Big Data for the Magnetic Field Variations in Solar–Terrestrial Physics and Their Wavelet Analysis, Chapter 19 in Knowledge Discovery in Big Data from Astronomy and Earth Observati…

Chapter · April 2020

CITATIONS
0

READS
543

3 authors:

Bozhidar Atanasov Srebrov
Bulgarian Academy of Sciences
69 PUBLICATIONS **55** CITATIONS
SEE PROFILE

Ognyan Kounchev
Bulgarian Academy of Sciences
136 PUBLICATIONS **554** CITATIONS
SEE PROFILE

Georgi Simeonov
Institute of Mathematics and Informatics at Bulgarian Academy of Sciences (IMI-B…
22 PUBLICATIONS **20** CITATIONS
SEE PROFILE

Knowledge Discovery in Big Data from Astronomy and Earth Observation

AstroGeoInformatics

EDITED BY
PETR ŠKODA
FATHALRAHMAN ADAM

ELSEVIER

Knowledge Discovery in Big Data from Astronomy and Earth Observation

This page intentionally left blank

Knowledge Discovery in Big Data from Astronomy and Earth Observation

AstroGeoInformatics

Edited by

Petr Škoda
Stellar Department
Astronomical Institute of the Czech Academy of Sciences
Ondřejov, Czech Republic

Fathalrahman Adam
Earth Observation Center
German Remote Sensing Data Center
DLR German Aerospace Center
Wessling, Germany

ELSEVIER

Elsevier
3251 Riverport Lane
St. Louis, Missouri 63043

Knowledge Discovery in Big Data from Astronomy and Earth Observation ISBN: 978-0-12-819154-5

Publisher: Candice Janco
Acquisitions Editor: Marisa LaFleur
Editorial Project Manager: Andrea Dulberger
Production Project Manager: Sreejith Viswanathan
Designer: Alan Studholme

CHAPTER 19

Big Data for the Magnetic Field Variations in Solar-Terrestrial Physics and Their Wavelet Analysis

BOZHIDAR SREBROV, ASSOC PROF, DR • OGNYAN KOUNCHEV, PROF, DR • GEORGI SIMEONOV, ASSISTANT

19.1 INTRODUCTION TO BIG MAGNETIC DATA IN SOLAR-TERRESTRIAL PHYSICS

According to the popular understanding of Big Data, in solar-terrestrial physics this would refer to large amounts of measured data satisfying the following properties:

1. their source is not homogeneous,
2. they have large dimension,
3. the size and the format of the data excess the capacity of the conventional tools to effectively capture, store, manage, analyze, and exploit them, and finally,
4. the data have a complex and dynamic relationship.

In the present, public and private institutions are increasingly facing Big Data challenges, and a wide variety of techniques have been developed and adapted to aggregate, manipulate, organize, analyze, and visualize them. The techniques currently applied in big astronomical and Earth observation data usually draw from several fields, including statistics, applied mathematics, and computer science, and institutions that intend to derive value from the data should adopt a flexible, reliable, and multidisciplinary approach. In particular, utilizing Big Data in solar-terrestrial physics and related analytics is expected to improve the performance of prediction mechanisms for extreme geomagnetic events as geomagnetic storms.

Big magnetic data in solar-terrestrial physics are collected at a tick-by-tick level, i.e., at higher frequency. Let us remark that the standard registrations of the geomagnetic field variations have an order of about 0.1–10 mHz (i.e., of periods 1.66 min until 2.77 hours). The analysis of geomagnetic data at higher frequency (here, high frequency is understood from the point of view of the standards in geomagnetism) uncovers the complex structure of irregularities and roughness (i.e., multifractal phenomena) due to huge amounts of microstructure noise. The nonhomogeneity characterized by multifractal phenomena is caused by a large number of instantaneous changes in the geomagnetic field due to geomagnetic storms and various sources of noises as, for example, the low-frequency plasma instability modes. Therefore, mining big geomagnetic data needs to intelligently extract information conveyed at different frequencies. At the present moment the registration of geomagnetic signals has the maximal frequency of seconds; however, the majority of the geomagnetic or ionosound data are still collected at maximal frequency 4 or 5 min, as we will see below.

With the classic assumption of data mining and statistical modeling, data are generated by certain unknown functions representing signals plus random noise (see, e.g., one of the bibles of modern pattern recognition (Bishop, 2006) and the references therein). Analyzing big magnetic data in solar-terrestrial physics is equivalent to extracting the systematic patterns (i.e., approximate the unknown function) conveyed in the data from noise, which is the standard approach of the classic signal processing theory brilliantly presented in the famous handbook Oppenheim and Schafer (2010).

The situation in the modeling of geomagnetic fields and in particular geomagnetic storms falls in the framework of analyzing jump events in Big Data, which has been recently thoroughly studied by various researchers. In particular, it has been studied in the context of financial time series, in the nice research of Sun and Meinl (2012); see also Sun and Yu (2015).

In geomagnetism, jumps are often caused by some unexpected large geomagnetic storms or by predictable changes in the sectorial structure of the interplanetary magnetic field (IMF). Traditional linear denoising methods (e.g., moving average) usually fail to capture these jumps accurately as these linear methods tend to blur them. On the other hand, nonlinear filters are

Knowledge Discovery in Big Data from Astronomy and Earth Observation. https://doi.org/10.1016/B978-0-12-819154-5.00031-X

not appropriate to smooth out these jumps sufficiently, because the patterns extracted by nonlinear filters are not stationary to present long-run dynamic information. The situation is rather similar to that in the case of financial time series, as noted in Sun and Meinl (2012) and in the references therein.

The present chapter deals with research in the global structure of the magnetic phenomena in solar-terrestrial physics, by systematically applying *wavelet analysis* to the available Big Data. We use the *continuous wavelet transform* (CWT) to analyze the structure and the dynamics of geomagnetic storms. The wavelet method is one of a number of multifractal spectrum computing methods and has proven to be a reliable in signal processing, as established in the classical monograph Mallat (2009). It has proven to be very suitable for analyzing time series analysis – for example, smoothing, denoising, and jump detection in diverse areas of science, finance, and economics; see, e.g., Percival and Walden (2000), Gençay et al. (2002), Hubbard (1998), Meinl and Sun (2012), Kounchev (2001).

An important advantage of the wavelet method in analyzing magnetic phenomena in solar-terrestrial physics, where factors of different scales interfere, is that it performs a multiresolution analysis (MRA), in other words, it allows us to analyze the data at different scales (each one associated with a particular frequency passband) at the same time. In this way, wavelets enable us to identify single events truncated in one frequency range as well as coherent structures across different scales. Many recent studies have applied wavelet methods in mining geophysical and geomagnetic data; some recent references are Xu et al. (2008), Xu (2011), Klausner et al. (2016a, 2016b, 2017), Schnepf et al. (2016).

Let us recall that there are in fact two versions of wavelet analysis: CWT and the *discrete wavelet transform* (DWT). In both of them, there is a large variety of wavelet functions by which one may perform the signal decomposition. A common approach in choosing the wavelet function is to use the shortest wavelet filter that can provide reasonable results; see, e.g., Percival and Walden (2000). The main challenge in performing wavelet analysis is to determine the combination of wavelet function, level of decomposition, and threshold rule to reach an optimal smoothness that generally improves the performance of classic models after denoising the data. We have provided a short appendix at the end of the chapter which explains briefly the technical details of wavelet analysis and our choice of wavelet function.

Another major advantage of wavelet analysis (either discrete or continuous) which makes it well adapted to the purposes of Big Data is that, similar to the classical Fourier analysis, there exist very fast algorithms for processing large amounts of data (Mallat, 2009).

Due to space restrictions, in the present chapter we have limited ourselves to just preliminary research of the correlations which exist among the data in the frequency domain (the coefficients of the CWT of the time series). Although the results obtained are very interesting and promising, we have not provided a more detailed statistical analysis (as, e.g., in Sun and Meinl, 2012, Sun and Yu, 2015) which would uncover much deeper and interesting connections between the different factors playing a role in the geomagnetic phenomena. We leave such analysis for follow-up research, which is in progress.

The structure of the chapter is as follows: in Section 19.2 we recall the big picture of the solar-terrestrial mechanism – in quiet and in stormy periods. This has to give an idea to the unexperienced reader about the complexity of the manifestation of a geomagnetic storm and the necessity to apply modern methods of machine/deep learning to Big Data, for a deeper understanding of the phenomena. In Section 19.2.3 we provide a short description of the different components of the ground geomagnetic field provided by Chapman's analysis – the global index D_{st} and local disturbance index DS, which show how complicated the structure of the geomagnetic field on Earth is. In Section 19.2.6 we provide basic information about two famous geomagnetic storms. In Section 19.2.7 we provide basic information about the different types of data used for analyzing the big picture of the magnetic phenomena in solar-terrestrial physics. In Section 19.3 we provide the results of the application of CWT to the three main types of data in the form of simultaneous time series (IMF data, ionospheric TEC data, geomagnetic data), in some quiet days and in days of geomagnetic storms. For every experiment carried out, we provide some empirical observations on the correlations between the CWT coefficients (the frequency domain) for simultaneous time series. In a forthcoming study we will apply the methods of statistical analysis for a more rigorous analysis of these observed correlations between the different types of data.

The large variety of data used, from both solar astronomy and Earth observations, makes our research a contribution to the newly developing area of *Astro-GeoInformatics*.

19.2 MECHANISM OF GENERATING STRONG GEOMAGNETIC STORMS (LONG-PERIOD GEOMAGNETIC FIELD VARIATIONS)

The main purpose of the present study is to analyze different sources of data in solar-terrestrial physics, and to discover correlations between the (relatively) high-frequency geomagnetic variations (wave packages with short period about 0.1 mHz until 10 mHz) which happen not only during geomagnetic storms but also in more quiet periods.

Before carrying out such an analysis we will provide a short outline of the global context of the geomagnetic picture in the solar-terrestrial interactions.

19.2.1 The Big Picture of Solar-Terrestrial Physics – Quiet and Disturbed Geomagnetic Phenomena

In the present section we will outline the big picture of the influence of the solar activity on the IMF, the ionospheric parameters, and the (ground) geomagnetic field.

First of all, we speak about events happening in the magnetosphere, i.e., in the region of space surrounding Earth where the dominant magnetic field is the magnetic field of Earth, rather than the magnetic field of interplanetary space. The magnetosphere is formed by the interaction of solar wind with the Earth's magnetic field. The Earth's magnetic field is continually changing as it is buffeted by the solar wind.

1. In quiet periods the Sun emits a flow of particles (solar wind) having a relatively constant intensity and speed, which start with appr. 50 km/sec and is accelerated to about approximately 360 km/sec close to Earth (which is obtained by the Parker model). The speed is accelerated by a mechanism which is still not known but most probably due to energy transfer in the solar wind.
2. In a disturbed state, the hyperactivity of the Sun causes coronal mass ejection (CME), which increases the amount of charged particles; they have a macrospeed of about 1500–2500 km/sec when they leave the Sun; this speed decreases to about 500–700 km/sec when they approach the Earth.
3. In the interplanetary medium there is a collisionless plasma where the particles are electrons and ions (protons); the *Advanced Composition Explorer (ACE) satellite* collects the data of the *interplanetary magnetic field* (IMF) at a distance of 1.5 million km from the Earth in this plasma (see Fig. 19.1); the most important is the Z component of the IMF called B_z (which is perpendicular to the ecliptic) which influences the formation of the storm; during the strong storms the component B_z is negative.
4. After that, the flow of particles approaches the magnetosphere (about 12,000–25,000 km), which is the belt of Van Allen (mathematical figure called torus), where they are caught by the Earth's magnetic field (the Earth's dipole); this looks like a cavern but they are mainly concentrated in the equatorial domain. The so-called *ring current* is formed in the Van Allen belts. In the Van Allen belts one does not have plasma, but there is a kinetic movement of the different particles.
5. Then at about 1000 km from the Earth they enter the plasmasphere and the ionosphere, which is a plasma (partially ionized gas), and the laws describing the plasma state start to work. There are collisions among the particles. Hence, more wave effects may arise compared to the collisionless plasma. The ionosound stations acquire the values of the ionospheric parameters at these heights. For more details, see the excellent exposition in the classical monograph (Mitchner and Kruger, 1973).
6. On quiet days a regular source of disturbances in the Earth's ionosphere is also the solar terminator (at sunrise and sunset).

In Fig. 19.1 we provide the overall picture of different sources of measurements which we have used in our study.

In Fig. 19.2 we provide the magnetic field structure.

19.2.2 Geomagnetic Storms

The geomagnetic storms are by their nature *long-period geomagnetic field variations*.

Geomagnetic storms are phenomena directly related to solar activity. They result from the interaction of the magnetosphere and the ionosphere with changes in the interplanetary conditions caused by closed interplanetary magnetic structures. These structures are formed in the active centers of the Sun's chromosphere as a result of a sudden impulse ejection of the substance in the quiet solar wind called CME.

The storm is mainly characterized by a decrease in the horizontal component H of the geomagnetic field, which encompasses the entire planet during a geomagnetic storm. At low latitudes, for a long time, a current system called the *ring current* is formed around the Earth at a distance of approximately 2 to 4 Earth radii. The nowadays idea of the size and strength of the ring current system gives us reason to view it as a toroid inside the magnetosphere (in the area of Van Allen belts) and formed by particles of the solar wind; see Fig. 19.2. The geomagnetic field in the magnetosphere captures parti-

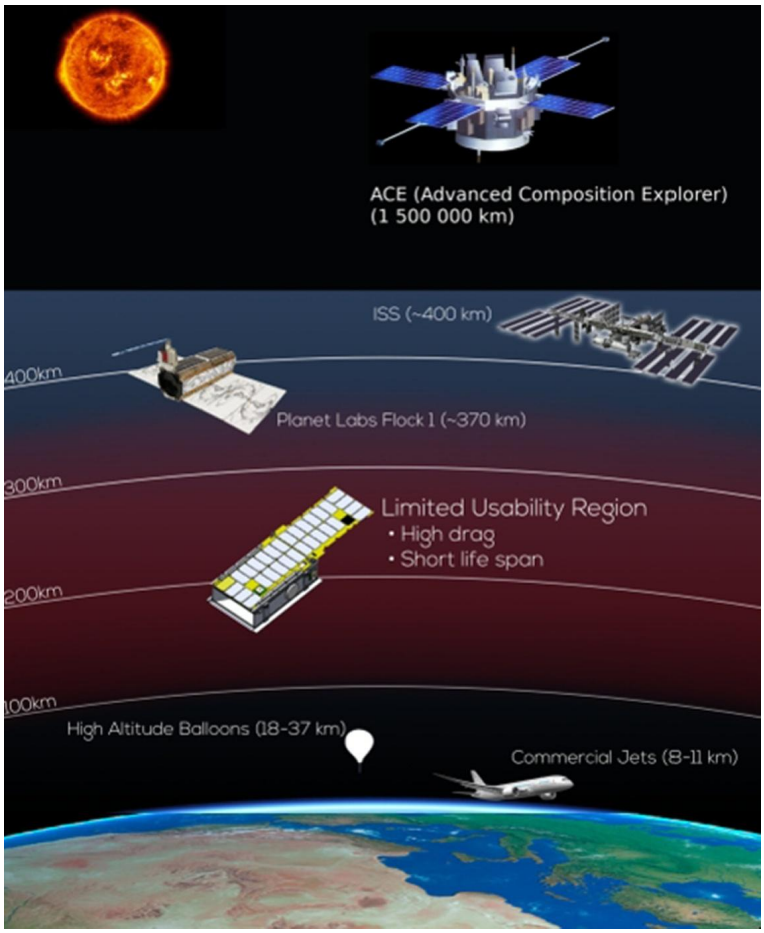

FIG. 19.1 Scheme of the interplanetary space, ionosphere, and Earth, from which we acquire data.

cles from the interplanetary plasma by creating an axisymmetric distribution of these particles in space. This leads to an amplification of the Van Allen radiation belt. The perturbed geomagnetic field associated with the so-called main storm phase in many cases is not axisymmetric. The analysis of the perturbed field shows that it can be represented as a sum of an axisymmetric part and an asymmetric component. This indicates that the proton belt is substantially asymmetric, especially in the early part of the main phase of the storm. Thus, asymmetry appears as an essential feature of ring current formation.

The intensive *proton belt* during a storm significantly deforms the magnetic field of the magnetosphere and changes its structure. In particular, we see that the domain of particle trapping approaches the Earth, in the interplanetary environment, and the polar ray oval shifts in the direction of the equator.

The intensity of the storm strongly depends on the geographic latitude. Thus, the decreasing of the horizontal component of the field during the storm in the different ground magnetic observatories is similar in shape but different in amplitude (see, e.g., Fig. 19.12 with the H component of the geomagnetic field from observatory PAG in Bulgaria, and Fig. 19.13 with the DS index of the geomagnetic field data from observatory SUA in Romania). At low latitudes, this amplitude is larger and decreases as the latitude increases. There is a difference in amplitude of the H component of the magnetic field for the different longitudes. This asymmetry of the field is related to the asymmetry of the ring current considered above.

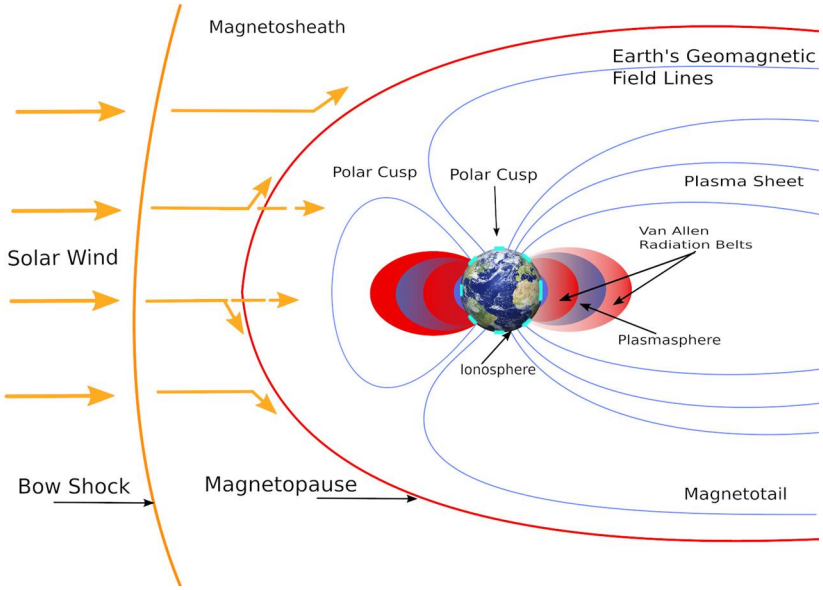

FIG. 19.2 Magnetic field structure.

Geomagnetic storms are very diverse, but they are subdivided into two main types, "standard" and "with sudden commencement." Geomagnetic storms of the second type are characterized by the absence of a pronounced sudden start. However in practice the main features of these storms during the main phase are like those of the standard type. Therefore, the initial contraction of the magnetosphere is not a prerequisite for the occurrence of the main phase of the storm.

For our purposes it is enough to mention that the type of Sun activity (solar chromospheric disturbances near or far from the solar equator) may cause different types of storms. The formation and spread of the interplanetary disturbed structure of plasma and the interplanetary magnetic field is a complex magnetohydrodynamic process that is essentially three-dimensional. The statistical analysis shows that more than 80% of the strong geomagnetic storms are associated with intensive processes in the active centers of the Sun. For average storms, this percentage is smaller and is in the range from 60% to 80%. Briefly, solar astronomy may provide important information which has to be taken into account if predictions are needed.

It happens that important Sun processes *do not have any impact* on the geomagnetic field. This can occur in a relatively small number of cases when the disturbed interplanetary structure does not affect the point of the Earth's orbit in which it is at that moment. The precise understanding of these phenomena requires the joint efforts of astronomers and geophysicists. All this indicates that knowing and predicting geomagnetic storms depends on knowing the propagation of interplanetary disturbances as a hydrodynamic process in three dimensions (Srebrov, 2003). For example, it has been found that the direction of the IMF vector B is essential to unlock the geomagnetic storm mechanism. In particular, if the Z component (denoted by B_z in the coordinate system where the axis Z is perpendicular to the ecliptic) is negative, then this indicates a possibility for a very strong geomagnetic storm.

It would be interesting for the unexperienced reader to hear about the parameters of a simulated *coronal mass ejection* (CME); see details in Srebrov (2003), where a magnetohydrodynamic model is numerically implemented: the release energy is $E = 6.0 \times 10^{22}$ J; the release mass is $M = 2.5 \times 10^{10}$ kg; the initial velocities of the ejected flow are the radial $u_d = 1500$ km/sec and the tangential is $v_d = 0$; the duration of CME is 180 seconds; and the angle of the small conic area associated with the CME is 27°; see a model with these realistic parameters simulated in Srebrov (2003). To understand how the Southern direction of the IMF is formed, we will look at the results of the computer modeling of the disturbance propagation in the interplanetary environment caused by the CME.

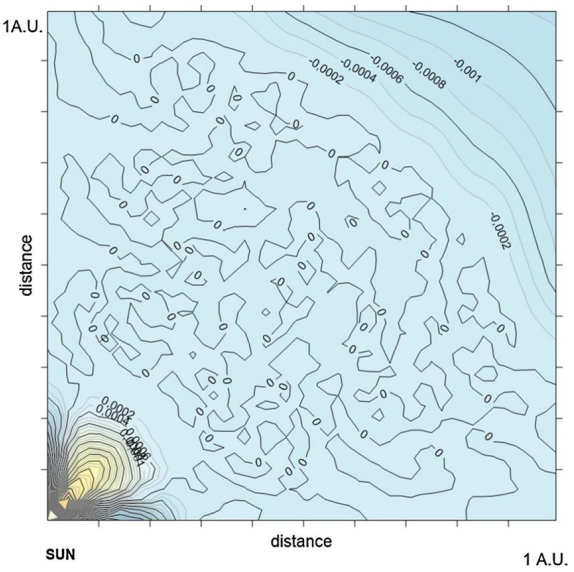

FIG. 19.3 Magnetic field components after CME – contour plot.

As more details are provided in Srebrov (2003), let us shortly describe the dynamics of the simulated disturbance propagation in the interplanetary medium. After the CME happens at $t = 0$ and has duration 3 min, about 40 hours later the disturbance reaches Earth's orbit. In order to provide a deeper perspective on the disturbance caused by the CME, we show the following figures. The magnetic field components are visualized in Fig. 19.3 (see also other similar figures in Srebrov, 2003).

Let us note that Fig. 19.3 visualizes the magnetic field components on the meridional plane. The meridional plane is defined by Sun–Earth as the x axis, and Z is vertical to the ecliptic; in it the tangential velocity v is the component of the velocity along the Z axis.

Further, in Figs. 19.4, 19.5, and 19.6 we provide the radial solar wind velocity U at different times: 10, 30, 50 hours after the CME; note that the value of U in the Figures is dimensionless.

Finally, in Fig. 19.7 we provide the contour plot of the *radial* velocity at 40 hours after the CME.

In Fig. 19.8 we provide the tangential velocity V of the disturbed solar wind at 40 hours after the CME; the values of V are dimensionless.

In Fig. 19.9 we provide the contour plot of the tangential velocity 40 hours after the CME.

As concerns the visualization of the magnetic field, we refer to the paper Srebrov (2003), where detailed figures of the disturbed tangential IMF B_z are provided. In

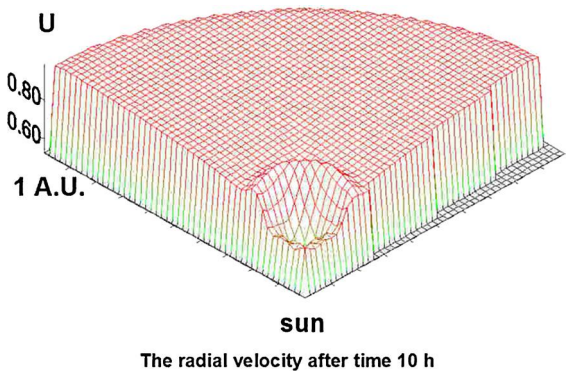

FIG. 19.4 Radial velocity after 10 h.

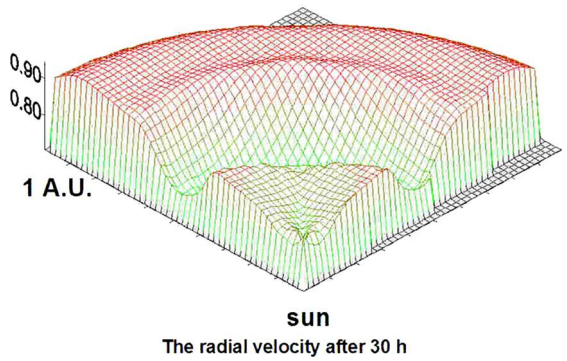

FIG. 19.5 Radial velocity after 30 h.

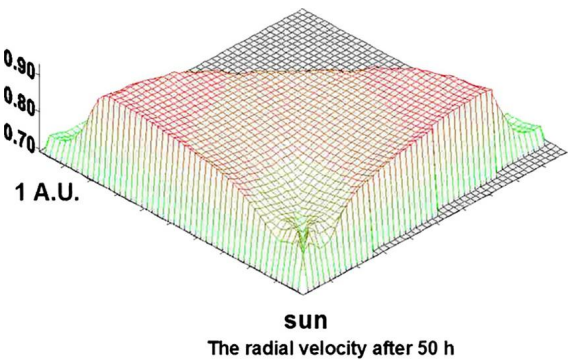

FIG. 19.6 Radial velocity after 50 h.

Srebrov (2003) it is seen that the tangential component of the magnetic field vector B_t (which in fact coincides with the B_z component in the geocentric Cartesian coordinate system) also has positive and negative values. Thus a structure is formed in which the direction of

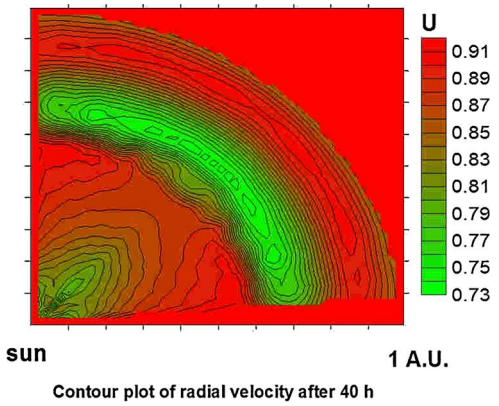

FIG. 19.7 Radial velocity after 40 h – contour plot.

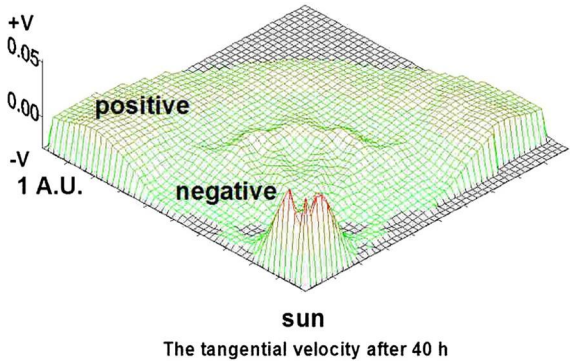

FIG. 19.8 Tangential velocity after 40 h.

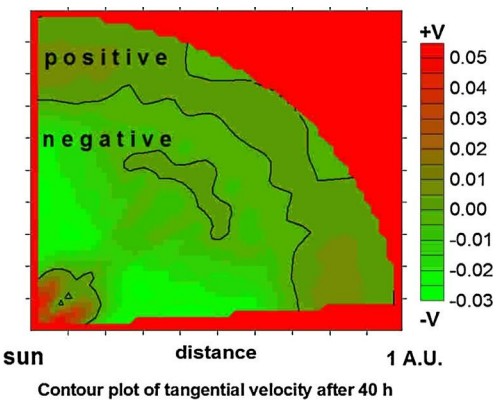

FIG. 19.9 Tangential velocity after 40 h – contour plot.

the disturbed magnetic field vector is changing to the North or to the South (in the same coordinate system mentioned above). As we know, these changes are the ones found to be the main cause of the geomagnetic storms at coupling of the disturbed IMF with the Earth's magnetosphere magnetic field. From the results of the simulations it is visible that the initial relatively closely disturbed conical area, caused by the CME, also expands in the tangential direction during the propagation, and has values which are comparable with the data gathered for example from the observations near the Earth orbit.

The above explanation justifies why the component B_z of the IMF is so important for the understanding of geomagnetic storms.

19.2.3 Ground Geomagnetic Field, and Geomagnetic Activity Index During a Storm

One of the classical models of geomagnetic storm is the Chapman model (Akasofu and Chapman, 1972). It describes the disturbance of the (ground) geomagnetic field variation during a geomagnetic storm.

According to Chapman's analysis (Akasofu and Chapman, 1972), if the time $t = 0$ denotes the sudden start of the storm, then at a certain time point t the disturbed magnetic field vector D measured at a point on Earth's surface (with components denoted by $D(H)$, $D(D)$, and $D(Z)$) can be expressed in a Fourier series expansion as follows:

$$D(\theta, \varphi, t) = c_0(\theta, t) + \sum_{n=-\infty}^{\infty} c_0(\theta, t) \sin(n\varphi + \alpha_n(\theta, t)),$$

where θ is the complement of the geomagnetic latitude to 90°, φ is the geomagnetic longitude, and α_n is the phase angle. The first and second terms in this expression represent the axially symmetrical component of the dipole axis and the asymmetric part of the disturbed field that varies with the longitude φ. These components are referred to as *storm-time variation* (D_{st}) and *local time-dependent disturbance* (DS), respectively, which contain daily regular variation S_r of type S_q and the variance from the asymmetric part of the ring current. So we can write

$$D = D_{st} + DS \tag{19.1}$$

where DS is a *geomagnetic index* that characterizes local storms. The variation of the horizontal component $D_{st}(H)$ of the D_{st} field is a function of the variable θ. It is larger at low latitudes. The declination $D_{st}(D)$ has little change during the geomagnetic storm. The vertical component $D_{st}(Z)$ also changes slightly compared to the horizontal component. Thus, the D_{st} variation is practically parallel to the Earth's surface except in the

areas of the polar cap, where the vertical component has larger positive changes. For that reason the main interest represents the horizontal component $D_{st}(H)$. Hence, we have the following formula for the *horizontal components*:

$$DS(H) = D(H) - D_{st}(H), \qquad (19.2)$$

where $D(H)$ is often denoted by $H(t)$ and is called *the horizontal component* of the field.

By neglecting $D_{st}(D)$ and $D_{st}(Z)$ as *small quantities*, it is apparent from the above analysis that the magnitude of the geomagnetic storm is determined mainly by the horizontal component $D_{st}(H)$, and this is the geomagnetic activity index during the storm. It describes the intensity of the symmetrical part of the circular current that occurs in the equatorial area of the magnetosphere during a magnetospheric storm. This is the global part of the geomagnetic storm index. It represents the mean value of the disturbed horizontal component of the geomagnetic field determined by data from several low-level observatories, distributed by geographic lengths. On quiet days, this variation may be around ±20 nT, but during a geomagnetic storm it reaches large negative values of the order of hundreds of nT.

Remark 19.1. The determination of the D_{st} *index* is provided every hour, and is practically determined by taking an average of the data through a consortium of geomagnetic observatories. The process of practical determination of this geomagnetic index is described in detail in the IAGA bulletin.

Remark 19.2. Geomagnetic field variation data in the past decades contained mean *hour* values; however, the present data contain mean *minute* values, and in all INTERMAGNET observatories, the registration is done with *mean second* values. This shows an increase in the information about the geomagnetic field and one may speak about a "Big Data" shift of the measurement paradigm.

19.2.3.1 The D_{st} Index During the 2003 Storm

In order to give an idea about the general form of the D_{st} index during the storm we provide the D_{st} data during the storm on 29 October 2003, downloaded from the World Data Center (WDC) for geomagnetism in Kyoto; the data available cover 3 days, every hour; see Fig. 19.10. In Fig. 19.10 we see that the D_{st} variation has two big decreases due to two different CMEs.

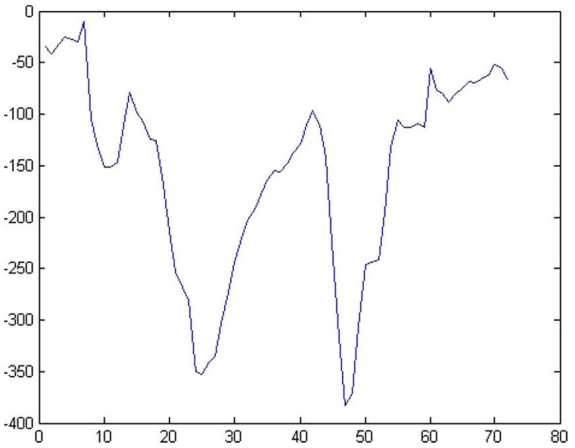

FIG. 19.10 The D_{st} index from 29 October until 1 November 2003.

19.2.4 Ionospheric Parameters From Ionospheric Sounding Stations

Let us recall that the ionosphere of the Earth can be seen as a *conductive layer*. Its motion induces an electromotive force $v \times B$. Here v is the vector of the *drift velocity* of the charged particles (not to be mixed with the tangential macrovelocity of the solar wind!), and B is the geomagnetic field vector in the *ionosphere*.

Let us mention that the ionosphere is influenced by different factors which cause some *long-period (low-frequency)* ionospheric parameter changes:

1. *Solar tidal* movements and movements, caused by the periodic warming of the atmosphere by the Sun, cause changes in the geomagnetic field, called quiet-solar variations. The latter, as noted above, are referred to as S_q. In turbulent conditions, we have a regular variation S_r associated with the local time, which differs from S_q due to the influence on the magnetospheric storm ionosphere. As with S_q, the conditions in the high atmospheric layers are particularly important and essential. For example, the concentration of the charged particles in the ionosphere, which depends on UV ionization, as well as the degree of invasion of charged particles in the ionosphere from the above areas, such as the plasmasphere and the magnetosphere.
2. The distribution in the ionosphere of the electromotive forces, determined by the *lunar tidal* movements of the atmosphere, is approximately fixed in relation to the moon. Thus, the induced current and the resulting magnetic field are also fixed for an observer on the moon. Each magnetic observatory performs one turn per day about this distribution, turning in

a circle defined by the latitude. Therefore, the stations register a variation of the geomagnetic field over time. This variation is called lunar-day and is denoted by L.

19.2.4.1 Data About the Parameters of the Ionospheric Plasma

In the present work we have used data for different parameters of *ionospheric plasma* in a specific local area, as TEC, F2, etc. In the experiments presented we have limited ourselves to the TEC data since it is considered as the most important of all parameters. For example, we use the ionosphere sounding data from ground ionospheric stations, which are chosen to be located near the (ground) geomagnetic field registration points. This allows by comparison of the two "signals" to seek for the presence or absence of a possible correlation between them. Thus, we have the possibility of identifying the origin of individual modes (wave packages) and groups of modes.

The *ionospheric* and geomagnetic data are synchronized in universal time so that we can monitor the presence or absence of simultaneity of the two signals. We have used data on ionospheric parameters with a sampling frequency comparable to the geomagnetic data, namely, every 5 minutes. This is very interesting in terms of decomposition of geomagnetic variations. Of course, the *modes* associated with extraionospheric origin can also be identified and this has already been commented (as for example the geomagnetic pulsations).

Let us note that in recent years, the *ionospheric* data registration also tends to increase the sampling frequency and the sampling has now reached in some stations a period of 5 minutes between two samples. Given that a large number of ionospheric plasma parameters are measured at these stations, we can assume that they may also be referred to as "Big Data" paradigm in this area.

For example, in Fig. 19.11 we provide a graph of the TEC data from the ionosound station in Athens, on a quiet day, 4 January 2018. The variation of the TEC data is shown. We see that the main trend is given by the daily variation of the TEC. The fast oscillations with periods less than 3 hours are observed during the whole day. At midday time we observe modes (wave packages) with a larger amplitude which are apparently of soliton type. This phenomenon has been studied in Belashov and Belashova (2015).

Remark 19.3. An interesting research of the short-period modes (wave packages) in the ionosphere was carried out by V. Belashov, who created the soliton model for their explanation; see Belashov and Belashova (2015) and references therein. Matching of the model to the TEC values is presented on p. 338. See also Srebrov et al. (2018) for confirming observations.

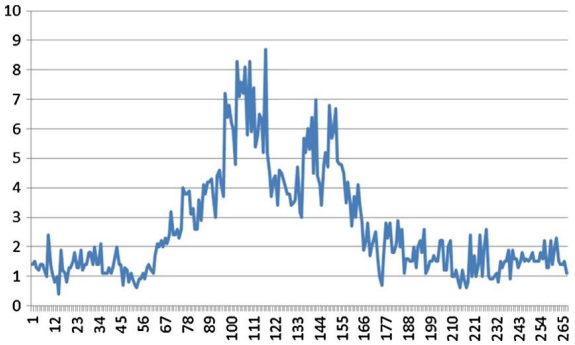

FIG. 19.11 Graph of the TEC data from the ionosound station in Athens, on 4 January 2018.

19.2.5 Emergence of Higher-Frequency Modes in the Ionospheric Parameters and in the IMF Which Are Related to the Ground Geomagnetic Field Variations

By now we have considered the long-periodic geomagnetic variation phenomena. Below we describe shortly the generation of short-period (high-frequency) geomagnetic variations caused by *low-frequency plasma instabilities*.

As already mentioned above, the plasma in the ionosphere, plasmasphere and the interplanetary collisionless plasma medium have considerable instability. Instability also occurs in the radiation belts of the Earth, i.e., in the magnetosphere, although most of the latter do not realize the conditions characterizing the medium as plasma. Some examples of cosmic plasma show the presence of instabilities that produce nonthermal waves and various distribution functions of kinetic particles parameters, especially in the ionosphere and the plasmasphere.

Instabilities can also cause a nonlinear effect of wave propagation in the ionosphere, plasmasphere, and interplanetary medium, and they also cause collisions and acceleration of fast particles in astrophysical plasma.

For the current work, it is important to note the low-frequency instabilities in terms of natural plasma frequencies (magnetohydrodynamic instabilities, fluid instabilities, and drift instabilities). They create the low-frequency modes in the geomagnetic field and in the IMF. The properties of the wave modes are strongly dependent on the frequency. For low-frequency

plasma modes, the circular frequency of the waves is much smaller than the natural frequencies as the plasma frequency and the cyclotron frequency of the plasma.

Practically, very often these *low-frequency* plasma modes coincide with the *higher-frequency* geomagnetic field variations considered by us. In this work, we use geomagnetic field registration data on the ground with periods more than 1 minute that are associated with *low-frequency modes* in the Earth's plasma cosmic environment, and also with ring current fluctuations in the magnetosphere.

In *plasma*, the macroinstabilities occur in the low-frequency mode and usually involve the magnetic field (not just the ground geomagnetic field!). Therefore, in the *present research* we use data about the variations of the magnetic field in different areas of the near-Earth space and those generated in the interplanetary space.

As concerns the *short-period (higher-frequency)* ionospheric parameter variations, let us recall that the variation can be decomposed, as a superposition of sources located in the magnetosphere and the ionosphere and also associated with the various tidal movements of the high atmosphere. The high-frequency (ground) geomagnetic field variations may be a result of short-period ionospheric parameters variations.

The above arguments justify the *necessity* to use data on the *ionosphere* status as well as data of the *interplanetary magnetic field* (IMF) and the ground geomagnetic field during the storm.

In the present work, in each case, we choose to study the behavior of the various magnetic field components that are associated with certain processes in the respective area of observation.

As we have mentioned above, the disturbances of the IMF and of the solar wind propagate in the interplanetary medium. Beyond that, in the interplanetary medium different *low-frequency* plasma modes are generated, as a result of the plasma instabilities. Data on IMF variations induced by macroinstability of the interplanetary plasma medium are recorded by satellites located outside the Earth's magnetosphere, as for example the *satellite ACE*, whose data we use. These data contain the values of the components of the magnetic vector as well as the values of its magnitude, measured by different instruments. In order to study these instabilities and related waves in the interplanetary medium, we use data for the IMF with data sampled every 4 minutes. Magnetic modes in these media and in this frequency range are associated with the so-called macroinstability.

Finally, *one of the main objectives of our study* is to identify modes with *short periods* (with frequencies much larger than the natural plasma frequencies), which are generated by microinstabilities, and are caught by the ground geomagnetic field registration and in the plasma ionospheric parameters.

19.2.6 The Strong Geomagnetic Storms in 2003 and 2017 to Be Analyzed

In the present research we apply wavelet analysis to analyze short-period geomagnetic field variations, ionospheric parameter variations, and IMF variations, during the manifestation of the two famous strong geomagnetic storms during the 23rd solar cycle in 2003, and during the 24th solar cycle in 2017.

19.2.6.1 The Storm in 2003

The data obtained from solar astronomy observations have provided the following report.

During the 23rd solar cycle in October and November 2003 there were two very strong storms. One started on October 29 and the other began on November 21. In the last 10 days of October 2003, the lean activity has gone to an extremely high level. On October 18, a large active region (AR), turning North of the solar equator, was designated by NOAA as AR 484. On October 28, AR 484 was located near the sub-Earth point of the solar disk 8° on the East of the central meridian and 16° North latitude. At 11:10 UTC AR 484 produced one of the largest solar flares for the current solar cycle. This flare was classified as X17 (peak X-ray flux 1.7×10^{-3} W/m^2).

An extreme CME with a radial plasma velocity of 2500 km/s was observed. The mass ejected from this CME was in the range of $1.4–2.1 \times 10^{13}$ kg, and the kinetic energy released was $4.2–6.4 \times 10^{25}$ J. The following day, 29 October, AR 484 again produced a large eruption. This peak was named X10 (X-ray flux 10^{-3} W/m^2) at 20:49 UTC. I was targeting the Earth halo at a speed of 2000 km/s and with a kinetic energy of 5.7×10^{25} J. The IMF reached about −50 nT; its normal value, in calm conditions, is 10 times lower. The shock wave of the event on 28 October was determined by the ACE satellite at 05:59 UTC. At 06:13 UTC an SSC pulse was registered, marking the beginning of the sixth storm by the registration stamp (since 1932). On 29 and 30 October the planetary index Kp reached a value of 9. The geomagnetic storm continued until 1 November and had a horizontal component down to around −400 nT. The highest value of the D_{st} index was registered on 30 October at 23:00 UT.

19.2.6.2 The Storm on 7 and 8 September 2017

The other storm considered in the present work is the one on 7 and 8 September 2017. It was one of the most flare-productive periods of the now-waning solar cycle 24. Solar AR2673 and AR2674 both matured to complex magnetic configurations as they transited the disk. AR2673 transformed from a simple sunspot on 2 September to a complex region with order-of-magnitude growth on 4 September, rapidly reaching beta-gamma-delta configuration. On subsequent days the region issued three X-class flares and multiple partial halo ejecta. Combined, the two active regions produced more than a dozen M-class flares. As a parting shot AR2673 produced (1) an X-9 level flare; (2) an associated moderate solar energetic particle event; and (3) a ground-level event, as it arrived at the solar West limb on 10 September. From 4–16 September the radiation environment at geosynchronous orbit was at minor storm level and 100 MeV protons were episodically present in geostationary orbit during that time frame. The early arrival of the CME associated with the 6 September X-9 flare produced severe geomagnetic storming on 7 and 8 September. The full set of events was bracketed by high-speed streams that produced their own minor-to-moderate geomagnetic storming.

19.2.7 Acquired Data for Short-Period Variations of the Geomagnetic Field, the Ionospheric Parameters, and the IMF

As we have explained in Section 19.2, the global picture of the geomagnetic phenomena is very complicated and dynamic. For that reason, for the explanation as well as for the prediction of its dynamics one needs to attract as much as possible observable data, which form the basis of our Big Data analysis.

We analyze high-frequency time series data from different sources. Let us recall that "high-frequency" registrations in geomagnetism are of the order of 0.1–10 mHz (i.e., of periods 1.66 min until 2.77 hours). The following three high-frequency time series were acquired:

1. T_G: Time series for the ground geomagnetic data (from ground geomagnetic observatories, 1 min sampling and 1 sec sampling), in nT. Our main objective is to seek for correlations in the wavelet coefficients of the CWT of the above time series which explain the dynamics of different geomagnetic phenomena.
2. T_{IP}: Time series for the ionospheric parameters – TEC (from ionospheric sounding stations, 5 min sampling), in TEC units.

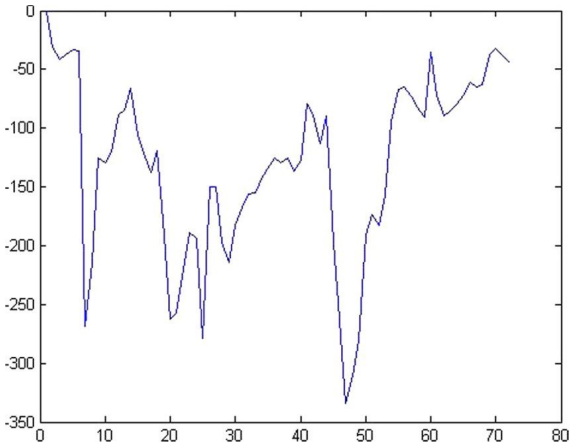

FIG. 19.12 The mean hour values of the H component registered at PAG, 29–31 October 2003.

3. T_{IMF}: Time series for the IMF (from the ACE satellite, 4 s sampling), in nT.

19.2.8 Data About the Strongly Disturbed Geomagnetic Field in October and November 2003 and September 2017

19.2.8.1 The H Component of the Geomagnetic Data From the Panagyurishte (PAG) Observatory

In Fig. 19.12 we visualize the variation of the H component (see formula (19.2)) registered at the geomagnetic observatory PAG during the geomagnetic storm on 29 October 2003. The data are mean hour values, registered from 0:00 on 29 October until 23:59 on 31 October.

19.2.8.2 The DS Index From the Surlary (SUA) Geomagnetic Data

In Fig. 19.13 we provide the data for the index DS variation (defined in formula (19.1)) during the storm on 29–31 October 2003, registered at the Surlary (SUA) geomagnetic observatory in Romania. The sampling of the data is 1 minute.

We see all details due to the fact that the data are provided every single minute.

Remark 19.4. It is questionable whether it is worth applying the CWT to the DS index given by $H(t) - D_{st}$ (see formulas (19.1) and (19.2)), or directly to the raw data of the H component, since we are in principle seeking for variations of H, but the D_{st} index is provided on an hourly basis. This tends to create artificial jump every hour.

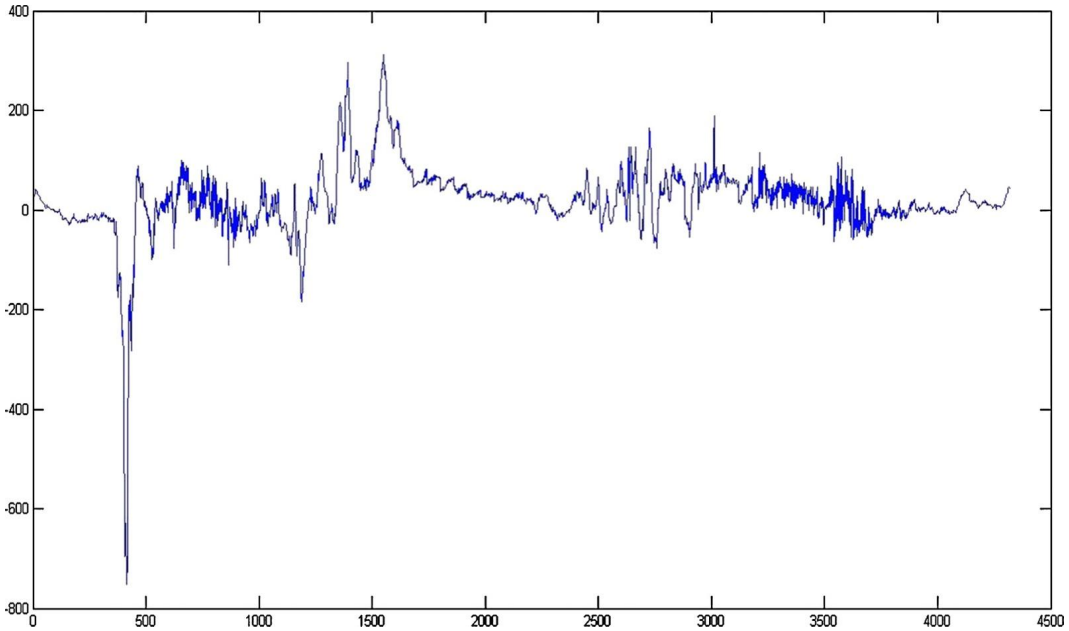

FIG. 19.13 The index DS variation during the period 29–31 October 2003, registered at the Surlary (SUA). The sampling is 1 minute.

19.3 EXPERIMENTS WITH WAVELET ANALYSIS AND CONCLUSIONS

In the experiments to follow, we are motivated by the interest to discover wave packages of short periods and their correlations in the three different sources of data which we have already discussed:

1. (ground) geomagnetic data, from geomagnetic observatories,
2. ionospheric data (TEC values, from ionosound stations),
3. IMF data (from the ACE satellite).

We would like to identify any kind of correlation and causality among them by applying the CWT method.

19.3.1 References on Applications of Wavelet Analysis to Geomagnetism

Before presenting the results of our experiments, we provide some references which might be useful to the reader.

Let us note that in a number of works (Mandrikova et al., 2014; Wei et al., 2004; Zossi de Artigas et al., 2008; Boudouridis and Zesta, 2007; Jach et al., 2006; Katsavrias et al., 2016; Xu et al., 2008), the variation of geomagnetic data (in particular of D_{st}) is analyzed by means of wavelet analysis. Recently, wavelet analysis of geomagnetic field perturbations was widely used in the study of tsunami waves (Klausner et al., 2011, 2014, 2016a, 2016b, 2017; Schnepf et al., 2016). In Jach et al. (2006) and Xu (2011), wavelet analysis of the geomagnetic field is used to define a new index, alternative to D_{st}, but on a minute basis.

19.3.2 Experiments With Data on a Quiet Day, 28 July 2018

In order to have control over the statistical behavior of the geomagnetic data during geomagnetic storms, we have taken data from quiet days from two geomagnetic observatories – the in situ repeat station at Balchik (Bulgaria) and SUA in Romania; the distance between them is about 190 km. This means that they have almost identical conditions, and it is well known that there are no strong magnetic variations from natural or artificial character in the regions.

19.3.2.1 *Visualization of the Wavelet Analysis*

We have provided a brief summary on the CWT in Appendix 19.A.

In the experiments below we perform CWT to time series $f(j)$, where $f(j) = F\left(\delta t_j\right)$ for some "continuous time series" $F(t)$, and $t_j\delta$ are the sampling times on a uniform mesh, $t_1, t_2, ..., t_N$. Here δ denotes the sampling interval, for example, we have $\delta = 1$ second, 4 sec-

FIG. 19.14 The H component with sampling 1 second at a repeat station in Balchik on 28 July 2018.

onds, 1 minute, 1 hour, etc. We visualize the absolute values $|W_\psi f(a,b)|$ of the CWT coefficients $W_\psi f(a,b)$ defined in formula (19.A.1). In all our experiments the shifts b (visualized on the X axis) run through the full set of indices $1, 2, ..., N$. On the other hand, the non-negative parameter a in formula (19.A.1), which denotes the scaling parameter (and has the meaning of *periodicity*), is interesting for us mainly for shorter intervals. Hence, in some experiments we consider only periods $a \le N_1$ for some maximal period $N_1 < N$. The parameter a is visualized on the Y axis. In order to get a better idea of the behavior of the CWT coefficients $W_\psi f(a,b)$, we find it instructive to have the visualization of $|W_\psi f(a,b)|$ both as a *heatmap* and as a *contour* map; see, for example, Fig. 19.15.

19.3.2.2 Experiments With Balchik Geomagnetic Data, 28 July 2018, 1 Second Data

The H component geomagnetic data were collected every second for 24 hours at a repeat station in Balchik (Bulgaria) on 28 July 2018.

We provide the graph of the time series of the data in Fig. 19.14. The daily variation of the field is clearly visible as the main trend, and also some rather permanent short-period variations.

Below, in Fig. 19.15, we provide the experiments, namely, the *heatmap* (top) and the *contour map* (bottom) of the absolute values of the CWT coefficients $|W_\psi f(a,b)|$ of the H component.

In Fig. 19.15 we see that in the CWT of the time series of the Balchik geomagnetic data, some very interesting details are identified. Around midday, some wave packages with periods about 100 min are clearly visible. They show the possibility for soliton-like oscillations, related to the solar terminator, theoretically studied in Belashov and Belashova (2015), and which has been recently observed in the wavelet analysis experiments with geomagnetic data in the paper Srebrov et al. (2018).

19.3.2.3 Experiment With SUA Geomagnetic Data on 28 July 2018

We have performed experiments with the time series formed by the H component of the geomagnetic data from SUA (Surlary, Romania), on a whole day, 28 July 2018; the sampling is every minute. The *heatmap* (top) and the *contour map* (bottom) of the CWT are provided in Fig. 19.16.

Remark 19.5. In Fig. 19.16 we see wave packages of 100 minutes during the whole day, and wave packages with periods below 180 minutes about midday. At midday one observes an intensive process of generation with a period between 100 and 150 minutes. Right at the same time interval (between 600 and 700 minutes) there are wave packages with period 30–50 minutes. One may suggest that the latter phenomenon may be generated by the solar terminator, and have a soliton structure (Belashov and Belashova, 2015).

Remark 19.6. There is a lot of similarity at the scale of 20–60 minutes between the CWT of Balchik data and the CWT of SUA data.

We compare the CWT of the Balchik data in Fig. 19.15 with the CWT of the SUA data in Fig. 19.16. We see that the wave packages with periods about 80 minutes are much more expressed in the Balchik second data than on the SUA minute data. This shows that the oscillation phenomena carry a persistent character. In particular, as was concluded in the paper Srebrov et al. (2018), we may suggest the existence of soliton-like patterns at the periods of 40 to 60 min.

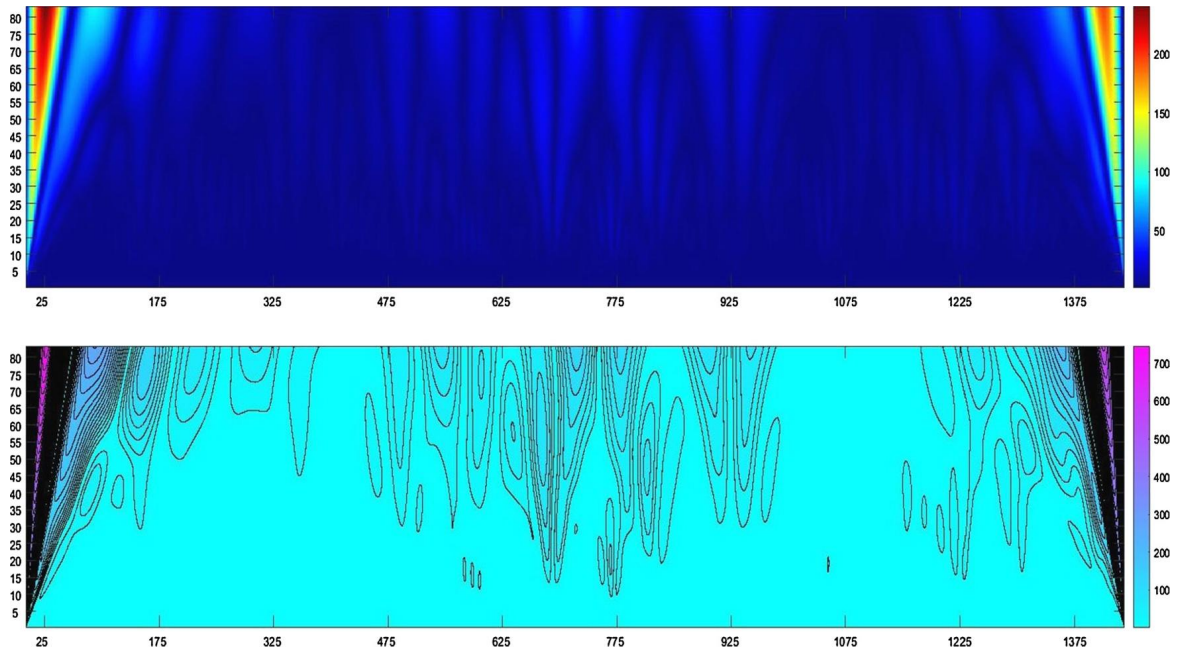

FIG. 19.15 (top) *Heatmap* and (bottom) *contour map* of the CWT of the H component of the Balchik data, on 28 July 2018.

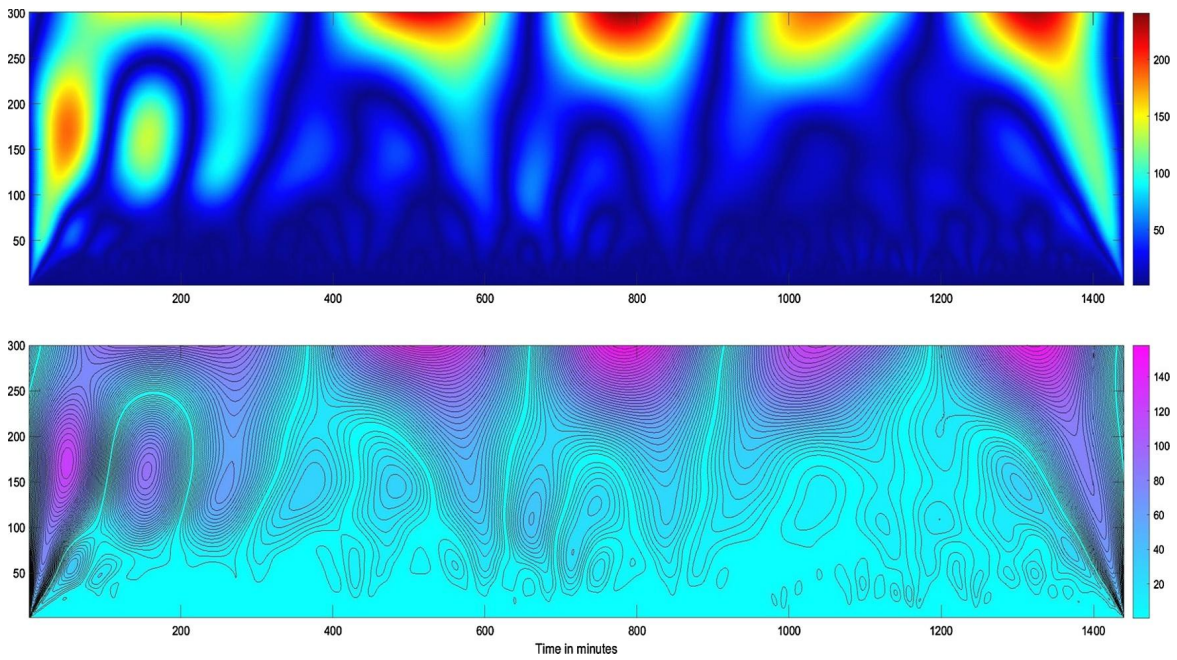

FIG. 19.16 (top) *Heatmap* and (bottom) *contour map* of the CWT of SUA data, 28 July 2018.

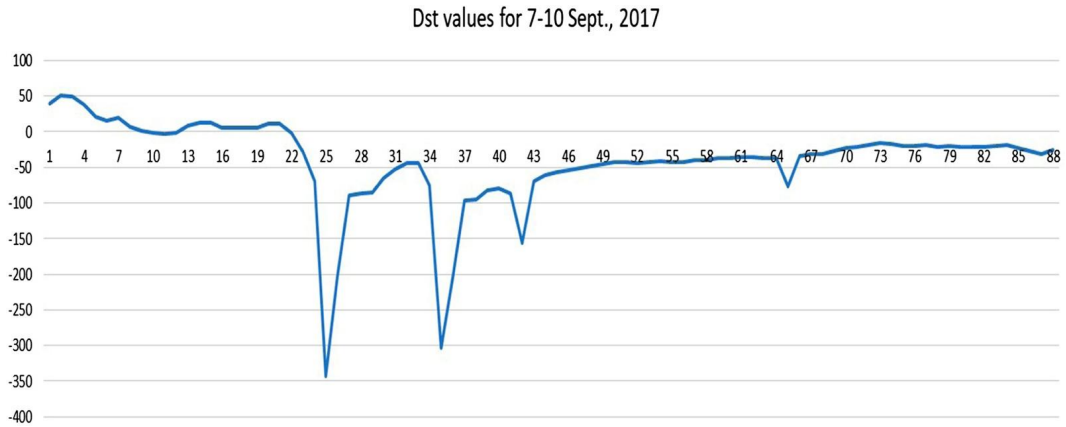

FIG. 19.17 The D_{st} graph for 7–10 September 2017.

Remark 19.7. Conclusions: Wave packages with periods from 10 to 100 min exist during quiet geomagnetic days, which are however predominant at midday.

19.3.3 Experiments With Data for the Geomagnetic Storm on 7 and 8 September 2017

In the following we provide a wavelet analysis of the data from ground geomagnetic field, ionospheric parameters, and IMF, collected during this strong geomagnetic storm.

First we provide the picture of the main trend which is determined by the D_{st} index.

19.3.3.1 The D_{st} for the Period 7–10 September 2017

Fig. 19.17 shows the D_{st} graph for the period.

As we have described the D_{st} index in Section 19.2.3, it shows a very unusual behavior after the storm of 7–10 September 2017.

Remark 19.8. In Fig. 19.17 we see the variation of the D_{st} index during the manifestation of the geomagnetic storm in September, 2017. In the figure we see the two decreases of the magnetic field, caused by two events on the Sun surface, and also a very long recovery phase of the storm during the period 9 and 10 September. This long recovery phase is most probably related to the lack of short-period variations in the geomagnetic records in the observatory on 9 and 10 September. The decay of the ring current cannot create significant geomagnetic variations on the ground in the region of the PAG observatory. However, as we have remarked above, the ionospheric macroinstabilities during these two days cannot create a variation in the ionospheric current system, which itself would create variations to be registered by this ground observatory.

19.3.3.2 Experiments With Geomagnetic Data From PAG, 7–10 September 2017, 1 min Data

We have acquired the H component of the geomagnetic data from the Panagyurishte (PAG) geomagnetic observatory. These data are every 1 minute sampling period. We provide the *heatmap* (top) and the *contour map* (bottom) of the CWT; see Fig. 19.18.

We see that during 7 and 8 September of the geomagnetic storm we may identify wave packages with periodicity 20–100 minutes. However, on 9 and 10 September, one cannot identify wave packages with periods in the interval 20–100 minutes.

19.3.3.3 Experiments With Ionospheric Data From Athens, 7–10 September 2017, 5 min Data

We have taken the ionosound TEC data from the ionosound station in Athens (ATN). The data are measured for four full days, 7–10 September 2017, every 5 minutes (this frequency is the modern standard for sampling of ionosounding data).

In Fig. 19.19 we provide the *heatmap* (top) and the *contour map* (bottom) of the CWT of the TEC time series.

We see that the short-period scales which are interesting for us really show regular patterns. For that reason we have restricted the scales *only to 50*, and we show the results in Fig. 19.20.

We see that during the geomagnetic storm we have a lot of wave packages and regular patterns. However

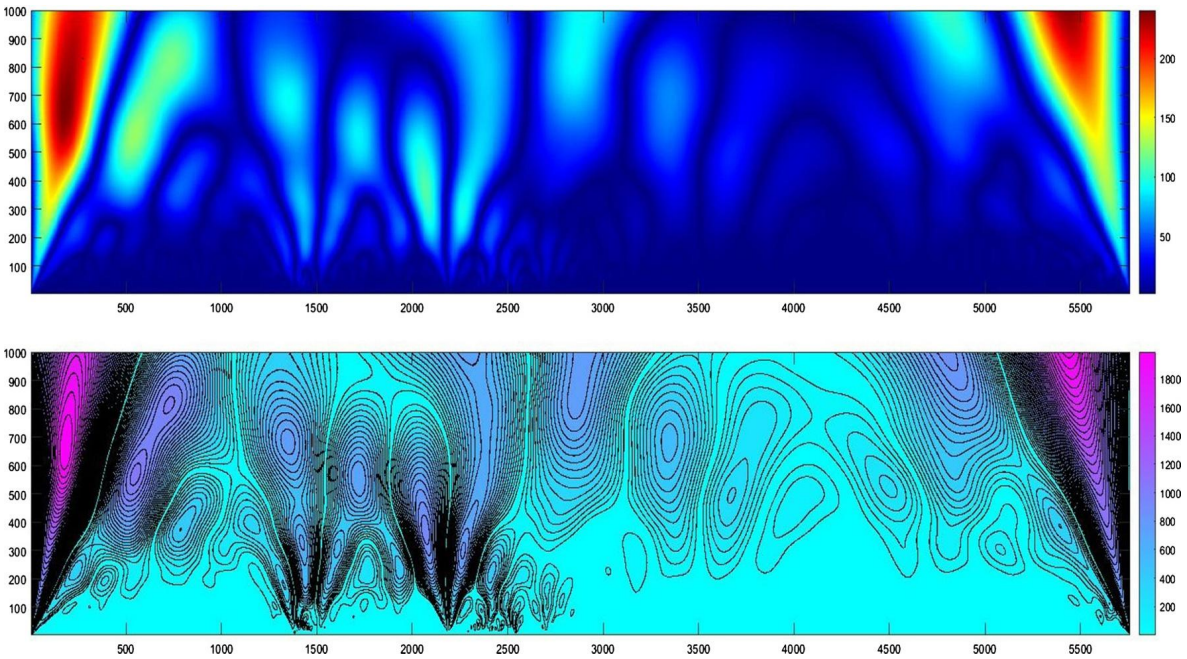

FIG. 19.18 (top) *Heatmap* and (bottom) *contour map* of the CWT for H component of PAG, 7–10 September 2017.

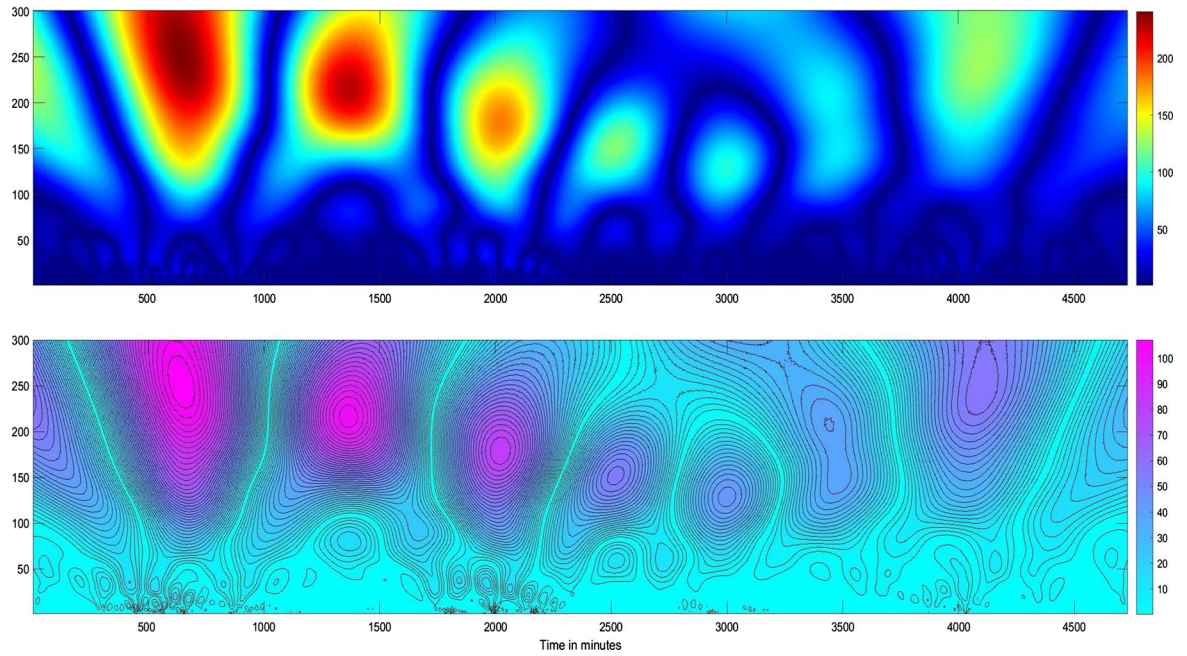

FIG. 19.19 (top) *Heatmap* and (bottom) *contour map* of the CWT of the TEC data, Athens, 7–10 September 2017.

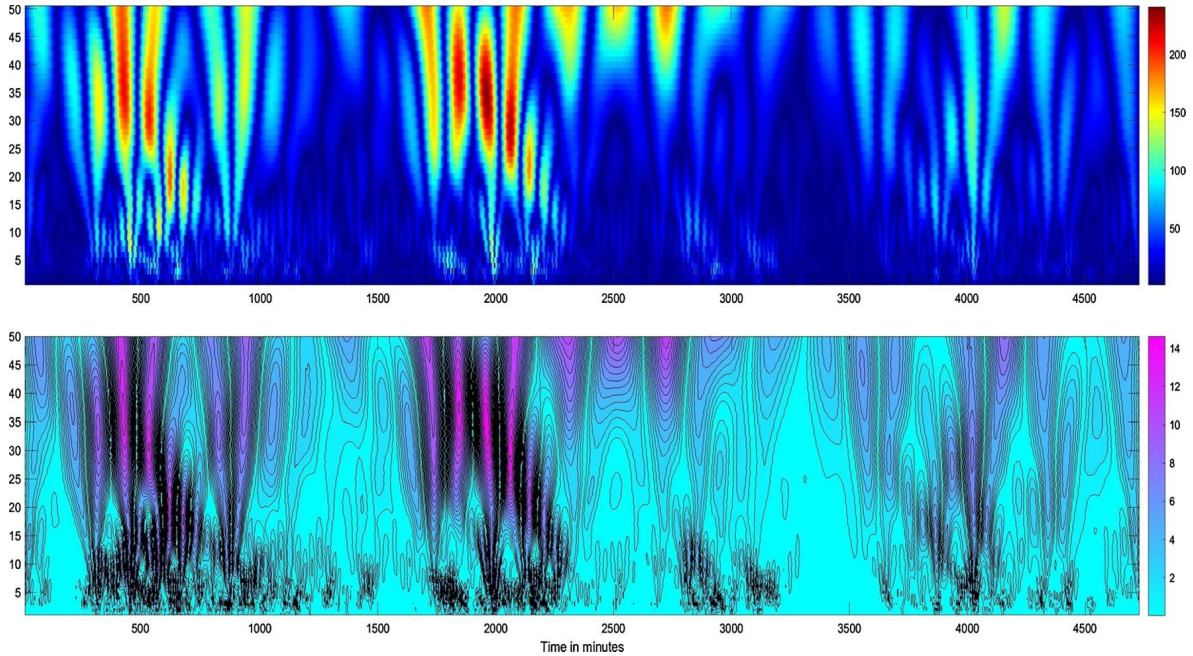

FIG. 19.20 (top) *Heatmap* and (bottom) *contour map* of the CWT of the TEC data, Athens, 7–10 September 2017, limited to short periods ≤ 50 minutes.

what is not less interesting, similar pattern appear during the two days 9 and 10 September (during the recovery phase of the storm), at midday time, having periods 20–50 minutes. This seems to be due to the solar terminator influence, as was suggested in Belashov and Belashova (2015), Schnepf et al. (2016).

19.3.3.4 Experiments With IMF Data From ACE Satellite, 7–10 September 2017, 4 min Data

We retrieved the B_z component of the IMF from the ACE satellite on 7–10 September 2017, every 4 min data.

In Fig. 19.21 we provide the CWT of the time series containing the values of B_z; again the *heatmap* is on the top and the *contour map* is on the bottom of the figure.

An interesting observation is that in the ACE data and in the ground PAG data one *cannot identify* any wave packages of periods 20–100 min after the storm, i.e., on 9 and 10 September 2017. This makes us believe that there is a correlation between the two observable values. This is in strong contrast to the ionospheric observations provided in Fig. 19.19 and Fig. 19.20, where such wave packages are available. This shows that in some cases the ionospheric plasma generates short-period modes with a relatively small amplitude. This would explain the lack of similar modes in the ground geomagnetic data of PAG. On the other hand, it is clear that on 9 and 10 September, these short-periodic modes are the result of eigenoscillations of the ionospheric plasma, which are not caused by the influence of the IMF.

19.3.4 Experiments With Data for the 2003 Strong Geomagnetic Storm

Since the geomagnetic storm in 2003 was unusually strong, it has become a handbook example for testing analysis tools. However, in 2003 there were not so many data available. In particular, the *ionosound* data are not available with the present sampling, but only hour data.

We provide below the results for wavelet analysis for the following data:

1. In 2003, for different (ground) geomagnetic ground observatories (SUA) we have 1 minute data for the H component.
2. ACE satellite data for IMF are available at frequency 4 min.

On the other hand, during this storm the ionospheric data are only *mean hour* which is *insufficient* for the present analysis.

First of all, we have the main trend of the magnetic fields provided by the D_{st} data in Fig. 19.10 above.

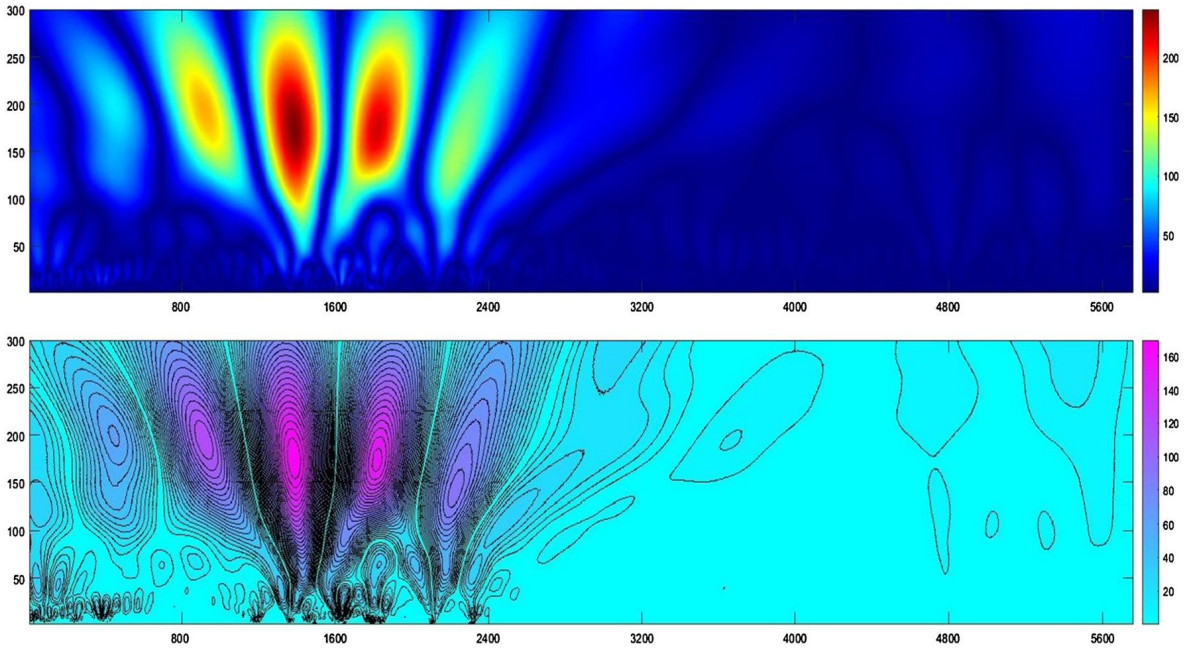

FIG. 19.21 (top) *Heatmap* and (bottom) *contour map*, for CWT of the B_z component of IMF, 7–10 September 2017.

19.3.4.1 Experiments With Geomagnetic Data From SUA, on 28 and 29 October 2003

One may analyze the original source data H component, or DS obtained after subtracting the D_{st} from the data; see formula (19.1). As we said above, since the D_{st} data are given every hour this may create artifacts every whole hour. We provide the contour plot of the CWT for the H component of the geomagnetic field from SUA, 28 and 29 October 2003, sampling 1 minute, in Fig. 19.22.

In Fig. 19.22, we may clearly identify short-period wave packages with periods below 3–4 hours, as well as with periods of 4–12 hours, but also with periods about 24 hours. The last are related with the main and recovery phase of the storm, i.e., with the ring current. The wave packages with periods below 3 hours may be related to the fluctuations of the ring current and the ionospheric instabilities generating such modes. Modes with periods below 100 minutes are generated mainly at midday, and may be related to the solar terminator and are eventually of soliton type, as was mentioned already above; see also Belashov and Belashova (2015), Srebrov et al. (2018).

In Fig. 19.23 we provide the CWT of the DS data.

19.3.4.2 Experiments With Spline Smoothing of D_{st}

We have provided some interesting experiments which show the effect of subtraction of D_{st} after smoothing the D_{st} with splines; see Fig. 19.24.

On the bottom of Fig. 19.24 we have the CWT of the H component and on the top we have the CWT of the $DS = H - \widetilde{D_{st}}$, where $\widetilde{D_{st}}$ is the smoothed D_{st} (which as mentioned is provided by the WDC of geomagnetism in Kyoto on an hourly basis).

This shows that one has to be careful when subtracting the D_{st} index (which is a step function) from the H component since this creates nonsmooth signal and the Fourier or wavelet analyses generate artificial frequencies.

19.3.4.3 Experiments With IMF Data, on 28 and 29 October 2003, 4 min Data

We have provided the CWT for the B_z component of the IMF in Fig. 19.25. Again the *heatmap* of the CWT is shown on top, while the *contour map* is shown on bottom.

In Fig. 19.25 we see various families of wave packages, which may be separated into two types: those with periods less than 100 minutes, and those with periods between 100–450 minutes. Their explanation is related

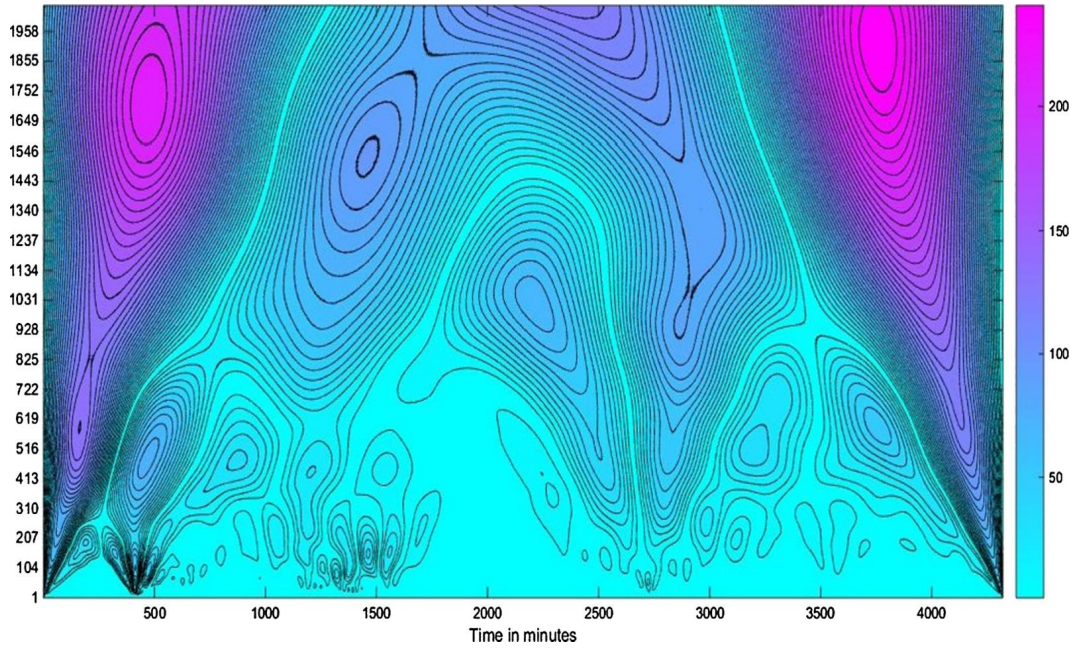

FIG. 19.22 Contour plot of CWT for the H component from SUA, 28 and 29 October 2003.

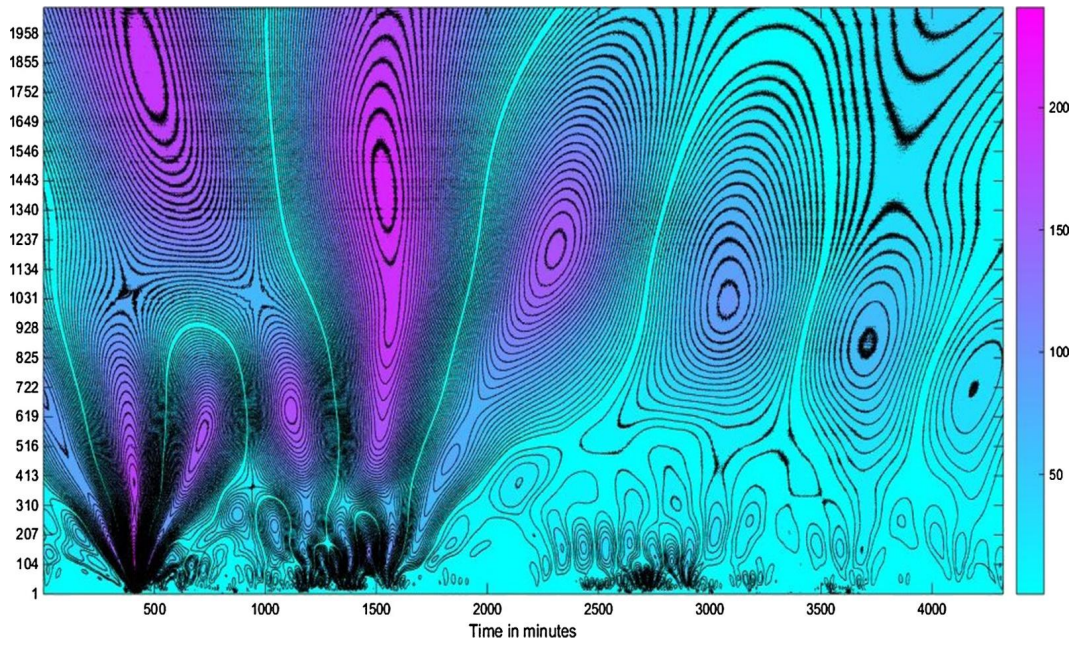

FIG. 19.23 Contour plot of the CWT for the DS index of the geomagnetic field from SUA, 28 and 29 October 2003.

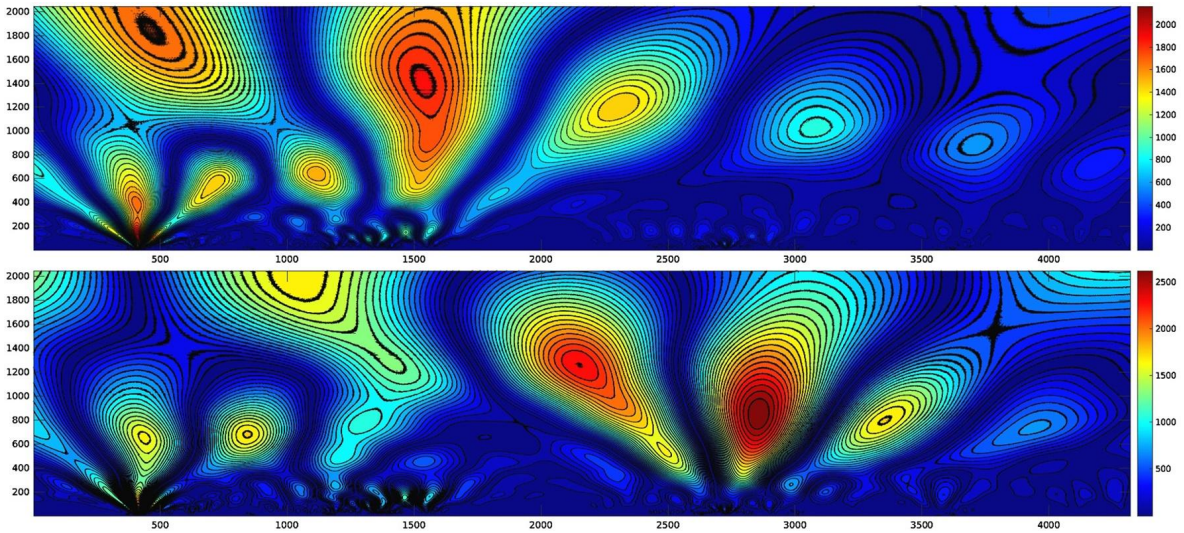

FIG. 19.24 (top) CWT of the $DS = H - \widetilde{D_{St}}$ and (bottom) CWT of the H component.

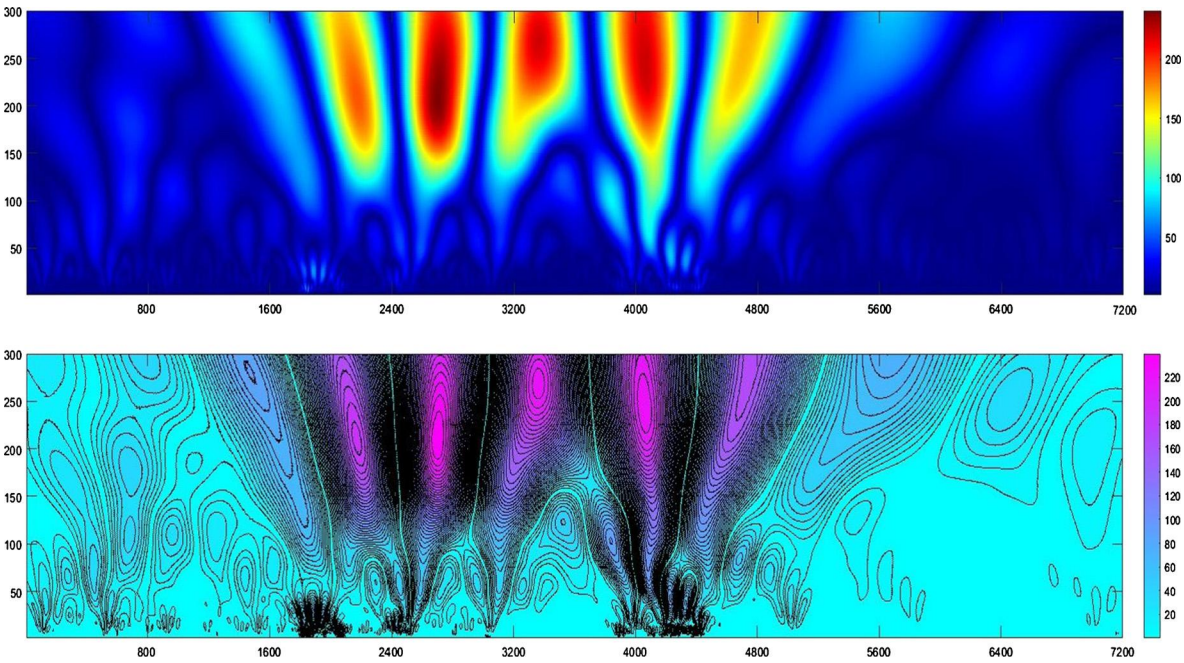

FIG. 19.25 (top) *Heatmap* of the CWT and (bottom) *contour map* for the B_z component of IMF, 28 and 29 October 2003.

to the complex structure of the disturbance of the B_z component during the strong geomagnetic storm.

19.4 CONCLUSIONS

1. The main objective of the present research is to apply wavelet analysis to Big Data in solar-terrestrial physics, for the investigation of short-period variations of the (ground) *geomagnetic field/ionospheric parameters* in a region with mean geographic latitude. Thus, by applying CWT to large amounts of heterogeneous data (geomagnetic field, ionospheric parameters, and IMF), we have identified modes (wave packages) with different periods, of the order of 20 to a few hundred minutes with a significant amplitude, which is enough to be registered by the equipment in the geomagnetic observatories.
   As is known, in the same range there exist the so-called geomagnetic pulsations, but they have a very low amplitude and exist for a short time only during the night hours for this geographic latitude. Unlike the geomagnetic pulsations, the short-period variations of our interest, have significant amplitude, and are identified in the present research; they are discovered during the whole day and may be divided into modes with periods less than 3 hours and modes which have a period greater than 3 hours.
2. Our analysis of the variations of the geomagnetic field, the ionospheric plasma parameters, and the IMF, has shown persistent short-term periodic events, as wave packages. The short-period modes (wave packages) of the variations which we have identified have a clear explanation (e.g., from plasma physics) and are caused by macroinstabilities in different domains of the near-Earth space environment.
3. We have identified the presence of modes with periods shorter than 3 hours, generated predominantly by the ionospheric plasma, but also similar modes which exist in the IMF. The ability of wavelet analysis to uncover multiresolution structure of the data gave us the possibility to identify short-periodic wave packages in the geomagnetic field variations, in IMF, and in the ionospheric parameters.
4. The present research represents a contribution to the newly developing area of AstroGeoInformatics due to the large spectrum of the analyzed phenomena which belong to solar-terrestrial physics.

ACKNOWLEDGMENTS

The authors thank the editors of this volume, Petr Škoda and Adam Fathalrahman, for their patience and assistance. Thanks extend to Aleksandra Nina (Belgrade) for discussions on the ionosphere. All authors thank the Project on Modern mathematical methods for Big Data, DH 02-13 and Project on Statistical methods for machine learning of data with complex structure, KP-06-N32-8 with Bulgarian NSF, and also the Project SatWebMare with ESA (in the PECS framework). Last but not least, all owe thanks to the COST action BigSkyEarth with EU. OK thanks the Alexander von Humboldt Foundation.

The services of the World Data Center for Geomagnetism, at the Kyoto University, Japan, the INTERMAGNET network, the ACE Science Center at Caltech, and the GIRO center, University of Massachusetts at Lowell, are gratefully acknowledged. MATLAB®/Octave code was written for the experiments.

APPENDIX 19.A WAVELET ANALYSIS AND ITS APPLICATIONS TO GEOMAGNETIC DATA

In the present research we have decided for continuous wavelet transform (CWT). We provide the essentials of the CWT and some useful references for the applications of wavelet analysis.

19.A.1 Technical Stuff

We will say that the integrable function ψ is a *wavelet function* if it satisfies the following properties:

1. the *admissibility* condition holds,

$$0 < C_\psi := \int_{-\infty}^{\infty} \frac{\left|\widehat{\psi}(\omega)\right|^2}{|\omega|} d\omega < \infty,$$

2. the *zero integral* condition holds,

$$\widehat{\psi}(0) = 0,$$

where $\widehat{\psi}(\omega)$ is the Fourier transform of ψ. This condition is equivalent to

$$\int \psi(\omega)\, d\omega = 0.$$

We consider only real-valued functions ψ.

Once the wavelet function ψ is fixed, then for every integrable function f (which is considered to represent the signal) which has a sufficient decay at ∞, and for every two real numbers $a, b \in \mathbb{R}$ with $a > 0$, we may define the CWT $W_\psi f(a, b)$ by putting

$$W_\psi f(a,b) := \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt. \quad (19.A.1)$$

The number a is called *scale*, and b is called *translation* (shift). Recall that the usual definition of the *frequency* k is then given by putting

$$k = \frac{1}{a}.$$

Unlike the usual Fourier transform, where the dimension of the variable of the signal $f(t)$ is transformed into the same-dimensional frequency domain, here we see that the CWT $W_\psi f(a,b)$ depends on two variables. We are able to reconstruct the original signal f from this representation, by means of the *Calderon* inversion formula (Jaffard et al., 2001; Mallat, 2009)

$$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{a}}W_\psi f(a,b)\psi\left(\frac{t-b}{a}\right)\frac{da}{a^2}db. \tag{19.A.2}$$

As in the DWT it is always the question to find some reasonable approximation in the Calderon formula which takes into account only the larger values of $\left|W_\psi f(a,b)\right|$, which will result in an approximation of the double integral in the equality in formula (19.A.2). Thus, the question is whether it is possible to use just a part of the integration domain. This may be achieved in different ways; one approach is to apply a threshold on the absolute value of the CWT $\left|W_\psi f(a,b)\right|$, say, $\varepsilon > 0$, and define the domain

$$D_\varepsilon = \left\{(a,b) : \left|W_\psi f(a,b)\right| < \varepsilon\right\}$$

and then consider the approximation integral

$$I_\varepsilon := \frac{1}{C_\psi}\int\!\!\int_{D_\varepsilon}\frac{1}{\sqrt{a}}W_\psi f(a,b)\psi\left(\frac{t-b}{a}\right)\frac{da}{a^2}db,$$

so that the remainder would satisfy

$$|f(t) - I_\varepsilon(t)| \le \delta \qquad \text{for all } t \in \mathbb{R}.$$

Unlike the discrete wavelet transform (DWT) here we do not have a clearly defined multiresolution analysis (MRA), and also there are no father wavelets (scaling functions). However, in general, one may use the wavelets in the discrete wavelet theory and apply them in CWT if they are smooth enough. The CWT is a very convenient tool to detect and characterize singularities in functions, in order to distinguish between noise and signal (see Jaffard et al., 2001; Mallat, 2009). In particular, one may use CWT to study fractal behavior of the signals.

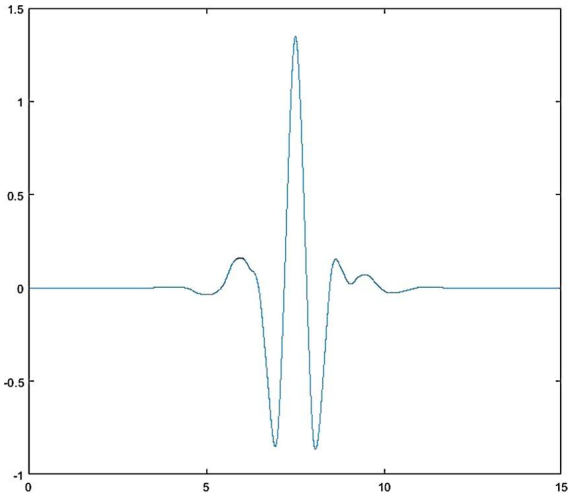

FIG. 19.26 The graph of the sym8 mother wavelet.

In a wide area of applications, people use CWT with a wavelet function ψ equal to the Mexican hat and the Morlet wavelet, although these two functions do not enjoy the usual scheme of MRA as introduced in Mallat (2009), Jaffard et al. (2001). In the present research, after numerous experimentations, we have decided for the *symlet* family of functions, which are a modified version of the Daubechies wavelet family *db*, since they enjoy increased symmetry (Addison, 2011). We have applied the *sym8* wavelet function, provided in Fig. 19.26.[1]

However, it is important to remark that the experiments with many other wavelets ψ have shown that the singularities which we detect by the *symlets* may be analyzed with the same success by applying the other wavelets; completely subjectively, we have found that *sym8* gives on average one of the best possible visual pictures. This fact shows that our observation is due to persistent physical events and may not be an artifact which is due to the particular wavelet which we choose.

19.A.2 CWT of Some Simple Functions

An important control of the wavelet analysis method is to consider the CWT of the simple jump functions, as for example *impulse trains*, which are sums of Dirac delta functions. It gives us idea about the behavior of the CWT of more complicated signals. A main reason to consider these impulse trains is the result of Belashov (Belashov and Belashova, 2015), who has very successfully modeled (better than the traditional IRI model)

[1] http://wavelets.pybytes.com/wavelet/sym8/.

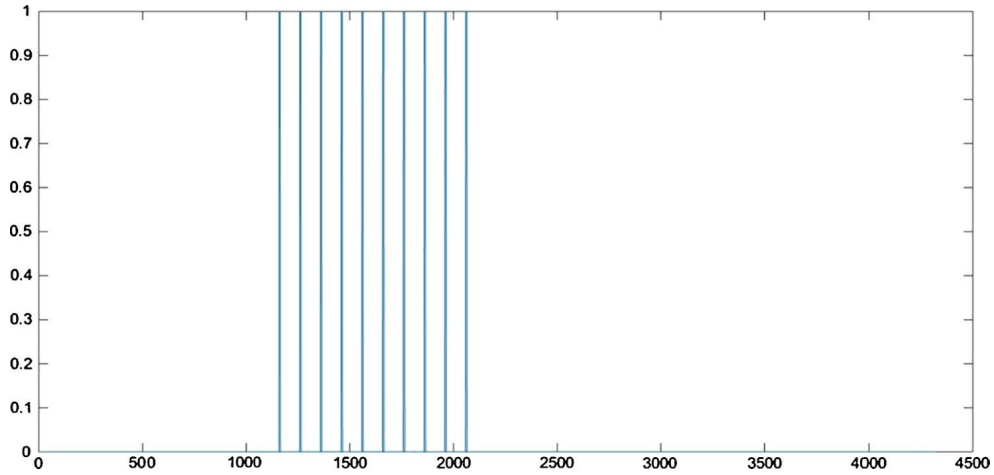

FIG. 19.27 Impulse train.

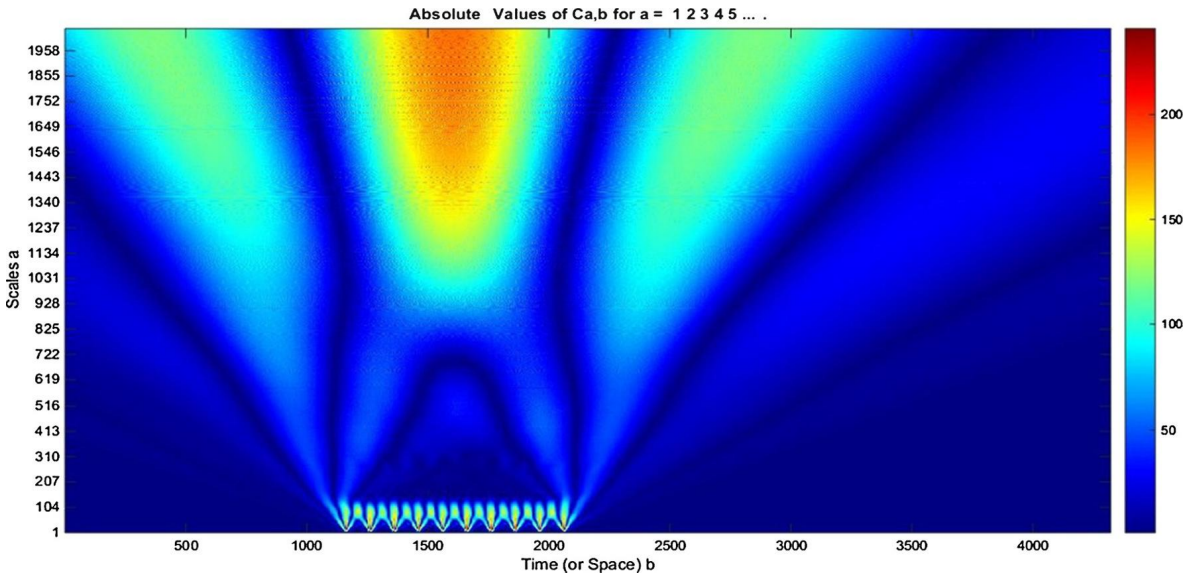

FIG. 19.28 CWT of the impulse train.

the ionospheric data by assuming that the disturbances of the ionosphere are related (even on quiet days!) to wave packages having a soliton character; we have already announced this resemblance of our analysis in Srebrov et al. (2018).

Fig. 19.27 shows the graph of a simple impulse train.

It has the CWT shown in Fig. 19.28. The y axis shows the length of the period (the scale a in the CWT $W_\psi(a,b)$ and the x axis shows the number b).

Remark 19.A.1. From Fig. 19.28 we see that a package of periodic pulses has considerable CWT $\left|W_\psi(a,b)\right|$ for lower periods $0 \leq a \leq 105$, but it also shows an "integral effect" and shows a considerable CWT $\left|W_\psi(a,b)\right|$ for longer periods $a \geq 1030$. The moral of this observation is that one has to be really careful when solving the inverse problem, i.e., when drawing conclusions about the singularities of the original signal $f(t)$ judging by the large-period behavior of the CWT $W_\psi(a,b)$.

Remark 19.A.2. Another interesting example is provided also on Wikipedia,[2] where CWT is provided of a frequency breakdown signal by using the *symlet* as a wavelet function with five vanishing moments.

REFERENCES

Addison, Paul S., 2011. The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, Second Edition. CRC Press.

Akasofu, S.-I., Chapman, Sydney, 1972. Solar-Terrestrial Physics. Clarendon Press, Oxford.

Belashov, Vasily Yu, Belashova, Elena S., 2015. Dynamics of IGW and traveling ionospheric disturbances in regions with sharp gradients of the ionospheric parameters. Advances in Space Research 56, 333–340.

Bishop, C., 2006. Pattern Recognition and Machine Learning. Springer.

Boudouridis, A., Zesta, E., 2007. Comparison of Fourier and wavelet techniques in the determination of geomagnetic field line resonances. Journal of Geophysical Research 112 (A08205).

Gençay, R., Selçuk, F., Whitcher, B., 2002. An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. Academic Press.

Hubbard, Barbara Burke, 1998. The World According to Wavelets: The Story of a Mathematical Technique in the Making. AK Peters Ltd.

Jach, A., Kokoszka, P., Sojka, J., Zhu, L., 2006. Wavelet-based index of magnetic storm activity. Journal of Geophysical Research 111, A09215.

Jaffard, S., Meyer, Y., Ryan, R.D., 2001. Wavelets. Tools for Science and Technology. SIAM.

Katsavrias, Ch., Hillaris, A., Preka-Papadema, P., 2016. A wavelet based approach to solar–terrestrial coupling. Advances in Space Research 57, 2234–2244.

Klausner, V., Almeida, T., de Meneses, F.C., Kherani, E.A., Pillat, V.G., Muella, M.T.A.H., 2016a. Chile2015: induced magnetic fields on the Z component by tsunami wave propagation. Pure and Applied Geophysics 173, 1463–1478.

Klausner, V., Almeida, T., De Meneses, F.C., Kherani, E.A., Pillat, V.G., Muella, M.T.A.H., 2017. Chile2015: induced magnetic fields on the Z component by tsunami wave propagation. In: The Chile-2015 (Illapel) Earthquake and Tsunami. Pageoph Topical Volumes. Birkhäuser, Cham, pp. 193–208.

Klausner, V., Domingues, M.O., Mendes, O., Papa, A.R.R., 2011. Tsunami effects on the z component of the geomagnetic field. Online version at https://arxiv.org/abs/1108.4893.

Klausner, Virginia, Kherani, Esfhan A., Muella, Marcio T.A.H., 2016b. Near- and far-field tsunamigenic effects on the Z component of the geomagnetic field during the Japanese event. Journal of Geophysical Research.

Klausner, V., Mendes, O., Domingues, M.O., Papa, A.R.R., Tyler, R.H., Frick, P., Kherani, Esfhan A., 2014. Advantage of wavelet technique to highlight the observed geomagnetic perturbations linked to the Chilean tsunami (2010). Journal of Geophysical Research: Space Physics 119, 3077–3093.

Kounchev, O., 2001. Multivariate Polysplines. Applications to Numerical and Wavelet Analysis. Academic Press/Elsevier.

Mallat, S., 2009. A Wavelet Tour of Signal Processing. Academic Press.

Mandrikova, O.V., Solovev, I.S., Zalyaev, T.L., 2014. Methods of analysis of geomagnetic field variations and cosmic ray data. Earth, Planets and Space, 66–148.

Meinl, Thomas, Sun, Edward, 2012. A nonlinear filtering algorithm based on wavelet transforms for high-frequency financial data analysis. Studies in Nonlinear Dynamics and Econometrics 16 (3), 1–24.

Mitchner, M., Kruger, Charles H., 1973. Partially Ionized Gases, 1st Edition. Wiley.

Oppenheim, A., Schafer, R., 2010. Discrete-Time Signal Processing. Pearson.

Percival, Donald B., Walden, Andrew T., 2000. Wavelet Methods for Time Series Analysis. Cambridge University Press.

Schnepf, N.R., Manoj, C., An, C., et al., 2016. Time-frequency characteristics of tsunami magnetic signals from four Pacific Ocean events. Pure and Applied Geophysics 173, 3935–3953.

Srebrov, B.A., 2003. MHD modeling of supersonic, super-alfvenic distrubances propagating in the interplanetary plasma and their relationship to the geospace environment. Advances in Space Research 31, 1413–1418.

Srebrov, B., Pashova, L., Kounchev, O., 2018. Study of local manifestations of G5 – extreme geomagnetic storms (29–31 October, 2003) in midlatitudes using geomagnetic data by continuous wavelet transforms. Comptes Rendus de L'Academie Bulgare Des Sciences 71.

Sun, Edward W., Meinl, Thomas, 2012. A new wavelet-based denoising algorithm for high-frequency financial data mining. European Journal of Operational Research 217, 589–599.

Sun, Edward W., Yu, Min-The, 2015. Generalized optimal wavelet decomposing algorithm for big financial data. International Journal of Production Economics 165, 194–214.

Wei, H.L., Billings, S.A., Balikhin, M., 2004. Analysis of the geomagnetic activity of the Dst index and self-affine fractals using wavelet transforms. Nonlinear Processes in Geophysics 11, 303–312.

Xu, Zhonghua, 2011. Study of geomagnetic disturbances and ring current variability during storm and quiet times using wavelet analysis and ground-based magnetic data from multiple stations. Online PhD thesis at https://digitalcommons.usu.edu/etd/984/.

Xu, Z., Zhu, L., Sojka, J.J., Kokoszka, P., Jach, A., 2008. An assessment study of the wavelet-based index of magnetic storm activity (WISA) and its comparison to the Dst index. Journal of Geophysical Research 70, 1579–1588.

Zossi de Artigas, M., Fernandez de Campra, P., Zotto, E.M., 2008. Geomagnetic disturbances analysis using discrete wavelets. Geofísica Internacional 47 (3), 257–263.

[2] https://en.wikipedia.org/wiki/Continuous_wavelet_transform.