

Метод моделирования и прогнозирования ионосферных данных на основе совмещения вейвлет-преобразования и моделей авторегрессии-проинтегрированного скользящего среднего

МАНДРИКОВА О.В.^{1,2}, ГЛУШКОВА Н.В.^{1,2}, ЖИВЕТЬЕВ И.В.¹

¹Институт космических исследований и распространения радиоволн ДВО РАН,
Россия

²Камчатский государственный технический университет, Россия

oksanam1@mail.kamchatka.ru, nv.glushkova@ya.ru

Аннотация

Работа направлена на создание средств анализа ионосферных параметров и выявления аномалий, возникающих в периоды ионосферных возмущений. В статье предложен *метод моделирования и анализа ионосферных параметров*, основанный на совмещении кратномасштабного анализа с моделями авторегрессии – проинтегрированного скользящего среднего. Метод позволяет выявлять закономерности в параметрах ионосферы, получать прогноз о вариациях, а также может быть использован для решения задачи заполнения пропусков в ионосферных параметрах с учетом их суточного и сезонного хода. Апробация метода выполнялась на данных критической частоты ионосферы f_oF2 и полного электронного содержания над Камчаткой и Магаданом (регистрацию данных выполняет ИКИР ДВО РАН).

Введение

Одной из важных задач анализа ионосферных данных является задача контроля состояния ионосферы и выделение и интерпретация аномалий, возникающих в периоды ионосферных возмущений [1-3]. Ионосферные аномалии могут быть обусловлены повышенной солнечной и сейсмической активностями [1-3]. Сложная структура ионосферных данных, наличие в них разномасштабных аномалий не позволяет использовать для их обработки традиционные методы, основанные на процедурах сглаживания [3]. Предложенный в статье *метод моделирования и анализа ионосферных параметров* основан на совмещении вейвлет-преобразования с моделями авторегрессии-проинтегрированного скользящего среднего (АРПСС). В основе метода лежит конструкция кратномасштабного анализа, позволяющая представить исходный временной ряд в виде разномасштабных компонент, подавить шум и выделить устойчивые характеристики структуры данных. Для построения общей параметрической конструкции, описывающей временной ход данных и позволяющей построить их прогноз, выделенные характерные компоненты моделируются методами АРПСС. Построенные многокомпонентные модели временных рядов f_oF2 и полного электронного содержания (ПЭС) над Камчаткой и Магаданом позволили изучить регулярные суточные и сезонные изменения параметров. На основе оценки остаточных ошибок полученных моделей выявлены аномалии, возникающие в периоды повышенной солнечной активности и в периоды сильных землетрясений на Камчатке.

Описание метода

2.1. Построение модели. В качестве базового пространства регистрируемых дискретных данных рассмотрим замкнутое пространство $V_j = \text{clos}_{L^2(R)}(2^j \phi(2^j t - k)) : k \in Z$ масштаба $j = 0$, порожденное скэйлинг-функцией $\phi \in L^2(R)$, где $L^2(R)$ – пространство Лебега [4]. На основе кратномасштабных разложений до уровня m , получим представление данных в виде [3]:

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

где $f[2^{-m} t] \in V_{-m}, g[2^j t] \in W_j$, W_j – пространство масштаба j , порожденное вейвлет-базисом $\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k)$; аппроксимирующая компонента $f[2^{-m} t] = \sum_k c_{-m,k} \phi_{-m,k}(t)$, где коэффициенты разложения $c_{-m,k} = \langle f, \phi_{-m,k} \rangle$, описывает тренд ряда; детализирующие компоненты $g[2^j t] = \sum_k d_{j,k} \Psi_{j,k}(t)$, где коэффициенты разложения $d_{j,k} = \langle f, \Psi_{j,k} \rangle : (j, k) \in I^j$, описывают разномасштабные детали; $e[2^j t] = \sum_k e_{j,k} \Psi_{j,k}$ – шумовые составляющие, где коэффициенты разложения $e_{j,k} = \langle f, \Psi_{j,k} \rangle : (j, k) \notin I^j$.

Для подавления шумовых составляющих $e[2^j t]$ и выделения компонент $f[2^{-m} t]$ и $g[2^j t]$, описывающих *устойчивые характеристики структуры* данных, применим следующие операции:

1. Восстановим каждую из полученных компонент (1) до исходного масштаба $j = 0$, получим восстановленные компоненты вида: $f_0^\mu(t) = \sum_k c_{0,k}^\mu \phi_{0,k}(t)$, $g_0^\mu(t) = \sum_k d_{0,k}^\mu \Psi_{0,k}(t)$, $e_0^\mu(t) = \sum_k e_{0,k}^\mu \Psi_{0,k}(t)$ где μ – номер компоненты.
2. Используя традиционный подход [5], определим модели из класса моделей АРПСС для аппроксимации каждой из полученных восстановленных компонент.
3. Выполним диагностические проверки полученных моделей компонент [5]. Если диагностические проверки модели компоненты подтверждают её адекватность данным, то будем считать, что модель компоненты готова к использованию и данная компонента, описывает *устойчивые характеристики* структуры данных.
4. Используя соотношение (1) объединим модели выделенных компонент в общую многокомпонентную конструкцию (остальные компоненты ряда из соотношения (1) будем считать шумовыми). Получим параметрическую *многокомпонентную конструкцию*, описывающую временной ход данных:

$$f_0(t) = \sum_{\mu=\overline{1,M}} \sum_{k=1, N_j^\mu} s_{j,k}^\mu(t) b_{j,k}^\mu(t), \quad (2)$$

где $s_{j,k}^\mu(t) = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k-l}^\mu(t) - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k-n}^\mu(t)$ – оценочное значение μ -ой компоненты, $\gamma_{j,l}^\mu p_j^\mu$ – параметры и порядок авторегрессионной модели μ -ой компоненты, $\omega_{j,k}^\mu(t) = \nabla^{\nu^\mu} \beta_{j,k}^\mu(t)$, $\beta_{j,k}^1 = c_{j,k}^1$, $\beta_{j,k}^\mu = d_{j,k}^\mu$, $\mu = \overline{2, M}$, ν^μ – порядок разности μ -ой компоненты, $h_j^\mu, \theta_{j,k}^\mu$ – порядок модели и параметры скользящего среднего модели μ -ой компоненты, $a_{j,k}^\mu$ – остаточные ошибки модели μ -ой компоненты, M – количество моделируемых компонент, описывающих *устойчивые характеристики структуры* данных, N_j^μ – длина μ -ой компоненты, $b_{j,k}^1 = \phi_{j,k}$ – скэйлинг-функция, $b_{j,k}^\mu = \Psi_{j,k}$, $\mu = \overline{2, M}$ – вейвлет-базис μ -ой компоненты, j – масштаб.

Замечание. Введенное соотношение (2) верно для любого масштаба j . Поэтому процедура идентификации модели может быть выполнена без операции восстановления компонент до исходного масштаба $j = 0$ (см. операция 1.). На основе изменения уровня разложения и применения различных вейвлет-функций может быть выбрана *наилучшая многокомпонентная модель временного ряда*.

Прогнозирование данных и обнаружение аномалий

Прогнозирование значения $s_{j,k+q}^\mu$, $q \geq 1$ (см. соотношение (2)) определяет прогноз $s_{j,k}^\mu$ в момент $t = k$ с упреждением q . Значение $s_{j,k+q}^\mu$ на основе полученной модели (2) определяется как $s_{j,k+q}^\mu(t) = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k+q-l}^\mu(t) - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k+q-n}^\mu(t)$. Остаточные ошибки μ -ой компоненты модели масштаба j определяются как разность между прогнозными и фактическими значениями данных в момент времени $t = k+q$: $a_{j,k+q}^\mu(t) = s_{j,k+q,\text{прогноз}}^\mu(t) - s_{j,k+q,\text{фактич}}^\mu(t)$. Если в данных возникает аномалия, то произойдет изменение их структуры и абсолютные значения остаточных ошибок возрастут. Поэтому *процедура выделения аномалий* может быть построена на оценке остаточных ошибок полученных моделей компонент при выполнении операции прогнозирования. *Обнаружение аномалий* в компоненте с номером μ масштаба j будем выполнять на основе проверки условия:

$$D_{U_j} = \frac{1}{U_j} \sum_{q=1}^{U_j} (a_{j,k+q}^\mu(t))^2 > T_{A_j}, \quad (3)$$

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

Результаты моделирования и анализа данных

В процессе исследований использовались часовые данные foF2 и двухчасовые данные ПЭС. Учитывая сезонный характер ионосферного процесса, данные предварительно были разделены на сезоны и моделировались отдельно. На основе изменения уровня разложения и применения различных вейвлет-базисов (использовались базисные функции семейства Добеши), для описания временного хода foF2 были получены разные модели вида (2) (см. замечание п.2.1). Путем минимизации остаточных ошибок полученных моделей, определена *наилучшая модель временного хода foF2*, которая в вейвлет-пространстве имеет следующий вид:

$$f_0(t) = f[2^{-3}t] + \sum_{j=-2}^{-3} g[2^j t] + e[2^{-1}t], \quad (4)$$

где компонента $f[2^{-3}t] = \sum_k c_{-3,k} \phi_{-3,k}(t)$ – описывает тренд ряда, компоненты $g[2^j t] = \sum_k d_{j,k} \Psi_{j,k}(t)$, $j = -2, -3$ – описывают разномасштабные детали, $e[2^{-1}t]$ – шумовая составляющая. Для полученных на основе соотношения (4) восстановленных компонент $f_0^1(t) = \sum_k c_{0,k}^1 \phi_{0,k}(t)$ и $g_0^\mu(t) = \sum_k d_{0,k}^\mu \Psi_{0,k}(t)$, $\mu = 2, 3$ были идентифицированы проинтегрированные авторегрессионные модели второго порядка. На рис.1 показаны результаты моделирования данных foF2 станции “Паратунка”. Анализ рис.1 показывает, что модель позволяет выполнять прогноз вариаций критической частоты с шагом упреждения пять часов. Сопоставление результатов моделирования с геомагнитными данными и данными

каталога землетрясений показало, что в периоды повышенной солнечной и сейсмической активности наблюдается увеличение ошибок моделей. На рис.2 представлены результаты моделирования foF2 и ПЭС. Выявление аномалий основано на формуле (3). Анализ рис.2 показывает, что накануне землетрясений наблюдается локальная аномальная особенность в ионосфере длительностью несколько часов (в данных foF2 30.06.05). В период землетрясений возникает крупномасштабная аномалия длительностью более суток (в данных foF2 и ПЭС).

Вывод

Предложенный в статье метод моделирования и анализа ионосферных параметров позволил построить модели временного хода foF2 и ПЭС и выявить аномальные особенности, возникающие в ионосфере в периоды возрастания солнечной и сейсмической активности на Камчатке.

Работа поддержана грантом РФФИ – ДВО РАН №11-07-98514-р_восток_a и грантом “У.М.Н.И.К.” - № 9633р/14207 от 30.08.2011. Данные сейсмического каталога любезно предоставлены Камчатским филиалом геофизической службы РАН (г. Петропавловск-Камчатский).

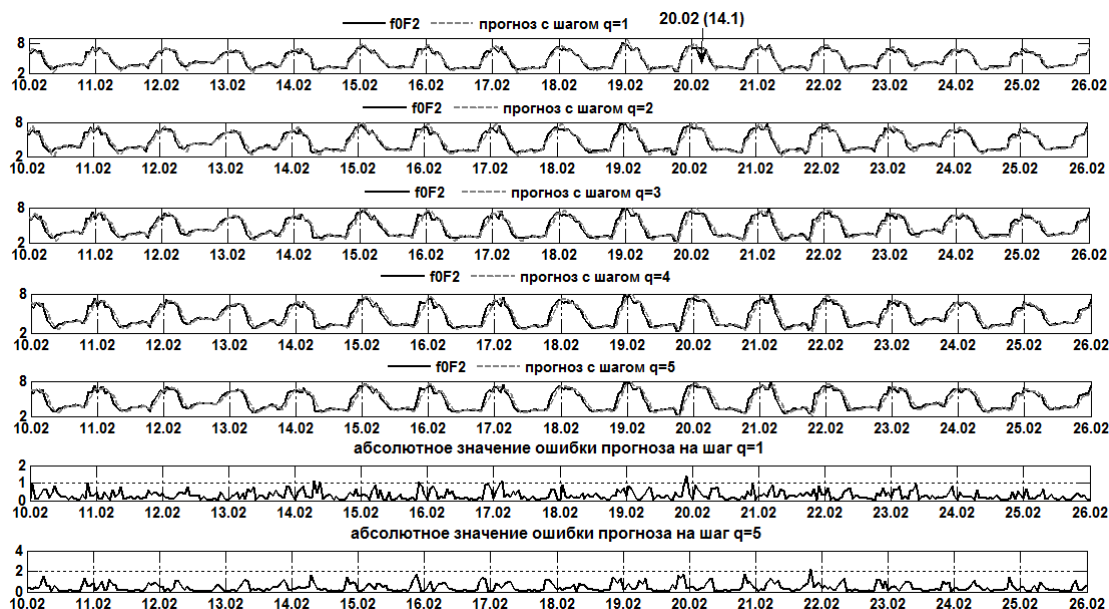


Рис. 1. Моделирование данных foF2 за период 09.02.2011-27.02.2011 гг.

Литература

1. Афраймович Э.Л. GPS- мониторинг верхней атмосферы Земли / Э.Л.Афраймович, Н.П. Перевалова - Иркутск: ГУ НУ РВХ ВСНЦ СО РАН, 2006.-480с.
2. Липеровская Е.В., Липеровский В.А., Похотелов О.А. О возмущениях в F-области ионосферы перед землетрясениями // Геофизические исследования. -2006. -№ 6. -С. 51-58.
3. S. Mallat A Wavelet tour of signal processing [пер. с англ.] -М.: Мир, 2005. -671 с.
4. Мандрикова О.В., Глушкова Н.В. Моделирование и анализ данных критической частоты ионосферы на основе совмещения вейвлет-преобразования и моделей авторе-

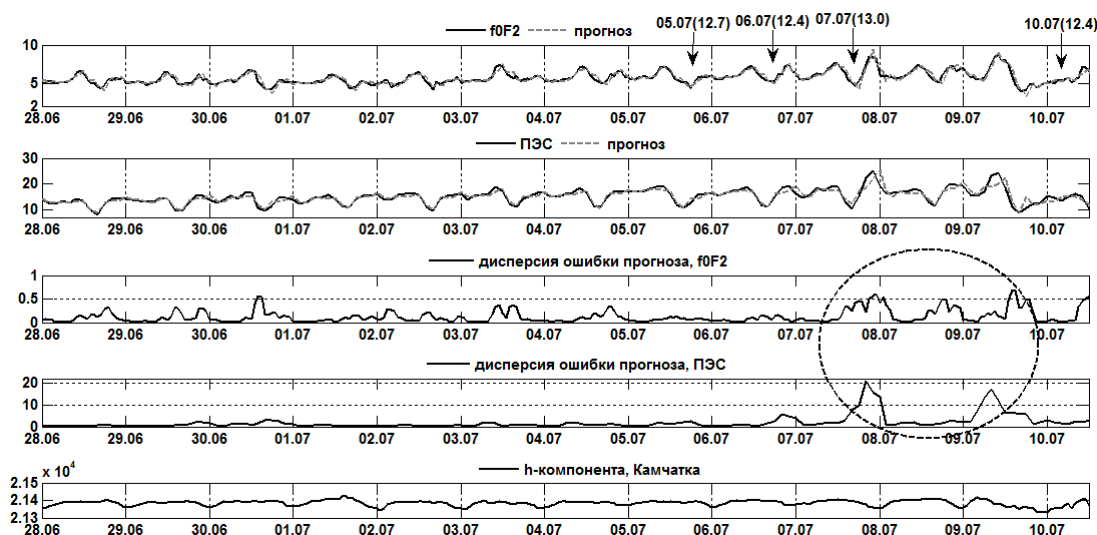


Рис. 2. Моделирование данных f_0F_2 и ПЭС за период 27.06.2005-10.07.2005гг. Стрелками отмечены землетрясения. Крупномасштабная аномалия выделена пунктиром.

грессии// 14-я Международная конференция “Цифровая обработка сигналов и ее применение – DSPA-2012”, Москва. Т. 2. -С. 346-349.

5. Бокс Дж., Дженкинс Г. Анализ временных рядов прогноз и управление. -М.: Мир, 1974. -604 с.

Method of modeling and forecasting of ionospheric data based on the combination of wavelet transform and autoregressive-integrated moving average models

Mandrikova O.V.^{1,2}, Glushkova N.V.^{1,2}, Zhivetiev I.V.¹

¹ Institute of Cosmophysical Researches and Radio Wave Propagation FEB RAS, Russia

² Kamchatka State Technical University, Russia

The work is focused on the development of technologies and software systems for the study of the time variations of ionospheric parameters and detection anomalies that can occur during periods of ionospheric disturbances. The paper describes a method of modeling of calm (background) variations of the critical frequency of the F2 layer of the ionosphere. The method is a combination of multiply-scaled analysis and methods of autoregressive-integrated moving average (ARIMA). This method allows identifying typical behavior of ionospheric parameters, making forecast with five-hour increments and detecting anomalies which occur during periods of ionospheric disturbances. Using multiply-scaled analysis allows us to decompose the time series into components and simplify the complex structure of the data. Computational algorithms developed on the basis of the method provide a selection of informative components and noise reduction; these informative components are modeled by ARIMA methods. Forecasting and analysis of residual errors of the model provides detection of anomalies which can occur during periods of ionospheric disturbances. The proposed method can be used to fill data gaps with respect to diurnal and seasonal variation.

To test the method, we used data of the critical frequency of the F2 layer of the ionosphere (registered at the station Paratunka, IKIR FEB RAS), as well as data of the total electron content in the ionosphere over Kamchatka. We detected some peculiarities associated with solar and seismic activity on Kamchatka using the method we developed.