Global thermosphere-ionosphere response to onset of 20 November 2003 magnetic storm

G. Crowley,1,2 C. L. Hackert,1 R. R. Meier,3 D. J. Strickland,4 L. J. Paxton,5 X. Pi,6 A. Mannucci,6 A. B. Christensen,7 D. Morrison,5 G. S. Bust,8,2 R. G. Roble,9 N. Curtis,1,2 and G. Wene10

Received 3 November 2005; revised 18 April 2006; accepted 17 July 2006; published 18 October 2006.

There is great interest in understanding how the thermosphere-ionosphere system responds to geomagnetic storms. New insights are possible using the new generation of fully coupled three-dimensional models, together with extensive ionospheric databases. The period of postsolar maximum geomagnetic storms in October and November 2003 were some of the largest storms ever recorded. In this paper, we explore how the thermosphere-ionosphere system responded to the onset of the 20 November 2003 geomagnetic storm, using the NCAR TIMEGCM. The model simulates dramatic changes in the thermospheric equatorward winds, O/N2, and corresponding ionospheric electron densities. The model is used as a framework to interpret an increase in the observed ionospheric total electron content, and F region electron density, in the European and North African sector, in terms of changes in the neutral gas. Corresponding compositional effects observed by the GUVI instrument on the TIMED satellite lend credence to the model results. We describe some of the important physical processes that will affect planning for the utilization of measurements from the Geospace investigations in NASA’s Living With a Star Program. The study illustrates the value of measuring both the neutral and ionized gases, of obtaining quasi-global views from imaging instruments, and the synergy between satellite data, ground-based measurements, and models.


I. Introduction

While ionospheric variability has been studied for many years (e.g., review by Proß [1980]), geomagnetic storms still defy our ability to understand, explain, or simulate their detailed effects on the ionosphere and thermosphere [Proß, 1987; Crowley et al., 1989a, 1989b, 1995; Burns et al., 1995; Fuller-Rowell et al., 1996; Craven et al., 1994; Strickland et al., 2001]. Although there is a significant amount of ionospheric electron density data available to define the ionospheric response to storms, supporting thermospheric and electric field data have always been relatively sparse. In addition, the available thermospheric and E field data were usually obtained at a different location and time than the ionospheric measurements, making it difficult to explain the cause of ionospheric effects. The storms during October and November 2003 are some of the largest on record and have provoked a great deal of interest in the geospace community. A first principles model of the coupled ionosphere-thermosphere system was used to simulate the onset of the positive phase of the storm around 1200 UT on 20 November 2003. The model reproduced several measured features of the storm and we use the model to explain the mechanisms that are most likely responsible for the observed effects.

3 Storm-time ionospheric enhancements (positive storms) can occur at different local times and different latitudes, and there are likely a number of different mechanisms responsible for their formation. Positive storm types were categorized by Proß [1995]. It is generally accepted that production and loss can be accounted for by solar illumination and neutral composition, while transport occurs via electric fields and neutral winds. However, there is still disagreement about the causes of specific F region enhancements. For example, Foster [1993] and Foster et al. [2004]
have drawn attention to frequent enhancements of the midlatitude ionosphere in the afternoon sector that appear to be driven by strong electric fields. The noon event described in the current study seems to have a different cause. The study illustrates the value of measuring both the neutral and ionized gases, of obtaining quasi-global views from imaging instruments, and the synergy between satellite data, ground-based measurements, and models.

2. Model Simulations

[4] Thermospheric General Circulation Models (TGCMs) were developed by the National Center for Atmospheric Research (NCAR) beginning in the early 1980s to study the global temperature, circulation, and chemical structure of the thermosphere and its response to solar and auroral activity. A three-dimensional (3-D) coupled thermosphere-ionosphere general circulation model (TI-GCM) was extended to include self-consistent electrodynamic interactions (TIE-GCM) between the ionosphere and thermosphere [Richmond et al., 1992]. The model now extends down to 30 km, to include the mesosphere and upper stratosphere and is known as TIME-GCM [Roble and Ridley, 1994]. The TIME-GCM code has a long heritage and is widely acknowledged as one of the premier upper atmosphere codes in existence. It predicts winds, temperatures, major and minor composition, electron densities and electrodynamic quantities globally from 30 km to about 600 km altitude. The current model uses a fixed geographic grid with a 5° × 5° horizontal resolution, and a vertical resolution of half a pressure scale height. The model time-step is typically 3 min. The inputs required by the TIME-GCM include the solar flux at 57 key wavelengths (parameterized by the F10.7 flux), auroral particle precipitation, high-latitude electric fields, and tides propagating up from below the 30 km lower boundary. The codes were initially developed at NCAR for a CRAY Supercomputer environment, but the TIME-GCM was ported to SwRI and can now run in a workstation environment or on a Beowulf system [Crowley et al., 1999].

[5] The November 2003 TIMEGC runs presented here are the same as those described by Meier et al. [2005]; the inputs included solar flux, tides, and high-latitude electric fields and particle precipitation. The size of the auroral oval and particle fluxes were driven by Hemispheric Power estimates from the DMSP and NOAA satellites on a cadence of about 15 min. The cross cap potential was represented by a Heelis et al. [1982] analytical model driven by the IMF By component. The cross-cap potential difference needed to drive the Heelis model was obtained from a Weimer [1996] empirical model driven by solar wind inputs. The only midlatitude electric fields included in the model are those from the neutral-wind dynamo. Tidal inputs are based on seasonal climatology from the GSWM model [Hagan et al., 1999].

3. Ionospheric Observations

[6] Figures 1a and 1c respectively, present global maps of the TEC obtained from GPS ground-based receiver data [Mannucci et al., 1998; Pi et al., 2000] for 1200 UT on 19 November 2003 (quiet day) and 1200 UT on 20 November 2003 (storm onset). Comparison of the two (Figure 1e) reveals that the TEC near noon over Europe and North Africa has dramatically increased by over 30 TEC units following storm onset, doubling in some areas. On the right side of the figure are corresponding model results in TEC units × 10^{16}. The model TEC was expected to be less than the observed TEC because the upper boundary of the model is near 700 km, and it does not capture plasmaspheric variations associated with the storm. However, on the dayside, the plasmaspheric contribution typically provides only about 20% of the measured TEC, explaining why the model is in such good agreement with the observed TEC. Figure 1f reveals that the model also simulated a global distribution of increases and decreases that are very similar to those observed. In the European sector, the model simulated a TEC increase that is slightly larger in extent and magnitude than the measured difference shown in Figure 1e.

[7] Figures 2a and 2c depict values of the corresponding peak electron density measured in the Northern Hemisphere F region at 1200 UT for the quiet and storm day, respectively. The maps of NmF2 were obtained by assimilation of multiple ionospheric data sets by the IDA3D (Ionospheric Data Assimilation Three Dimensional) algorithm of Bust et al. [2004]. IDA3D is an objective analysis algorithm, based upon three-dimensional variation (3DVAR) data assimilation. In this case, IDA3D ingested ground-based GPS TEC from a large number of stations distributed throughout North America, Alaska, Greenland, and Europe. The University of Texas at Austin (ARL:UT) Greenland tomographic array [Bust et al., 2004] collected data from numerous satellite passes during this period and provided ground-based beacon tomography TEC estimates. In addition, numerous ionosondes measured NmF2 and HmF2, and electron densities were also available for ingestion from the Sondrestrom, Greenland incoherent scatter radar. Finally, DMSP satellite measurements of in situ electron density at 840 km and space-based GPS occultation TEC from the CHAMP satellite were also assimilated. The IDA3D NmF2 shown in Figure 2c for the storm day increases in the same region as the TEC of Figure 1. Figure 2e shows the difference between the quiet and stormtime NmF2 distributions, revealing that NmF2 is two or three times larger on the storm day in the European and North African region around 0° longitude and 30°–60° latitude, in the same region as the TEC increase of Figure 1. This suggests that much of the increase in TEC is caused by an increase in the F region electron density, rather than purely a topside or plasmaspheric density increase. Figures 2b, 2d, and 2f show the corresponding model results. The TIMEGCM overestimates the value of NmF2 at low latitudes on both the quiet and storm days (a problem that will probably be improved as the dynamo code is improved), but the modeled increase of about 10^{12} el/cm^2 is comparable with that observed, although slightly broader as noted for the TEC.

4. GUVI Observations

[8] In an ideal world, there would be direct observations of the thermospheric drivers to help explain the ionospheric behavior and to validate the model predictions. For the present study, the only thermospheric data available were
Figure 1. Measured total electron content for (a) QUIET, (c) STORM, (e) difference (STORM-QUIET), and (b), (d), (f) corresponding model results.
Figure 2
from the GUVI instrument [Christensen et al., 2003] on the NASA-TIMED satellite. GUVI provides remote sensing observations of the column-integrated O/N2 ratio (called $\Sigma$O/N2 here). GUVI measures a narrow swath below the satellite and is unable to provide a global snapshot. Instead, a global picture must be accumulated over about 15 orbits, which takes about a day. The TIMED orbit precesses slowly, so over a few days, the data are all collected within a small range of local times. On 19–21 November the nadir data were all collected close to local noon. The upper panel of Figure 3 shows the $\Sigma$O/N2 data from GUVI for days 323–325, plotted as a function of latitude and longitude (note that time runs from right to left). These GUVI data were discussed in more detail by Meier et al. [2005].

Comparison of the data from Day 323 and 324 reveals that reduced $\Sigma$O/N2 values observed in the northern hemisphere were depleted at increasingly equatorward locations, beginning over Scandinavia. Equatorward of the depletions, enhanced values of $\Sigma$O/N2 were observed. A black ellipse on Day 324 indicates the European and North African regions near 1200 UT that are the focus of this paper. Within the ellipse, the values of $\Sigma$O/N2 on the storm day are generally larger than on the preceding quiet day. The lower panel shows the model predictions in the same format, indicating that the model simulates the GUVI data reasonably well during the first 18 hours of the storm. In particular, the model simulates the progression of observed increases and decreases of $\Sigma$O/N2 between Scandinavia and Florida, including the changes over Europe and North Africa. The GUVI data, the model simulation, and the compositional storm response were discussed in more detail by Meier et al. [2005].

5. Discussion

[9] Figures 1 and 2 showed that the model simulated broad increases in TEC and NmF2 over the European and North African sectors at 1200 UT, in reasonable agreement with the observations. We now explore the changes in other model parameters from quiet to storm times to investigate the reason for the TEC and foF2 increases.

Figure 3. Column integrated $\Sigma$O/N2 ratio: (a) measured by GUVI; (b) predicted by TIMEGCM. The vertical red lines indicate day boundaries. Note that time runs from right to left.

Figure 2. IDA3D NmF2 for (a) QUIET, (c) STORM, (e) difference (STORM-QUIET), and (b), (d), (f) corresponding model results.
Figure 4a shows the predicted change in the height of the maximum $F$ region electron density. Note we limit the value of the peak height plotted here to altitudes above 200 km, and the unphysical-looking contours in the auroral zone and southern hemisphere during nighttime arise when the maximum electron density occurs in the $E$ region. The red ellipse highlights the region of interest for this paper, and shows that the height of the $F$ region peak on the storm day increased by 130 km to 250 km over the quiet day values in Europe and North Africa. The factors contributing to this height rise are the meridional wind (Figure 4b), electrodynamic drift (Figure 4c), and the exospheric temperature (Figure 4d). The meridional wind in the European sector was poleward on the quiet day (not shown), but strongly equatorward during the storm. The difference plot (Figure 4b) reveals large equatorward (negative values) wind differences of 150 m/s to 330 m/s within the red ellipse, which tended to raise the $F$ region during the storm. The shape of the wind contours was reflected in the hmF2 differences (Figure 4a), indicating that the wind made a major contribution to the increase in hmF2. The vertical $E \times B$ drift from dynamo processes in this region is small, both on the quiet and storm days, according to the model, and the difference in Figure 4c is also close to zero. Thus a strong $E \times B$ drift is not required to explain the observed increase of TEC and NmF2. The nighttime temperature near the equator doubled (from 850 K to about 1500 K) due to adiabatic compression associated with the wind surges from both hemispheres. Even within the ellipse, the temperature increased by 100–400 K (15–45%), indicating that the thermal
expansion of the thermosphere must also contribute to the increased hmF2 in that region.

[11] The equatorward wind blows plasma along magnetic field lines and thus to greater altitudes at midlatitudes so that the ionospheric plasma moves into an area where the ratio of [O] to [N2] is greater. Consequently, recombination slows relative to production, leading to greater electron densities. Figure 5 shows the model predictions of [O]/[N2] at the F region peak at 1200 UT for both quiet-time (Figure 5a) and during the storm (Figure 5b), together with their difference (Figure 5c). These figures show that the [O]/[N2] at the F region peak increased dramatically from less than 10 to values between 10 and 100 in the European and North African sectors, and within the ellipse was everywhere larger during the storm.

[12] While there are no in situ data to confirm the model predictions of [O]/[N2] in the F region, the GUVI instrument provided measurements of the column integrated [O]/[N2] ratio. Figure 5d shows the model predictions of the global changes in ΣO/N2 at 1200 UT during the storm day. The column integrated values mimic the changes shown in Figure 5c for the model [O]/[N2] at the F region peak. The reasonable agreement between the GUVI observations and the model ΣO/N2 in Figure 3 lends confidence that the model is producing a reasonable simulation of the storm compositional response, especially near 1200 UT during the onset of the storm.

6. Conclusions

[13] We have studied the ionospheric response to the onset of a storm during 20 November 2003. The TEC and peak F region electron density were both significantly enhanced in the European and North African sectors at 1200 UT on 20 November (Day 324). Simulation of the period using the TIMEGCM suggests that these midlati-
tude ionospheric changes were dominated by changes in the neutral gas that were forced by high-latitude heating processes associated with the storm. Specifically, the model suggests that neutral winds drive the ionospheric plasma to greater altitudes, where the ratio of production to loss is greater, leading to enhanced electron densities in the sunlit sector. Unfortunately, there is little neutral atmospheric data to validate the model predictions, or to directly inform the mechanisms causing the observed changes. GUVI observations of the column-integrated O/N\textsubscript{2} ratio suggest the model is simulating this part of the storm event with reasonable accuracy. The model contained only dynamo electric fields, and omitted midlatitude electric fields caused by penetration from high latitudes and those caused by inner magnetospheric processes [e.g., Huang et al., 2006]. However, the excellent agreement between the model and data suggests that electric fields are unimportant for this particular timeframe in this event (other times and other events undoubtedly contain large E field effects).

[14] Ionospheric storms have been studied for many years. Enormous amounts of ionospheric data are currently available from ground-based instruments, so the ionospheric response to storms is being observed better than ever before. Unfortunately, it is not possible to rigorously test the many theories of what drives the ionospheric behavior because corresponding measurements of the drivers are usually lacking: namely, the thermospheric neutral composition, winds, temperature, and electric fields. What are needed are simultaneous, collocated measurements of the comprehensive set of parameters. Such data have not been forthcoming in the past, showing that it is not sufficient to rely on serendipity to provide the required overlap of measurements in space and time.

[15] The Ionosphere-Thermosphere Storm Probes mission, planned for launch in 2015 as part of NASA’s Living With a Star Program, will address this problem with comprehensive in situ measurements of the neutral parameters listed above, together with the electric fields. This mission will consist of two satellites in 450 km circular orbits, separated by 1 to 2 hours of local time so that the important longitudinal variation of the storm drivers can also be captured. In addition to the two low-altitude satellites, a global imager placed in a geostationary or Molniya orbit would provide valuable information on the global composition changes accompanying the storms. The choice between Molniya and geostationary orbits is a difficult one because they have different benefits. The Molniya orbit provides the ability to stare at a broad area for several hours, and two satellites can provide continuous coverage over a range of fixed local times. However, it is not ideal for low-latitude studies and favors a single hemisphere (usually Northern). In contrast, geosynchronous orbit provides the ability to stare continuously at a fixed geographic location, but it favors low latitudes, has very little access to high latitudes, and is subject to local time variations.

[16] This study illustrates the urgent requirement to measure both the neutral and ionized gases in order to understand global ionospheric storm behavior. It also demonstrates the value in obtaining quasi-global views from imaging instruments, and the synergy between satellite data, ground-based measurements, and models. In the future, we anticipate further improvements in the model, particularly our simulation of the recovery phase on Day 321. We also anticipate using GUVI limb data together with ionospheric data to study later phases of the storm.

[17] Acknowledgments. This work was supported by the TIMED/GUVI project through NASA grant NAG5-5001 to The Aerospace Corporation and NASA Task ICR at the JHU Applied Physics Laboratory. G.C. was also supported by NASA grants NAG5-10059, NAG5-13613, and NNG04GN04G (LWS). This material also draws upon work supported by the National Science Foundation under grant ATM-0228458 and ATM-0228467. We are also grateful to Stephen Mende for a helpful discussion of the relative merits of Molniya and geostationary orbits for imagers.

[18] Lou-Chuang Lee thanks the reviewers for their assistance in evaluating this paper.

References


---

G. S. Bust, G. Crowley, and N. Curtis, Atmospheric and Space Technology Research Associates, 11118 Quail Pass, San Antonio, TX 78249, USA. (gbust@astraspace.net; gcrowley@astraspace.net; ncurtis@astraspace.net)

A. B. Christensen, Northrop Grumman Corporation, Corporate Headquarters, 1840 Century Park East, Los Angeles, CA 90067-2199, USA. (andy.christensen@nbg.com)

C. L. Hackert, Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78238-5166, USA. (chackert@swri.org)

A. Mannucci and X. Pi, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA. (tony.mannucci@jpl.nasa.gov; xiaoping.pi@jpl.nasa.gov)

R. R. Meier, School of Computational Sciences, George Mason University, Fairfax, VA, USA. (rmeier@gmu.edu)

D. Morrison and L. J. Paxton, Johns Hopkins University, Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723-6099, USA. (daniel.morrison@jhuapl.edu; larry.paxton@jhuapl.edu)

D. J. Strickland, Computational Physics Inc., 8001 Braddock Rd., Suite 210, Springfield, VA 22151, USA. (dstrick@cpi.com)

R. G. Roble, NCAR, High Altitude Observatory, 3450 Mitchell Lane, Boulder, CO 80301, USA. (roble@hao.ucar.edu)

G. Wene, University of Texas at San Antonio, 6900 Loop 1604 West, San Antonio, TX 78249, USA. (gwene@utsa.edu)