


FULL PAPER

Open Access



Improvements and validation of the IRI UP method under moderate, strong, and severe geomagnetic storms

Alessio Pignalberi^{1*} , Marco Pietrella², Michael Pezzopane² and Rolando Rizzi¹

Abstract

This work describes several important improvements made to the International Reference Ionosphere UPdate (IRI UP) method, and a careful validation of its performances under disturbed conditions. The IRI UP method has been improved developing an algorithm capable to properly filter wrongly autoscaled ionosonde data to be assimilated, avoiding the use of these in the assimilation process. Furthermore, the preliminary quality check used to choose the variogram model in the Universal Kriging method has been replaced with a new quality check routine (NQCR), based on statistical tests carried out using the variables Q_1 , Q_2 , and cR , built on variogram's residuals. NQCR objectively identifies the best variogram model from which to get more reliable effective indices maps to be ingested in the IRI model to obtain updated f_oF2 and h_mF2 maps. IRI UP has been applied on 30 different time intervals, between January 1, 2004, and December 31, 2016, characterized by moderate, strong, and severe geomagnetic conditions, over the European region. A statistical comparison between IRI UP and IRI at the truth sites located at Fairford (51.7°N, 1.5°W, UK) and San Vito (40.6°N, 17.8°E, Italy), for f_oF2 and h_mF2 , has been performed. From the statistical validation clearly emerges how IRI UP, for f_oF2 , performs significantly better than IRI, for each of the 30 geomagnetic storms considered. Regarding h_mF2 , IRI UP performances are lower than those for f_oF2 , although still better than IRI ones. In the light of the results achieved in this investigation, the IRI UP method represents an interesting approach to Space Weather forecast in the ionospheric domain.

Keywords: IRI UP, Universal Kriging, Ionospheric data assimilation, Nowcasting maps

Introduction

Space Weather events can significantly affect the functioning of radio systems, with effects that can be rapid (immediately after the event) or delayed (a few days after the event). High-frequency (HF) point-to-point communications exploit the terrestrial ionosphere and represent a valid and complementary alternative to satellite communications. The ionosphere is a complex dynamic propagation environment, and this makes HF communications problematic even during not so intense Space Weather events, because of its intrinsic variability (e.g., Kouris et al. 1998; Kouris and Fotiadis 2002). Extreme

events occurring on the Sun, such as flares and coronal mass ejections, produce important variations of the corpuscular (solar wind) and electromagnetic [ultraviolet (UV) and X-ray emissions] component arriving on the Earth, which have a deep impact on the magnetosphere–ionosphere–atmosphere system (Zolesi and Cander 2014). As a consequence of such events, under particular conditions [i.e., when the north-south component B_z of the Interplanetary Magnetic Field is orientated southward ($B_z < 0$)], moderate, strong, and severe geomagnetic storms can occur giving rise to ionospheric storms, since the dynamics and structure of the ionosphere are significantly affected by the geomagnetic activity (Buonsanto 1999). Under ionospheric storm conditions, empirical climatological models, like the International Reference Ionosphere (IRI) model (Bilitza et al. 2014, 2017; Bilitza and Reinisch 2008), are not able to properly predict the

*Correspondence: alessio.pignalberi2@unibo.it

¹ Dipartimento di Fisica e Astronomia, Università di Bologna "Alma Mater Studiorum", Bologna, Italy

Full list of author information is available at the end of the article

ionospheric behavior, in particular during severe geomagnetic storms, as recently demonstrated for instance by Pignalberi et al. (2016). Moreover, ionospheric storms give rise to an abnormal behavior of the HF operative band, with important repercussions on the reliability of radio communications at both regional and global scale. In these circumstances, in order to select the best frequencies to be used for HF communications, timely and accurate information about the ionospheric channel is required. That is why, as established by the IRI community (Bilitza et al. 2011, 2014, 2017), real-time sparse ionosonde data should be used in conjunction with climatological models in order to get a picture of the ionospheric plasma variability as near as possible to the real conditions.

Currently, this is made possible thanks to the modern ionospheric stations which are equipped with software that provides, in almost real time, the electron density profile and the ionospheric characteristics among which, from the radio propagative point of view, the most important are the critical frequency of the F2 ionospheric layer, f_oF2 , and the propagation factor $M(3000)F2$.

Assimilating F2 layer characteristics and electron density profiles, obtained from ionosonde measurements, into background ionospheric models, it was possible to develop nowcasting models which have been repeatedly proved to be effective in providing a better comprehensive specification of the ionosphere (Angling and Khattatov 2006; Thompson et al. 2006; Decker and McNamara 2007; McNamara et al. 2007, 2008, 2010, 2011; Nava et al. 2011; Shim et al. 2011; Galkin et al. 2012; Pezzopane et al. 2011, 2013; Pietrella 2015; Pignalberi et al. 2018a, b).

Nowcasting models provide f_oF2 , $M(3000)F2$, $hmF2$, and electron density maps, which have great value from the operative point of view, because their use is aimed at maximizing the reliability and availability of HF services, especially under very disturbed ionospheric conditions as those occurring during geomagnetic and ionospheric storms.

In the light of these considerations, it is easy to understand how important is to supply maps as reliable as possible. To this regard, a method, called International Reference Ionosphere Update (IRI UP), has been recently developed by Pignalberi et al. (2018a, b) as a valid alternative to both the Simplified Ionospheric Regional Model UPdating (SIRMUP) model (Zolesi et al. 2004; Tsagouri et al. 2005) and the IRI-SIRMUP-P (ISP) model (Pezzo-
pane et al. 2011, 2013).

The IRI UP method, which has the potentiality to work over any area where an appropriate number of autoscaled ionosonde measurements are available, assimilates ionosonde f_oF2 and $M(3000)F2$ data and, relying on the Universal Kriging geostatistical interpolation

technique (Kitanidis 1997), produces maps of effective values $IG_{12\text{eff}}$ of the 12-months smoothed ionospheric index IG_{12} [derived by the IG index as calculated by Liu et al. (1983)], and maps of effective values $R_{12\text{eff}}$ of the 12-month smoothed sunspots number R_{12} (Houminer et al. 1993).

$IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps are then taken as input by the IRI climatological model to update the ionospheric background, thus providing an instantaneous two-dimensional (2-D) mapping of f_oF2 and $M(3000)F2$ and, hence, of $hmF2$ in the region of application. Its validation was early conducted for a case study taking into account only the severe St. Patrick geomagnetic storm occurred on March 17, 2015 (Pignalberi et al. 2018a, b).

It must be kept in mind that the goodness of IRI UP nowcasting maps depends mainly on two important issues: (a) f_oF2 and $M(3000)F2$ data, autoscaled in some reference stations, constitute a discrete dataset which must be as reliable as possible because this dataset is used to compute the values of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ at each reference station; (b) the Kriging interpolation method which, starting from the discrete values of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$, provides $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ nowcasting maps over the area under consideration, depends crucially on the choice of the variogram model which fits the experimental one. Therefore, in order to get a proper interpolation and consequently reliable $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps, it is extremely important that selected variogram models have a good quality.

With regard to these two issues, in the work of Pignalberi et al. (2018a, b), the reliability of f_oF2 and $M(3000)F2$ autoscaled data was based on a quality check of ionograms: for the reference stations equipped with the Automatic Real-Time Ionogram Scaler with True height analysis (ARTIST) software (Reinisch and Huang 1983; Reinisch et al. 2005; Galkin and Reinisch 2008), only ionograms with a Confidence Score (CS) greater than 75 (see http://www.ursi.org/files/CommissionWebsites/INAG/web-73/confidence_score.pdf) were selected; for the reference stations equipped with the Autoscala software (Scotto and Pezzopane 2002, 2008; Pezzopane and Scotto 2005, 2007; Scotto 2009; Scotto et al. 2012), a careful visual inspection was adopted to select the reliable ionograms. Furthermore, to avoid unrealistic $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps, a preliminary quality check of the variogram model, based on the exponent s of the *power* variogram model was also implemented.

In this paper, which represents the continuation of the work of Pignalberi et al. (2018a, b), the authors extended the investigation about the IRI UP method performance under disturbed geomagnetic conditions, selecting 30 time intervals characterized by moderate, strong, and severe geomagnetic storms occurred from January 1,

2004, to December 31, 2016. The validation of the IRI UP method is carried out comparing the IRI UP performance with the one of the IRI model with the STORM option “ON,” at the two truth sites of Fairford (51.7°N, 1.5°W, UK) and San Vito (40.6°N, 17.8°E, Italy).

Moreover, an appropriate filter (F) was implemented in the IRI UP method to automatically discard those ionosonde data clearly wrong (spikes) which, once assimilated, would affect negatively the ionospheric modeling. Its application turned out to be more effective than CS and visual inspection methods.

In addition, a new quality check routine (NQCR), based on statistical tests carried out using the statistical variables Q_1 , Q_2 , and cR , built on the residuals [differences between the observed (experimental variogram) and the modeled (variogram model) semivariance values], has also been proposed with the intention of replacing the quality check based on the exponent s of the *power* variogram model. This allowed to select more objectively, and with a more acceptable degree of confidence, the “best” variogram model to be used.

The data used and periods under study are presented in “[Data used and periods under study](#)” section. The description of the F algorithm implemented in the IRI UP method is provided in “[Description of the filter implemented to select ionosonde data](#)” section. A recall to the Kriging interpolation method is outlined in “[The Kriging interpolation method: a brief recall](#)” section. The description of the NQCR procedure is provided in “[On the choice of the best variogram model in the Universal Kriging procedure: a new quality check routine \(NQCR\)](#)” section. The validation of the IRI UP method and results obtained applying the NQCR procedure are the subject of “[Validation of IRI UP method including the F algorithm and the NQCR procedure: some results](#)” section. The discussion about the results and possible future developments is given in “[Discussion, conclusions, and future developments](#)” section.

Data used and periods under study

The data used in this study consist of: (a) A_p geomagnetic index data (Rostoker 1972); (b) K_p geomagnetic index data (Menvielle and Berthelier 1991); (c) f_oF2 and $M(3000)F2$ data.

Geomagnetic indices were downloaded from the OMNIWeb Data Explorer—NASA site at <https://omniweb.gsfc.nasa.gov/form/dx1.html>.

f_oF2 and $M(3000)F2$ data were downloaded from the interactive ionogram scaling software, SAO Explorer, developed at the University of Massachusetts Lowell Center for Atmospheric Research (UMLCAR) (<http://ulcar.uml.edu/SAO-X/SAO-X.html>) (Khmyrov et al. 2008; Reinisch and Galkin 2011).

In particular, f_oF2 and $M(3000)F2$ values from Rome and Gibilmanna were autoscaled from the ionograms recorded by an AIS-INGV ionosonde (Zuccheretti et al. 2003), and those from Warsaw were autoscaled from the ionograms recorded by a VISRC2 ionosonde (Pezzopane et al. 2009).

The ARTIST system was instead applied to autoscale f_oF2 and $M(3000)F2$ data from the ionograms recorded by digisondes (Bibl and Reinisch 1978) installed in the remaining ionospheric stations.

Table 1 provides an overview of the ionospheric stations considered, highlighting the ionosonde type, the autoscaling software, and the state of the station. The state “assimilated” means that f_oF2 and $M(3000)F2$ hourly values are assimilated to update the background model, while the state “used as test site” indicates that f_oF2 and $M(3000)F2$ hourly values are used only to test the IRI UP performance. The ionospheric stations available for data assimilation are listed for each year in Table 2. Figure 1 shows the geographical distribution of the ionospheric stations under consideration.

The IRI UP performance was evaluated at the testing stations of Fairford and San Vito over 30 time intervals characterized by geomagnetic storms occurred between January 1, 2004, and December 31, 2016. Such period embraces part of the past 23rd solar cycle (August 1996–December 2008) and the current one. A_p daily mean values > 50 and K_p maximum values $> 5^+$ (see <https://www.spaceweatherlive.com/en/help/the-kp-index>), as recorded during the main phase day of the storm, are the thresholds adopted to discard minor storms, thus selecting only the moderate, strong, and severe ones.

Specifically, each time interval was selected considering the day before the main phase of the storm (which in the selected cases is always a quiet day), the day of the main phase, and the next 4 days of the recovery phase, for a total number of 6 days for each time interval. Nevertheless, when the storm is characterized by substorms with A_p daily mean > 50 , then from the last substorm occurred, 4 days of the recovery phase are further considered; this means that, in these cases, the period under study can be characterized by more than 6 days. Some information about the geomagnetic storms considered in this study is summarized in Table 3.

Description of the filter implemented to select ionosonde data

Instead of reading the CS index from SAO files (Khmyrov et al. 2008), and visually inspecting the ionograms autoscaled by Autoscala, a new algorithm (F) was implemented in the IRI UP method to select ionosonde data. This has a dual objective: (a) to improve the level of

Table 1 European ionosonde network involved in the study

Ionospheric stations	Lat (°)	Lon (°)	Ionosonde type	Autoscaling software	State
Athens (Ath)	38.0°N	23.5°E	Digisonde DPS-4D	ARTIST 5	Assimilated
Chilton (Chi)	51.5°N	0.6°W	Digisonde DPS-1	ARTIST 4	Assimilated
Dourbes (Dou)	50.1°N	4.6°E	Digisonde DPS-4D	ARTIST 5	Assimilated
El Arenosillo (EIA)	37.1°N	6.7°W	Digisonde DPS-4D	ARTIST 5	Assimilated
<i>Fairford (Fair)</i>	51.7°N	1.5°W	Digisonde DPS-4D	ARTIST 5	<i>Used as test site</i>
Gibilmanna (Gib)	37.9°N	14.0°E	AIS-INGV	Autoscala 4.1	Assimilated
Juliusruh (Jul)	54.6°N	13.4°E	Digisonde DPS-4D	ARTIST 5	Assimilated
Moscow (Mos)	55.5°N	37.3°E	Digisonde DPS-4	ARTIST 5	Assimilated
Nicosia (Nic)	35.0°N	33.2°E	Digisonde DPS-4D	ARTIST 5	Assimilated
Pruhonice (Pru)	50.0°N	14.6°E	Digisonde DPS-4D	ARTIST 5	Assimilated
Rome (Rom)	41.8°N	12.5°E	AIS-INGV	Autoscala 4.1	Assimilated
Roquetes (Roq)	40.8°N	0.5°E	Digisonde DPS-4D	ARTIST 5	Assimilated
<i>San Vito (SaV)</i>	40.6°N	17.8°E	Digisonde DPS-4D	ARTIST 5	<i>Used as test site</i>
Warsaw (War)	52.2°N	21.1°E	VISRC2	Autoscala 4.1	Assimilated

List of ionospheric stations considered in the study and their geographical coordinates. The ionosonde type, the autoscaling software, and the state of the station are also highlighted. In italics the stations used as test site

Table 2 Ionosonde stations temporal availability

Year	Number of available ionospheric stations for assimilation	Available ionospheric stations for assimilation
2004	8	Ath, Chi, Dou, EIA, Jul, Pru, Rom, Roq
2005	8	Ath, Chi, Dou, EIA, Jul, Pru, Rom, Roq
2006	7	Ath, Chi, Dou, EIA, Jul, Pru, Rom
2007	8	Ath, Chi, Dou, EIA, Jul, Pru, Rom, Roq
2008	10	Ath, Chi, Dou, EIA, Jul, Mos, Nic, Pru, Rom, Roq
2009	11	Ath, Chi, Dou, EIA, Gib, Jul, Mos, Nic, Pru, Rom, Roq
2010	10	Ath, Chi, Dou, EIA, Jul, Mos, Nic, Pru, Rom, Roq
2011	10	Ath, Chi, Dou, EIA, Jul, Mos, Nic, Pru, Rom, Roq
2012	12	Ath, Chi, Dou, EIA, Gib, Jul, Mos, Nic, Pru, Rom, Roq, War
2013	12	Ath, Chi, Dou, EIA, Gib, Jul, Mos, Nic, Pru, Rom, Roq, War
2014	12	Ath, Chi, Dou, EIA, Gib, Jul, Mos, Nic, Pru, Rom, Roq, War
2015	12	Ath, Chi, Dou, EIA, Gib, Jul, Mos, Nic, Pru, Rom, Roq, War
2016	12	Ath, Chi, Dou, EIA, Gib, Jul, Mos, Nic, Pru, Rom, Roq, War

Ionospheric stations available for data assimilation for each year, from 2004 to 2016

reliability of assimilated data; (b) to speed up the selection process of data which are going to be assimilated.

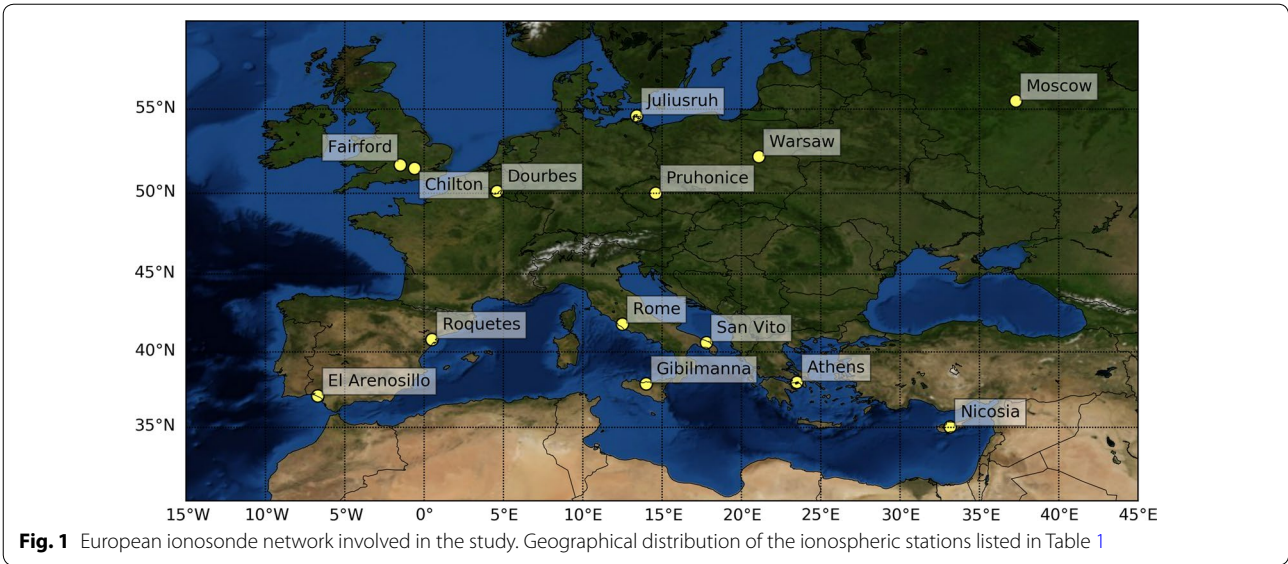
The way the F algorithm works is as follows. In order to understand whether a value of $foF2$ or $M(3000)F2$ at a given hour hr [$foF2_{hr}$ and $M(3000)F2_{hr}$, respectively] has been correctly autoscaled, $foF2$ and $M(3000)F2$ data, autoscaled in the previous 15 days at the same considered hour, are taken into account to calculate the mean values $\bar{m}_{foF2,hr,15\text{preddays}}$ and $\bar{m}_{M(3000)F2,hr,15\text{preddays}}$ and the corresponding standard deviations $sd_{foF2,hr,15\text{preddays}}$ and $sd_{M(3000)F2,hr,15\text{preddays}}$. Specifically, standard deviations are calculated when time series have a number N of available data greater than 5; otherwise, they are fixed to 0.5 MHz and 0.15 for $foF2$ and $M(3000)F2$, respectively.

Subsequently, if $sd_{foF2,hr,15\text{preddays}} \geq 0.5$ MHz, $foF2_{hr}$ is considered reliable if the following inequalities are fulfilled

$$\bar{m}_{foF2,hr,15\text{preddays}} - 5sd_{foF2,hr,15\text{preddays}} \leq foF2_{hr} \leq \bar{m}_{foF2,hr,15\text{preddays}} + 5sd_{foF2,hr,15\text{preddays}}; \tag{1}$$

otherwise, if $sd_{foF2,hr,15\text{preddays}} < 0.5$ MHz, $foF2_{hr}$ is considered reliable if the following inequalities are fulfilled

$$\bar{m}_{foF2,hr,15\text{preddays}} - 5(0.5) \leq foF2_{hr} \leq \bar{m}_{foF2,hr,15\text{preddays}} + 5(0.5). \tag{2}$$



The inequalities (2) are considered because it may happens that when very quiet days occur, the previous 15 days are characterized by $foF2$ values very close to each other; in these cases, the value of $sd_{foF2,hr,15predays}$ would be too small and, consequently, the inequalities (1) would constitute a too selective filter.

The values of $foF2_{hr}$ considered reliable are then used to calculate IG_{12eff} values at the hour hr in the corresponding assimilated ionosonde station (Pignalberi et al. 2018a, b).

Likewise, if $sd_{M(3000)F2,hr,15predays} > 0.15$, $M(3000)F2_{hr}$ is considered reliable if the following inequalities are fulfilled

$$\begin{aligned} & \bar{m}_{M(3000)F2,hr,15predays} - 5sd_{M(3000)F2,hr,15predays} \\ & \leq M(3000)F2_{hr} \leq \bar{m}_{M(3000)F2,hr,15predays} \\ & + 5sd_{M(3000)F2,hr,15predays}; \end{aligned} \tag{3}$$

otherwise, if $sd_{M(3000)F2,hr,15predays} < 0.15$, $M(3000)F2_{hr}$ is considered reliable if the following inequalities are fulfilled

$$\begin{aligned} & \bar{m}_{M(3000)F2,hr,15predays} - 5(0.15) \leq foF2_{hr} \\ & \leq \bar{m}_{M(3000)F2,hr,15predays} + 5(0.15). \end{aligned} \tag{4}$$

As already said for $foF2$, the inequalities (4) replace the inequalities (3) which would constitute a too selective filter.

The values of $M(3000)F2_{hr}$ considered reliable are then used to calculate R_{12eff} at the hour hr (Pignalberi et al. 2018a, b).

The aim of the proposed F algorithm is to remove spikes, that is values of $foF2$ and $M(3000)F2$ which are

evidently wrongly autoscaled. Figure 2 shows an example of the effectiveness of the filter applied on $foF2$ and $M(3000)F2$ data recorded at Fairford over the period October 24–29, 2016 (storm number 30 of Table 3).

The Kriging interpolation method: a brief recall

In this section some basic concepts concerning the Kriging Interpolation Method (KIM) are recalled and, at the same time, the fundamental notions on which the experimental variogram is based are provided, in order to improve the understanding of next sections.

The KIM estimates the value $\hat{z}(\mathbf{x}_0)$ at a given point \mathbf{x}_0 through a linear combination of n measurements of the variable z taken at locations with spatial coordinates $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, i.e.,

$$\hat{z}(\mathbf{x}_0) = \sum_{i=1}^n \lambda_i z(\mathbf{x}_i), \tag{5}$$

where the bold letter stands for the array of coordinates of the measurements locations.

Therefore, the problem consists in selecting a set of coefficients $\lambda_1, \lambda_2, \dots, \lambda_n$ that fulfill the conditions of *unbiasedness* and *minimum variance* (see for details Pignalberi et al. 2018a, b).

The fundamental brick of the KIM is the experimental variogram from which it is possible to get indications about the spatial correlations between the measurements.

Generally speaking, if we consider a relatively small number n of measurements $z(\mathbf{x}_1), z(\mathbf{x}_2), \dots, z(\mathbf{x}_n)$ as it can be assumed in this investigation (because the number of reference stations is at most 12), it is possible to form

Table 3 Storm time periods analyzed in this study

Storm number	Main phase day [dd mm yyyy]	A_p	K_p max	Analyzed period [dd-dd mm-mm yyyy]	R_{12}
Solar cycle 23					
1	22 01 2004	64	7	21–26 01 2004 (Str)	80.1
2	27 07 2004	186	9 –	22–31 07 2004 (Sev)	64.8
3	10 11 2004	161	9 –	06–14 11 2004 (Sev)	56.6
4	18 01 2005	84	8 –	16–25 01 2005 (Sev)	54.5
5	05 04 2005	50	7	04–09 04 2005 (Str)	49.3
6	08 05 2005	91	8 +	07–12 05 2005 (Sev)	45.0
7	15 05 2005	87	8 +	14–19 05 2005 (Sev)	45.0
8	30 05 2005	90	8 –	29–4 05–06 2005 (Sev)	45.0
9	12 06 2005	54	7 +	11–16 06 2005 (Str)	44.5
10	23 06 2005	50	7	22–27 06 2005 (Str)	44.5
11	10 07 2005	57	6 +	09–14 07 2005 (Mod)	44.6
12	24 08 2005	102	9 –	23–28 08 2005 (Sev)	41.9
13	11 09 2005	101	8 –	10–16 09 2005 (Sev)	39.4
14	14 04 2006	65	7	13–18 04 2006 (Str)	27.0
15	15 12 2006	94	8 +	14–19 12 2006 (Sev)	20.2
Solar cycle 24					
16	05 04 2010	55	8 –	04–09 04 2010 (Sev)	20.8
17	09 03 2012	87	8	08–13 03 2012 (Sev)	98.3
18	15 07 2012	78	7	14–19 07 2012 (Str)	84.5
19	17 03 2013	72	7 –	16–21 03 2013 (Str)	84.4
20	01 06 2013	58	7	30–05 05–06 2013 (Str)	90.6
21	29 06 2013	50	6 +	28–02 06–07 2013 (Mod)	90.6
22	02 10 2013	58	8 –	01–06 10 2013 (Sev)	107.0
23	17 03 2015	108	8 –	16–21 03 2015 (Sev)	82.2
24	22 06 2015	57	8 –	21–26 06 2015 (Sev)	72.1
25	27 08 2015	52	6 +	26–31 08 2015 (Mod)	66.4
26	09 09 2015	60	7	08–13 09 2015 (Str)	65.9
27	07 10 2015	74	7 +	06–11 10 2015 (Str)	64.3
28	20 12 2015	70	7 –	19–24 12 2015 (Str)	57.8
29	08 05 2016	70	6 +	07–12 05 2016 (Mod)	36.0
30	25 10 2016	57	6 +	24–29 10 2016 (Mod)	31.5

The main phase day (MPD), A_p daily mean, and K_p max geomagnetic indices values as recorded during the MPD, analyzed periods, and the R_{12} solar activity index are shown for each considered moderate (Mod), strong (Str), and severe (Sev) geomagnetic storm. Periods characterized by significant substorms are highlighted in italics

$\frac{n(n-1)}{2}$ pairs of measurements, to define the distance of each pair $h_k = |\mathbf{x}_k - \mathbf{x}_k'|$, and hence the corresponding semivariance $\gamma(h_k)$ as follows

$$\gamma(h_k) = \frac{1}{2} [z(\mathbf{x}_k) - z(\mathbf{x}_k')]^2, \tag{6}$$

where $k = 1, \dots, \frac{n(n-1)}{2}$ refers to each pair of measurements.

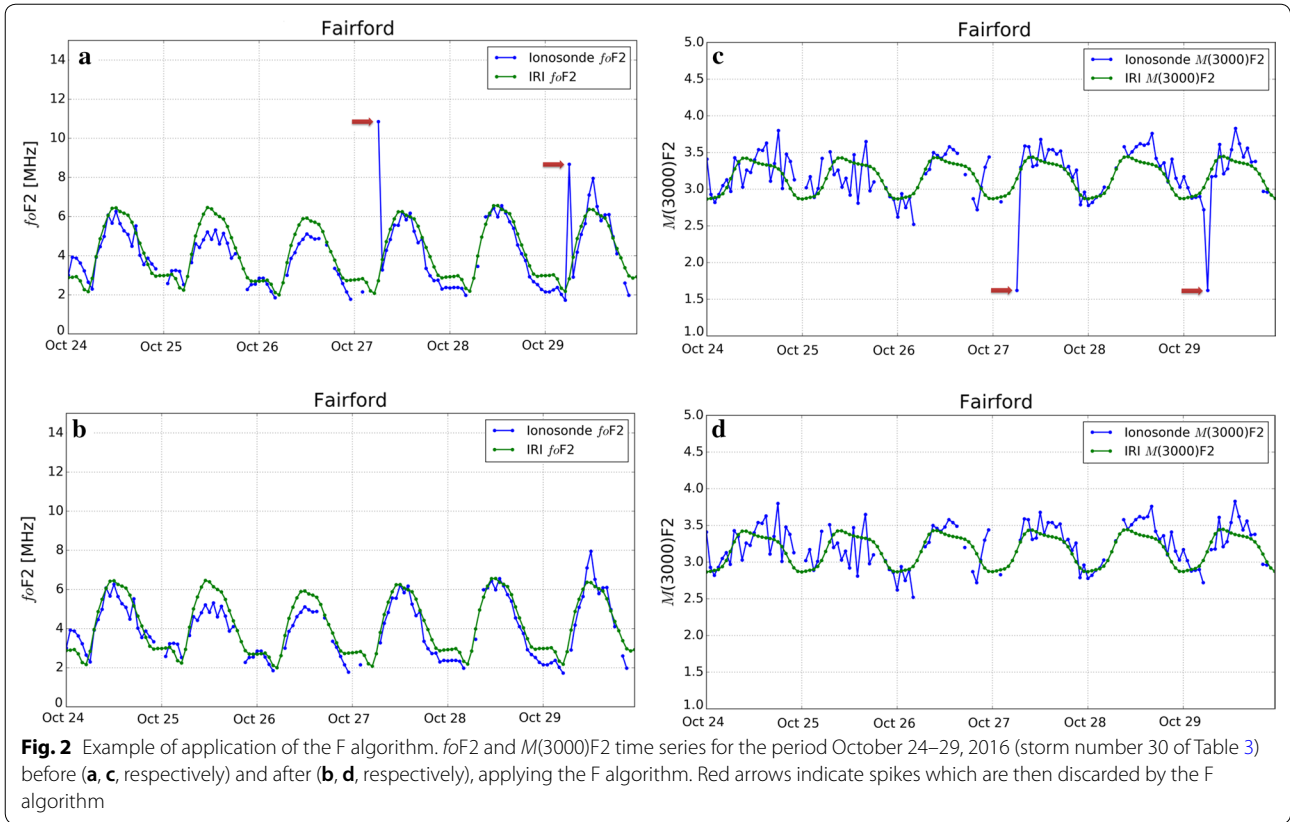
The plot formed by the $\frac{n(n-1)}{2}$ points of coordinates $(h_k, \gamma(h_k))$ constitutes the experimental variogram.

For cases where a huge number n of measurements is considered, it would be better to arrange the $\frac{n(n-1)}{2}$ pairs of measurements in K bins having all the same width

$W = \frac{h_{\max} - h_{\min}}{K}$, being h_{\min} and h_{\max} the absolute minimum and maximum distance among the $\frac{n(n-1)}{2}$ pairs of measurements.

The lower and upper limit of each bin is defined through an iterative procedure which starts from h_{\min} (the lower limit of the first bin) and ends with h_{\max} (the upper limit of the last bin). For example, the first, second, third and last bins correspond to the intervals $[h_{\min}, h_{\min} + W)$, $[h_{\min} + W, h_{\min} + 2W)$, $[h_{\min} + 2W, h_{\min} + 3W)$ and $[h_{\min} + (K - 1)W, h_{\max}]$, respectively.

As each bin can contain a different number, $N_{K'}$ of pairs of measurements, the distance pertinent to each bin is



then defined as the average value, \bar{h}_K , of the distances h_k “falling” inside the bin

$$\bar{h}_K = \frac{\sum_{k=1}^{N_K} h_k}{N_K}. \tag{7}$$

As a consequence, the semivariance associated with a given bin is the average value of the semivariances $\gamma(h_k)$ “falling” inside that bin

$$\bar{\gamma}_K = \frac{\sum_{k=1}^{N_K} \gamma(h_k)}{N_K} = \frac{1}{2(N_K)} \sum_{k=1}^{N_K} [z(\mathbf{x}_k) - z(\mathbf{x}'_k)]^2; \tag{8}$$

the plot obtained with the K points (one for each bin) of coordinates $(\bar{h}_K, \bar{\gamma}_K)$ constitutes the experimental variogram.

It is hence possible to “build” an experimental variogram based on the assimilated measurements recorded in the reference stations. It depends essentially on how the reference stations are distributed over the considered area. Data assimilation from many reference stations located close to each other is important to catch the small-scale spatial structures, because it populates the

part of the variogram near the origin. The assimilation of data from reference stations placed far from each other is instead fundamental for the description of the large-scale spatial behavior.

Once the variogram is “built,” there is the need to find a mathematical function which fits the experimental data. The mathematical expressions that can be used to fit the experimental variogram correspond to five commonly used variogram models: *linear* and *power* (non-stationary models), *gaussian*, *spherical*, and *exponential* (stationary models).

The distinction between stationary and non-stationary models depends on their behavior at distances comparable to the size of the domain. When the experimental variogram presents a steady trend around a value, called *sill* (σ^2), as the distance increases, it is possible to define a length scale, called *range* (α), at which the *sill* is obtained. α describes the spatial scale from which two measurements of the variable are no longer correlated. The variogram models characterized by the parameters σ^2 and α are called stationary. The value at the origin, $\gamma(0) = c_0$, indicates a discontinuity because the experimental variogram does not converge to zero as the distance decreases. The parameter c_0 , called *nugget*, is a term describing how well the microscale variability is represented. The

mathematical functions describing the various variogram models are the following:

1. *Linear*:

$$\gamma(h) = c_0 + \vartheta h, \tag{9}$$

where $\vartheta > 0$ is the slope;

2. *Power*:

$$\gamma(h) = c_0 + \vartheta h^s, \tag{10}$$

where $\vartheta > 0$ is the scale and $0 < s < 2$ is the exponent;

3. *Gaussian*:

$$\gamma(h) = c_0 + (\sigma^2 - c_0) \left[1 - e^{-\left(\frac{7h}{4\alpha}\right)^2} \right], \tag{11}$$

exhibiting a parabolic behavior around the origin;

4. *Spherical*:

$$\gamma(h) = \begin{cases} c_0 + (\sigma^2 - c_0) \left[\frac{3h}{2\alpha} + \frac{1}{2} \left(\frac{h}{\alpha} \right)^3 \right] & \text{for } 0 \leq h \leq \alpha, \\ \sigma^2 & \text{for } h > \alpha \end{cases}, \tag{12}$$

exhibiting a linear behavior around the origin;

5. *Exponential*:

$$\gamma(h) = c_0 + (\sigma^2 - c_0) \left[1 - e^{-\left(\frac{3h}{\alpha}\right)} \right], \tag{13}$$

exhibiting a linear behavior around the origin.

It is important to keep in mind that each variogram model embeds, with a different degree of reliability, the information concerning the small- and large-scale behavior of the parameter under study. That is why the selection of the variogram model plays a fundamental role in determining the quality and reliability of the prediction map over the area under consideration.

The observation at the point of coordinate \mathbf{x}_0 can be represented as

$$z(\mathbf{x}_0) = m(\mathbf{x}_0) + \varepsilon(\mathbf{x}_0), \tag{14}$$

where $m(\mathbf{x}_0)$ is the deterministic part of $z(\mathbf{x}_0)$ representing the large-scale spatial variability, while $\varepsilon(\mathbf{x}_0)$ is the stochastic part of $z(\mathbf{x}_0)$ which describes, particularly for stationary variogram models, the small-scale spatial variability. In the Ordinary Kriging method, $m(\mathbf{x}_0) = m$, i.e., it is a constant which does not depend on the spatial coordinates. Since the electron density presents values at mid-low latitudes higher than the ones at mid-high latitudes, there exists a latitudinal spatial gradient characterizing the ionospheric characteristics which we are

going to describe. This fact makes the Universal Kriging method (UKM) particularly suitable to describe the variability of the ionospheric characteristics under study, because UKM takes into account also the spatial gradients by means of additional terms, so that the term $m(\mathbf{x}_0)$, also called *drift* part, is written as

$$m(\mathbf{x}_0) = \sum_{j=1}^p f_j(\mathbf{x}_0) \beta_j, \tag{15}$$

where $f_1(\mathbf{x}_0), \dots, f_p(\mathbf{x}_0)$ are functions of spatial coordinates generally known, and β_1, \dots, β_p are the so called *drift coefficients*, which are usually unknown.

In our specific case the term $m(\mathbf{x}_0)$ is written as:

$$m(\mathbf{x}_0) = A + \varphi_0 \cdot B + \lambda_0 \cdot C, \tag{16}$$

being φ_0 and λ_0 the longitude and latitude of the point \mathbf{x}_0 , which are known, while the coefficients A , B , and C must be determined. The combination of Eq. (15) with Eq. (16) gives $f_1(x_i) = 1$, $f_2(x_i) = \varphi_0$, and $f_3(x_i) = \lambda_0$.

Combining Eq. (14) with Eq. (15), we get

$$z(\mathbf{x}_0) = \sum_{j=1}^p f_j(\mathbf{x}_0) \beta_j + \varepsilon(\mathbf{x}_0). \tag{17}$$

With regard to the issue dealing with the stochastic part $\varepsilon(\mathbf{x}_0)$, being the topic out of the context of this work, we invite the interested reader to refer to Section 3.3 of Pignalberi et al. (2018a) and to Kitanidis (1997).

On the choice of the best variogram model in the Universal Kriging procedure: a new quality check routine (NQCR)

As already explained by Pignalberi et al. (2018a, b), reliable *foF2*, *M(3000)F2*, and *hmF2* maps can be obtained when the IRI model is updated with a realistic $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ representation. Such representation is obtained applying the UKM on a discrete set of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ values, obtained at the locations of selected reference stations. It must be pointed out that the UKM provides a variogram model which, in principle, should be the one that best fits the experimental variogram (i.e., the variogram built on the discrete set of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ values).

Therefore, the problem of how to choose among the possible variogram models described in “[The Kriging interpolation method: a brief recall](#)” section is of crucial importance, because this choice greatly affects the goodness of the $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps and, consequently, the capability of delivering an accurate and trustworthy *foF2*, *M(3000)F2*, and *hmF2* modeling. To this regard, in the work of Pignalberi et al. (2018a, b) a routine for

the preliminary quality check of the variogram model, based on the evaluation of the exponent s ($0 < s < 2$) of the power variogram model, was developed in the IRI UP method.

In cases for which s is close to 0, the variogram model is an approximately straight horizontal line, while in cases for which s tends to 2 the variogram model has an approximately parabolic behavior at both small and large scales (Kitanidis 1997). For $s \ll 1$ the semivariance does not vary with the distance, as it should be expected; therefore, these cases correspond to unrealistic situations which would lead to unrealistic maps of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$. For this reason, for the preliminary quality check of the variogram model, we have set a very low threshold value for s ($s_{\text{thrs}} = 0.1$) discarding those variograms for which $s < s_{\text{thrs}}$. Examples of variogram models and related maps of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ discarded by the IRI UP method are shown in Fig. 3. The meaning of the associated statistics parameters Q_1 , Q_2 , and cR written in the legend will be clarified in the next sections.

It should be noted that $s < s_{\text{thrs}}$ does not provide an objective criterion, because s_{thrs} is the same for each variogram model. That is why in this work the IRI UP method has been updated by the NQCR procedure to choose the best variogram model, through appropriate statistical tests.

Q_1 statistics

Residuals are the differences between observations and model predictions. In statistical modeling (regression, time series, analysis of variance, and geostatistics), the parameter estimation and the model validation depend heavily on the examination of residuals.

We can define the variable

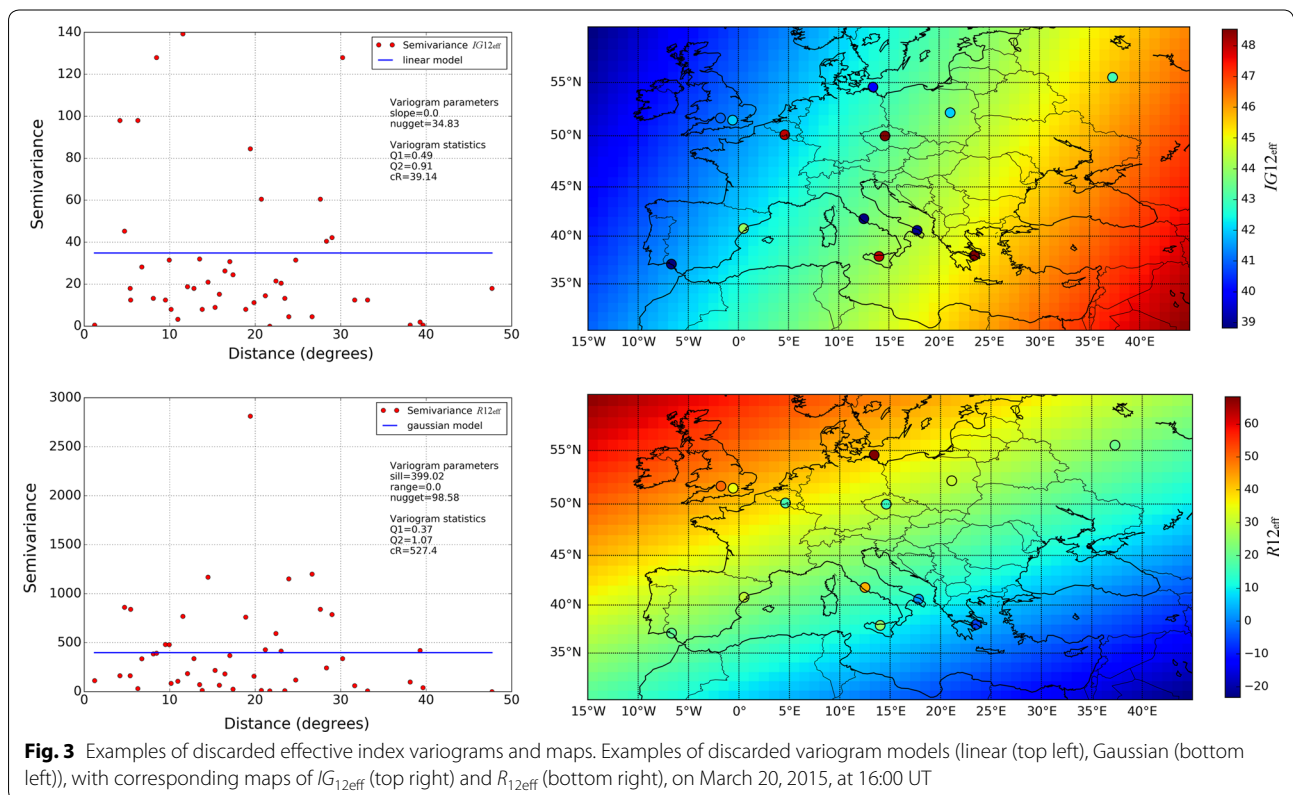
$$Q_1 = \frac{1}{n-1} \sum_{k=2}^n \varepsilon_k, \tag{18}$$

where n is the number of observations and $\varepsilon_k = \frac{\delta_k}{S_k}$ are the normalized residuals, being δ_k the common residuals and S_k the variance of their distribution. It can be proven (see Kitanidis 1997) that Q_1 is a statistical variable which follows the normal distribution with a probability density function (PDF) given by

$$f(Q_1) = \frac{1}{\sqrt{\frac{2\pi}{n-1}}} e^{-\left(\frac{Q_1^2}{\frac{2}{n-1}}\right)}, \tag{19}$$

with a mean value $m = 0$ and a variance $\sigma^2 = \frac{1}{n-1}$.

Therefore, we have a probability of about 95% that Q_1 is ranged in the interval



$$-2\sigma < Q_1 < +2\sigma \quad \text{i.e.} \quad |Q_1| < \frac{2}{\sqrt{n-1}}, \quad (20)$$

and a probability of about 68% that Q_1 is ranged in the interval

$$-\sigma < Q_1 < +\sigma \quad \text{i.e.} \quad |Q_1| < \frac{1}{\sqrt{n-1}}. \quad (21)$$

An example of the PDF and the cumulative probability function (CPF) of the variable Q_1 , obtained for $n = 12$, is shown in Fig. 4.

When the condition (20) is fulfilled, the variogram model under examination is accepted and it is significant from the statistical point of view with a probability of about 95%; this implies that there exists a 5% of probability to accept an incorrect variogram model. However, a 5% cutoff is customary in statistics. Therefore, when (20) is met, we can assume that the variogram model has passed the Q_1 statistical test.

Q_2 statistics

Another test that can be carried out to test the goodness of the variogram model is that relying on the statistical variable Q_2 defined as

$$Q_2 = \frac{1}{n-1} \sum_{k=2}^n \varepsilon_k^2. \quad (22)$$

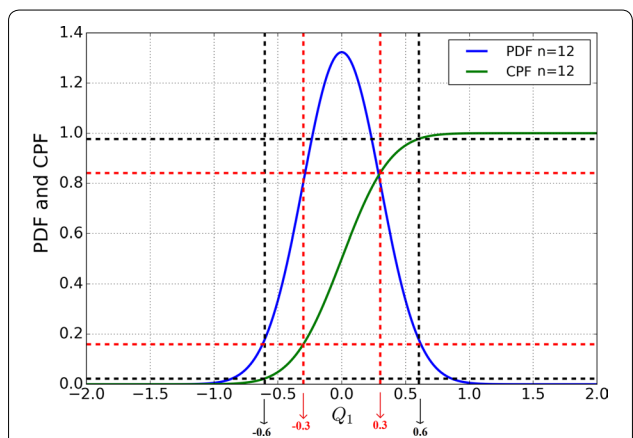


Fig. 4 Probability density function and cumulative probability function for the Q_1 variable. Example of PDF (solid blue line) and CPF (solid green line) for Q_1 , for $n = 12$. Red and black arrows (corresponding to red and black dashed vertical lines) indicate, respectively, the thresholds $\pm 1\sigma$ and $\pm 2\sigma$. Black and red dashed horizontal lines represent the values referred to the CPF corresponding to the conditions (20) and (21) which, for the considered case, provide, respectively, the numerical solutions $-0.60 < Q_1 < +0.60$ and $-0.30 < Q_1 < +0.30$

It can be proven (see Kitanidis 1997) that Q_2 is a statistical variable whose PDF is

$$f(Q_2) = \frac{(n-1)^{\frac{n-1}{2}} Q_2^{\frac{n-3}{2}} e^{-\left[\frac{(n-1)Q_2}{2}\right]}}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)}, \quad (23)$$

with a mean value $m \rightarrow 1$, for $n \rightarrow \infty$, and a variance $\sigma^2 = \frac{2}{n-1}$.

An example of the PDF and CPF of the variable Q_2 , obtained for $n = 12$, is shown in Fig. 5.

Therefore, according to Fig. 5, if the value of Q_2 given by Eq. (22) is included in the interval

$$L < Q_2 < U, \quad (24)$$

the variogram model under examination is significant from the statistical point of view with a probability of 95%. When this condition is met, we could say that the variogram model has passed the Q_2 statistical test to the usual confidence level of 5%, in the sense that there exists, however, a 5% probability that an incorrect variogram model is accepted.

It is worth noting that the form of PDF and CPF depends on the number n of assimilated ionosonde data and, consequently, the same stands for the values of the two thresholds (L, U) and (L^{1st}, U^{3rd}).

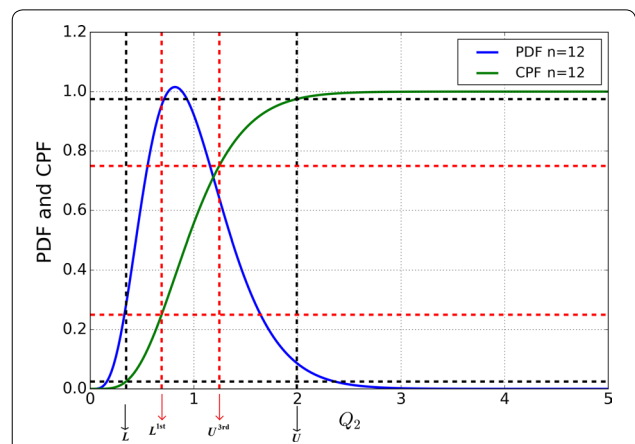


Fig. 5 Probability density function and cumulative probability function for the Q_2 variable. Example of PDF and CPF for Q_2 for $n = 12$. The black dashed horizontal lines intersect the CPF at the two points of coordinates ($L, 0.025$) and ($U, 0.975$), while the red dashed horizontal lines intersect the CPF at the two points of coordinates ($L^{1st}, 0.25$) and ($U^{3rd}, 0.75$), where L^{1st} and U^{3rd} identify the first and third quartile, respectively. As a consequence, the black and red arrows mark the intervals of the variable Q_2 corresponding to a statistical significance of 95% and 50%, respectively

The *cR* criterion

The residuals are particularly important in evaluating how closely the variogram model fits the data, since smaller residuals imply a better fit. To construct stable (i.e., less affected by random error) criteria for the choice of the best variogram model, we may also use the stabilized geometric mean of the residuals' variance (S_k), subject to the constraint $Q_2 = 1$ (see Kitanidis 1997), i.e., the parameter

$$cR = e^{\left(\frac{1}{n-1} \sum_{k=2}^n \ln(S_k^2)\right)}. \tag{25}$$

The condition

$$cR = \text{minimum}, \tag{26}$$

allows to choose among the various variogram models.

Q_1 , Q_2 , and *cR* statistical test: some results

The statistical criteria (20), (24), and (26) described in the previous sections constitute the NQCR procedure implemented in the IRI UP method.

For the selection of the “best” variogram model, one could be tempted to consider only the *cR* criterion, leaving out the criteria (20) and (24). Nevertheless, from a preliminary investigation conducted over a large number of variogram models, we realized that if only the *cR* criterion were applied, several variogram models which do not satisfy the criteria (20) and/or (24) would be accepted.

For these reasons, we decided to proceed to the selection of the variogram through an iterative procedure that NQCR applies following the flowchart depicted in Fig. 6. The figure shows just an example of how the NQCR procedure can be applied on each of the epochs (dd/mm/yyyy/hh) listed in Table 3 and illustrates, in general terms, the various steps carried out in order to select that variogram model which, to an acceptable degree of confidence, fits the data more reliably than the other ones.

Figure 7 shows some examples of spherical and linear variogram models which have met the requirements (20), (24), and (26) and that therefore have been selected to get a statistically significant $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ modeling and, consequently, a reliable mapping of *foF2* and $M(3000)F2$ and, hence, of *hmF2*. Note that the variograms reported in Fig. 7, matching the requirements (20), (24), and (26), automatically fulfill also the previous quality check $s < s_{\text{thrs}}$.

Some examples of Q_1 , Q_2 , and *cR* time series, for linear and spherical variogram models used to obtain $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps, are shown in Figs. 8 and 9, respectively. For the Q_1 statistical test, each case exceeding the threshold (20) is rejected, as well as for the Q_2 statistical test each case not included in the interval (24) is refused. Note that the epochs characterized by the greatest values of *cR* are those following the main phase of the storm.

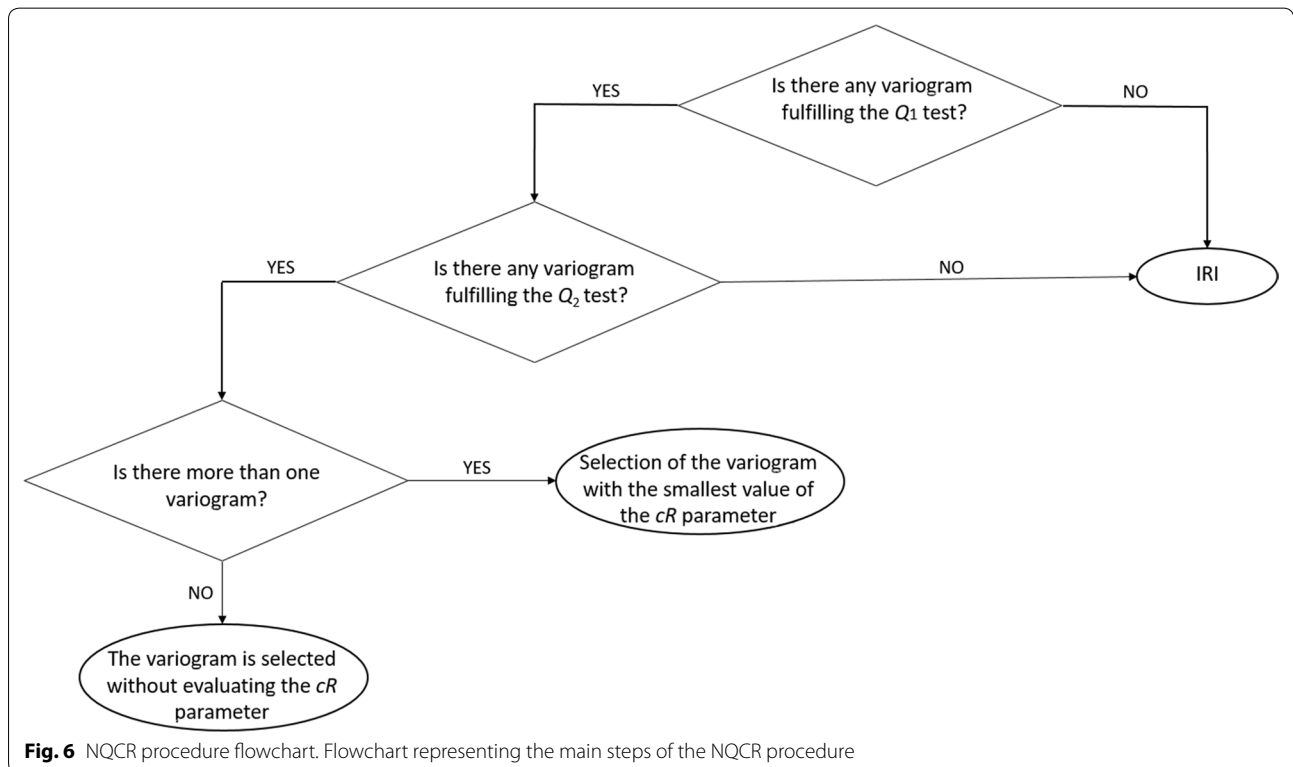
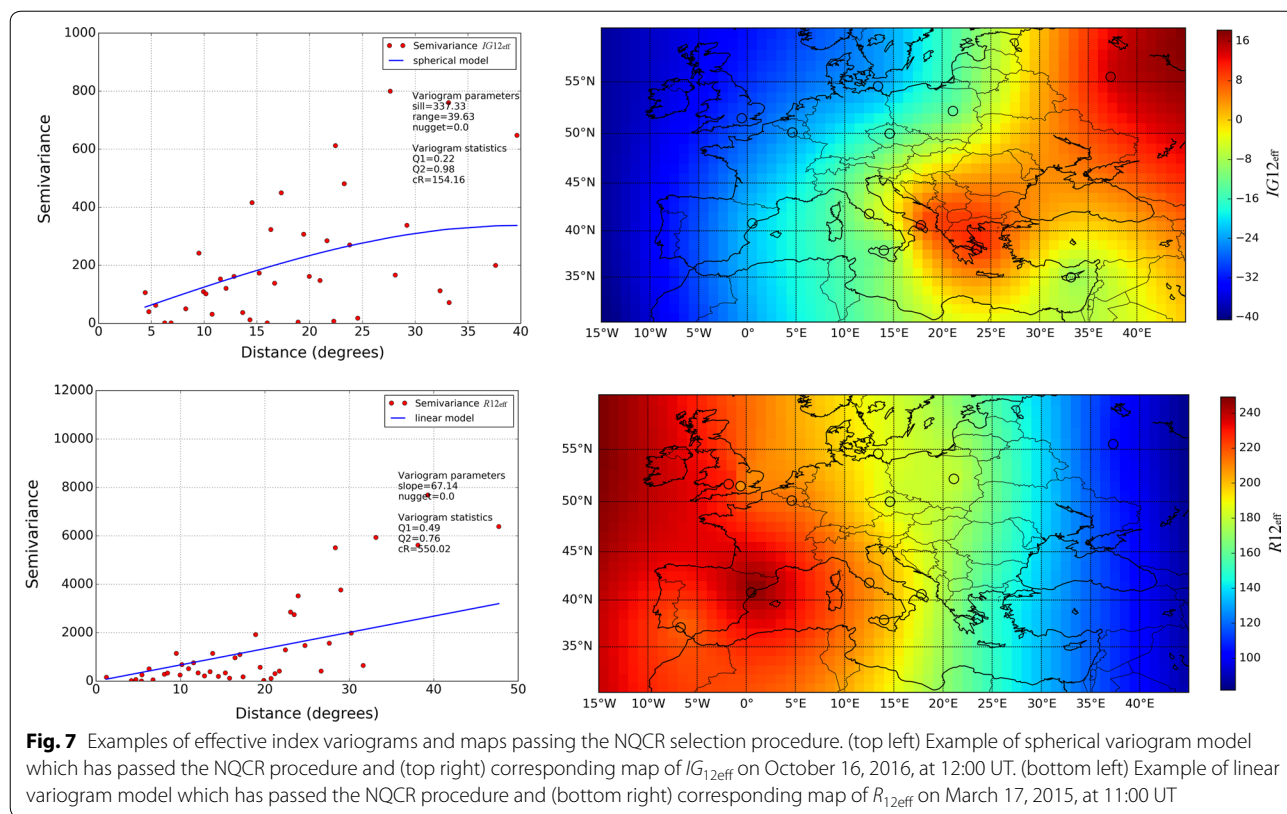


Fig. 6 NQCR procedure flowchart. Flowchart representing the main steps of the NQCR procedure



Blue, red and green dots represent the values of Q_1 , Q_2 , and cR , respectively, computed by Eqs. (18), (22), and (25). We want to stress once again here the fact that for the Q_1 and Q_2 statistics the thresholds depend on the number n of $foF2$ and $M(3000)F2$ values available in the reference stations which, at a given hour, have been considered according to the criteria described in “Description of the filter implemented to select ionosonde data” section. Obviously, the value of n can be different from epoch to epoch because, at a given moment, it is possible that a station is not working and/or the ionospheric characteristics of a station are wrongly autoscaled.

This explains why the continuous black lines representing the Q_1 and Q_2 thresholds in Figs. 8 and 9 are not flat. The implementation of the NQCR procedure constitutes then an important difference with respect to the preliminary quality check based only on the exponent s for which, whatever is the epoch and variogram model under study, the threshold value s_{thrs} is not a function of n but it is fixed to 0.1.

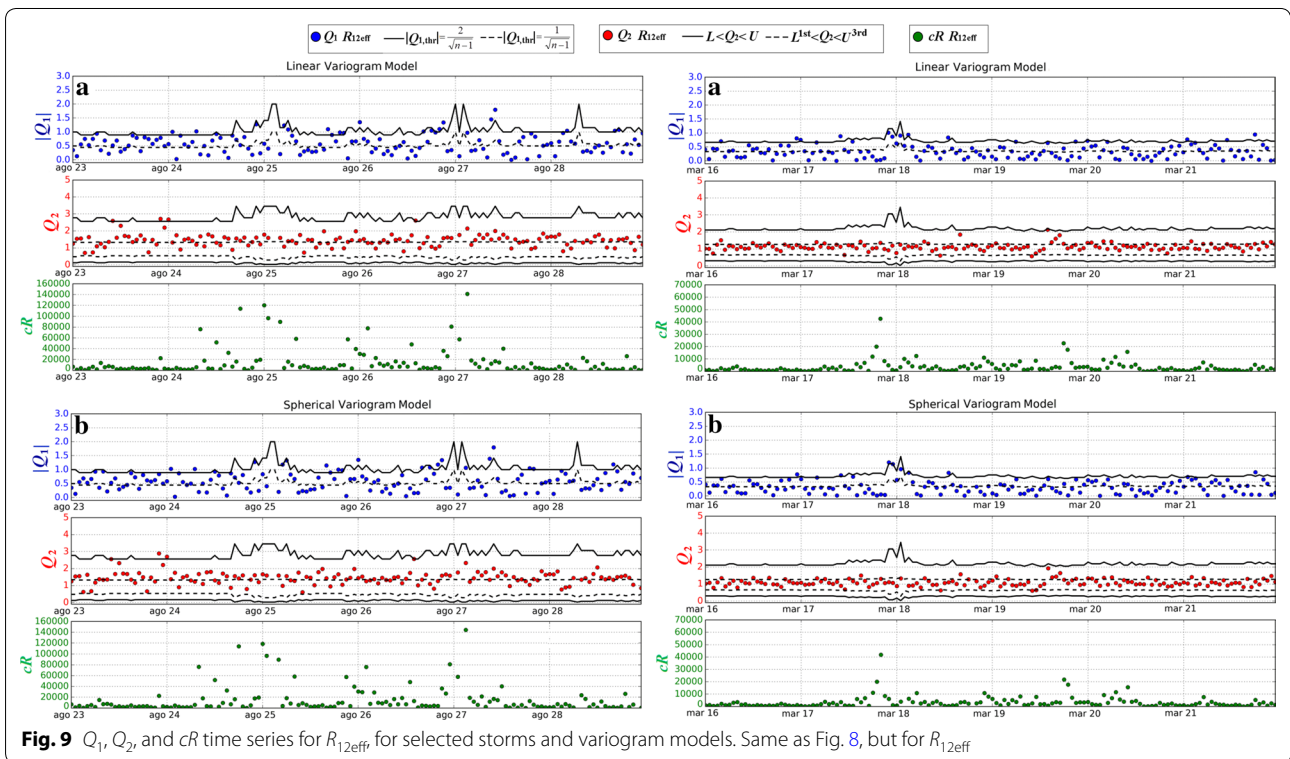
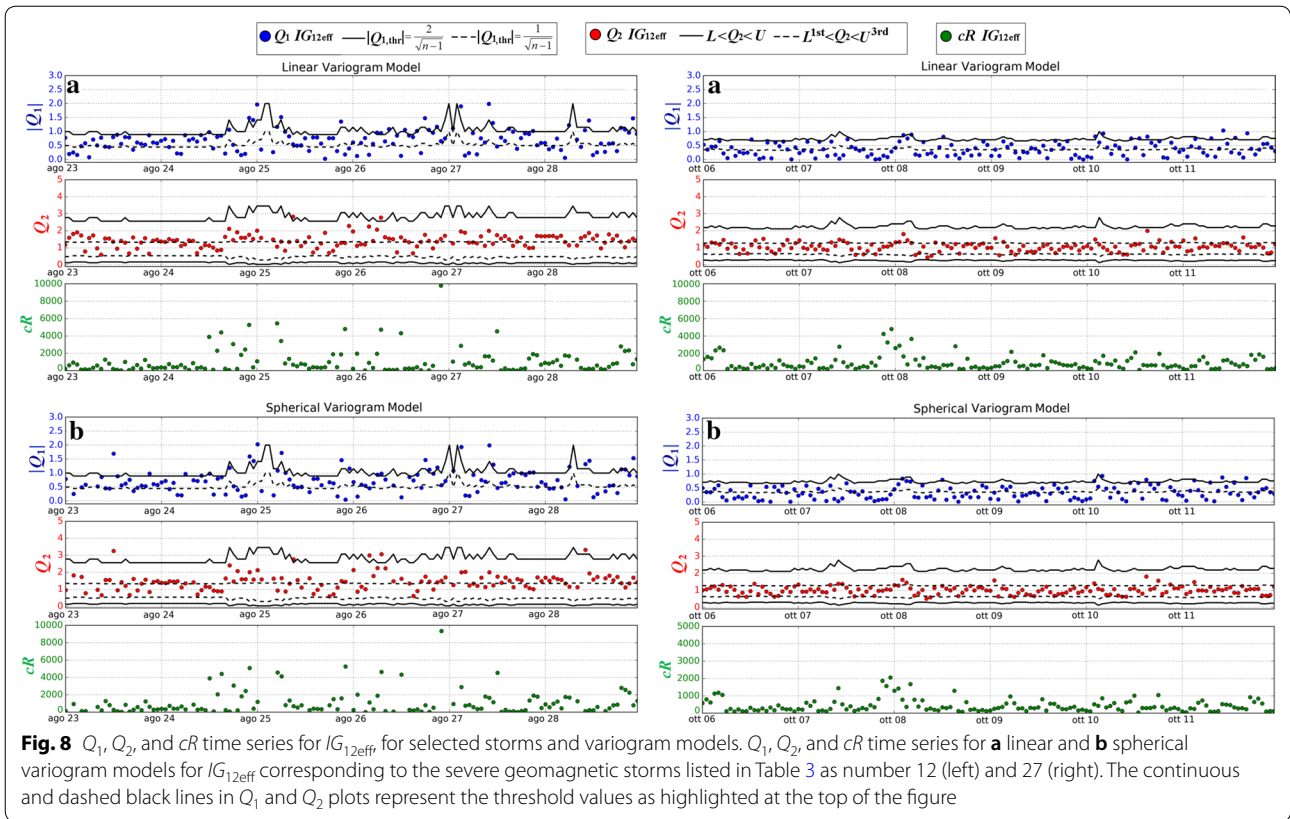
Validation of IRI UP method including the F algorithm and the NQCR procedure: some results

The IRI UP method, embedding the F algorithm and the NQCR procedure, as described, respectively, in “Description of the filter implemented to select ionosonde data”

and “On the choice of the best variogram model in the Universal Kriging procedure: a new quality check routine (NQCR)” sections, has been systematically tested over the 30 disturbed time intervals listed in Table 3, in order to investigate its performance during moderate, strong, and severe geomagnetic storms.

For each epoch, the testing procedure follows four steps:

1. the variogram models which have passed the Q_1 and Q_2 statistical tests (as, for example, those fitting the $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ experimental variograms of Fig. 7) and the cR criterion are considered;
2. applying the UKM for each selected variogram model, $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps are calculated over the European area depicted in Fig. 1;
3. $foF2$ and $M(3000)F2$ maps are calculated giving as input to the IRI model the $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ maps calculated in 2); then, applying the empirical formula which relates $hmF2$ to $foF2$ and $M(3000)F2$ (Bilitza et al. 1979), also $hmF2$ maps are obtained;
4. from $foF2$ and $hmF2$ maps, values at the truth sites of Fairford and San Vito are extracted and compared with corresponding measurements.



For each storm the following statistical parameters are calculated:

$$\begin{aligned} &\text{Root Mean Square Error, RMSE}_X \\ &= \sqrt{\frac{\sum_{i=1}^N (X_{\text{modeled},i} - X_{\text{ionosonde},i})^2}{N}}, \end{aligned} \quad (27)$$

where N is the number of epochs that constitute the considered storm, X stands for $foF2$ or $hmF2$, the subscript modeled stands for IRI UP or IRI predicted values (the IRI model is considered with the storm option “ON”), while the subscript ionosonde refers to values recorded by the ionosonde;

$$\begin{aligned} &\text{Normalized Root Mean Square Error,} \\ &\text{NRMSE}_X = \frac{\text{RMSE}_X}{\bar{X}_{\text{ionosonde}}}, \end{aligned} \quad (28)$$

where $\bar{X}_{\text{ionosonde}}$ is the arithmetic mean over time of $X_{\text{ionosonde}}$ values;

$$\begin{aligned} &\text{Pearson Correlation Coefficient, } \rho_X \\ &= \frac{\text{cov}(X_{\text{modeled}}, X_{\text{ionosonde}})}{\sigma_{X_{\text{modeled}}} \sigma_{X_{\text{ionosonde}}}}, \end{aligned} \quad (29)$$

where $\text{cov}()$ is the covariance, while $\sigma_{X_{\text{modeled}}}$ and $\sigma_{X_{\text{ionosonde}}}$ are the standard deviations;

$$\begin{aligned} &\text{Mean Delta, MD}_X = \frac{\sum_{i=1}^N (X_{\text{modeled},i} - X_{\text{ionosonde},i})}{N}. \end{aligned} \quad (30)$$

In addition, also the percentage of discarded maps is calculated:

$$\begin{aligned} \text{\%discarded} = \frac{\text{totalmaps} - \text{goodmaps}}{\text{totalmaps}}, \end{aligned} \quad (31)$$

where goodmaps is the number of $IG_{12\text{eff}}$ or $R_{12\text{eff}}$ maps passing the first two steps of the NQCR procedure, i.e., Q_1 , and Q_2 tests, and totalmaps is the total potential number of maps.

Figures 10 and 11 show some examples of comparison between IRI and IRI UP, in terms of the aforementioned statistical quantities carried out at Fairford and San Vito for each storm listed in Table 3, for both $foF2$ and $hmF2$.

Figure 12 shows, for each storm individually and for the whole group of storms, the percentage of discarded variogram models calculated by Eq. (31).

Finally, the MD_X quantities provided by Eq. (30), taken as absolute values (to avoid that positive and negative values around zero could cancel each other), have been used to calculate the following mean value:

$$\begin{aligned} \text{Mean}(\text{MD}_X \text{ Absolute Value}), \text{MMD}_X \text{AV} = \sum_{j=1}^{30} \frac{|\text{MD}_{X,j}|}{30}. \end{aligned} \quad (32)$$

Tables 4 and 5 summarize the statistical results calculated by Eqs. (27)–(29) and Eq. (32) at the truth sites of Fairford and San Vito for $foF2$ and $hmF2$, respectively, in the following three cases: (a) IRI UP method running with a fixed variogram model (like in Pignalberi et al. 2018a) and considering only those cases passing the first two steps of the NQCR procedure, namely Q_1 and Q_2 tests; (b) IRI background model; (c) IRI UP method embedding the complete NQCR procedure.

The winning percentage of each variogram model is computed for each single storm and for the complete storm set, evaluating the following parameters

$$\begin{aligned} \text{\%single_storm},i = \frac{n_{ss,i}}{N_{ss}}, \end{aligned} \quad (33)$$

$$\begin{aligned} \text{\%all_storms},i = \frac{\sum_{ss=1}^{30} n_{ss,i}}{\sum_{ss=1}^{30} N_{ss}}, \end{aligned} \quad (34)$$

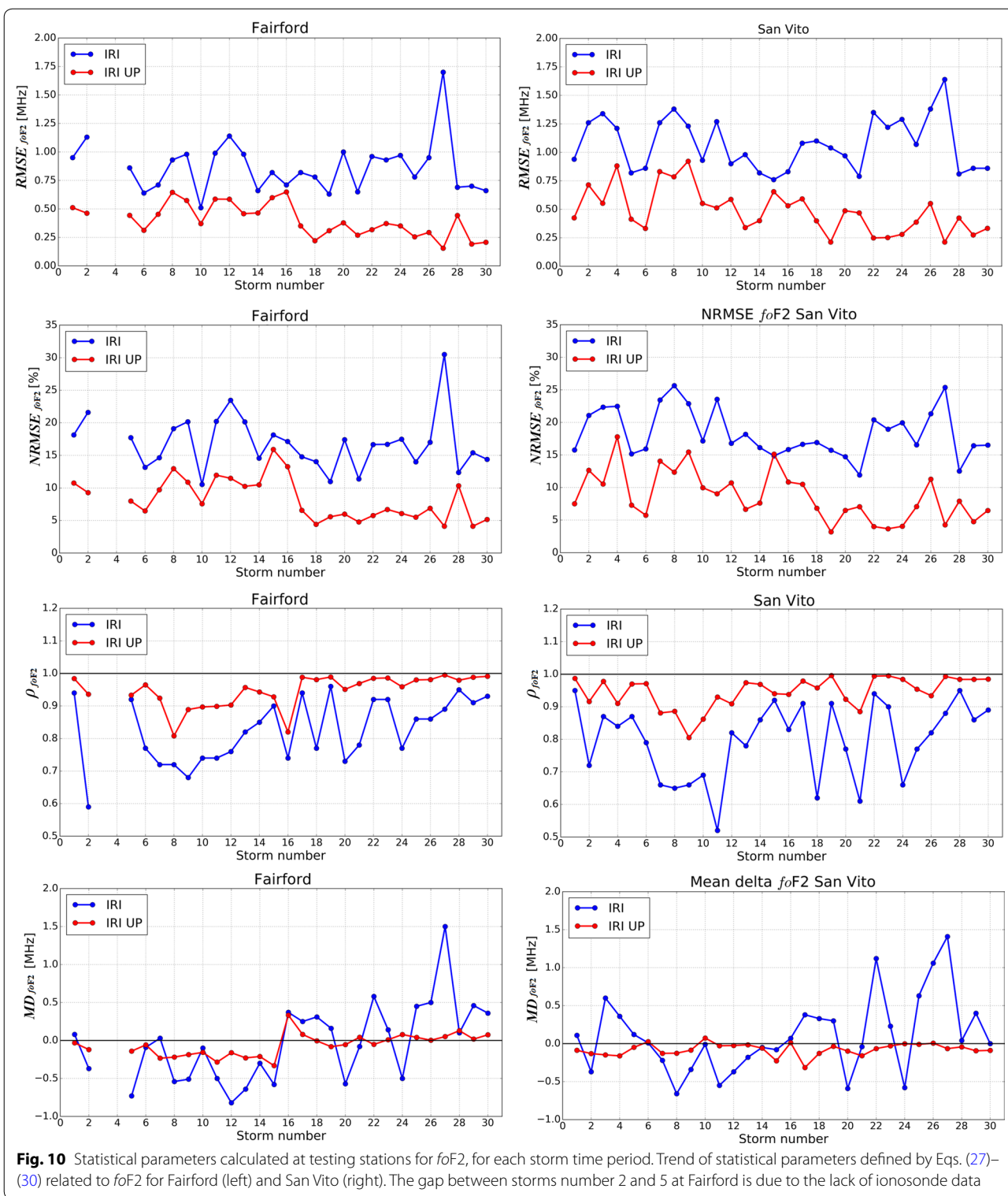
where i is the index running on the 5 possible variogram models, ss is the index running on the considered storms, $n_{ss,i}$ is the number of times the i variogram model is declared as the “winner” by the NQCR for the storm ss , and N_{ss} is the total number of epochs included in each single storm ss .

Figures 13a and 14a show the winning percentage of each variogram model for each single storm, for $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ respectively; Figs. 13b and 14b show the same percentage for the complete storm set.

Discussion, conclusions, and future developments

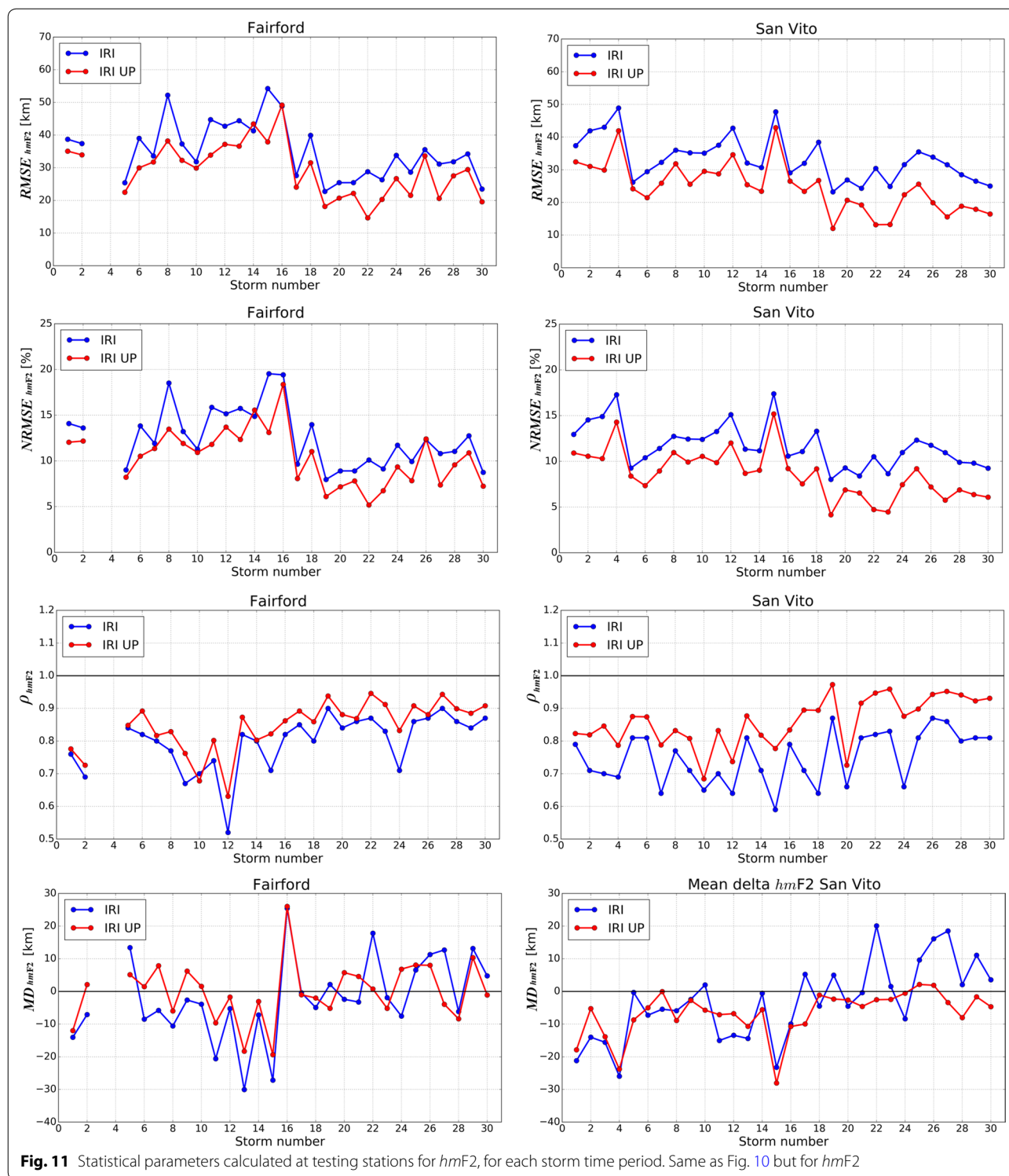
In this investigation the IRI UP method was upgraded by applying the F algorithm and the NQCR procedure to select the best variogram model.

The F algorithm described by Eqs. (1)–(4) has proven to be very effective in disregarding ionosonde data which, once assimilated, would affect badly the modeling of $IG_{12\text{eff}}$ and $R_{12\text{eff}}$, leading to unrealistic $foF2$ and $M(3000)F2$ maps and, consequently, to unlikely $hmF2$ maps (Pignalberi et al. 2018a, b). It must be noted that considering five standard deviations and thresholds values for the standard deviation equal to 0.5 MHz, for $foF2$, and 0.15, for $M(3000)F2$, are subjective choices, aiming to remove especially those measurements which are clearly out of range (spikes), as shown in Fig. 2.



Using residuals it is possible to define the statistical parameters Q_1 and Q_2 along with their PDF and CPE, which are sketched in Figs. 4 and 5, respectively, for a number of reference stations equal to 12. In these figures,

the two thresholds $|\sigma|$ and $|2\sigma|$ for Q_1 , and $[L-U]$ and $[L^{1st}-L^{3rd}]$ for Q_2 , have been deliberately taken very different from each other, to highlight how the choice of the threshold value can have an important impact on



the number of discarded variogram models. In fact, in Figs. 8 and 9, for the storm number 12 (August 23–28, 2005), it clearly emerges that the number of rejected variogram models is relatively large when the threshold is lowered from $|\sigma|$ to $|\sigma|$ (for the statistical test Q_1) and

from $[L-U]$ to $[L^{1st}-U^{3rd}]$ (for the statistical test Q_2). The threshold effect is however much less evident in the case of the storm number 27 (October 6–11, 2015), for which a limited number of variograms are discarded when reducing the threshold.

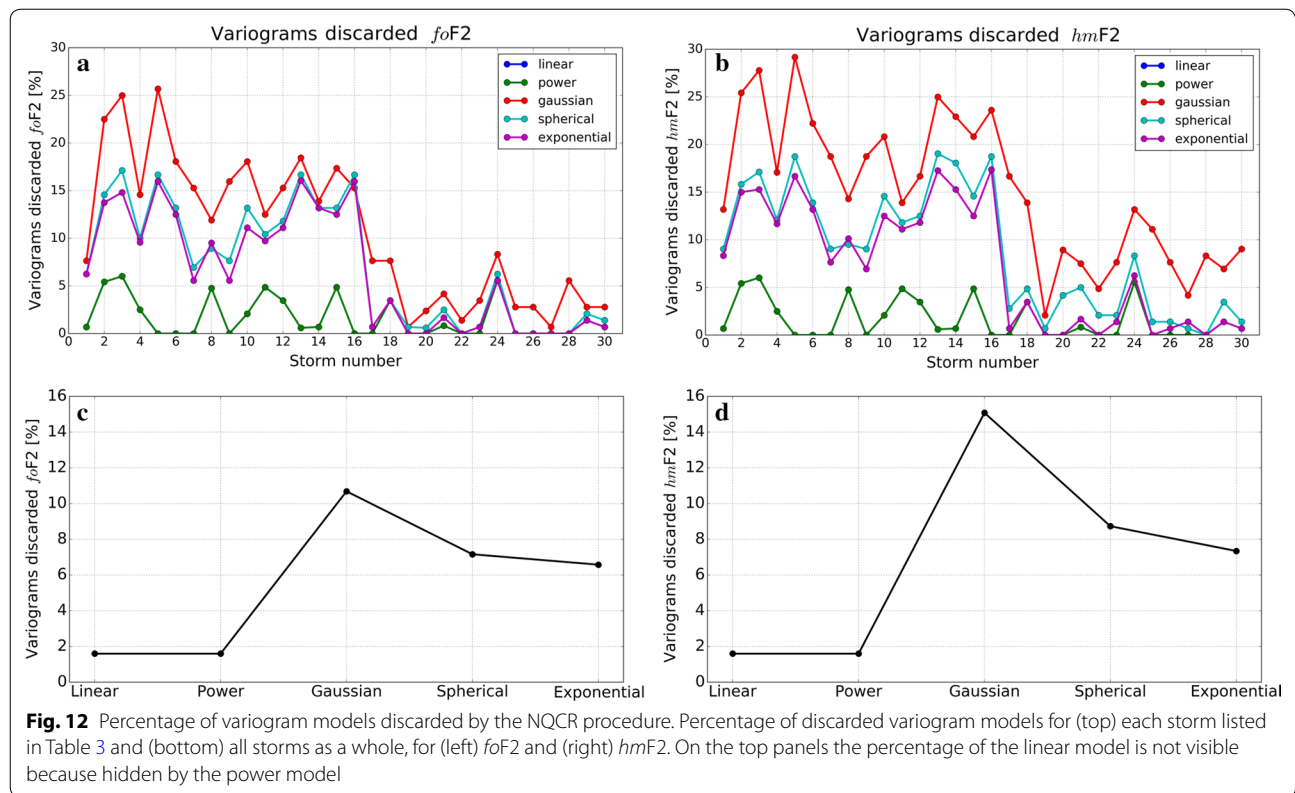


Table 4 Statistical validation of *foF2* as modeled by IRI UP and IRI

Station	Ionospheric characteristic	Model	RMSE (MHz)	NRMSE (%)	ρ	MMD _x AV [MHz]
Fairford	<i>foF2</i>	IRI UP— <i>linear</i>	0.474	9.40	0.927	0.137
		IRI UP— <i>power</i>	0.471	9.37	0.929	0.139
		IRI UP— <i>Gaussian</i>	0.467	9.28	0.920	0.152
		IRI UP— <i>spherical</i>	0.464	9.23	0.926	0.146
		IRI UP— <i>exponential</i>	0.461	9.14	0.927	0.144
		IRI	0.865	16.85	0.824	0.415
		<i>IRI UP</i>	0.401	8.25	0.947	0.122
San Vito	<i>foF2</i>	IRI UP— <i>linear</i>	0.558	9.74	0.927	0.088
		IRI UP— <i>power</i>	0.555	9.69	0.927	0.090
		IRI UP— <i>Gaussian</i>	0.538	9.38	0.922	0.083
		IRI UP— <i>spherical</i>	0.539	9.39	0.927	0.088
		IRI UP— <i>exponential</i>	0.538	9.36	0.927	0.090
		IRI	1.075	18.37	0.797	0.374
		<i>IRI UP</i>	0.485	8.70	0.946	0.084

Statistical results for *foF2* obtained at the two truth sites of Fairford and San Vito for: IRI, IRI UP running with a fixed variogram model (IRI UP—*variogram model* chosen in the table), and IRI UP embedding the NQCR procedure (italics)

This is probably due to the different numbers of available reference stations used in the assimilation, which for the period October 6–11, 2015 ($n = 12$), is larger than that for the period August 23–28, 2005 ($n = 8$), thus

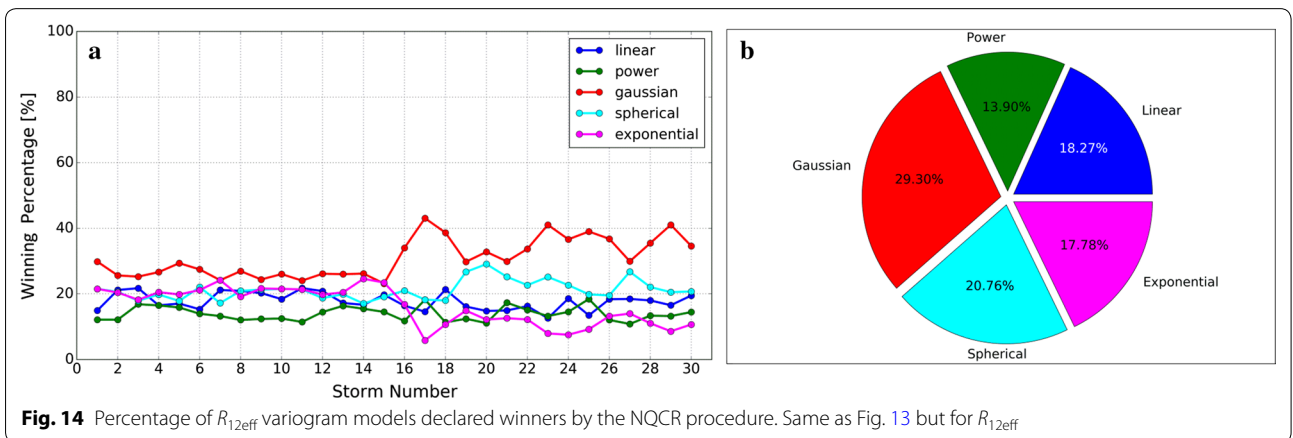
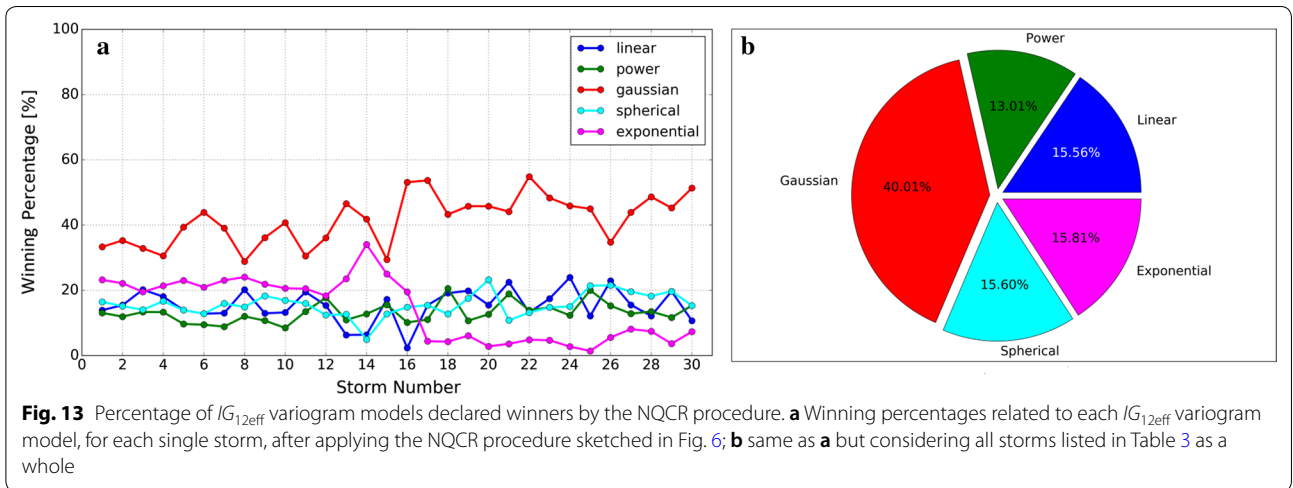
allowing a better representation of the spatial gradients over the area under study.

When the distribution of semivariance values calculated through Eq. (6) is such that the experimental

Table 5 Statistical validation of hmF2 characteristic as modeled by IRI UP and IRI

Station	Ionospheric characteristic	Variogram model	RMSE (km)	NRMSE (%)	ρ	MMD _x AV (km)
Fairford	hmF2	IRI UP—linear	31.567	11.31	0.817	7.361
		IRI UP—power	31.552	11.27	0.819	7.376
		IRI UP—Gaussian	30.334	10.83	0.796	7.457
		IRI UP—spherical	30.744	10.97	0.811	7.388
		IRI UP—exponential	30.844	11.00	0.811	7.479
		IRI	35.241	12.57	0.797	9.879
		IRI UP	29.369	10.44	0.845	6.844
San Vito	hmF2	IRI UP—linear	30.132	10.62	0.798	7.469
		IRI UP—power	30.197	10.64	0.799	7.674
		IRI UP—Gaussian	28.193	10.19	0.780	8.162
		IRI UP—spherical	29.805	10.50	0.783	8.053
		IRI UP—exponential	29.685	10.46	0.792	7.901
		IRI	33.252	11.70	0.749	9.577
		IRI UP	24.671	8.62	0.859	6.959

Same as Table 4 but for hmF2



variogram cannot be adequately represented by any of the variogram models defined in “[The Kriging interpolation method: a brief recall](#)” section, we are forced to discard the variogram modeling because it would lead to unrealistic maps. Figure 3 shows some examples of discarded variogram models for $IG_{12\text{eff}}$ and $R_{12\text{eff}}$; in these two cases the fitting function is an approximately straight horizontal line whatever is the considered variogram model, and the corresponding maps are not able to reproduce $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ values over the reference stations. This is due to the very large nugget values ($c_0 = 34.8$ for $IG_{12\text{eff}}$ and $c_0 = 98.6$ for $R_{12\text{eff}}$) describing the microscale variability. This implies that colored dots, marking the reference stations, are in contrast with the colors characterizing the regions around the reference stations, which means that the maps are not realistic since do not match the measured data.

On the other hand, if in Fig. 3 we look at the statistical parameters related to these two experimental variograms, the associated values of Q_1 (0.49 for $IG_{12\text{eff}}$, 0.37 for $R_{12\text{eff}}$) and Q_2 (0.91 for $IG_{12\text{eff}}$, 1.07 for $R_{12\text{eff}}$) are such that the two experimental variograms do not pass the Q_1 and Q_2 statistical test defined in “[Q₁ statistics](#)” and “[Q₂ statistics](#)” sections, and hence, they must be rejected along with their corresponding maps.

On the contrary, when the nugget effect is not so relevant and the distribution of semivariance values can be well described by most variogram models, the experimental variograms produce realistic maps. This is what happens for the case shown in Fig. 7 where the null nugget effect ($c_0 = 0$ for $IG_{12\text{eff}}$ and $R_{12\text{eff}}$), and the associated values of Q_1 (0.22 for $IG_{12\text{eff}}$, 0.49 for $R_{12\text{eff}}$) and Q_2 (0.98 for $IG_{12\text{eff}}$, 0.76 for $R_{12\text{eff}}$) are such that the two experimental variograms pass both the Q_1 and the Q_2 statistical test. In this case, the corresponding maps show $IG_{12\text{eff}}$ and $R_{12\text{eff}}$ values, over the reference stations, compatible with those of the regions close to the reference stations.

The preliminary quality check based on the exponent s sets always the same threshold value $s_{\text{thrs}} = 0.1$ whatever is the epoch and variogram model under study, without taking into account that the goodness of the variogram model depends also on the number n of available ionosonde data which are going to be considered in the UKM. This significant limitation characterizing the first version of IRI UP is overcome because, through the NQCR procedure, the quality check now depends on the n value that can change from epoch to epoch. Moreover, another essential aspect that should be considered when NQCR is applied is that, from a scientific point of view, “winning” variogram models are more reliable than the ones which have passed only the $s \geq s_{\text{thrs}}$ test. As it is easy to realize looking at Figs. 8 and 9, the number of variogram models discarded by the NQCR procedure depends on

the established thresholds values. As a general rule, if, at a given epoch, the number of reliable ionosonde data to be assimilated is relatively large, we can choose a more selective threshold, thus providing ionospheric characteristics maps with a high confidence level. In the event that the number of reliable ionosonde data is lower, we have to increase the threshold and in this case a map can be provided, but at a lower confidence level. It is clear that when there are very few reliable data to be assimilated, we cannot provide a statistically significant updated map. In this case the IRI UP method is not applicable and we rely on the IRI background map.

IRI UP and IRI prediction maps of the ionospheric characteristics $foF2$ and $M(3000)F2$, relative to 30 geomagnetic storms occurred between January 1, 2004, and December 31, 2016, are used to generate IRI UP and IRI prediction maps of $hmF2$ using Bilitza et al. (1979). The values of $foF2$ and $hmF2$ extracted at the two truth sites of Fairford and San Vito from the corresponding IRI UP and IRI prediction maps have been compared with the measurements, to compare IRI UP and IRI performance.

The obtained results confirm those shown in Pignalberi et al. (2018a, b) for the St. Patrick geomagnetic storm; in fact, as shown in Figs. 10 and 11, as well as in Tables 4 and 5, they indicate that IRI UP performs significantly better than IRI, for all the 30 considered cases.

Moreover, results of Tables 4 and 5 show that when IRI UP is applied deciding a priori the variogram model, and then considering only those cases passing the first two steps of the NQCR procedure, a clear difference among the various variogram models does not emerge.

In fact, NRMSE values for $foF2$ and $hmF2$ range, respectively, between 9.14 and 9.40% and between 10.83 and 11.31% at Fairford and between 9.36 and 9.74% and between 10.19 and 10.64% at San Vito, while MMD_X AV values for $foF2$ and $hmF2$ range, respectively, between 0.137 and 0.152 MHz and between 7.361 and 7.479 km at Fairford and between 0.083 and 0.090 MHz and between 7.469 and 8.162 km at San Vito.

When the IRI UP method runs with the complete NQCR procedure, its performance shows a noticeable improvement at the considered truth sites, for both $foF2$ and $hmF2$. In fact, in this case, NRMSE values for $foF2$ and $hmF2$ are 8.25% and 10.44% at Fairford and 8.70% and 8.62% at San Vito, while MMD_X AV values for $foF2$ and $hmF2$ are 0.122 MHz and 6.844 km at Fairford and 0.084 MHz and 6.959 km at San Vito.

NRMSE and MMD_X AV values obtained by the IRI UP method embedding the NQCR procedure, systematically smaller than the ones obtained by both the IRI UP method running with a fixed variogram model and the IRI model, prove that the NQCR procedure embedded in

the IRI UP method is effective in providing more precise and accurate results.

In general, from Figs. 10 and 11 it emerges that IRI UP performs slightly better for *foF2* than for *hmF2*. This happens for different reasons:

1. *foF2* assimilated data are generally more reliable than the corresponding *M(3000)F2* data. This is mostly due to the fact that ionograms can be characterized by multiple reflections of the F2 layer, and when this happens an autoscaling program can be misled and the second-order reflection can be identified as the real trace; in this case the *foF2* value is usually not affected by a significant error, while the *M(3000)F2* is significantly wrongly scaled (Scotto and Pezzopane 2008);
2. the spherical harmonic expansion used by IRI to describe the *foF2* and *M(3000)F2* spatial behavior (see Eq. (1) in Pignalberi et al. 2018a) stops when the maximum order of the harmonics is equal to 76, for *foF2*, and to 49, for *M(3000)F2*, which means that *foF2* maps present a higher spatial resolution than *M(3000)F2* ones;
3. *hmF2* is calculated applying the empirical formula of Bilitza et al. (1979), and this implies that the error characterizing *hmF2* depends on the errors relative to *foF2*, *M(3000)F2*, *foE* (the E-layer critical frequency), and $R_{12\text{eff}}$; therefore, the error propagation leads to an error associated with *hmF2* which is intrinsically larger than that of *foF2*;
4. last but not least, *foF2* predictions are based on the $IG_{12\text{eff}}$ index, which is an ionospheric index because it is “built” just starting from *foF2* values recorded at several ionospheric stations (Liu et al. 1983), while *M(3000)F2* predictions, which come into play to calculate *hmF2*, are not based on an ionospheric index, but on $R_{12\text{eff}}$.

Another positive aspect of the IRI UP method is that the experimental variogram has a higher spatial variability the larger is the number n of data assimilated from the reference stations, so that corresponding maps are statistically more reliable and are not discarded. This situation is clear from the results of Fig. 12 where, for each kind of variogram model, a decreasing trend of the number of discarded maps is observed starting from the storm number 17, corresponding to 2012, i.e., the year from which the number of available ionospheric stations maximizes (see Table 2).

It is also to be noted that stationary variogram models (gaussian, spherical, and exponential) are more sensitive than non-stationary ones to the number n of assimilated data; this is probably due to their more complex

mathematical formulation which requires a greater value of n to represent more adequately the spatial correlations between measurements, for every spatial scale. In fact, considering all storms as a whole, it results that percentages of rejected stationary variograms for *foF2* (Fig. 12c) and *hmF2* (Fig. 12d) are greater than those related to non-stationary variograms (linear and power). This result is probably due to the cumulative effect of the first 16 storms listed in Table 3, which are relative to years characterized by a low value of n . Nevertheless, in Fig. 12a, b, the trend observed starting from the storm number 17, corresponding to the years for which n is increased ($n = 12$), suggests that the percentages of stationary and non-stationary discarded variogram models may converge to similar values as the number of stations increases.

The winning percentages shown in Figs. 13b and 14b indicate that among the linear, power, spherical, and exponential variogram models there is not a clear predominance of one model over another, and that the gaussian variogram model shows the higher percentages, $\approx 40\%$ and 29% , for $IG_{12\text{eff}}$ and $R_{12\text{eff}}$, respectively. This fact is surprising, because if we consider Fig. 12 the Gaussian model is the one most rejected. This means that the Gaussian variogram model passes more difficult the Q_1 and Q_2 statistical tests, but when this happens, it is more likely to be the best according to the NQCR procedure. A fact that clearly emerges also when each single storm is considered (Figs. 13a, 14a).

It is worth noting that the achieved results have been obtained without explicitly considering the hour of the day. In fact, with regard to the future developments, a very important aspect that will have to be considered is that ionospheric characteristics depend inherently on the hour of the day. At the solar terminator (hours around sunrise and sunset), regardless of the large-scale latitudinal spatial gradients, the electron density spatial distribution manifests also large longitudinal gradients on small spatial scales. On the contrary, under ionospheric stationary conditions (hours around noon and midnight) the electron density spatial distribution does not show large longitudinal differences and hence the ionospheric variability is characterized by small gradients on both large and small spatial scale. These considerations imply that the choice of the variogram model should depend on the hour of the day.

In principle, using the UKM method described in “The Kriging interpolation method: a brief recall” section, the large-scale spatial behavior of the considered ionospheric characteristic is represented by the term m of Eq. (14), which allows to characterize also those sectors that are far from reference stations where the data assimilation takes place, without necessarily using non-stationary variogram models (Pignalberi et al. 2018a, b).

This means that stationary variogram models (Gaussian, spherical, and exponential) could describe better the situations at the sunrise and sunset hours, when the ionospheric characteristic shows large gradients on small spatial scale.

Vice versa, in the hours around noon, when the small-scale spatial gradients are not so important, non-stationary variogram models (linear and power) are likely more indicated.

In the light of these considerations, a careful and detailed study, aimed to investigate how the hour of the day affects the choice of the variogram model, is of crucial importance to get useful clues in order to improve further the goodness of the variogram model selection and consequently the quality of prediction maps of the main ionospheric characteristics.

The results achieved in this investigation prove, however, that reliable and trustworthy updated maps of the main ionospheric characteristics can be provided with a satisfactory degree of confidence, especially under moderate, strong and severe geomagnetic storm conditions. This means that IRI UP method, embedding the F algorithm and the NQCR procedure, represents an interesting approach to Space Weather forecast in the ionospheric domain, for any region characterized by an adequately distributed network of ionosondes.

Abbreviations

ARTIST: Automatic Real-Time Ionogram Scaler with True height; CPF: cumulative probability function; CS: Confidence Score; HF: high frequency; IRI: International Reference Ionosphere; IRI UP: International Reference Ionosphere UPdate; ISP: IRI-SIRMUP-P; KIM: Kriging interpolation method; MD: Mean Delta; MMDAV: Mean Mean Delta absolute value; MPD: main phase day; NQCR: new quality check routine; NRMSE: normalized root mean square error; PDF: probability density function; RMSE: root mean square error; SIRMUP: Simplified Ionospheric Regional Model UPdating; UKM: Universal Kriging method; UT: universal time; UV: ultraviolet.

Authors' contributions

AP conceived the study, developed the IRI UP method improvements described in the paper, and made the statistical analysis needed to validate the new version of the method. MPI drafted the manuscript, gave important insights about the developed statistical procedures, and actively participated to the discussion of the results. MPe participated in the discussion of the statistical results and helped to draft the manuscript. RR gave insights about the statistical analysis, participated in the discussion of the statistical results, and revised the manuscript. All authors read and approved the final manuscript.

Author details

¹ Dipartimento di Fisica e Astronomia, Università di Bologna "Alma Mater Studiorum", Bologna, Italy. ² Istituto Nazionale di Geofisica e Vulcanologia, 00143 Rome, Italy.

Acknowledgements

This publication uses data from 14 ionospheric observatories in Europe, made available via the public access portal of the Digital Ionogram Database of the Global Ionosphere Radio Observatory in Lowell, MA. The authors are indebted to observatory directors and ionosonde operators for heavy investments of their time, effort, expertise, and funds needed to acquire and provide measurement data to academic research. The IRI team is acknowledged for developing and maintaining the IRI model and for giving access to the corresponding Fortran code via the IRI Web site (<http://irimodel.org/>).

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Ionosonde data used in this study are publicly available at the Digital Ionogram Database (<http://ulcar.uml.edu/DIDBase/>) and can be freely downloaded by means of the SAO Explorer software developed by the University of Massachusetts, Lowell (<http://ulcar.uml.edu/SAO-X/SAO-X.html>). Geomagnetic indices were downloaded from the OMNIWeb Data Explorer—NASA site at <https://omniweb.gsfc.nasa.gov/form/dx1.html>. The IRI model Fortran code is available via the IRI Web site (<http://irimodel.org/>). The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

This work is funded by the Department of Physics and Astronomy of the University of Bologna via a doctoral scholarship for the Geophysics Doctorate School. The Istituto Nazionale di Geofisica e Vulcanologia (INGV) in Rome made available human and technological resources, and working space, needed to carry out this work.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 1 August 2018 Accepted: 8 November 2018

Published online: 19 November 2018

References

- Angling MJ, Khattatov B (2006) Comparative study of two assimilative models of the ionosphere. *Radio Sci* 41:RS5520. <https://doi.org/10.1029/2005RS003372>
- Bibl K, Reinisch BW (1978) The universal digital ionosonde. *Radio Sci* 13:519–530. <https://doi.org/10.1029/RS013i003p00519>
- Bilitza D, Reinisch BW (2008) International reference ionosphere 2007: improvements and new parameters. *Adv Space Res* 42(4):599–609. <https://doi.org/10.1016/j.asr.2007.07.048>
- Bilitza D, Sheikh M, Eyfrig R (1979) A global model for the height of the F2-peak using M3000 values from the CCIR numerical map. *Telecommun J* 46:549–553
- Bilitza D, McKinnell LA, Reinisch B, Fuller-Rowell T (2011) The International Reference Ionosphere today and in the future. *J Geod* 85:909–920. <https://doi.org/10.1007/s00190-010-0427-x>
- Bilitza D, Altadill D, Zhang Y, Mertens C, Truhlik V, Richards P, McKinnell LA, Reinisch B (2014) The International Reference Ionosphere 2012—a model of international collaboration. *J Space Weather Space Clim* 4:A07. <https://doi.org/10.1051/swsc/2014004>
- Bilitza D, Altadill D, Truhlik V, Shubin V, Galkin I, Reinisch B, Huang X (2017) International Reference Ionosphere 2016: from ionospheric climate to real-time weather predictions. *Space Weather* 15:418–429. <https://doi.org/10.1002/2016SW001593>
- Buonsanto MJ (1999) Ionospheric storms—a review. *Space Sci Rev* 88:563–601
- Decker DT, McNamara LF (2007) Validation of ionospheric weather predicted by global assimilation of ionospheric measurements (GAIM) models. *Radio Sci* 42:RS4017. <https://doi.org/10.1029/2007RS003632>
- Galkin IA, Reinisch BW (2008) The new ARTIST 5 for all digisondes. In: Ionosonde Network Advisory Group Bulletin, in: IPS Radio and Space Services, Surry Hills, NSW, Australia, vol 69, pp 1–8. <http://www.ips.gov.au/IPSHo sted/INAG/web-69/2008/artist5-inag.pdf>
- Galkin IA, Reinisch BW, Huang X, Bilitza D (2012) Assimilation of GIRO data into a real-time IRI. *Radio Sci* 47:7. <https://doi.org/10.1029/2011RS004952>
- Houminer Z, Bennett JA, Dyson PL (1993) Real-time ionospheric model updating. *J Electr Eng Aust* 13(2):99–104
- Khmyrov GM, Galkin IA, Kozlov AV et al (2008) Exploring digisonde ionogram data with SAO-X and DIDBase. In: Proceedings of AIP conference radio sounding and plasma physics, vol 974, pp 175–185. <https://doi.org/10.1063/1.2885027>

- Kitanidis PK (1997) Introduction to geostatistics: application to hydrogeology. Cambridge University Press, Cambridge
- Kouris SS, Fotiadis DN (2002) Ionospheric variability: a comparative statistical study. *Adv Space Res* 29(6):977–985. [https://doi.org/10.1016/S0273-1177\(02\)00045-5](https://doi.org/10.1016/S0273-1177(02)00045-5)
- Kouris SS, Fotiadis DN, Xenos TD (1998) On the day-to-day variation of f_oF_2 and $M(3000)F_2$. *Adv Space Res* 22(6):873–876. [https://doi.org/10.1016/S0273-1177\(98\)00116-1](https://doi.org/10.1016/S0273-1177(98)00116-1)
- Liu RY, Smith PA, King JW (1983) A new solar index which leads to improved f_oF_2 predictions using the CCIR Atlas. *Telecommun J* 50:408–414
- McNamara LF, Decker DT, Welsh JA, Cole DG (2007) Validation of the Utah State University global assimilation of ionospheric measurements (GAIM) model predictions of the maximum usable frequency for a 3000 km circuit. *Radio Sci* 42:RS3015. <https://doi.org/10.1029/2006RS003589>
- McNamara LF, Baker CR, Decker DT (2008) Accuracy of USU-GAIM specifications of f_oF_2 and $M(3000)F_2$ for a worldwide distribution of ionosonde locations. *Radio Sci* 43:RS1011. <https://doi.org/10.1029/2007RS003754>
- McNamara LF, Retterer JM, Baker CR, Bishop GJ, Cooke DL, Roth CJ, Welsh JA (2010) Longitudinal structure in the CHAMP electron densities and their implications for global ionospheric modeling. *Radio Sci* 45:RS2001. <https://doi.org/10.1029/2009RS004251>
- McNamara LF, Bishop GJ, Welsh JA (2011) Assimilation of ionosonde profiles into a global ionospheric model. *Radio Sci* 46:RS2006. <https://doi.org/10.1029/2010RS004457>
- Menvielle M, Berthelier A (1991) The K-derived planetary indices: description and availability. *Rev Geophys* 29(3):415–432
- Nava B, Radicella SM, Azpillicueta F (2011) Data ingestion into NeQuick 2. *Radio Sci* 46(6):RS0D17. <https://doi.org/10.1029/2010RS004635>
- Pezzopane M, Scotto C (2005) The INGV software for the automatic scaling of f_oF_2 and $MUF(3000)F_2$ from ionograms: a performance comparison with ARTIST 4.01 from Rome data. *J Atmos Sol Terr Phys* 67(12):1063–1073. <https://doi.org/10.1016/j.jastp.2005.02.022>
- Pezzopane M, Scotto C (2007) The automatic scaling of critical frequency f_oF_2 and $MUF(3000)F_2$: a comparison between Autoscala and ARTIST 4.5 on Rome data. *Radio Sci* 42:RS4003. <https://doi.org/10.1029/2006RS003581>
- Pezzopane M, Scotto C, Tomasik E, Krashenninikov I (2009) Autoscala: an aid for different ionosondes. *Acta Geophys* 58(3):513–526. <https://doi.org/10.2478/s11600-009-0038-1>
- Pezzopane M, Pietrella M, Pignatelli A, Zolesi B, Cander LR (2011) Assimilation of autoscaled data and regional and local ionospheric models as input sources for real-time 3-D International Reference Ionosphere modeling. *Radio Sci* 46:RS5009. <https://doi.org/10.1029/2011RS004697>
- Pezzopane M, Pietrella M, Pignatelli A, Zolesi B, Cander LR (2013) Testing the three-dimensional IRI-SIRMUP-P mapping of the ionosphere for disturbed periods. *Adv Space Res* 52(10):1726–1736. <https://doi.org/10.1016/j.asr.2012.11.028>
- Pietrella M (2015) A software package generating long term and near real time predictions of the critical frequencies of the F2 layer over Europe and its applications. *Int J Geosci* 6(4):373–387. <https://doi.org/10.4236/ijg.2015.64029>
- Pignalberi A, Pezzopane M, Tozzi R, De Michelis P, Coco I (2016) Comparison between IRI and preliminary swarm langmuir probe measurements during the St. Patrick storm period. *Earth Planets Space* 68(1):93. <https://doi.org/10.1186/s40623-016-0466-5>
- Pignalberi A, Pezzopane M, Rizzi R, Galkin IA (2018a) Effective solar indices for ionospheric modeling: a review and a proposal for a real-time regional IRI. *Surv Geophys* 39:125–167. <https://doi.org/10.1007/s10712-017-9438-y>
- Pignalberi A, Pezzopane M, Rizzi R, Galkin IA (2018b) Correction to: Effective solar indices for ionospheric modeling: a review and a proposal for a real-time regional IRI. *Surv Geophys* 39:169. <https://doi.org/10.1007/s10712-017-9453-z>
- Reinisch BW, Galkin IA (2011) Global ionospheric radio observatory (GIRO). *Earth Planets Space* 63(4):377–381. <https://doi.org/10.5047/eps.2011.03.001>
- Reinisch BW, Huang X (1983) Automatic calculation of electron density profiles from digital ionograms: 3. Processing of bottom side ionograms. *Radio Sci* 18(3):477–492. <https://doi.org/10.1029/RS018i003p00477>
- Reinisch BW, Huang X, Galkin IA, Paznukhov V, Kozlov A (2005) Recent advances in real-time analysis of ionograms and ionospheric drift measurements with digisondes. *J Atmos Sol Terr Phys* 67(12):1054–1062. <https://doi.org/10.1016/j.jastp.2005.01.009>
- Rostoker G (1972) Geomagnetic indices. *Rev Geophys Space Phys* 10:157
- Scotto C (2009) Electron density profile calculation technique for Autoscala ionogram analysis. *Adv Space Res* 44:756–766. <https://doi.org/10.1016/j.asr.2009.04.037>
- Scotto C, Pezzopane M (2002) A software for automatic scaling of f_oF_2 and $MUF(3000)F_2$ from ionograms. In: Proceedings of the XXVII general assembly of the international union of radio science, 17–24 August, Maastricht, The Netherlands. International Union of Radio Science, CD-ROM, Ghent
- Scotto C, Pezzopane M (2008) Removing multiple reflections from the F2 layer to improve Autoscala performance. *J Atmos Sol Terr Phys* 70(15):1929–1934. <https://doi.org/10.1016/j.jastp.2008.05.012>
- Scotto C, Pezzopane M, Zolesi B (2012) Estimating the vertical electron density profile from an ionogram: on the passage from true to virtual heights via the target function method. *Radio Sci* 47:RS1007. <https://doi.org/10.1029/2011RS004833>
- Shim JS, Kuznetsova M, Rastatter L et al (2011) CEDAR electrodynamic thermosphere ionosphere (ETI) challenge for systematic assessment of ionosphere/thermosphere models: nmF_2 , hmF_2 , and vertical drift using ground-based observations. *Space Weather* 9:S12003. <https://doi.org/10.1029/2011SW000727>
- Thompson DC, Scherliess L, Sojka JJ, Schunk RW (2006) The Utah State University Gauss–Markov Kalman filter of the ionosphere: the effect of slant TEC and electron density profile data on model fidelity. *J Atmos Sol Terr Phys* 68(9):947–958. <https://doi.org/10.1016/j.jastp.2005.10.011>
- Tsagouri I, Zolesi B, Belehaki A, Cander LR (2005) Evaluation of the performance of the real-time updated simplified ionospheric regional model for the European area. *J Atmos Sol-Terr Phys* 67(12):1137–1146. <https://doi.org/10.1016/j.jastp.2005.01.012>
- Zolesi B, Cander LR (2014) Ionospheric prediction and forecasting. Springer, Berlin
- Zolesi B, Belehaki A, Tsagouri I, Cander LR (2004) Real-time updating of the simplified ionospheric regional model for operational applications. *Radio Sci* 39:RS2011. <https://doi.org/10.1029/2003RS002936>
- Zuccheretti E, Tutone G, Sciacca U, Bianchi C, Arokiasamy BJ (2003) The new AIS-INGV digital ionosonde. *Ann Geophys* 46(4):647–659. <https://doi.org/10.4401/ag-4377>

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com