CEDAR Electrodynamics Thermosphere Ionosphere (ETI) Challenge for systematic assessment of ionosphere/thermosphere models: NmF2, hmF2, and vertical drift using ground-based observations

J. S. Shim,[1] M. Kuznetsova,[2] L. Rastätter,[2] M. Hesse,[2] D. Bilitza,[2] M. Butala,[3] M. Codrescu,[4] B. Emery,[5] B. Foster,[5] T. Fuller-Rowell,[4] J. Huba,[6] A. J. Mannucci,[3] X. Pi,[3] A. Ridley,[7] L. Scherliess,[8] R. W. Schunk,[8] P. Stephens,[3] D. C. Thompson,[9] L. Zhu,[8] D. Anderson,[10] J. L. Chau,[11] J. J. Sojka,[8] and B. Rideout[12]

Received 2 September 2011; revised 22 October 2011; accepted 27 October 2011; published 31 December 2011.

[1] Objective quantification of model performance based on metrics helps us evaluate the current state of space physics modeling capability, address differences among various modeling approaches, and track model improvements over time. The Coupling, Energetics, and Dynamics of Atmospheric Regions (CEDAR) Electrodynamics Thermosphere Ionosphere (ETI) Challenge was initiated in 2009 to assess accuracy of various ionosphere/thermosphere models in reproducing ionosphere and thermosphere parameters. A total of nine events and five physical parameters were selected to compare between model outputs and observations. The nine events included two strong and one moderate geomagnetic storm events from GEM Challenge events and three moderate storms and three quiet periods from the first half of the International Polar Year (IPY) campaign, which lasted for 2 years, from March 2007 to March 2009. The five physical parameters selected were NmF2 and hmF2 from ISRs and LEO satellites such as CHAMP and COSMIC, vertical drifts at Jicamarca, and electron and neutral densities along the track of the CHAMP satellite. For this study, four different metrics and up to 10 models were used. In this paper, we focus on preliminary results of the study using ground-based measurements, which include NmF2 and hmF2 from Incoherent Scatter Radars (ISRs), and vertical drifts at Jicamarca. The results show that the model performance strongly depends on the type of metrics used, and thus no model is ranked top for all used metrics. The analysis further indicates that performance of the model also varies with latitude and geomagnetic activity level.

Citation: Shim, J. S., 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, doi:10.1029/2011SW000727.

[1]Goddard Planetary Heliophysics Institute, University of Maryland Baltimore County, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.
[4]Space Weather Prediction Center, NOAA, Boulder, Colorado, USA.
[5]High Altitude Observatory, National Center for Atmospheric Research, Boulder, Colorado, USA.
[6]Plasma Physics Division, Naval Research Laboratory, Washington, D. C., USA.
[7]Space Physics Research Laboratory, University of Michigan, Ann Arbor, Michigan, USA.
[8]Center for Atmospheric and Space Sciences, Utah State University, Logan, Utah, USA.
[9]Air Force Research Laboratory, Albuquerque, New Mexico, USA.
[10]Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, Colorado, USA.
[11]Radio Observatorio de Jicamarca, Instituto Geofisico del Peru, Lima, Peru.
[12]Haystack Observatory, Massachusetts Institute of Technology, Westford, Massachusetts, USA.

1. Introduction

[2] Knowledge of the ionospheric parameters, such as NmF2, hmF2 and plasma drift velocity, is critical for specification and forecasting of the ionosphere. In order to not only specify the ionosphere but also understand the physics behind the dynamics of the ionosphere, many ionospheric models have been developed over the past several decades. The ionospheric models can be broadly divided into three main categories: empirical, physics-based theoretical, and data assimilation models. Empirical models provide an average behavior of the ionosphere. Although empirical models are based on limited observational data, they are widely used because of their relative simplicity. Over the last three decades, first-principles physics-based models of the ionosphere have been developed [*Schunk et al.*, 2002], and they require numerous input parameters associated with coupling processes to the thermosphere, plasmasphere and magnetosphere. Therefore, the accuracy of the ionospheric model outputs strongly depends on the accuracy of the input drivers. Recently, data assimilation methods have been used for ionosphere monitoring and forecasting that can be attributed to more available ionospheric observations [*Richmond and Kamide*, 1988; *Schunk et al.*, 2002; *Scherliess et al.*, 2004; *Wang et al.*, 2004]. However, each of them has their own limitations for reproducing ionospheric weather.

[3] In order to better understand strengths and limitations of ionospheric models and to further improve the capability for specification and forecasting of the ionosphere, it is crucial to evaluate the models systematically and quantitatively. There have been many model validation studies, not only of the ionosphere [*Anderson et al.*, 1998], but also of the other regions of geospace [*Spence et al.*, 2004; *Pulkkinen et al.*, 2010, 2011; *Rastätter et al.*, 2011]. However, a more systematic quantitative metric study is needed for assessing long-term model improvements objectively. We performed for the first time metric studies for various Ionosphere/Thermosphere models including empirical, physic-based, coupled and data assimilation models. This work constituted a main part of the CEDAR (Coupling, Energetics, and Dynamics of Atmospheric Regions) Electrodynamics Thermosphere Ionosphere (ETI) Challenge. The Challenge was initiated in order to evaluate the current state of ionosphere models and to assess their improvements over time. For the Challenge, nine time periods (three events from four GEM events, three moderate storms and three quiet periods from the March 2007 to March 2008 time-frame, which is the first half of the International Polar Year (IPY) from March 2007 to March 2009) were selected to compare model outputs with observations. Electron and neutral densities along the CHAMP (Challenging Minisatellite Payload) trajectories, NmF2 and hmF2 from ISRs, and vertical drift at Jicamarca were chosen for the physical parameters to be tested. Model outputs and observational data used for the Challenge will be permanently posted on the Community Coordinated Modeling Center (CCMC) website for the space science communities to use. The CCMC is supporting the Challenge using their experience with the GEM community and the metric tools available at the CCMC.

[4] In this paper, we present the preliminary results of the Challenge including the comparison between model results and the measurements of the NmF2 and hmF2 from four ISR stations and vertical drift at Jicamarca for the nine selected time periods. The results from the comparison of space-based measurements and model simulations (electron density and neutral density along the CHAMP trajectories, and NmF2 and hmF2 from the CHAMP and COSMIC satellites) will be presented in a companion paper, which will be followed by the next report on climatological study covering one year of the IPY.

2. Setup of the Challenge

[5] In order to evaluate the performance of ionospheric models for different geomagnetic activity levels, nine time periods were selected. Three events were chosen among four events selected for the GEM (Geospace Environment Modeling) Challenge [*Pulkkinen et al.*, 2010, 2011; *Rastätter et al.*, 2011] (see Table 1). The events were sorted into three levels of geomagnetic activity according to the magnitude of Kp index. Strong storm events were defined when $Kp\_max \geq 7$ during the time interval. Moderate storm events and quiet periods corresponded to $4 \leq Kp\_max < 7$ and $Kp\_max < 4$, respectively. One of the three GEM events E.2001.243 was sorted as a moderate storm event, and the other two GEM events (E.2006.348 and E.2005.243) as strong storm events. We selected three moderate storm events and three geomagnetically quiet periods during IPY from March 2007 to March 2008. Figure 1 shows Kp values for three GEM events and one moderate storm and one quiet period. Each event lasted no longer than three consecutive days.

[6] For each event, we compared the model outputs with measurements for (1) vertical drift at Jicamarca (11.95°S, 283.13°E), (2) NmF2, and (3) hmF2 from Incoherent Scatter Radar (ISR) Stations (Millstone Hill, EISCAT Svalbard, Poker Flat, and Sondrestrom). Not all models contributed to all nine events, as indicated below.

3. Models and Model Output Submissions

[7] For the study, we used the model outputs submitted by model developers through the CCMC online submission interface developed for this and other model validation studies. We also used model outputs generated by the CCMC using ionosphere-thermosphere (IT) models hosted at the CCMC [*Webb et al.*, 2009]. Table 2 shows the submissions of outputs using various models, which include empirical and physics based ionosphere models, IT coupled models, and data assimilation models. Multiple output submissions from one model using different input drivers and/or different boundary conditions were distinguished by unique model setting identifier, for example,

Table 1. Events Studied in the Coupling, Energetics, and Dynamics of Atmospheric Regions (CEDAR) Electrodynamics Thermosphere Ionosphere (ETI) Challenge

	Event Name	Date (DOY) and Time	F10.7	Kp_max
GEM events	E.2006.348	2006/12/14 (doy 348) 12:00 UT to 12/16 (doy 350) 00:00 UT	91	8
	E.2001.243	2001/08/31 (doy 243) 00:00 UT to 09/01 (doy 244) 00:00 UT	192	4
	E.2005.243	2005/08/31 (doy 243) 10:00 UT to 09/01 (doy 244) 12:00 UT	86	7
Moderate storms	E.2007.091	2007/04/01 (doy 091) 00:00 UT to 04/02 (doy 092) 12:00 UT	72	5
	E.2007.142	2007/05/22 (doy 142) 12:00 UT to 05/25 (doy 145) 00:00 UT	74	6
	E.2008.059	2008/02/28 (doy 059) 12:00 UT to 03/01 (doy 061) 12:00 UT	69	5
Quiet periods	E.2007.079	2007/03/20 (doy 079) 00:00 UT to 03/22 (doy 081) 00:00 UT	72	1
	E.2007.190	2007/07/09 (doy 190) 00:00 UT to 07/10 (doy 191) 00:00 UT	80	0
	E.2007.341	2007/12/07 (doy 341) 00:00 UT to 12/09 (doy 343) 00:00 UT	80	1

1_SAMI3_HWM07 and 1_SAMI3_HWM93, 1_TIE-GCM and 2_TIE-GCM, and 1_GITM and 3_GITM. The model setting identifier marked with asterisk in Table 2 denotes that model results are submitted by the CCMC.

[8] For the comparison of the vertical drift, six submissions of model simulations from four models, IRI, SAMI3, GITM and TIE-GCM were analyzed. Ten submissions using eight models (IRI, SAMI3, USU-IFM, CTIPe, GITM, TIE-GCM, JPL-GAIM and USU-GAIM) were included for NmF2 and hmF2 comparison. Brief descriptions of the models (and the submissions of model outputs) are provided in sections 3.1–3.8.

3.1. IRI (1_IRI)

[9] The International Reference Ionosphere (IRI) based on all available data sources [*Bilitza*, 1990, 2001, 2004; *Bilitza and Reinisch*, 2008] is the most comprehensive and widely used empirical model for the ionosphere. For a given location, time, and date, IRI provides monthly averages of electron density, electron temperature, ion temperature, and ion composition (O^+, H^+, He^+, N^+, NO^+, O_2^+, and cluster ions) in the altitude range from 50 km to 2000 km. Additional parameters given by IRI include the Total Electron Content (TEC), the occurrence probability for Spread-F and also for the F1-region, and the equatorial vertical ion drift. The major data sources are the worldwide network of ionosondes, incoherent scatter radars, topside sounders, and in situ instruments on several satellites and rockets.

3.2. SAMI3 (1_SAMI3_HWM07 and 1_SAMI3_HWM93)

[10] SAMI3 (Sami3 is Also a Model of the Ionosphere) is a three-dimensional global ionospheric model based on the two-dimensional model SAMI2 [*Huba et al.*, 2000, 2008]. SAMI3 calculates the plasma and chemical evolution of seven ion species (H^+, He^+, N^+, O^+, N_2^+, NO^+ and O_2^+) in the altitude range 85 km to 20,000 km. The ion temperature equation is solved for three ion species (H^+, He^+, and O^+) as well as the electron temperature equation. An offset, tilted dipole geomagnetic field is used, and the plasma is modeled from hemisphere to hemisphere. High altitude boundary conditions are not needed, since a complete ionospheric flux tube is modeled. The neutral composition and temperature are provided by NRLMSISE00 model [*Picone et al.*, 2002], and the neutral winds are obtained from the HWM models [*Hedin et al.*, 1991; *Drob et al.*, 2008]. SAMI3 uses a unique, nonorthogonal, nonuniform, fixed grid. The grid is designed to optimize the numerical mesh so that the spatial resolution decreases with increasing altitude.

[11] Two submissions for the vertical drifts using SAMI3 were used (see Table 2). 1_SAMI3_HWM07 and 1_SAMI3_HWM93 are SAMI3 runs with the neutral wind model HWM07 and HWM93, respectively. The two SAMI3 simulations presented here are for two quiet events (E.2007.079 and E.2007.190) and for two moderate storms (E.2007.091 and E.2007.142). In addition to the two submissions for the vertical drifts made by model developer, we used NmF2 and hmF2 data from 1_SAMI3_HWM93 submitted by the CCMC.

3.3. USU-IFM (1_USU-IFM)

[12] Utah State University (USU)-Ionosphere Forecast Model (IFM) is a three-dimensional, high-resolution, multi-ion model of the global ionosphere [*Schunk et al.*, 1997] that is based on the USU Time-Dependent Ionosphere Model (TDIM) [*Schunk*, 1988; *Sojka*, 1989]. The model covers the altitude range from 90 to 1600 km and all latitudes and longitudes. The spatial and temporal resolutions of the IFM are flexible. For this study, 3° × 7.5° latitude-longitude grid was selected. IFM is based on a numerical solution of the continuity, momentum, and energy equations of multiple ion species. The equations are solved along magnetic field lines for individual convecting flux tubes of plasma, and the 3-D nature of the model is obtained by following a large number of plasma flux tubes. The model accounts for the displacement between the geomagnetic and geographic poles. The IFM model uses F10.7, 90 days average F10.7, daily Ap and eight 3-h Kp indices. The IFM also uses empirical inputs for the neutral atmosphere and magnetosphere parameters needed by the model, e.g., neutral composition, neutral wind, electric field, auroral precipitation, solar EUV, and resonantly scattered radiation. The outputs of the IFM include 3-D distributions of electrons and various ion species, electron and ion temperatures, TEC, NmF2, hmF2, and other auxiliary ionospheric plasma parameters.

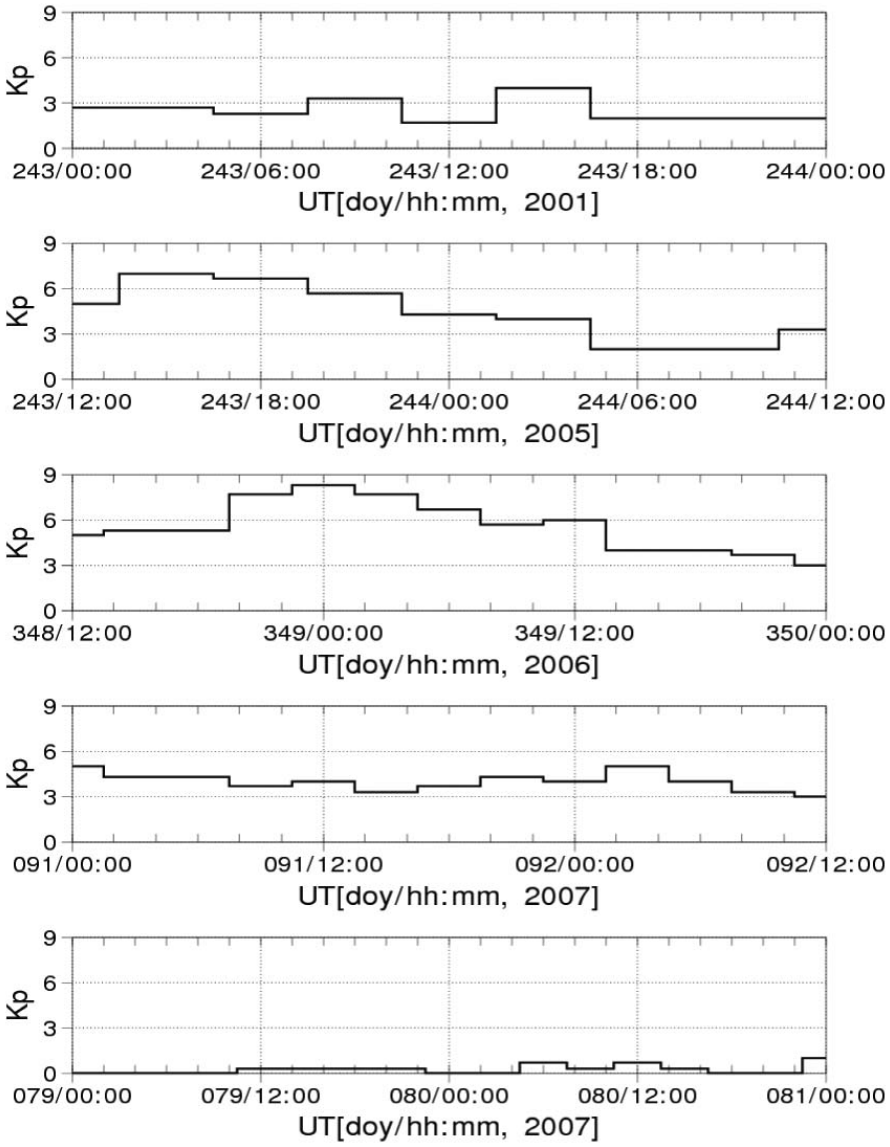

Figure 1. Kp values for five of the studied events.

IFM is used as a first-principle physical model in the USU Gauss-Markov Kalman filter model, USU-GAIM, which is described in Section 3.8.

3.4. CTIPe (1_CTIPE)

[13] The Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics (CTIPe) model [*Codrescu et al.*, 2000; *Millward et al.*, 2001] is a physics-based nonlinear, coupled thermosphere-ionosphere-plasmasphere numerical model. The model consists of four distinct components: global thermosphere, high-latitude ionosphere, mid and low-latitude ionosphere/plasmasphere, and electrodynamical calculation of the global dynamo electric field. All four components of the CTIPe are run concurrently and are fully coupled with respect to energy, momentum, and continuity. The thermospheric code simulates the time-dependent global structure of the wind vector, temperature, and density of the neutral thermosphere. The spatial resolutions are 2° in latitude and 18° in longitude, and the vertical direction is divided into 15 levels in logarithm of pressure extending from a lower boundary of 1 Pa at 80 km altitude to a upper boundary of about 3×10^{-7} Pa, which corresponds to about 400∼500 km altitude. The high-latitude ionosphere convection model calculates field-aligned ion velocity components from the field-aligned momentum equation. The ionosphere is computed self-consistently with the thermosphere pole-ward of 23° latitude in both hemispheres. The plasmasphere model solves coupled equations of continuity, momentum and energy balance along many closed flux tubes

Table 2. Models Submitted for the CEDAR Challenge

Model Setting ID	Model/Version	Resolution (Latitude × Longitude × Altitude)
	Empirical Model	
1_IRI[a]	IRI-2007, empirical ionospheric model	(50 km < alt < 2000 km)
	Physics-Based Model	
1_SAMI3_HWM07	SAMI3 with the neutral wind model HWM07	120 × 90 × 160 (90 km < alt < 20,000 km)
1_SAMI3_ HWM93 (1_SAMI3_ HWM93[a]: only for NmF2 and hmF2 at Millstone Hill)	SAMI3 with the neutral wind model HWM93	120 × 90 × 160 (90 km < alt < 20,000 km)
1_USU-IFM[a]	IFM driven by F10.7, Kp and empirical inputs for the thermosphere parameters	60 × 49 × 73 (90 km < alt < 1600 km)
	Coupled Model	
1_CTIPE[a]	CTIPe driven by *Weimer* [2005]	91 × 20 × 15 (∼90 km < alt < 500 km)
1_GITM	GITM	25 × 50 ×13
3_GITM	GITM with different O^+ reaction rates and collision frequencies, lower boundary conditions, equatorial electrojet, and potential solver from those for 1_GITM	25 × 50 ×13
1_TIE-GCM[a]	TIE-GCM1.93 driven by *Heelis et al.* [1982]	36 × 72 × 29 (∼90 km < alt < 500 km)
2_TIE-GCM	TIE-GCM1.94 driven by *Weimer* [2005] with dynamic critical colatitudes	36 × 72 × 29 (∼90 km < alt < 500 km)
	Data Assimilation Model	
1_JPL-GAIM	USC/JPL GAIM with ground-based GPS data (−55° < lat < 55°) and COSMIC data	60 × 36 × every 40 km altitude
1_USU-GAIM[a]	USU-GAIM23 with GPS TEC observations from up to 400 ground stations (−60° < lat < 60°)	44 × 24 × 83 (90 km < alt < 1400 km)

[a]Model results are submitted by the CCMC using the models hosted at the CCMC. Different model setups are referred as different model setting identification number.

concurrently. The magnetospheric input to the model is based on the statistical models of auroral precipitation and electric fields described by *Fuller-Rowell and Evans* [1987] and *Weimer* [2005], respectively.

3.5. GITM (1_GITM and 3_GITM)

[14] The Global Ionosphere Thermosphere Model (GITM) is a three-dimensional spherical code that solves for the coupled ionosphere and thermosphere [*Ridley et al.*, 2006]. GITM explicitly solves for the neutral densities of O, O_2, $N(^2D)$, $N(^2P)$, $N(^4S)$, N_2, NO, H, and He; and ion species O^+ (^4S), O^+ (^2D), O^+ (^2P), O_2^+, N^+, N_2^+ NO^+, H^+, and He^+. Unlike other global ionosphere-thermosphere models, GITM uses an altitude grid instead of a pressure grid. The number of grid points in a simulation can be specified by users. The model also can be run in one dimension, in which a single latitude and longitude are modeled and horizontal transport and gradients are ignored. GITM does not assume hydrostatic equilibrium so that significant vertical flows can develop self-consistently. GITM is coupled to a large number of models of the high-latitude ionospheric electrodynamics. It can also be coupled with a global magnetohydrodynamic (MHD) model [*Powell et al.*, 1999] of the magnetosphere, since GITM is part of University of Michigan's Space Weather Modeling Framework [*Tóth et al.*, 2005]. This allows investigation of the coupling between the thermosphere, ionosphere and the magnetosphere systems [e.g., *Ridley et al.*, 2003].

[15] For the study, two model output submissions (1_GITM and 3_GITM) were made. They differed in the reaction rates of $O^+ + O_2/NO/N_2 \rightarrow O + O_2^+/NO^+/N_2^+$, the equatorial electrojet and potential solver, lower boundary conditions, and the collision frequencies for O^+ to neutral species.

3.6. TIE-GCM (1_TIE-GCM and 2_TIE-GCM)

[16] The NCAR (National Center for Atmospheric Research) TIE-GCM (Thermosphere-Ionosphere- Electrodynamics General Circulation Model) [*Roble et al.*, 1988; *Richmond et al.*, 1992] is a comprehensive, first-principles, three-dimensional, nonlinear representation of the coupled thermosphere and ionosphere system that includes a self-consistent solution of the low-latitude electric field. The model solves the three-dimensional momentum, energy and continuity equations for neutral and ion species at each time step, using a semi-implicit, fourth-order, centered finite difference scheme, on each pressure surface in a staggered vertical grid. It has 29 constant-pressure levels in the vertical, extending from approximately 97 km to 500 km in intervals of one-half scale height, and a 5° × 5° latitude-longitude grid, in its base configuration. The TIE-GCM uses an imposed electric

field model such as the *Weimer* [2005] or *Heelis et al.* [1982] electric potential models at high magnetic latitudes above a critical magnetic co-latitude, and a dynamo solution equatorwards of a second critical colatitude which are set to constants or to dynamic values as a function of magnetic activity. At the lower boundary, atmospheric tides are specified using the Global Scale Wave Model (GSWM) [*Hagan et al.*, 1999].

[17] Two versions of model run, 1_TIE-GCM and 2_TIE-GCM, are submitted by using TIE-GCM1.93 driven by *Heelis et al.* [1982] with constant critical colatitudes and TIE-GCM1.94 driven by *Weimer* [2005] with dynamic critical colatitudes, respectively.

3.7. JPL/USC-GAIM (1_JPL-GAIM)

[18] Global Assimilative Ionospheric Model (GAIM) developed by The University of Southern California (USC) and the Jet Propulsion Laboratory (JPL) is a physics-based 3-D data assimilation model [*Wang et al.*, 2004; *Pi et al.*, 2003]. The JPL/USC-GAIM solves the conservation of mass and momentum equations for plasma, which account for production, loss, and transport of the major ionization species in the F region (O^+). The JPL/USC-GAIM numerically solves for the ion and electron density state and key drivers such as equatorial electrodynamics, neutral winds, and ion production terms using the following well-developed empirical models for various input parameters: thermospheric densities and winds (MSIS [*Hedin*, 1991], HWM [*Hedin et al.*, 1996]), solar EUV [*Tobiska*, 1991], electric fields [*Fejer et al.*, 1991; *Heppner and Maynard*, 1987; *Scherliess and Fejer*, 1999], and electron energy precipitation flux [*Fuller-Rowell and Evans*, 1987]. It uses both four-dimensional variational analysis (4DVAR) and Kalman filter techniques to ingest multiple data sources, which include line-of-sight TEC measurements made from ground-based GPS receiver networks and space-borne GPS receivers, satellite UV limb scans, and ionosondes.

[19] Submission 1_JPL_GAIM for the study was generated by assimilating ground based GPS TEC data (30 s cadence) and COSMIC TEC data (10 s cadence). The ground data from about 200 stations were only assimilated if they came from a site located between $\pm 55°$ geomagnetic latitude. The spatial resolutions of the model are 10° in longitude, 3° in latitude, and 40 km in altitude. The geomagnetic field used in the model was based on the International Geomagnetic Reference Field (IGRF), which has high fidelity below $\pm 55°$ geomagnetic latitude.

3.8. USU-GAIM

[20] Physics-based data assimilation model of the ionosphere, Global Assimilation of Ionospheric Measurements (GAIM) was developed at Utah State University [*Schunk et al.*, 2004; *Scherliess et al.*, 2004, 2006]. This model is a Gauss-Markov Kalman Filter (GMKF) model, which uses a physics-based model of the ionosphere and a Kalman filter as a basis for assimilating a diverse set of real-time (or near real-time) measurements. The physics-based model is the USU-IFM (see section 3.3). The GMKF model assimilates bottom-side N_e profiles from a variable number of ionosondes, slant TEC from a variable number of ground GPS/TEC stations, in situ N_e from four DMSP satellites, and line-of-sight UV emissions measured by satellites. With the GMKF model, the ionospheric densities obtained from the IFM constitute a background ionospheric density field on which perturbations are superimposed based on the available data sources and their errors. The primary output from the USU GAIM model is a time-dependent 3-dimensional global electron density distribution. The USU GAIM also provides Vertical Total Electron Content (TEC) obtained from the leveled and bias corrected slant TEC values, which are assimilated by the model.

[21] The versions of the USU GAIM currently hosted at the CCMC only uses GPS observations distributed within $\pm 60°$ geographic latitude, therefore, the plasma densities at high latitude are provided by the IFM without assimilating any data. We used the USU-GAIM results generated by the CCMC for this study.

4. Data

[22] The vertical drifts derived from the 150-km echoes measured with the Jicamarca Unattended Long-Term studies of the Ionosphere and Atmosphere (JULIA) Coherent Scatter Radar were used as ground truth [*Hysell et al.*, 1997; *Chau and Woodman*, 2004]. The JULIA radar located at the Jicamarca, Peru observes daytime vertical $E \times B$ velocity. The data are recorded every five minutes during the daytime (07:00–18:00 LT). The average of the error in vertical drift for E.2006.348 strong storm event is about 0.7 m/s. For the other events (E.2001.243, E.2007.142, and E.2008.059), the average value of the error does not exceed 1 m/s. The vertical drifts from JULIA are only available during the day from about 0700 to 1800 LT (1200–2300 UT). In order to increase data coverage, we also used vertical drifts obtained from Jicamarca-Piura ΔH magnetometer measurements [*Anderson et al.*, 2004] also during the daytime only for four events for which there were no JULIA data available. It is well known that the difference in the magnitudes of the horizontal component (H) between a magnetometer placed directly on the magnetic equator and one displaced a few degrees away provides the vertical drift in the F region ionosphere. We used modeled and observed vertical drifts with 5-min temporal resolution.

[23] NmF2 and hmF2 data for this study were provided by Millstone Hill (42.62°N, 288.51°E; 54°N), EISCAT Svalbard (78.09°N, 16.02°E; 74°N), Poker Flat (65.13°N, 212.53°E; 65°N), and Sondrestrom (66.99°N, 309.05°E; 75°N) ISR stations (geographic latitude, longitude and geomagnetic latitude of each station are in the parenthesis). For the study, we used Gridded ISR data over Millstone Hill that were converted to overhead (vertical line-of-sight) fits from the various directions every 15 min using a specified height

grid from 100 to 548 km (http://madrigal.haystack.mit.edu/madrigal/) [*Holt et al.*, 2002]. Sojka provided the data from the three other ISRs that were used in Ionospheric Challenges of the IPY (International Polar Year) [*Sojka et al.*, 2007]. The cadence of NmF2, hmF2 from the four ISRs was about 15 min. To calculate the accuracy of the model output, the modeled NmF2 and hmF2 with different output frequency were averaged over 15 min.

5. Metrics

[24] To quantify the model performance, four different metrics were used. The term 'metric' refers to functions that produce one real number for each set of observed and modeled physical parameters.

5.1. Root-Mean Square (RMS) Difference

[25] One of the most meaningful and widely used ways to evaluate model performance is to calculate root-mean square difference between the model estimates and observations defined as

$$RMS = \sqrt{\frac{\sum (x_{obs} - x_{\mathrm{mod}})^2}{N}}$$

where x_{obs} and x_{mod} are observed and modeled values, respectively. RMS difference has the same unit as observed and modeled values, x_{obs} and x_{mod}. Perfect model predictions have RMS differences of 0. Therefore, the closer the RMS error is to 0, the more accurate the model is.

5.2. Prediction Efficiency

[26] Prediction efficiency, one of the skill scores against the mean of observations, is also commonly used to describe performance of models:

$$PE = 1 - \frac{RMS_{\mathrm{mod}}}{RMS_{ref}} = 1 - \sqrt{\frac{\sum (x_{obs} - x_{\mathrm{mod}})^2 / N}{\sum (x_{obs} - \langle x_{obs} \rangle)^2 / N}}$$

where x_{obs} and x_{mod} are again observed and modeled values, respectively, and $\langle x_{obs} \rangle$ is the mean value of the observed measurements. In this study, the mean value of the observations $\langle x_{obs} \rangle$ was considered as a reference model instead of using any empirical model. The prediction efficiency ranges from negative infinity to 1. A prediction efficiency of 1 implies perfect model performance, while a prediction efficiency of 0 means that the model performance is as accurate as the mean of the observed data. Negative value indicates that the observed mean is a better predictor than the model. In order to take account of local time dependence of the vertical drift, NmF2 and hmF2, the daytime prediction efficiency (0600–1800 LT) and the nighttime prediction efficiency (1800–0600 LT) were obtained separately using daytime and nighttime mean values of observations during any given event.

5.3. Ratio of the Maximum Change in Amplitudes and Ratio of the Maximum Amplitudes

[27] The root mean square error and prediction efficiency measure how well observed data and modeled values are correlated with each other. Metrics based on ratio are used to quantify the model capability to produce peak values or short-term variations during a certain period of time, even though performance of model is poor in term of the RMS error and/or prediction efficiency. Selected two types of ratio were the ratio of the maximum change (difference between maximum and minimum values; max - min) and the ratio of the maximum (max) values of models to those of observations during a certain time interval:

$$ratio(\max - \min) = \frac{(x_{\mathrm{mod}})_{\max} - (x_{\mathrm{mod}})_{\min}}{(x_{obs})_{\max} - (x_{obs})_{\min}},$$

$$ratio(\max) = \frac{(x_{\mathrm{mod}})_{\max}}{(x_{obs})_{\max}}$$

where $(x_{obs})_{max}$ and $(x_{mod})_{max}$ are the maximum values of the observed and modeled signals during a certain time window. Perfect models have a ratio of 1. The ratio of max-min and the ratio of max larger than 1 overestimate maximum variations and maximum values. Note that the two ratios depend on the length of time window; the choice of time interval will be discussed in section 6.1.

6. Results

6.1. Vertical Drifts at Jicamarca

[28] Figure 2 shows examples of the observed and modeled vertical drifts at Jicamarca during the daytime for E.2006.348 storm event (Figure 2a), and for E.2007.190 quiet time event (Figure 2b). In Figure 2, the black curves represent observation data, and the color curves correspond to averaged modeled values over every 5 min. Note that similar plots for all CEDAR Challenge events can be created through the CCMC's online metrics tools at http://ccmc.gsfc.nasa.gov.

[29] Figure 2 provides only qualitative model performance on predicting vertical drifts during the day (0700∼1700). However, there is a notable difference in performance of the models. For example, during the E.2007.190 event, models including 1_SAMI3_HWM07, 1_SAMI3_HWM93, and 3_GITM produce larger temporal variation of the vertical drift than the actual observations, whereas 1_TIE-GCM, 2_TIE-GCM, and 1_IRI produce smaller variation than the observed vertical drifts. The observed values seem to match better with the results of the empirical model 1_IRI than with the results from the physics-based models (see Figure 2b). During the storm event, however, physics-based coupled model such as 2_TIE-GCM performs better in producing dynamics or peak values than the IRI model (see Figure 2a). It is fairly obvious from Figure 2 that prediction of vertical drift

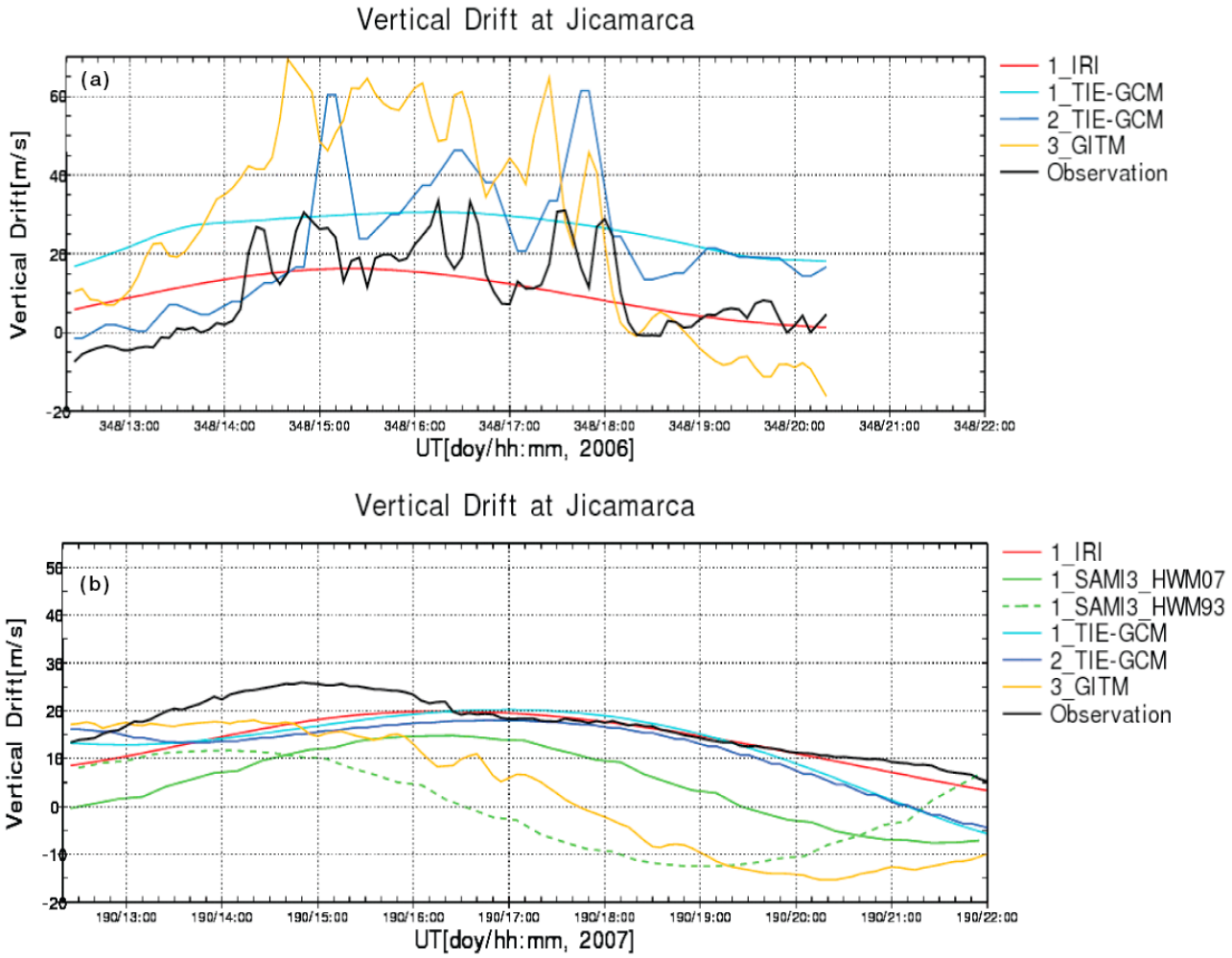

Figure 2. Examples of the observed (black curves) and modeled (color curves) vertical drifts at Jicamarca during the daytime for the (a) E.2006.348 event and (b) E.2007.190 event.

during the geomagnetic storm is challenging. Other than that, we can only draw limited conclusions on model performance from Figure 2.

[30] In order to make definite comparisons, we quantified the model performance by using metrics. Figure 3 shows the model ranking based on four different metrics for vertical drift at Jicamarca. As mentioned in the previous section, the ratio depends on the length of time window selected. To find out the appropriate length of the time window, three different time windows of 1, 4 and 7 h were taken. The event E.2006.348 (Figure 2a) was selected for this test since Figure 2 clearly shows that one model, 2_TIE-GCM for this event, is better than the other models such as 1_IRI and 1_TIEGCM, which can be compared with the results for different time windows. It is found that the ratio of max-min of 2_TIEGCM is closest to 1, while those of 1_IRI and 1_TIEGCM are farther away from 1 with the time window of 1 h (not shown here). The opposite holds true for the 7-h time window. Thus, 1-h time window length was selected to calculate the ratio of maximum minus minimum values (max-min) and ratio of maximum values (max).

[31] Figure 3 shows RMS error (Figure 3a) and prediction efficiency PE (Figure 3b), and the ratio of max-min (Figure 3c) and the ratio of max (Figure 3d). Squares indicate the average values for strong storm cases (E.2005.243 and E.2006.348); circles and triangles denote the average of moderate storms (E.2001.243, E.2007.142, and E.2008.059) and the average of the quiet periods (E.2007.079, E.2007.190, and E.2007.341), respectively. Crosses are the average of all three geomagnetic activity levels. The ranking of the model performance is arranged by the final average values. The best performing model is located at the extreme left. In Figures 3c and 3d the models located above (below) the thin black horizontal line, which indicates a ratio of 1, overestimate (underestimate) maximum variations and/or maximum values. Note that neither all modeled data nor all observational measurements are

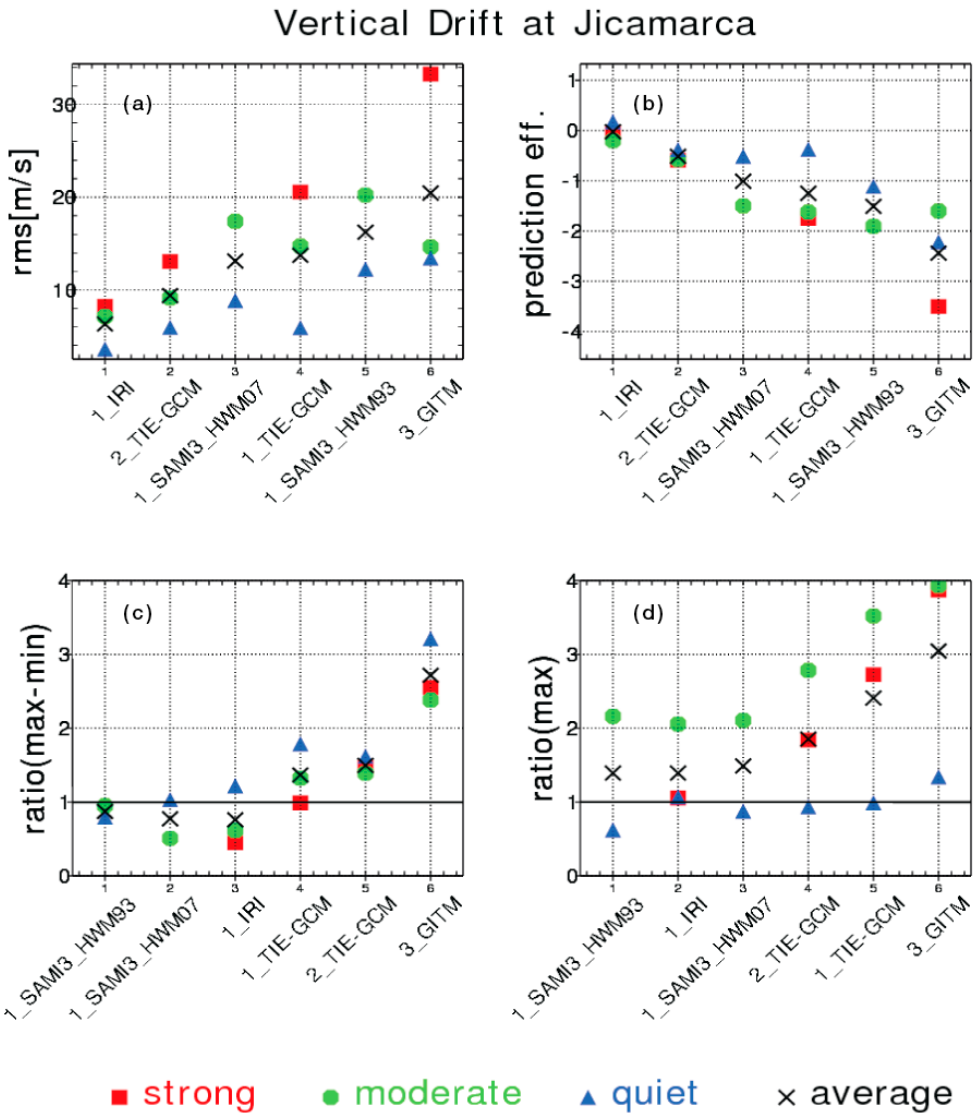

Figure 3. Model ranking for predicting the vertical drifts at Jicamarca based on (a) RMS error, (b) prediction efficiency, (c) ratio of max-min, and (d) ratio of max. Squares denote the average values for strong storm cases (E.2005.243 and E.2006.348); circles and triangles indicate the average of moderate storms (E.2001.243, E.2007.142, and E.2008.059) and the average of quiet periods (E.2007.079, E.2007.190, and E.2007.341). Crosses are the average of the all three geomagnetic activity levels, and the rankings of the models performance are arranged by the final average values. The best performing model is located at extreme left.

available for all events. For example, 1_SAMI_HWM07 and 1_SAMI_HWM93 do not have data for the two strong storm cases, and there are no observations for the event E.2007.091. Therefore, comparing the averaged performance taken over the events requires caution.

[32] The RMS differences for the vertical drift increase as a function of geomagnetic activity. All models show worst performance for strong storm periods in terms of RMS error. In terms of both RMS error and prediction efficiency, the empirical model 1_IRI ranks the top followed by the coupled model 2_TIE-GCM. For all events, all models except for 1_IRI during the quiet time periods have negative prediction efficiency, which is calculated using only daytime observations available. This suggests that the mean of the observations is more accurate than modeled vertical drifts at Jicamarca during daytime.

[33] As for the ratio of max-min and ratio of max, 1_SAMI3_HWM93 gives us the most accurate vertical drifts, but the model ranking based on the average values

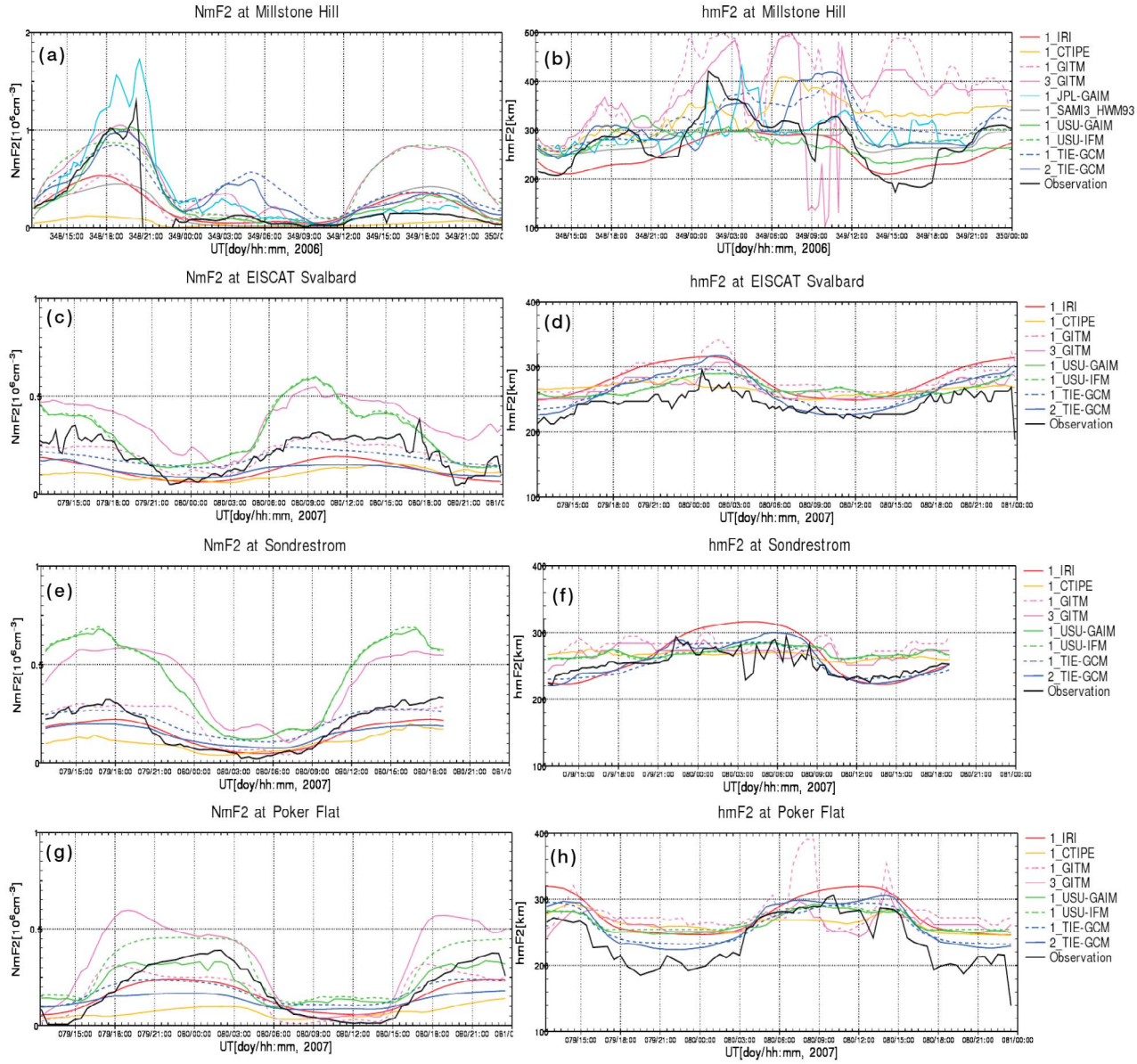

Figure 4. Observed (black curves) and modeled (color curves) NmF2 and hmf2 at (and b) Millstone Hill for a strong storm case, E.2006.348, event and at (c and d) EISCAT Svalbard, (e and f) Sondrestrom, and (g and h) Poker Flat for the E.2007.079 event.

over all events available changes when only one event or only one geomagnetic activity level is considered. For example, 1_SAMI3_HWM07, 1_SAMI3_HWM93 and 1_TIEGCM show the best performance in producing max-min for quiet periods, moderate and strong storm cases, respectively. Also, in terms of the ratio of max, 1_IRI performs better than 1_SAMI3_HWM93 for each geomagnetic level. Therefore, it should be noted that a model, which both overestimates and underestimates the max-min and max, could have a ratio closest to 1 in an average sense.

[34] It is evident from Figure 3d that maximum values of the modeled vertical drifts and observations show good agreements during the quiet periods, while the models tend to overestimate maximum values during the storms.

6.2. NmF2 and hmf2

[35] Figure 4 shows examples of the observed and modeled NmF2 and hmf2. Results for Millstone Hill are shown for a strong storm case, E.2006.348 event (Figures 4a and 4b). Results are shown for E.2007.079, one of the geomagnetically quiet periods, at EISCAT Svalbard

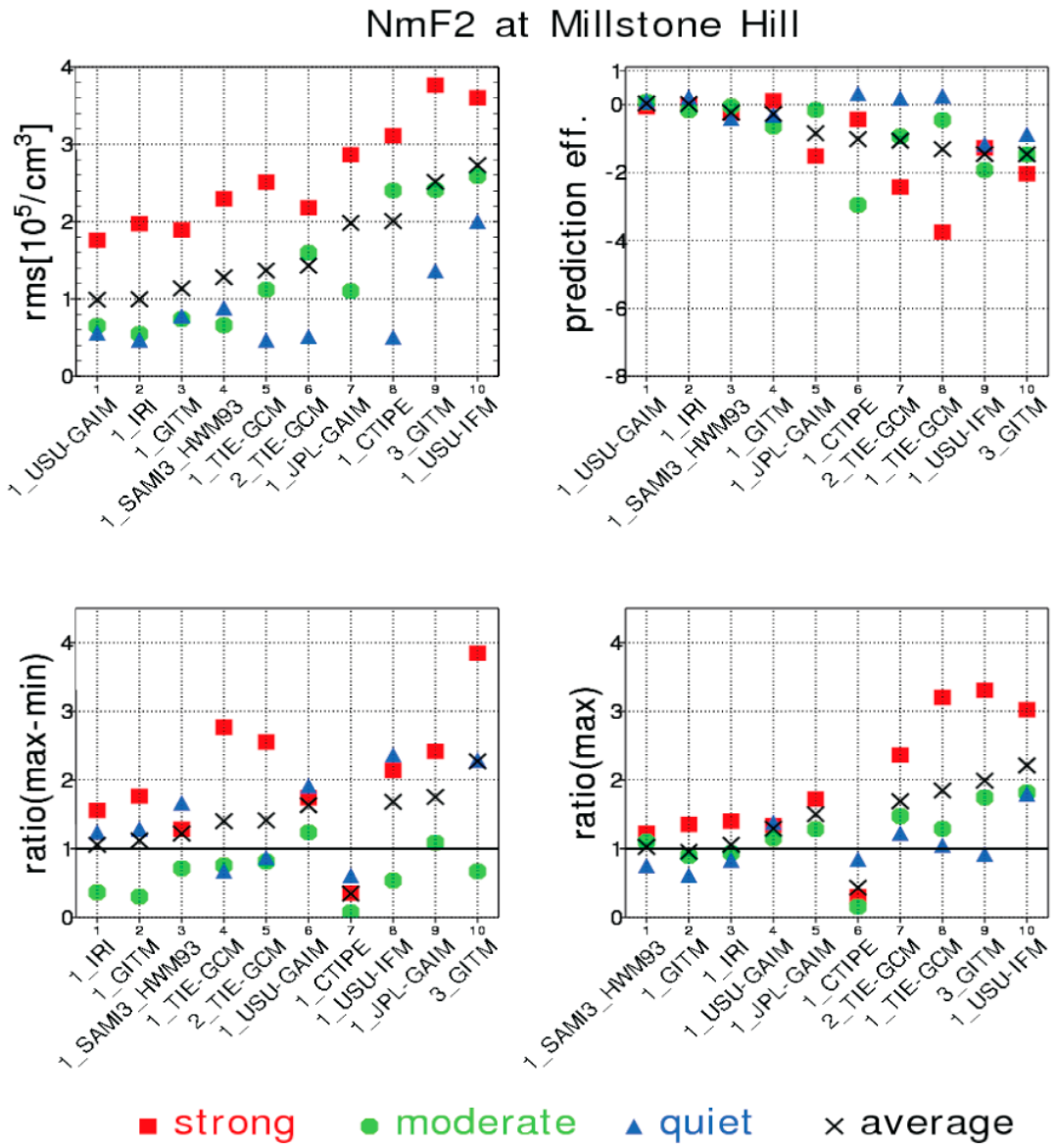

Figure 5. Same as Figure 3 but for predicting the NmF2 at Millstone Hill. Squares and triangles denote for strong storm cases (E.2006.348) and for quiet periods (E.2007.341). Circles indicate the average of moderate storms (E.2007.142 and E.2008.059).

(Figures 4c and 4d), Sondrestrom (Figures 4e and 4f), and Poker Flat (Figures 4g and 4h). Although the model results for most cases show similar diurnal variations as the observations, models such as 1_USU-GAIM, 1_USU-IFM and 3_GITM produce larger NmF2 variation than measurements, while 1_CTIPe produces smaller variation than the observed values at EISCAT Svalbard and at Sondrestrom. For the geomagnetically quiet condition, the results from the empirical model 1_IRI show better agreements with the observations than the results from the physics-based models (see Figures 4c–4h). On the other hand, the physics-based models perform better in producing the peak values during the storm event than the IRI model (see Figures 4a and 4b). Figure 4 also clearly indicates that the prediction of NmF2 and hmF2 during geomagnetic storms is as challenging as the prediction of vertical drifts.

[36] Figures 5–8 show the model ranking using four different metrics for NmF2 and hmF2. Table 3 shows available measurements from the four ISR stations for each event. In Figure 5, capability of models to produce NmF2 at Millstone Hill at middle latitudes is compared. During the storm, all 10 model output submissions show largest RMS error of predicted NmF2 that ranges between about

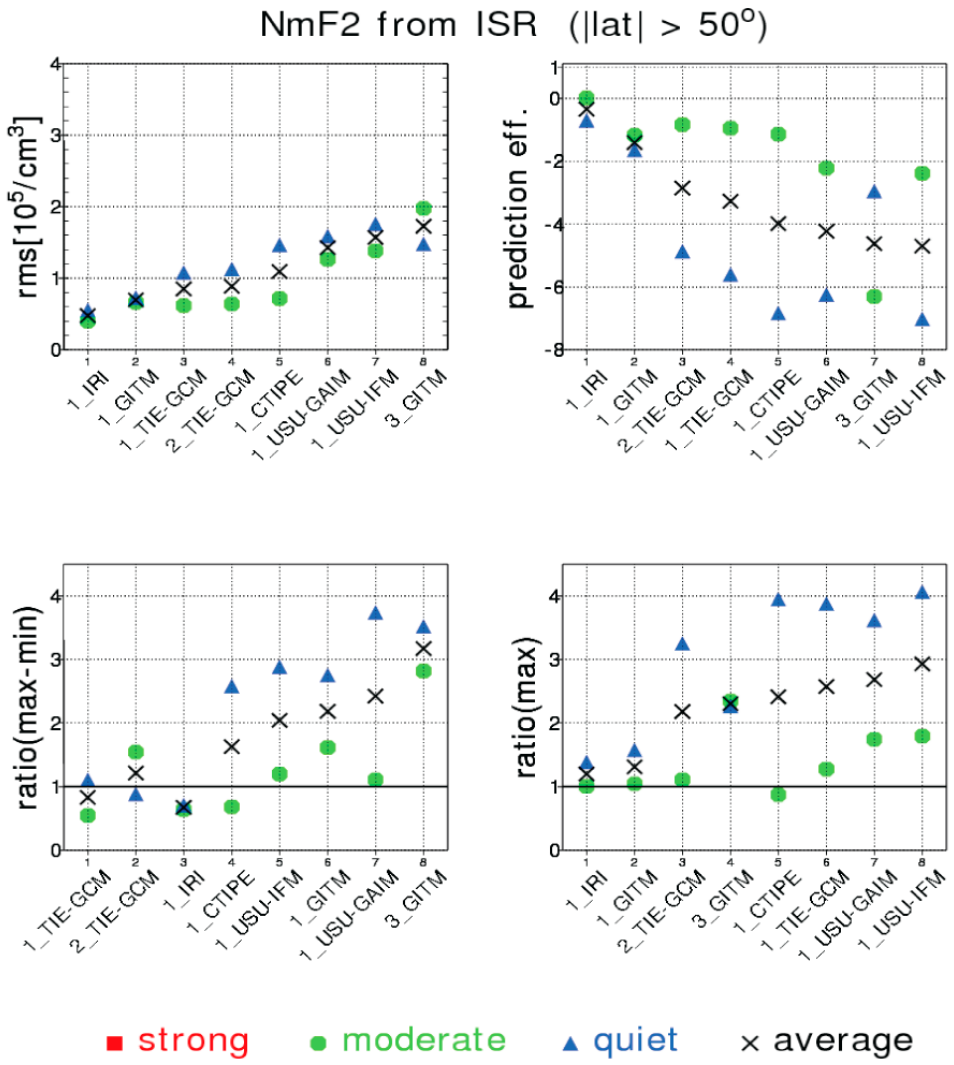

Figure 6. Same as Figure 3 but for predicting the NmF2 in high latitudes. Circles and triangles denote the average of moderate storms (E.2007.091, E.2007.142, and E.2008.059) and the average of quiet periods (E.2007.079, E.2007.190, and E.2007.341).

2×10^5 and 4×10^5 /cm^3. The models that are ranked higher show no significant differences in RMS error between quiet periods and moderate storm events. However, relatively low ranked models show distinct differences in RMS error between quiet periods and moderate storm events and have larger error for moderate storms than for quiet periods. The data assimilation model, 1_USU-GAIM is ranked at the top followed by IRI in both RMS error and prediction efficiency. The model rankings based on the RMS error and prediction efficiency are similar but not the same. The ranking for prediction efficiency is obtained based on the average prediction efficiency for daytime (0600–1800 LT) and nighttime (1800–0600 LT), while the RMS error does not have local time dependence. However, there is no difference in model rankings for vertical drifts prediction based on the RMS and PE in Figure 3, since only daytime vertical drifts were considered for the study. At middle latitudes, the majority of models tend to overestimate max-min, which indicates short-term (1 hr) temporal variation, and max of NmF2 during the storm event except for 1_CTIPe. On the other hand, models have a tendency to underestimate max-min during moderate storms. Note that 1_JPL-GAIM did not provide Millstone Hill results for the quiet period, and only includes one moderate storm case (E.2008.059).

[37] The model performance of predicting NmF2 at high latitudes is shown in Figure 6, which displays the average performance of the models taken over the three ISR stations (EISCAT Svalbard, Poker Flat and Sondrestrom) at high latitudes (geomagnetic latitude >60°) for only medium

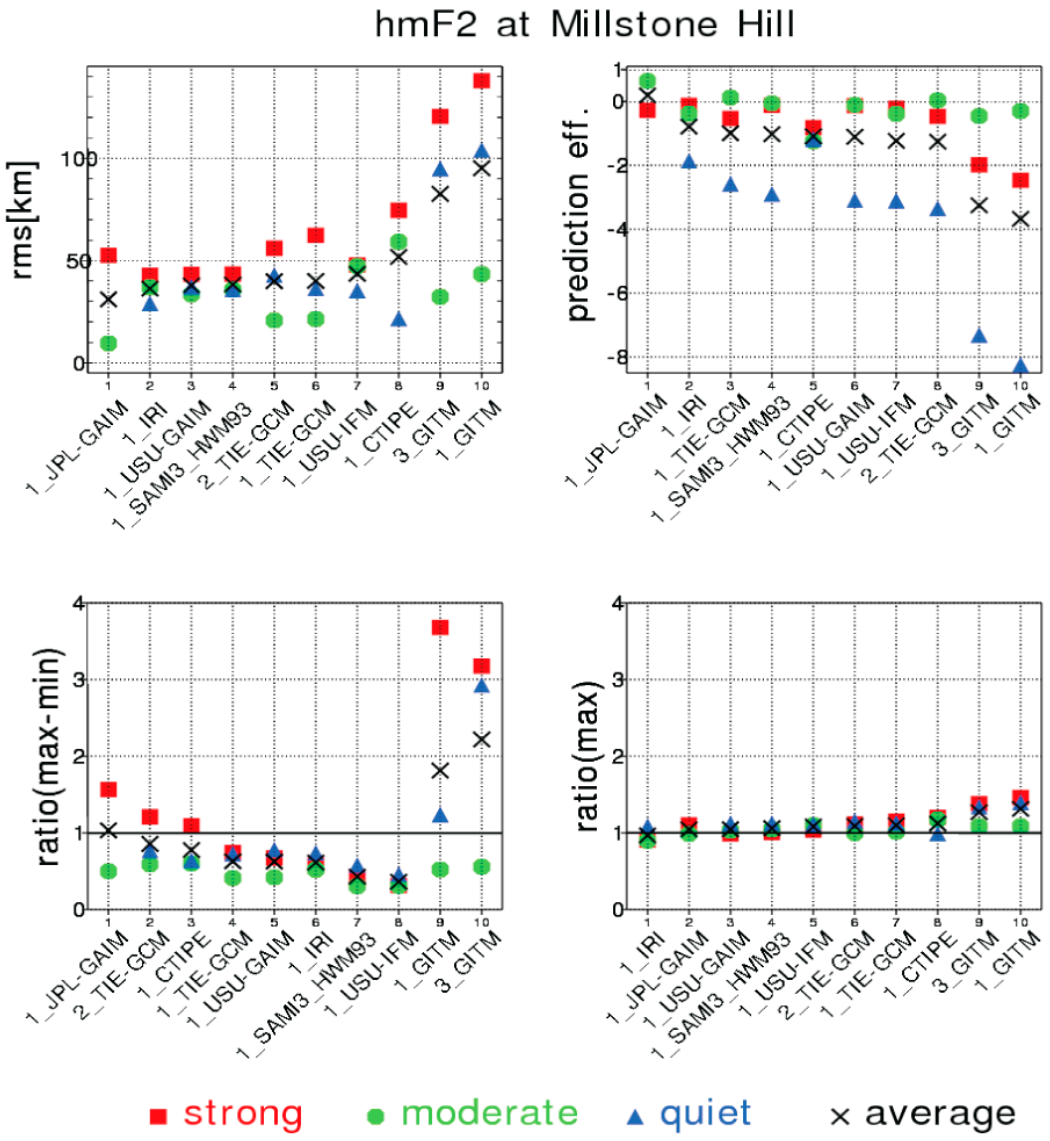

Figure 7. Same as Figure 3 but for predicting the hmF2 at Millstone Hill. Squares and triangles denote for strong storm cases (E.2006.348) and for quiet periods (E.2007.341). Circles indicate the average of moderate storms (E.2007.142 and E.2008.059).

storm and quiet events due to absence of data for storm events. At high latitudes, the models show notable differences in performance compared to middle latitudes (Figure 5). First, all models except for 3_GITM produce smaller RMS error during moderate storms than during quiet times at high latitudes. Second, except for 1_IRI during moderate storms, all models fail to provide positive prediction efficiency. It indicates that the mean value of the observations is a better predictor than the modeled NmF2 at high latitudes. Next, 1_USU_GAIM now ranks near the bottom as opposed to near or at the top for middle latitudes in Figure 5. The reason why the data assimilation model 1_USU-GAIM at high latitudes shows worse performance than in middle latitudes is that the versions of the USU-GAIM currently hosted at the CCMC only use GPS TEC data between −60 and +60 geographic latitudes. At high latitudes, 1_IRI and 1_GITM are ranked higher in terms of RMS error, PE, and ratio of max, but are ranked relatively lower based on the ratio of (max-min). Finally, all models except 3_GITM show larger negative PE for quiet periods than for moderate storms. Also, all models have worse PE at high latitudes compared to middle latitudes, especially during the quiet periods. Note that 1_SAMI3_HWM93 has no data for NmF2 at high latitudes, which is not covered by the model. Also, there are no data from 1_JPL-GAIM.

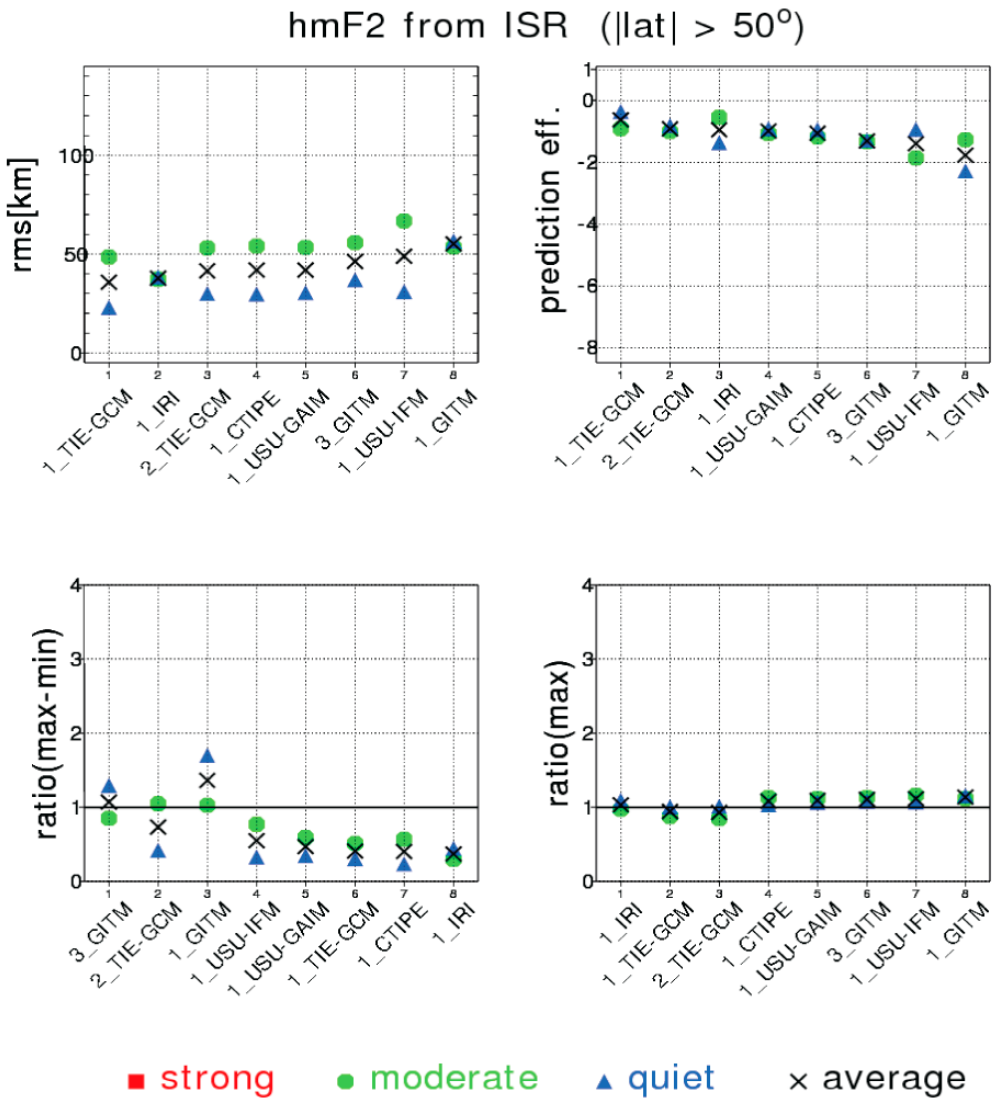

Figure 8. Same as Figure 3 but for predicting the hmF2 in high latitudes. Circles and triangles denote the average of moderate storms (E.2007.091, E.2007.142, and E.2008.059) and the average of quiet periods (E.2007.079, E.2007.190, and E.2007.341).

Table 3. NmF2 and hmF2 Observations Available During the Selected Events

Events		Measurements
Strong storms	E.2006.348	Millstone Hill
	E.2005.243	no data available
Moderate storms	E.2001.243	no data available
	E.2007.091	EISCAT Svalbard, Poker Flat
	E.2007.142	Millstone Hill, EISCAT Svalbard, Poker Flat, Sondrestrom
	E.2008.059	Millstone Hill, Sondrestrom
Quiet periods	E.2007.079	EISCAT Svalbard, Poker Flat, Sondrestrom
	E.2007.190	EISCAT Svalbard
	E.2007.341	Millstone Hill, EISCAT Svalbard, Poker Flat, Sondrestrom

[38] Figure 7 and 8 present the models ranking of performance in predicting hmF2 at Millstone Hill (middle latitude) and in high latitudes. In Figure 7, dependency of the model performance on geomagnetic activity is also shown, but the dependency is not linear. For example, some models have larger RMS error for quiet periods than moderate storms, although all models have the largest RMS error for strong storms. Furthermore, all models have worst PE for hmF2 in the middle latitude for the quiet periods. The models exhibit a tendency to underestimate the difference (max-min) for most cases. In terms of the ratio of max, all models show good agreements with the measurements. The data assimilation models (1_JPL-GAIM, 1_USU-GAIM) and the empirical model (1_IRI) are ranked higher than the others, but differences in scores among the models ranked above average are not substantial. The differences become smaller as the latitude increases in Figure 8. In high latitudes, RMS error for hmF2 ranges from about 20 to 30 km for geomagnetic quiet time and from about 50 to 70 km for moderate storms. PE scores approximately range from -2 to 0. The ratio of max is about 1, and the ratio of max-min is less than 2 in most cases. 1_TIE-GCM, 2_TIE-GCM and IRI are ranked either at the top and or near top. 3_GITM is ranked the top with respect to ratio of max-min and demonstrates more accuracy in predicting hmF2 in high latitudes than in middle latitudes.

7. Discussions and Conclusion

[39] The performance of Ionosphere/Thermosphere models in predicting NmF2, hmF2 and vertical drifts has been quantified using four different metrics for three different geomagnetic activity levels. For this IT model validation study, nine events were selected and divided into three geomagnetic levels according to maximum Kp value during the events (see Table 1). JULIA vertical drifts and NmF2 and hmF2 from ISRs at one station in middle latitude (Millstone Hill) and three stations (EISCAT Svalbard, Poker Flat and Sondrestrom) in high latitude regions were used as ground truth for comparison with a total of 10 model submissions (see Table 2). To quantify the model performance, we calculated four different skill scores including RMS error, prediction efficiency against the mean of the observations, as well as the ratio of max-min and ratio of max with a 1-h time window. The capabilities of models to produce the selected physical parameters were ranked by the average skill scores over all events for which the measurements were available (see Table 3). Therefore, it is important to note that not all model submissions had data for all events when comparing the model performance. Also, ISR data were not available from all ISR stations for all events

[40] The model performance shows an evident dependence on geomagnetic activity. RMS error increases as geomagnetic activity levels increase for the majority of the models in most cases with the exception of NmF2 and hmF2 RMS errors, which are larger for quiet periods than for moderate storm events. This nonlinearity is possibly due to the simultaneous changes in the ionospheric drivers such as electric fields, neutral wind, neutral composition, and neutral temperature, which differ from storm to storm, depending on the energy input to the ionosphere from the magnetosphere. The models that are ranked higher appear to have smaller difference in RMS error between different geomagnetic levels. However, the dependence of the prediction efficiency on geomagnetic activity is not as clear as the behavior of the RMS error, and even the opposite behavior is true for some cases. For example, for predicting NmF2 in high latitude and hmF2 in middle latitude, most models show worse performance for quiet periods than for strong and/or moderate storms. In addition, in terms of ratio of max-min and ratio of max, the models tend to overestimate max-min and max during the strong storm events. Most models also overestimate maximum vertical drift during the moderate storms, while they show good agreements with observed maximum vertical drift during the quiet periods.

[41] It is clearly seen that model performance varies with metric selections and latitude so that none of models rank at the top for all metrics used. In general, the empirical model 1_IRI, which provides long-term climate trends rather than short-term variations in weather, shows better performance than the other models in terms of RMS error and PE, while the other models produce better ratio of max-min than 1_IRI. This suggests that physics-based models produce relatively better short-term variations of the ionospheric physical parameters during storms and even quiet periods. The data assimilation models 1_USU-GAIM and 1_JPL-GAIM are ranked top in their predictions of NmF2 and hmF2 at Millstone Hill (located at middle latitudes) with respect to RMS and PE. However, they are outranked by the empirical model 1_IRI and the coupled model 1_TIE-GCM in high latitudes due to the limited latitude coverage of data used for the data assimilation models.

[42] From the comparisons among the same types of models, it is found that one of the two data assimilation models, 1_USU-GAIM produces better NmF2 for all cases, better hmF2 during strong storms, and worse hmF2 during moderate storms than 1_JPL-GAIM in middle latitudes. Two physics-based ionospheric models, 1_SAMI3_HWM93 and 1_USU-IFM show similar performance in producing hmF2. However, 1_SAMI3_HWM93 shows better performance in producing NmF2 than 1_USU-IFM in middle latitudes. It appears that 2_TIE-GCM, 1_TIE-GCM, and 1_CTIPE among five coupled model submissions perform similarly in producing NmF2 in middle latitudes during quiet periods and in high latitudes during moderate storms. The three submissions also show similar performance in predicting hmF2 at high latitudes. In most cases, they show better performance than 1_GITM and 3_GITM. However, 1_GITM is the best for NmF2 in middle latitudes during storms and in high latitudes during quiet periods in terms of RMS errors and PE, and 3_GITM for hmF2 in high latitudes during quiet periods in terms ratio of max-min. As mentioned above,

1_CTIPE, 2_TIE-GCM and 1_TIE-GCM show similar performance. For most of the other cases, however, the TIE-GCM results have better performance than 1_CTIPE, except for NmF2 during strong storms in terms of PE and ratios, and hmF2 during quiet periods in terms of RMS and PE in middle latitudes.

[43] Improvements of model performance are pronounced when enhanced and/or more complex input drivers are used. For example, 1_SAMI3_HWM07 using upgraded Horizontal Wind Model shows better performance in predicting vertical drifts in terms of RMS error and PE, although it is not as good in terms of ratio of max-min during moderate storm events compared to 1_SAMI3_HWM93 driven by the earlier Horizontal Wind Model. During the storm events, 2_TIE-GCM (driven by Weimer electric potential with dynamic critical crossover colatitudes) shows better performance than 1_TIE-GCM (driven by Heelis electric potential with constant critical crossover colatitudes) in producing all three physical parameters in terms of all metrics. However, 2_TIE-GCM does not show systematic improvements for moderate storm and quiet periods. Also, there are no systematic improvements when 1_GITM and 3_GITM are compared. 1_GITM has better scores for NmF2, whereas 3_GITM has better scores for hmF2.

[44] This is the first systematic study to quantify the performance of various Ionosphere/Thermosphere models. The results of this study provide a baseline from which performance of new models can be evaluated, although there were limitations in this study such as relatively short duration of the events considered, temporal coverage of data, and uncertainty of the measurements. The results from the comparison of measurements and model simulations of the electron density and neutral density along the CHAMP trajectories will be presented in a companion paper, which will be followed by the next report on a climatological study covering one year of the IPY. 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 communities to use in the future.

[45] **Acknowledgments.** The Jicamarca Radio Observatory is a facility of the Instituto Geofisico del Peru operated with support from National Science Foundation (NSF) award AGS-0905448 through Cornell University. The Millstone Hill incoherent scatter radar is supported by the NSF. This study made use of the CEDAR Data Base at the National Center for Atmospheric Research, which is supported by the NSF.

References

Anderson, D., A. Anghel, J. Chau, and O. Veliz (2004), Daytime vertical E × B drift velocities inferred from ground-based magnetometer observations at low latitudes, *Space Weather*, *2*, S11001, doi:10.1029/2004SW000095.

Anderson, D. N., et al. (1998), Intercomparison of physical models and observations of the ionosphere, *J. Geophys. Res.*, *103*, 2179–2192, doi:10.1029/97JA02872.

Bilitza, D. (1990), International Reference Ionosphere 1990, *NSSDC/WDC-A-R&S 90–22*, 155 pp., Natl. Space Sci. Data Cent., Greenbelt, Md., Nov.

Bilitza, D. (2001), International Reference Ionosphere 2000, *Radio Sci.*, *36*(2), 261–275, doi:10.1029/2000RS002432.

Bilitza, D. (2004), A correction for the IRI topside electron density model based on Alouette/ISIS topside sounder data, *Adv. Space Res.*, *33*, 838–843, doi:10.1016/j.asr.2003.07.009.

Bilitza, D., and B. W. Reinisch (2008), International reference ionosphere 2007: Improvements and new parameters, *Adv. Space Res.*, *42*, 599–609, doi:10.1016/j.asr.2007.07.048.

Chau, J. L., and R. F. Woodman (2004), Daytime vertical and zonal velocities from 150-km echoes: Their relevance to *F*-region dynamics, *Geophys. Res. Lett.*, *31*, L17801, doi:10.1029/2004GL020800.

Codrescu, M. V., T. J. Fuller-Rowell, J. C. Foster, J. M. Holt, and S. J. Cariglia (2000), Electric field variability associated with the Millstone Hill electric field model, *J. Geophys. Res.*, *105*, 5265–5273, doi:10.1029/1999JA900463.

Drob, D. P., et al. (2008), An empirical model of the Earth's horizontal wind fields: HWM07, *J. Geophys. Res.*, *113*, A12304, doi:10.1029/2008JA013668.

Fejer, B. G., E. R. de Paula, S. A. Gonzales, and R. F. Woodman (1991), Average vertical and zonal *F* region plasma drifts over Jicamarca, *J. Geophys. Res.*, *96*, 13,901–13,906, doi:10.1029/91JA01171.

Fuller-Rowell, T. J., and D. S. Evans (1987), Height-integrated Pedersen and Hall conductivity patterns inferred from the TIROS-NOAA satellite data, *J. Geophys. Res.*, *92*(A7), 7606–7618, doi:10.1029/JA092iA07p07606.

Hagan, M. E., M. D. Burrage, J. M. Forbes, J. Hackney, W. J. Randel, and X. Zhang (1999), GSWM-98: Results for migrating solar tides, *J. Geophys. Res.*, *104*, 6813–6827, doi:10.1029/1998JA900125.

Hedin, A. E. (1991), Extension of the MSIS thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, *96*, 1159–1172, doi:10.1029/90JA02125.

Hedin, A. E., et al. (1991), Revised global model of thermosphere winds using satellite and ground-based observations, *J. Geophys. Res.*, *96*, 7657, doi:10.1029/91JA00251.

Hedin, A. E., et al. (1996), Empirical wind model for the upper, middle, and lower atmosphere, *J. Atmos. Terr. Phys.*, *58*, 1421–1447, doi:10.1016/0021-9169(95)00122-0.

Heelis, R. A., J. K. Lowell, and R. W. Spiro (1982), A model of the high-latitude ionospheric convection pattern, *J. Geophys. Res.*, *87*, 6339, doi:10.1029/JA087iA08p06339.

Heppner, J. R., and N. C. Maynard (1987), Empirical high-latitude electric-field models, *J. Geophys. Res.*, *92*, 4467–4489, doi:10.1029/JA092iA05p04467.

Holt, J. M., S.-R. Zhang, and M. J. Buonsanto (2002), Regional and local ionospheric models based on Millstone Hill incoherent scatter radar data, *Geophys. Res. Lett.*, *29*(8), 1207, doi:10.1029/2002GL014678.

Huba, J., G. Joyce, and J. Fedder (2000), Sami2 is another model of the ionosphere (SAMI2): A new low-latitude ionosphere model, *J. Geophys. Res.*, *105*, 23,035, doi:10.1029/2000JA000035.

Huba, J. D., G. Joyce, and J. Krall (2008), Three-dimensional equatorial spread F modeling, *Geophys. Res. Lett.*, *35*, L10102, doi:10.1029/2008GL033509.

Hysell, D. L., M. F. Larsen, and R. F. Woodman (1997), JULIA radar studies of electric fields in the equatorial electrojet, *Geophys. Res. Lett.*, *24*(13), 1687–1690, doi:10.1029/97GL00373.

Millward, G. H., I. C. F. Müller-Wodrag, A. D. Aylward, T. J. Fuller-Rowell, A. D. Richmond, and R. J. Moffett (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, *J. Geophys. Res.*, *106*, 24,733–24,744, doi:10.1029/2000JA000342.

Pi, X., C. Wang, G. A. Hajj, G. Rosen, B. D. Wilson, and G. J. Bailey (2003), Estimation of E × B drift using a global assimilative ionospheric model: An observation system simulation experiment, *J. Geophys. Res.*, *108*(A2), 1075, doi:10.1029/2001JA009235.

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

Powell, K. G., P. L. Roe, T. J. Linde, T. I. Gombosi, and D. L. D. Zeeuw (1999), A solution-adaptive upwind scheme for ideal magnetohydrodynamics, *J. Comput. Phys.*, *154*, 284, doi:10.1006/jcph.1999.6299.

Pulkkinen, A., L. Rastätter, M. Kuznetsova, M. Hesse, A. Ridley, J. Raeder, H. J. Singer, and A. Chulaki (2010), Systematic evaluation

of ground and geostationary magnetic field predictions generated by global magnetohydrodynamic models, *J. Geophys. Res.*, *115*, A03206, doi:10.1029/2009JA014537.

Pulkkinen, A., et al. (2011), Geospace Environment Modeling 2008–2009 Challenge: Ground magnetic field perturbations, *Space Weather*, *9*, S02004, doi:10.1029/2010SW000600.

Rastätter, L., M. M. Kuznetsova, A. Vapirev, A. Ridley, M. Wiltberger, A. Pulkkinen, M. Hesse, and H. J. Singer (2011), Geospace Environment Modeling 2008–2009 Challenge: Geosynchronous magnetic field, *Space Weather*, *9*, S04005, doi:10.1029/2010SW000617.

Richmond, A. D., and Y. Kamide (1988), Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique, *J. Geophys. Res.*, *93*(A6), 5741–5759, doi:10.1029/JA093iA06p05741.

Richmond, A. D., E. C. Ridley, and R. G. Roble (1992), A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, *19*, 601–604, doi:10.1029/92GL00401.

Ridley, A. J., T. I. Gombosi, D. L. D. Zeeuw, C. R. Clauer, and A. D. Richmond (2003), Ionospheric control of the magnetospheric configuration: Neutral winds, *J. Geophys. Res.*, *108*(A8), 1328, doi:10.1029/2002JA009464.

Ridley, A. J., Y. Deng, and G. Toth (2006), The global ionosphere-thermosphere model, *J. Atmos. Sol. Terr. Phys.*, *68*, 839–864, doi:10.1016/j.jastp.2006.01.008.

Roble, R. G., E. C. Ridley, A. D. Richmond, and R. E. Dickinson (1988), A coupled thermosphere/ionosphere general circulation model, *Geophys. Res. Lett.*, *15*, 1325–1328, doi:10.1029/GL015i012p01325.

Scherliess, L., and B. G. Fejer (1999), Radar and satellite global equatorial F region vertical drift model, *J. Geophys. Res.*, *104*, 6829–6842, doi:10.1029/1999JA900025.

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

Scherliess, L., R. W. Schunk, J. J. Sojka, D. C. Thompson, and L. Zhu (2006), Utah State University Global Assimilation of Ionospheric Measurements Gauss-Markov Kalman filter model of the ionosphere: Model description and validation, *J. Geophys. Res.*, *111*, A11315, doi:10.1029/2006JA011712.

Schunk, R. W. (1988), A mathematical model of the middle and high latitude ionosphere, *Pure Appl. Geophys.*, *127*, 255, doi:10.1007/BF00879813.

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

Schunk, R. W., L. Scherliess, and J. J. Sojka (2002), Ionospheric specification and forecast modeling, *J. Spacecr. Rockets*, *39*(2), 314–324, doi:10.2514/2.3815.

Schunk, R. W., et al. (2004), Global Assimilation of Ionospheric Measurements (GAIM), *Radio Sci.*, *39*, RS1S02, doi:10.1029/2002RS002794.

Sojka, J. J. (1989), Global-scale physical model of the F region ionosphere, *Rev. Geophys.*, *27*, 371, doi:10.1029/RG027i003p00371.

Sojka, J. J., R. Schunk, T. van Eyken, J. Kelly, C. Heinselman, and M. McCready (2007), Ionospheric challenges of the International Polar Year, *Eos Trans. AGU*, *88*(15), 171, doi:10.1029/2007EO150003.

Spence, H., D. Baker, A. Burns, T. Guild, C.-L. Huang, G. Siscoe, and R. Weigel (2004), Center for integrated space weather modeling metrics plan and initial model validation results, *J. Atmos. Sol. Terr. Phys.*, *66*, 1499–1507, doi:10.1016/j.jastp.2004.03.029.

Tobiska, W. K. (1991), Revised solar extreme ultraviolet flux model, *J. Atmos. Terr. Phys.*, *53*, 1005–1018, doi:10.1016/0021-9169(91)90046-A.

Tóth, G., et al. (2005), Space weather modeling framework: A new tool for the space science community, *J. Geophys. Res.*, *110*, A12226, doi:10.1029/2005JA011126.

Wang, C., G. Hajj, X. Pi, I. G. Rosen, and B. Wilson (2004), Development of the global assimilative ionospheric model, *Radio Sci.*, *39*, RS1S06, doi:10.1029/2002RS002854.

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

Weimer, D. R. (2005), Improved ionospheric electrodynamic models and application to calculating Joule heating rates, *J. Geophys. Res.*, *110*, A05306, doi:10.1029/2004JA010884.

D. Anderson, Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, 325 Broadway, Boulder, CO 80305, USA.

D. Bilitza, M. Hesse, M. Kuznetsova, and L. Rastätter, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

M. Butala, A. J. Mannucci, X. Pi, and P. Stephens, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA.

J. L. Chau, Radio Observatorio de Jicamarca, Instituto Geofisico del Peru, Apartado 13-0207, Lima, Peru.

M. Codrescu and T. Fuller-Rowell, Space Weather Prediction Center, NOAA, 325 Broadway, Boulder, CO 80305, USA.

B. Emery and B. Foster, High Altitude Observatory, National Center for Atmospheric Research, Boulder, CO 80301, USA.

J. Huba, Plasma Physics Division, Naval Research Laboratory, Code 6790, Washington, DC 20375-5320, USA.

B. Rideout, Haystack Observatory, Massachusetts Institute of Technology, Westford, MA 01886, USA.

A. Ridley, Space Physics Research Laboratory, University of Michigan, 2455 Hayward St., Ann Arbor, MI 48109-2143, USA.

L. Scherliess, R. W. Schunk, J. J. Sojka, and L. Zhu, Center for Atmospheric and Space Sciences, Utah State University, 4405 Old Main Hill, Logan, UT 84322-4405, USA.

J. S. Shim, Goddard Planetary Heliophysics Institute, University of Maryland Baltimore County, NASA Goddard Space Flight Center, Bldg. 21, 251, Mail Code 674, Greenbelt, MD 20771, USA. (jasoon.shim@nasa.gov)

D. C. Thompson, Air Force Research Laboratory, Albuquerque, NM 87117, USA.