

Ionospheric parameter modelling and anomaly discovery by combining the wavelet transform with autoregressive models

Oksana V. Mandrikova^{1,2}, Nadegda V. Fetisova (Glushkova)^{1,2}, Riad Taha Al-Kasasbeh^{3,*}, Dmitry M. Klionskiy⁴, Vladimir V. Geppener⁴, Maksim Y. Ilyash⁵

¹ Institute of Cosmophysical Research and Radio Wave Propagation (far-eastern branch of the Russian Academy of Sciences), Paratunka, Kamchatka region, Russian Far East, Russian Federation

² Kamchatka State Technical University, Petropavlovsk-Kamchatsky, Russian Federation

³ Al-balqa Applied University, Amman, Jordan

⁴ Saint Petersburg Electrotechnical University "LETI", Saint Petersburg, Russian Federation

⁵ Saint Petersburg State University, Saint Petersburg, Russian Federation

Article history

Received January 21, 2015; accepted August 10, 2015.

Subject classification:

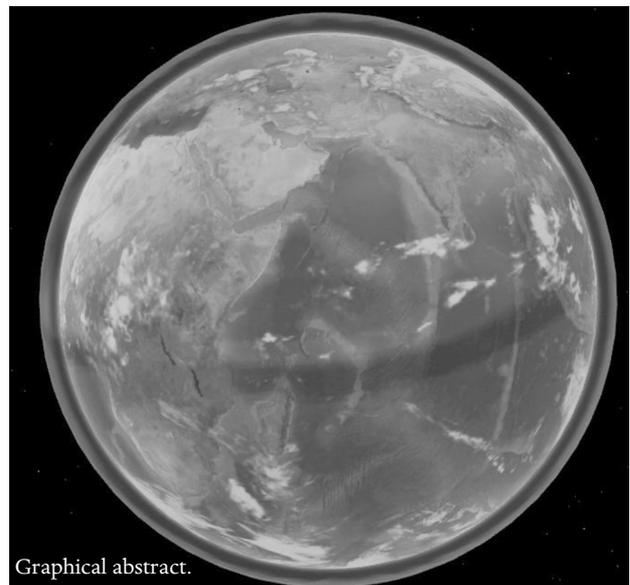
Ionospheric parameters, Perturbation, Earthquake, Anomaly, Prediction, Wavelet transform, Autoregressive integrated moving average model, Critical frequency data, Total electron content data.

ABSTRACT

The paper is devoted to new mathematical tools for ionospheric parameter analysis and anomaly discovery during ionospheric perturbations. The complex structure of processes under study, their a-priori uncertainty and therefore the complex structure of registered data require a set of techniques and technologies to perform mathematical modelling, data analysis, and to make final interpretations. We suggest a technique of ionospheric parameter modelling and analysis based on combining the wavelet transform with autoregressive integrated moving average models (ARIMA models). This technique makes it possible to study ionospheric parameter changes in the time domain, make predictions about variations, and discover anomalies caused by high solar activity and lithospheric processes prior to and during strong earthquakes. The technique was tested on critical frequency foF2 and total electron content (TEC) datasets from Kamchatka (a region in the Russian Far East) and Magadan (a town in the Russian Far East). The mathematical models introduced in the paper facilitated ionospheric dynamic mode analysis and proved to be efficient for making predictions with time advance equal to 5 hours. Ionospheric anomalies were found using model error estimates, those anomalies arising during increased solar activity and strong earthquakes in Kamchatka.

1. Introduction

One of the most important tasks of ionospheric data analysis is connected with ionospheric state control along with anomaly discovery with subsequent interpretation of anomalies occurring during ionospheric perturbations [Afraimovich et al. 2000, Afraimovich et al. 2001, Nakamura et al. 2007, Pervak et al. 2008,



Hamoudi et al. 2009, Liming et al. 2011, Mandrikova et al. 2012a, Mandrikova et al. 2012b]. Ionospheric anomalies may be caused by increased solar activity; in seismically active regions they can also be encountered during the periods of high seismic activity [Mandrikova et al. 2012a, 2012b]. The Earth's ionosphere is the area of the atmosphere extending from 80 km to 1000 km and affecting radio wave propagation [Maruyama et al. 2011, Nakamura et al. 2009, Watthanasangmechai et al. 2012]. The dynamic mode of the ionosphere is determined by the distribution of its parameters over altitude, atmospheric density, chemical composition, and

spectral characteristics of solar radiation. Ionospheric studies are mostly based on analyzing environmental parameters, which can dramatically change with altitude and depend on a solar activity cycle, geomagnetic conditions, geographical coordinates and contain characteristic diurnal and seasonal changes [Afraimovich et al. 2000, Afraimovich et al. 2001, Nakamura et al. 2009, Watthanasangmechai et al. 2012, Danilov 2013, Roux 2012]. Among the main parameters characterizing the ionospheric state are variations of critical frequency of F2-layer of the ionosphere (foF2) and total electron content. Ionospheric critical frequency data (foF2) are registered by vertical radio sounding using an impulse radar (ionosonde) and these data are represented as a time series. Variations of total electron content are obtained using readings of double-frequency ground GPS receivers [Afraimovich et al. 2000]. Variations of foF2 and total electron content characterizing electronic concentration of the ionosphere largely depend on diurnal and seasonal variations of solar radiation, solar activity cycles, and geomagnetic activity [Afraimovich et al. 2000, Nakamura et al. 2007, Hamoudi et al. 2009, Mandrikova et al. 2012a, Mandrikova et al. 2012b]. Abrupt fluctuations of the ionospheric electron density arising due to increased solar and/or geomagnetic activity lead to abnormal behavior of ionospheric critical frequency variations and total electron content variations [Afraimovich et al. 2000, Nakamura et al. 2007, Hamoudi et al. 2009, Mandrikova et al. 2012a, Mandrikova et al. 2012b].

Ionospheric state control and anomaly extraction are important tasks of ionospheric parameter analysis [Afraimovich et al. 2000, Afraimovich et al. 2001, Liu et al. 2008, Nakamura et al. 2009, Roux 2012, Watthanasangmechai et al. 2012, Danilov 2013, Ezquer et al. 2014, Zhao et al. 2014]. Ionospheric parameters affect many spheres of our life and can have a negative effect on satellite systems and radio communication propagation. An ionospheric anomaly is a very complex phenomenon in time and space [Afraimovich et al. 2000, Afraimovich et al. 2001, Danilov 2013] and they mostly arise in periods of increased solar and geomagnetic activities [Afraimovich et al. 2000, Afraimovich et al. 2001, Danilov 2013, Roux 2012]. In seismically active regions they can also be observed in periods of increased seismic activity [Afraimovich et al. 2000, Afraimovich et al. 2001, Nakamura et al. 2007, Nakamura et al. 2009, Watthanasangmechai et al. 2012, Klimenko et al. 2012]. The registered ionospheric parameters contain anomalies manifesting themselves as a significant deviation (rise or decrease) from the background level. The main reason for anomalies in lower ionospheric layers (regions E and D) is ionization rate variation due to cor-

puscular invasion. Variations in region F are caused by indirect factors - neutral composition and dynamic processes [Prolls 1995]. Morphology and physics of ionospheric phenomena were described in many previous works and review papers [Prolls 1995, Rees 1995, Buonsanto 1999, Mendillo 2006, Danilov 2013].

The ionospheric structure, being variable and heterogeneous, is normally studied by variation analysis of environmental parameters registered by special equipment. Lack of knowledge on the structure of registered parameters and an insufficient number of formal models for parameter representation make the task in question rather complicated to deal with. Despite the fact that contemporary technologies of near-Earth space monitoring are developing rather steadily and rapidly, the capabilities to control and predict the ionospheric state are still rather limited [Nakamura et al. 2007, Mandrikova et al. 2012a, Mandrikova et al. 2012b].

The complex structure of ionospheric parameter variations results in the conventional techniques being inefficient for any kind of mathematical modelling or computer simulation. The conventional techniques mentioned above are mostly based on smoothing procedures and do not allow us to explore subtle features of data [Nakamura et al. 2007, Pervak et al. 2008, Hamoudi et al. 2009, Liming et al. 2011, Mandrikova et al. 2012a, Mandrikova et al. 2012b], the latter usually containing important information about the processes under study. Low sensitivity of the registered parameters is the reason why their processing and analysis have become even more complicated and time-consuming. One of the popular approaches to data simulation and analysis is the application of ARIMA models [Kay and Marple 1981, Huang et al. 2013]. They make it possible to reveal steady characteristics of data and some possible ways of data prediction. Multiple practical research has confirmed the capability of this mathematical apparatus to deal with various applied problems in the area of geophysics [Box and Jenkins 1970, Basseville and Nikiforov 1993, Huang et al. 2013]. Nonetheless, ARIMA methods have certain restrictions on working with particular kinds of data [Nakamura et al. 2009, Huang et al. 2013].

Wavelet-based approaches have been broadly applied to multiple areas of signal processing and related fields in the last twenty years [Al-Kasasbeh 2004, Mandrikova and Bogdanov 2007, Zhang et al. 2008, Li et al. 2009, Vargas-Bracamontes et al. 2009, Beltrán and Ponce de León 2010, Hafez et al. 2010, Mandrikova et al. 2010, Hafez and Ghamry 2011, Liming et al. 2011, Mandrikova et al. 2011, Wang et al. 2011, Gwal et al. 2012, Mandrikova et al. 2012b, Mandrikova et al. 2012c, Bakhshi et al. 2013, Castillo et al. 2013, Han and Chang 2013, Mandrikova et al. 2013a]. These methods encompass some

elements of the approximation theory [Donoho and Johnstone 1998, Mallat 1999] and filtering [Mitra 1998] and in many cases they help us to solve the tasks in question [Vargas-Bracamontes et al. 2009, Mandrikova et al. 2012a, Mandrikova et al. 2012b, Mandrikova et al. 2012c, Mandrikova et al. 2013a]. Owing to a wide range of existing wavelet bases with compact support, wavelets provide special opportunities for detailed studies of complex data structure. Fast computational algorithms for wavelet decomposition [Mallat 1999, Percival and Walden 2000, Daubechies 2001] undoubtedly pose a substantial advantage for real-time signal processing, which is indispensable for quick data analysis.

The suggested technique of ionospheric parameter simulation and analysis is based on combining the wavelet transform with ARIMA models. The underlying idea is concerned with employing multiresolution analysis [Mallat 1999, Percival and Walden 2000, Daubechies 2001], which acts to decompose the original dataset into a finite number of multiscale components with different time resolutions. Application of the wavelet transform to foF2 data in order to extract trend and fluctuations was also suggested in the work by Roux [2012] and showed its efficiency for finding connections between ionospheric parameters and solar and geomagnetic activities. The components obtained by decomposition have a more simple structure compared to the original data. This paper is focused on ARIMA methods used for approximation. This approach has been applied by the authors since 2003 and was successfully adapted for anomaly extraction in subsoil radon. Identification of *component models* includes noise suppression and determination of stable data characteristics. In this paper we suggest joining multiscale component models together so that they will make a *general parametric construction* characterizing data variability in time and allowing us to make predictions. Using residual error estimates for final prediction we have developed a numerical algorithm for anomaly extraction.

Multicomponent time series models of ionospheric critical frequency foF2 (data were registered by the University of Cosmophysical Research and Radio Wave Propagation, Paratunka village, Kamchatka region, Russian Far East) and TEC (data were obtained from bi-frequency ground-based GPS receivers) [Afraimovich et al. 2000, Afraimovich et al. 2001, Erdođan and Arslan 2009] in Kamchatka and Magadan were constructed. They confirmed the efficiency of the suggested technique and made it possible for scientists to analyze regular diurnal and seasonal parameter changes. Using the estimate of deviation from the background level ionospheric anomalies were extracted with a duration ranging from several dozens of minutes to several hours.

These anomalies arise during ionospheric perturbations. Our detailed analysis has shown that the anomalies described above occur during the increased solar activity and strong earthquakes in Kamchatka (several seismic events of energy class $k > 12.5$ have been analyzed to make this conclusion).

2. Suggested technique

2.1. Model construction

Consider a closed space $V_j = \text{clos}_{L^2(\mathbb{R})} 2^j \varphi 2^j t - k : k \in \mathbb{Z}$ with scale $j = 0$ formed by a scaling function $\varphi \in L^2\mathbb{R}$. Using multiresolution decompositions to level m we can represent the original data the following way [Mandrikova et al. 2011, 2012c, 2013a]:

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

where:

- $f 2^{-m} t \in V_{-m}$, $g 2^j t \in W_j$;
- W_j is a space of scale j formed by wavelet basis $\Psi_{j,k}(t) = 2^{j/2} \Psi(2^j t - k)$;
- $f 2^{-m} t = \sum_k c_{-m,k} \varphi_{-m,k}(t)$ is an approximating component;
- $c_{-m,k} = \langle f, \varphi_{-m,k} \rangle$ are decomposition coefficients characterizing trend;
- $g 2^j t = \sum_k d_{j,k} \Psi_{j,k}(t)$ are detailed components;
- $d_{j,k} = \langle f, \Psi_{j,k} \rangle : (j,k) \in I^j$ are decomposition coefficients representing multiscale details;
- $e 2^j t = \sum_k e_{j,k} \Psi_{j,k}$ are noise components;
- $e_{j,k} = \langle f, \Psi_{j,k} \rangle : (j,k) \notin I^j$ are decomposition coefficients.

Expression (1) can be used for representing a time series as a sum of multiscale components $f 2^{-m} t$, $g 2^j t$, and $e 2^j t$, where $j = \overline{-1, -m}$. In order to suppress noise components $e 2^j t$ and extract components $f 2^{-m} t$ and $g 2^j t$ characterizing stable characteristics of data the following sequence of steps is needed:

(1) Reconstruction of the components $f 2^{-m} t$, $g 2^j t$, and $e 2^j t$, $j = \overline{-1, -m}$ to the original scale $j = 0$ and calculation of the reconstructed components $f_0^\mu(t) = \sum_k c_{0,k}^\mu \varphi_{0,k}(t)$, $g_0^\mu(t) = \sum_k d_{0,k}^\mu \Psi_{0,k}(t)$, $e_0^\mu(t) = \sum_k e_{0,k}^\mu \Psi_{0,k}(t)$, where μ is a component number.

(2) Application of a conventional approach [Basville and Nikiforov 1993] for determining an ARIMA model in order to approximate each reconstruction component $f_0^\mu(t) = \sum_k c_{0,k}^\mu \varphi_{0,k}(t)$, $g_0^\mu(t) = \sum_k d_{0,k}^\mu \Psi_{0,k}(t)$, $e_0^\mu(t) = \sum_k e_{0,k}^\mu \Psi_{0,k}(t)$.

(3) Performance of diagnostic tests for component models. If diagnostic tests for a component model confirm its adequacy, the component model reflects stable characteristics of data (according to ARIMA model theory).

(4) Combination of the extracted component models to create a *general parametric construction* (all other components from Expression (1) will be considered as

noise components). The outcome of this combination is a *parametric multicomponent model* representing data changes in the time domain:

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

where:

- $s_{j,k}^\mu(t) = \sum_{l=1}^{p_j^\mu} \gamma_{j,l}^\mu \omega_{j,k-l}^\mu t - \sum_{n=1}^{h_j^\mu} \theta_{j,n}^\mu a_{j,k-n}^\mu(t)$ is an estimated value of the μ -th component;

- $\gamma_{j,l}^\mu$ are autoregressive parameters of the μ -th component;

- $\omega_{j,k}^\mu t = \nabla^{v^\mu} \beta_{j,k}^\mu(t)$, $\beta_{j,k}^1 = c_{j,k}^1$, $\beta_{j,k}^\mu = d_{j,k}^\mu$, $\mu = \overline{2, \overline{M}}$, $v^{\mu-}$ is a difference order of the μ -th component;

- p_j^μ is an autoregressive model order of the μ -th component;

- h_j^μ , $\theta_{j,k}^\mu$ are model orders and parameters of moving average of the μ -th component model;

- $a_{j,k}^\mu$ are residual errors of the μ -th component model;

- M is the number of modelled components representing stable characteristics of data;

- N_j^μ is the length of the μ -th component;

- $b_{j,k}^1 = \varphi_{j,k}$ is a scaling function;

- $b_{j,k}^\mu = \Psi_{j,k}^\mu$, $\mu = \overline{2, \overline{M}}$ is a wavelet basis of the μ -th component;

- j is scale.

Remark 1. Expression (2) is valid for any scale j . Thus, identification of the suggested model can be carried out without reconstructing the components to the original scale $j=0$ (see step 1 above). By changing a decomposition level (see Expression (1)) we can obtain various models of type (2) for representing a time series. Minimization of the residual error leads to the *best multicomponent model of a time series*. The resulting estimate can be further improved by applying various wavelets.

2.2. Data prediction and anomaly discovery

Prediction of $s_{j,k+q}^\mu$, $q \geq 1$ (see Expression (2)) determines the predicted value $s_{j,k}^\mu$ at the moment $t=k$ with time advance q . Value $s_{j,k+q}^\mu$ can be found via model (2):

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

The residual errors of the μ -th component at scale j are determined as a difference between predicted and actual values of data at time point $t=k+q$:

$$a_{j,k+q}^\mu(t) = s_{j,k+q, \text{predicted}}^\mu(t) - s_{j,k+q, \text{actual}}^\mu(t).$$

Model (2) represents regular changes in approximated data. If an anomaly has occurred, the data structure will be modified, but the absolute residual errors

will rise. Anomaly extraction can be performed using a residual error estimate of component models when performing predictions.

Anomaly discovery in a component with number μ at scale j will be based on the following condition:

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

where T_{A_j} is a threshold given in advance (this threshold defines the presence of anomalies on scale j in data), U_j is the window length on scale j .

3. Experimental research

The experimental research was conducted using critical frequency foF2 data collected in Kamchatka and Magadan. The critical frequency foF2 data contain missing values, which might impede further simulation and detailed analysis. To reduce errors it was decided to choose time periods with the smallest number of misses. Due to seasonal features of ionospheric processes, the data were first split into seasonal sets and then simulated separately. Below there is a detailed description of the critical frequency data simulation for winter and summer seasons.

The model construction was based on hourly values of foF2 critical frequency data for the period from 1968 to 2013. These data were obtained by vertical radio sounding of the ionosphere. Simulation of quiet variations of ionospheric parameters was performed using time intervals of a relatively calm geomagnetic field (intervals where the resulting K-index ΣK did not exceed 24). These intervals did not contain any strong seismic events (intervals where there were no earthquakes with the energy class $Ks \geq 12$, which happened no farther than 300 km from the station of the ionospheric radio sounding).

Taking into account the seasonal variability of an ionospheric process we performed separate data simulation for different seasons. Tables 1 and 2 contain time intervals for identification and model estimation. The solar activity was estimated using the value of radiation corresponding to a wavelength of f10.7. If $f10.7 < 100$ the activity was taken as low, otherwise the activity was considered to be high.

By changing the decomposition level (see (1)) and applying various wavelet functions (Daubechies wavelets were taken) different models of type (2) were obtained for representing time changes in foF2. Model identification and diagnostics for different time series components (steps 2,3 in 2.1) have shown that the original time series foF2 and their approximating components of the 1st and 2nd decomposition levels have a complex

High solar activity	
winter season	summer season
20.12.1968–09.01.1969	31.07.1969–20.08.1969
03.01.1970–22.01.1970	30.07.1970–16.08.1970
11.02.1970–02.03.1970	07.06.1979–24.06.1979
03.01.1982–20.01.1982	20.07.1989–10.08.1989
01.01.1991–23.01.1991	10.07.1990–28.07.1990
30.11.2000–19.12.2000	27.06.1998–15.07.1998
19.12.2011–08.01.2012	20.06.2001–07.07.2001
29.12.2012–24.01.2013	18.06.2002–06.07.2002
	28.05.2002–18.06.2002

Low solar activity	
09.01.1984–29.01.1984	24.06.1983–14.07.1983
11.02.1985–27.02.1985	26.05.1983–12.06.1983
22.01.1987–10.02.1987	20.06.1984–10.07.1984
29.12.2005–14.01.2006	28.06.2004–17.07.2004
15.01.2006–04.02.2006	13.08.2007–01.09.2007
09.02.2008–27.02.2008	

Table 1. Time intervals of foF2 data used for model identification.

High solar activity	
winter season	summer season
15.12.1970–29.12.1970	23.06.1969–05.07.1969
07.02.1981–21.02.1981	04.06.1971–21.06.1971
07.02.2002–25.02.2002	12.06.1989–29.06.1989
05.12.2011–18.12.2011	13.08.2000–26.08.2000
30.01.2012–11.02.2012	03.08.2002–17.08.2002
04.02.2013–18.02.2013	06.07.2002–18.07.2002

Low solar activity	
13.12.1983–29.12.1983	25.06.1987–06.07.1987
06.01.1993–25.01.1993	09.08.2010–22.08.2010
05.01.2007–18.01.2007	

Table 2. Time intervals of foF2 data used for model diagnostics.

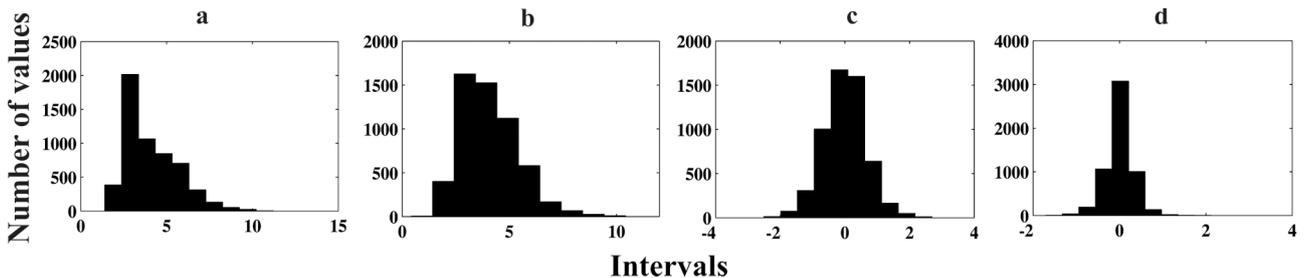
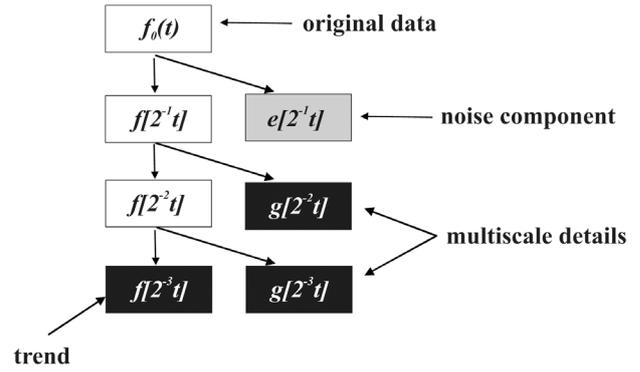

 Figure 2. Distribution estimates using histograms: (a) histogram of the original time series foF2; (b,c,d) histograms of the components $f_0^1(t)$, $g_0^3(t)$, $g_0^2(t)$.


Figure 1. Approximated components in the wavelet domain (high-lighted in black).

structure and cannot be approximated by an ARIMA model. The reason is non-zero correlation of model residuals. Residual error minimization is the way of finding the *best time series model* foF2, which includes 3 (three) components for winter and summer seasons. These components have the following form in the wavelet domain:

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

where:

- $f 2^{-3}t = \sum_k c_{-3k} \varphi_{-3k}(t)$ characterizes trend;
- $g 2^j t = \sum_k d_{j,k} \Psi_{j,k}^\mu(t)$, $j = -2, -3$ are components representing multiscale details;
- $e 2^{-1}t$ is a noise component.

Both of the reconstructed components $f_0^1(t) = \sum_k c_{0,k}^1 \varphi_{0,k}(t)$ and $g_0^\mu(t) = \sum_k d_{0,k}^\mu \Psi_{0,k}^\mu(t)$, $\mu = 2, 3$ have a more simple structure than the original time series $f(t)$. In Figure 2 we can see the distribution estimates of the registered time series foF2 and the components $f_0^1(t)$, $g_0^\mu(t)$, $\mu = 2, 3$. The model parameter estimation is performed for each extracted component.

Each reconstructed component $f_0^1(t) = \sum_k c_{0,k}^1 \varphi_{0,k}(t)$ ($j = -3$ before reconstruction, time periods are 8 and longer) and $g_0^\mu(t) = \sum_k d_{0,k}^\mu \Psi_{0,k}^\mu(t)$, $\mu = 2, 3$ ($j = -2-3$ before reconstruction, time periods range from 2 to 4 hours and from 4 to 8 hours, respectively) was used for identifying ARIMA models of the 2nd order.

In Table 3 there are estimated model parameters

Analyzed period	$f_0^1(t)$		$g_0^3(t)$		$g_0^2(t)$	
	Parameter		Parameter		Parameter	
	1st	2nd	1st	2nd	1st	2nd
27.06.2005–10.07.2005 (Kamchatka)	1.013	-0.291	0.828	-0.339	0.443	-0.572
16.01.2006–04.02.2006 (Kamchatka)	1.012	-0.278	0.829	-0.344	0.371	-0.622
09.02.2011–27.02.2011 (Kamchatka)	1.009	-0.266	0.805	-0.355	0.437	-0.473
16.01.2006–04.02.2006 (Magadan)	1.009	-0.267	0.821	-0.348	0.366	-0.624
General parameters (summer and winter seasons)	1.01	-0.27	0.81	-0.35	0.35	-0.65

Table 3. Component model parameters.

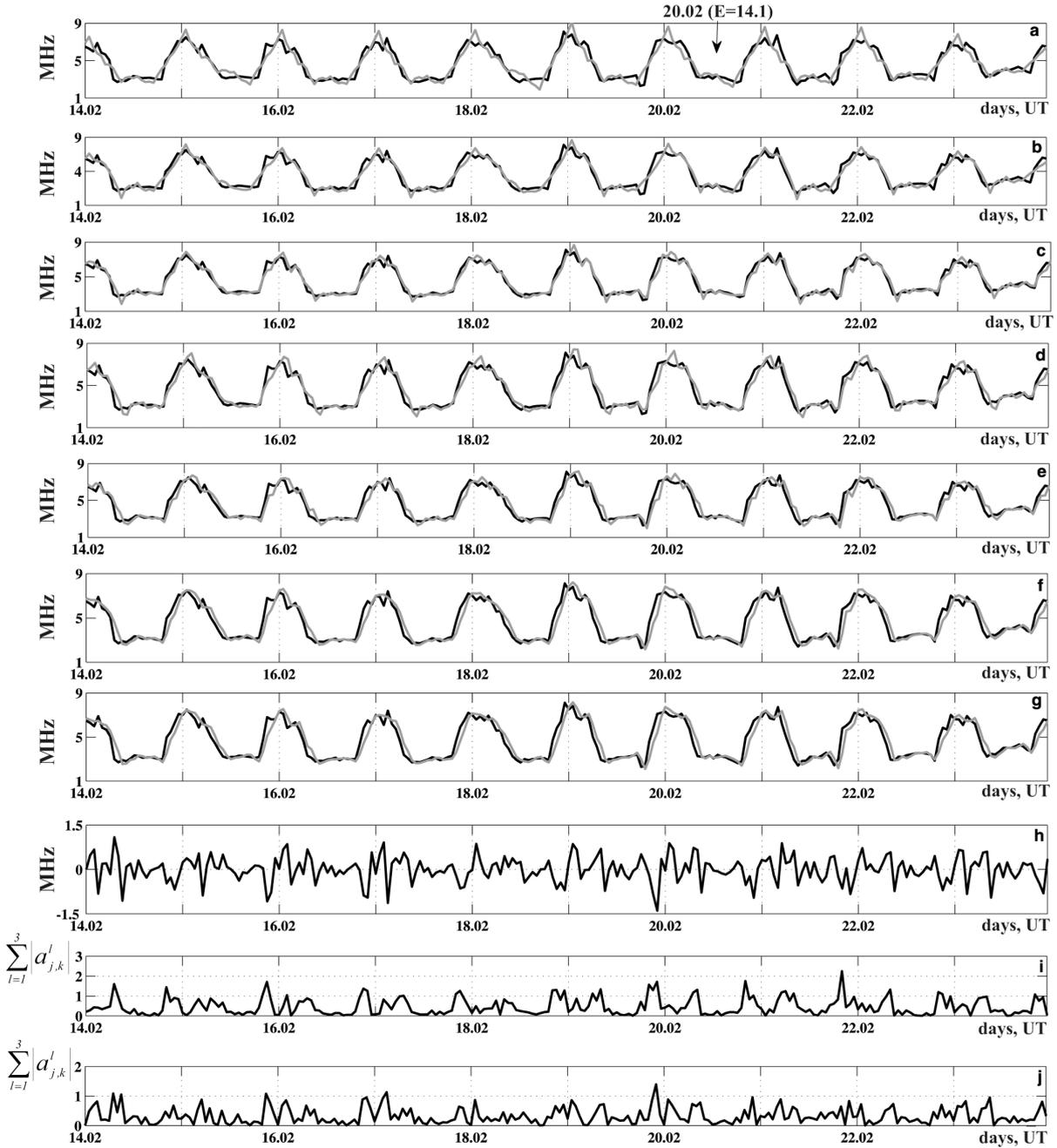


Figure 3. Simulation and prediction results for foF2 data. Black: original data foF2; grey: prediction. (a) Prediction for $f_0^1(t)$ with time advance $q = 1$; (b) prediction of the sum $f_0^1(t)$ and $g_0^3(t)$ with time advance $q = 1$; (c) prediction of the sum $f_0^1(t)$, $g_0^3(t)$, and $g_0^2(t)$ with time advance $q = 1$; (d,e,f,g) prediction of the sum $f_0^1(t)$, $g_0^3(t)$, and $g_0^2(t)$ with time advance $q = 2, 5, 5$, respectively; (h) prediction error with time advance $q = 1$; (i) absolute prediction error with time advance $q = 1$; (j) absolute prediction error with time advance $q = 5$. The arrow indicates a seismic event in Kamchatka (energy class $k = 14.1$).

of these components for data with different time periods. Due to close parameters of the corresponding components a general time series model was determined for foF2 for winter and summer seasons:

$$f_0(t) = \sum_{\mu=1,3} \sum_{k=1,N} s_{0,k}^{\mu}(t) b_{0,k}^{\mu} t, \quad (5)$$

where

- $s_{0,k}^1(t) = (1.01 \cdot \omega_{0,k-1}^1 - 0.27 \omega_{0,k-2}^1) + a_{0,k}^1(t)$; $s_{0,k}^2(t) = (0.35 \cdot \omega_{0,k-1}^2 - 0.65 \omega_{0,k-2}^2) + a_{0,k}^2(t)$; $s_{0,k}^3(t) = (0.81 \cdot \omega_{0,k-1}^3 - 0.35 \omega_{0,k-2}^3) + a_{0,k}^3(t)$;
- $B^l \omega_{0,k}^{\mu}(t) = \omega_{0,k-l}^{\mu}(t)$;
- $a_{0,k}^{\mu}(t)$ are residual errors of the μ -th component model;
- $\omega_{0,k-l}^{\mu} t = \nabla \beta_{0,k}^{\mu}(t)$;
- $\beta_{0,k}^1 = c_{0,k}^1$;
- $\beta_{0,k}^{\mu} = d_{0,k}^{\mu}$, $\mu = 2, 3$;
- $a_{0,k}^{\mu}$ are residual errors of the μ -th component model;
- N is time series length;
- $b_{0,k}^1 = \varphi_{0,k}$ is a scaling function;
- $b_{0,k}^{\mu} = \Psi_{0,k}^{\mu}$, $\mu = 2, 3$ is a wavelet basis of the μ -th component model.

Figure 3 shows the simulation and prediction results for foF2 data from the Paratunka observatory. These results were found by model (5) (the period is 14.02.2011–23.02.2011). Figure 4 also illustrates the simulation error estimates for foF2 data obtained by ARIMA methods and the suggested technique. These estimates were calculated using foF2 data from Paratunka for the following periods: 13.07.1969–21.08.1969, 26.07.1970–01.09.1970, 26.07.1972–22.07.1972, 31.05.1979–03.07.1979, 26.05.1983–11.06.1983, 22.05.1984–18.06.1984, 21.07–

16.08.1987, 21.07.1989–16.08.1989, 30.07–21.08.1992, 30.07.1999–21.08.1999, 13.06.2000–28.06.2000, 18.06.2002–06.07.2002, 27.06–10.07.2005.

Analysis of Figures 3 and 4 confirms the efficiency of the suggested technique and shows that the multi-component model is suitable for making a prediction for ionospheric parameters with time advance equal to 5 hours. At some time points there are rises in prediction errors (Figure 3h,i,j). In order to analyze the prediction errors, our simulation results were compared with geomagnetic data (H-component of the geomagnetic field was used), which were needed for estimating the geomagnetic perturbation intensity for the analyzed time periods (the geomagnetic perturbation intensity was estimated using the technique expounded in [Mandrikova et al. 2013b]), $E_b = \sum_a |W_{\psi} f_{b,a}|$, $W_{\psi} f_{b,a} := |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \Psi(\frac{t-b}{a}) dt$, $f \in L^2(\mathbb{R})$, $a, b \in \mathbb{R}$, $a \neq 0$. The results provided in Figure 5 illustrate the fact that prediction error increase associated with an ionospheric anomaly is observed prior to a strong earthquake in Kamchatka (20.02.2011, energy class E=14.1, the anomaly is shown in Figure 5 by the dashed line).

Testing the suggested technique on TEC data has confirmed possible connection between ionospheric anomalies and seismic events in Kamchatka [Mandrikova et al. 2013c]. Such effects in the ionosphere with a duration from several minutes to several hours arising in periods of increased seismic activity in different seismically active regions are also specified in a number of earlier works [Afraimovich et al. 2000, Afraimovich et al. 2001, Nakamura et al. 2007, Nakamura et al. 2009,

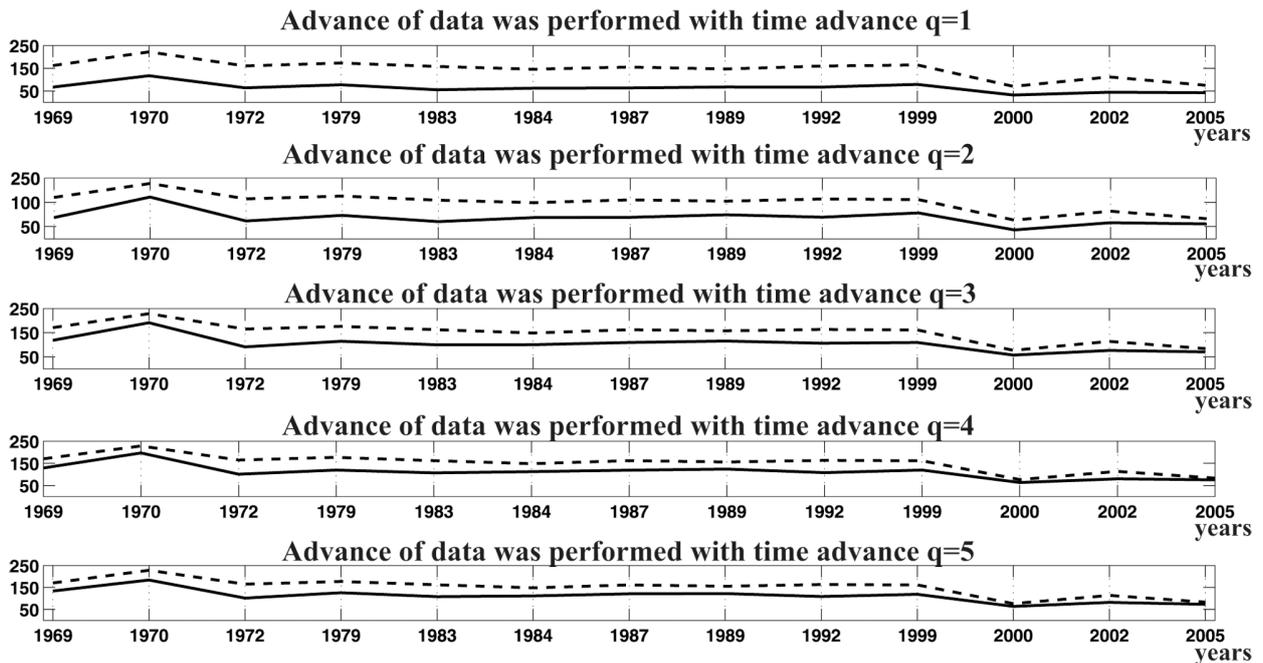


Figure 4. Simulation error for foF2 data. Solid line: using the wavelet transform; dashed line: without using the wavelet transform. Solid line is $E_{win}^{MCM} = \sum_{k=1}^K \sum_{j=1}^3 (a_{j,k}^1)^2$, K is the number of observations. Dashed line is $E_{win}^{ARIMA} = \sum_{k=1}^K e_k^2$, e_k is ARIMA model errors, K is the number of observations.

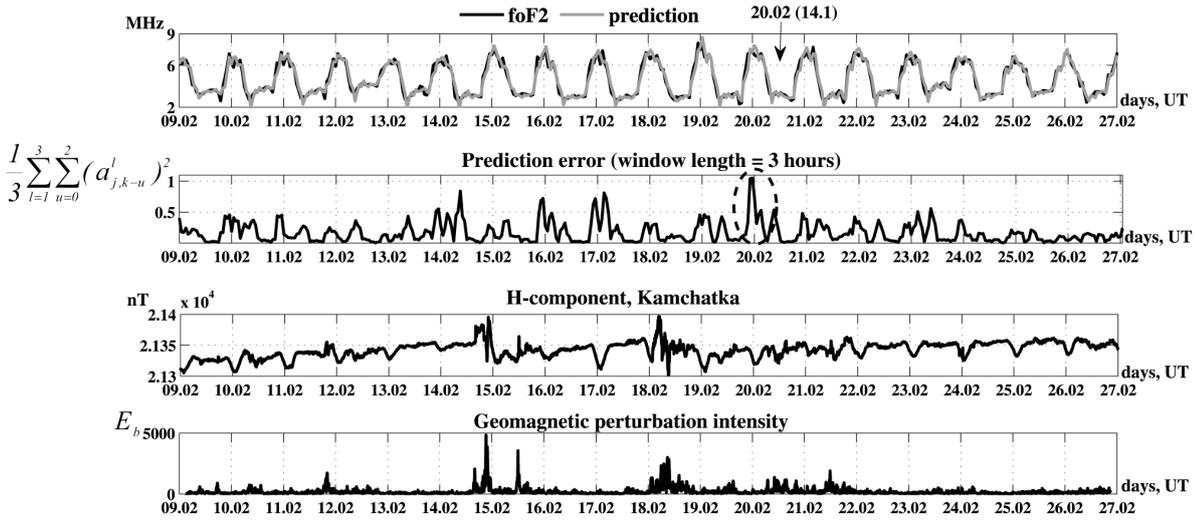


Figure 5. Detailed analysis of foF2 data for the period 09.02.2011–27.02.2011: prediction was performed using model (5) with time advance $q=1$; the arrow indicates a seismic event.

Klimenko et al. 2012, Watthanasangmechai et al. 2012]. The issues connected with the nature and origin of such effects are still to be studied more thoroughly. During the increased geomagnetic activity we can also observe a prediction error rise. Ionospheric anomalies extracted in periods of strong and moderate magnetic storms were large-scale ones and their duration was equal to one day or even longer, which is caused by significant changes of electronic concentration level relative to the background. They reached their maximum intensity during minor geomagnetic perturbations. It was also possible to observe small-scale anomalies connected

with local fluctuations of the electron density of the ionosphere. The results described in the paper agree with those published before [Mendillo 2006, Roux 2012, Danilov 2013]. Such a complex behavior of the ionosphere is illustrated in Figures 5, 7 and 8. Ionospheric anomalies discovered in periods of solar flares and ejection of coronal material mostly had small scales and arose against a background of weakly disturbed and quiet geomagnetic field. The morphology and physics of such a behavior are expounded in review papers [Prolls 1995, Rees 1995, Buonsanto 1999, Danilov and Laštovička 2001, Danilov 2013]. The statistical analysis

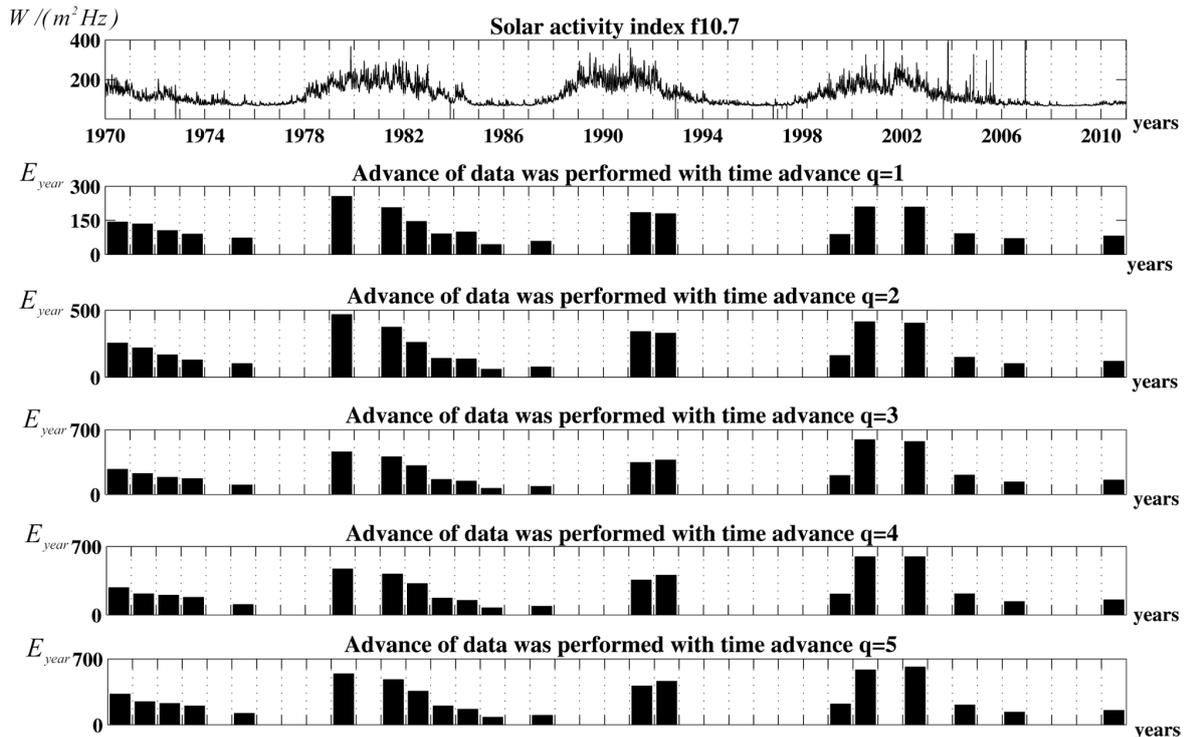


Figure 6. Prediction error estimates for foF2 data for different years. $E_{year} = \sum_{k=1}^K \sum_{j=1}^3 (a_{j,k}^1)^2$ values are indicated on the axes oy, $K=384$ is the number of observations per year.

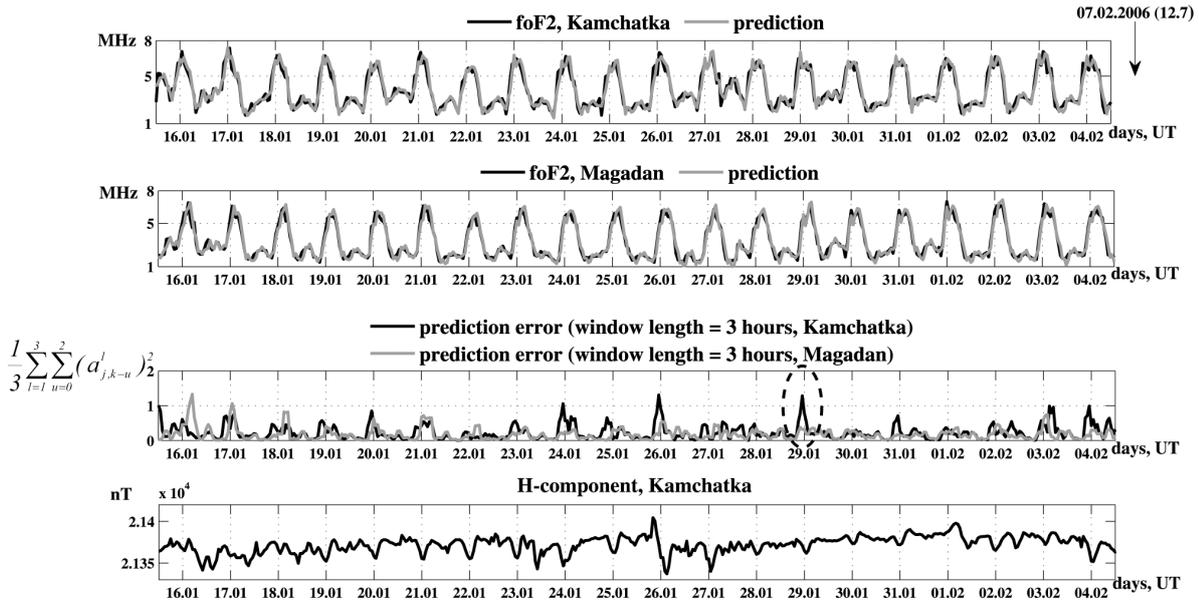


Figure 7. Simulation results for foF2 data from Kamchatka and Magadan for the period 16.01.2006–04.02.2006: prediction was made using model (5) with $q = 1$; the arrow indicates a seismic event in Kamchatka (energy class $k = 12.7$).

of foF2 data for different years has shown a substantial connection between anomaly intensity and anomaly occurrence frequency in the ionosphere, on the one hand, and solar and geomagnetic activities, on the other hand. This conclusion agrees with works [Afraimovich et al. 2000, Afraimovich et al. 2001, Roux 2012]. During the periods of increased solar activity the general error level goes up; when anomalies occur the deviation from the background level rises dramatically. We can confirm this by the results shown in Figure 6.

The joint analysis of the simulation results (data from Kamchatka and Magadan) reveals some general trend. The H-component of the geomagnetic field (Paratunka observatory) is shown in Figure 7 with the aim of analyzing the geomagnetic activity. When the

geomagnetic activity is increased, the prediction error rises for data from Kamchatka and Magadan. At some time points the prediction errors for data from Kamchatka exceed significantly those for data from Magadan (this period is shown in Figure 7 by the dashed line), which is probably connected with seismic events in Kamchatka.

Due to wavelet transform properties [Mandrikova et al. 2010] and Remark 1 given earlier in this paper, we performed ARIMA model identification to approximate the components $f2^{-3}t = \sum_k c_{-3,k} \varphi_{-3,k}(t)$ and $g2^{-3}t = \sum_k d_{-3,k} \psi_{-3,k}(t)$. The approximation was made without data reconstruction to the original scale $j = 0$. Figure 8 illustrates the simulation of these components. The comparison with Kp -index [Helmholtz-Centre Potsdam: In-

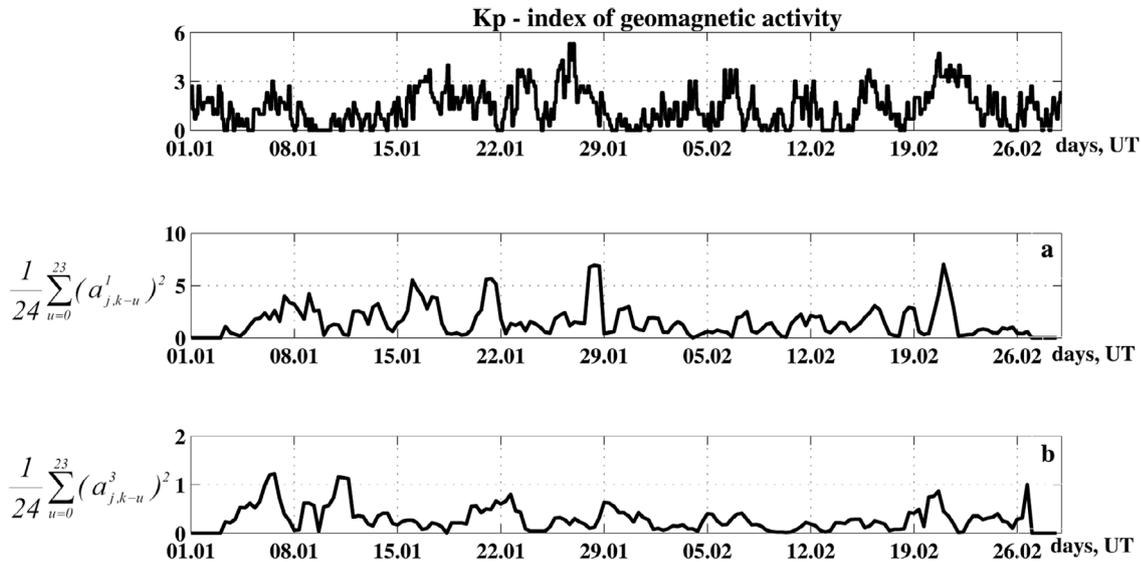


Figure 8. Simulation results for foF2 data from Kamchatka for the period 01.01.2006–26.02.2006. (a) Prediction error for the component $f2^{-3}t$ in the sliding time window equal to 24 hours; (b) prediction error for the component $g2^{-3}t$ in the sliding window equal to 24 hours.

dices of Global Geomagnetic Activity; Bartels 1957a, Bartels 1957b, Joselyn 1970, Mayaud 1980, Golovkov et al. 1989, Rangarajan 1989, Menvielle and Berthelier 1991, Sucksdorff et al. 1991, Menvielle et al. 1995, Siebert and Meyer 1996, Jach et al. 2006, Nowozynski 2007, Xu et al. 2008] values shows the connection between the prediction error and the geomagnetic activity level, which accords with the results discussed above.

4. Conclusions

The suggested technique of ionospheric parameter simulation and analysis based on combining the wavelet transform with ARIMA models was used to approximate trend in critical frequency data from Kamchatka and Magadan. The experiments have confirmed the efficiency of the suggested technique and its advantages for studying regular changes in ionospheric parameters and making predictions with time advance equal to 5 hours.

The residual error estimates of the models were employed for discovering ionospheric anomalies during ionospheric perturbations with a duration from several dozens of minutes to several hours. The changes in ionospheric parameters have different scales and arise prior to and during strong earthquakes in Kamchatka.

The variation analysis of the parameters during the periods with different solar activity and comparison of simulation results with geomagnetic data have shown the connection between anomaly intensity and anomaly occurrence frequency, on the one hand, and solar and geomagnetic activities, on the other hand. During the periods of increased geomagnetic activity there is a significant rise of deviation from the background level due to anomaly occurrence.

Acknowledgements. The data for research were kindly provided by Kamchatka branch of Geophysical Survey of the Russian Academy of Sciences (GS RAS). The paper and the research are supported by the grant of the Russian Science Foundation no. 14-11-00194 (MOV) and Contract no. 02.G25.31.0058 dated 12.02.2013 (Ministry of Education and Science of the Russian Federation).

References

- Afraimovich, E.L., E.A. Kosogorov, K.S. Palamarchouk, N.P. Perevalova and A.V. Plotnikov (2000). The use of GPS arrays in detecting the ionospheric response during rocket launchings, *Earth, Planets, and Space*, 52 (11), 1061-1066.
- Afraimovich, E.L., N.P. Perevalova, A.V. Plotnikov and A.M. Uralov (2001). The shock-acoustic waves generated by the earthquakes, *Annales Geophysicae*, 19 (4), 395-409.
- Al-Kasasbeh, R.T. (2004). Wavelet-based method for EEG artifacts classification, In: *Proceedings of IFMBE*, Kuala Lumpur, Malaysia, 157-159.
- Bakhshi, A.D., S. Bashir, S.I. Shah and M.A. Maud (2013). Non-linear trend estimation of cardiac repolarization using wavelet thresholding for improved T-wave alternans analysis, *Digital Signal Processing*, 23 (4), 1197-1208.
- Bartels, J. (1957a). The technique of scaling indices K and Q of geomagnetic activity, *Annals of the International Geophysical Year*, 4, 215-226.
- Bartels, J. (1957b). The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields, *Annals of the International Geophysical Year*, 4, 227-236.
- Basseville, M., and I.V. Nikiforov (1993). *Detection of Abrupt Changes - Theory and Application*, Prentice-Hall, Inc., Englewood Cliffs, N.J.; ISBN 0-13-126780-9.
- Beltrán, J.R., and J. Ponce de León (2010). Estimation of the instantaneous amplitude and the instantaneous frequency of audio signals using complex wavelets, *Digital Signal Processing*, 90 (12), 3093-3109.
- Box, G., and G. Jenkins (1970). *Time series analysis: Forecasting and control*, San Francisco, Holden-Day.
- Buonsanto, M.J. (1999). Ionospheric storms - A review, *Space Science Reviews*, 88, 563-601.
- Castillo, E., D.P. Morales, G. Botella, A. García, L. Parrilla and A.J. Palma (2013). Efficient wavelet-based ECG processing for single-lead FHR extraction, *Digital Signal Processing*, 23 (6), 1897-1909.
- Danilov, A.D., and J. Laštovi ka (2001). Effects of geomagnetic storms on the ionosphere and atmosphere, *International Journal of Geomagnetism and Aeronomy*, 2, 209-224.
- Danilov, A.D. (2013). Ionospheric F-region response to geomagnetic disturbances. *Advances in Space Research*, 52 (3), 343-366.
- Daubechies, I. (2001). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics.
- Donoho, D., and I. Johnstone (1998). Minimax estimation via wavelet, *Annals of Statistics*, 879-921.
- Erdoğan, H., and N Arslan (2009). Identification of vertical total electron content by time series analysis, *Digital Signal Processing*, 19 (4), 740-749.
- Ezquer, R.G., J.L. López, L.A. Scidá, M.A. Cabrera, B. Zolesi, C. Bianchi, M. Pezzopane, E. Zuccheretti and M. Mosert (2014). Behaviour of ionospheric magnitudes of F2 region over Tucumán during a deep solar minimum and comparison with the IRI 2012 model predictions, *Journal of Atmospheric and Solar-Terrestrial Physics*, 107, 89-98.
- Golovkov, V.P., V.O. Papitashvili and N.E. Papitashvili (1989). Automated calculation of the K-indices using the method of natural orthogonal components, *Ge-*

- omagnetizm i Aeronomiia, 29, 667-670 (in Russian).
- Gwal, A.K., S. Rubeena, P. Gopal and J.K. Santosh (2012). Study of seismic precursors by wavelet analysis, *Research Journal of Engineering Sciences*, 1 (4), 48-52.
- Hafez, A.G., T.A. Khan and T. Kohda (2010). Clear P-wave arrival of weak events and automatic onset determination using wavelet filter banks, *Digital Signal Processing*, 20 (3), 715-723.
- Hafez, A.G., and E. Ghamry (2011). Automatic detection of geomagnetic sudden Commencements via Time Frequency Clusters, *Advances Space Research*, 48 (9), 1537-1544.
- Hamoudi, M., N. Zaourar, R. Mebarki, L. Briquieu and M. Parrot (2009). Wavelet analysis of ionospheric disturbances, In: *EGU General Assembly 2009, Geophysical Research Abstracts*, vol. 11, EGU2009-8523.
- Han, X.H., and X.M. Chang (2013). An intelligent noise reduction method for chaotic signals based on genetic algorithms and lifting wavelet transforms, *Information Sciences*, 218, 103-118.
- Helmholtz-Centre Potsdam: Indices of Global Geomagnetic Activity. Helmholtz-Centre Potsdam GFZ, German Research Centre for Geosciences; <http://www.gfz-potsdam.de/en/section/earths-magnetic-field/services/kp-index/>
- Huang, J., M. Korolkiewicz, M. Agrawal and J. Boland (2013). Forecasting solar radiation on an hourly time scale using a Coupled AutoRegressive and Dynamical System (CARDS) model, *Solar Energy*, 87, 136-149.
- Jach, A., P. Kokoszka, J. Sojka and L. Zhu (2006). Wavelet-based index of magnetic storm activity, *Journal of Geophysical Research*, 111, A09215; doi:10.1029/2006 ja011635.
- Joselyn, J.A. (1970). A real-time index of geomagnetic activity, *Journal of Geophysical Research*, 75, 2777-2780.
- Kay, S., and S. Marple (1981). Spectrum analysis - a modern perspective, *Proceedings of the IEEE*, 69 (11), 1380-1419.
- Klimenko, M.V., V.V. Klimenko, I.E. Zakharenkova and I.V. Karpov (2012). Modeling of local disturbance formation in the ionosphere electron concentration before strong earthquakes, *Earth, Planets and Space*, 64 (6), 441-450.
- Li, H., T. Yi, M. Gu and L. Huo (2009). Evaluation of earthquake-induced structural damages by wavelet transform, *Progress in Natural Science*, 19 (4), 461-470.
- Liming, H.E., W.U. Lixin, L.I.U. Shanjun and M.A. Baodong (2011). Seismo-ionospheric anomalies detection based on integrated wavelet, In: *Geoscience and Remote Sensing Symposium (IGARSS)*, 2011 IEEE International (July 24-29, 2011), 1834-1837; doi:10.1109/IGARSS.2011.6049479.
- Liu, L., W. Wan, M.-L. Zhang and B. Zhao (2008). Case study on total electron content enhancements at low latitudes during low geomagnetic activities before the storms, *Annales Geophysicae*, 26 (4), 893-903.
- Mallat, S. (1999). *A Wavelet tour of signal processing*, Academic Press.
- Mandrikova, O.V., and V.V. Bogdanov (2007). Multi-structural models of geophysical signals, In: *8th International Conference on pattern recognition and image analysis: new information technologies (PRIA-8-2007, October 8-12, 2007): Proceedings*, Yoshkar-Ola, vol. 3, 81-84.
- Mandrikova, O.V., V.V. Geppener, T.S. Goreva and D.M. Klienskiy (2010). Multicomponent simulation and forecasting of a complex natural signal, In: *International conference "Intelligent Information Processing" IIP-8 (Cyprus, Paphos, October 17-23, 2010)*.
- Mandrikova, O.V., I.S. Solovjev, V.V. Geppener and D.M. Klienskiy (2011). New wavelet-based approach intended for the analysis of subtle features of complex natural signals, *Pattern Recognition and Image Analysis: New Information Technologies*, 21 (2), 293-296.
- Mandrikova, O.V., Yu.A. Polozov and T.L. Zaliaev (2012a). Methods of analysis and interpretation of ionospheric critical frequency foF2 data based on wavelet transform and neural networks, In: *European Seismological Commission 33rd General Assembly (GA ESC 2012, August 19-24, 2012)*; http://www.esc2012-moscow.org/files/GA_ESC_2012-Program_13.08.2012.xls.
- Mandrikova, O.V., Yu.A. Polozov, V.V. Bogdanov and E.A. Zhizhikina (2012b). Method of detection of abnormal features in ionosphere critical frequency data on the basis of wavelet transformation and neural networks combination, *Journal of Software Engineering and Applications*, 5 (12B), 181-187; doi:10.4236/jsea.2012.512b035.
- Mandrikova, O.V., I.S. Solovjev, V.V. Geppener and D.M. Klienskiy (2012c). Analyzing Subtle Features of Natural Time Series by Means of a Wavelet-Based Approach, *Pattern Recognition and Image Analysis*, 22 (2), 323-332.
- Mandrikova, O.V., I.S. Solovyev, V.V. Geppener, D.M. Klienskiy and R.T. Al-Kasasbeh (2013a). Analysis of the Earth's magnetic field variations on the basis of a wavelet-based approach, *Digital Signal Processing*, 23, 329-339.
- Mandrikova, O.V., V.V. Bogdanov and I.S. Solovjev (2013b). Wavelet Analysis of Geomagnetic Field Data, *Geomagnetism and Aeronomy*, 53 (2), 268-273.
- Mandrikova, O.V., N.V. Glushkova and I.V. Zhivetiev

- (2013c). Method of modelling and forecasting of ionospheric data based on the combination of wavelet transform and autoregressive-integrated moving average model, In: *Solar terrestrial relations and physics of earthquakes precursors*, VI Intern. Conf., v. Paratunka Kamchatka region (September 9-13, 2013), 29-30.
- Maruyama, T., T. Tsugawa, H. Kato, A. Saito, Y. Otsuka and M. Nishioka (2011). Ionospheric multiple stratifications and irregularities induced by the 2011 off the Pacific coast of Tohoku Earthquake, *Earth, Planets and Space*, 63, 869-873.
- Mayaud, P.N. (1980). Derivation, Meaning and Use of Geomagnetic Indices, Geophysical Monograph 22, American Geophysical Union, Washington D.C.
- Mendillo, M. (2006). Storms in the ionosphere: Patterns and processes for total electron content, *Reviews of Geophysics*, 44, RG4001; doi:10.1029/2005RG000193.
- Menvielle, M., and A. Berthelier (1991). The K-derived planetary indices: description and availability, *Reviews of Geophysics*, 29 (3), 415-432.
- Menvielle, M., N. Papitashvili, L. Häkkinen and C. Sucksdorff (1995). Computer production of K indices: review and comparison of methods, *Geophysical Journal International*, 123, 866-886.
- Mitra, S.K. (1998). *Digital Signal Processing: A Computer-Based Approach*, New York, McGraw-Hill.
- Nakamura, M.I., T. Maruyama and Y. Shidama (2007). Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan, *Earth Planets Space*, 59, 1231-1239.
- Nakamura, M., T. Maruyama and Y. Shidama (2009). Using a neural network to make operational forecasts of ionospheric variations and storms at Kokubunji, Japan, *Journal of the National Institute of Information and Communications Technology*, 56, 391-406.
- Nowozynski, K. (2007). Calculate geomagnetic activity K indices using the Adaptive Smoothing method; <http://www.intermagnet.org>.
- Percival, D.B., and A.T. Walden (2000). *Wavelet Methods for Time Series Analysis* (Cambridge Series in Statistical and Probabilistic Mathematics), Cambridge University Press.
- Pervak, S., V. Choliy and V. Taradiy (2008). Spectral Analysis of the Ionospheric Irregularities from GPS Observations, In: *Proc. of the 17th Annual Conf. of Doctoral Students - WDS 2008* (Prague, June 3-6, 2008), Part II: Physics of Plasmas and Ionized Media, 189-191.
- Prolls, G. (1995). Ionospheric F-region storms, In: H. Volland (ed.), *Handbook of Atmospheric Electrodynamics*, CRC Press, Boca Raton, vol. 2, 195-248.
- Rangarajan, G.K. (1989). Indices of geomagnetic activity, In: J.A. Jacobs (ed.), *Geomagnetism*, 3, Academic Press, London, 323-384.
- Rees, D. (1995). Observations and modeling of ionospheric and thermospheric disturbances during major geomagnetic storms: A review, *Journal of Atmospheric and Terrestrial Physics*, 57, 1433.
- Roux, S.G. (2012). A scale dependent wavelet-based cross-correlation analysis, *Journal of Atmospheric and Solar-Terrestrial Physics*, 90-91, Special Issue, 186-197.
- Siebert, M., and J. Meyer (1996). Geomagnetic Activity Indices, In: W. Dieminger et al. (eds.), *The Upper Atmosphere*, Springer, Berlin Heidelberg, 887-911.
- Sucksdorff, C., R. Pirjola and L. Häkkinen (1991). Computer production of K-values based on linear elimination, *Geophysical Transactions*, 36, 333-345.
- Vargas-Bracamontes, D.M., F. Alejandro Nava and G.A. Reyes-Dávila (2009). Time-scale wavelet patterns related to the 1998-1999 eruptions of the Colima volcano, and their possible implications for eruption forecasting, *Journal of Volcanology and Geothermal Research*, 184 (3-4), 271-284.
- Wang, J., Y. Xie, C. Zhu and X. Xu a (2011). Solar Radiation Prediction Based on Phase Space Reconstruction of Wavelet Neural Network, *Procedia Engineering*, 15, 4603-4607.
- Wattanasangmechai, K., P. Supnithi, S. Lerkvaranyu, T. Tsugawa, T. Nagatsuma and T. Maruyama (2012). TEC prediction with neural network for equatorial latitude station in Thailand, *Earth, Planets and Space*, 64 (6), 473-483.
- Xu, Z., L. Zhu, J. Sojka, P. Kokoszka and A. Jach (2008). An assessment study of the wavelet-based index of magnetic storm activity (WISA) and its comparison to the Dst index, *Journal of Atmospheric and Solar-Terrestrial Physics*, 70, 1579- 1588.
- Zhang, M., L. Fu, H. Li and G. Wang (2008). Harmonic retrieval in complex noises based on wavelet transform, *Digital Signal Processing*, 18 (4), 534-542.
- Zhao, X., B. Ning, L. Liu and G. Song (2014). A prediction model of short-term ionospheric foF2 based on AdaBoost, *Advances in Space Research*, 53 (3), 387-394.

Corresponding author: Riad Taha Al-Kasasbeh,
Al-balqa Applied University, Amman, Jordan;
email: rjordanjo@mail.ru.

© 2015 by the Istituto Nazionale di Geofisica e Vulcanologia. All rights reserved.