Article

Towards a Real-Time Description of the Ionosphere: A Comparison between International Reference Ionosphere (IRI) and IRI Real-Time Assimilative Mapping (IRTAM) Models

Alessio Pignalberi * , **Marco Pietrella and Michael Pezzopane**

Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Rome, Italy;
marco.pietrella@ingv.it (M.P.); michael.pezzopane@ingv.it (M.P.)
* Correspondence: alessio.pignalberi@ingv.it

Abstract: This paper focuses on a detailed comparison, based on the F2-layer peak characteristics foF2 and hmF2, between the International Reference Ionosphere (IRI), which is a climatological empirical model of the terrestrial ionosphere, and the IRI Real-Time Assimilative Mapping (IRTAM) procedure, which is a real-time version of IRI based on data assimilation from a global network of ionosondes. To perform such a comparison, two different kinds of datasets have been considered: (1) foF2 and hmF2 as recorded by 40 ground-based ionosondes spread all over the world from 2000 to 2019; (2) foF2 and hmF2 from space-based COSMIC/FORMOSAT-3 radio occultation measurements recorded from 2006 to 2018. The aim of the paper is to understand whether and how much IRTAM improves IRI foF2 and hmF2 outputs for different locations and under different diurnal, seasonal, solar and magnetic activity conditions. The main outcomes of the study are: (1) when ionosonde observations are considered for validation, IRTAM significantly improves the IRI foF2 modeling both in accuracy and precision, while a slight improvement in the IRI hmF2 modeling is observed for specific locations and conditions; (2) when COSMIC observations are considered for validation, no noticeable improvement is observed from the IRTAM side for both foF2 and hmF2. Indeed, IRTAM can improve the IRI foF2 description only nearby the assimilated ionosonde locations, while the IRI hmF2 description is always more accurate and precise than IRTAM one.

Keywords: International Reference Ionosphere (IRI); IRI Real-Time Assimilative Mapping (IRTAM); foF2; hmF2; ionosondes data; COSMIC/FORMOSAT-3 radio occultation data; Space Weather

1. Introduction

Space Weather events can have a deep negative impact on the technological systems, such as power systems, satellites, Global Positioning System (GPS), pipelines, and communication cables. The damages suffered by these systems, on which our society is nowadays greatly dependent, besides leading to very high costs, can also significantly affect human life [1,2]. Therefore, in this context, in recent years the near real-time specification of the ionosphere has become more and more important to nowcast and possibly mitigate the adverse consequences of Space Weather events. For this purpose, several models able to assimilate real-time ionospheric measurements have been recently proposed [3–16].

Some climatological models already existent were adapted for the quasi real-time assimilation of ionospheric data. One outstanding example is the IRI-based Real-Time Assimilative Model (IRTAM) [17,18] that, by ingesting ionosonde-derived F2-layer peak parameters values, updates the underlying empirical global climatological knowledge of the ionosphere provided by the International Reference Ionosphere (IRI) model [19,20], thus supplying a global real-time representation of the ionosphere. The assimilation of real-time measurements in a background empirical model is one of the most applied and fruitful methodologies for the real-time specification of the ionospheric electron density. In

this context, the knowledge of the large-scale climatological behavior of the ionosphere provided by the underlying background empirical model is complemented with the small-scale weather information provided by real-time assimilated data. The effectiveness and quality of such data-assimilation procedures is critically dependent on the applied algorithm, on the quality, spatial distribution, and availability of assimilated data, and of course on the underlying background empirical model.

Empirical models, such as IRI, are based on analytical formulations whose numerical coefficients are obtained on the basis of the underlying datasets; as a consequence, when new datasets are released, it is of utmost importance to validate the model against new data and eventually recalculate the model's coefficients with the inclusion of the newest data. This validation and recalculation scheme is an ongoing process for empirical models and leads to the continuous improvement of the model itself. Over the years, IRI underwent many validation studies and comparisons with other ionospheric models [21–28]; on the contrary, validations of the IRTAM model are restricted to the works by Vesnin [29] and Galkin et al. [18] for specific locations and conditions. Due to the ever-growing importance that IRTAM is gaining as the most used and affirmed real-time specification of the ionosphere, it is important to validate its performances against large and different datasets to quantify the improvement made by IRTAM in the description of ionospheric weather when compared to the climatological representation made by IRI.

In the present paper, a global validation of the ionosphere F2-layer peak characteristics as modeled by IRI and IRTAM is presented. Specifically, the IRI and IRTAM models, the latter assimilating both the F2-layer ordinary critical frequency (*fo*F2) and the F2-layer peak height (*hm*F2) from ground-based ionosondes, have been validated according to two different datasets: (1) *fo*F2 and *hm*F2 ground-based ionosonde observations recorded from 1 January 2000 to 31 December 2019 at 40 ionospheric stations spread in both hemispheres; (2) *fo*F2 and *hm*F2 derived from Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC/FORMOSAT-3) radio occultation (RO) observations from 22 April 2006 to 31 December 2018. In order to assess the performances of both IRI and IRTAM, different statistical metrics have been estimated. Corresponding results are represented in the form of grids of values as a function of the local time (LT) and month of the year, for three different levels of solar activity, for the different ionosonde locations. Moreover, the spatial variation of the calculated statistical values has been investigated through the COSMIC dataset. Comprehensive statistical results are provided for the entire ionosonde and COSMIC datasets as distribution of residuals, density plots, and residuals deviation ratio values, allowing us to draw a complete picture of IRI and IRTAM performances in the description of the F2-layer peak characteristics. As far as we know, it is the first time that IRI and IRTAM are cross-validated on the basis of such a large dataset covering very different conditions and locations. Moreover, the use of *fo*F2 and *hm*F2 datasets from different measurement techniques, such as ionosonde and radio occultation, represents an added value in the validation process because it allows us to validate IRTAM against independent data (i.e., COSMIC RO data) and evaluate how much IRTAM is tied to the assimilated data from ionosondes.

A brief description of both IRI and IRTAM models will be provided in Section 2. An overview of the two different datasets used for validation and some information about the runs of IRI and IRTAM models are given in Section 3. The statistics metrics, the binning procedures, and the graphical representation of the results are the subject of Section 4. The validation results for *fo*F2 are described in Sections 5 and 6, while those for *hm*F2 are described in Sections 7 and 8; the validation shown in Sections 5 and 7 is based on ionosonde data, while that shown in Sections 6 and 8 is based on COSMIC RO data. Final analyses and considerations are the subject of Section 9, while the conclusive remarks are outlined in Section 10.

2. IRI and IRTAM Models: A Brief Recall

2.1. IRI

IRI is a project started in 1968 by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI) with the aim to develop an international standard for the terrestrial ionosphere. In April 2014, IRI became the official International Standardization Organization (ISO) standard for the ionosphere [20]. The IRI model is empirical, based on ground and space data, and provides monthly medians of the electron density, electron temperature, ion temperature, and ion composition in the altitude range of 60 km to 2000 km [19]. Additionally, it provides the vertical total electron content (vTEC) from the lower boundary to a user-specified upper boundary. Other IRI outputs include the vertical ion drift near the magnetic equator, the F1-layer and spread-F occurrence probability, and the representation of auroral boundaries.

The F2-layer peak plasma frequency *fo*F2, and the corresponding maximum electron density *Nm*F2, are two of the most important parameters when modeling the ionosphere and are related by the formula $Nm\text{F2} = 1.24 \times 10^{10}(fo\text{F2})^2$, where *Nm*F2 and *fo*F2 are, respectively, expressed in m^{-3} and MHz. Both parameters are very important for a reliable characterization of the ionosphere for both scientific and applicative purposes [2,30].

IRI proposes two options for modeling them: the model recommended by the Consultative Committee on International Radio (CCIR) of the International Telecommunication Union (ITU) and the model developed by a special URSI working group [31]. Both models use the same mathematical functions but are based on different datasets of ionosonde data and different methods to fill data gaps, especially over the ocean areas. The CCIR model is recommended mainly for land regions, while URSI is recommended when the region under investigation includes large ocean areas. Specifically, both models are based on a procedure of numerical mapping proposed by Jones and Gallet [32–34], which is based on a Fourier time series describing the diurnal variation of monthly medians of *fo*F2 observed at each of the ionosonde stations considered to develop the model (about 150 in total). Legendre special functions (see [35] for further details) are then used to represent the variation of the Fourier coefficients with geographic coordinates and modip, the modified dip coordinate introduced by Rawer [36] to better describe the magnetic field dependence of ionospheric parameters. As driver index both models use the 12-month running mean (IG_{12}) of the ionosonde-based Ionospheric Global (*IG*) index introduced by Liu et al. [37]. To describe the global daily behavior of *fo*F2, both CCIR and URSI models require 988 coefficients because the diurnal trend is described through a 6th order Fourier time series, then 13 time coefficients, and each time coefficient undergoes a harmonic spatial expansion to the 9th order, then 76 spatial coefficients. This reasoning is valid for a specific month. The seasonal variability is added by sorting data as a function of the month of the year; then, 12 datasets of 988 coefficients are obtained. Finally, the solar activity variation is described by calculating these coefficients for two levels of solar activity, $IG_{12} = 0$ for low solar activity and $IG_{12} = 100$ for high solar activity, and then performing a linear interpolation between them (see [14] for further details).

Since the IRI-2001 version, the *fo*F2 modeling is also accompanied by a storm option developed by Fuller-Rowell et al. [38] and Araujo-Pradere et al. [39,40] to represent the ionosphere under magnetically disturbed conditions. It is based on the 33 h prior history of the a_p magnetic index and gives reliable results, especially at mid latitudes.

Concerning the *hm*F2 modeling, IRI proposes three options. The first one has been developed by Bilitza et al. [41] with *M*(3000)F2 ionosonde data and is based on the anti-correlation between *hm*F2 and the propagation factor *M*(3000)F2 [35], with *M*(3000)F2 values from CCIR [35] or URSI [31] mapping procedures (see [14] for further details). The second and the third ones have been recently developed, respectively, by Altadill et al. [42] with *hm*F2 ionosonde data, and by Shubin et al. [43,44] with *hm*F2 ionosonde and COSMIC RO data, and they are both based on the spherical harmonic formalism.

2.2. Real-Time IRI and the IRTAM Method

In recent years, different data-assimilation techniques have been applied with the intention of improving the IRI output. The aim is to move from the climatological representation provided by the standard IRI model to a description of the ionospheric weather conditions based on the ingestion of real-time measurements.

Bilitza et al. [45] used worldwide ionosonde data from 1986 to 1989 to obtain equivalent *IG* indices. Komjathy et al. [46], Hernandez-Pajares et al. [47], Ssessanga et al. [48], and Habarulema and Ssessanga [49] used Global Navigation Satellite System data to determine R_{12} and IG_{12} equivalent indices. Recently, Pignalberi et al. [14,50,51] proposed a new data-assimilation method, based on ionosonde data, to update the IRI model in the European region through the calculation of an effective IG_{12}; their procedure has been recently updated by assimilating also vTEC values [52]. Pezzopane et al. [11,12] and Pietrella et al. [15] assimilating ionosonde F2-layer peak parameter measurements first determined an effective sunspot number which is used by the Simplified Ionospheric Regional Model [16], and then applied an interpolation technique to assimilate into IRI the full electron density profile recorded by ionosondes.

The assimilation can be carried out as a post-processing activity, without requiring a real-time analysis; however, in recent years, big steps forward have been taken towards a real-time IRI, performing a real-time assimilation. With regard to this, very good results have been achieved by Galkin et al. [17,18], who proposed the IRTAM method (http://giro.uml.edu/RTAM, accessed on 30 July 2021). IRTAM assimilates real-time *fo*F2, *hm*F2, B_0, and B_1 measurements from the worldwide network of Digisonde stations (the Global Ionospheric Radio Observatory—GIRO) and uses the URSI procedure [31] to represent the difference between data and model and update the 988 coefficients of the corresponding harmonic expansion for the specific time of the assimilation. As a final step, IRTAM generates near real-time maps of *fo*F2, *hm*F2, B_0, and B_1 every 15 min.

IRTAM uses the URSI procedure [31] to describe both *fo*F2 and *hm*F2; however, while for *fo*F2 this is exactly that of IRI, for *hm*F2 this was specifically implemented for IRTAM according to the work of Brunini et al. [53], and differs from any IRI *hm*F2 formulation. The Brunini et al. [53] procedure uses the URSI one to directly map the *hm*F2 values, that are those modeled by the Bilitza et al. [41] formulation using the *M*(3000)F2 values provided by the URSI mapping procedure. Practically, this process is nothing else than a simple *hm*F2 re-mapping, with the purpose to standardize the formalism between *fo*F2 and *hm*F2. By virtue of this, the IRTAM *hm*F2 modeled values cannot be considered as a direct updating of the IRI *hm*F2 values. Moreover, in this paper, IRI *hm*F2 values are those output by the Shubin et al. [44] default option, whose formulation differs from the IRTAM one and is also based on different datasets. On the contrary, the IRTAM *fo*F2 modeled values are a direct update of the IRI *fo*F2 values, because both models rely on the same formalism.

IRTAM is a four-dimensional data assimilation method because it does not merely assimilate the current data from the GIRO Digisonde network at the time of assimilation but takes into account also the prior 24 h history of the ionosphere at the assimilated station. This approach has the advantage of increasing the robustness of IRTAM by smoothing out data jitter, outliers, and low-confidence values through the diurnal Fourier analysis [17,18]. Moreover, this approach is really suited to IRI because the time and spatial variations are strictly connected through the 988 coefficients [14]. The IRTAM data-assimilation algorithm is named NECTAR (Non-linear Error Compensation Technique for Associative Restoration, [18]). As a first step, for each assimilated station, NECTAR considers the 24 h values recorded by the ionosonde prior to the assimilation time and calculates the differences between observed and modeled (by IRI) values; the procedure is the same for both *fo*F2 and *hm*F2. The 24 h time series of detrended values is used to describe the corresponding diurnal trend trough a 6th order Fourier analysis (the same as that of IRI) plus a linear term, for a total of 14 time coefficients. The second step concerns the spatial interpolation of the detrended diurnal coefficients at locations different from assimilating sites. This task is accomplished by NECTAR through a recurrent Hopfield neural network

optimizer [54], which is used as a spatial interpolator to smoothly spread the information from the assimilating sites to the entire global grid. Once the global grid of detrended diurnal coefficients is obtained, the third step consists in retrieving the 76 spatial coefficients to be used as correction terms in the Jones and Gallet spatial harmonic expansion. The output of the NECTAR method is then 14 (diurnal) × 76 (spatial) correction coefficients to be added to the original ones. The original coefficients are intended to represent the climate behavior of the ionosphere; as a consequence, the correction coefficients calculated by IRTAM describe the departures from the climatological behavior of the ionosphere.

3. Measured and Modeled Data Used for Validation

3.1. Observations from Ground-Based Ionosondes

Observations of the F2-layer peak ionospheric characteristics, foF2 and hmF2, measured by 40 ground-based ionosonde stations, located at different latitudes in both hemispheres, during the last two solar cycles (from 1 January 2000 to 31 December 2019), are considered as reference. The selected ionosonde stations are listed in Table 1 with the corresponding geographic coordinates, quasi-dipole (QD) magnetic latitude [55], modip [36], and time coverage of the dataset. Figure 1 illustrates the spatial distribution of the 40 ionosonde stations.

Table 1. The ionosonde stations considered for validation along with the corresponding geographic coordinates, QD magnetic latitude, modip, and time coverage of the dataset.

Number	Ionosonde (Country)	Geographic Latitude [°]	Geographic Longitude [°]	Quasi-Dipole Latitude [°]	Modip [°]	Years Dataset
1	Anyang (South Korea)	37.4° N	126.9° E	31.0° N	46.3° N	2000–2009
	I-Cheon (South Korea)	37.1° N	127.5° E	30.7° N	46.0° N	2010–2019
2	Ascension Island (UK)	7.9° S	14.4° W	19.1° S	34.3° S	2000–2019
3	Athens (Greece)	38.0° N	23.5° E	31.9° N	46.7° N	2002–2019
4	Boa Vista (Cape Verde)	2.8° N	60.7° W	10.6° N	19.5° N	2013–2019
5	Boulder (USA)	40.0° N	105.3° W	38.1° N	53.0° N	2004–2019
6	Cachoeira Paulista (Brazil)	22.7° S	45.0° W	18.8° S	32.2° S	2000–2019
7	Chilton (U.K.)	51.5° N	0.6° W	47.7° N	55.6° N	2000–2019
8	Dourbes (Belgium)	50.1° N	4.6° E	45.8° N	54.8° N	2001–2019
9	Dyess AFB (USA)	32.4° N	99.8° W	41.5° N	49.2° N	2000–2009
10	Eielson (USA)	64.7° N	147.1° W	64.9° N	64.1° N	2012–2019
11	El Arenosillo (Spain)	37.1° N	6.7° W	30.5° N	44.8° N	2000–2019
12	Fortaleza (Brazil)	3.9° S	38.4° W	7.1° S	13.6° S	2001–2019
13	Gakona (USA)	62.4° N	145.0° W	63.0° N	62.8° N	2000–2019
14	Goose Bay (Canada)	53.3° N	60.3° W	60.2° N	58.9° N	2000–2010
15	Grahamstown (South Africa)	33.3° S	26.5° E	41.9° S	50.3° S	2000–2019
16	Guam (USA)	13.6° N	144.9° E	6.1° N	12.3° N	2012–2019
17	Hermanus (South Africa)	34.4° S	19.2° E	42.5° S	51.0° S	2008–2019
18	Jicamarca (Peru)	12.0° S	76.8° W	0.2° N	0.4° N	2000–2019
19	Juliusruh (Germany)	54.6° N	13.4° E	50.7° N	57.7° N	2001–2019
20	King Salmon (USA)	58.4° N	156.4° W	56.8° N	59.8° N	2000–2012
21	Kwajalein (Marshall Islands)	9.0° N	167.2° E	4.1° N	7.6° N	2004–2013
22	Learmonth (Australia)	21.8° S	114.1° E	29.6° S	44.9° S	2001–2019
23	Louisvale (South Africa)	28.5° S	21.2° E	38.3° S	49.7° S	2000–2019
24	Millstone Hill (USA)	42.6° N	71.5° W	51.8° N	54.2° N	2000–2019

Table 1. *Cont.*

| Number	Ionosonde (Country)	Geographic Latitude [°]	Geographic Longitude [°]	Quasi-Dipole Latitude [°]	Modip [°]	Years Dataset
25	Moscow (Russia)	55.5° N	37.3° E	51.5° N	58.6° N	2008–2019
26	Nicosia (Cyprus)	35.0° N	33.2° E	29.2° N	44.6° N	2008–2019
27	Nord Greenland (Greenland)	81.4° N	17.5° W	81.0° N	75.2° N	2006–2013
28	Norilsk (Russia)	69.2° N	88.0° E	64.7° N	67.5° N	2002–2015
29	Point Arguello (USA)	34.8° N	120.5° W	40.2° N	48.8° N	2000–2019
30	Port Stanley (Falkland Islands)	51.6° S	57.9° W	38.7° S	48.3° S	2000–2019
31	Pruhonice (Czech Republic)	50.0° N	14.6° E	45.4° N	54.9° N	2004–2019
32	Ramey (Puerto Rico)	18.5° N	67.1° W	27.5° N	38.9° N	2000–2019
33	Rome (Italy)	41.8° N	12.5° E	35.9° N	49.3° N	2000–2019
34	Roquetes (Spain)	40.8° N	0.5° E	34.7° N	48.2° N	2000–2019
35	San Vito (Italy)	40.6° N	17.8° E	34.6° N	48.6° N	2000–2019
36	Sao Luis (Brazil)	2.6° S	44.2° W	2.9° S	5.0° S	2000–2019
37	Sondrestrom (Greenland)	67.0° N	50.9° W	72.2° N	65.8° N	2000–2012
38	Thule (Greenland)	77.5° N	69.2° W	84.5° N	72.7° N	2000–2014
39	Tromso (Norway)	69.6° N	19.2° E	66.5° N	66.6° N	2000–2018
40	Wallops Island (USA)	37.9° N	75.5° W	47.8° N	52.0° N	2000–2019
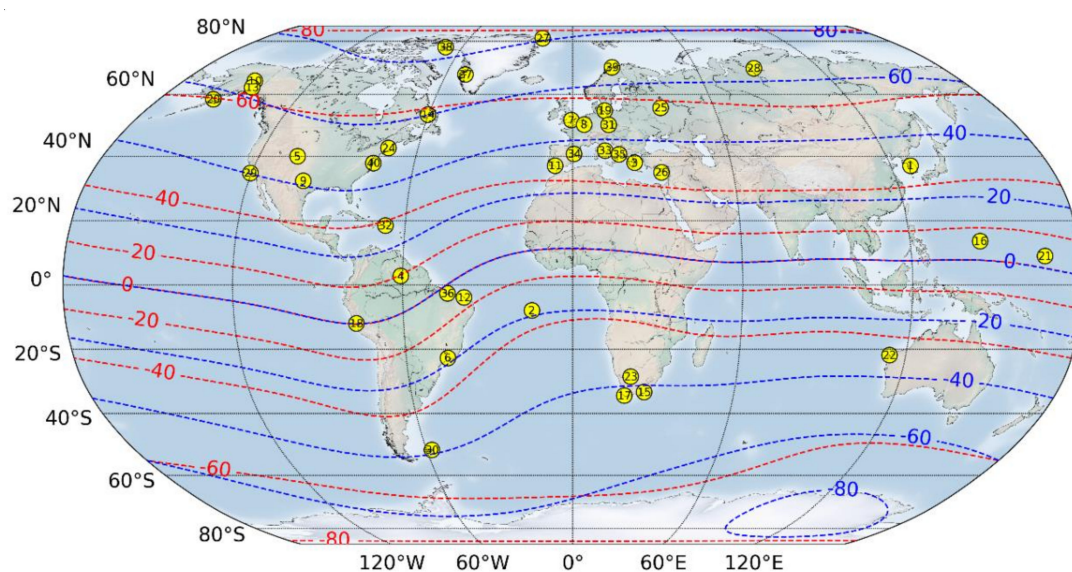
Figure 1. Global view of the ionosonde stations listed in Table 1. The yellow circles depict the ionosonde stations location with the corresponding identification number (Table 1, first column). Blue dashed curves depict the QD magnetic parallels, while red ones depict the modip parallels.

Ionosonde measured *fo*F2 and *hm*F2 values were downloaded from the Digital Ionogram DataBASE [56]. For each station, ionograms were recorded by DPS Digisondes [57], and autoscaled by the Automatic Real-Time Ionogram Scaler with True height analysis (ARTIST) software [58]. ARTIST flags the reliability of autoscaled parameters through the Confidence Score (C-Score) parameter [59] ranging from 0 to 100. For this study, only the most reliable values were considered, namely those with C-Score $\geq$ 75. *fo*F2 and *hm*F2 time series have a fifteen-minute time sampling (at minutes 0, 15, 30, and 45 of each Universal Time (UT) hour) according to the sounding repetition rate of most of the ionosondes. In

Figure 2a, the percentage of available *fo*F2 and *hm*F2 values per year, also considering the applied filtering based on the C-score value, is graphically represented for each ionosonde station. By virtue of the fifteen-minute time sampling, 100% of available values per year correspond to 35,040 (35,136 for a leap year).
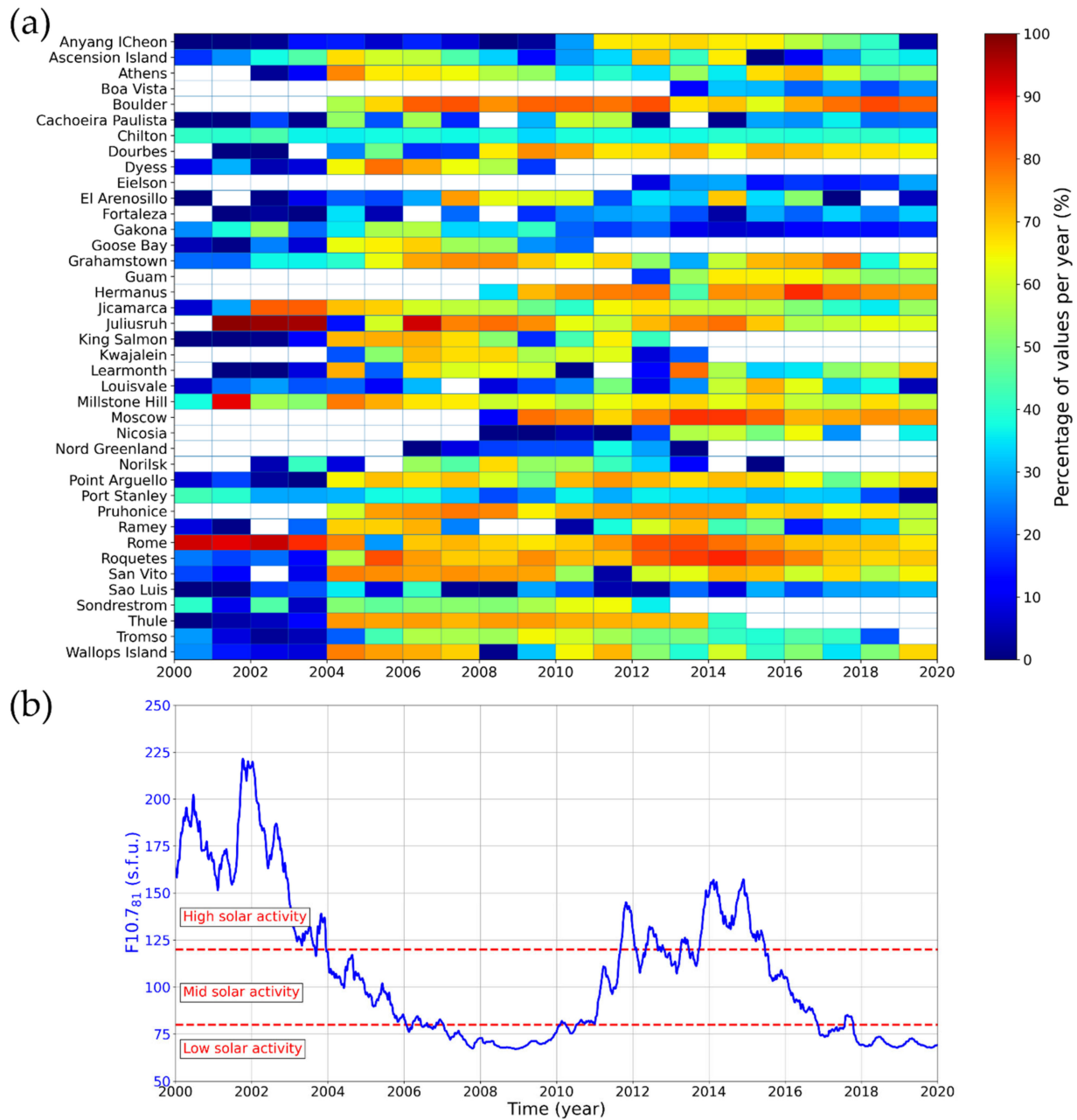
Figure 2. (**a**) Percentage of available ionosonde observations per year and ionosonde station. (**b**) F10.7$_{81}$ time series from 2000 to 2019. Red horizontal dashed lines identify the three ranges chosen for the solar activity description.

The considered dataset includes the last two solar cycles as depicted by the F10.7$_{81}$ solar index, i.e., the 81-day running mean of the F10.7 solar index, in Figure 2b. F10.7 is the solar radio flux at 10.7 cm wavelength (2800 MHz) [60] and represents one of the most used solar activity proxies for ionospheric modeling. In particular, its 81-day running mean was used in order to smooth its short-time variability. F10.7 daily data were downloaded at the NASA's OMNIWeb Data Explorer website (https://omniweb.gsfc.nasa.gov/form/dx1.html, accessed on 30 July 2021). Three solar activity ranges are considered here to evaluate the solar activity dependence:

- Low solar activity (LSA): $F10.7_{81} < 80$ s.f.u. (solar flux units, 1 s.f.u. $= 10^{-22}$ Wm^{-2} Hz^{-1}).
- Mid solar activity (MSA): 80 s.f.u. $\leq F10.7_{81} < 120$ s.f.u.
- High solar activity (HSA): $F10.7_{81} \geq 120$ s.f.u.

The $F10.7_{81}$ solar index thresholds were selected by considering the solar activity level experienced in the last two solar cycles, and on the basis of the available datasets of ionosonde and COSMIC derived observations.

3.2. Observations from Space-Based COSMIC/FORMOSAT-3 Satellites

Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC/FORMOSAT-3, hereafter COSMIC) was a six low-Earth-orbit microsatellites constellation launched on 15 April 2006. The mission was a collaborative project between the National Space Organization in Taiwan and the University Corporation for Atmospheric Research in the United States. COSMIC satellites were deployed into a circular orbit, with 72° of inclination, at about 800 km of altitude and a separation angle of 30° in longitude between neighboring satellites [61]. Each satellite carries a GPS RO receiver capable of measuring the phase delay of radio waves from GPS satellites as they are occulted by the Earth's atmosphere, thus providing an accurate determination of the ionospheric vertical electron density profile up to the COSMIC satellite altitude. COSMIC RO data were downloaded from the COSMIC Data Analysis and Archive Center (CDAAC, https://data.cosmic.ucar.edu/gnss-ro/cosmic1/, accessed on 30 July 2021).

Specifically, COSMIC retrieved *fo*F2 and *hm*F2 ionospheric characteristics from 22 April 2006 to 31 December 2018 were considered in this study. For this time range, a total of 3,626,729 COSMIC electron density profiles were available, and the most reliable ones were selected by applying the filtering procedure described in the "Methods" section of Pignalberi et al. [62]. As a consequence of such a filtering procedure, the COSMIC dataset that was used in this study for validation purposes was reduced to 1,791,602 profiles.

3.3. IRI and IRTAM Models Runs

In this study, we focus on the F2-layer peak characteristics modeled by the IRI-2016, which is the current version of the IRI model. Specifically, IRI was run for the same time periods and locations covered by the ionosondes and COSMIC datasets by using the IRI Fortran code available at the IRI website (http://irimodel.org/, accessed on 30 July 2021). In this way, a one-to-one comparison between measured and modeled *fo*F2 and *hm*F2 values is guaranteed.

IRI *fo*F2 values were modeled through the URSI coefficients [31], while *hm*F2 values were modeled through the Shubin et al. [44] option. Moreover, the IRI model was run with the storm option [39,40] "ON" to take into account the magnetic activity disturbance effect on modeled *fo*F2 and *hm*F2 values.

Likewise, the IRTAM procedure was also run for the same time periods and locations covered by the ionosondes and COSMIC datasets. The IRTAM runs were made by using the IRTAM Fortran package available on the Global Assimilative Model of Bottomside Ionosphere Timeline (GAMBIT) Consortium website (http://giro.uml.edu/GAMBIT/, accessed on 30 July 2021), with the application of the GAMBIT coefficients (https://ulcar.uml.edu/GAMBIT/GambitCoefficients/, accessed on 30 July 2021) for modeling both *fo*F2 and *hm*F2. GAMBIT coefficients are the URSI ones corrected by IRTAM through the NECTAR method (see Section 2.2) on the basis of assimilated data. Since GAMBIT coefficients are available from the beginning of 2000, the validation analysis here presented is restricted to years 2000–2019. It has to be pointed out that the number of stations assimilated by IRTAM has changed over the years due to the number of available GIRO Digisondes able to stream information in real-time to GAMBIT. This number has changed from about 10 at the beginning of 2000 to about 60 at the end of 2019. As a consequence, we should expect different performance of IRTAM for different years; specifically, an increase of its performance as time goes by.

4. Methodologies of Analysis

4.1. Statistical Metrics Adopted in the Validation Process

To evaluate the IRI and IRTAM models' performances, different statistical metrics were calculated between measured and modeled *fo*F2 and *hm*F2 values. Specifically, the mean of the residuals (*Res. Mean*) between measured and modeled (by IRI and IRTAM) values, the root mean square error (*RMSE*), the normalized root mean square error (*NRMSE*), and the Pearson correlation coefficient (*R*) are the considered statistical metrics:

$$Res.Mean \ = \frac{\sum_{i=1}^{N}(xx\text{F2}_{\text{measured},i} - xx\text{F2}_{\text{modeled},i})}{N}, \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(xx\text{F2}_{\text{measured},i} - xx\text{F2}_{\text{modeled},i})^2}{N}}, \tag{2}$$

$$NRMSE \ [\%] = \frac{\text{RMSE}(xx\text{F2}_{\text{measured}}, xx\text{F2}_{\text{modeled}})}{\overline{xx\text{F2}_{\text{measured}}}} \cdot 100, \tag{3}$$

$$R = \frac{\text{cov}(xx\text{F2}_{\text{measured}}, xx\text{F2}_{\text{modeled}})}{\sigma_{xx\text{F2}_{\text{measured}}}\sigma_{xx\text{F2}_{\text{modeled}}}} \in [-1, 1], \tag{4}$$

where $xx = fo$ when considering *fo*F2 and $xx = hm$ when considering *hm*F2. The subscript "modeled" refers to values obtained through IRI or IRTAM, while the subscript "measured" refers to values observed by either ionosondes or COSMIC satellites; the index i runs on the N values of the time series. $\overline{xx\text{F2}_{\text{measured}}}$ is the arithmetic mean of $xx\text{F2}_{\text{measured}}$, cov is the covariance between modeled and measured values, and σ the corresponding variance. The unit of measurement of both *Res. Mean* and *RMSE* can be either MHz or km, for *fo*F2 and *hm*F2, respectively. It is important to point out that *Res. Mean* highlights the accuracy of modeled values, while *RMSE* and *NRMSE* are indicators of the corresponding precision.

Concerning the validation with the COSMIC dataset also the root mean square percentage error (*RMSPE*) was considered:

$$RMSPE \ [\%] = \sqrt{\frac{\sum_{i=1}^{N}\left(\frac{xx\text{F2}_{\text{measured},i} - xx\text{F2}_{\text{modeled},i}}{xx\text{F2}_{\text{measured},i}} \cdot 100\right)^2}{N}}. \tag{5}$$

Differently from *NRMSE*, when calculating *RMSPE*, the normalization of the residuals between measured and modeled values is performed point-by-point before calculating the root mean square value. The use of *RMSPE* is more suitable than that of *NRMSE* for datasets with very different absolute values for different diurnal, seasonal, or solar activity conditions. In principle, this is true for both datasets, the COSMIC and the ionosonde ones. Nevertheless, the COSMIC dataset is global, that is it varies with the geographical coordinates, while the location of ionosondes is fixed. This is why *RMSPE* is suited especially for the COSMIC dataset.

In order to evaluate the deviations between models and observations, the residuals deviation ratio parameter R_{cw} [18] was also computed:

$$R_{\text{cw}} = \left|\frac{\varepsilon_{\text{c}}}{\varepsilon_{\text{w}}}\right| = \left|\frac{xx\text{F2}_{\text{IRI}} - xx\text{F2}_{\text{measured}}}{xx\text{F2}_{\text{IRTAM}} - xx\text{F2}_{\text{measured}}}\right|, \tag{6}$$

where ε_{c} represents the residuals between "climatological" (i.e., IRI modeled) and measured values, and ε_{w} are the corresponding residuals between "weather" (i.e., IRTAM modeled) and measured values.

The statistical parameter R_{cw} allows us to directly compare IRI and IRTAM models and easily evaluate the corresponding performances. In fact, by definition, $R_{\text{cw}} = 1$ when both models are on par in terms of their accuracy, while $R_{\text{cw}} > 1$ when IRTAM is more accurate than IRI and $R_{\text{cw}} < 1$ when IRI is more accurate than IRTAM.

4.2. Data Binning

Both *fo*F2 and *hm*F2 exhibit variations at very different temporal and spatial scales, that are also a consequence of the varying solar and magnetic conditions. To highlight the main climatological diurnal, seasonal, spatial, solar and magnetic activity variations embedded in the data, both modeled and measured values were binned according to different schemes that are here reported in detail.

The diurnal variation was studied in a twofold way:

1. Data were binned as a function of the LT with fifteen minute-wide bins (96 bins in total).
2. Data were collected in three separate diurnal sectors as a function of the solar zenith angle (SZA):
   - Daytime: SZA $\leq 80°$.
   - Solar terminator: $80° <$ SZA $< 100°$.
   - Nighttime: SZA $\geq 100°$.

The seasonal variation has been studied in a threefold way:

1. (only for ionosondes) Data were binned as a function of the month of the year (12 bins in total).
2. (only for COSMIC) Data were binned as a function of the day of the year (doy) in bins five-day wide (73 bins in total).
3. Data were collected in four bins representative of the four seasons. Specifically, bins centered at equinoxes and solstices:
   - March Equinox: $35 \leq$ doy ≤ 125.
   - June Solstice: $126 \leq$ doy ≤ 217.
   - September Equinox: $218 \leq$ doy ≤ 309.
   - December Solstice: doy ≤ 34 OR doy ≥ 310.

The solar activity variation was studied by selecting the three levels of solar activity (low, mid, and high) defined in Section 3.1 on the basis of the $F10.7_{81}$ solar index, and by binning data as a function of $F10.7_{81}$ in bins two s.f.u.-wide.

The magnetic activity variation was studied by selecting three levels of magnetic disturbance on the basis of the global a_p index [63]:

- Quiet magnetic activity: $a_p < 20$ nT.
- Moderate magnetic activity: 20 nT $\leq a_p < 100$ nT.
- Disturbed magnetic activity: $a_p \geq 100$ nT.

a_p values were downloaded from NASA's Space Physics Data Facility of the Goddard Space Flight Center (https://spdf.gsfc.nasa.gov/pub/data/omni/high_res_omni/, accessed on 30 July 2021).

For what concerns the spatial variation, in the case of the ionosonde dataset this is directly related to the station's geographic position. Conversely, for the COSMIC dataset, this was investigated using geographic coordinates, the magnetic QD latitude and modip. Specifically:

1. Data were binned as a function of the geographic coordinates in bins that are 2.5°-wide in latitude and 5°-wide in longitude.
2. Data were binned as a function of the modip latitude in bins that are 2.5°-wide.
3. Data were binned according to three ranges of modip values:
   - Low modip: $-30° <$ modip $< 30°$.
   - Mid modip: $-60° \leq$ modip $\leq -30°$ AND $30° \leq$ modip $\leq 60°$.
   - High modip: modip $< -60°$ AND modip $> 60°$.

4.3. Graphical Representation of the Statistical Results

Several graphical representations, depending also on the dataset used for validation, were adopted to highlight at best the different climatological variations exhibited by the data.

The following graphical representations were used for both the ionosonde and the COSMIC datasets:

- Statistical distribution of residuals between measured and modeled (by IRI and IRTAM) *fo*F2 and *hm*F2 values.
- Density plots between measured and modeled *fo*F2 and *hm*F2 values, with the corresponding best linear fit line.

The following graphical representations were used only for the ionosondes dataset:

- Grids of *Res. Mean*, *RMSE*, and *NRMSE* values between measured and modeled *fo*F2 and *hm*F2 values, as a function of the LT and the month of the year, for the three different levels of solar activity defined in Section 3.1.

Differently, in the validation analysis based on the COSMIC observations, IRI and IRTAM *RMSPE* binned values, and corresponding differences (IRTAM-IRI), have been plotted as a function of:

- Geographic latitude vs. geographic longitude.
- Modip vs. LT hour.
- Modip vs. doy.
- Modip vs. F10.7$_{81}$.

Finally, statistical probability distributions of the residuals deviation ratio parameter R_{cw} were calculated for the entire ionosonde and COSMIC *fo*F2 and *hm*F2 datasets.

5. Validation Results for *fo*F2 Based on Ground-Based Ionosonde Observations

5.1. Statistics on the Full Dataset

The full *fo*F2 ionosonde dataset includes 10,133,987 observations obtained cumulating measurements recorded from 2000 to 2019 at the 40 ionospheric stations under consideration. This dataset has been used to calculate the two-dimensional density plots of IRI-*fo*F2 vs. ionosonde-*fo*F2 and IRTAM-*fo*F2 vs. ionosonde-*fo*F2, along with the histograms of corresponding residuals, as shown in Figure 3. The comparison between IRI and IRTAM results highlights that overall IRTAM turns out to be more accurate and precise than IRI. In fact, the statistical distribution of IRTAM residuals around the zero value is less widespread than the IRI one, and *Res. Mean*$_{IRTAM}$ = 0.005 MHz is smaller than *Res. Mean*$_{IRI}$ = 0.051 MHz. Furthermore, the comparisons *RMSE*$_{IRTAM}$ = 0.674 MHz vs. *RMSE*$_{IRI}$ = 1.002 MHz and *NRMSE*$_{IRTAM}$ = 12.270% vs. *NRMSE*$_{IRI}$ = 18.258% show also that the IRI *fo*F2 modeling precision is significantly improved by IRTAM. This point is also reinforced by the IRTAM density plot which appears less scattered than the IRI one (R_{IRTAM} = 0.959 vs. R_{IRI} = 0.908).

Similar conclusions can be made also when IRI and IRTAM performances are evaluated on both the cumulative *fo*F2 time series binned according to Section 4.2, and to the three different level of solar activity defined in Section 3.1, as highlighted by Table 2, and the *fo*F2 time series available for each considered ionospheric station, as highlighted by Table 3. In fact, Table 2 shows values of *Res. Mean*, *RMSE* and *NRMSE* that are always lower for IRTAM, which means that IRTAM models *fo*F2 better than IRI during daytime, nighttime, solar terminator hours, in all seasons, and during different solar and magnetic activity conditions. The improvement made by IRTAM is quite consistent and, more importantly, does not show any particular diurnal, seasonal, or solar activity dependence (as also evidenced by [29]), i.e., IRTAM outperforms IRI in all the conditions. The assimilation of real-time data is particularly efficient during magnetically disturbed conditions where the ionosphere exhibits very steep spatial gradients of the electron density distribution and very fast time variations.
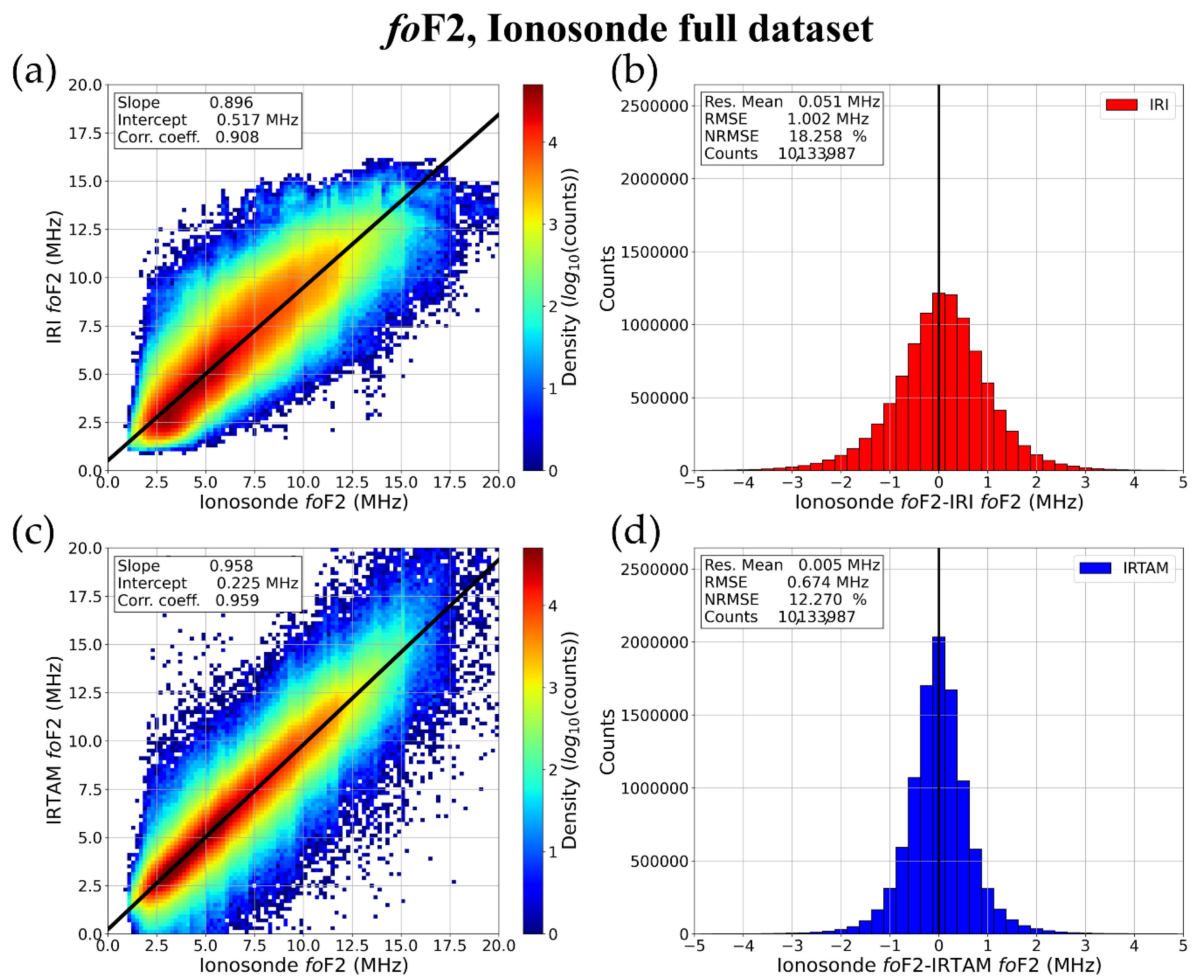
Figure 3. Density plots between measured (*x*-axis) and modeled (*y*-axis) *fo*F2 values, for (**a**) IRI and (**c**) IRTAM. Measured values are those from the entire ionosonde dataset. The black solid lines represent the best linear fit. Corresponding slope and intercept values, along with the correlation coefficient, are reported in the upper left box of each plot. Statistical distributions of the resultant residuals for (**b**) IRI and (**d**) IRTAM. Corresponding *Res. Mean*, *RMSE*, and *NRMSE* values, along with the total number of counts on which the statistics is based, are reported in the upper left box of each plot.

Table 2. *Res. Mean, RMSE, NRMSE,* and *R* values calculated for the IRI and IRTAM models on the basis of the *fo*F2 ground-based ionosonde observations, for the bins defined in Section 4.2, according to the three levels of solar activity defined in Section 3.1, and for the full dataset (bottom row). The number of counts on which the statistics were calculated is reported in the rightmost column.

Ionosonde Stations Dataset	Model	Res. Mean [MHz]	RMSE [MHz]	NRMSE [%]	R	Counts
Daytime	IRI	−0.048	1.026	14.861	0.887	4,091,455
	IRTAM	−0.015	0.676	9.789	0.952	
Nighttime	IRI	0.125	1.003	23.790	0.842	4,198,762
	IRTAM	0.013	0.674	15.983	0.930	
Solar terminator	IRI	0.106	0.946	18.011	0.885	1,843,770
	IRTAM	0.032	0.668	12.721	0.944	
March Equinox	IRI	0.179	1.040	18.129	0.914	2,710,599
	IRTAM	0.015	0.685	11.934	0.963	

Table 2. *Cont.*

Ionosonde Stations Dataset	Model	Res. Mean [MHz]	RMSE [MHz]	NRMSE [%]	R	Counts
June Solstice	IRI	−0.033	0.919	17.038	0.880	2,255,233
	IRTAM	0.003	0.646	11.967	0.942	
September Equinox	IRI	−0.002	1.008	18.216	0.906	2,644,252
	IRTAM	0.000	0.665	12.008	0.959	
December Solstice	IRI	0.046	1.205	19.493	0.918	2,523,903
	IRTAM	0.002	0.695	13.210	0.962	
LSA	IRI	0.130	0.853	19.015	0.879	3,782,579
	IRTAM	−0.028	0.590	13.151	0.936	
MSA	IRI	0.046	0.993	18.117	0.890	3,801,672
	IRTAM	0.007	0.665	12.135	0.953	
HSA	IRI	−0.057	1.202	17.180	0.896	2,549,736
	IRTAM	0.050	0.792	11.327	0.957	
Quiet magnetic activity	IRI	0.056	0.981	17.957	0.910	9,147,468
	IRTAM	0.006	0.655	11.995	0.960	
Moderate magnetic activity	IRI	0.015	1.167	20.322	0.890	959,046
	IRTAM	−0.007	0.814	14.171	0.949	
Disturbed magnetic activity	IRI	−0.065	1.682	27.107	0.826	27,473
	IRTAM	−0.015	1.174	18.932	0.921	
Full dataset	IRI	0.051	1.002	18.258	0.908	10,133,987
	IRTAM	0.005	0.674	12.270	0.959	

Table 3. Same as Table 2, but for each of the *fo*F2 datasets relative to the 40 ionosonde stations listed in Table 1.

Ionosonde (Country)	Model	Res. Mean [MHz]	RMSE [MHz]	NRMSE [%]	R	Counts
Anyang and I-Cheon (South Korea)	IRI	−0.164	0.891	14.310	0.926	201,558
	IRTAM	−0.104	0.656	10.514	0.961	
Ascension Island (UK)	IRI	−0.217	1.683	21.457	0.870	264,916
	IRTAM	0.289	1.352	17.240	0.933	
Athens (Greece)	IRI	−0.770	1.503	30.079	0.845	313,227
	IRTAM	−0.179	0.661	13.235	0.942	
Boa Vista (Cape Verde)	IRI	0.364	1.634	18.484	0.868	60,295
	IRTAM	0.326	1.031	11.655	0.952	
Boulder (USA)	IRI	0.137	0.794	17.257	0.900	420,504
	IRTAM	−0.013	0.564	12.255	0.949	
Cachoeira Paulista (Brazil)	IRI	−0.193	1.416	21.012	0.895	158,881
	IRTAM	0.101	0.944	14.010	0.955	
Chilton (U.K.)	IRI	0.111	0.851	16.945	0.914	269,815
	IRTAM	0.037	0.641	12.765	0.951	
Dourbes (Belgium)	IRI	0.272	0.781	15.842	0.912	328,740
	IRTAM	0.097	0.488	9.907	0.965	
Dyess AFB (USA)	IRI	0.197	0.988	18.951	0.895	142,345
	IRTAM	−0.004	0.729	13.979	0.936	
Eielson (USA)	IRI	0.200	0.741	14.212	0.878	54,960
	IRTAM	0.040	0.436	8.366	0.956	
El Arenosillo (Spain)	IRI	−0.149	0.987	17.429	0.908	206,362
	IRTAM	−0.167	0.653	11.531	0.963	

Table 3. *Cont.*

Ionosonde (Country)	Model	Res. Mean [MHz]	RMSE [MHz]	NRMSE [%]	R	Counts
Fortaleza (Brazil)	IRI	−0.271	1.453	19.349	0.844	115,951
	IRTAM	−0.151	0.965	12.847	0.937	
Gakona (USA)	IRI	−0.033	0.789	16.965	0.886	188,948
	IRTAM	−0.128	0.568	12.198	0.945	
Goose Bay (Canada)	IRI	0.225	0.784	16.981	0.882	138,063
	IRTAM	−0.138	0.619	13.394	0.926	
Grahamstown (South Africa)	IRI	0.263	0.923	17.213	0.937	389,105
	IRTAM	0.010	0.541	10.087	0.976	
Guam (USA)	IRI	−0.200	1.232	15.569	0.880	150,696
	IRTAM	0.036	0.923	11.661	0.938	
Hermanus (South Africa)	IRI	0.380	0.938	17.344	0.930	297,534
	IRTAM	0.053	0.496	9.173	0.977	
Jicamarca (Peru)	IRI	−0.245	1.258	17.802	0.866	379,962
	IRTAM	0.068	0.828	11.719	0.945	
Juliusruh (Germany)	IRI	0.140	0.822	16.308	0.923	473,533
	IRTAM	0.042	0.557	11.049	0.965	
King Salmon (USA)	IRI	0.131	0.771	18.513	0.859	180,833
	IRTAM	0.065	0.583	14.008	0.924	
Kwajalein (Marshall Islands)	IRI	−0.338	1.346	19.961	0.857	160,301
	IRTAM	0.032	0.953	14.135	0.933	
Learmonth (Australia)	IRI	0.170	0.951	16.597	0.904	262,583
	IRTAM	−0.124	0.629	10.985	0.962	
Louisvale (South Africa)	IRI	0.346	1.050	17.327	0.936	191,705
	IRTAM	0.065	0.565	9.329	0.979	
Millstone Hill (USA)	IRI	0.093	0.879	16.280	0.918	441,837
	IRTAM	0.011	0.641	11.861	0.957	
Moscow (Russia)	IRI	0.163	0.768	15.995	0.920	301,375
	IRTAM	0.052	0.481	10.020	0.968	
Nicosia (Cyprus)	IRI	−0.296	0.938	15.395	0.928	110,598
	IRTAM	−0.116	0.645	10.591	0.965	
Nord Greenland (Greenland)	IRI	0.066	0.771	20.584	0.745	47,039
	IRTAM	−0.191	0.653	17.428	0.839	
Norilsk (Russia)	IRI	0.115	0.737	17.516	0.836	145,454
	IRTAM	−0.051	0.357	8.481	0.963	
Point Arguello (USA)	IRI	0.158	0.907	16.580	0.913	377,422
	IRTAM	0.017	0.573	10.481	0.965	
Port Stanley (Falkland Islands)	IRI	−0.512	1.247	23.169	0.887	212,867
	IRTAM	−0.040	0.745	13.854	0.953	
Pruhonice (Czech Republic)	IRI	0.274	0.775	15.527	0.913	396,878
	IRTAM	0.092	0.498	9.971	0.963	
Ramey (Puerto Rico)	IRI	0.019	1.239	19.490	0.870	231,544
	IRTAM	0.003	0.786	12.366	0.950	
Rome (Italy)	IRI	−0.100	0.900	15.741	0.924	520,519
	IRTAM	−0.089	0.711	12.450	0.953	
Roquetes (Spain)	IRI	−0.013	0.829	15.010	0.923	443,895
	IRTAM	−0.069	0.601	10.889	0.961	

Table 3. *Cont.*

Ionosonde (Country)	Model	Res. Mean [MHz]	RMSE [MHz]	NRMSE [%]	R	Counts
San Vito (Italy)	IRI	0.097	0.808	15.058	0.920	372,074
	IRTAM	0.020	0.602	11.211	0.956	
Sao Luis (Brazil)	IRI	−0.157	1.403	19.002	0.843	122,965
	IRTAM	−0.045	0.992	13.433	0.928	
Sondrestrom (Greenland)	IRI	0.540	0.977	20.860	0.812	200,111
	IRTAM	0.138	0.650	13.864	0.885	
Thule (Greenland)	IRI	0.610	0.967	22.536	0.767	278,374
	IRTAM	0.135	0.561	13.072	0.886	
Tromso (Norway)	IRI	−0.009	0.764	18.261	0.845	259,000
	IRTAM	−0.134	0.529	12.659	0.932	
Wallops Island (USA)	IRI	0.211	0.870	16.679	0.912	321,218
	IRTAM	0.068	0.628	12.051	0.953	

Table 3 instead shows the statistical parameters (1–4) calculated for the *fo*F2 time series recorded in each of the 40 ionosonde stations listed in Table 1. This table highlights how the IRI *fo*F2 modeling accuracy and precision are improved by IRTAM in practically every ionosonde station regardless the location. Since the number and distribution of the stations assimilated by IRTAM has changed over the years, and critically depends on the real-time availability of data, it is difficult to distinguish between assimilated and non-assimilated stations. This is why we extended our analysis also to COSMIC data that are completely independent of ionosondes' ones and are not assimilated by IRTAM.

5.2. Diurnal, Seasonal, and Solar Activity Statistics for Different Zonal Sectors

In this section, we are going to show the IRI and IRTAM performances for 5 of the 40 ionospheric stations listed in Table 1. These five ionospheric stations are selected according to their modip value, and then considering the QD magnetic latitude (QD lat.) as representative of different sectors that are: (10° S, 10° N), equatorial latitudes; (10° N, 30° N) and (10° S, 30° S), low latitudes; (30° N, 55° N) and (30° S, 55° S), middle latitudes; (55° N, 75° N) and (55° S, 75° S), high latitudes; (75° N, 90° N) and (75° S, 90° S], polar cap latitudes. More importantly, within these sectors, the selection criterion takes into account those stations characterized by very long time series of data, in order to guarantee more statistically significant results. The diurnal and seasonal *fo*F2 modeling performance of both IRI and IRTAM are investigated in terms of the statistical parameters (1–3), which are shown in the form of grids as a function of LT (*x*-axis, with fifteen-minute-wide bins) and month of the year (*y*-axis), for low, mid, and high solar activity, according to the three ranges highlighted in Section 3.1.

Specifically, the five selected ionospheric stations are:

- Jicamarca (equatorial station, QD lat. = 0.2° N, Figures 4–6): it was selected because among the stations with low modip (modip = 0.4° N) it is the one that presents the longest dataset (379,962 measurements); it has also the peculiarity of laying right above the magnetic equator.
- Ascension Island (low-latitude station, QD lat. = 19.1° S, Figures 7–9): it was chosen because among the low-latitude stations characterized by a mid-low modip (modip = 34.3° S) it is the one that presents the longest dataset (264,916 measurements); it has also the particularity to lay over the southern equatorial anomaly crest.
- Rome (mid-latitude station, QD lat. = 35.9° N, Figures 10–12): it was chosen because among the mid-latitude stations characterized by a mid modip (modip = 49.3° N) it is the one that presents the longest dataset, which is also the longest dataset among the 40 considered stations (520,519 measurements).

- Sondrestrom (high latitude, QD lat. = 72.2° N, Figures 13–15): it was preferred among the stations with high modip (modip = 65.8° N) even if it presents a dataset that is a little bit shorter (200,111 measurements) than that of Tromso (259,000 measurements). The reason for this choice lies in the fact that Sondrestrom is significantly higher in latitude than Tromso (QD lat. = 66.5° N) and, thus, more representative of the auroral latitudes.
- Thule (polar cap, QD lat. = 84.5° N, Figures 16–18): it was chosen among the stations with very high modip (modip = 72.7° N) because it provides a dataset (278,374 measurements) substantially larger than that of Nord Greenland (47,039 measurements), and also because of its proximity to the north pole.

Figure 4. Grids of (**a**,**b**) *Res. Mean*, (**c**,**d**) *RMSE*, and (**e**,**f**) *NRMSE* values calculated between *fo*F2 values modeled by (**a**,**c**,**e**) IRI and (**b**,**d**,**f**) IRTAM, and those measured by the Jicamarca ionosonde for LSA. Monthly values (*y*-axis) are binned in fifteen-minute-wide bins in LT (*x*-axis). The white color highlights bins with less than ten counts for which values were not calculated because they were considered not statistically significant.

Jicamarca *fo*F2 statistics
Mid solar activity
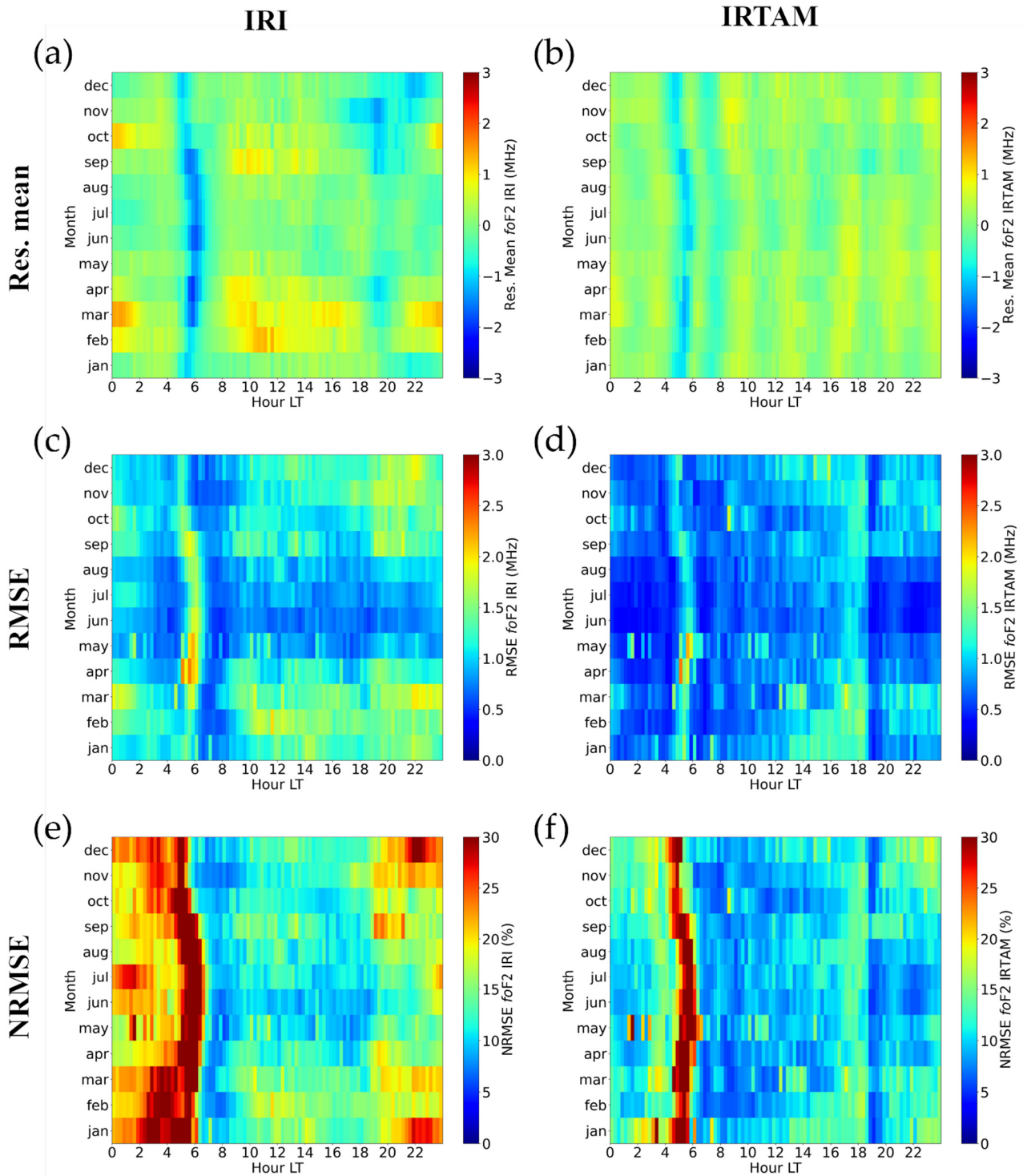
Figure 5. Same as Figure 4 but for MSA level.
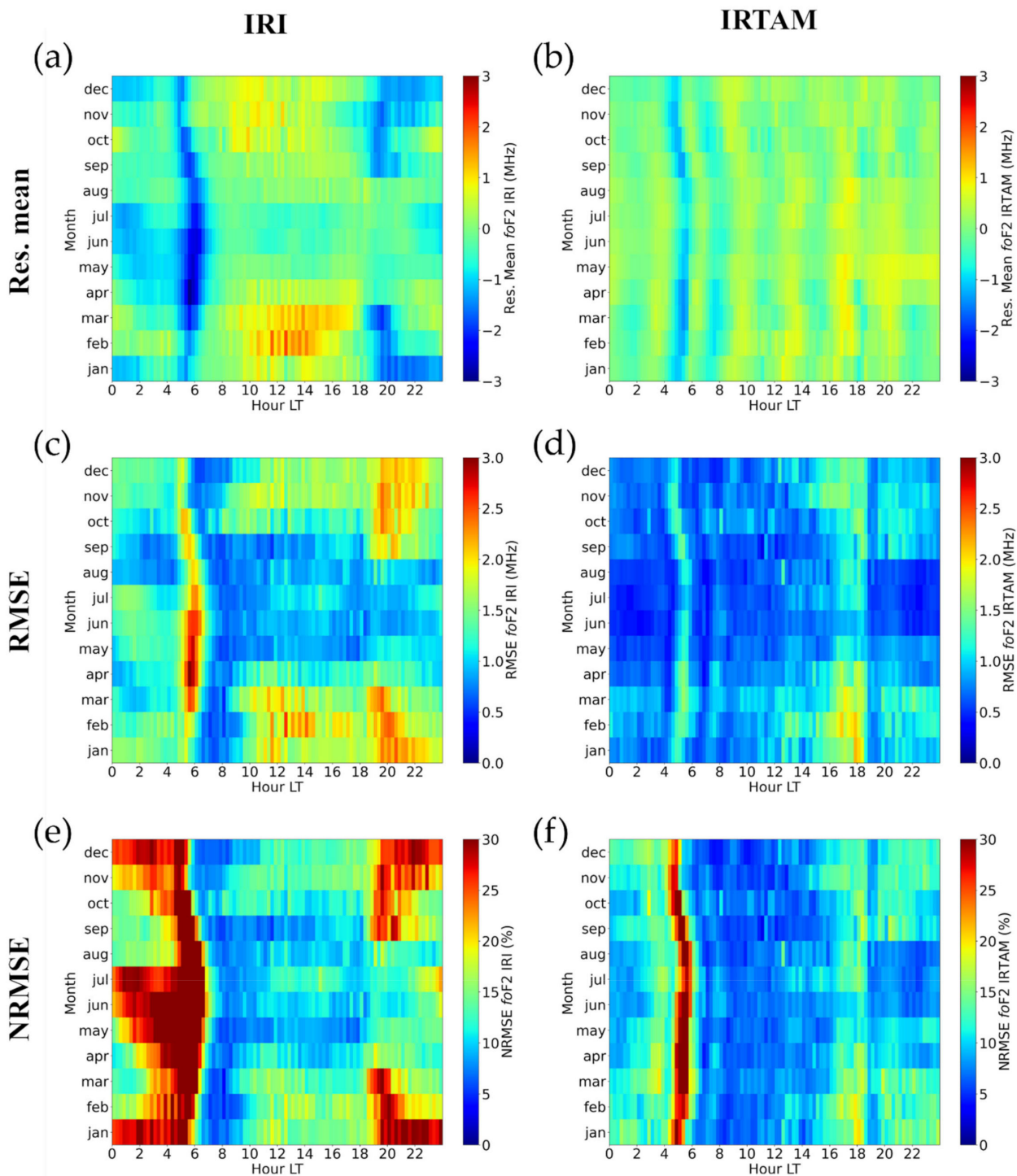
Figure 6. Same as Figure 4 but for HSA level.

Ascension Island *fo*F2 statistics
Low solar activity
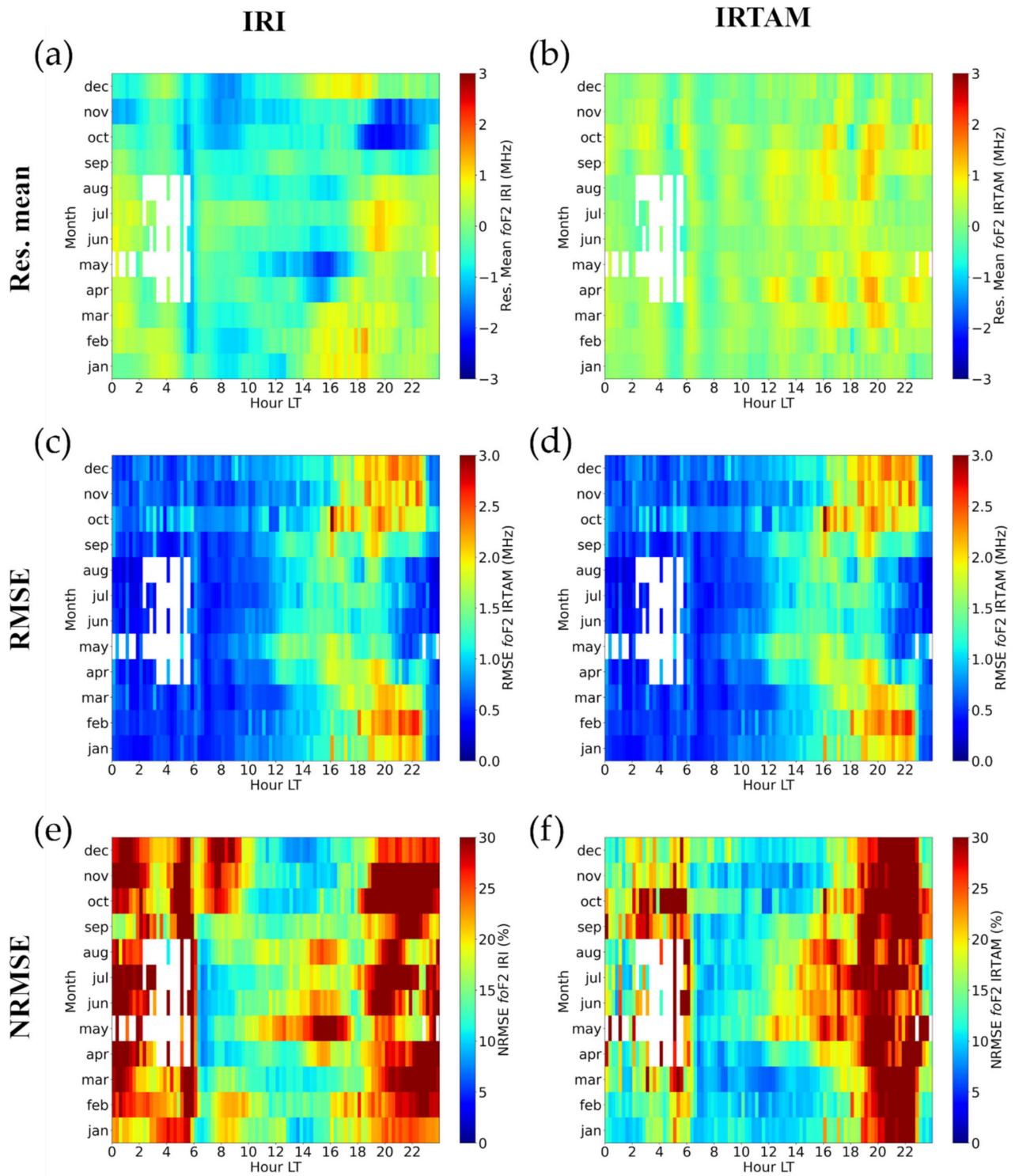
Figure 7. Same as Figure 4 but for Ascension Island.

Ascension Island *fo*F2 statistics
Mid solar activity

IRI   **IRTAM**
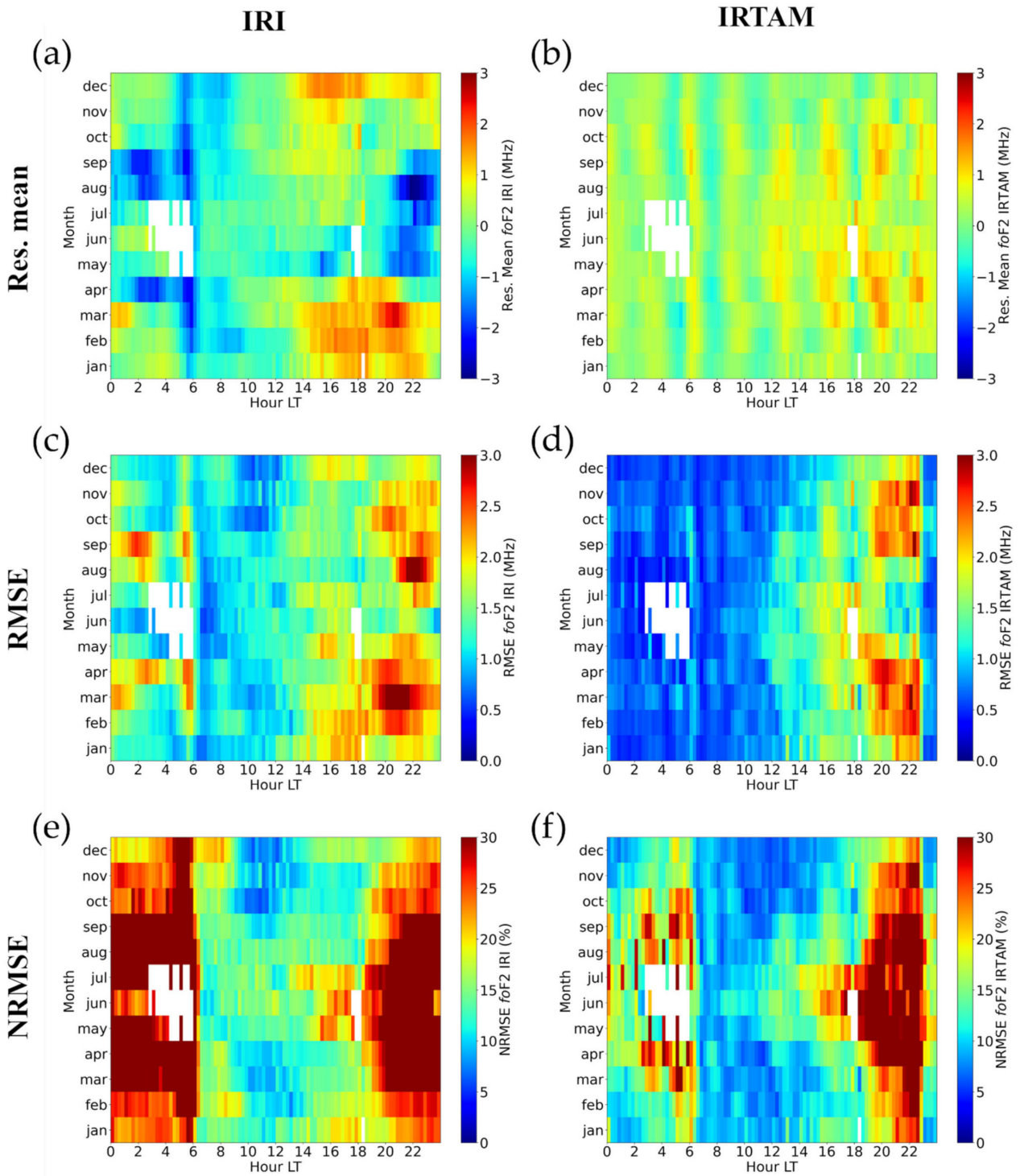
Figure 8. Same as Figure 7 but for MSA level.

Ascension Island *fo*F2 statistics
High solar activity
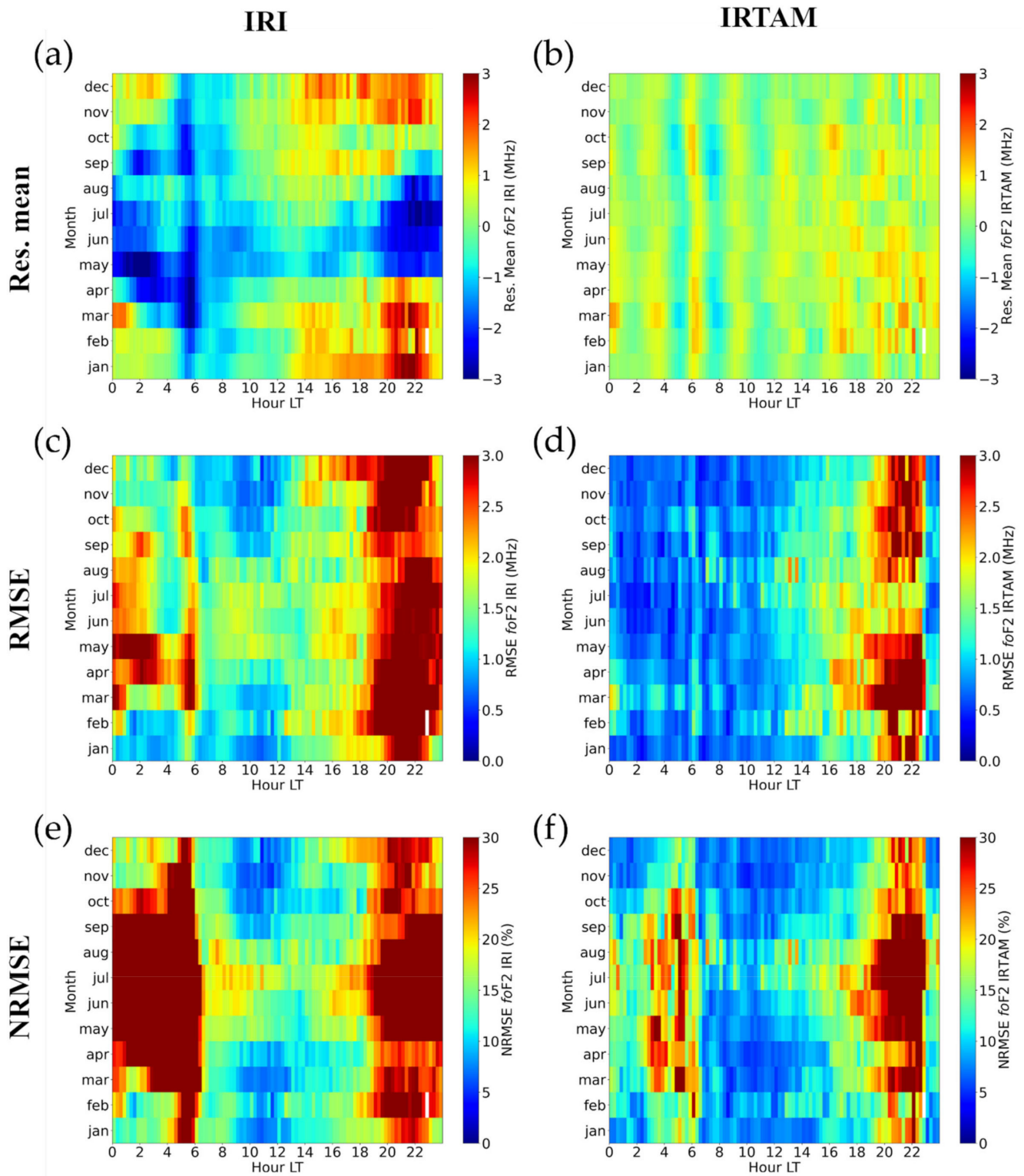
Figure 9. Same as Figure 7 but for HSA level.

Rome *fo*F2 statistics
Low solar activity
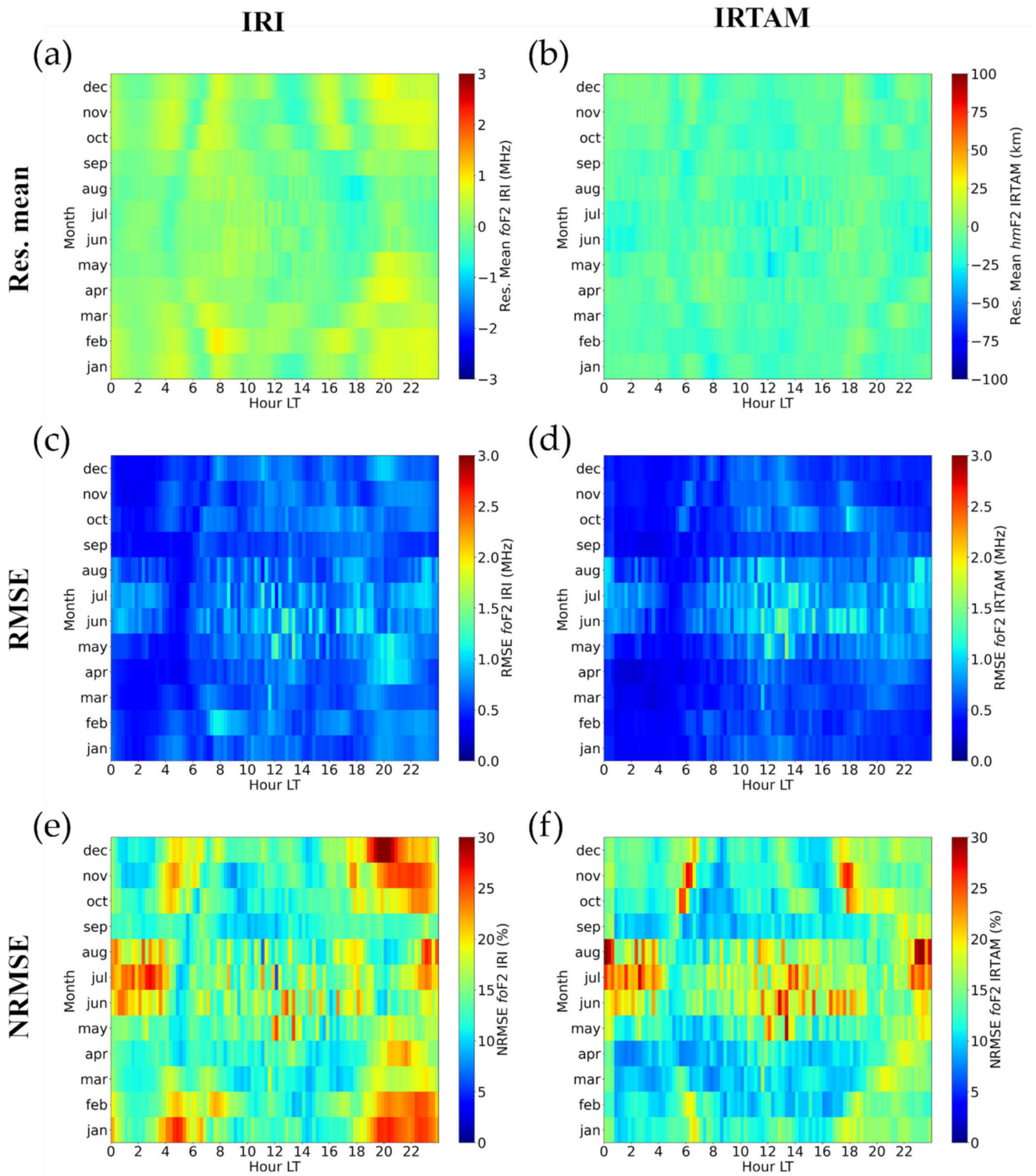
Figure 10. Same as Figure 4 but for Rome.

Rome *fo*F2 statistics
Mid solar activity

IRI **IRTAM**
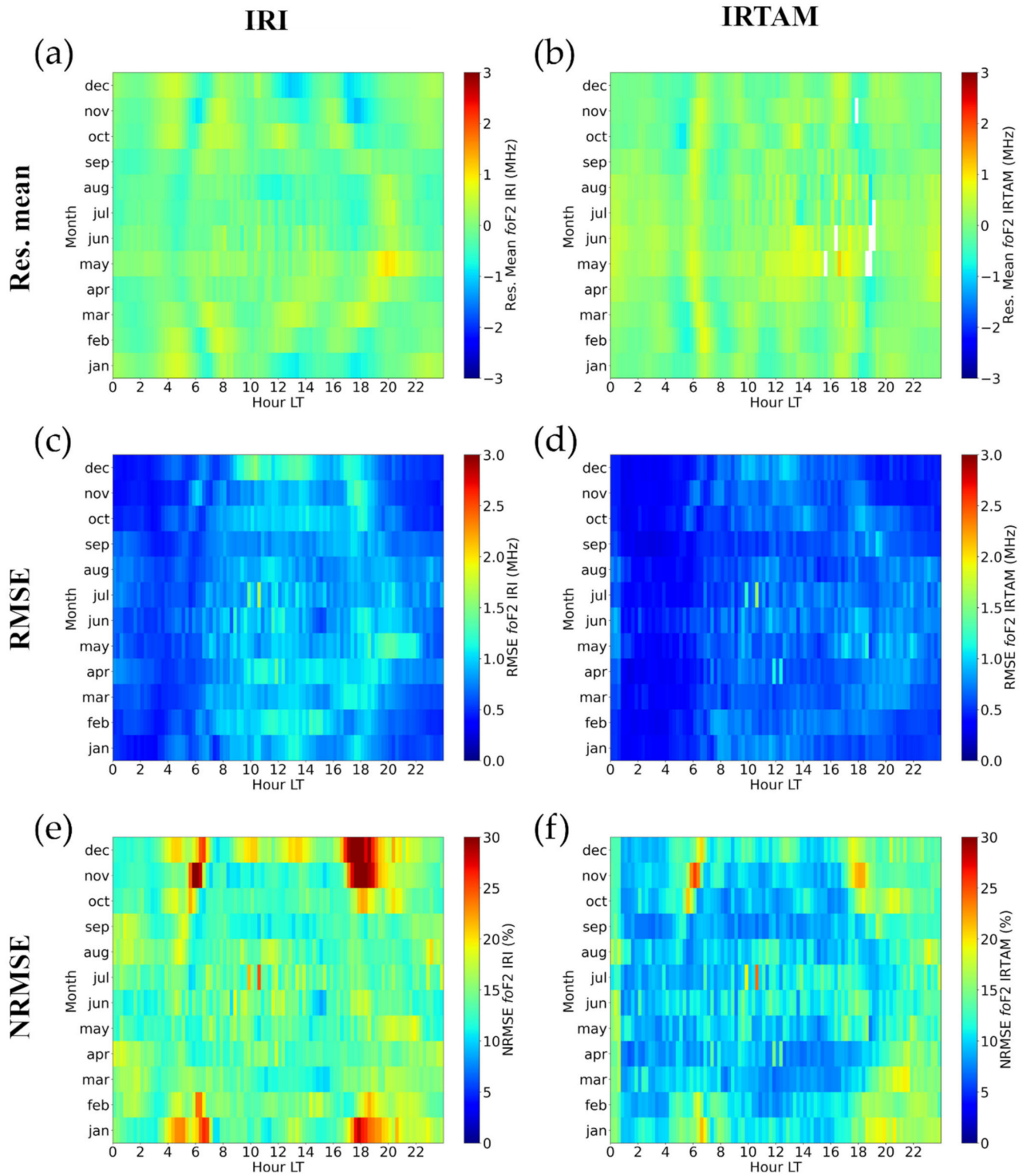
Figure 11. Same as Figure 10 but for MSA level.

Rome *fo*F2 statistics
High solar activity

IRI

IRTAM
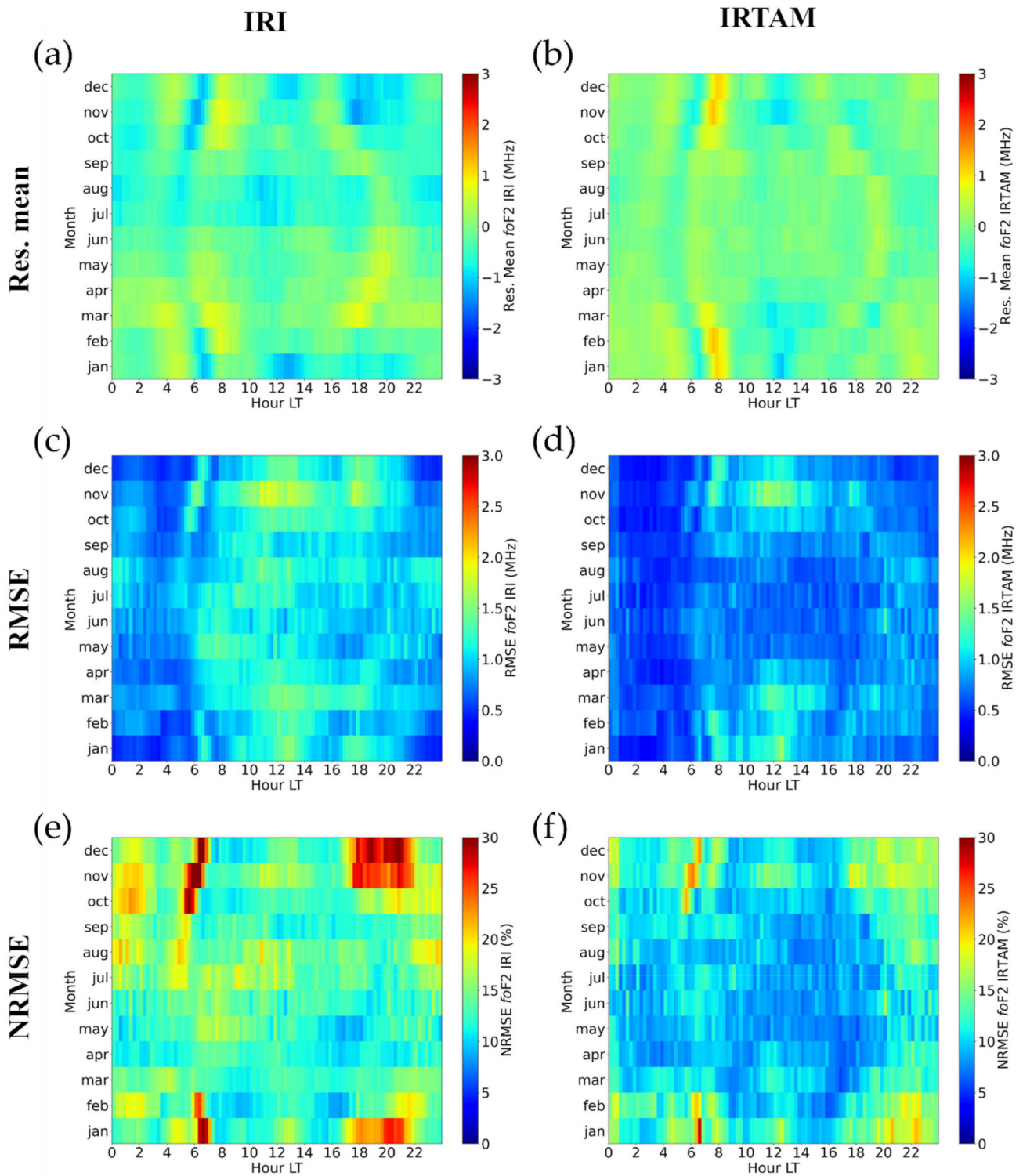
Figure 12. Same as Figure 10 but for HSA level.

Sondrestrom *fo*F2 statistics
Low solar activity
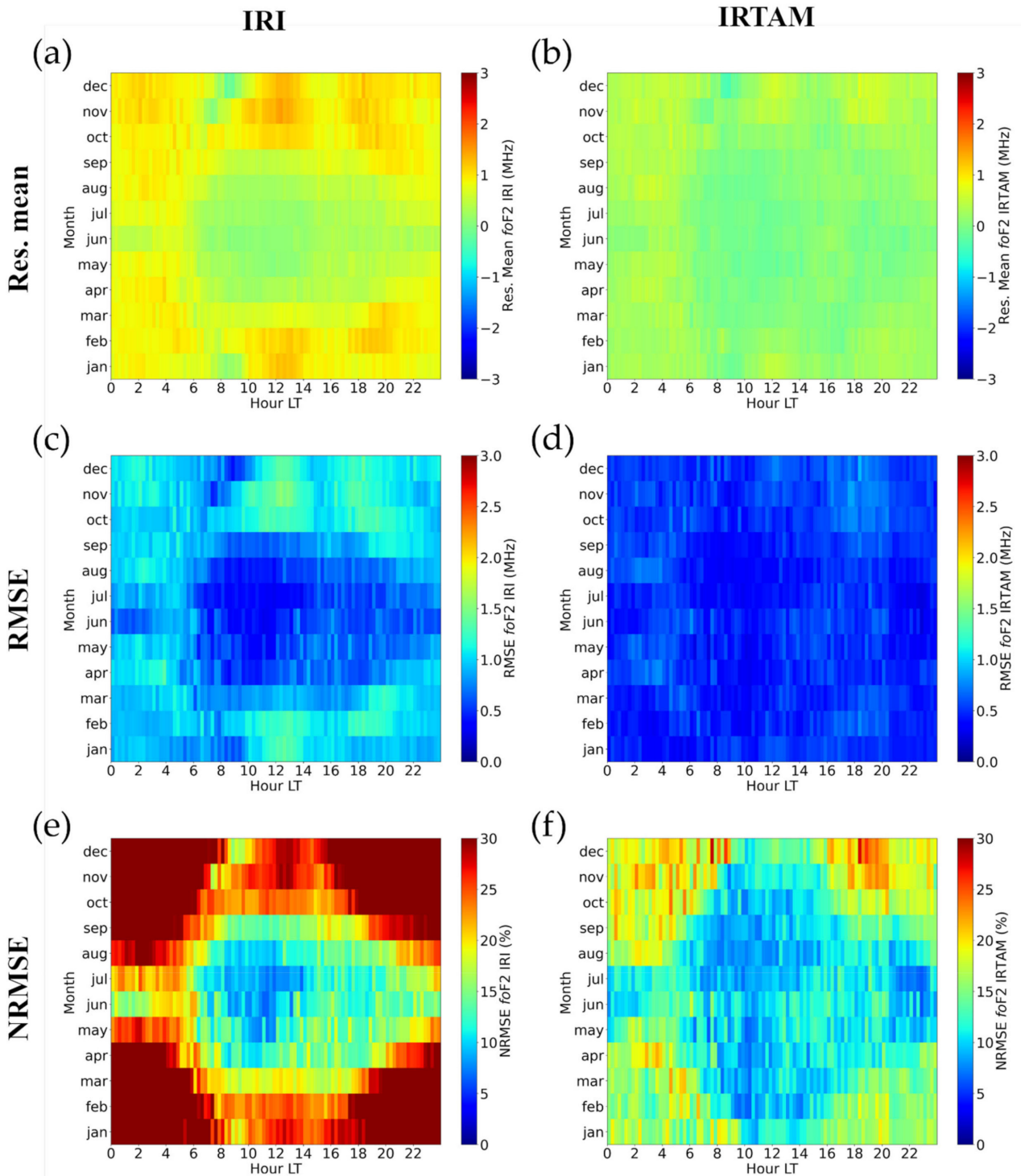
Figure 13. Same as Figure 4 but for Sondrestrom.

Sondrestrom *fo*F2 statistics
Mid solar activity
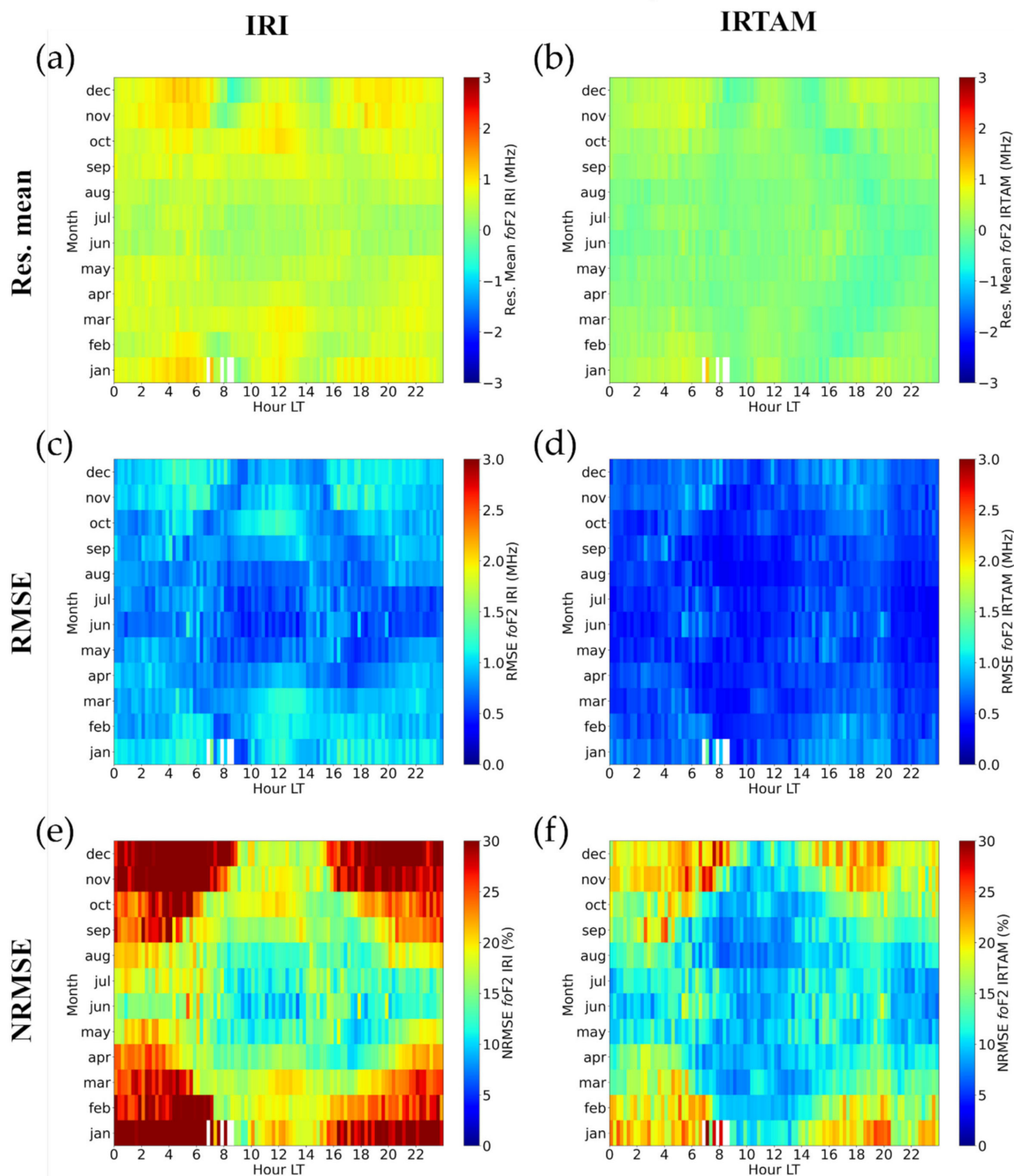
Figure 14. Same as Figure 13 but for MSA level.

Sondrestrom *fo*F2 statistics
High solar activity

IRI

IRTAM
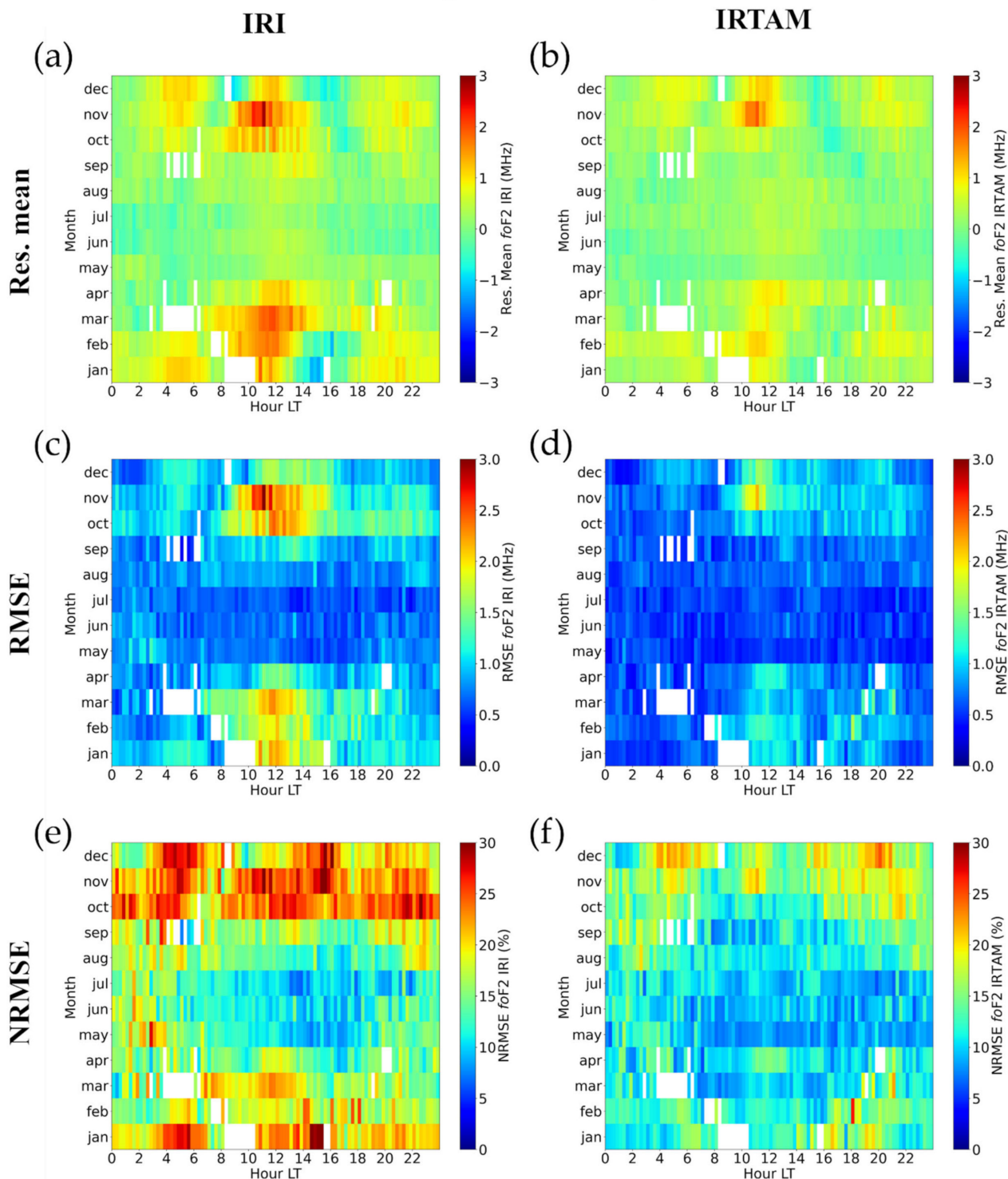
Figure 15. Same as Figure 13 but for HSA level.

Thule *fo*F2 statistics
Low solar activity

IRI
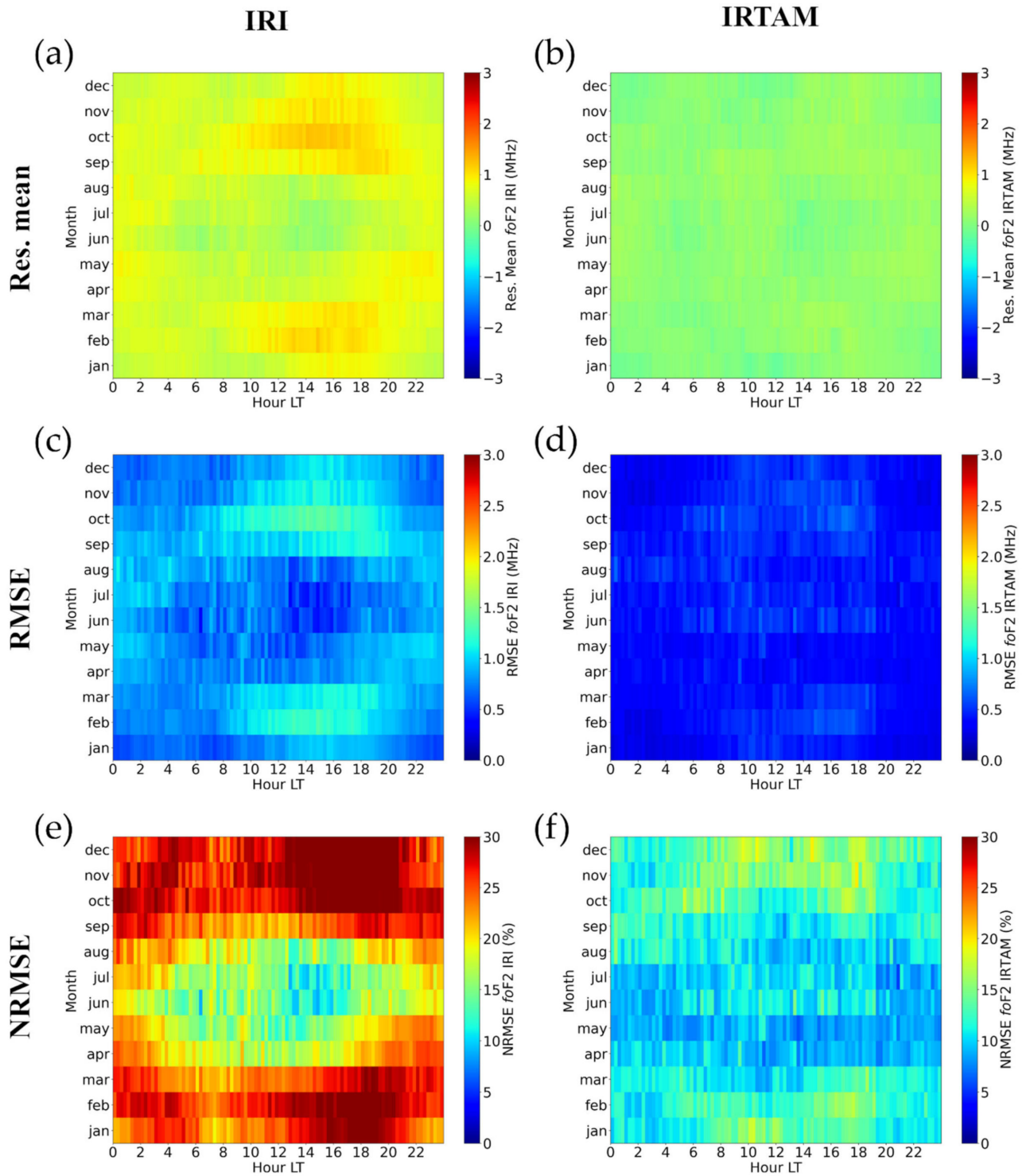
IRTAM

Figure 16. Same as Figure 4 but for Thule.
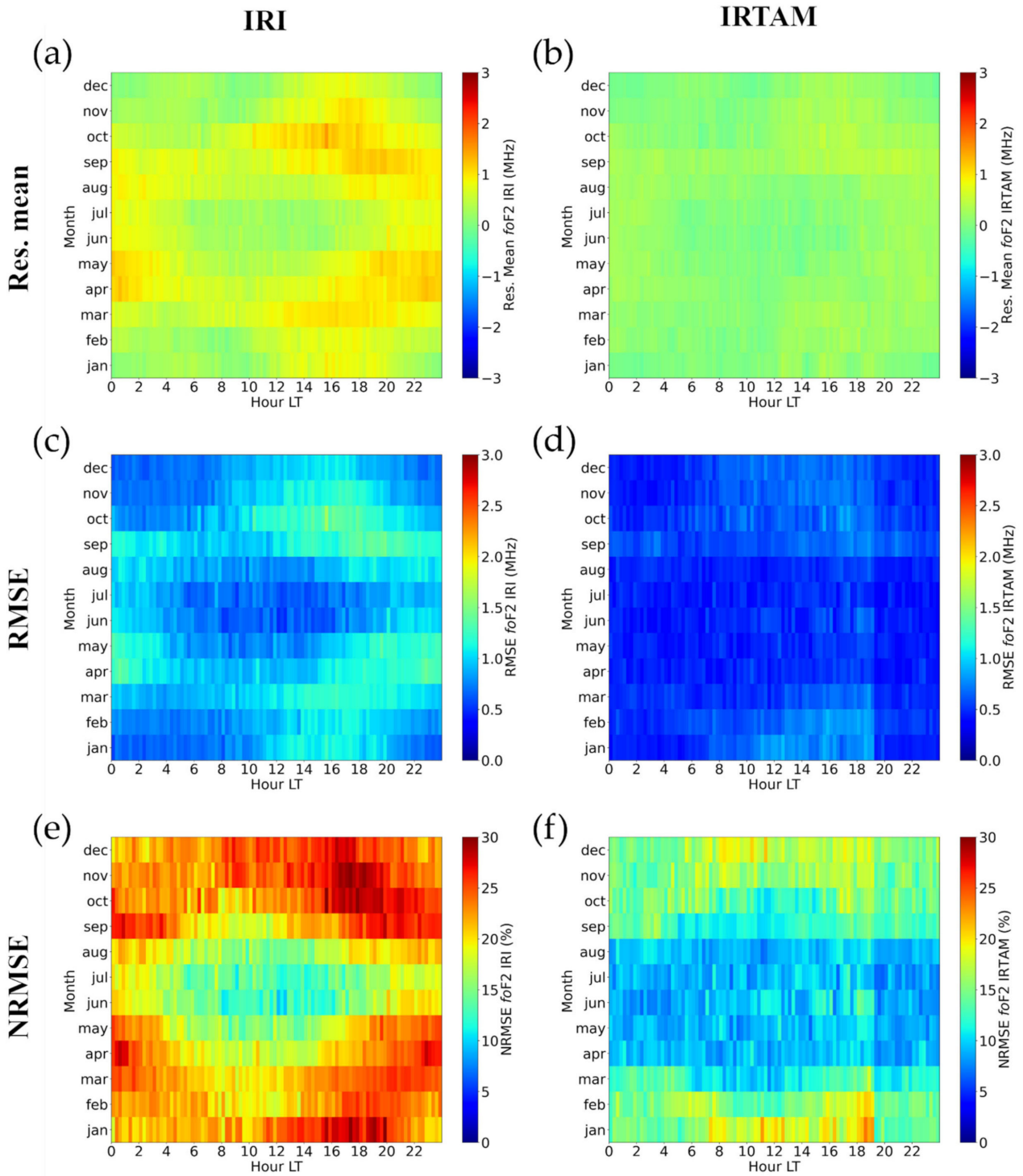
Figure 17. Same as Figure 16 but for MSA level.

Thule *fo*F2 statistics
High solar activity
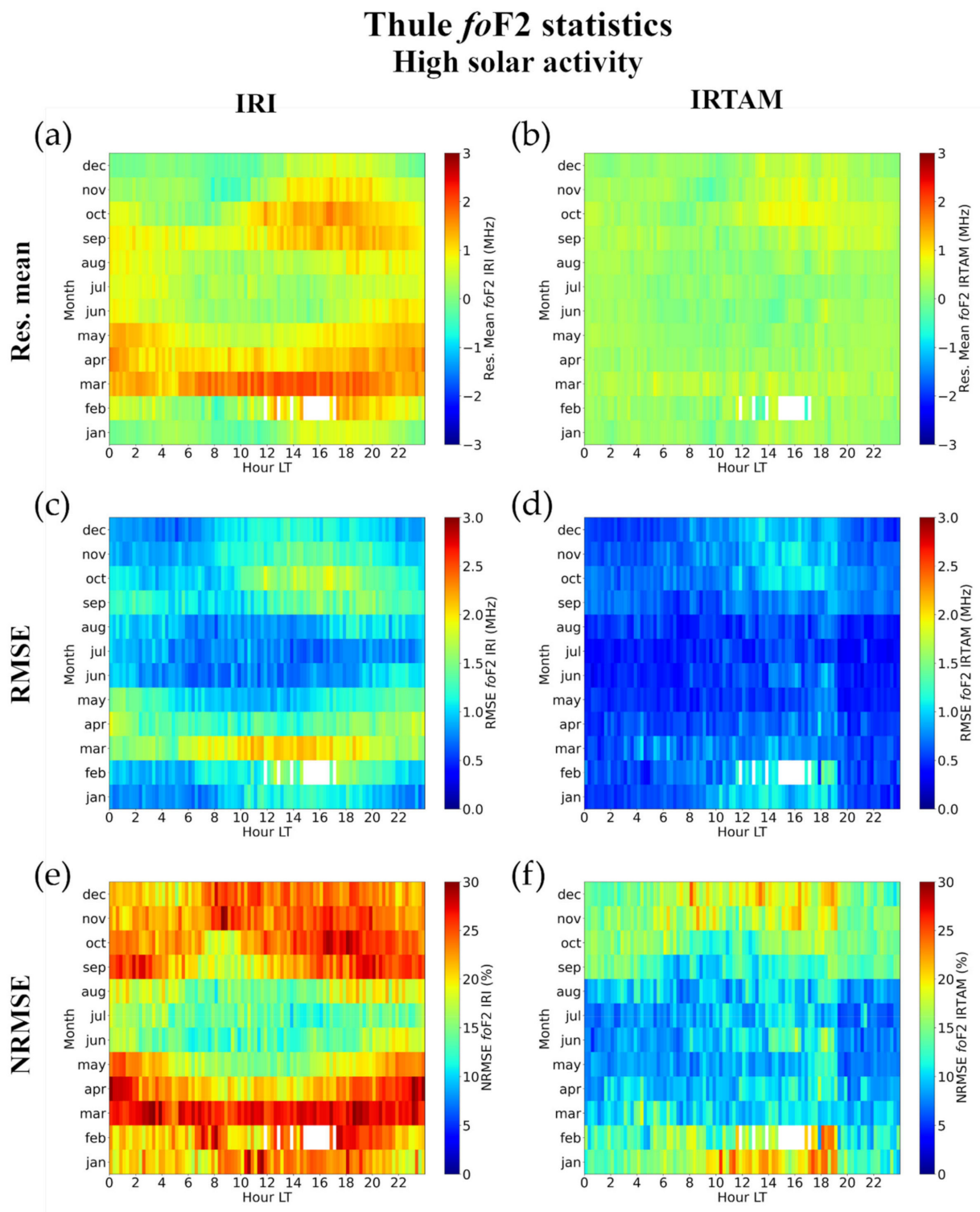
IRI   **IRTAM**

Figure 18. Same as Figure 16 but for HSA level.

In order to simplify the IRI-IRTAM comparison to the reader, all figures have been generated with the same scale. Since the main goal of the paper is the comparison between IRI and IRTAM performances, we will focus on the differences between the statistical metrics provided by IRI and IRTAM, without going into detail about the corresponding physical and theoretical explanation of the observed trends.

Figures 4–6 show that at Jicamarca for LSA, IRTAM models *fo*F2 more accurately than IRI, especially during pre-sunrise and post-sunset hours. An improvement in the *fo*F2

modeling accuracy made by IRTAM is still observed for MSA around sunrise hours for equinoctial and summer months and during daytime hours for some equinoctial months. For both LSA and MSA, at sunrise both models show high percentages of error (at least 25%). Nevertheless, during nighttime, a significant reduction in the percentage error made by IRTAM can be appreciated, which means a good improvement in the *fo*F2 modeling precision. For HSA, around sunrise in equinoctial and summer months, IRTAM performs better than IRI in terms of accuracy. The relatively high absolute errors (at least 2.0 MHz) made by IRI in equinoctial and summer months are strongly smoothed by IRTAM. Except for the hours around sunrise, where the percentage error remains pretty high (at least 25%), a very strong reduction in the percentage error made by IRTAM is observed in all seasons, which means a considerable improvement in the *fo*F2 modeling precision.

Figures 7–9 show that at Ascension Island overall, for LSA, during daytime hours and for all seasons, IRI models *fo*F2 less accurately than IRTAM. A very similar behavior of the two models is observed for the absolute error. However, during the central hours of the day, IRTAM provides a greater *fo*F2 modeling precision, being its percentage errors lower than those of IRI. Overall, for both MSA and HSA, the comparison between IRI and IRTAM *Res. Mean* grids suggests that IRI is less accurate than IRTAM. The IRTAM model reduces both the absolute and the percentage error, during the first hours of the night and daytime hours, which highlights an improvement in the IRI *fo*F2 modeling precision. Both IRI and IRTAM provide bad performance during post-sunset hours, for which extremely high percentage errors, up to 30% and more, are observed.

Figures 10–12 show that at Rome for LSA, a better performance in terms of *fo*F2 modeling accuracy is clearly observed in favour of IRTAM, while *RMSE* and *NRMSE* grids show very similar patterns, thus indicating that the precision with which IRI and IRTAM model *fo*F2 is comparable. Overall, for MSA, IRI and IRTAM performances are comparable in terms of accuracy. For HSA, a greater accuracy is provided by IRTAM, during pre-sunrise hours and daytime hours for summer months. For both MSA and HSA, the IRTAM model lowers the values of the absolute and percentage errors, which means that an improvement in the *fo*F2 modeling precision is achieved.

Figures 13–18 show that at Sondrestrom and Thule for the three considered solar activity levels a remarkable reduction in the *Res. Mean*, *RMSE*, and *NRMSE* values is observed when passing from IRI to IRTAM, which means that a significant improvement in the *fo*F2 modeling is achieved in terms of both accuracy and precision. The improvement brought by IRTAM is remarkable at equinoctial and winter months.

From a visual inspection of *fo*F2 grids obtained for the three solar activity ranges, some general considerations on how the IRI and IRTAM performances depend on solar activity can be made: no clear *Res. Mean* trend related to the solar activity for both IRI and IRTAM is found; IRI and IRTAM models show an increasing *RMSE* trend as the solar activity increases; a clear decreasing *NRMSE* trend is detected as the solar activity increases at Jicamarca, Ascension Island, and Rome for IRTAM, and at Thule and Sondrestrom for IRI. For the other cases it is difficult to establish a clear solar activity dependence. However, patterns characterized by higher *NMRSE* are almost always observed for lower solar activity.

6. Validation Results for *fo*F2 Based on Radio Occultation Observations

The full COSMIC dataset, comprising 1,791,602 selected electron density profiles measured from 2006 to 2018, was used to retrieve reliable *fo*F2 values for calculating the two-dimensional density plots of IRI-*fo*F2 vs. COSMIC-*fo*F2 and IRTAM-*fo*F2 vs. COSMIC-*fo*F2, along with the histograms of corresponding residuals, as shown in Figure 19. Differently from Figure 3, Figure 19 shows that IRTAM does not improve the *fo*F2 modeling accuracy of IRI. In fact, the IRTAM statistical distribution of residuals around the zero value, very similar to the IRI one, shows a value of *Res. Mean*$_{\text{IRTAM}}$ = 0.095 MHz that is a little bit greater than *Res. Mean*$_{\text{IRI}}$ = 0.067 MHz.
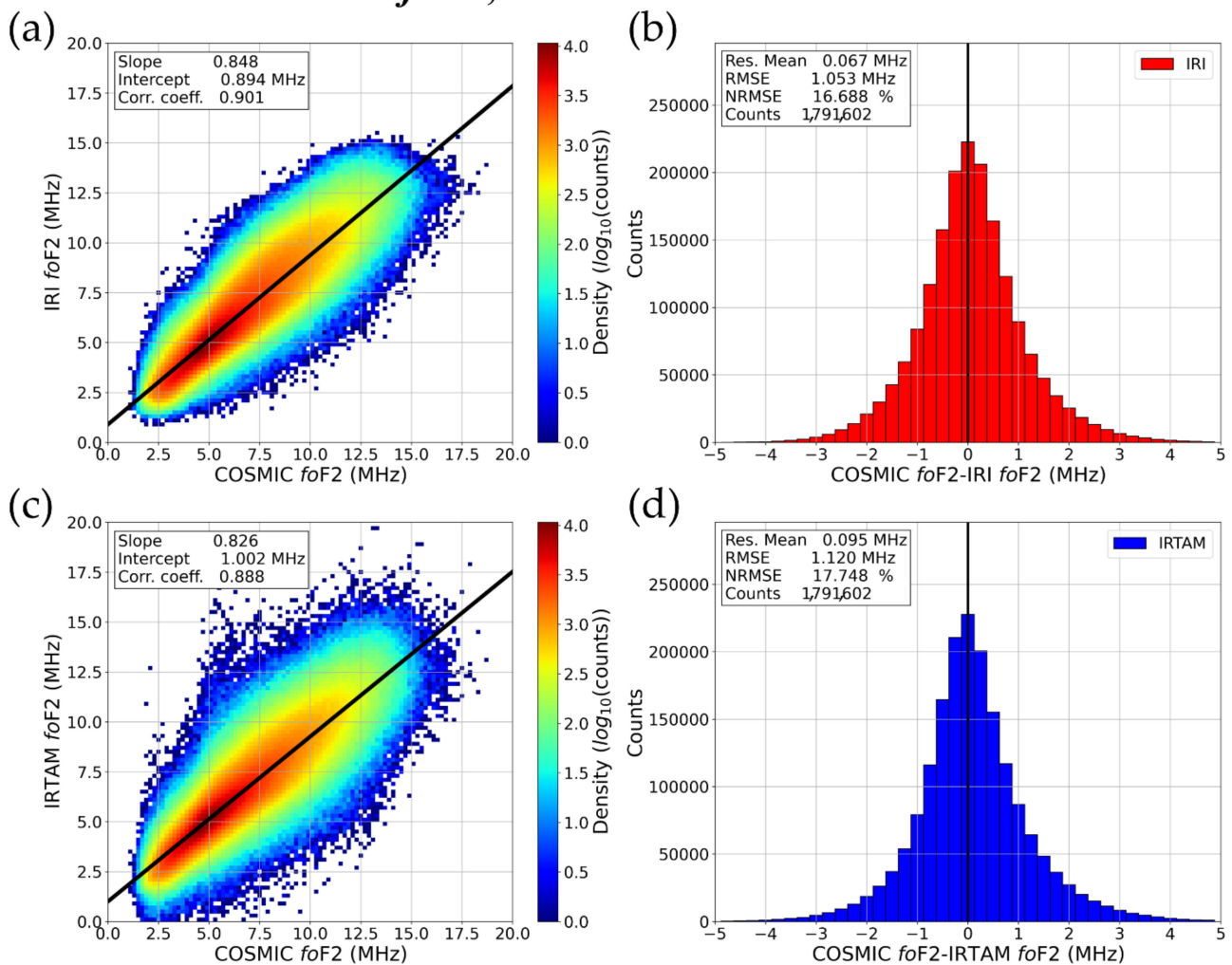
Figure 19. Density plots between measured (*x*-axis) and modeled (*y*-axis) *fo*F2 values, for (**a**) IRI and (**c**) IRTAM. Measured values are those from the entire COSMIC dataset. The black solid lines represent the best linear fit. Corresponding slope and intercept values, along with the correlation coefficient, are reported in the upper left box of each plot. Statistical distributions of the residuals between *fo*F2 values measured by COSMIC and corresponding values modeled by (**b**) IRI and (**d**) IRTAM. Corresponding *Res. Mean*, *RMSE*, and *NRMSE* values, along with the total number of counts on which the statistics is based, are reported in the upper left box of each plot.

Moreover, the comparisons $RMSE_{IRTAM} = 1.120$ MHz vs. $RMSE_{IRI} = 1.053$ MHz and $NRMSE_{IRTAM} = 17.748\%$ vs. $NRMSE_{IRI} = 16.668\%$ show a slight worsening of the *fo*F2 modeling precision made by IRTAM. This is also supported by the IRTAM density plot which appears more scattered than the IRI one ($R_{IRTAM} = 0.888$ vs. $R_{IRI} = 0.901$). However, given that the statistical quantities calculated with the two models do not differ significantly, we can claim that the IRI and IRTAM performances based on the COSMIC *fo*F2 data can be considered somewhat comparable.

The statistical results reported in the following Table 4 show that only in a few cases is *Res. Mean*$_{IRTAM}$ < *Res. Mean*$_{IRI}$, while in all cases $RMSE_{IRI} < RMSE_{IRTAM}$, $NRMSE_{IRI}$ < $NRMSE_{IRTAM}$, and $R_{IRI} > R_{IRTAM}$. This means that the same concerns held for the entire COSMIC *fo*F2 dataset (Figure 19) hold also when IRI and IRTAM performances are evaluated on the COSMIC *fo*F2 time series binned according to the procedures described in Section 4.2, and considering the three levels of solar activity defined in Section 3.1.
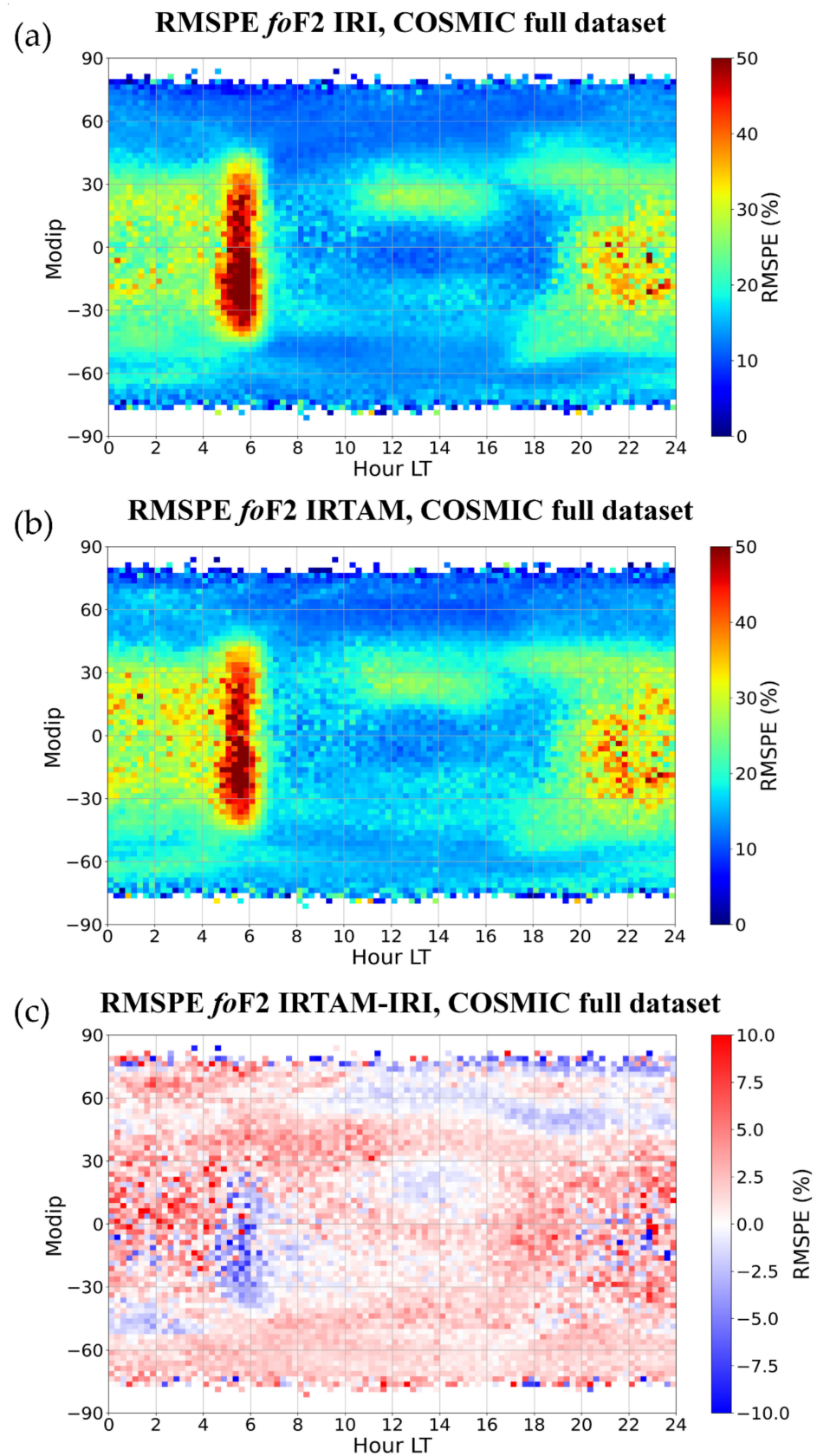
Table 4. *Res. Mean*, *RMSE*, *NRMSE*, and *R* values calculated for the IRI and IRTAM models on the basis of the *fo*F2 COSMIC observations, for the bins defined in Section 4.2, according to the three levels of solar activity defined in Section 3.1, and for the full dataset (bottom row). The number of counts on which the statistics were calculated is reported in the rightmost column.

| COSMIC Dataset	Model	Res. Mean [MHz]	RMSE [MHz]	NRMSE [%]	R	Counts
Daytime	IRI	0.012	1.058	14.841	0.891	998,544
	IRTAM	0.048	1.127	15.801	0.875	
Nighttime	IRI	0.126	1.076	21.781	0.856	422,509
	IRTAM	0.172	1.157	23.424	0.834	
Solar terminator	IRI	0.146	1.014	17.872	0.883	370,549
	IRTAM	0.136	1.059	18.672	0.871	
March Equinox	IRI	0.332	1.157	17.291	0.911	454,118
	IRTAM	0.257	1.188	17.754	0.902	
June Solstice	IRI	−0.068	0.917	15.715	0.898	439,417
	IRTAM	−0.021	0.985	16.870	0.881	
September Equinox	IRI	−0.044	1.044	16.724	0.908	447,872
	IRTAM	0.056	1.080	17.304	0.900	
December Solstice	IRI	0.041	1.077	16.665	0.888	450,195
	IRTAM	0.085	1.210	18.721	0.859	
LSA	IRI	−0.085	0.899	17.064	0.879	750,993
	IRTAM	−0.233	1.000	18.972	0.859	
MSA	IRI	0.080	1.075	16.357	0.891	608,461
	IRTAM	0.180	1.092	16.613	0.890	
HSA	IRI	0.312	1.252	16.142	0.891	432,148
	IRTAM	0.547	1.337	17.237	0.890	
Quiet magnetic activity	IRI	0.061	1.043	16.561	0.902	1,660,108
	IRTAM	0.090	1.111	17.643	0.888	
Moderate magnetic activity	IRI	0.138	1.176	18.046	0.895	129,917
	IRTAM	0.159	1.232	18.897	0.885	
Disturbed magnetic activity	IRI	0.165	1.561	21.840	0.841	1577
	IRTAM	0.251	1.537	21.500	0.851	
Low Modip	IRI	−0.422	1.373	17.518	0.872	267,449
	IRTAM	−0.254	1.481	18.902	0.844	
Mid Modip	IRI	0.139	1.036	16.496	0.903	1,230,835
	IRTAM	0.168	1.100	17.521	0.890	
High Modip	IRI	0.208	0.746	14.746	0.861	293,318
	IRTAM	0.107	0.766	15.129	0.842	
Full dataset	IRI	0.067	1.053	16.688	0.901	1,791,602
	IRTAM	0.095	1.120	17.748	0.888	

In order to provide a comprehensive picture of how IRI and IRTAM models behave when the COSMIC *fo*F2 dataset is considered for validation, global maps of IRI and IRTAM *RMSPE* values were calculated along with the corresponding maps of the differences ($RMSPE_{IRTAM} - RMSPE_{IRI}$). It is important to note that for validating both models in terms of spatial, diurnal, seasonal, and solar activity variability, the maps were calculated binning IRI, IRTAM and (IRTAM-IRI) *RMSPE* values as a function of: geographic latitude vs. geographic longitude, modip vs. LT hour, modip vs. doy, and modip vs. F10.7$_{81}$ (see Figures 20–23).
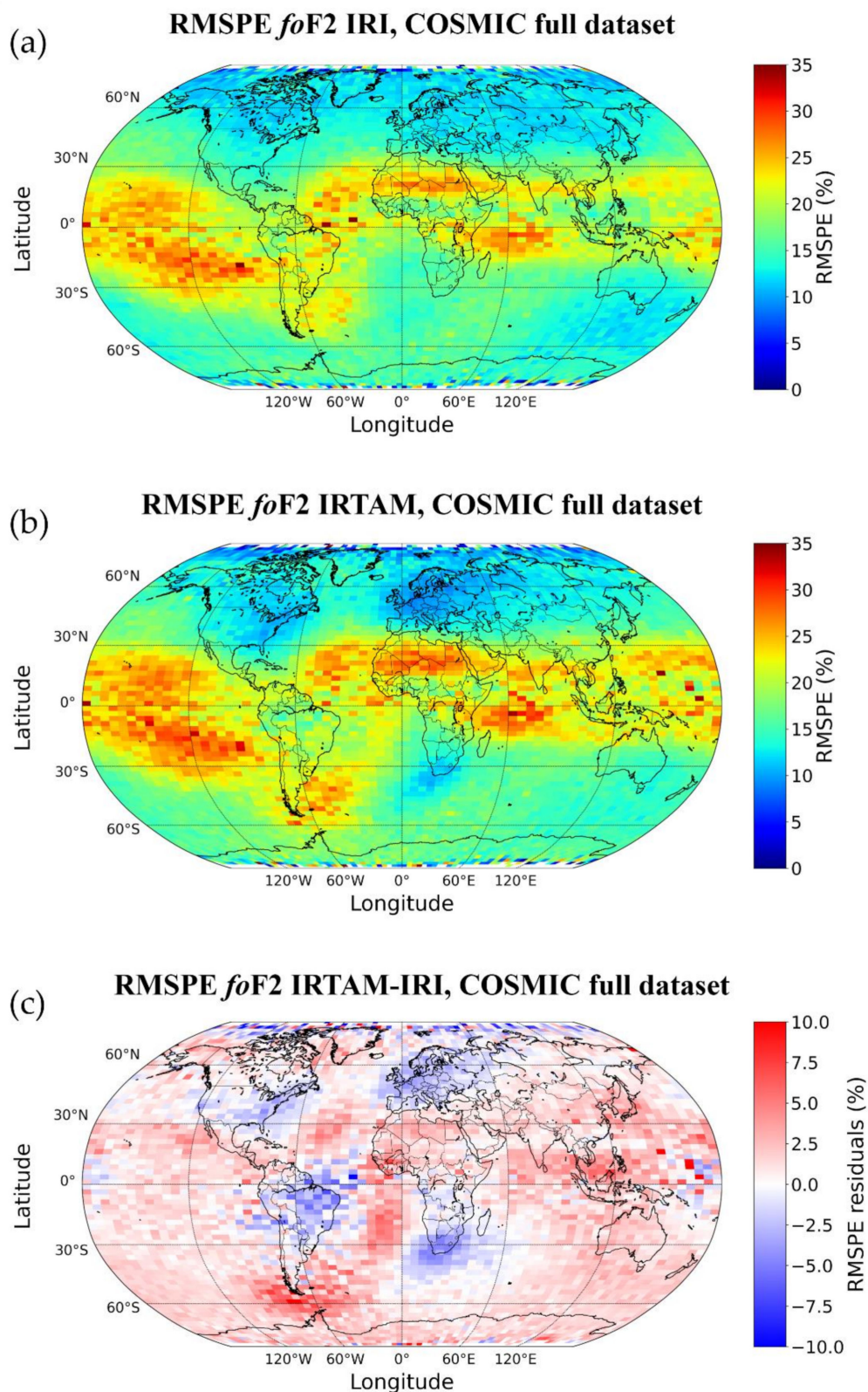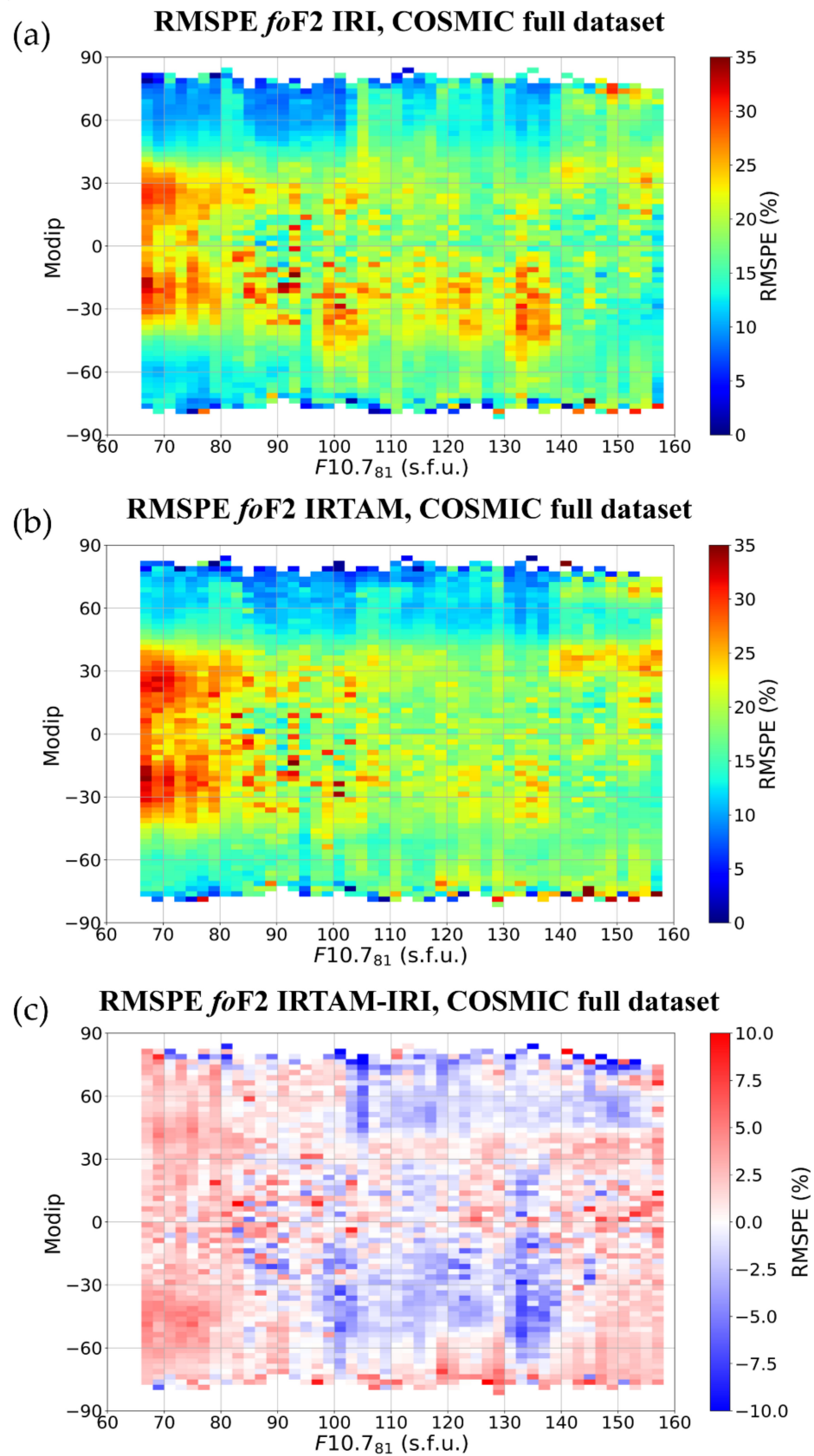
Figure 20. Maps of binned *RMSPE* values calculated between *fo*F2 values measured by COSMIC satellites and modeled by (**a**) IRI and (**b**) IRTAM, and (**c**) the corresponding map of differences. Data were binned as a function of the geographic latitude (2.5°-wide bins) and longitude (5°-wide bins).

(a)

RMSPE *fo*F2 IRI, COSMIC full dataset

(b)

RMSPE *fo*F2 IRTAM, COSMIC full dataset

(c)

RMSPE *fo*F2 IRTAM-IRI, COSMIC full dataset

Figure 21. Grids of binned *RMSPE* values calculated between *fo*F2 measured by COSMIC satellites and modeled by (**a**) IRI and (**b**) IRTAM, and (**c**) the corresponding grid of differences. Data were binned as a function of the LT (fifteen minute-wide bins) and modip (2.5°-wide bins).
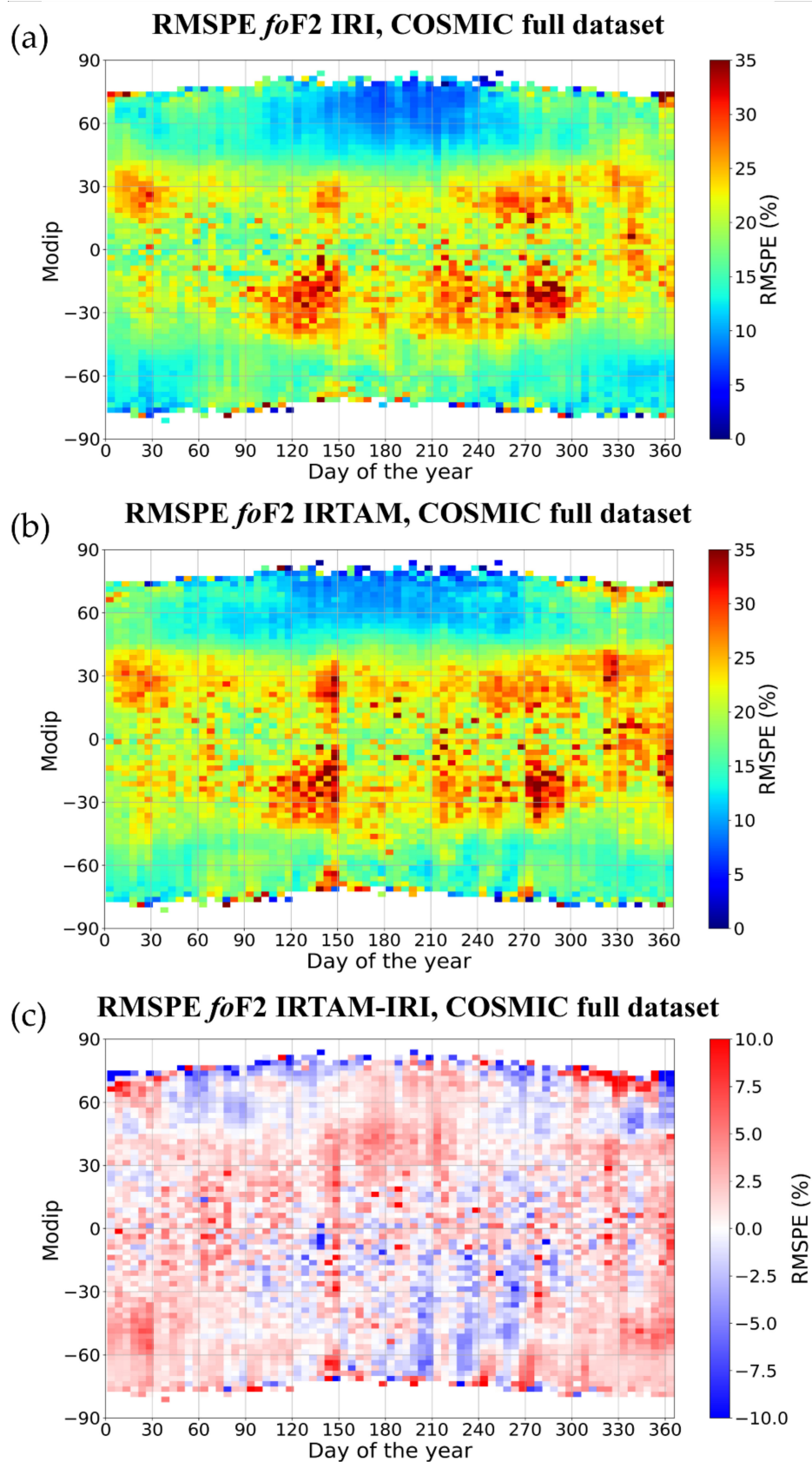
Figure 22. Same as Figure 21 but in this case data were binned as a function of the day of the year (five-day wide bins) and modip (2.5°-wide bins).

(a) **RMSPE *fo*F2 IRI, COSMIC full dataset**

(b) **RMSPE *fo*F2 IRTAM, COSMIC full dataset**

(c) **RMSPE *fo*F2 IRTAM-IRI, COSMIC full dataset**

Figure 23. Same as Figure 21 but in this case data were binned as a function of the $F10.7_{81}$ solar activity index (two s.f.u.-wide bins) and modip (2.5°-wide bins).

Figure 20 highlights significant differences between the two models. Specifically, the actual ionospheric conditions are better represented by IRTAM over those areas where a high concentration of ionosondes is available for data assimilation. Specifically, a clear improvement in the *fo*F2 modeling is observed over Europe, South Africa, North America, and Brazil, where the *RMSPE* percentage error is remarkably reduced when IRTAM is run. Vice versa, over the oceans, deserts (e.g., Sahara, Gobi, Kalahari) and Antarctica as well, a better performance from the IRI side is observed. These circumstances can explain why the results of Figure 19 show a global balance between the two models. As expected from an assimilative model, IRTAM improves considerably IRI predictions over those regions where it is possible to assimilate a large number of data. However, a little unexpected is the observed situation that IRTAM does not seem to reproduce the IRI background in areas very far away from the assimilation sites, which is a constraint that it is usually requested by an assimilative model (see, for example, [11,14]). In fact, since IRTAM works on the residuals between measured values and IRI values through NECTAR (see Section 2.2), it should correct the background model close to assimilation sites and relax to the background far away from assimilation sites. This is not the case represented in Figure 20. A possible explanation of this could lay on the global nature of the correction made by IRTAM to URSI coefficients. In fact, even if the diurnal correction is maximum at assimilation sites (Section 2.2), corrected diurnal coefficients are used to calculate the corrected spatial coefficients through the Jones and Gallet procedure; as a consequence, the correction affects the spatial harmonic functions that have a global nature. Because of this, the effect of the correction at one assimilation site can influence also a very distant location. On the one hand, this is a clear advantage because the correction has a global nature and is robust against sporadic data outliers but, on the other hand, the quality of the procedure critically depends on the spatial distribution of assimilation sites. It is clear from Figure 20 that the current Digisondes' global distribution is an important constraint for IRTAM. This accounts for the IRTAM performances that are better at mid latitudes than at low latitudes, as visible in Figures 20–23.

Figure 21 shows that IRI and IRTAM *RMSPE* maps obtained as a function of modip and local time are very similar: (a) both IRI and IRTAM show the best performance during daytime hours at northern high latitudes; (b) a worsening of both IRI and IRTAM performance is detected during the central hours of the day around 30° N; (c) the highest absolute and percentage errors occur in the latitudinal band (30° S, 30° N) around sunrise hours. This last circumstance was somehow expected from the IRI side because, being an empirical climatological model, IRI can have some limitations in describing the variability of the ionospheric characteristics when they present wide variations in a relatively small time window, as is the case at solar terminator hours. Nonetheless, the bottom panel of Figure 21 tells us that between 30° S and 30° N IRI performance was noticeably improved by IRTAM around sunrise hours. This fact constitutes a clear evidence of how data assimilation can play a key role in improving *fo*F2 modeling in certain situations. Both models exhibit quite large errors during nighttime at mid-low latitudes, as already evidenced by Jicamarca and Ascension Islands grids (Figures 4–9).

Figure 22 shows that from a seasonal point of view the performances of IRI and IRTAM are similar. Both models show most of the problems at low/equatorial latitudes throughout the year, especially in correspondence to the crests of the equatorial ionospheric anomaly, mostly the southern one. For what concerns the northern crest the output of IRI seems to be better than the IRTAM one. At mid and high latitudes, the seasonal dependence of the error associated with both models is remarkable, with lowest errors during summer months and highest during winter months.

Figure 23 shows the IRI and IRTAM *RMSPE* maps obtained as a function of modip and the solar activity. This figure highlights a clear improvement of the IRI performance made by IRTAM only for MSA conditions. The poor results observed for LSA from the IRI side were expected, because the conditions characterizing the solar minimum in 2008/2009 [64,65], much lower and prolonged than earlier minima, represented a chal-

lenge for IRI [66–68]. At the same time, it is somewhat unexpected that IRTAM cannot improve the IRI output for LSA, even though the model assimilates *fo*F2 values that well represent the very low solar activity conditions of that period. As seen before, most of the improvement made by IRTAM is narrowed to mid latitudes.

7. Validation Results for *hm*F2 Based on Ground-Based Ionosonde Observations

7.1. Statistics on the Full Dataset

The full *hm*F2 ionosonde dataset comprises 10,133,987 measurements obtained cumulating time series recorded from 2000 to 2019 at the 40 considered ionospheric stations. This dataset has been used to calculate the two-dimensional density plots of IRI-*hm*F2 vs. ionosonde-*hm*F2 and IRTAM-*hm*F2 vs. ionosonde-*hm*F2, along with the histograms of corresponding residuals, as shown in Figure 24.
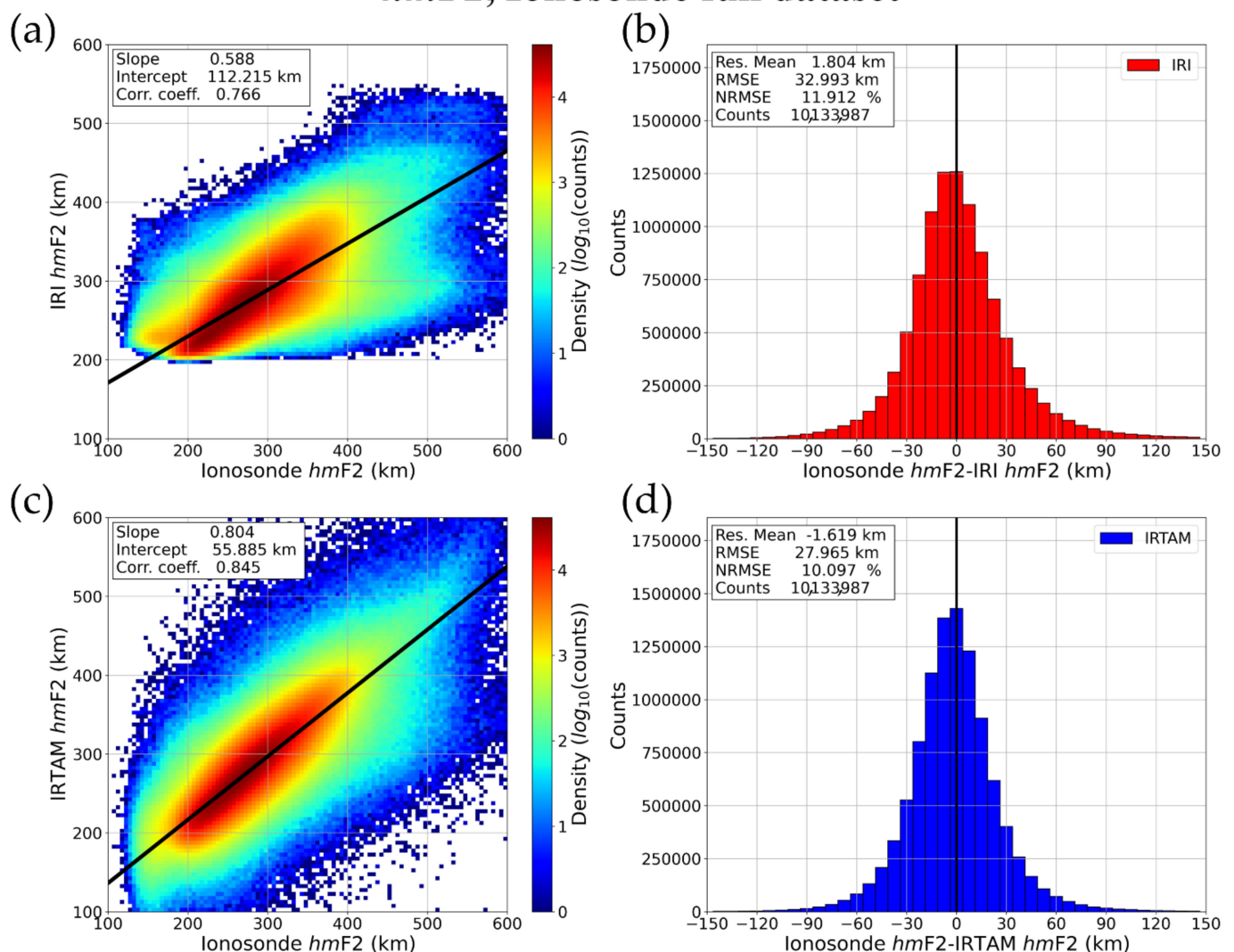
Figure 24. Density plots between measured (*x*-axis) and modeled (*y*-axis) *hm*F2 values, for (**a**) IRI and (**c**) IRTAM. Measured values are those from the entire ionosonde dataset. The black solid lines represent the best linear fit. Corresponding slope and intercept values, along with the correlation coefficient, are reported in the upper left box of each plot. Statistical distributions of the resultant residuals for (**b**) IRI and (**d**) IRTAM. Corresponding *Res. Mean*, *RMSE*, and *NRMSE* values, along with the total number of counts on which the statistics is based, are reported in the upper left box of each plot.

The comparison between IRI and IRTAM highlights that the IRTAM model slightly improves the *hm*F2 modeling accuracy. In fact, the IRTAM and IRI statistical distributions of residuals around the zero value are very similar and, on the other hand, *Res. Mean*$_{\text{IRTAM}}$ = −1.619 km is just a little bit smaller than *Res. Mean*$_{\text{IRI}}$ = 1.804 km. Additionally, the comparisons *RMSE*$_{\text{IRTAM}}$ = 27.965 km vs. *RMSE*$_{\text{IRI}}$ = 32.993 km and *NRMSE*$_{\text{IRTAM}}$ = 10.097% vs. *RMSE*$_{\text{IRI}}$ = 11.912% testify a slight improvement of the *hm*F2 modeling precision made by IRTAM. This point is also supported by the IRTAM density plot and the correlation coefficient (R_{IRTAM} = 0.845 vs. R_{IRI} = 0.766).

Compared to the *fo*F2 results (ionosonde dataset, Figure 3), in this case the improvement made by IRTAM is lower. However, we have to consider that, for what concerns *hm*F2, IRTAM cannot be considered as a direct updating of IRI because of the Brunini procedure [53] for mapping *hm*F2 (see Section 2.2) and, more importantly, because we applied the current default *hm*F2 IRI option, i.e., that of Shubin et al. [44]. Hence, IRI and IRTAM *hm*F2 models have to be considered as completely unrelated.

Overall, the statistical results reported in Table 5 show that IRTAM models *hm*F2 better than IRI in terms both of accuracy and precision, although the degree of improvement is lower than that obtained for *fo*F2 (Table 2).

Table 5. *Res. Mean, RMSE, NRMSE*, and *R* values calculated for the IRI and IRTAM models on the basis of the *hm*F2 ground-based ionosondes observations, for the bins defined in Section 4.2, according to the three levels of solar activity defined in Section 3.1, and for the full dataset (bottom row). The number of counts on which the statistics were calculated is reported in the rightmost column.

Ionosonde Stations Dataset	Model	Res. Mean [km]	RMSE [km]	NRMSE [%]	R	Counts
Daytime	IRI	−6.736	27.844	11.021	0.807	4,091,455
	IRTAM	−3.142	24.745	9.794	0.855	
Nighttime	IRI	10.293	38.663	12.625	0.580	4,198,762
	IRTAM	0.428	31.143	10.170	0.745	
Solar terminator	IRI	1.424	29.298	11.088	0.722	1,843,770
	IRTAM	−2.901	27.034	10.231	0.784	
March Equinox	IRI	2.693	30.997	11.096	0.788	2,710,599
	IRTAM	−1.338	26.225	9.388	0.859	
June Solstice	IRI	−5.670	34.761	12.768	0.743	2,255,233
	IRTAM	−2.196	29.699	10.908	0.825	
September Equinox	IRI	1.607	30.758	11.154	0.784	2,644,252
	IRTAM	−1.798	26.638	9.660	0.849	
December Solstice	IRI	7.735	35.608	12.723	0.765	2,523,903
	IRTAM	−1.217	29.494	10.539	0.844	
LSA	IRI	1.581	32.216	12.481	0.725	3,782,579
	IRTAM	−2.231	28.479	11.034	0.804	
MSA	IRI	4.591	33.831	12.081	0.736	3,801,672
	IRTAM	−0.750	28.283	10.100	0.829	
HSA	IRI	−2.019	32.867	10.943	0.757	2,549,736
	IRTAM	−2.006	26.689	8.886	0.853	
Quiet magnetic activity	IRI	0.682	31.877	11.593	0.772	9,147,468
	IRTAM	−1.580	27.402	9.966	0.845	
Moderate magnetic activity	IRI	11.754	41.017	13.914	0.720	959,046
	IRTAM	−1.983	32.279	10.950	0.832	
Disturbed magnetic activity	IRI	28.087	66.786	20.793	0.595	27,473
	IRTAM	−1.683	45.640	14.210	0.807	
Full dataset	IRI	1.804	32.993	11.912	0.766	10,133,987
	IRTAM	−1.619	27.965	10.097	0.845	

Table 6 shows the statistical parameters (1–4) calculated for the hmF2 time series recorded in each of the 40 ionosonde stations listed in Table 1. Differently from the foF2 results shown in Table 3, the number of cases for which $Res.Mean_{IRI} < Res.Mean_{IRTAM}$ and for which $Res.Mean_{IRTAM} < Res.Mean_{IRI}$ is practically the same. This circumstance is in accordance with the very similar IRI (1.804 km) and IRTAM (-1.619 km) $Res.Mean$ absolute values calculated when the whole hmF2 ionosonde dataset is considered (bottom row of Table 5). This fact points out that the two models are practically equivalent in terms of accuracy. Nonetheless, the IRTAM $RMSE$ and $NRMSE$ values are always lower than the IRI ones, with the exception of the stations of Anyang and I-Cheon, Kwajalein, and Rome. This fact highlights that IRTAM models hmF2 with a precision better than that of IRI.

Table 6. Same as Table 5, but for each of the hmF2 datasets relative to the 40 ionosonde stations listed in Table 1. The number of counts on which the statistics were calculated is reported in the rightmost column.

Ionosonde (Country)	Model	Res. Mean [km]	RMSE [km]	NRMSE [%]	R	Counts
Anyang and I-Cheon (South Korea)	IRI	−0.211	29.869	10.995	0.779	201,558
	IRTAM	−1.389	30.911	11.379	0.794	
Ascension Island (UK)	IRI	0.826	34.775	12.186	0.638	264,916
	IRTAM	−2.158	32.434	11.366	0.784	
Athens (Greece)	IRI	−4.951	33.230	12.370	0.716	313,227
	IRTAM	−6.015	26.230	9.765	0.851	
Boa Vista (Cape Verde)	IRI	−6.362	35.615	11.520	0.736	60,295
	IRTAM	3.007	26.703	8.637	0.852	
Boulder (USA)	IRI	4.083	29.890	11.126	0.763	420,504
	IRTAM	−3.033	25.579	9.522	0.836	
Cachoeira Paulista (Brazil)	IRI	16.579	45.903	15.790	0.499	158,881
	IRTAM	8.994	32.261	11.097	0.789	
Chilton (U.K.)	IRI	−3.192	29.583	10.891	0.831	269,815
	IRTAM	−7.850	28.018	10.315	0.861	
Dourbes (Belgium)	IRI	0.887	25.512	9.526	0.845	328,740
	IRTAM	−1.512	21.726	8.112	0.889	
Dyess AFB (USA)	IRI	3.456	36.564	13.355	0.708	142,345
	IRTAM	−5.363	32.746	11.961	0.785	
Eielson (USA)	IRI	−15.912	28.692	11.911	0.759	54,960
	IRTAM	−8.381	24.055	9.986	0.817	
El Arenosillo (Spain)	IRI	11.753	30.989	10.907	0.792	206,362
	IRTAM	7.118	24.843	8.744	0.866	
Fortaleza (Brazil)	IRI	3.462	41.126	13.206	0.725	115,951
	IRTAM	9.718	34.957	11.225	0.832	
Gakona (USA)	IRI	−12.830	36.554	14.376	0.729	188,948
	IRTAM	−15.801	35.878	14.110	0.784	
Goose Bay (Canada)	IRI	5.420	36.757	13.826	0.792	138,063
	IRTAM	3.144	34.465	12.964	0.801	
Grahamstown (South Africa)	IRI	2.539	25.744	9.616	0.764	389,105
	IRTAM	−2.788	23.069	8.617	0.825	
Guam (USA)	IRI	−15.517	36.903	12.338	0.700	150,696
	IRTAM	−2.786	27.130	9.071	0.860	
Hermanus (South Africa)	IRI	3.261	21.424	8.029	0.827	297,534
	IRTAM	−6.104	20.670	7.746	0.862	

Table 6. *Cont.*

Ionosonde (Country)	Model	Res. Mean [km]	RMSE [km]	NRMSE [%]	R	Counts
Jicamarca (Peru)	IRI	4.135	41.582	12.738	0.777	379,962
	IRTAM	−0.082	29.664	9.087	0.901	
Juliusruh (Germany)	IRI	1.661	27.765	10.207	0.837	473,533
	IRTAM	−1.090	23.896	8.785	0.880	
King Salmon (USA)	IRI	31.973	62.509	20.815	0.696	180,833
	IRTAM	16.755	42.827	14.261	0.832	
Kwajalein (Marshall Islands)	IRI	5.016	38.853	12.565	0.581	160,301
	IRTAM	−3.969	40.814	13.199	0.747	
Learmonth (Australia)	IRI	8.493	32.307	11.521	0.727	262,583
	IRTAM	−2.128	29.816	10.633	0.805	
Louisvale (South Africa)	IRI	1.622	25.004	9.063	0.782	191,705
	IRTAM	−4.845	22.754	8.247	0.849	
Millstone Hill (USA)	IRI	−8.326	31.833	11.904	0.779	441,837
	IRTAM	−6.702	27.867	10.421	0.834	
Moscow (Russia)	IRI	−3.617	22.775	8.747	0.840	301,375
	IRTAM	−3.259	19.542	7.505	0.888	
Nicosia (Cyprus)	IRI	−3.840	26.716	10.018	0.784	110,598
	IRTAM	−3.482	23.631	8.861	0.853	
Nord Greenland (Greenland)	IRI	−4.422	41.196	15.602	0.351	47,039
	IRTAM	−5.385	38.360	15.528	0.530	
Norilsk (Russia)	IRI	−0.618	27.228	10.994	0.622	145,454
	IRTAM	−2.677	17.375	7.015	0.873	
Point Arguello (USA)	IRI	10.669	35.686	12.619	0.716	377,422
	IRTAM	5.099	26.358	9.321	0.854	
Port Stanley (Falkland Islands)	IRI	−2.350	34.875	12.103	0.805	212,867
	IRTAM	−4.516	28.228	9.796	0.878	
Pruhonice (Czech Republic)	IRI	−1.091	24.158	9.111	0.851	396,878
	IRTAM	−2.741	22.586	8.518	0.874	
Ramey (Puerto Rico)	IRI	4.184	33.092	11.533	0.732	231,544
	IRTAM	1.407	25.776	8.983	0.851	
Rome (Italy)	IRI	−5.319	27.841	10.189	0.844	520,519
	IRTAM	−9.187	28.701	10.503	0.850	
Roquetes (Spain)	IRI	3.402	23.623	8.551	0.864	443,895
	IRTAM	0.473	21.542	7.798	0.891	
San Vito (Italy)	IRI	7.930	29.538	10.645	0.786	372,074
	IRTAM	5.167	26.355	9.497	0.844	
Sao Luis (Brazil)	IRI	11.516	43.761	13.375	0.694	122,965
	IRTAM	6.016	34.662	10.594	0.827	
Sondrestrom (Greenland)	IRI	1.676	45.625	16.934	0.537	200,111
	IRTAM	−3.123	40.476	15.023	0.673	
Thule (Greenland)	IRI	12.306	47.456	16.678	0.308	278,374
	IRTAM	5.532	35.084	12.330	0.696	
Tromso (Norway)	IRI	−3.295	33.955	13.577	0.551	259,000
	IRTAM	−6.586	27.803	11.118	0.749	
Wallops Island (USA)	IRI	−0.615	28.794	10.395	0.764	321,218
	IRTAM	0.698	25.755	9.298	0.833	

7.2. Diurnal, Seasonal, and Solar Activity Statistics Variations for Different Zonal Sectors

The same criteria described in Section 5.2 for the selection of the ionospheric stations and the validation of the results based on ground-based *fo*F2 ionosonde observations, were also adopted to validate the results of IRI and IRTAM models for *hm*F2. Therefore, as done in Section 5.2, the statistical grids for the same five ionospheric stations previously selected are shown in Figures 25–39 for *hm*F2.
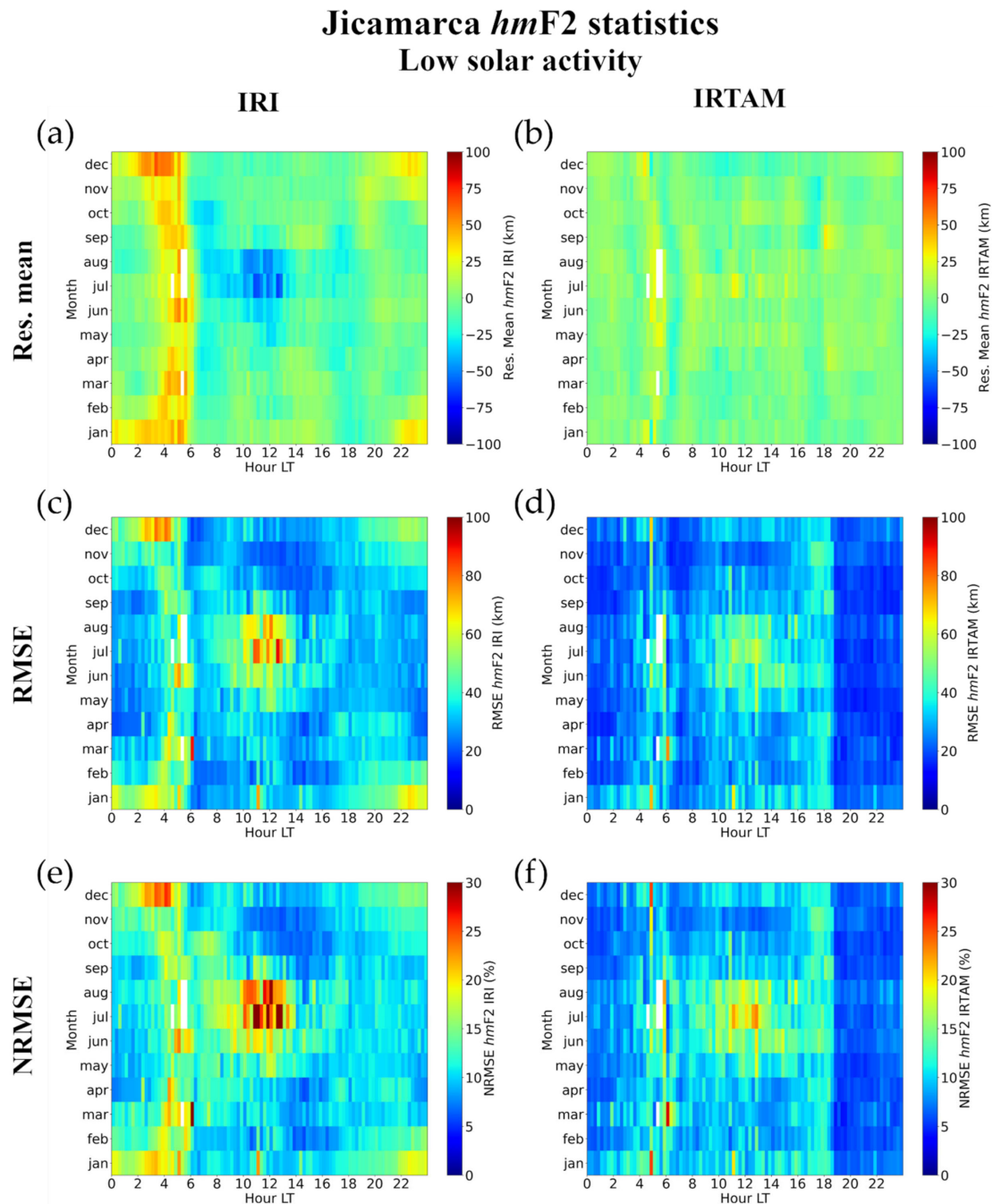
Figure 25. Grids of (**a**,**b**) *Res. Mean*, (**c**,**d**) *RMSE*, and (**e**,**f**) *NRMSE* values calculated between *hm*F2 values modeled by (**a**,**c**,**e**) IRI and (**b**,**d**,**f**) IRTAM, and those measured by the Jicamarca ionosonde for LSA. Monthly values (*y*-axis) are binned in fifteen minute-wide bins in LT (*x*-axis). The white color highlights bins with less than ten counts for which values were not calculated because considered not statistically significant.

Jicamarca *hm*F2 statistics
Mid solar activity
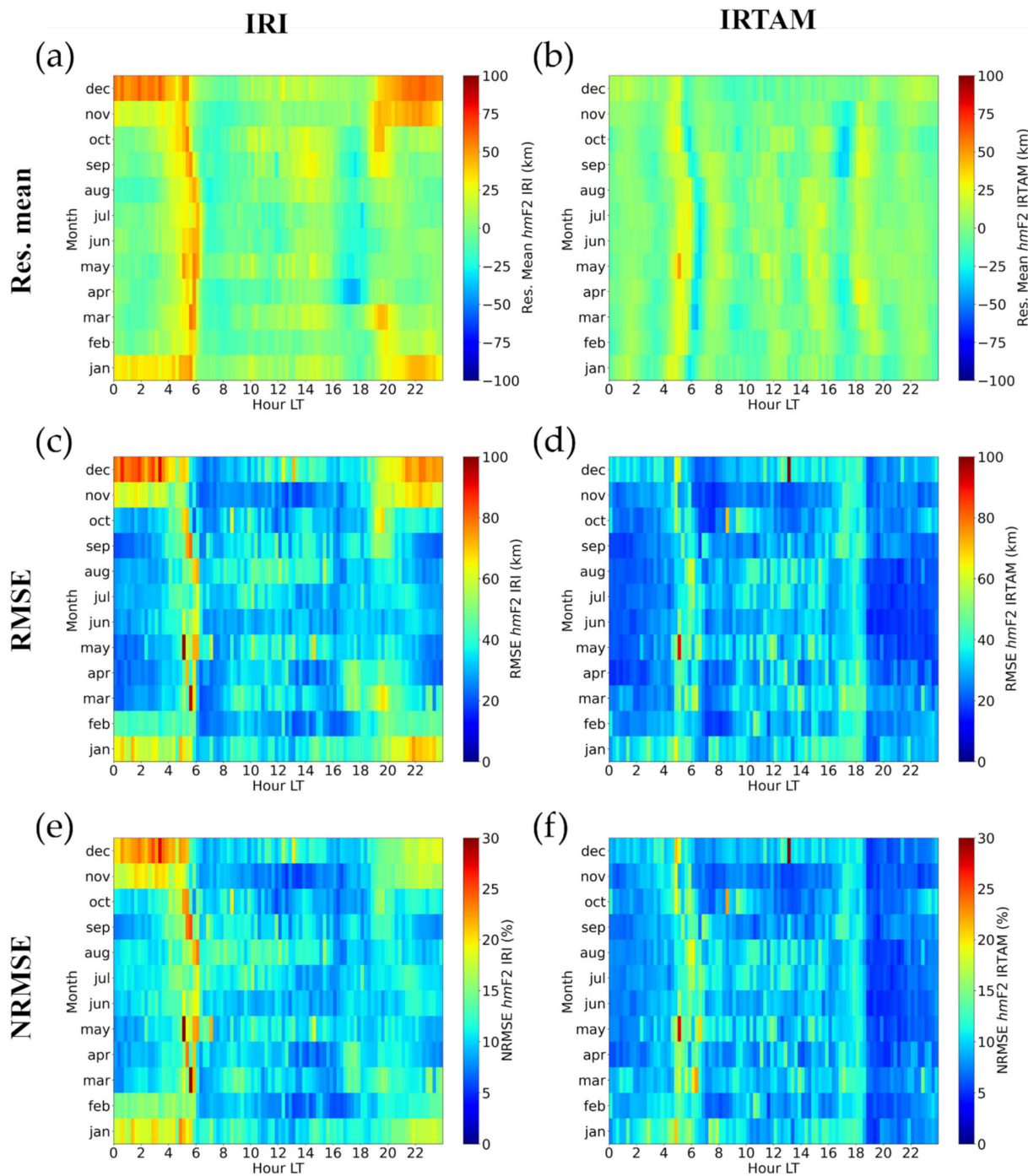
Figure 26. Same as Figure 25 but for MSA level.

Jicamarca *hm*F2 statistics
High solar activity

IRI   IRTAM
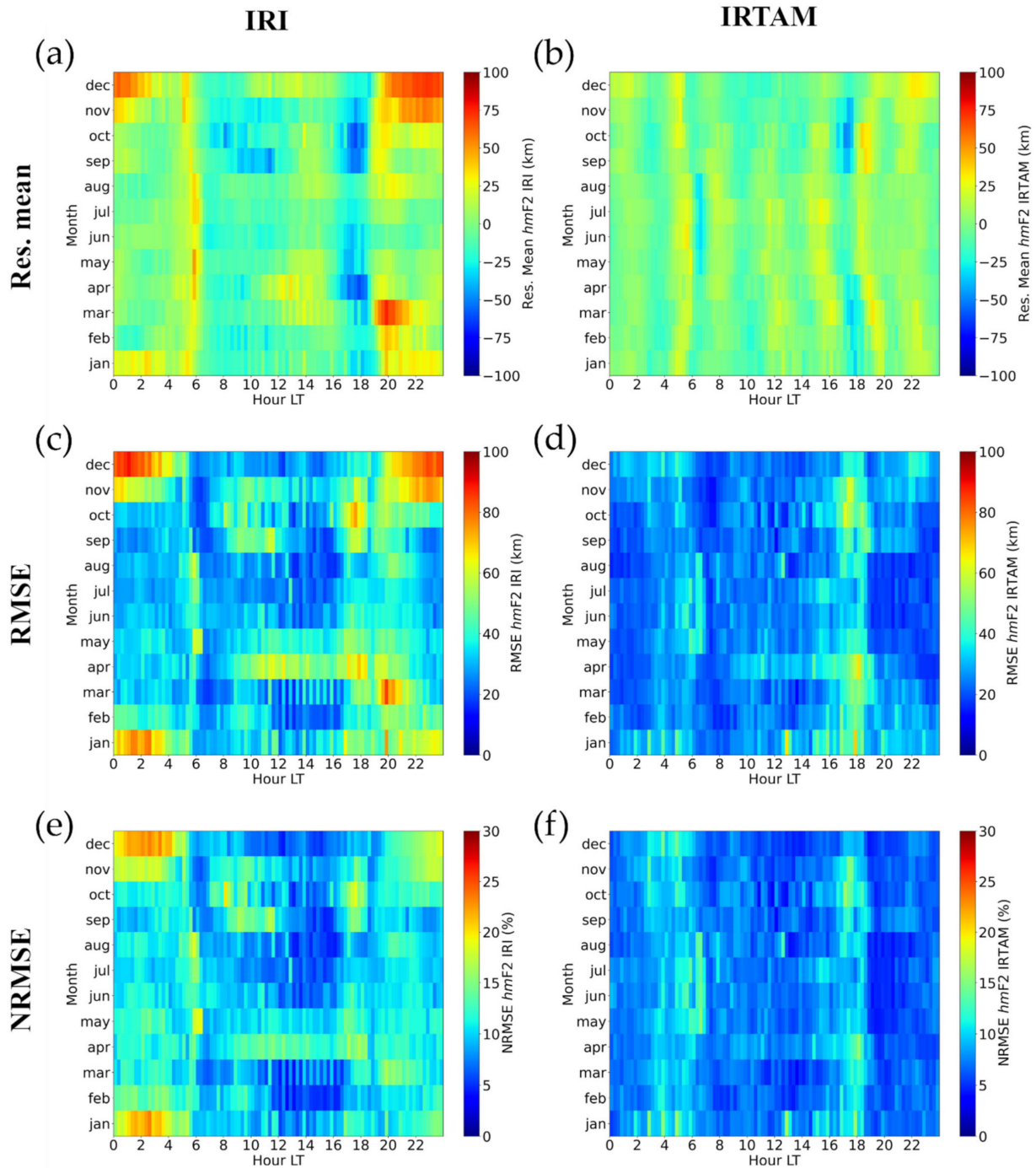
Figure 27. Same as Figure 25 but for HSA level.

Ascension Island *hm*F2 statistics
Low solar activity
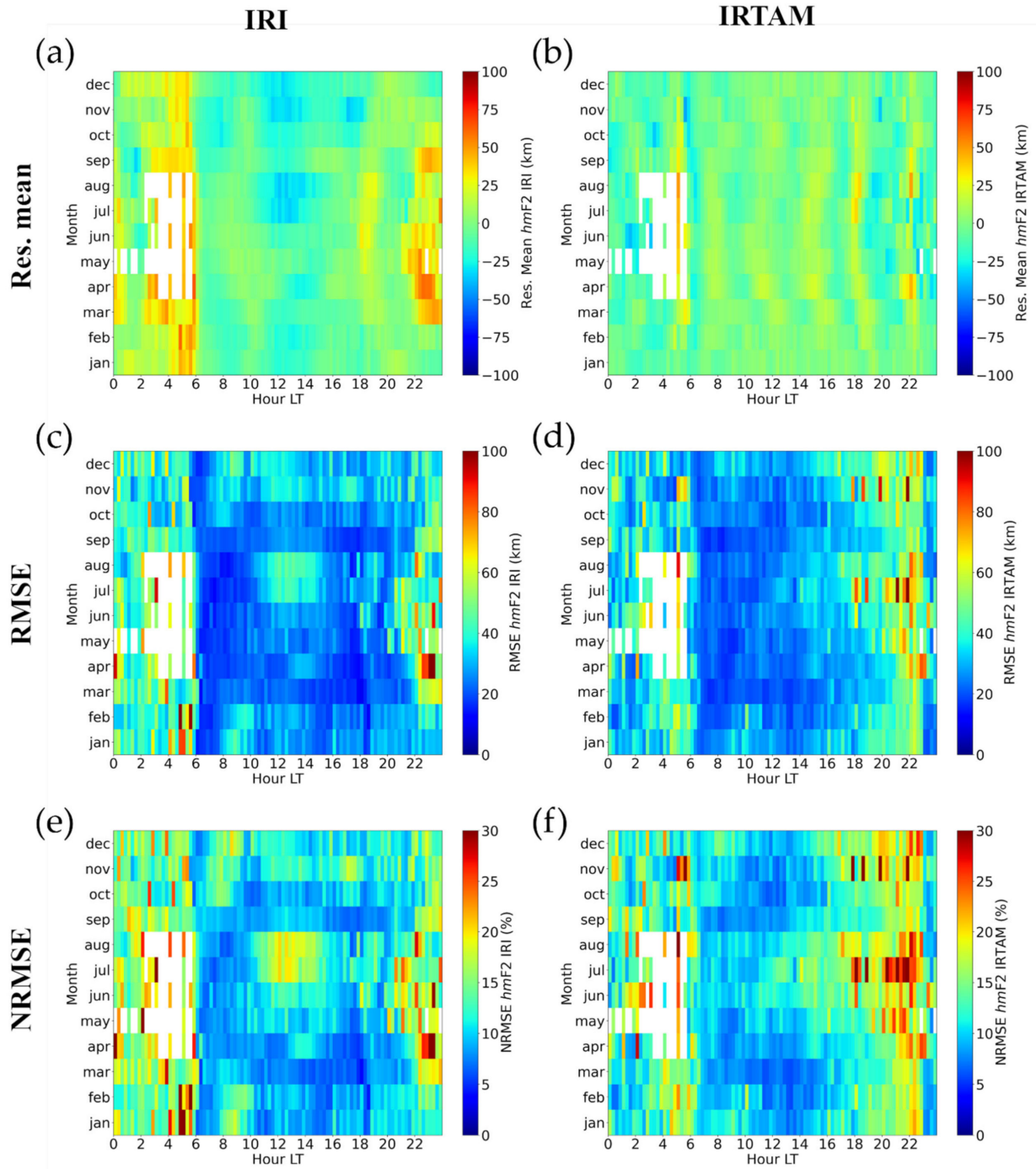
Figure 28. Same as Figure 25 but for Ascension Island.

Ascension Island *hm*F2 statistics
Mid solar activity

IRI   **IRTAM**
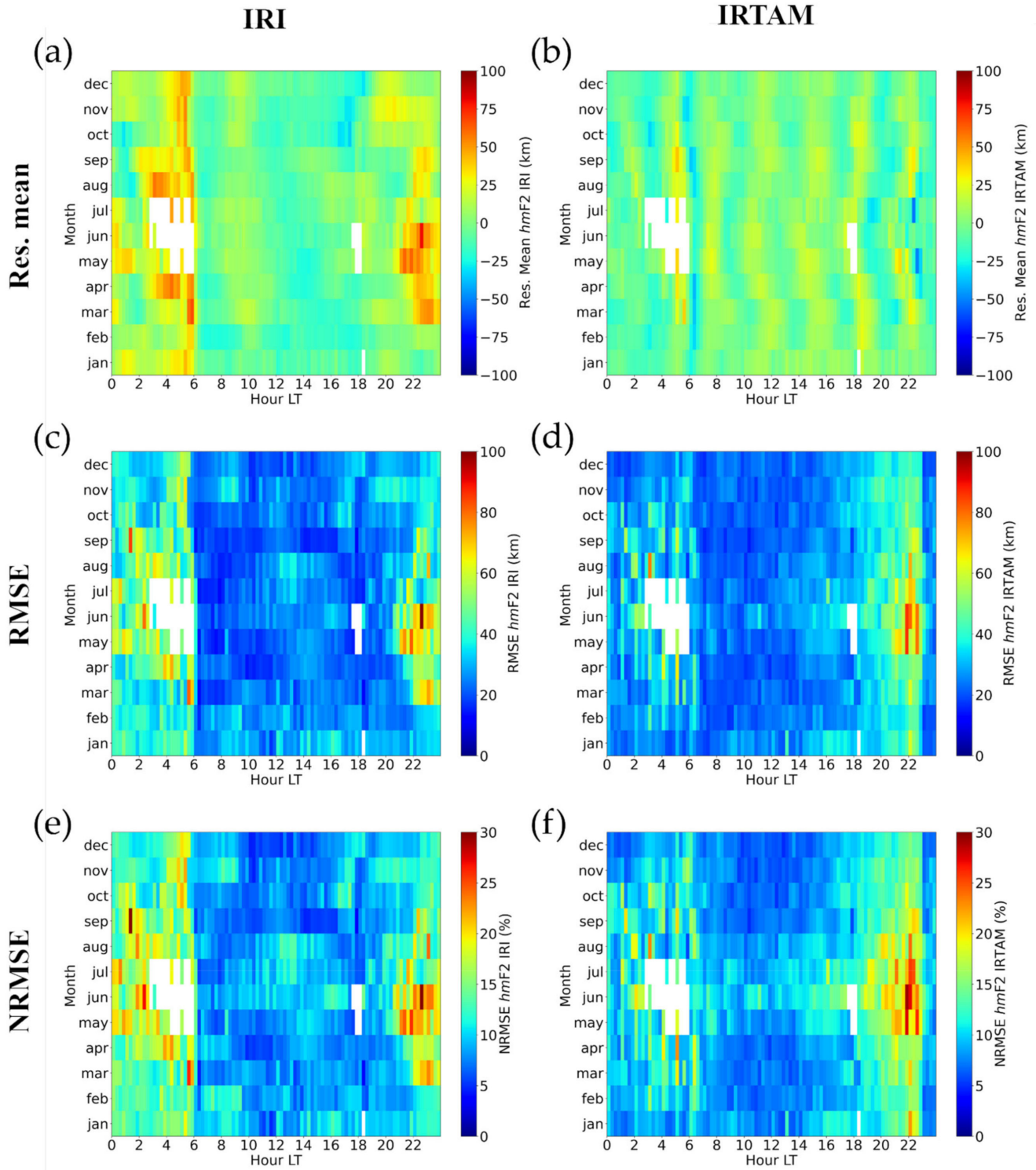
Figure 29. Same as Figure 28 but for MSA level.

Ascension Island *hm*F2 statistics
High solar activity

IRI

IRTAM
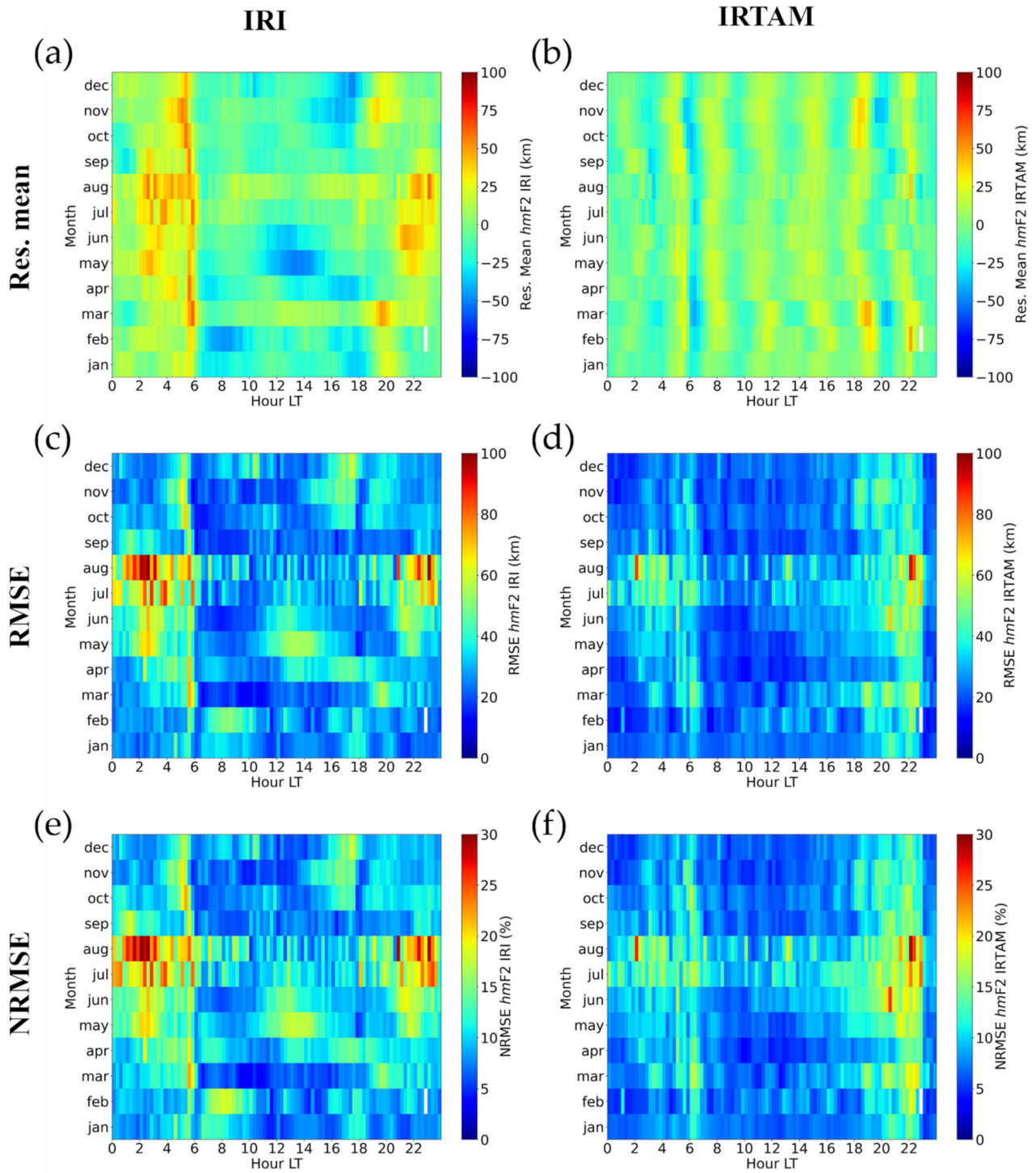
Figure 30. Same as Figure 28 but for HSA level.

Rome *hm*F2 statistics
Low solar activity
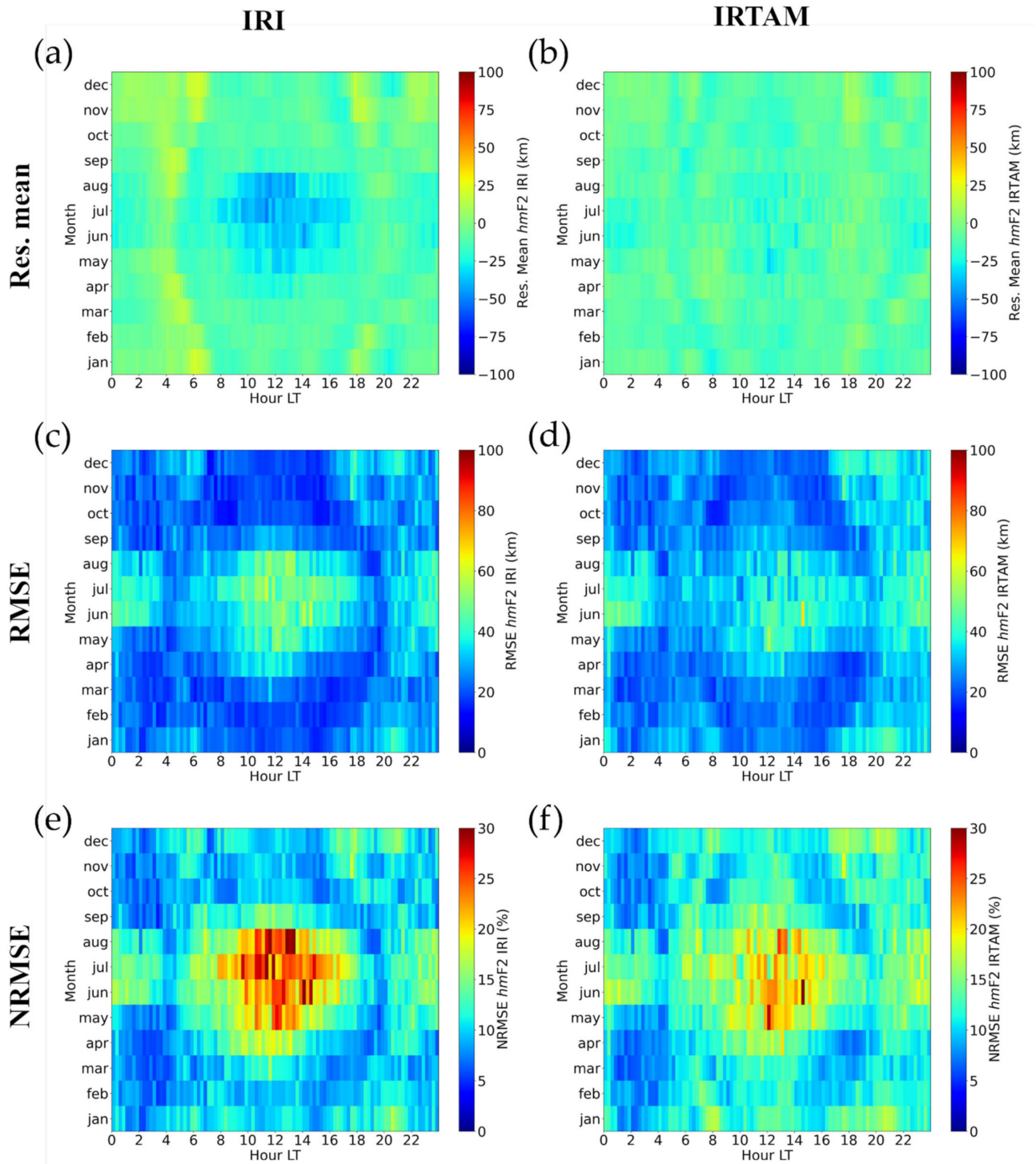
Figure 31. Same as Figure 25 but for Rome.

Rome *hm*F2 statistics
Mid solar activity
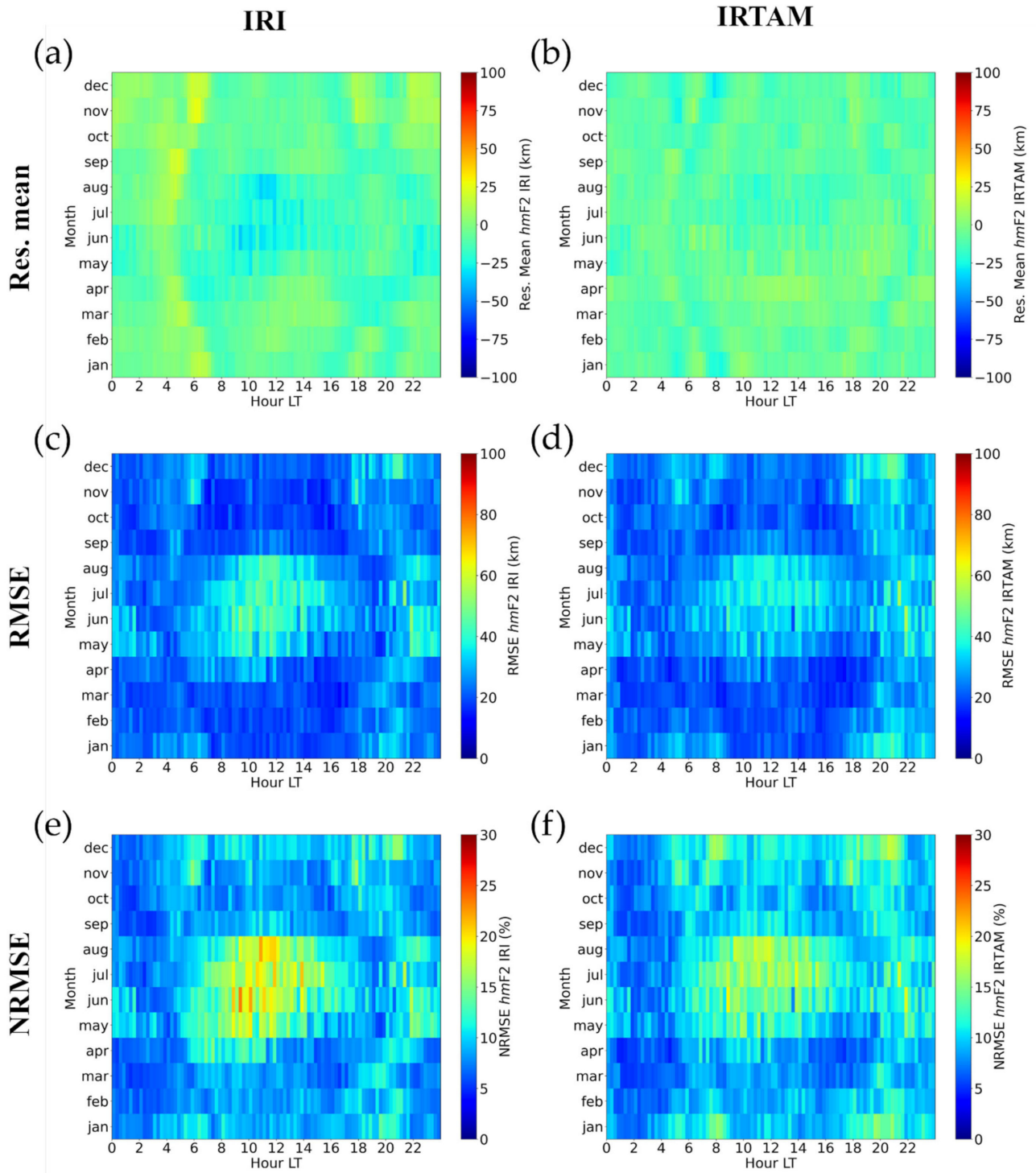
Figure 32. Same as Figure 31 but for MSA level.

Rome *hm*F2 statistics
High solar activity
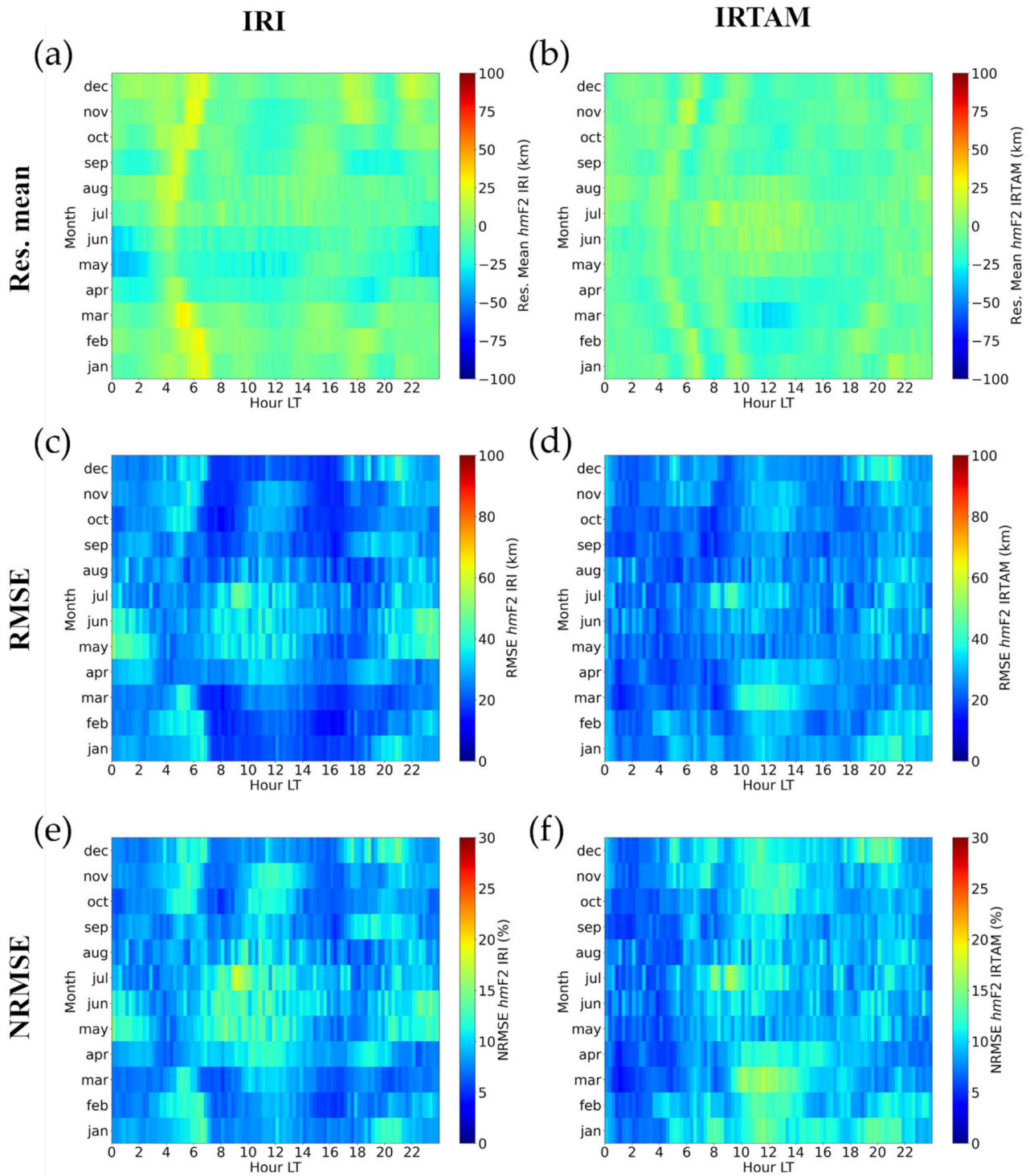
Figure 33. Same as Figure 31 but for HSA level.

Sondrestrom *hm*F2 statistics
Low solar activity

IRI

IRTAM
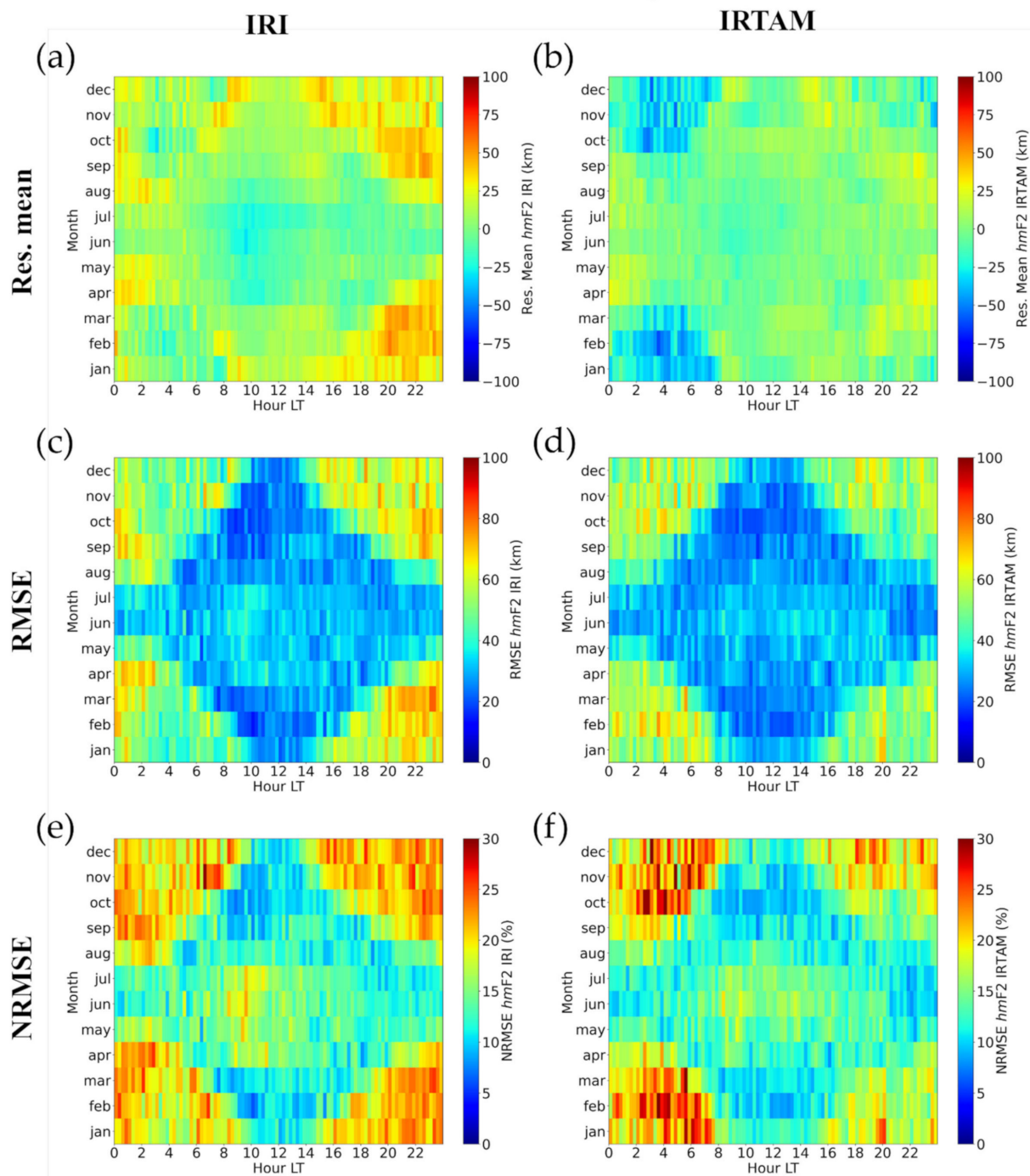
Figure 34. Same as Figure 25 but for Sondrestrom.

Sondrestrom *hm*F2 statistics
Mid solar activity

IRI

IRTAM
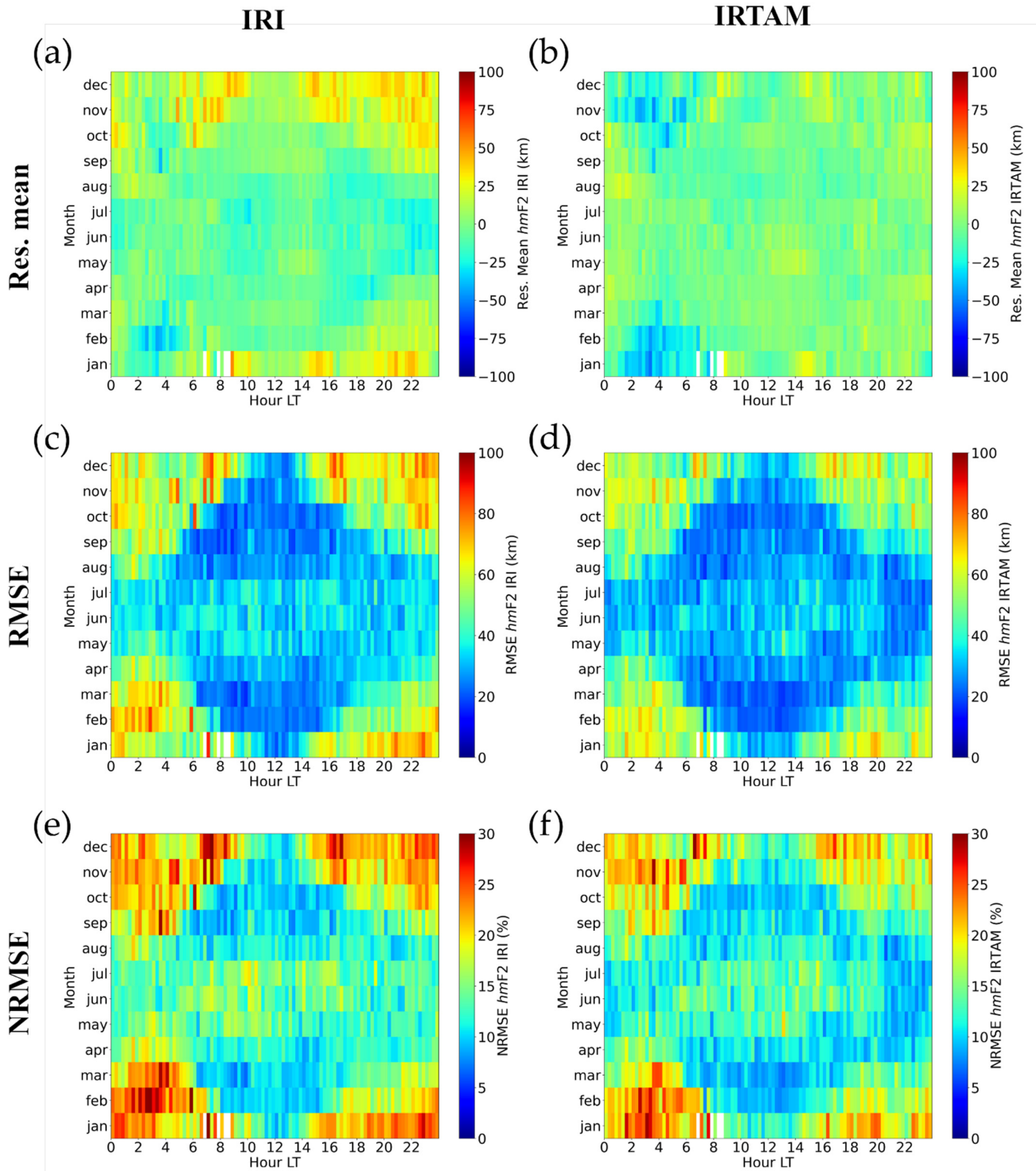
Figure 35. Same as Figure 34 but for MSA level.

Sondrestrom *hm*F2 statistics
High solar activity

IRI

IRTAM
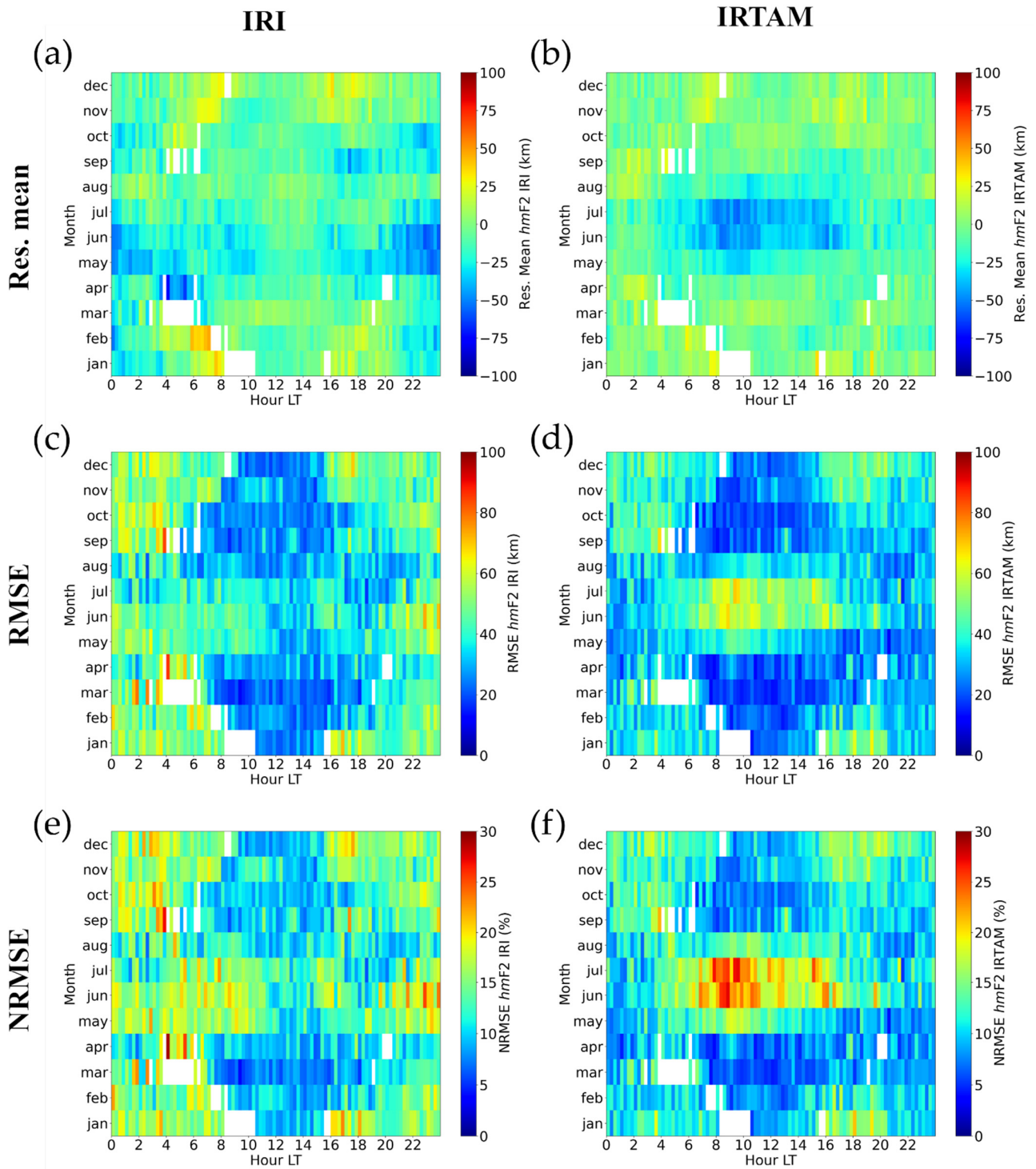
Figure 36. Same as Figure 34 but for HSA level.

Thule *hm*F2 statistics
Low solar activity

IRI

IRTAM
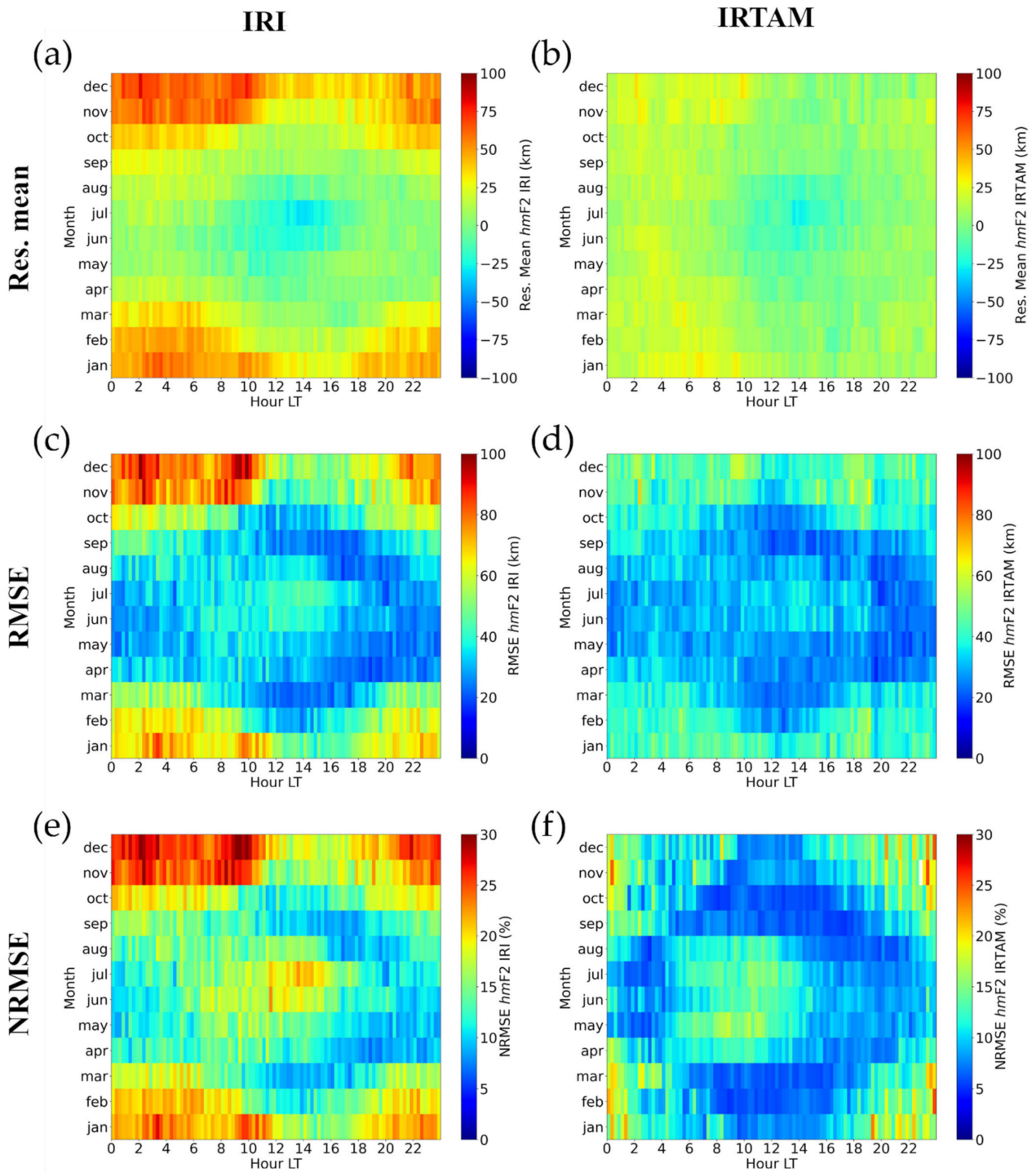
Figure 37. Same as Figure 25 but for Thule.

Thule *hm*F2 statistics
Mid solar activity
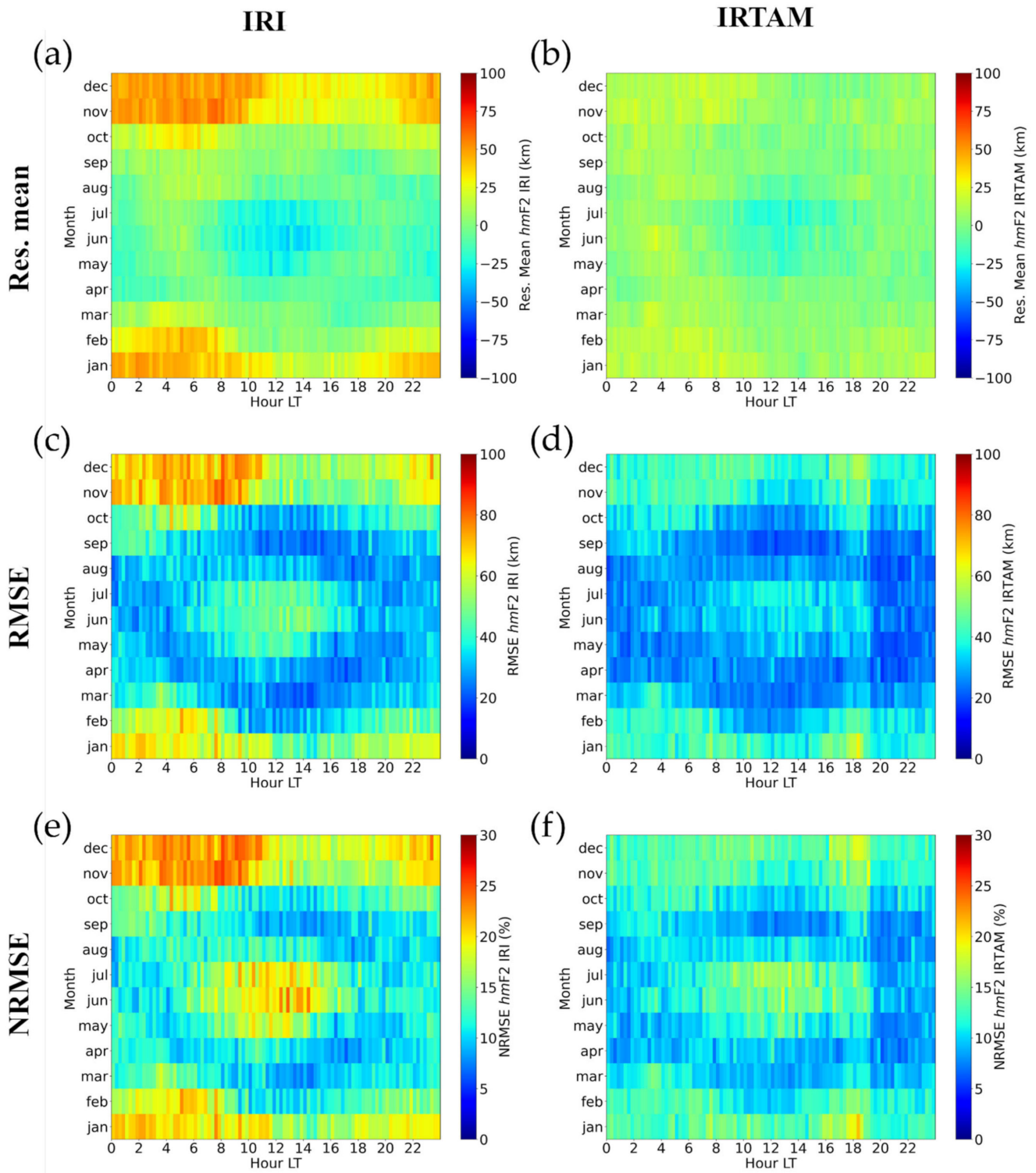
Figure 38. Same as Figure 37 but for MSA level.

Thule *hm*F2 statistics
High solar activity
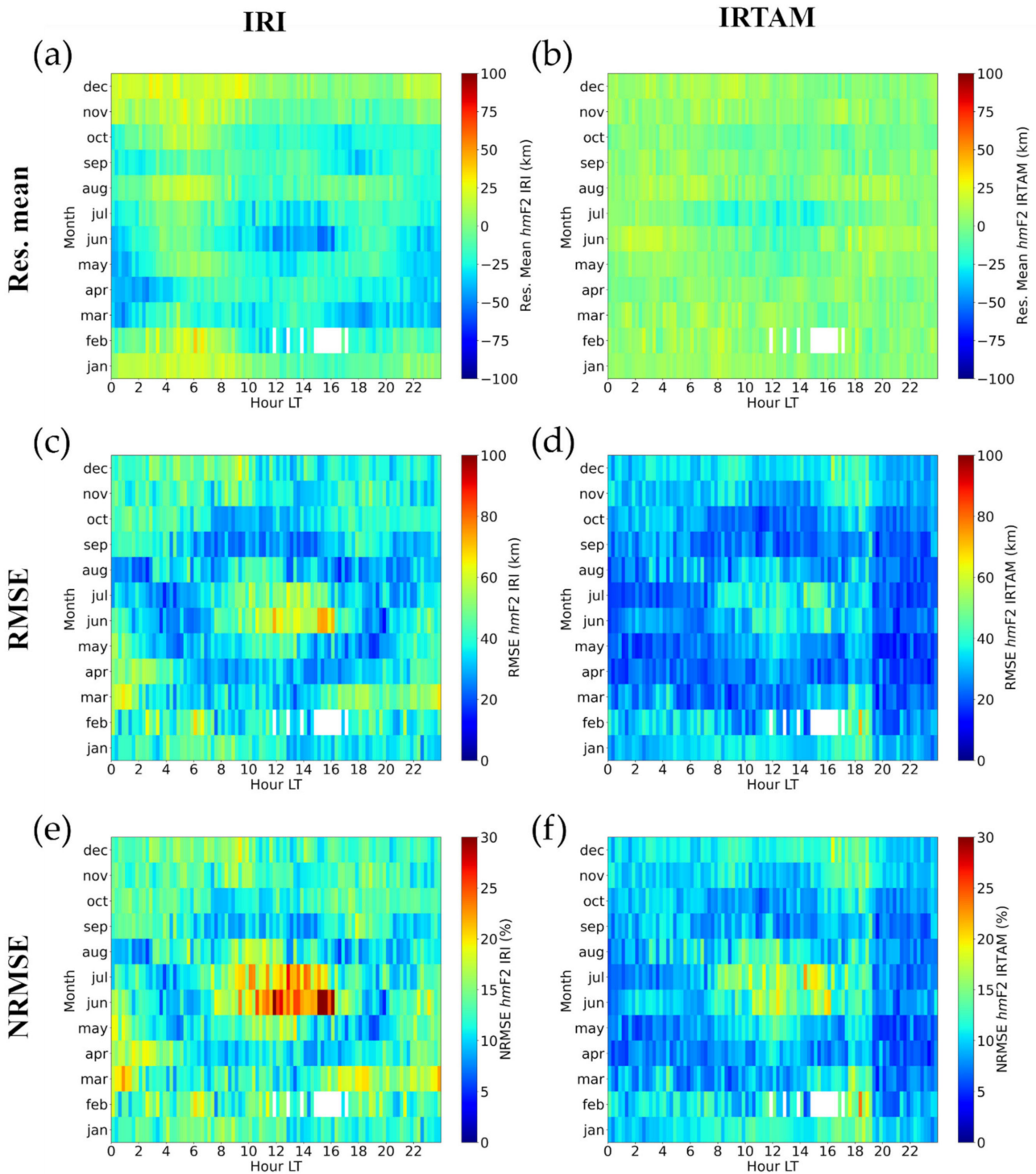
Figure 39. Same as Figure 37 but for HSA level.

Figures 25–27 show that at Jicamarca, for the three considered solar activity levels, IRTAM presents lower values of *Res. Mean*, *RMSE*, and *NRMSE* than IRI, which means that IRTAM improves the IRI *hm*F2 modeling in terms of both accuracy and precision. This improvement is observed mostly around sunrise and post-sunset hours. Moreover, IRTAM reduces the error made by IRI in the *hm*F2 modeling around noon in summer.

Figures 28–30 show that at Ascension Island, for the three considered solar activity levels, the IRI *hm*F2 modeling accuracy is improved by IRTAM especially during nighttime hours. Overall, for LSA, the *hm*F2 modeling precision of IRI and IRTAM can be considered equivalent during pre-sunrise and daytime hours, since the *RMSE* and *NRMSE* grids at these hours show similar patterns. IRI instead shows a better output during post-sunset hours. For MSA and HSA, concerning the modeling precision, a better performance of IRI is observed during pre-sunrise hours.

Figures 31–33 show that overall at Rome, for the three considered solar activity levels, IRTAM improves the IRI modeling accuracy, in particular for sunrise and daytime hours in summer. Concerning the modeling precision, both models are somewhat equivalent independently of solar activity. Nonetheless, a slight improvement made by IRTAM is found during the central hours of the day for summer months.

Figures 34–36 show that at Sondrestrom, for LSA and MSA, the IRI *hm*F2 modeling accuracy during post-sunset hours is remarkably improved by IRTAM, in particular for winter and equinoctial months. Nevertheless, a worsening of IRTAM performance is observed during pre-sunrise hours in winter. For HSA, a drop of the *hm*F2 modeling accuracy is observed from the IRTAM side during daytime hours in summer. For LSA and MSA, *RMSE* and *NRMSE* grids show very similar patterns, thus indicating that the precision with which IRI and IRTAM model *hm*F2 is on the whole comparable. Nevertheless, it must be pointed out that a clear improvement when passing from IRI to IRTAM is observed during post-sunset hours in winter.

Figures 37–39 show that at Thule, for the three considered solar activity levels, a considerable improvement in the *hm*F2 modeling accuracy is achieved by IRTAM during the whole day especially in winter and equinoctial months. Overall, *RMSE* and *NRMSE* grids show that IRTAM improves the *hm*F2 modeling precision, independently of solar activity. In particular, the improvement is appreciated during the whole day in winter and during daytime hours in summer.

From a visual examination of *hm*F2 grids achieved for the three solar activity ranges, some generic conclusions on how the IRI and IRTAM performances depend on solar activity can be drawn: no clear *Res. Mean* trend related to the solar activity for both IRI and IRTAM is found; neither IRI nor IRTAM show a clear *RMSE* trend related to the solar activity; for both models, a decreasing *NRMSE* trend is observed as the solar activity increases.

8. Validation Results for *hm*F2 Based on Radio Occultation Observations

The full COSMIC dataset, comprising 1,791,602 electron density profiles measured from 2006 to 2018, was used to retrieve reliable COSMIC *hm*F2 values for calculating the two-dimensional density plots of IRI-*hm*F2 vs. COSMIC-*hm*F2 and IRTAM-*hm*F2 vs. COSMIC-*hm*F2, along with the histograms of corresponding residuals, as shown in Figure 40. The figure shows that IRTAM slightly improves the *hm*F2 modeling accuracy because *Res. Mean*$_\text{IRTAM}$ = 0.140 km is smaller than *Res. Mean*$_\text{IRI}$ = 3.728 km.
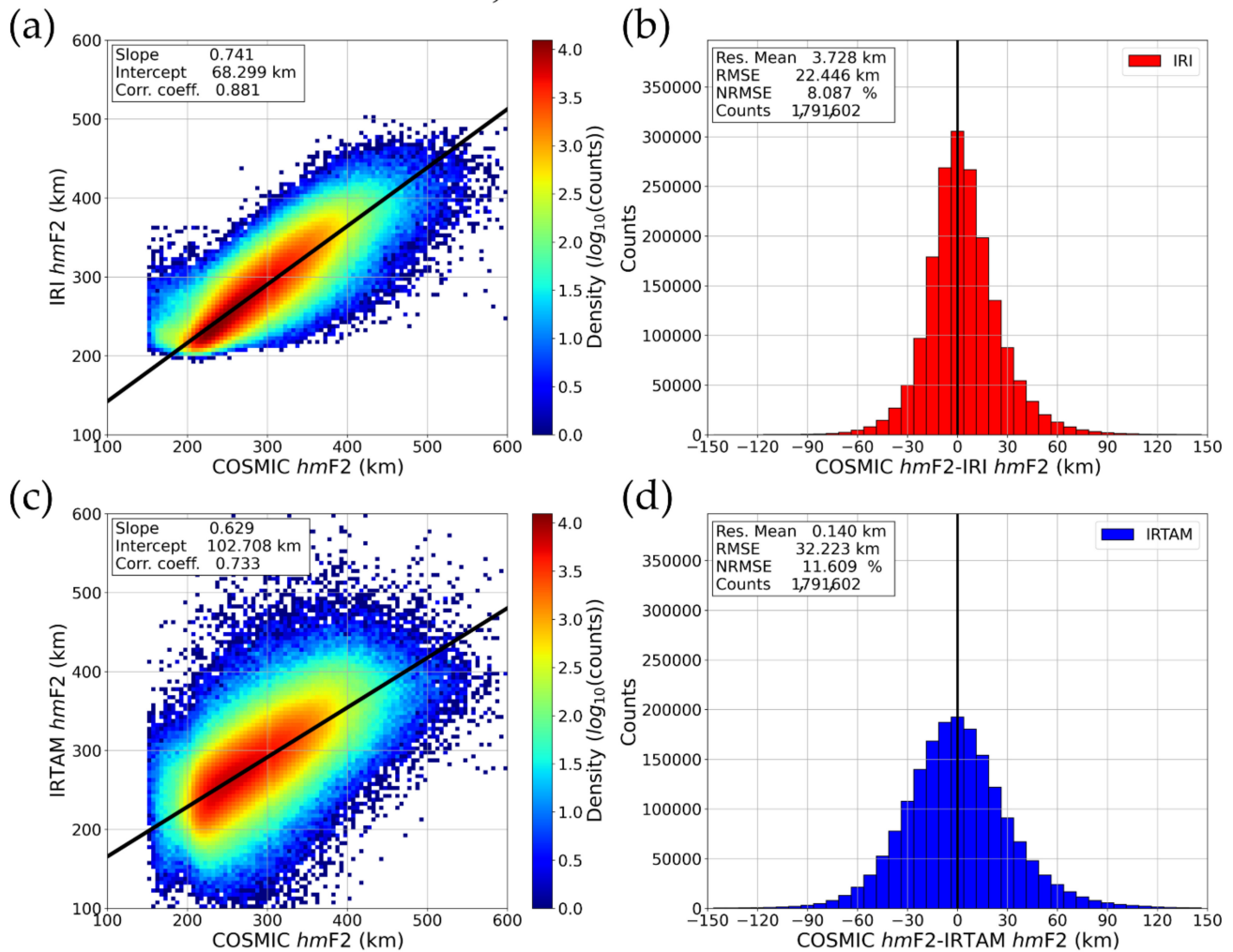
Figure 40. Density plots between measured (*x*-axis) and modeled (*y*-axis) *hm*F2 values, for (**a**) IRI and (**c**) IRTAM. Measured values are those from the entire COSMIC dataset. The black solid lines represent the best linear fit. Corresponding slope and intercept values, along with the correlation coefficient, are reported in the upper left box of each plot. Statistical distributions of the resultant residuals for (**b**) IRI and (**d**) IRTAM. Corresponding *Res. Mean*, *RMSE*, and *NRMSE* values, along with the total number of counts on which the statistics is based, are reported in the upper left box of each plot.

The same cannot be said for the *hm*F2 modeling precision. In fact, the comparisons $RMSE_{IRTAM}$ = 32.223 km vs. $RMSE_{IRI}$ = 22.446 km and $NRMSE_{IRTAM}$ = 11.609% vs. $NRMSE_{IRI}$ = 8.087% show that the IRTAM absolute and percentage errors are by far higher than those of IRI. Moreover, the IRTAM density plot appears much more scattered than that of IRI (R_{IRTAM} = 0.733 vs. R_{IRI} = 0.881).

The statistical results reported in Table 7 show that in some cases *Res. Mean*$_{IRTAM}$ < *Res. Mean*$_{IRI}$, while in many cases $RMSE_{IRTAM} > RMSE_{IRI}$, $NRMSE_{IRTAM} > NRMSE_{IRI}$, and $R_{IRTAM} < R_{IRI}$, which means that the same considerations made for the full COSMIC *hm*F2 dataset hold also when IRI and IRTAM performances are evaluated according to the procedures described in Section 4.2, considering the three levels of solar activity defined in Section 3.1.

Table 7. *Res. Mean*, *RMSE*, *NRMSE*, and *R* values calculated for the IRI and IRTAM models on the basis of the *hm*F2 COSMIC observations, for the bins defined in Section 4.2, according to the three levels of solar activity defined in Section 3.1, and for the full dataset (bottom row). The number of counts on which the statistics were calculated is reported in the rightmost column.

| COSMIC Dataset	Model	Res. Mean [km]	RMSE [km]	NRMSE [%]	R	Counts
Daytime	IRI	2.766	20.065	7.498	0.898	998,544
	IRTAM	−0.502	31.423	11.742	0.734	
Nighttime	IRI	6.659	28.618	9.352	0.775	422,509
	IRTAM	2.581	36.554	11.945	0.588	
Solar terminator	IRI	2.980	20.424	7.509	0.863	370,549
	IRTAM	−0.915	28.912	10.630	0.699	
March Equinox	IRI	4.827	23.133	8.296	0.871	454,118
	IRTAM	0.621	30.988	11.114	0.748	
June Solstice	IRI	1.623	21.718	7.928	0.885	439,417
	IRTAM	4.796	32.985	12.041	0.730	
September Equinox	IRI	2.934	21.552	7.864	0.884	447,872
	IRTAM	−2.540	30.864	11.261	0.743	
December Solstice	IRI	5.465	23.296	8.222	0.882	450,195
	IRTAM	−2.225	33.969	11.990	0.713	
LSA	IRI	1.291	19.834	7.677	0.863	750,993
	IRTAM	−4.387	31.950	12.367	0.655	
MSA	IRI	5.375	23.349	8.255	0.859	608,461
	IRTAM	0.503	31.365	11.088	0.719	
HSA	IRI	5.645	25.248	8.318	0.854	432,148
	IRTAM	7.495	33.846	11.151	0.721	
Quiet magnetic activity	IRI	2.749	21.283	7.699	0.890	1,660,108
	IRTAM	−0.432	31.686	11.462	0.737	
Moderate magnetic activity	IRI	15.854	33.382	11.454	0.808	129,917
	IRTAM	7.206	38.091	13.069	0.681	
Disturbed magnetic activity	IRI	35.519	61.099	19.366	0.571	1577
	IRTAM	19.862	56.750	17.987	0.543	
Low Modip	IRI	5.568	30.485	9.539	0.805	267,449
	IRTAM	11.235	41.462	12.973	0.648	
Mid Modip	IRI	3.666	21.098	7.746	0.874	1,230,835
	IRTAM	−0.829	30.338	11.139	0.722	
High Modip	IRI	2.313	19.031	7.289	0.856	293,318
	IRTAM	−5.912	30.208	11.570	0.627	
Full dataset	IRI	3.728	22.446	8.087	0.881	1,791,602
	IRTAM	0.140	32.223	11.609	0.733	

Analogously to what was carried out in Section 6 for *fo*F2, a global picture of the IRI and IRTAM behavior relative to the COSMIC *hm*F2 dataset is shown in Figures 41–44.
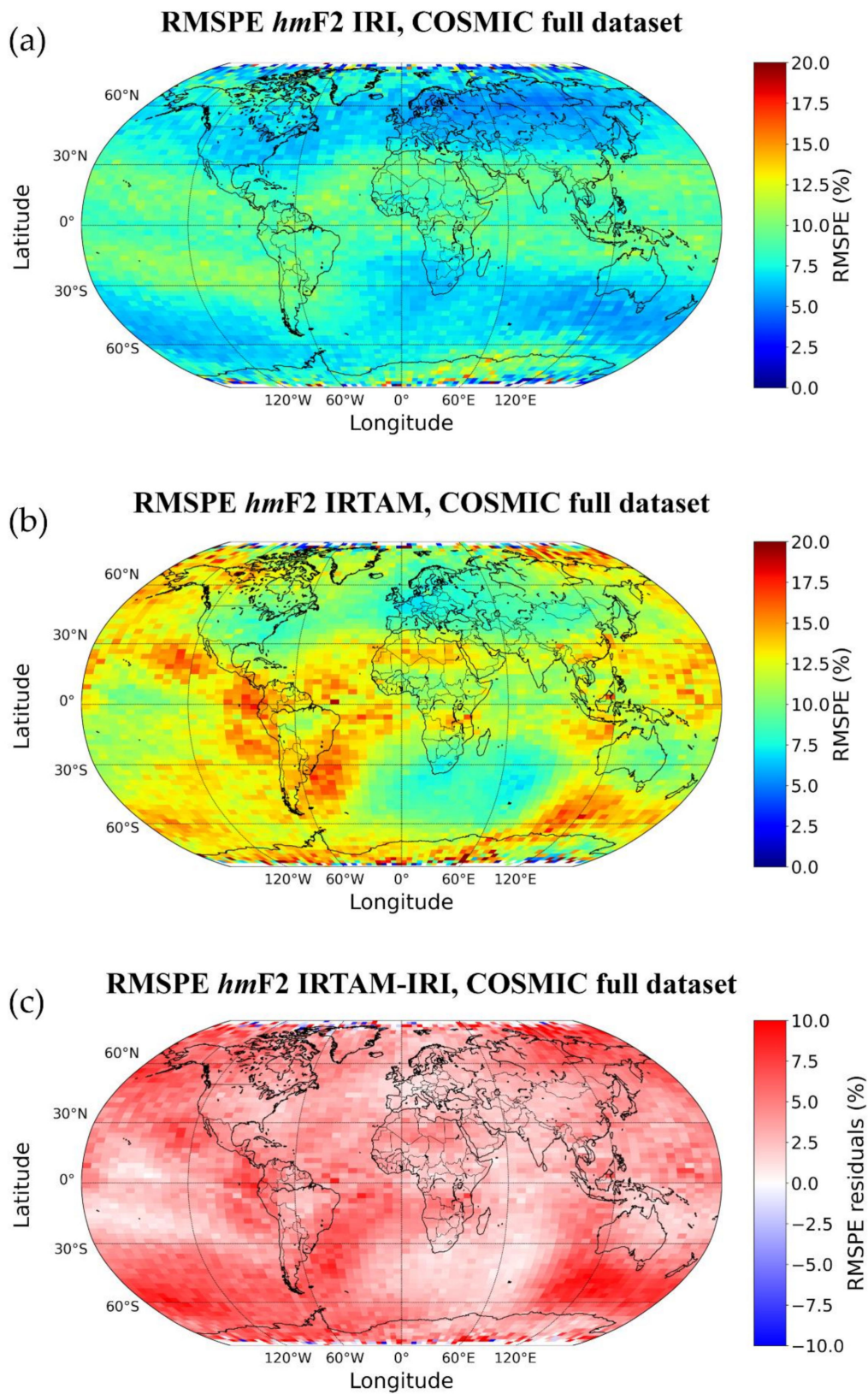
Figure 41. Maps of binned *RMSPE* values calculated between *hm*F2 values measured by COSMIC satellites and modeled by (**a**) IRI and (**b**) IRTAM, and (**c**) the corresponding map of differences. Data were binned as a function of the geographic latitude (2.5°-wide bins) and longitude (5°-wide bins).
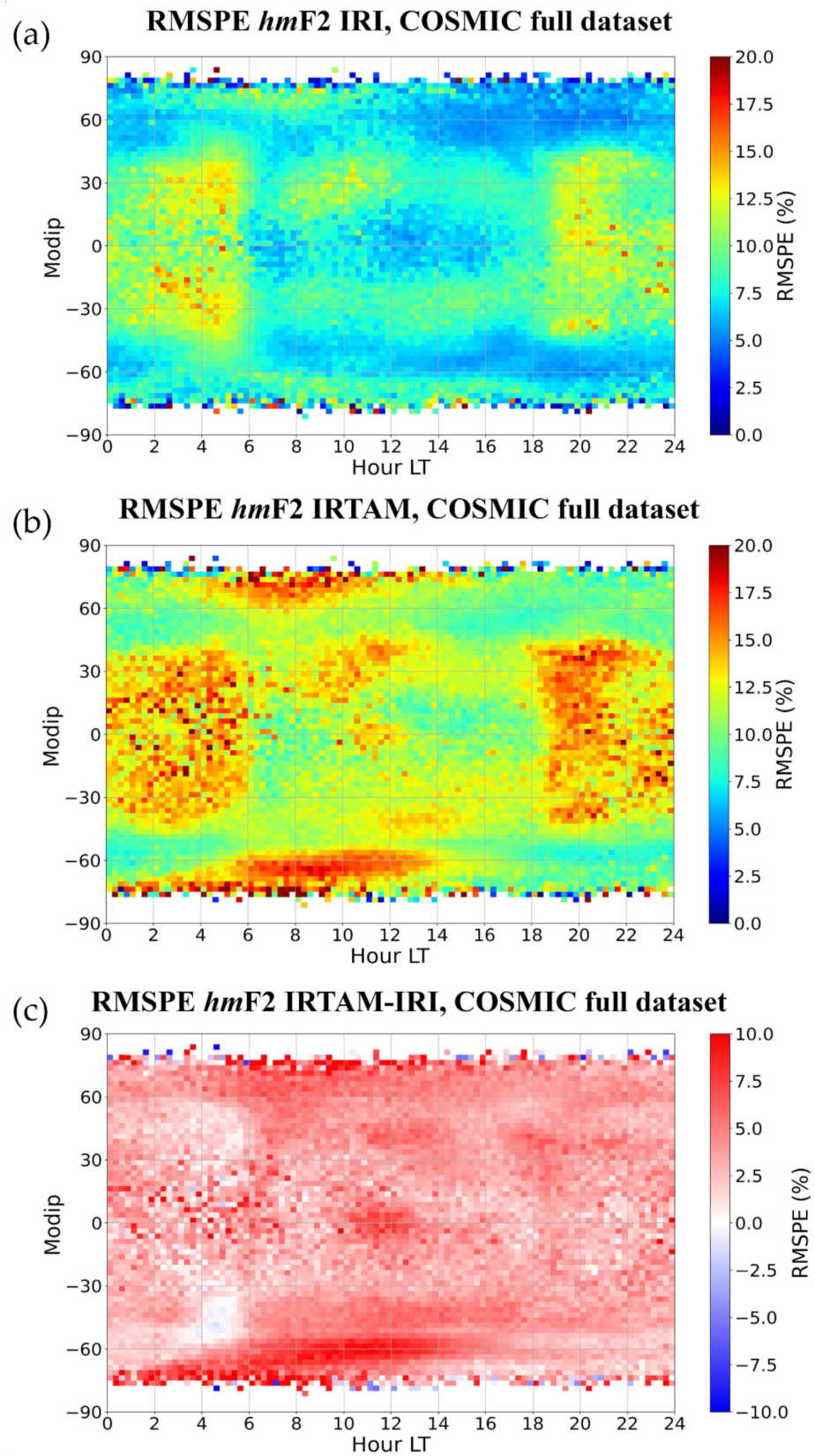
Figure 42. Grids of binned *RMSPE* values calculated between *hm*F2 values measured by COSMIC satellites and modeled by (**a**) IRI and (**b**) IRTAM, and (**c**) the corresponding grid of differences. Data were binned as a function of the LT (fifteen minute-wide bins) and modip (2.5°-wide bins).
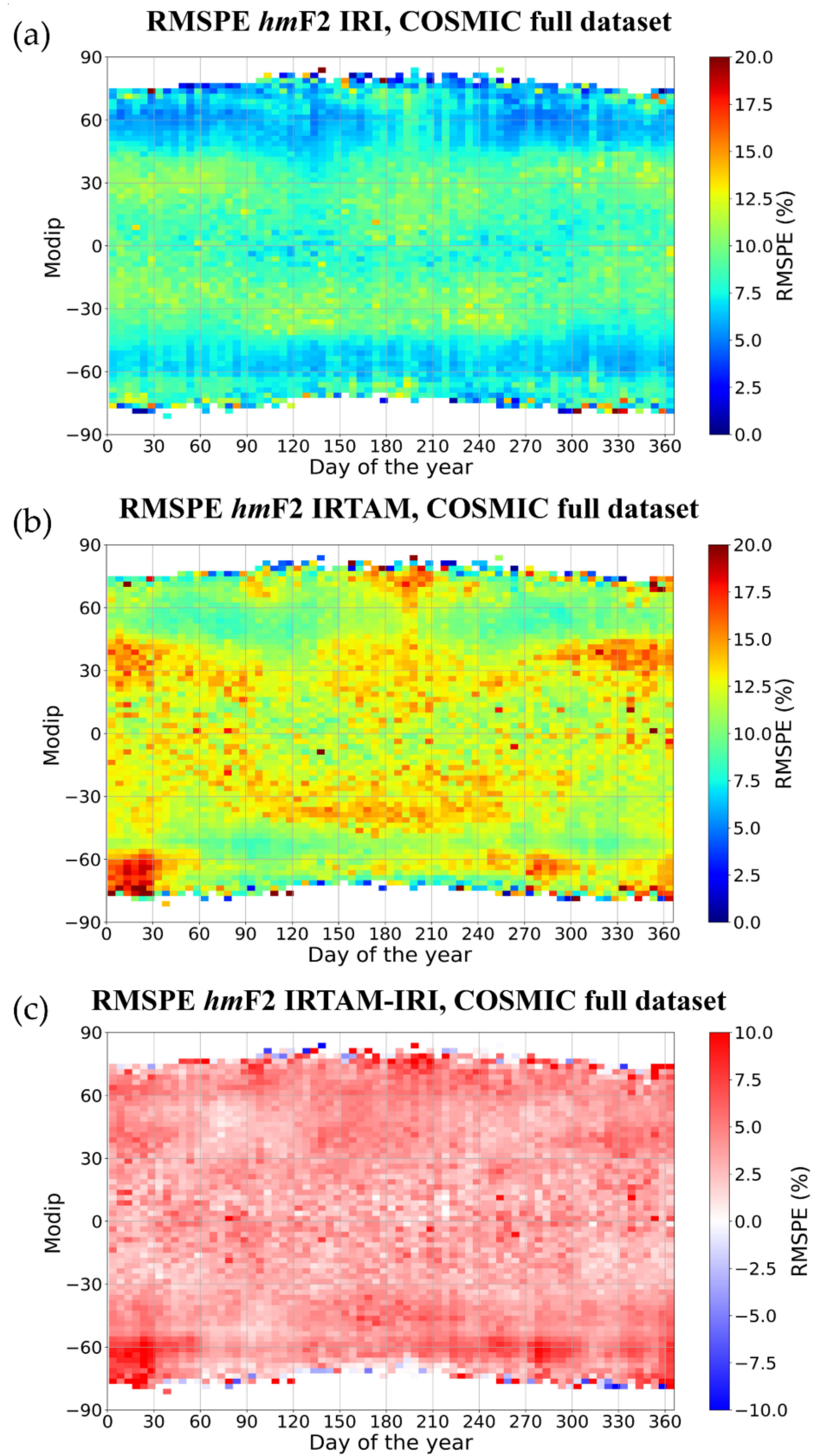
Figure 43. Same as Figure 42 but in this case data were binned as a function of the day of the year (fiveday wide bins) and modip (2.5°-wide bins).
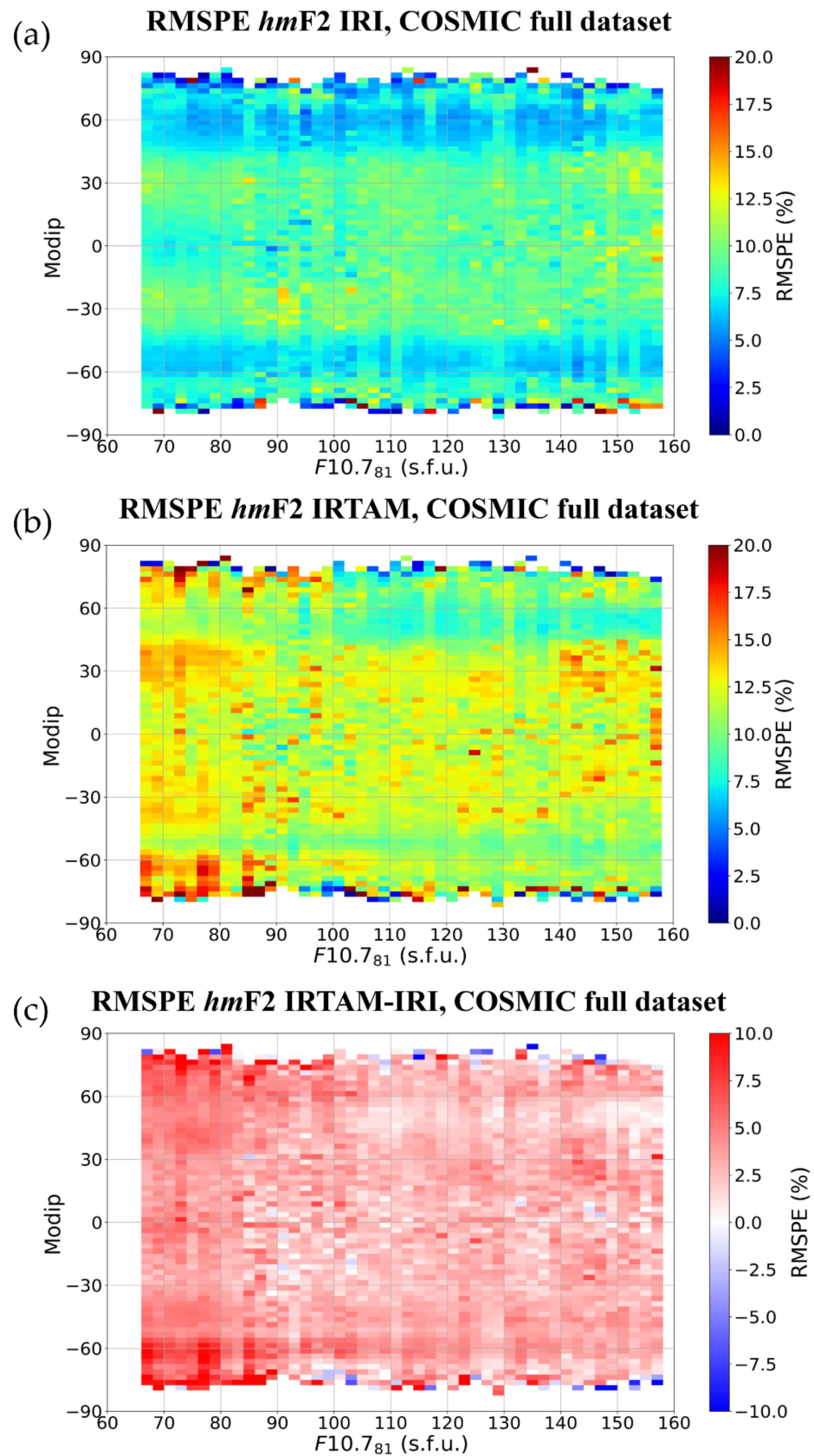
Figure 44. Same as Figure 42 but in this case data were binned as a function of the F10.7$_{81}$ solar activity index (2 sfu-wide bins) and modip (2.5°-wide bins).

Figure 40 shows that, differently from *fo*F2, the values of $RMSE_{IRI}$ and $NRMSE_{IRI}$ calculated for *hm*F2 are relatively smaller than those of $RMSE_{IRTAM}$ and $NRMSE_{IRTAM}$. This is reflected also in the distribution of residuals that in the case of IRTAM appears more widespread around the zero than that of IRI, and this has a strong impact on the maps shown in Figures 41–44. In fact, these maps clearly highlight how IRTAM, independently of the analysis (spatial, diurnal, seasonal, and dependent on solar activity), does not improve the *hm*F2 modeling made by IRI. This result might be affected to a some extent by the fact that IRI *hm*F2 values have been calculated through the Shubin et al. [44] model, which is partly based on electron density profiles collected by COSMIC between 2006 and 2012.

9. Final Analyses and Comparisons between IRI and IRTAM

The results shown in the previous sections are here summarized through the residual deviation ratio R_{cw} defined in Section 4. R_{cw} is a statistical parameter that in general is very suitable to assess definitively the performance of one model over another. To this end, the IRI and IRTAM performances are evaluated analyzing the distributions (in a logarithmic scale) of the residuals' deviation ratio calculated on both the full *fo*F2 ionosonde/COSMIC dataset (Figure 45) and the full *hm*F2 ionosonde/COSMIC dataset (Figure 46).
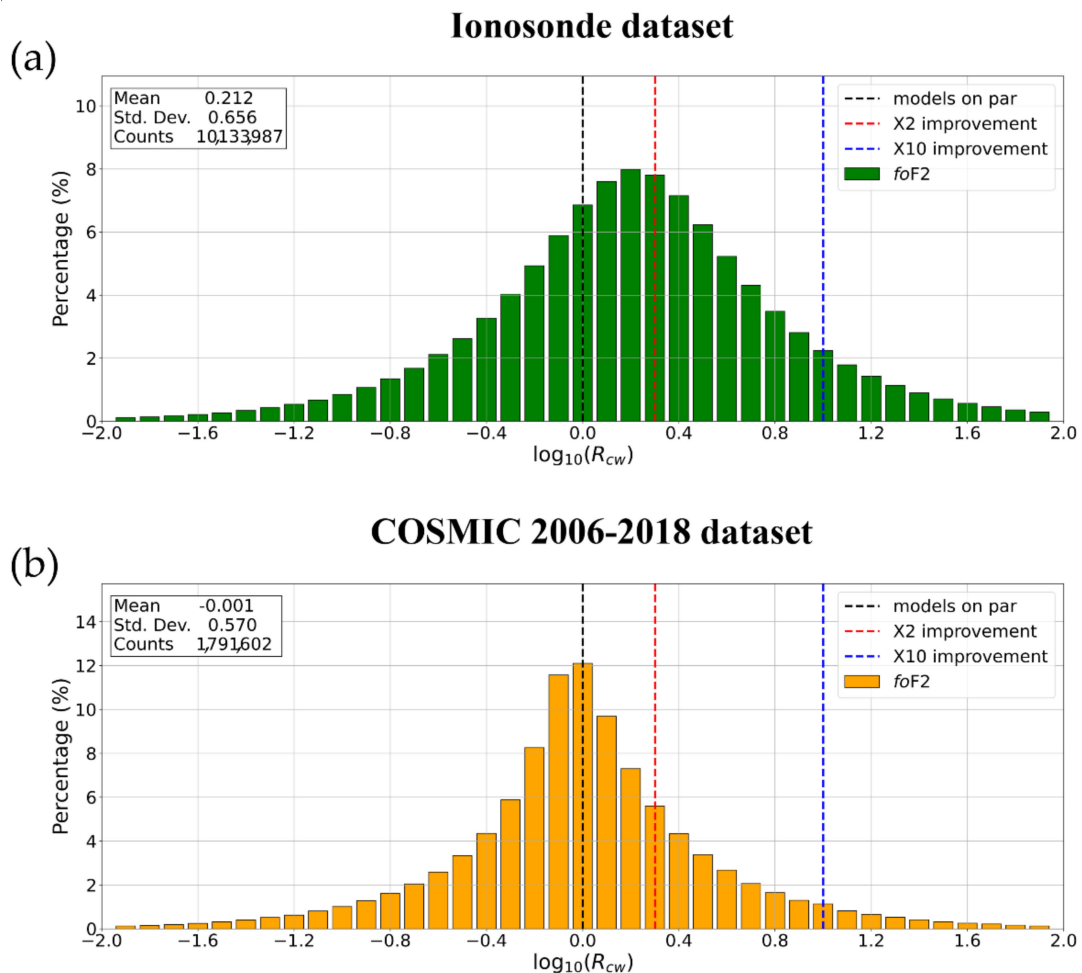
Figure 45. Probability distributions of the residuals' deviation ratio in a logarithmic scale, $\log_{10}(R_{cw})$, between IRI and IRTAM models calculated on the (**a**) entire ionosonde and (**b**) COSMIC *fo*F2 datasets. The mean, standard deviation, and counts values are reported in the upper left corner of each plot. The dashed vertical lines indicate, respectively, when the models are on par (in black), IRTAM improves IRI by a factor of 2 (in red), and IRTAM improves IRI by a factor of 10 (in blue).

Ionosonde dataset

(a)

COSMIC 2006-2018 dataset

(b)
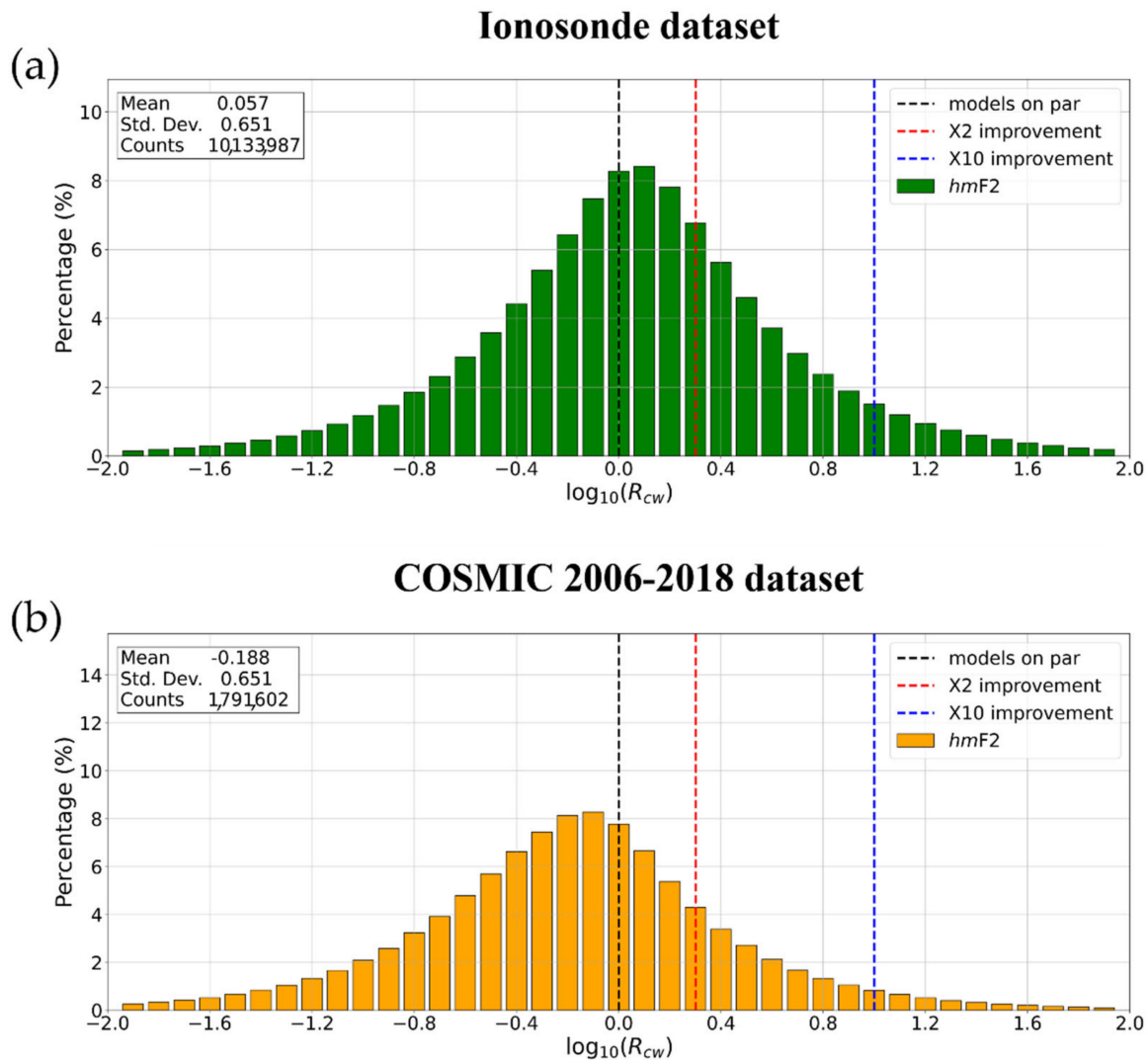
Figure 46. Same as Figure 45 but for *hm*F2.

The $\log_{10}(R_{cw})$ distribution shown in Figure 45a is clearly "shifted" towards the positive values. Specifically, the mean value of the distribution equal to +0.212 highlights that overall IRTAM performs better than IRI by a factor of about 1.6 when considering the *fo*F2 ionosonde dataset. Instead, the $\log_{10}(R_{cw})$ distribution shown in Figure 45b is quite symmetric with respect to the zero value. This fact is strongly supported by the mean value of the distribution, which is equal to −0.001. Therefore, the IRI and IRTAM performances can be considered equivalent when considering the COSMIC *fo*F2 dataset.

The $\log_{10}(R_{cw})$ distribution shown in Figure 46a is quite symmetric around the zero. Specifically, the mean value of the distribution equal to +0.057 (which corresponds to an improvement factor of about 1.14) highlights that IRTAM and IRI provide quite comparable outputs when considering the *hm*F2 ionosonde dataset. The $\log_{10}(R_{cw})$ distribution shown in Figure 46b is clearly "shifted" towards the negative values. Specifically, the mean value of the distribution equal to −0.188 points out that IRI performs better than IRTAM when considering the *hm*F2 COSMIC dataset by a factor of about 1.5.

The same analysis based on $\log_{10}(R_{cw})$ distribution was applied by Galkin et al. [18] on a dataset of *fo*F2 values recorded by 59 ionosondes during May–June 2019. They found results very similar to the ones shown in Figure 45a, with IRTAM improving IRI by a factor of about two (about 0.3 in the logarithmic scale of Figure 45). The slight differences are due on the one hand to the fact that to test IRTAM Galkin et al. [18] used only data from assimilated stations, while in this study we used also non-assimilated stations, and on

the other hand to the larger extension of our dataset covering different seasons and solar activity levels. Results similar to those of Galkin et al. [18] were obtained also by Vesnin [29] by considering a larger dataset covering one solar cycle, but again using only data from assimilated stations to test the model. Vesnin [29] investigated the IRTAM performance also for hmF2 and found that IRTAM improved IRI by a factor of about 1.8. However, in that analysis the oldest Bilitza et al. [41] hmF2 IRI option was used as comparison. When using the newest Shubin et al. [44] default IRI hmF2 option, we find much lower differences between IRTAM and IRI, thus confirming the very important step forward made by IRI about the hmF2 modeling, as on the other hand recently outlined by different authors [69–71].

Figures 45 and 46 confirm the general picture outlined by the analyses described in Sections 5–8, i.e., the comparison with ionosonde data highlights how IRTAM significantly improves the foF2 prediction made by IRI, while for hmF2 the performances are quite similar between the two models. Since IRTAM assimilates both foF2 and hmF2 from the GIRO network, we would have expected a similar improvement also in the hmF2 prediction. Besides the obvious differences due to the application of the newest Shubin et al. [44] IRI hmF2 default option, two important points need to be highlighted. First, the IRTAM hmF2 description is based on the mapping procedure introduced by Brunini et al. [53], which introduces residuals in the range from -10 to 10 km when compared to the original hmF2 values obtained from the Bilitza et al. [41] formulation. The second point is inherent to the hmF2 derivation from ionograms. In fact, while foF2 is a parameter that is directly obtained from ionograms as the maximum ordinary frequency reflected by the ionosphere, hmF2 has to be derived trough a mathematical inversion procedure that, starting from critical frequencies measured at different virtual heights, allows obtaining the electron density values at real heights [72]. This inversion procedure is sensitive to different possible error sources due to the E-valley presence, the interpretation of the F2-layer cusp made by ARTIST, and in general the quality of the ionogram echo traces. All of these matters may represent possible sources of error, that are estimated in the order of 10 km [73]. From the above considerations, it clearly emerges that to obtain reliable hmF2 values is more difficult than to get foF2 ones, even with data assimilation. This is a point that requires further improvements and refining of both the data assimilation procedure and the measurement technique itself.

The comparison with the F2-layer peak characteristics derived from COSMIC RO showed a general worsening of the IRTAM performances in comparison with the IRI ones. Specifically, the comparison with the COSMIC dataset (Figures 20 and 41) highlighted that IRTAM improves the IRI foF2 and hmF2 predictions mainly in regions characterized by a dense ionosonde network. This suggests the extent to which IRTAM is tied to data assimilated by ionosondes and to the corresponding spatial distribution. Since assimilated data are used by IRTAM to update the coefficients of the spherical harmonic analysis underlying the IRI description, we would have expected that the improvements were not restricted to the ionosondes' locations but would embrace at least the whole latitudinal sector where assimilated ionosondes are located. Moreover, the IRTAM description should fade towards that of the IRI in regions where the effect of the assimilated data can be considered negligible. These two points are very important, impacting on IRTAM global performances, and need to be deeply investigated for future versions of IRTAM. Currently, IRTAM assimilates data from about 60 GIRO Digisondes. With the ever-increasing number of available Digisondes, able to provide real-time data, and as a consequence of their more homogeneous spatial distribution, a continuous improvement of the IRTAM performance is expected. However, even if in the future the availability of ionosonde data should increase for both the time resolution and the spatial coverage, the fact that IRTAM is so tied to the underlying IRI model (i.e., to the URSI formalism) represents a limit for the improvement and development of IRTAM itself. In fact, IRTAM through NECTAR, to minimize at assimilation sites the mismatch between measured and IRI modeled values, calculates the corrections to be added to the URSI coefficients, but the order of the diurnal

and spatial harmonics is left unchanged. This means that steep spatial gradients and fast time variations that are below the limits that can be resolved with the current spatial and temporal resolution of the URSI formalism would not be represented by IRTAM, even with an increased availability of assimilated data. Since steep spatial gradients and fast time variations are customary under specific Space Weather conditions, and the aim of data-assimilation methods is the reliable representation of such conditions, this poses serious limitations for the IRTAM model that its developers should bear in mind in the future.

10. Conclusions

In the present paper, we compared the IRI and IRTAM models; the latter, being a real-time version of IRI, is based on the assimilation of ionosonde measurements. In order to assess the performance of the two models, two different datasets have been considered: (1) foF2 and hmF2 from ground-based ionosonde observations; (2) foF2 and hmF2 from space-based COSMIC RO observations. Through different analyses and comparison methodologies, we highlighted the main performances exhibited by both IRI and IRTAM for different locations and under different diurnal, seasonal, solar and magnetic activity conditions.

The main results of the study are:

- When ionosonde observations are considered for validation, IRTAM improves significantly the IRI foF2 modeling while it slightly improves the IRI hmF2 modeling.
- When COSMIC observations are considered for validation, IRTAM improves neither the IRI foF2 modeling nor the IRI hmF2 modeling.

These results highlight that IRTAM, in contrast to most of assimilation models, has ample room for improvement. The points that in our opinion deserve specific attention are: the bad performance of IRTAM when modeling foF2 at low latitudes; the global hmF2 modeling made by IRTAM which is often unreliable, especially in areas far away from the assimilating sites, where the representation made by IRTAM is at times really different from that of the IRI background; the fact that IRTAM performances are too dependent on the assimilated ionosondes location.

The improvement in the near real-time specification of the ionospheric F2-layer peak characteristics is becoming more and more important nowadays for telecommunication purposes and for Space Weather applications in general. For example, Hartman et al. [74] have recently applied IRTAM as the Floating Potential Measurement Unit (FPMU) back-up system that will be used to support the International Space Station (ISS) program. IRTAM foF2 maps were used by Froń et al. [75] to provide global maps of the ionospheric equivalent slab thickness (τ) parameter that are delivered through the GAMBIT Explorer software (http://giro.uml.edu/GAMBIT, accessed on 30 July 2021). An improved real-time specification of τ on a global basis is very important because this parameter describes the shape of the ionospheric electron density profile; thus, an improved specification of τ can help empirical models such as IRI in the description of the profile shape, especially the topside part [76–79].

Since in the incoming years the applications based on a near real-time specification of the ionospheric conditions will increase in number, an ever more reliable and robust representation of the ionosphere will become of outstanding importance. This is why near real-time data-assimilation models such as IRTAM need continuous improvement and refining, on the one hand to improve the climatological description of the ionosphere made by IRI, and on the other hand to pave the way for a reliable ionospheric weather description.

Author Contributions: Conceptualization, A.P., M.P. (Michael Pezzopane); methodology, A.P.; data curation, A.P.; investigation, all authors; validation, all authors; formal analysis, A.P.; writing—original draft preparation, A.P., M.P. (Marco Pietrella); writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Abbreviations

The following abbreviations are frequently used in this manuscript:

a_P	Planetary 3-h-range a_P magnetic index
C-Score	ARTIST Confidence Score
CCIR	Consultative Committee on International Radio
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate/FORMOSAT-3 satellites
doy	Day of the year
F10.7	Daily solar radio flux at 10.7 cm
$F10.7_{81}$	81-day running mean of the daily F10.7
*fo*F2	Ordinary critical frequency of the F2-layer
GAMBIT	Global Assimilative Model of Bottomside Ionosphere Timeline
GIRO	Global Ionospheric Radio Observatory
GPS	Global Positioning System
*hm*F2	Height of the F2-layer peak
HSA	High solar activity
IRI	International Reference Ionosphere
IRTAM	IRI Real-Time Assimilative Mapping

LSA	Low solar activity
LT	Local Time
M(3000)F2	Ionospheric propagation factor
Modip	Modified dip latitude
MSA	Mid solar activity
NECTAR	Non-linear Error Compensation Technique for Associative Restoration
NRMSE	Normalized root mean square error
QD	Quasi Dipole
R	Pearson correlation coefficient
R_{cw}	Residuals deviation ratio parameter
Res. Mean	Mean of residuals
RMSE	Root mean square error
RMSPE	Root mean square percentage error
RO	Radio occultation
SZA	Solar zenith angle
vTEC	vertical Total Electron Content
URSI	International Union of Radio Science
UT	Universal Time

References

1. Moldwin, M. *An Introduction To Space Weather*; Cambridge University Press: Cambridge, MA, USA, 2008. [CrossRef]
2. Cander, L.R. *Ionospheric Space Weather*; Springer Nature: Cham, Switzerland, 2019. [CrossRef]
3. Schunk, R.W.; Scherliess, L.; Sojka, J.J.; Thompson, D.C.; Anderson, D.N.; Codrescu, M.; Minter, C.; Fuller-Rowell, T.J.; Heelis, R.A.; Hairston, M.; et al. Global assimilation of ionospheric measurements (GAIM). *Radio Sci.* **2004**, *39*. [CrossRef]
4. Angling, M.J.; Khattatov, B. Comparative study of two assimilative models of the ionosphere. *Radio Sci.* **2006**, *41*. [CrossRef]
5. Decker, D.T.; McNamara, L.F. Validation of ionospheric weather predicted by global assimilation of ionospheric measurements (GAIM) models. *Radio Sci.* **2007**, *42*, RS4017. [CrossRef]
6. McNamara, L.F.; Decker, D.T.; Welsh, J.A.; Cole, D.G. 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.* **2007**, *42*, RS3015. [CrossRef]
7. McNamara, L.F.; Bishop, G.J.; Welsh, J.A. Assimilation of ionosonde profiles into a global ionospheric model. *Radio Sci.* **2011**, *46*, RS2006. [CrossRef]
8. Buresova, D.; Nava, B.; Galkin, I.; Angling, M.; Stankov, S.M.; Coisson, P. Data ingestion and assimilation in ionospheric models. *Ann. Geophys.* **2009**, *52*, 235–253. [CrossRef]
9. Nava, B.; Coisson, P.; Radicella, S.M. A new version of the NeQuick ionosphere electron density model. *J. Atmos. Sol. Terr. Phys.* **2008**, *70*, 1856–1862. [CrossRef]
10. Nava, B.; Radicella, S.M.; Azpilicueta, F. Data ingestion into NeQuick 2. *Radio Sci.* **2011**, *46*. [CrossRef]
11. Pezzopane, M.; Pietrella, M.; Pignatelli, A.; Zolesi, B.; Cander, L.R. Assimilation of autoscaled data and regional and local ionospheric models as input sources for real-time 3-D international reference ionosphere modeling. *Radio Sci.* **2011**, *46*, 5009. [CrossRef]
12. Pezzopane, M.; Pietrella, M.; Pignatelli, A.; Zolesi, B.; Cander, L.R. Testing the three-dimensional IRI-SIRMUP-P mapping of the ionosphere for disturbed periods. *Adv. Space Res.* **2013**, *52*, 1726–1736. [CrossRef]
13. Shim, J.S.; Kuznetsova, M.; Rastatter, L.; Hesse, M.; Bilitza, D.; Butala, M.; Codrescu, M.; Emery, B.; Foster, B.; Fuller-Rowell, T.; et al. CEDAR electrodynamics thermosphere ionosphere (ETI) challenge for systematic assessment of ionosphere/thermosphere models: NmF2, hmF2, and vertical drift using ground-based observations. *Space Weather* **2011**, *9*, S12003. [CrossRef]
14. Pignalberi, A.; Pezzopane, M.; Rizzi, R.; Galkin, I. Effective solar indices for ionospheric modeling: A review and a proposal for a real-time regional IRI. *Surv. Geophys.* **2018**, *39*, 125–167. [CrossRef]
15. Pietrella, M.; Pignalberi, A.; Pezzopane, M.; Pignatelli, A.; Azzarone, A.; Rizzi, R. A comparative study of ionospheric IRIEup and ISP assimilative models during some intense and severe geomagnetic storms. *Adv. Space Res.* **2018**, *61*, 2569–2584. [CrossRef]
16. Pietrella, M.; Pezzopane, M.; Zolesi, B.; Cander, L.R.; Pignalberi, A. The simplified ionospheric regional model (SIRM) for HF prediction: Basic theory, its evolution and applications. *Surv. Geophys.* **2020**, *41*, 1143–1178. [CrossRef]
17. Galkin, I.A.; Reinisch, B.W.; Huang, X.; Bilitza, D. Assimilation of GIRO data into a real-time IRI. *Radio Sci.* **2012**, *47*. [CrossRef]
18. Galkin, I.A.; Reinisch, B.W.; Vesnin, A.M.; Bilitza, D.; Fridman, S.; Habarulema, J.B.; Veliz, O. Assimilation of sparse continuous near-Earth weather measurements by NECTAR model morphing. *Space Weather* **2020**, *18*, e2020SW002463. [CrossRef]
19. Bilitza, D.; Altadill, D.; Truhlik, V.; Shubin, V.; Galkin, I.; Reinisch, B.; Huang, X. International reference ionosphere 2016: From ionospheric climate to real-time weather predictions. *Space Weather* **2017**, *15*, 418–429. [CrossRef]
20. Bilitza, D. IRI the international standard for the ionosphere. *Adv. Radio Sci.* **2018**, *16*, 1–11. [CrossRef]

21. Lei, J.; Syndergaard, S.; Burns, A.G.; Solomon, S.C.; Wang, W.; Zeng, Z.; Roble, R.G.; Wu, Q.; Kuo, Y.-H.; Holt, J.M.; et al. Comparison of COSMIC ionospheric measurements with ground-based observations and model predictions: Preliminary results. *J. Geophys. Res.* **2007**, *112*, A07308. [CrossRef]

22. Damboldt, T.; Suessmann, P. Information document on the analysis and validity of present ITU foF2 and M (3000) f2 maps. *Int. Telecommun. Union* **2011**. question ITU-R 212-1/3. Available online: http://www.itu.int/md/R07-WP3L-C-0086/en (accessed on 30 July 2021).

23. Shim, J.S.; Tsagouri, I.; Goncharenko, L.; Rastaetter, L.; Kuznetsova, M.; Bilitza, D.; Codrescu, M.; Coster, A.J.; Solomon, S.C.; Fedrizzi, M.; et al. Validation of ionospheric specifications during geomagnetic storms: TEC and foF2 during the 2013 March storm event. *Space Weather* **2018**, *16*, 1686–1701. [CrossRef]

24. Shim, J.S.; Kuznetsova, M.; Rastätter, L.; Hesse, M.; Bilitza, D.; Butala, M.; Codrescu, M.; Emery, B.; Foster, B.; Fuller-Rowell, T.; et al. CEDAR electrodynamics thermosphere ionosphere (ETI) challenge for systematic assessment of ionosphere/thermosphere models: Electron density, neutral density, NmF2, and hmF2 using space based observations. *Space Weather* **2012**, *10*, S10004. [CrossRef]

25. Pedatella, N.M.; Yue, X.; Schreiner, W.S. Comparison between GPS radio occultation electron densities and in situ satellite observations. *Radio Sci.* **2015**, *50*, 518–525. [CrossRef]

26. Pignalberi, A.; Pezzopane, M.; Tozzi, R.; De Michelis, P.; Coco, I. Comparison between IRI and preliminary Swarm Langmuir probe measurements during the St. Patrick storm period. *Earth Planets Space* **2016**, *68*. [CrossRef]

27. Tsagouri, I.; Goncharenko, L.; Shim, J.S.; Belehaki, A.; Buresova, D.; Kuznetsova, M.M. Assessment of current capabilities in modeling the ionospheric climatology for space weather applications: foF2 and hmF2. *Space Weather* **2018**, *16*, 1930–1945. [CrossRef]

28. Cai, X.; Burns, A.G.; Wang, W.; Coster, A.; Qian, L.; Liu, J.; Solomon, S.C.; Eastes, R.W.; Daniell, R.E.; McClintock, W.E.; et al. Comparison of GOLD nighttime measurements with total electron content: Preliminary results. *J. Geophys. Res. Space Phys.* **2020**, *125*, e2019JA027767. [CrossRef]

29. Vesnin, A.M. Validation of F2 Layer Peak Height and Density of Real-Time International Reference Ionosphere. Master's Thesis, University of Massachusetts Lowell, Lowell, MA, USA, 2014. Available online: https://ulcar.uml.edu/GAMBIT/Vesnin-Master-thesis-2014.pdf (accessed on 30 July 2021).

30. Zolesi, B.; Cander, L.R. *Ionospheric Prediction and Forecasting*; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]

31. Rush, C.M.; Fox, M.; Bilitza, D.; Davies, K.; McNamara, L.; Stewart, F.G.; PoKempner, M. Ionospheric mapping-an update of foF2 coefficients. *Telecommun. J.* **1989**, *56*, 179–182.

32. Jones, W.B.; Gallet, R.M. Representation of diurnal and geographical variations of ionospheric data by numerical methods. *Telecommun. J.* **1962**, *29*, 129–149. Available online: https://nvlpubs.nist.gov/nistpubs/jres/66D/jresv66Dn4p419_A1b.pdf (accessed on 30 July 2021).

33. Jones, W.B.; Gallet, R.M. Representation of diurnal and geographic variations of ionospheric data by numerical methods, II. Control of instability. *Telecommun. J.* **1965**, *32*, 18–28.

34. Jones, W.B.; Graham, R.P.; Leftin, M. *Advances in Ionospheric Mapping by Numerical Methods*; ESSA Technical Report ERL107-ITS75; US Department of Commerce: Boulder, CO, USA, 1969. Available online: https://www.govinfo.gov/content/pkg/GOVPUB-C13-4811417984235af8236e56a8e5d5d483/pdf/GOVPUB-C13-4811417984235af8236e56a8e5d5d483.pdf (accessed on 30 July 2021).

35. *CCIR Atlas of Ionospheric Characteristics Report 340*; Consultative Committee on International Radio, International Telecommunication Union: Geneva, Switzerland, 1967.

36. Rawer, K. Propagation of decameter waves (HF Band). In *Meteorological and Astronomical Influences on Radio Wave Propagation*; Landmark, B., Ed.; Academic Press: New York, NY, USA, 1963; pp. 221–250.

37. Liu, R.; Smith, P.; King, J. A new solar index which leads to improved foF2 predictions using the CCIR atlas. *Telecomm. J.* **1983**, *50*, 408–414.

38. Fuller-Rowell, T.J.; Araujo-Pradere, E.; Codrescu, M.V. An empirical ionospheric storm-time correction model. *Adv. Space Res.* **2000**, *25*, 139–146. [CrossRef]

39. Araujo-Pradere, E.A.; Fuller-Rowel, T.J.; Codrescu, M.V. STORM: An empirical storm-time ionospheric correction model 1. Model description. *Radio Sci.* **2002**, *37*. [CrossRef]

40. Araujo-Pradere, E.A.; Fuller-Rowell, T.J. STORM: An empirical storm-time ionospheric correction model 2. Validation. *Radio Sci.* **2002**, *37*, 1071. [CrossRef]

41. Bilitza, D.; Sheik, N.; Eyfrig, R. A global model for the height of the F2-peak using M3000 values from the CCIR numerical map. *Telecommun. J.* **1979**, *46*, 549–553.

42. Altadill, D.; Magdaleno, S.; Torta, J.M.; Blanch, E. Global empirical models of the density peak height and of the equivalent scale height for quiet conditions. *Adv. Space Res.* **2013**, *52*, 1756–1769. [CrossRef]

43. Shubin, V.N.; Karpachev, A.T.; Tsybuly, K.G. Global model of the F2 layer peak height for low solar activity based on GPS radio-occultation data. *J. Atmos. Sol. Terr. Phys.* **2013**, *104*, 106–115. [CrossRef]

44. Shubin, V.N. Global median model of the F2-layer peak height based on ionospheric radio-occultation and ground based digisonde observations. *Adv. Space Res.* **2015**, *56*, 916–928. [CrossRef]

45. Bilitza, D.; Bhardwaj, S.; Koblinsky, C. Improved IRI predictions for the GEOSAT time period. *Adv. Space Res.* **1997**, *20*, 1755–1760. [CrossRef]

46. Komjathy, A.; Langley, R.; Bilitza, D. Ingesting GPS-derived TEC data into the international reference ionosphere for single frequency radar altimeter ionospheric delay corrections. *Adv. Space Res.* **1998**, *22*, 793–802. [CrossRef]

47. Hernandez-Pajares, M.; Juan, J.; Sanz, J.; Bilitza, D. Combining GPS measurements and IRI model values for space weather specification. *Adv. Space Res.* **2002**, *29*, 949–958. [CrossRef]

48. Ssessanga, N.; Kim, Y.H.; Kim, E.; Kim, J. Regional optimization of the IRI-2012 output (TEC, foF2) by using derived GPS-TEC. *J. Korean Phys. Soc.* **2015**, *66*, 1599–1610. [CrossRef]

49. Habarulema, J.B.; Ssessanga, N. Adapting a climatology model to improve estimation of ionosphere parameters and subsequent validation with radio occultation and ionosonde data. *Space Weather* **2016**, *15*, 84–98. [CrossRef]

50. Pignalberi, A.; Pezzopane, M.; Rizzi, R.; Galkin, I. Correction to: Effective solar indices for ionospheric modeling: A review and a proposal for a real-time regional IRI. *Surv. Geophys.* **2018**, *39*, 169. [CrossRef]

51. Pignalberi, A.; Pietrella, M.; Pezzopane, M.; Rizzi, R. Improvements and validation of the IRI UP method under moderate, strong, and severe geomagnetic storms. *Earth Planets Space* **2018**, *70*, 1–22. [CrossRef]

52. Pignalberi, A.; Habarulema, J.B.; Pezzopane, M.; Rizzi, R. On the development of a method for updating an empirical climatological ionospheric model by means of assimilated vTEC measurements from a GNSS receiver network. *Space Weather* **2019**, *17*, 1131–1164. [CrossRef]

53. Brunini, C.; Conte, J.F.; Azpilicueta, F.; Bilitza, D. A different method to update monthly median hmF2 values. *Adv. Space Res.* **2013**, *51*, 2322–2332. [CrossRef]

54. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 2554–2558. [CrossRef]

55. Laundal, K.M.; Richmond, A.D. Magnetic Coordinate Systems. *Space Sci. Rev.* **2017**, *206*, 27. [CrossRef]

56. Reinisch, B.W.; Galkin, I.A. Global Ionospheric Radio Observatory (GIRO). *Earth Planets Space* **2011**, *63*, 377–381. [CrossRef]

57. Bibl, K.; Reinisch, B.W. The universal digital ionosonde. *Radio Sci.* **1978**, *13*, 519–530. [CrossRef]

58. Galkin, I.A.; Reinisch, B.W. The new ARTIST 5 for all digisondes. In *Ionosonde Network Advisory Group (INAG) Bulletin*, 69th ed.; International Radio Science Union: Ghent, Belgium, 2008; Available online: http://www.ursi.org/files/CommissionWebsites/INAG/web-69/2008/artist5-inag.pdf (accessed on 30 July 2021).

59. Galkin, I.A.; Reinisch, B.W.; Huang, X.; Khmyrov, G.M. Confidence score of ARTIST-5 ionogram autoscaling. In *Ionosonde Network Advisory Group (INAG) Bulletin*, 73rd ed.; International Radio Science Union: Ghent, Belgium, 2013; Available online: http://www.ursi.org/files/CommissionWebsites/INAG/web-73/confidence_score.pdf (accessed on 30 July 2021).

60. Tapping, K.F. The 10.7 cm solar radio flux (F10.7): F10.7. *Space Weather* **2013**, *11*, 394–406. [CrossRef]

61. Anthes, R.A.; Bernhardt, P.A.; Chen, Y.; Cucurull, L.; Dymond, K.F.; Ector, D.; Healy, S.B.; Ho, S.-P.; Hunt, D.C.; Kuo, Y.-H.; et al. The COSMIC/FORMOSAT-3 mission: Early results. *Bull. Am. Meteorol. Soc.* **2008**, *89*, 313–333. [CrossRef]

62. Pignalberi, A.; Pezzopane, M.; Nava, B.; Coisson, P. On the link between the topside ionospheric effective scale height and the plasma ambipolar diffusion, theory and preliminary results. *Sci. Rep.* **2020**, *10*, 17541. [CrossRef]

63. Rostoker, G. Geomagnetic indices. *Rev. Geophys. Space Phys.* **1972**, *10*, 935–950. [CrossRef]

64. Solomon, S.C.; Woods, T.N.; Didkovsky, L.V.; Emmert, J.T.; Qian, L. Anomalously low solar extreme-ultraviolet irradiance and thermospheric density during solar minimum. *Geophys. Res. Lett.* **2010**, *37*, 16. [CrossRef]

65. Perna, L.; Pezzopane, M. foF2 vs solar indices for the Rome station: Looking for the best general relation which is able to describe the anomalous minimum between cycles 23 and 24. *J. Atmos. Space Phys.* **2016**, *146*, 13–21. [CrossRef]

66. Lühr, H.; Xiong, C. The IRI 2007 model overestimates electron density during the 23/24 solar minimum. *Geophys. Res. Lett.* **2010**, *37*, L23101. [CrossRef]

67. Ezquer, R.G.; López, J.L.; Scidá, L.A.; Cabrera, M.A.; Zolesi, B.; Bianchi, C.; Pezzopane, M.; Zuccheretti, E.; Mosert, M. Behavior of ionospheric magnitudes of F2 region over Tucumán during a deep solar minimum and comparison with the IRI2012 model predictions. *J. Atmos. Sol. Terr. Phys.* **2014**, *107*, 89–98. [CrossRef]

68. Perna, L.; Pezzopane, M.; Ezquer, R.; Cabrera, M.; Baskaradas, J.A. NmF2 trends at low and mid latitudes for the recent solar minima and comparison with IRI-2012 model. *Adv. Space Res.* **2017**, *60*, 363–374. [CrossRef]

69. Arıkan, F.; Sezen, U.; Gulyaeva, T.L. Comparison of IRI-2016 F2 layer model parameters with ionosonde measurements. *J. Geophys. Res. Space Phys.* **2019**, *124*, 8092–8109. [CrossRef]

70. Mengist, C.K.; Yadav, S.; Kotulak, K.; Bahar, A.; Zhang, S.-R.; Seo, K.-H. Validation of International Reference ionosphere model (IRI-2016) for F-region peak electron density height (hmF2): Comparison with Incoherent Scatter Radar (ISR) and ionosonde measurements at Millstone Hill. *Adv. Space Res.* **2020**, *65*, 2773–2781. [CrossRef]

71. Huang, H.; Moses, M.; Volk, A.E.; Abu Elezz, O.; Kassamba, A.A.; Bilitza, D. Assessment of IRI-2016 hmF2 model options with digisonde, COSMIC and ISR observations for low and high solar flux conditions. *Adv. Space Res.* **2021**. [CrossRef]

72. Titheridge, J.E. *Ionogram Analysis with the Generalised Program Polan, Rep. UAG-93*; World Data Center A for Solar-Terrestrial Physics: Boulder, CO, USA, 1985.

73. Chen, C.F.; Reinisch, B.W.; Scali, J.L.; Huang, X.; Gamache, R.R.; Buonsanto, M.J.; Ward, B.D. The accuracy of ionogram-derived N(h) profiles. *Adv. Space Res.* **1994**, *14*, 43–46. [CrossRef]

74. Hartman, W.A.; Schmidl, W.D.; Mikatarian, R.; Galkin, I. Correlation of IRTAM and FPMU data confirming the application of IRTAM to support ISS Program safety. *Adv. Space Res.* **2019**, *63*, 1838–1844. [CrossRef]

75. Froń, A.; Galkin, I.; Krankowski, A.; Bilitza, D.; Hernández-Pajares, M.; Reinisch, B.; Li, Z.; Kotulak, K.; Zakharenkova, I.; Cherniak, I.; et al. Towards cooperative global mapping of the ionosphere: Fusion feasibility for IGS and IRI with global climate VTEC maps. *Remote Sens.* **2020**, *12*, 3531. [CrossRef]

76. Themens, D.R.; Jayachandran, P.T.; Bilitza, D.; Erickson, P.J.; Häggström, I.; Lyashenko, M.V.; Reid, B.; Varney, R.H.; Pustovalova, L. Topside electron density representations for middle and high latitudes: A topside parameterization for E-CHAIM based on the NeQuick. *J. Geophys. Res. Space Phys.* **2018**, *123*, 1603–1617. [CrossRef]

77. Dos Santos Prol, F.; Themens, D.R.; Hernández-Pajares, M.; de Oliveira Camargo, P.; de Assis Honorato Muella, M.T. Linear vary-chap topside electron density model with topside sounder and radio-occultation data. *Surv. Geophys.* **2019**, *40*, 277. [CrossRef]

78. Pezzopane, M.; Pignalberi, A. The ESA swarm mission to help ionospheric modeling: A new NeQuick topside formulation for mid-latitude regions. *Sci. Rep.* **2019**, *9*, 12253. [CrossRef]

79. Pignalberi, A.; Pezzopane, M.; Themens, D.R.; Haralambous, H.; Nava, B.; Coïsson, P. On the analytical description of the topside ionosphere by NeQuick: Modeling the scale height through COSMIC/FORMOSAT-3 selected data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2020**, *13*, 1867–1878. [CrossRef]