Space Weather

RESEARCH ARTICLE
10.1029/2018SW002034

Special Section:
Space Weather Capabilities Assessment

Validation of Ionospheric Specifications During Geomagnetic Storms: TEC and foF2 During the 2013 March Storm Event

J. S. Shim[1], I. Tsagouri[2], L. Goncharenko[3], L. Rastaetter[4], M. Kuznetsova[4], D. Bilitza[5], M. Codrescu[6], A. J. Coster[3], S. C. Solomon[7], M. Fedrizzi[6], M. Förster[8], T. J. Fuller-Rowell[6], L. C. Gardner[9], J. Huba[10], A. A. Namgaladze[11], B. E. Prokhorov[8], A. J. Ridley[12], L. Scherliess[9], R. W. Schunk[9], J. J. Sojka[9], and L. Zhu[9]

[1]The Catholic University of America, NASA GSFC, Greenbelt, MD, USA, [2]National Observatory of Athens, Penteli, Greece, [3]Haystack Observatory, Westford, MA, USA, [4]NASA GSFC, Greenbelt, MD, USA, [5]Department of Physics and Astronomy, George Mason University, Fairfax, Virginia, USA, [6]NOAA SWPC, Boulder, CO, USA, [7]High Altitude Observatory, NCAR, Boulder, CO, USA, [8]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam, Germany, [9]Center for Atmospheric & Space Sciences, Utah State University, Logan, UT, USA, [10]Plasma Physics Division, Naval Research Laboratory, Washington, DC, USA, [11]Department of Physics, Murmansk Arctic State University, Murmansk, Russia, [12]Space Physics Research Laboratory, University of Michigan, Ann Arbor, MI, USA

Abstract To address challenges of assessing space weather modeling capabilities, the Community Coordinated Modeling Center is leading a newly established *International Forum for Space Weather Modeling Capabilities Assessment*. This paper presents preliminary results of validation of modeled foF2 (F_2 layer critical frequency) and TEC (total electron content) during the first selected 2013 March storm event (17 March 2013). In this study, we used eight ionospheric models ranging from empirical to physics-based, coupled ionosphere-thermosphere and data assimilation models. The quantities we considered are TEC and foF2 changes and percentage changes compared to quiet time background, and the maximum and minimum percentage changes. In addition, we considered normalized percentage changes of TEC. We compared the modeled quantities with ground-based observations of vertical Global Navigation Satellite System TEC (provided by Massachusetts Institute of Technology Haystack Observatory) and foF2 data (provided by Global Ionospheric Radio Observatory) at the 12 locations selected in middle latitudes of the American and European-African longitude sectors. To quantitatively evaluate the models' performance, we calculated skill scores including correlation coefficient, root-mean square error (RMSE), ratio of the modeled to observed maximum percentage changes (yield), and timing error. Our study indicates that average RMSEs of foF2 range from about 1 MHz to 1.5 MHz. The average RMSEs of TEC are between ~5 and ~10 TECU (1 TEC Unit = 10^{16} el/m^2). dfoF2[%] RMSEs are between 15% and 25%, which is smaller than RMSE of dTEC[%] ranging from 30% to 60%. The performance of the models varies with the location and metrics considered.

Key Points:
- foF2/TEC and foF2/TEC changes during a storm predicted by eight ionosphere models were compared with GIRO foF2 and GPS TEC measurements
- Skill scores (e.g., correlation coefficient, RMSE, yield, and timing error) were calculated
- Model performance strongly depends on the quantities considered, the type of metrics used, and the location considered

Supporting Information:
- Supporting Information S1

Correspondence to:
J. S. Shim,
jasoon.shim@nasa.gov

Citation:
Shim, J. S., Tsagouri, I., Goncharenko, L., Rastaetter, L., Kuznetsova, M., Bilitza, D., et al. (2018). Validation of ionospheric specifications during geomagnetic storms: TEC and foF2 during the 2013 March storm event. *Space Weather*, *16*, 1686–1701. https://doi.org/10.1029/2018SW002034

Received 29 JUL 2018
Accepted 23 SEP 2018
Accepted article online 10 OCT 2018
Published online 5 NOV 2018

1. Introduction

In the past few decades, our general understanding of average behavior of the ionosphere during both quiet and disturbed times and the physical processes responsible for it has been established. However, there is still a lack of knowledge of how solar eruptive phenomena affect the interplanetary space and the near-Earth space environment, including the Earth's upper atmosphere. Modeling is crucial to not only advancing our understanding of causes and consequences of the space weather but also mitigating storm impacts on our modern daily lives. Therefore, ionosphere-thermosphere (IT) models have been developed to predict intensity, occurrence, and duration of ionospheric space weather effects on both ground- and space-based systems, including satellite communication and navigation systems, electric power grids, and pipelines.

To address the needs and challenge of assessment of our current knowledge about space weather effects on IT system and current state of IT modeling capabilities, the Community Coordinated Modeling Center (CCMC) has been supporting community-wide model validation projects, including Coupling, Energetics and Dynamics of Atmospheric Regions (CEDAR) and Geospace Environment Modeling (GEM)-CEDAR modeling challenges. The CCMC initiated the CEDAR electrodynamics thermosphere ionosphere challenge in 2009 focusing on the evaluation of basic IT system parameters modeled, such as electron and neutral densities, the ionospheric F_2 layer peak electron density (NmF2) and peak height (hmF2), and vertical drift (Shim

et al., 2011, 2012, 2014) during both quiet and storm periods. The CEDAR-GEM challenge that followed in 2011 focuses on assessing the prediction of geomagnetic storm impacts on the IT, including Joule heating (Rastätter et al., 2016), neutral density, and total electron content (TEC) (Shim, Rastätter, et al., 2017). Shim, Rastätter, et al. (2017) reported the results of evaluation of TEC prediction during the 2006 December storm (14–15 December) in eight 5°-wide longitude sectors using 15 simulations obtained from 8 ionospheric models.

Since last year, the CCMC has been leading a newly established *International Forum for Space Weather Modeling Capabilities Assessment* to define metrics to assess the current state of space weather modeling capabilities and to track the scientific progress in models that feed into operations. As part of this international effort, IT Working Group consisting of four working teams was formed. One of them, *Ionosphere Plasmasphere Density Working Team*, focuses on the evaluation of storm time TEC and F_2 layer critical frequency (foF2) predictions, which are the most important ionospheric characteristic parameters to assess space weather impact on radio propagation. The Ionosphere Plasmasphere Density Working Team performed its first validation study of foF2 and TEC predictions using various ionosphere/thermosphere models. This study is the first quantitative assessment of both parameters based on a set of metrics, unlike a number of validation studies done previously (Anderson et al., 1998; Araujo-Pradere et al., 2007; Burns et al., 2008; Feltens et al., 2011; Fuller-Rowell, Codrescu, et al., 2000; Orús et al., 2002, 2003; Perlongo et al., 2018; Zhu et al., 2006).

In this paper, we present preliminary results of the evaluation of TEC and foF2 predictions during the 2013 March storm (17 March, DOY [day of year] 076), which is one of the storm events selected by the IT Working Group to investigate differences and similarities in the performance of the models. We evaluated performance of eight IT models in middle latitudes first, which are known to be the best understood ionospheric region largely due to the relatively simple physics and reasonably good coverage of measurements. Moreover, the good coverage of measurements in this region is an additional advantage for model-data comparisons that we attempt here. In section 2, we briefly describe the Global Navigation Satellite System (GNSS) TEC and foF2 measurements, models, and metrics used for this study. In section 3, the results of model-data comparisons and performance of the models are presented. Finally, the summary and conclusions are in section 4.

2. Methods

2.1. Observations

In this study, we first selected 12 ionosonde stations in middle latitudes of the American and European-African longitude sectors (see Figure 1 and Table 1) including 4 stations located in the Southern Hemisphere (SH) in order to investigate longitudinal dependence and hemispheric asymmetry of the performance of the models. We used foF2 observations provided by Global Ionosphere Radio Observatory (GIRO) (http://giro.uml.edu/, Reinisch & Galkin, 2011) and GNSS vertical TEC (vTEC) data provided by Massachusetts Institute of Technology Haystack Observatory (http://cedar.openmadrigal.org/, http://cedar.openmadrigal.org/cgi-bin/gSimpleUIAccessData.py) (Rideout & Coster, 2006). For TEC model-data comparison, at any given location, we used 15 min × 3° latitude × 5° longitude running average every 5 min using 1° latitude × 1° longitude × 5 min-resolution TEC data. The averaged error of the TEC data over all 12 locations for this selected time interval, DOY 076, including 30 quiet days, is about 1 TECU (1 total electron content unit = 10^{16} el/m^2). The number of data points used at each location is 288. For GIRO foF2 observations, we used 45-min running average of 15-min resolution data at all stations except for Port Stanley for which we used 60-min running average due to relatively coarse data coverage. The number of data points at Port Stanley is 48, the other locations have more than 90 data points with a maximum of 96.

2.2. Models

The simulations used in this study were obtained from the latest version of the models available at the CCMC (Webb et al., 2009), with the exception of UAM simulation submitted by the model developer. The models used are IRI (International Reference Ionosphere) storm model, empirical; SAMI3 (Sami3 is also a model of the ionosphere), IFM (ionosphere forecast model), physics-based ionosphere; CTIPe (Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics) model GITM (Global Ionosphere Thermosphere Model), TIE-GCM (Thermosphere Ionosphere Electrodynamics General Circulation Model), UAM-P (Upper Atmosphere Model - Potsdam version), physics-based coupled IT; and USU-GAIM (Utah State University Global Assimilation of Ionospheric Measurement), physics-based ionosphere data assimilation model.

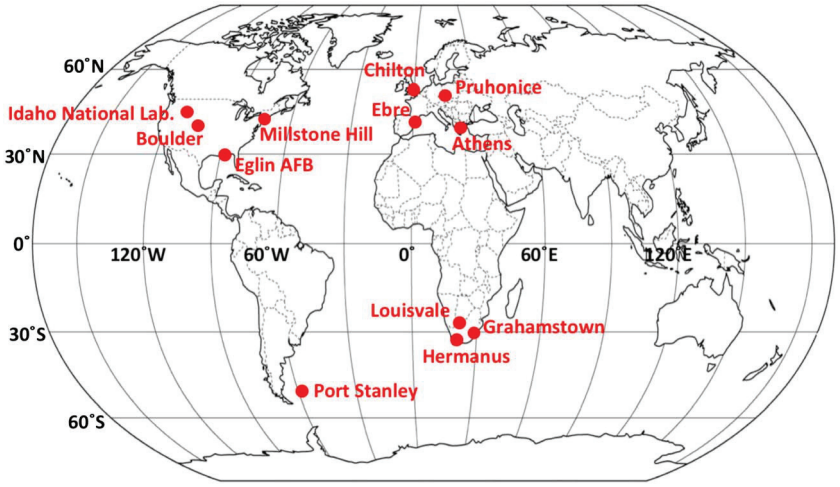

Figure 1. Twelve stations selected for the study.

Table 2 shows the version of the models, input data used for the simulations, and models used for lower boundary forcing and high-latitude electrodynamics. We used unique model setting identifiers to distinguish the simulations used in this study from those, obtained by using different version, input drivers, and/or different boundary conditions, used in our previous studies (Shim et al., 2011, 2012, 2014; Shim, Rastätter, et al., 2017).

Additional information on the models and model setting identifiers is available in Shim et al. (2011, please refer to all references included) and at http://ccmc.gsfc.nasa.gov/challenges/GEM-CEDAR/tags_list.php.

It is worth pointing out the factors in each model simulation that influence foF2 and TEC responses to geomagnetic storms. The IRI storm model (Fuller-Rowell et al., 1999; Fuller-Rowell, Araujo-Pradere, et al., 2000), based on the large number of storms covered by the long record of data from the worldwide network of ionosondes, describes the ratio of storm time foF2 to monthly average foF2 in a first-order approach that captures the summer hemisphere midlatitude ionospheric response. 4_IRI was driven by the time history of the 3-hr ap index over the preceding 30 hr with a weighting function deduced from physically based modeling. The two physics-based ionosphere model simulations, 1_IFM and 1_SAMI3, used the same empirical inputs for the neutral atmosphere and E × B drift. There are notable differences between them: the version of the SAMI3 currently hosted at the CCMC does not include high-latitude ionospheric electrodynamics (HLIE, e.g., the auroral precipitation and the convection electric field pattern), while 1_IFM used empirical models for the HLIE; 1_SAMI3 includes the plasmasphere, whereas 1_IFM does not. All four coupled IT models can be coupled to various models of the HLIE. For the simulations used in this study, 1_UAM-P was driven by field-aligned current. The other three simulations, 11_CTIPE, 11_TIE-GCM, and 6_GITM used the same empirical model for high-latitude electric potential (Weimer-2005. Weimer, 2005); however, to drive Weimer-2005, 11_CTIPE and 6_GITM used the interplanetary magnetic field and solar wind speed and density from the Advanced Composition Explorer 1-min resolution data (http://www.srl.caltech.edu/ACE/ASC), while 11_TIE-GCM used the 15-min trailing average (lagged by 5 min and sampled at 5 min) derived from the 1-min OMNI data set (http://omniweb.gsfc.nasa.gov); each of them was driven by a different model for the auroral precipitation. 6_GITM used Flare Irradiance Spectral Model (http://lasp.colorado.edu/lisird/data/fism) solar extreme ultraviolet (EUV) irradiance (Chamberlin et al., 2007), while the others used $F_{10.7}$ to specify the solar EUV flux using EUV flux model for aeronomic calculations (Richards et al., 1994). 1_UAM-P and 11_CTIPE include the plasmasphere, while 6_GITM and 11_TIE-GCM (Solomon et al., 2018) do not. The data assimilation model simulation, 1 _USU-GAIM, currently used only GNSS TEC data between −60 and +60 geographic latitudes.

2.3. Metrics

As the quiet-time reference (TEC_quiet) to quantify impacts of the storm, we used 30-day median value at a given time. The 30 days consist of 15 days before (1–15 March 2013) and 15 days after (22 March to 5 April 2013) the storm. The quantities considered to assess how well the models produce foF2 and TEC during

Table 1
Twelve Selected Locations

Ionosonde stations			Geographic Coord.		Geomagnetic Coord.
Name		Code	Latitude (°N)	Longitude (°E)	Latitude (°N)
American longitude	Millstone Hill	MH453	42.60	288.50	52.10
	Idaho National Lab	IF843	43.81	247.32	50.64
	Boulder	BC840	40.00	254.70	47.80
	Eglin AFB	EG931	30.50	273.50	39.81
	Port Stanley	PSJ5J	−51.60	302.10	−41.8
European-African longitude	Chilton	RL052	51.50	359.40	53.35
	Pruhonice	PQ052	50.00	14.60	49.3
	Ebre	EB040	40.80	0.50	42.79
	Athens	AT138	38.00	23.50	36.17
	Louisvale	LV12P	−28.50	21.20	−28.51
	Hermanus	HE13N	−34.42	19.22	−33.96
	Grahamstown	GR13L	−33.30	26.50	−34.10

the storm are, at any given location, (1) shifted values by subtracting a minimum value of the quiet time reference in 24 hr (e.g., TEC*[DOY, UT] = TEC [DOY, UT] − minimum of TEC_quiet = TEC [DOY, UT] − minimum 30-day median); (2) changes due to the storm with respect to the quiet time reference (e.g., dTEC [DOY, UT] = TEC [DOY, UT] − TEC_quiet[UT]); and (3) percentage changes (e.g., dTEC[%][DOY, UT] = 100*dTEC [DOY, UT]/TEC_quiet[UT]).

Potential systematic uncertainties in the models and observations and baseline differences between the models and between models and observations can be reduced by considering the shifted values and changes from the their own quiet time background values. In addition, the impact of differing upper boundaries for TEC calculations is likely reduced by using these quantities, since plasmaspheric TEC variations with geomagnetic activity are negligible in middle latitudes (Shim, Jee, et al., 2017).

For TEC validation, we also considered normalized percentage changes of TEC (dTEC[%]_norm = [dTEC[%] − ave_dTEC[%]]/std_dTEC[%]), where ave_dTEC[%] is the average of percentage changes of TEC (dTEC[%]) at a given time and at a given location over the quiet 30 days described above, and std_dTEC[%] is the standard deviation of the average percentage change. The normalized percentage change, dTEC[%]_norm, has the advantage to exclude seasonal, local time, and latitudinal dependences of TEC variability by normalizing the percentage variation using its statistical standard deviation and thus should have mainly storm-induced variation (Nishioka et al., 2017). Figure 2 shows an example of dTEC[%] (in red) and dTEC[%]_norm (in blue) at Athens on DOY 076 (between 06 UT and 22 UT). The difference between local daytime (around 12 UT) and nighttime (around 21 UT) in dTEC[%] peak value has disappeared in dTEC[%]_norm.

To quantify the performance of the models, we calculated the following skill scores: including (1) correlation coefficient (CC); (2) root-mean-square error (RMSE $=\sqrt{\frac{\sum(x_{obs}-x_{mod})^2}{N}}$, where x_{obs} and x_{mod} are observed and modeled values); (3) the ratio of the peak of modeled percentage change to that of the observed one (Yield $=\frac{(x_{mod})_{max}}{(x_{obs})_{max}}$), for example, [max. (min.) dTEC[%] of the model simulations]/[max. (min.) dTEC[%] of GNSS vTEC for positive (negative) phase] during the main phase of the storm; and (4) timing error, which is the difference between the modeled peak time and observed peak time (TE = t_peak_model − t_peak_obs.).

CC measures how well the observed and modeled values are linearly correlated (in phase) with each other, and RMSE measures how different the values are on average over the time interval considered. We calculated CC and RMSE for the error values below 95th percentile in order to remove the effect of outliers of the predicted values, which is one of the main concerns of using CC and RMSE. In addition to CC and RMSE, we considered yield and timing error (TE) to measure the models' capability to capture peak disturbances during the storm.

3. Performance of the Models in Predictions of foF2 and vTEC on 17 March 2013

There have been a number of extensive studies on geospace responses to the 2013 March storm (see Figure 3 for solar wind and geomagnetic conditions) together with the 2015 March storm (Zhang et al., 2017, please

Space Weather

10.1029/2018SW002034

Table 2
Models Submitted for This Study

Model setting ID	Model version	Drivers: Input data	Drivers: Models used for thermosphere, tides from lower boundary, and high-latitude electrodynamics		Upper boundary for TEC calculation/resolution
Empirical model					
4_IRI [a]	IRI-2016 (Bilitza et al., 2014, 2017; Fuller-Rowell et al., 1999; Fuller-Rowell, Araujo-Pradere, & Codrescu, 2000)	$F_{10.7}$, R12, IG12, and ap	URSI foF2 model, NeQuick for topside Ne, and NRLMSISE00 neutral composition (Picone et al., 2002)		~2,000 km
Physics-based ionosphere model					
1_SAMI3[a]	SAMI3 (Huba et al., 2000, 2008)	$F_{10.7}$ and ap	NRLMSISE00 neutral composition, HWM93 wind (Drob et al., 2008; Hedin et al., 1991), and E × B drift (Scherliess & Fejer, 1999)		~20,000 km, 1.5°lat × 4°lon
1_ IFM[a]	IFM (Schunk et al., 1997, 2002)	$F_{10.7}$, daily Ap, and Kp	MSIS-90 neutral composition (Hedin, 1991), HWM neutral wind (Hedin et al., 1991), Scherliess and Fejer E × B drift, high-latitude electric fields (Heppner & Maynard, 1987), and Hardy auroral precipitation (Hardy et al., 1985)		~1,600 km, 3°lat × 7.5°lon
Physics-based coupled ionosphere-thermosphere model					
			Tides	High-latitude electrodynamics	
11_CTIPE[a]	CTIPe_3.2 (Codrescu et al., 2000; Millward et al., 2001)	$F_{10.7}$, ACE IMF data and solar wind speed and density, NOAA POES hemispheric power data	migrating diurnal and semidiurnal tides	Weimer-2005 high-latitude electric potential (Weimer, 2005), Fuller -Rowell and Evans auroral precipitation (1987)	~2,000 km, 2°lat × 18°lon
6_GITM[a]	GITM_2.5 (Ridley et al., 2006)	FISM solar EUV irradiance, ACE IMF data, and solar wind speed and density	GSWM (Hagan et al., 1999) migrating diurnal and semidiurnal tides	Weimer-2005 high-latitude electric potential, Ovation auroral precipitation (Newell et al., 2009; Newell & Gjerloev, 2011)	~600 km, 2.5°lat × 5°lon
11_TIE-GCM[a]	TIE-GCM2.0 (Richmond et al., 1992; Roble et al., 1988; Solomon et al., 2012)	$F_{10.7}$, OMNI IMF data, and solar wind speed and density	GSWM migrating diurnal and semidiurnal tides	Weimer-2005 high-latitude electric potential, Roble and Ridley auroral precipitation (1987)	~600 km, 5°lat × 5°lon
1_UAM-P	Upper Atmosphere Model (UAM) - Potsdam Version (Namgaladze et al., 1988, 1991; Prokhorov et al., 2018)	$F_{10.7}$, Ap, Kp, AE, IMF data, and solar wind speed and density	with no tides	FACs using Iijima and Potemra Model (IPM) (Iijima & Potemra, 1976) and the Hardy auroral precipitation	~20,200 km 1.5–2.5°lat (irregular) × 10°lon
Physics-based data assimilation model					
1_USU-GAIM[a]	USU-GAIM2.3(Scherliess et al., 2004, 2006; Schunk et al., 2004)	*F*10.7, daily Ap, and Kp, GPS TEC observations from more than 350 ground stations ($-60° <$ latitude $< 60°$)	The IFM background physics-based ionosphere model, the same models used for IFM		~1,400 km, 3~5°lat (irregular) × 15°lon

Note. CCMC = Community Coordinated Modeling Center; TEC = total electron content; foF2 = F_2 layer critical frequency; SAMI-3 = Sami3 is also a model of the ionosphere; ACE = Advanced Composition Explorer; IMF = interplanetary magnetic field; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics Model; FISM = Flare Irradiance Spectral Model; GITM = Global Ionosphere Thermosphere Model; FACs = field-aligned currents; EUV = extreme ultraviolet; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version; HWM = Horizontal Wind Model; URSI = International Union of Radio Science; MSIS = Mass Spectrometer Incoherent Scatter; GSWM = Global Scale Wave Model; GPS = Global Positioning System.
[a]The model results are submitted by the CCMC using the models hosted at the CCMC.

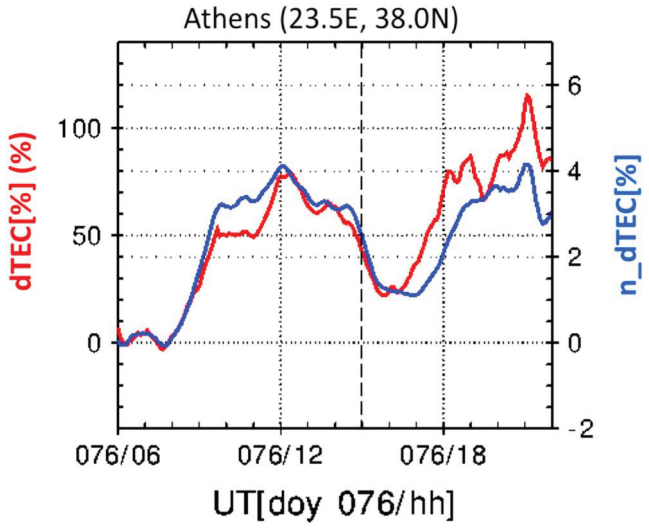

Figure 2. Example of dTEC[%] (red) and n_dTEC[%] (blue) at Athens between 06 UT and 22 UT on DOY 076. DOY = day of year.

refer to the references included). Figure 4 shows foF2 and TEC during the main phase of the 2013 (in red) along with the 30-day medians (in blue) at the 12 locations. As reported by Yue et al. (2016), the storm-enhanced density in middle-high latitudes and its source, the positive ionospheric storm effects, in the lower latitudes were observed around 12 UT in the European (EU) sector and around 20 UT in the North American (NA) sector, while the negative storm effects on the west side of the NA sector were also observed (at Idaho and Boulder) in the right two columns in Figure 4. In the SH, the positive TEC changes were observed in both South America (SAM) and South Africa (SAF). The foF2 responses to the storm are similar to the TEC responses with larger change in TEC at the selected locations except for Chilton and Grahamstown in which there were noticeable increases in TEC while foF2 hardly increases. We found that most model simulations do not reproduce the difference between eastern and western parts of the NA sector; for example, TEC increases at Millstone Hill and decreases at Idaho and Boulder around 20 UT, and foF2 and TEC show different responses in Chilton and Grahamstown (see Figure S1 in supporting information).

Figure 5 shows scatterplots of the observed (*x* axis) and modeled (*y* axis) shifted foF2 (foF2* in the first and third columns) and percentage change of foF2 (dfoF2[%] in the second and fourth columns) during the storm (17 March 2013) for all 12 locations grouped into 4 sectors: NA (denoted in green), Europe (EU, blue), SAF (red), and SAM (black). For most simulations, after shifting by subtracting the minimum of 30-day median, the modeled foF2* agrees better with the observed one than before shifting (see Figure S2 in the supporting information). For example, before shifting most foF2 data points of 1_GITM and 1_USU-GAIM are below and above the line with slope 1 (black solid line), respectively. This indicates that 1_GITM underestimates foF2, while 1_USU-GAIM overestimates it. They each produce the smallest (0.8 MHz) and largest (4.5 MHz) averaged minimum of 30-day median over all 12 locations compared to the observed one (3.6 MHz, see Figure S1). Most models appear unable to produce foF2* larger than about 7 MHz during the storm. The modeled dfoF2[%] shows less agreement with the observed values than the modeled foF2* does. The modeled dfoF2 (not shown here) and dfoF2[%] show similar relationship with the observed values.

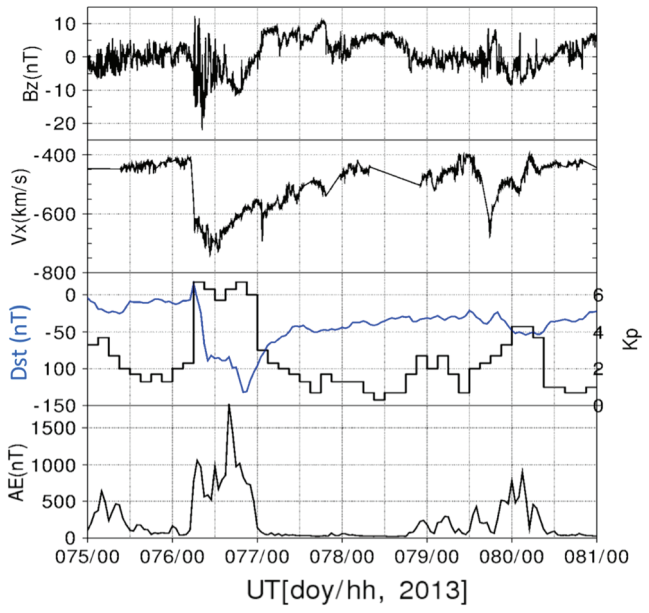

Figure 3. IMF Bz, Vx, Dst (blue), Kp (black), and AE values for the 2013 March event. IMF = interplanetary magnetic field.

In Figure 6, we show scatterplots of the observed (*x* axis) and modeled (*y* axis) shifted TEC and percentage change of TEC; different colors are associated with different sectors (same as Figure 5). It is found that the averaged minimum of 30-day median of the observed TEC is about 5 TECU, and the lowest and highest modeled averages are 0.2 TECU (6_GITM) and 7.8 TECU (1_UAM-P). The relationships between the modeled and observed TEC* are similar to those of foF2* with the tendency for most of the simulations to underestimate TEC* larger than around 30–40 TECU. This tendency is more pronounced in 11_CTIPE and 6_GITM simulations.

Most models appear to not reproduce the large enhancements of dTEC[%] (about 200%) at Port Stanley in South America. 1_USU-GAIM agrees better with GNSS vTEC than the other simulations do due to the fact that 1_USU-GAIM assimilated GNSS observations, which are included to produce the Massachusetts Institute of Technology TEC data set. Like the predicted dfoF2[%], the modeled dTEC[%] shows less agreement with the observed value than the modeled TEC* does. From now on, foF2 and TEC will represent shifted foF2 (foF2*) and shifted TEC (TEC*), respectively.

3.1. Correlation Coefficient

To quantitatively assess the models' performance of TEC and foF2 predictions, we first calculated CC between the modeled and observed foF2 and

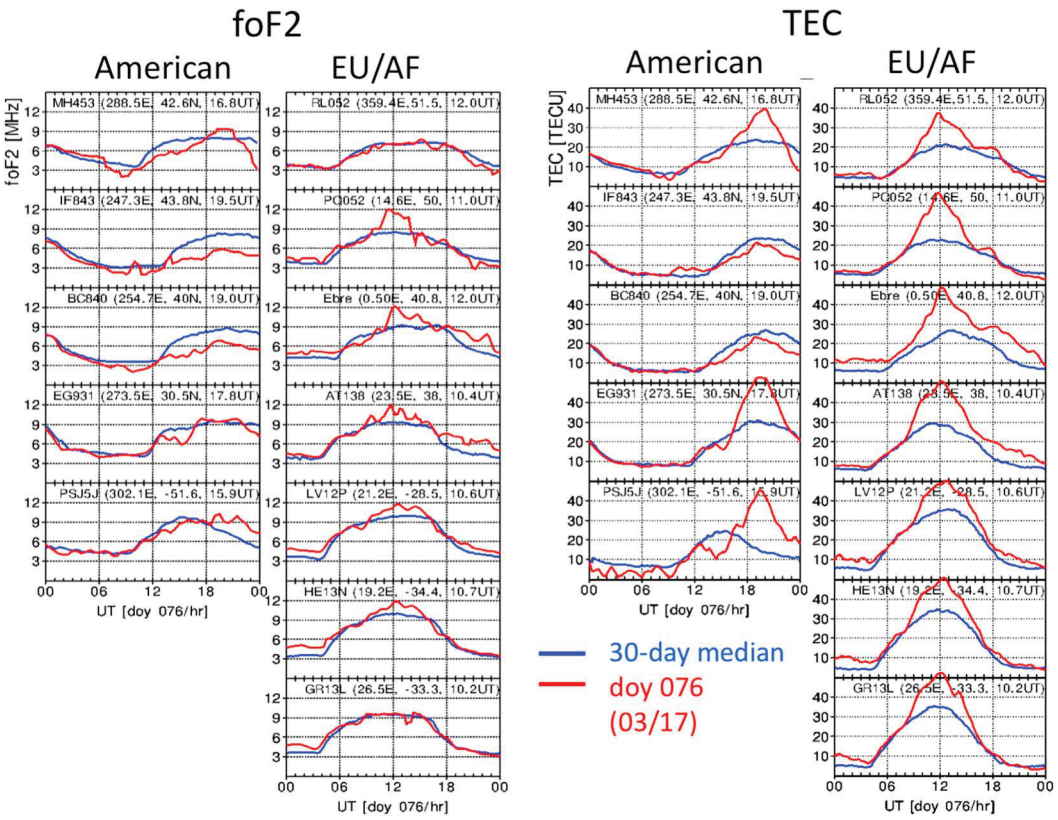

Figure 4. Observed foF2 and vertical TEC at 12 locations on DOY 76, 2013 (red) and 30-day median values (blue). TEC = total electron content; DOY = day of year; foF2 = F_2 layer critical frequency; EU = Europe; AF = Africa.

TEC for DOY 076. In Figure 7, the simulations were first grouped based on the type of the model and then arranged alphabetically in each group (foF2 in the left panel and TEC in the right panel). Four CC were displayed for each simulation. First three scores correspond to the average CC over EU, NA, and SH (SH refers to SAF and SAM combined), and the last one is the average of all 12 locations. Different colors denote different quantities: Blue denotes shifted foF2 and TEC, green and red the change and percentage changes, and black the normalized percentage change. The closer the circles are to the horizontal line of 1, the better the model performs. The modeled foF2 and TEC (blue dots) are highly correlated with the observed values.

The average CC values, over all 12 locations, between the modeled and observed foF2 (TEC) are about 0.8–0.9 (0.7–0.95). However, the modeled changes (green), percentage changes (red), and normalized percentage changes (black only applicable for TEC) are much less correlated (closer to uncorrelated) with the observed values, especially for dfoF2 and dfoF2[%] ($\sim -0.1 <$ average CC < 0.4). There is no big difference between dTEC[%] and dTEC[%]_norm based on the average values for each simulation.

For the changes and percentage changes, the physics-based coupled IT and data assimilation models have better CC than the ionosphere models have. The differences in CC among locations are smaller for foF2 and TEC than for the changes and percentage changes. Most simulations, except for 6_GITM, show lower CC for dfoF2 and dTEC in NA. It seems to be caused by the opposite responses to the storm in the eastern (positive phase) and western (negative phase) parts of NA not being captured, producing either positive (most physics-based models) or negative (4_IRI) storm in all locations of NA instead. 6_GITM shows the opposite responses relatively well although it underestimated the magnitude of the change.

4_IRI produces CC comparably to the physics-based models that perform best for both foF2 (11_CTIPE) and TEC (11_TIE-GCM) predictions, although it does not for dTEC and dTEC[%].

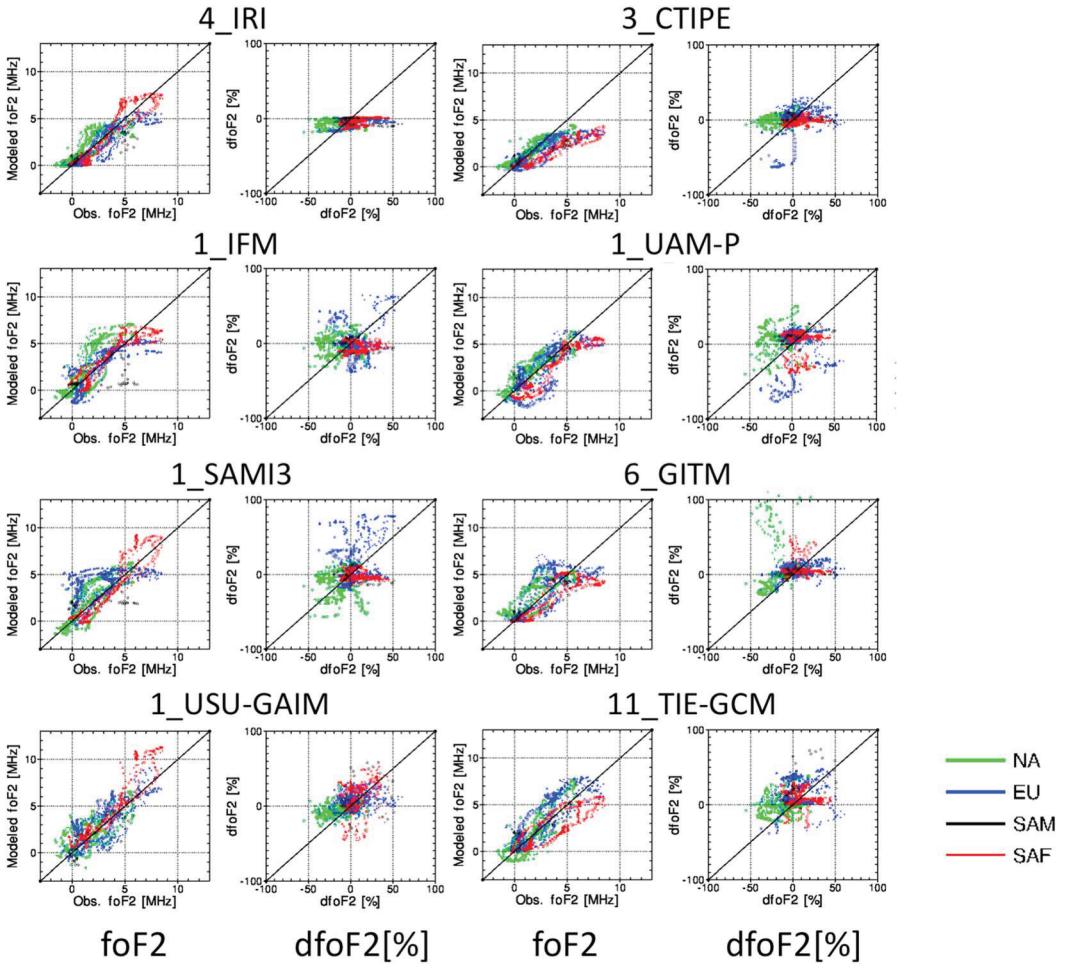

Figure 5. Scatterplots of the observed (*x* axis) and modeled (*y* axis) shifted foF2 (foF2* in the first and third columns) and percentage change of foF2 (dfoF2[%] in the second and fourth columns) during the storm (17 March 2013) for all 12 locations: North American sector (denoted in green), European sector (EU, blue), South Africa (SAF, red), and South America (SAM, black). foF2 = F_2 layer critical frequency; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics; GITM = Global Ionosphere Thermosphere Model; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version.

To compare the two physics-based ionosphere models, 1_IFM performs about the same as 1_SAMI3 for all the considered quantities based on the average CC, but 1_SAMI3 shows less correlation with the observed foF2 and TEC in EU, compared to 1_IFM.

Four physics-based coupled IT models also produce similar average CC. However, 11_CTIPE (11_TIE-GCM) shows high correlation for all three sectors for foF2 (TEC). Except for 6_GITM, the other three simulations show lower CC for dfoF2 and dTEC in NA than the other sectors. 11_TIE-GCM and 6_GITM show the most dependence of CC for dTEC and dfoF2 on location, respectively.

3.2. Root-Mean-Square Error

Figure 8 shows RMSE of foF2 and dfoF2, in the left panel, and TEC and dTEC, in the right panel. For foF2 (blue) and dfoF2 (green) predictions, average RMSE values range from ~1 (1_IRI and 6_GITM) to 1.5 MHz (1_SAMI3 and 1_UAM-P). 11_CTIPE has the lowest and highest RMSE for foF2: 0.7 in NA and 2 MHz in SH.

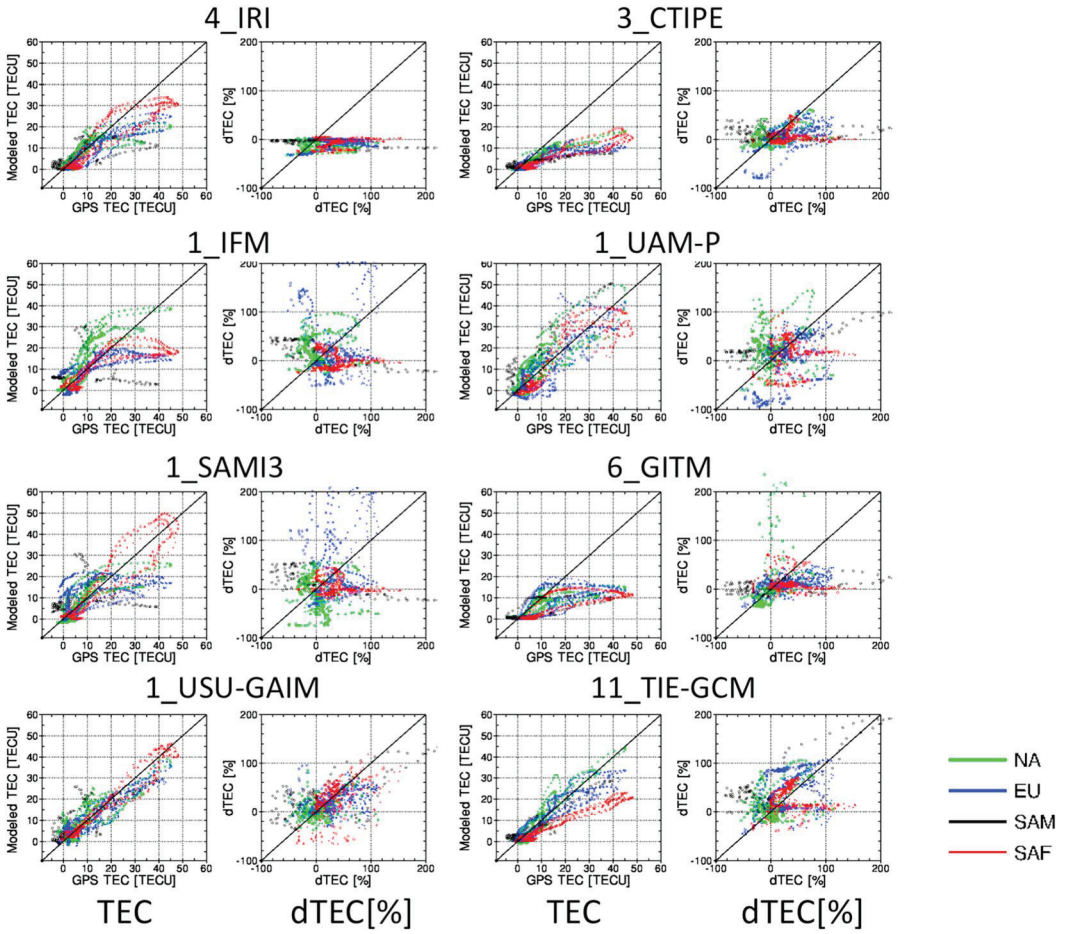

Figure 6. Same as Figure 5 but for shifted TEC (TEC*) and percentage change of TEC (dTEC[%]).TEC = total electron content; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics Model; GITM = Global Ionosphere Thermosphere; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version.

Average RMSE of TEC are between 5 (1_USU-GAIM) and 10 TECU (11_CTIPE and 6_GITM). However, the minimum RMSE for TEC is 3 TECU in NA (11_TIE-GCM), and the maximum RMSE is about 13 TECU in SH (11_CTIPE and 6_GITM). For dTEC, average RMSE varies from about 4 (1_USU-GAIM) and 7 TECU (1_UAM-P).

As seen in CC, RMSE also varies highly with location. Most models appear to predict foF2 and/or TEC better in NA and worse in SH (e.g., 11_CTIPE has its lowest of 5 TECU in NA and highest of 13 TECU in SH). Some models have the same tendency of better performance in NA for dTEC (e.g., 4_IRI, 11_CTIPE, and 6_GITM), while most models show the opposite tendency for dfoF2 prediction (e.g., 1_IFM, 1_SAMI3, 11_CTIPE, 6_GITM, 11_TIE-GCM, and 1_UAM-P). However, 4_IRI and 1_USU-GAIM that performed best for foF2 and TEC predictions, respectively, show less location dependence of RMSE than the others for foF2 and TEC predictions.

Shim, Rastätter, et al. (2017) found that the longitudinally averaged RMSE for both TEC* and dTEC (using TEC one day prior to the storm as a quiet time reference) in middle latitude during 2006 December event ranges from about 2 to 8 TECU, which are about 2–3 TECU smaller than average RMSE for 2013 March event. Overall, most models perform somewhat worse for 2013 March event than for 2006 December event. This is possibly partly attributed to TEC prediction at the four stations in the South Atlantic Anomaly (SAA) region (see Figure S3), which is not considered by Shim, Rastätter, et al. (2017). In the SAA region, TEC enhancement occurred due to the energy input from the outer radiation belts (Dmitriev et al., 2017); however, most models do not predict the large enhancement of TEC (e.g., dTEC[%] at Port Stanley) as we mentioned above.

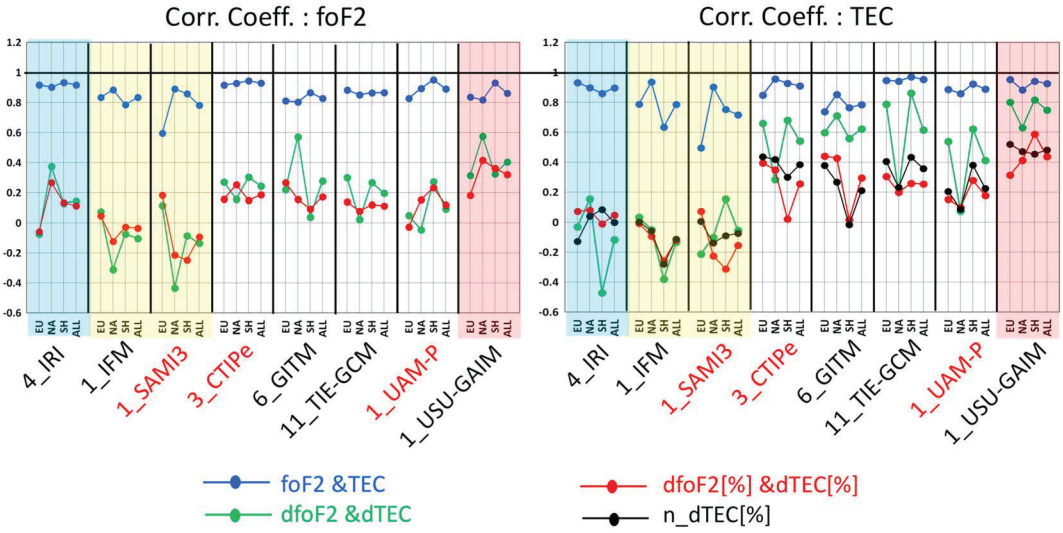

Figure 7. Correlation coefficient (CC) between modeled and observed foF2 (left panel) and TEC (right panel). Four CC were displayed for each simulation: Average CC over Europe (EU), North America (NA), Southern Hemisphere (SH), and the average over all 12 locations, from left to right. Different colors denote different quantities. Blue denotes shifted foF2 and TEC, green and red the change and percentage changes, and black the normalized percentage change. The simulations were first grouped by the type of the model and then arranged alphabetically in each group. The simulations in red include the plasmasphere. TEC = total electron content; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics; GITM = Global Ionosphere Thermosphere Model; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version; foF2 = F_2 layer critical frequency.

Based on the average RMSE, for TEC and dTEC, 1_SAMI3 (1_IFM) tends to perform worse in EU (SH) than the other sectors. Among the four physics-based coupled models, 11_TIE-GCM predicts TEC and dTEC better, 11_CTIPE and 6_GITM perform better for dfoF2, and 1_UAM-P predicts foF2 better.

Figure 9 shows RMSE of percentage changes of foF2 (blue) and TEC (red) and normalized percentage changes of TEC (black). RMSEs of dfoF2[%] based on the average over the 12 locations are between 15% and 25%, which is smaller than RMSE of dTEC[%] ranging from 30% to 60%. Difference in the performance among locations is more noticeable in dTEC[%] and dTEC[%]_norm than in dfoF2[%]. All simulations, except 6_GITM, have smaller RMSE of dTEC[%] in NA than in the other two sectors. 4_IRI produces comparable dfoF2[%] and dTEC[%] to those of the physics-based models for most cases. However, 4_IRI has larger RMSE in dTEC[%]_norm than the others do possibly due to the small standard deviation of dTEC[%]. It needs to be noted that IRI represents average conditions of the ionosphere and has a limitation to model relatively short-term disturbances during geomagnetic activities.

Differences in the performance based on RMSE of dTEC[%] among the simulations appear do not match those based on RMSE of dTEC[%]_norm. For example, 1_SAMI3 shows slightly larger average RMSE of dTEC[%] but smaller RMSE of dTEC[%]_norm than 1_IFM. 11_TIE-GCM shows smaller average RMSE of dTEC[%] but larger RMSE of dTEC[%]_norm than the other three coupled models. This is probably due to the difference in the normalization factor, standard deviation of dTEC[%], between the simulations.

3.3. Yield and Timing Error

In order to measure how well the models capture the degree of TEC and foF2 disturbances during the main phase, we calculated yield and TE of dfoF2[%], dTEC[%], and dTEC[%]_norm. In most of the 12 locations two peaks were observed as seen in Figure 2; therefore, we considered two time intervals, 06–15 UT and 15–22 UT, that correspond to about 06–17 LT and 15–24 LT in European-African longitude sector and about 22–11 LT and 08–18 LT in American longitude sector (see Figure S4 for yield and TE for all locations for each model). In each time interval, one yield value and one TE value were calculated as defined in section 2.3.

The yield is positive if the model correctly predicts the storm phase. In each sector, we calculate average yield and average TE over the number of stations where the yield is positive. Table 3 shows the total number of

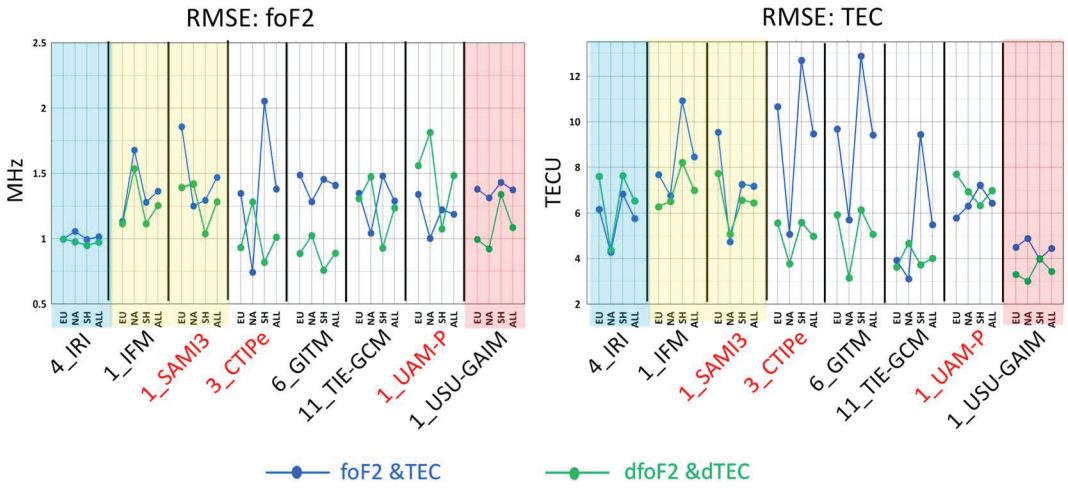

Figure 8. Same as Figure 7 but for RMSE of shifted foF2 and TEC, and changes of foF2 and TEC. TEC = total electron content; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics; GITM = Global Ionosphere Thermosphere Model; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version; foF2 = F_2 layer critical frequency; RMSE = root-mean square error.

stations where the models have correct storm phase, either positive or negative. The numbers in bold are the highest among the simulations. The coupled IT models and data assimilation model predict the storm phase better than the other three simulations, 4_IRI, 1_IFM, and 1_SAMI3.

In Figure 10, we present average yield (left) and average of absolute values of TE (right) over the two time intervals: dfoF2[%] in blue, dTEC[%] in red, and dTEC[%]_norm in black. Concerning the average of all 12 locations, yields of dfoF2[%] and dTEC[%] range from 0.5 to 1.5 with larger variation for dTEC[%] with location. For dTEC[%]_norm, yield values are between 0.8 and 3. Some of the simulations give better ratio for dTEC[%]_norm (closer to the horizontal line of 1) than for dTEC[%] (e.g., 11_CTIPE, 1_SAMI3, and 1_UAM-P); some of them give worse ratio for dTEC[%]_norm (e.g., 11_TIE-GCM in EU and 4_IRI in NA, which shows the largest difference in yield between dTEC[%] and dTEC[%]_norm). In most cases, 11_CTIPE and 6_GITM appear to underestimate percentage changes, while 1_UAM-P and 11_TIE-GCM tend to overestimate.

Average timing errors of dfoF2[%] and dTEC[%]_norm are between 1 and 4 hr, and TE of dTEC[%] are about 1–5 hr. 1_USU-GAIM has smaller TE for dTEC[%] (~1.5 hr) than for dfoF2[%] (~3 hr). The physics-based models have 1–2 hr-average TE for dfoF2[%], with the smallest of ~0.5 hr in EU predicted by 11_CTIPE, and 1.5–2.5 hr-average TE for dTEC[%] and dTEC[%]_norm.

1_SAMI3 gives slightly better yield and TE for most cases than 1_IFM and shows better yield and worse TE of dTEC[%]_norm compared to dTEC[%].

4. Summary and Conclusions

We performed a systematic assessment of TEC and foF2 predictions of IT models during the 2013 March storm event (DOY 076). We compared modeled foF2 and TEC obtained from eight ionospheric models with the ground-based GIRO foF2 and GNSS vTEC data for the selected 12 locations in middle latitudes of the American and European-African longitude sectors.

The quantities considered for model-data comparison are (1) shifted foF2 and TEC (e.g., TEC*) that were obtained by subtracting minimum of quiet time reference, 30-day median; (2) foF2 and TEC changes (e.g., dTEC) with respect to the quiet time reference; (3) percentage changes of foF2 and TEC (e.g., dTEC[%]); and (4) normalized percentage change of TEC (dTEC[%]_norm) by standard deviation of dTEC[%].

To quantify the performance of the models, we calculated skill scores, including CC, RMSE, the ratio of maximum (minimum) percentage change of the models to the observation (yield), and the TE (peak time

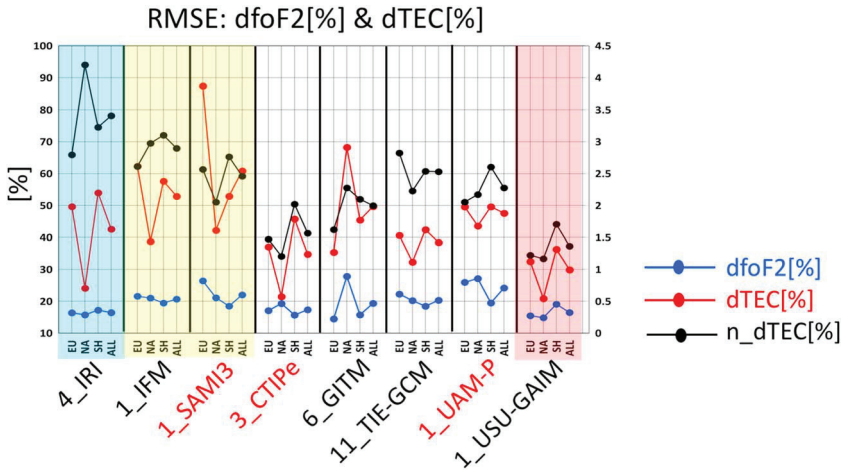

Figure 9. Same as Figure 7 but for RMSE of percentage change of foF2 and TEC, and normalized percentage change. Blue denotes dfoF2[%], red dTEC[%], and black n_dTEC[%]. TEC = total electron content; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics; GITM = Global Ionosphere Thermosphere Model; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version; foF2 = F_2 layer critical frequency; RMSE = root-mean square error.

difference between the simulations and the observations = t_peak_model – t_peak_obs). The skill scores were calculated for the three sectors: EU, NA, and SH to investigate the longitudinal and hemispheric dependence of the performance of the models.

Comparisons of the eight simulations indicate that most simulations tend to underestimate enhanced foF2 and TEC and appear to not reproduce large enhancements of dTEC[%] (e.g., about 200% TEC increase at Port Stanley in the SAA region). Most of them also appear to not capture opposite responses to the storm in the eastern and western parts of NA. This is what in part causes lower CC for dfoF2 and dTEC in NA. However, with respect to RMSE, most models tend to predict foF2 and/or TEC better in NA and worse in SH. Some models have the same tendency of better performance in NA for dTEC (e.g., 4_IRI, 11_CTIPE, and 6_GITM). Averaged RMSEs of foF2, over all 12 locations, range from about 1 to 1.5 MHz. The average RMSEs of TEC are between about 5 and 10 TECU. dfoF2[%] RMSEs are between 15% and 25%, which is smaller than RMSE of dTEC[%] ranging from 30% to 60%. In terms of CC, the physics-based coupled IT (11_CTIPE, 6_GITM, 11_TIE-GCM, and 1_UAM-P) and data assimilation (1_USU-GAIM) models predict the changes and percentage changes better than the three ionosphere models (4_IRI, 1_IFM, and 1_SAMI3). In addition, the coupled IT models and data assimilation model predict the storm phase better than the other three simulations, 4_IRI, 1_IFM, and 1_SAMI3.

As rather expected based on the fact that the empirical model IRI represents average conditions of the ionosphere, the IRI performs comparably to the physics-based models concerning RMSE and CC, which measure

Table 3
Number of Locations Where the Models Correctly Predict Negative or Positive Phase

	Time interval	4_IRI	1_IFM	1_SAMI3	11_CTIPE	6_GITM	11_TIEGCM	1_UAM-P	1_USU-GAIM
dfoF2	06–15 UT	7	5	7	8	5	**9**	8	**9**
	15–22 UT	6	6	6	**10**	7	6	8	8
dTEC	06–15 UT	2	8	7	9	**10**	8	**10**	**10**
	15 22 UT	2	6	9	7	**12**	10	7	11

Note. IRI = International Reference Ionosphere; IFM = ionosphere forecast model; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics Model; GITM = Global Ionosphere Thermosphere Model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; UAM-P = Upper Atmosphere Model - Potsdam version.

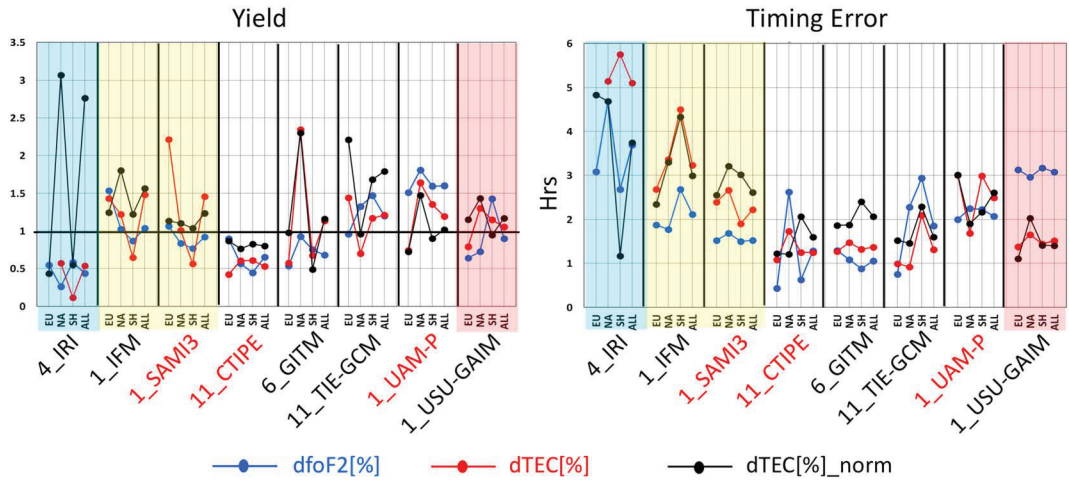

Figure 10. Same as Figure 7 but for yield (ratio) and absolute of timing error (|TE| = |t_peak_model − t_peak_obs|). TEC = total electron content; SAMI-3 = Sami3 is also a model of the ionosphere; CTIPE = Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics; GITM = Global Ionosphere Thermosphere Model; TIE-GCM = Thermosphere Ionosphere Electrodynamics General Circulation Model; IFM = ionosphere forecast model; USU-GAIM = Utah State University Global Assimilation of Ionospheric Measurement; IRI = International Reference Ionosphere; UAM-P = Upper Atmosphere Model - Potsdam version; foF2 = F_2 layer critical frequency.

the performance of the models on average over the time interval considered. 4_IRI produces the smallest RMSE for foF2 prediction with less location dependence. However, the IRI has a limitation in modeling relatively short-term disturbances during geomagnetic activities; therefore, it performs relatively poorly in its prediction of the maximum and minimum of percentage changes (yields and timing error, especially in NA, for dTEC[%]_norm). Compared to the physics-based model simulations, 4_IRI has larger RMSE in dTEC[%]_norm due to the small standard deviation of dTEC[%]. 1_USU-GAIM, which assimilated only GNSS TEC data, performs best for TEC prediction in most cases with the least location dependence of RMSE. 1_USU-GAIM makes better predictions of TEC-related quantities than foF2-related ones. This suggests that the integration of additional observations, such as ionosonde profiles, is necessary to make better prediction of foF2.

To compare the same types of models, two physics-based ionosphere models, 1_IFM and 1_SAMI3, show similar performance in terms of average CC and RMSE. 1_SAMI3 shows slightly worse CC and RMSE for foF2 and TEC in EU, and slightly better yield and TE for most cases. Overall, the average performances of the four physics-based coupled IT models over the 12 locations are comparable to each other. Among the four simulations, only 6_GITM captures the opposite responses in the eastern and western parts of NA relatively well, although it underestimates the magnitude of the change. Therefore, unlike the other simulations, 6_GITM produces better CC for dfoF2/dTEC in NA than in the other two sectors. Based on the average RMSE, 11_TIE-GCM predicts TEC and dTEC better, 11_CTIPE and 6_GITM perform better for dfoF2, and 1_UAM-P predicts foF2 better. In terms of CC, 11_CTIPE and 11_TIE-GCM show high correlation for all three sectors for foF2 and TEC, respectively. In most cases, 11_CTIPE and 6_GITM tend to underestimate percentage changes, while 1_UAM-P and 11_TIE-GCM overestimate them. The differences in performance among these four simulations could be caused by inherent differences among the models, for example, different methods to solve for chemistry and advection, and different ways to treat eddy diffusion and vertical transport (Fuller-Rowell et al., 1996; Perlongo et al., 2018; Prokhorov et al., 2018; Ridley et al., 2006; Solomon et al., 2012). In addition, the performance differences could also be caused by a combination of different input data and different models used for lower boundary forcing and high-latitude electrodynamics, since each simulation was obtained by using its default input data and drivers. To investigate the actual causes of these differences, detailed studies will be required, which are beyond the scope of this paper.

Furthermore, our findings suggest that in order to accurately model IT disturbances during geomagnetic storms, it is crucial to advance our knowledge on the coupling of IT with magnetosphere to correctly describe penetration electric fields and enhancement of energy deposition, which leads to composition changes, winds, and electron density changes (Huba et al., 2017; Maruyama et al., 2005). Theoretical modeling will

help us better understand the processes responsible for the observed characteristics and features during disturbed conditions, and this will improve our space weather forecasting capabilities.

This is the first systematic study to evaluate current capability of modeling foF2 and TEC during geomagnetic storms, performed by Ionosphere Plasmasphere Density Working Team of the International Forum for Space Weather Modeling Capabilities Assessment. The findings of this study will provide a baseline for future validation studies using new models and improved models, along with earlier results (Shim et al., 2011, 2012, 2014; Shim, Rastätter, et al., 2017) obtained through CEDAR electrodynamics thermosphere ionosphere and GEM-CEDAR modeling challenges. We will extend our study to include more geomagnetic storm events selected by the IT Working Group, including the 2015 March and 2017 September events, to investigate differences and similarities in the performance of the models. In addition, we will also include foF2 and TEC predictions in high and low latitudes.

Acknowledgments
The vTEC data were provided by MIT Haystack Observatory and can be obtained through CEDAR Madrigal database (http://cedar.openmadrigal.org). We thank the operators of the digisondes for sharing their data through http://giro.uml.edu/. Data from the South African ionosonde network are made available through the South African National Space Agency (SANSA), who is acknowledged for facilitating and coordinating the continued availability of the data. This work is supported by grants from the National Science Foundation (NSF) Space Weather Program. This model validation study is supported by the Community Coordinated Modeling Center (CCMC) at the Goddard Space Flight Center. Data processing and research at MIT Haystack Observatory are supported by cooperative agreement AGS-1242204 between the U.S. National Science Foundation and the Massachusetts Institute of Technology. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Model output and observational data used for the study will be permanently posted at the CCMC website (http://ccmc.gsfc.nasa.gov) and provided as a resource for the space science community to use in the future.

References

Anderson, D. N., Buonsanto, M. J., Codrescu, M., Decker, D., Fesen, C. G., Fuller-Rowell, T. J., et al. (1998). Intercomparison of physical models and observations of the ionosphere. *Journal of Geophysical Research*, *103*(A2), 2179–2192. https://doi.org/10.1029/97JA02872

Araujo-Pradere, E. A., Fuller-Rowell, T. J., Spencer, P. S. J., & Minter, C. F. (2007). Differential validation of the US-TEC model. *Radio Science*, *42*, RS3016. https://doi.org/10.1029/2006RS003459

Bilitza, D., Altadill, D., Truhlik, V., Shubin, V., Galkin, I., Reinisch, B., & Huang, X. (2017). International Reference Ionosphere 2016: From ionospheric climate to real-time weather predictions. *Space Weather*, *15*, 418–429. https://doi.org/10.1002/2016SW001593

Bilitza, D., Altadill, D., Zhang, Y., Mertens, C., Truhlik, V., Richards, P., et al. (2014). The International Reference Ionosphere 2012—A model of international collaboration. *Journal of Space Weather and Space Climate*, *4*, 1–12. https://doi.org/10.1051/swsc/2014004

Burns, A. G., Wang, W., Wiltberger, M., Solomon, S. C., Spence, H., Killeen, T. L., et al. (2008). An event study to provide validation of TING and CMIT geomagnetic middle-latitude electron densities at the F_2 peak. *Journal of Geophysical Research*, *113*, A05310. https://doi.org/10.1029/2007JA012931

Chamberlin, P. C., Woods, T. N., & Eparvier, F. G. (2007). Flare Irradiance Spectral Model (FISM): Daily component algorithms and results. *Space Weather*, *5*, S07005. https://doi.org/10.1029/2007SW000316

Codrescu, M. V., Fuller-Rowell, T. J., Foster, J. C., Holt, J. M., & Cariglia, S. J. (2000). Electric field variability associated with the Millstone Hill electric field model. *Journal of Geophysical Research*, *105*(A3), 5265–5273. https://doi.org/10.1029/1999JA900463

Dmitriev, A. V., Suvorova, V., Klimenko, M. V., Klimenko, V. V., Ratovsky, K. G., Rakhmatulin, R. A., & Parkhomov, V. A. (2017). Predictable and unpredictable ionospheric disturbances during St. Patrick's Day magnetic storms of 2013 and 2015 and on 8–9 March 2008. *Journal of Geophysical Research: Space Physics*, *122*, 2398–2423. https://doi.org/10.1002/2016JA0232

Drob, D. P., Emmert, J. T., Crowley, G., Picone, J. M., Shepherd, G. G., Skinner, W., et al. (2008). An empirical model of the Earth's horizontal wind fields: HWM07. *Journal of Geophysical Research*, *113*, A12304. https://doi.org/10.1029/2008JA013668

Feltens, J., Angling, M., Jackson-Booth, N., Jakowski, N., Hoque, M., Hernández-Pajares, M., et al. (2011). Comparative testing of four ionospheric models driven with GPS measurements. *Radio Science*, *46*, RS0D12. https://doi.org/10.1029/2010RS004584

Fuller -Rowell, T. J., & Evans, D. S. (1987). Height-integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data. *Journal of Geophysical Research*, *92*(A7), 7606–7618.

Fuller-Rowell, T., Araujo-Pradere, E., & Codrescu, M. (2000). An empirical ionospheric storm-time correction model. *Advances in Space Research*, *25*(1), 139–148. https://doi.org/10.1016/S0273-1177(99)00911-4

Fuller-Rowell, T., Codrescu, M., Araujo-Pradere, E., & Kutiev, I. (1999). Progress in developing a storm-time ionospheric correction model. *Advances in Space Research*, *22*(6), 821–827. https://doi.org/10.1016/S0273-1177(98)00105-7

Fuller-Rowell, T. J., Codrescu, M. C., & Wilkinson, P. (2000). Quantitative modeling of the ionospheric response to geomagnetic activity. *Annales de Geophysique*, *18*(7), 766–781. https://doi.org/10.1007/s00585-000-0766-7

Fuller-Rowell, T. J., Rees, D., Quegan, S., Moffett, R. J., Codrescu, M. V., & Millward, G. H. (1996). A coupled thermosphere-ionosphere model (CTIM). In R. W. Schunk (Ed.), *Handbook of Ionospheric Models* (pp. 217–238). Logan, Utah: Utah State University.

Hagan, M. E., Burrage, M. D., Forbes, J. M., Hackney, J., Randel, W. J., & Zhang, X. (1999). GSWM-98: Results for migrating solar tides. *Journal of Geophysical Research*, *104*(A4), 6813–6828. https://doi.org/10.1029/1998JA900125

Hardy, D. A., Gussenhoven, M. S., & Holeman, E. (1985). A statistical model of auroral electron precipitation. *Journal of Geophysical Research*, *90*(A5), 4229–4248. https://doi.org/10.1029/JA090iA05p04229

Hedin, A. E. (1991). Extension of the MSIS thermospheric model into the middle and lower atmosphere. *Journal of Geophysical Research*, *96*(A2), 1159–1172. https://doi.org/10.1029/90JA02125

Hedin, A. E., Biondi, M. A., Burnside, R. G., Hernandez, G., Johnson, R. M., Killeen, T. L., et al. (1991). Revised global model of thermospheric winds using satellite and ground-based observations. *Journal of Geophysical Research*, *96*(A5), 7657–7688. https://doi.org/10.1029/91JA00251

Heppner, J. P., & Maynard, N. C. (1987). Empirical high latitude electric field models. *Journal of Geophysical Research*, *92*(A5), 4467–4489. https://doi.org/10.1029/JA092iA05p04467

Huba, J., Joyce, G., & Fedder, J. (2000). Sami2 is Another Model of the Ionosphere (SAMI2): A new low-latitude ionosphere model. *Journal of Geophysical Research*, *105*(A10), 23,035–23,053. https://doi.org/10.1029/2000JA000035

Huba, J. D., Joyce, G., & Krall, J. (2008). Three-dimensional equatorial spread F modeling. *Geophysical Research Letters*, *35*, L10102. https://doi.org/10.1029/2008GL033509

Huba, J. D., Sazykin, S., & Coster, A. (2017). SAMI3-RCM simulation of the 17 March 2015 geomagnetic storm. *Journal of Geophysical Research: Space Physics*, *122*, 1246–1257. https://doi.org/10.1002/2016JA023341

Iijima, T., & Potemra, T. A. (1976). Field-aligned currents in the dayside cusp observed by Triad. *Journal of Geophysical Research*, *8*, 5971–5979. https://doi.org/10.1029/JA081i034p05971

Maruyama, N., Richmond, A. D., Fuller-Rowell, T. J., Codrescu, M. V., Sazykin, S., Toffoletto, F. R., et al. (2005). Interaction between direct penetration and disturbance dynamo electric fields in the storm-time equatorial ionosphere. *Geophysical Research Letters*, *32*, L17105. https://doi.org/10.1029/2005GL023763

Millward, G. H., Müller-Wodrag, I. C. F., Aylward, A. D., Fuller-Rowell, T. J., Richmond, A. D., & Moffett, R. J. (2001). An investigation into the influence of tidal forcing on F region equatorial vertical ion drift using a global ionosphere-thermosphere model with coupled electrodynamics. *Journal of Geophysical Research*, *106*(A11), 24,733–24,744. https://doi.org/10.1029/2000JA000342

Namgaladze, A. A., Korenkov, Y. N., Klimenko, V. V., Karpov, I. V., Bessarab, F. S., Surotkin, V. A., et al. (1988). Global model of the thermosphere-ionosphere-protonosphere system. *Pure and Applied Geophysics*, *127*(2/3), 219–254. https://doi.org/10.1007/BF00879812

Namgaladze, A. A., Korenkov, Y. N., Klimenko, V. V., Karpov, I. V., Surotkin, V. A., & Naumova, N. M. (1991). Numerical modeling of the thermosphere-ionosphere-protonosphere system. *Journal of Atmospheric and Terrestrial Physics*, *53*(11/12), 1113–1124. https://doi.org/10.1016/0021-9169(91)90060-K

Newell, P. T., & Gjerloev, J. W. (2011). Substorm and magnetosphere characteristic scales inferred from the SuperMAG auroral electrojet indices. *Journal of Geophysical Research*, *116*, A12232. https://doi.org/10.1029/2011JA016936

Newell, P. T., Sotirelis, T., & Wing, S. (2009). Diffuse, monoenergetic, and broadband aurora: The global precipitation budget. *Journal of Geophysical Research*, *114*, A09207. https://doi.org/10.1029/2009JA014326

Nishioka, M., Tsugawa, T., Jin, H., & Ishii, M. (2017). A new ionospheric storm scale based on TEC and foF2 statistics. *Space Weather*, *15*, 228–239. https://doi.org/10.1002/2016SW001536

Orús, R., Hernández-Pajares, M., Juan, J. M., Sanz, J., & García-Fernández, M. (2002). Performance of different TEC models to provide GPS ionospheric corrections. *Journal of Atmospheric and Solar - Terrestrial Physics*, *64*(18), 2055–2062.

Orús, R., Hernández-Pajares, M., Juan, J. M., Sanz, J., & García-Fernández, M. (2003). Validation of the GPS TEC maps with TOPEX data. *Advances in Space Research*, *31*(3), 621–627.

Perlongo, N. J., Ridley, A. J., Cnossen, I., & Wu, C. (2018). A year-long comparison of GPS TEC and global ionosphere-thermosphere models. *Journal of Geophysical Research: Space Physics*, *123*, 1410–1428. https://doi.org/10.1002/2017JA024411

Picone, J. M., Hedin, A. E., Drob, D. P., & Aikin, A. C. (2002). NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research*, *107*(A12), 1468. https://doi.org/10.1029/2002JA009430

Prokhorov, B. E., Förster, M., Lesur, V., Namgaladze, A. A., Holschneider, M., & Stolle, C. (2018). Modeling of the ionospheric current system and calculating its contribution to the Earth's magnetic field. Magnetic fields. In *Magnetic fields the solar system, astrophysics and space science library* (Vol. 448, pp. 263–292). Cham: Springer. https://doi.org/10.1007/978-3-319-64292-5_10

Rastätter, L., Shim, J. S., Kuznetsova, M. M., Kilcommons, L. M., Knipp, D. J., Codrescu, M., et al. (2016). GEM-CEDAR challenge: Poynting flux at DMSP and modeled Joule heat. *Space Weather*, *14*, 113–135. https://doi.org/10.1002/2015SW001238

Reinisch, B., & Galkin, I. (2011). Global Ionospheric Radio Observatory (GIRO). *Earth, Planets and Space*, *63*(4), 377–381. https://doi.org/10.5047/eps.2011.03.001

Richards, P. G., Fennelly, J. A., & Torr, D. G. (1994). EUVAC: A solar EUV flux model for aeronomic calculations. *Journal of Geophysical Research*, *99*(A5), 8981–8992. https://doi.org/10.1029/94JA00518

Richmond, A. D., Ridley, E. C., & Roble, R. G. (1992). A thermosphere/ionosphere general circulation model with coupled electrodynamics. *Geophysical Research Letters*, *19*(6), 601–604. https://doi.org/10.1029/92GL00401

Rideout, W., & Coster, A. (2006). Automated GPS processing for global total electron content data. *GPS Solutions*, *10*(3), 219–228. https://doi.org/10.1007/s10291-006-0029-5

Ridley, A. J., Deng, Y., & Toth, G. (2006). The global ionosphere-thermosphere model. *Journal of Atmospheric and Solar - Terrestrial Physics*, *68*(8), 839–864.

Roble, R. G., & Ridley, E. C. (1987). An auroral model for the NCAR thermospheric general circulation model (TGCM). *Annales Geophysicae*, *5A*, 369–382.

Roble, R. G., Ridley, E. C., Richmond, A. D., & Dickinson, R. E. (1988). A coupled thermosphere/ionosphere general circulation model. *Geophysical Research Letters*, *15*(12), 1325–1328. https://doi.org/10.1029/GL015i012p01325

Scherliess, L., & Fejer, B. G. (1999). Radar and satellite global equatorial *F*-region vertical drift model. *Journal of Geophysical Research*, *104*(A4), 6829–6842. https://doi.org/10.1029/1999JA900025

Scherliess, L., Schunk, R. W., Sojka, J. J., & Thompson, D. C. (2004). Development of a physics-based reduced state Kalman filter for the ionosphere. *Radio Science*, *39*, RS1S04. https://doi.org/10.1029/2002RS002797

Scherliess, L., Schunk, R. W., Sojka, J. J., Thompson, D. C., & Zhu, L. (2006). Utah State University global assimilation of ionospheric measurements Gauss-Markov Kalman filter model of the ionosphere: Model description and validation. *Journal of Geophysical Research*, *111*, A11315. https://doi.org/10.1029/2006JA011712

Schunk, R. W., Scherliess, L., & Sojka, J. J. (2002). Ionospheric specification and forecast modeling. *Journal of Spacecraft and Rockets*, *39*(2), 314–324.

Schunk, R. W., Scherliess, L., Sojka, J. J., Thompson, D. C., Anderson, D. N., Codrescu, M., et al. (2004). Global Assimilation of Ionospheric Measurements (GAIM). *Radio Science*, *39*, RS1S02. https://doi.org/10.1029/2002RS002794

Schunk, R. W., Sojka, J. J., & Eccles, J. V. (1997). Expanded capabilities for the ionospheric forecast model, Rep. AFRL-VS-HA-TR-98-0001, Air Force Res. Lab., Hanscom Air Force Base, Mass., December.

Shim, J. S., Jee, G., & Scherliess, L. (2017). Climatology of plasmaspheric total electron content obtained from Jason 1 satellite. *Journal of Geophysical Research: Space Physics*, *122*, 1611–1623. https://doi.org/10.1002/2016JA023444

Shim, J. S., Kuznetsova, M., Rastätter, L., Bilitza, D., Butala, M., Codrescu, M., et al. (2014). Systematic evaluation of Ionosphere/Thermosphere (IT) models: CEDAR Electrodynamics Thermosphere Ionosphere (ETI) challenge (2009-2010). In *Modeling the ionosphere-thermosphere system*, *Geophysical Monograph Series* (pp. 145–160). Washington, DC: American Geophysical Union.

Shim, J. S., Kuznetsova, M., Rastätter, L., Bilitza, D., Butala, M., Codrescu, M., et al. (2012). 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*, *10*, S10004. https://doi.org/10.1029/2012SW000851

Shim, J. S., Kuznetsova, M., Rastätter, L., Hesse, M., Bilitza, D., Butala, M., et al. (2011). CEDAR Electrodynamics Thermosphere Ionosphere (ETI) challenge for systematic assessment of ionosphere/thermosphere models: NmF2, hmF2, and vertical drift using ground-based observations. *Space Weather*, *9*, S12003. https://doi.org/10.1029/2011SW000727

Shim, J. S., Rastätter, L., Kuznetsova, M., Bilitza, D., Codrescu, M., Coster, A. J., et al. (2017). CEDAR-GEM challenge for systematic assessment of ionosphere/thermosphere models in predicting TEC during the 2006 December storm event. *Space Weather*, *15*, 1238–1256. https://doi.org/10.1002/2017SW001649

Solomon, S. C., Burns, A. G., Emery, B. A., Mlynczak, M. G., Qian, L., Wang, W., et al. (2012). Modeling studies of the impact of high-speed streams and co-rotating interaction regions on the thermosphere-ionosphere. *Journal of Geophysical Research*, *117*, A00L11. https://doi.org/10.1029/2011JA017417

Solomon, S. C., Qian, L., & Mannucci, A. J. (2018). Ionospheric electron content during solar cycle 23. *Journal of Geophysical Research: Space Physics*, *123*, 5223. https://doi.org/10.1029/2018JA025464

Webb, P. A., Kuznetsova, M. M., Hesse, M., Rastaetter, L., & Chulaki, A. (2009). Ionosphere-thermosphere models at the Community Coordinated Modeling Center. *Radio Science*, *44*, RS0A34. https://doi.org/10.1029/2008RS004108

Weimer, D. R. (2005). Improved ionospheric electrodynamic models and application to calculating Joule heating rates. *Journal of Geophysical Research*, *110*, A05306. https://doi.org/10.1029/2004JA010884

Yue, X., Wan, W., Liu, L., Liu, J., Zhang, S., Schreiner, W. S., et al. (2016). Mapping the conjugate and corotating storm-enhanced density during 17 March 2013 storm through data assimilation. *Journal of Geophysical Research: Space Physics*, *121*, 12,202–12,210. https://doi.org/10.1002/2016JA023038

Zhang, S.-R., Zhang, Y., Wang, W., & Verkhoglyadova, O. P. (2017). Geospace system responses to the St. Patrick's Day storms in 2013 and 2015. *Journal of Geophysical Research: Space Physics*, *122*, 6901–6906. https://doi.org/10.1002/2017JA024232

Zhu, L., Schunk, R. W., Jee, G., Scherliess, L., Sojka, J. J., & Thompson, D. C. (2006). Validation study of the ionosphere forecast model using the TOPEX total electron content measurements. *Radio Science*, *41*, RS5S11. https://doi.org/10.1029/2005RS003336