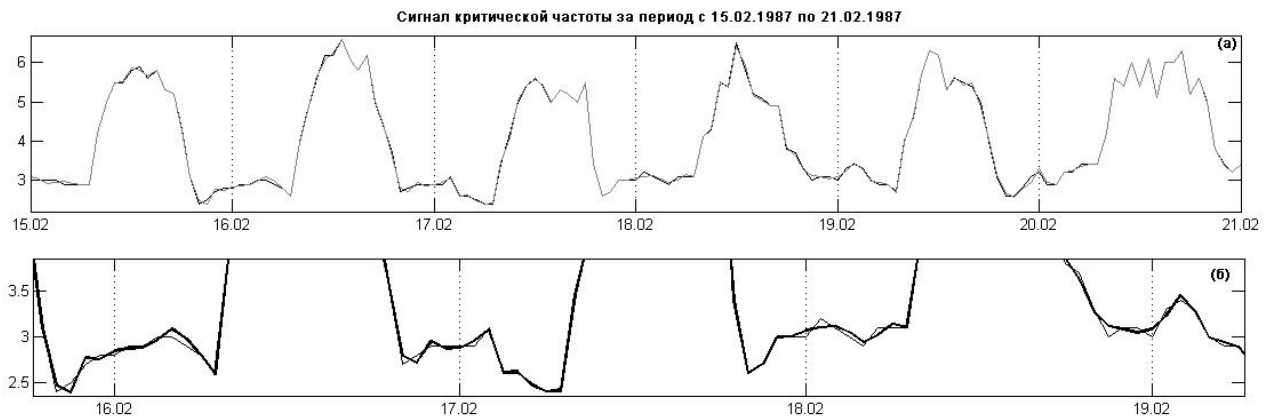


Рис. 1. Результаты подавления шумовых составляющих в данных foF2:

(а) – масштаб 100 % (черным цветом – исходный сигнал, серым – сигнал без шума),
(б) – масштаб 230 % (тонкая линия – исходный сигнал, толстая линия – сигнал без шума)

В табл. 2 показаны параметры моделей восстановленных компонент, полученные на основе описанных выше операций для данных зимнего периода времени.

Таблица 2

Параметры моделей восстановленных компонент

Анализируемый период	Восстановленная аппроксимирующая компонента		Восстановленная детализирующая компонента 3 уровня разложения		Восстановленная детализирующая компонента 2 уровня разложения	
	первый параметр	второй параметр	первый параметр	второй параметр	первый параметр	второй параметр
04.01.70-05.02.70	1.01	-0.27	0.82	-0.34	0.38	-0.61
07.02.83-23.02.83	1.01	-0.27	0.83	-0.34	0.33	-0.68
01.01.91-26.01.91	1.01	-0.27	0.81	-0.35	0.31	-0.69
01.12.00-22.12.00	1.01	-0.27	0.83	-0.34	0.38	-0.61
21.12.03-03.02.04	1.01	-0.27	0.82	-0.34	0.39	-0.60
08.02.11-27.02.11	1.01	-0.27	0.81	-0.35	0.44	-0.47

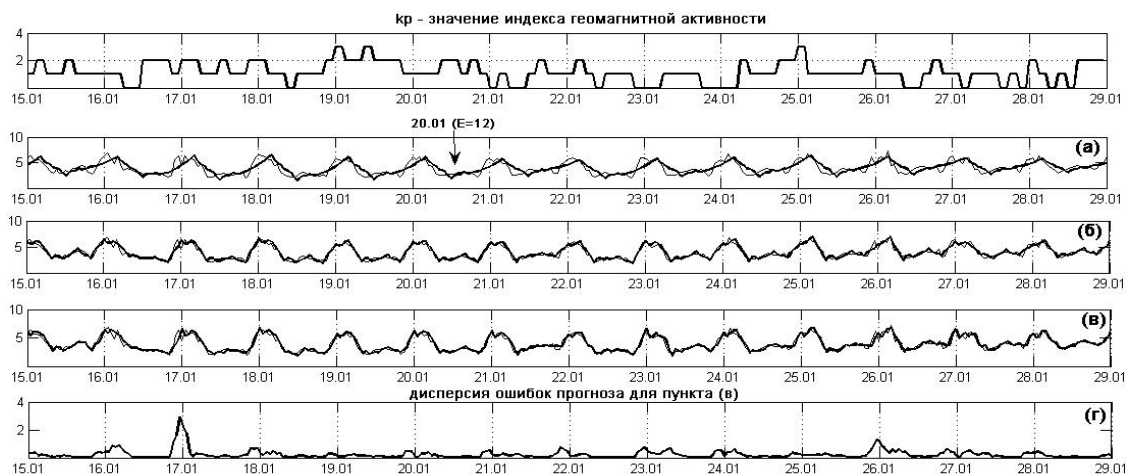


Рис. 2. Результаты моделирования ионосферных данных за период 15.01.2011-29.01.2011. Сигнал foF2 (тонкая линия), прогнозные значения восстановленных компонент (толстая линия). (а) – сигнал foF2, аппроксимирующая компонента; (б) – сигнал foF2, совмещение аппроксимирующей компоненты и детализирующей компоненты 3-го уровня разложения; (в) – сигнал foF2, совмещение аппроксимирующей компоненты и детализирующих компонент 3-го и 2-го уровней разложения

Результаты моделирования каждой восстановленной компоненты и процесс их совмещения в общую многокомпонентную конструкцию показаны на рисунке 2. Также на рисунке 2 представлены значения Кр-индекса, характеризующего солнечную активность. Процедура выделения аномалий в данных была основана на проверке условия (4). Анализ графиков на рисунке 2 подтверждает эффективность предлагаемого метода моделирования ионосферных данных и показывает наличие аномалии, возникающей в ионосфере за три дня до сейсмического события.

Выводы. В работе предложен метод многокомпонентного моделирования ионосферных данных, основанный на совмещении вейвлет-преобразования и авторегрессионных моделей. Выполнено моделирование, анализ данных и выделены аномалии, возникающие в периоды повышенной сейсмической активности на Камчатке (анализировались события энергетического класса с $k \geq 12$ в радиусе $R \sim 200$ км от П.-Камчатского).

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.» – № 9633р/14207 от 30.08.2011.

Литература

1. Э.Л. Афрамович, Н.П. Перевалова GPS-мониторинг и верхней атмосферы Земли – Иркутск:ГУ НЦ РВХ ВСНЦ СО РАМН, 2006. – 480с.
2. Марпл.-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ. -М.: Мир, 1990.
3. S. Mallat A Wavelet tour of signal processing [пер. с англ.] -М.: Мир, 2005. -671 с.
4. Мандрикова О.В., Глушкова Н.В. Метод моделирования данных критической частоты на основе совмещения вейвлет-преобразования и моделей авторегрессии-проинтегрированного скользящего среднего // Научные ведомости Белгородского государственного университета: Белгород. № 19. 2011.с.106-112.
5. Мандрикова О.В., Горева Т.С. Метод идентификации структурных компонентов сложного природного сигнала на основе вейвлет-пакетов. // Цифровая обработка сигналов.- Москва 2010 г.-№1.- с.45-50.

DETECTION AND ANALYSIS OF ANOMALIES IN DATA OF THE CRITICAL FREQUENCY OF THE IONOSPHERE BASED ON THE COMBINATION OF THE WAVELET TRANSFORM AND AUTOREGRESSIVE MODELS

N.V. GLUSHKOVA¹
O.V. MANDRICOVA²

¹⁾ Institute of Cosmophysical Research- es and Radio Wave Propagation, Petropavlosk-Kamchatsky

*²⁾ Kamchatka State Technical
University, Petropavlosk-Kamchatsky*

e-mail: nv.glushkova@yandex.ru, oksanam1@mail.kamchatka.ru

Described a method of multicomponent modeling iono- spheric data, based on combining wavelet-transform with the au- toregressive models, which allows to perform data analysis and selection of anomalous features. Approbation the method and based on it algorithms modeled data of the critical frequency f_{OF2} , recorded at the station "Paratunka" (Kamchatka Peninsula). Mod- eling identified features associated with solar activity, as well as appearing in periods of strong earthquakes in Kamchatka.

Keywords: wavelet transform, autoregressive model, critical frequency, anomalies.