Overcooling in the upper thermosphere during the recovery phase of the 2003 October storms

Jiuhou Lei,1,2 Alan G. Burns,3 Jeffrey P. Thayer,2 Wenbin Wang,3 Martin G. Mlynczak,4 Linda A. Hunt,5 Xiankang Dou,1 and Eric Sutton6

Received 10 July 2011; revised 6 February 2012; accepted 7 February 2012; published 10 March 2012.

[1] Infrared radiative emissions by carbon dioxide (CO2) and nitric oxide (NO) are the major cooling mechanisms of the lower thermosphere. During geomagnetically active periods, the NO density and cooling rate in the auroral regions increase significantly as a result of particle precipitation and Joule heating. Previous studies have shown that the time for NO density to recover to quiet time levels is longer than that of the thermosphere temperature or density recovery. This study explores the implications of these different recovery rates for the post-storm thermosphere. Thermosphere densities retrieved from the CHAMP and GRACE accelerometer measurements and NO cooling rates measured by TIMED/SABER are used to examine their variations during the post-storm period of the October 2003 geomagnetic storms. It was found that thermosphere densities at both CHAMP and GRACE altitudes recovered rapidly and continuously decreased below the quiet time densities during the post-storm period, especially at middle latitudes. Compared with the quiet time values, the maximum depletion in the CHAMP and GRACE densities after the storm is about 23–36%, and the estimated decrease of thermospheric temperature is as large as 70–110 K. Our analysis suggests that the elevated NO cooling rate, resulting from the slower recovery of NO densities in the post-storm period, is a plausible cause for this apparent post-storm overcooling of the thermosphere.


1. Introduction

[2] Solar radiation, Joule dissipation and particle precipitation are the primary heating sources for the upper atmosphere, whereas infrared radiative (IR) emissions by carbon dioxide (CO2) and nitric oxide (NO) and heat conduction are its major cooling mechanisms [Roble et al., 1987]. As the energy input from solar radiation and high latitude processes increase, the upper thermosphere tends to become hotter. Meanwhile, enhanced CO2 and NO cooling offsets this increase of thermosphere temperature, which is called the thermospheric ‘thermostat’ [Mlynczak et al., 2003], controlling the structure of the thermospheric temperature, density and composition.

[3] Statistical analysis has shown that geomagnetic energy contributes roughly 20% of the total energy input of the upper atmosphere, but can rise up to two-thirds of the energy budget during storm time [Knipp et al., 2005]. NO density, and its concomitant IR cooling, in the auroral zone increases significantly as a result of the enhancement of particle precipitation and Joule heating [e.g., Mlynczak et al., 2003, 2005, 2008; Barth et al., 2009; Lu et al., 2010]. In addition, Burns et al. [1989] and Maeda et al. [1989, 1992] demonstrated that the NO cooling plays an important role in damping the enhanced temperature and mass density associated with the elevated geomagnetic activity to recover back to the pre-storm level. They found that the time scales of thermospheric temperature (or neutral composition) recovery are around half a day. However, the recovery time for NO density is around 1 day [Solomon et al., 1999; Lu et al., 2010], which is usually longer than the time scales of the thermosphere temperature or density recovery. That is, when thermosphere temperature and mass density recover to its quiet time values, the NO density or cooling rate is still higher than the pre-storm level. Consequently, it is possible that the thermosphere might cool in the post-storm period to temperatures that are below those of the pre-storm thermosphere. We call this the “overcooling effect.”
we would expect overcooling to be strongest in such events.

gest storms in the solar cycle 23 occurred during this time and
val is selected as our study period because two of the stron-
phases of the October 2003 geomagnetic storms. This inter-
mass density and the NO cooling rates during the recovery
are used to examine the variations of both thermospheric
sphere using Broadband Emission Radiometry) instrument
from TIMED/SABER (Thermosphere Ionosphere Meso-
2004]).

Weirs et al. 2002] and GRACE accelerometer measurements using standard
methods [Sutton et al., 2007] are utilized in this study. CHAMP and GRACE satellites flying in the near-circular
orbits with inclinations of 87.3° and 89.5° respectively, provide thermosphere density measurements from pole to pole. The GRACE satellites consist of two identical satellites GRACE-A and GRACE-B flying about 220 km apart and they provide similar variations in thermosphere density. Thus only GRACE-A data are used for this study. On the basis of the error estimation given by Sutton et al. [2007], we found that in this event the errors in derived densities are larger near the poles and during the disturbances, and are around 10% due largely to neglecting neutral wind speeds. The errors from calibrations are less than 4–5% for CHAMP, and much less for GRACE during most of this period. The latitudinal pattern of these errors is not consistent with the overcooling features to be described later in the paper.

[6] The SABER instrument on board the TIMED satellite monitors the thermospheric energy budget by measuring infrared emissions from NO at 5.3 µm and from CO₂ at 15 µm [Mlynczak, 2019]. SABER scans the Earth’s limb from 400 km tangent height down to the hard earth surface, recording approximately 1600 profiles of limb radiance on each day. Since the CO₂ density is not sensitive to the magnetospheric energy input [Mlynczak et al., 2008], only globally distributed vertical profiles between 100 and 200 km of the SABER NO cooling rate are used to interpret thermosphere mass density recovery. Note that The SABER observations during the study period covered the latitudinal range from 51.9°S to 83.9°N, and the median local times of the NO data on the dayside and nightside are 17.37 and 2.41, respectively.

3. Observations

[7] Two severe geomagnetic storms were triggered by coronal mass ejections (CMEs) reaching the Earth on October 29 and 30, 2003 (two of the “Halloween” storms), a period when the Sun was extremely active [Skoug et al., 2004]. Figures 1a–1c shows the time series of the interplanetary magnetic field (IMF) Bz from the ACE satellite, geomagnetic activity index Kp and ring current index Dst during October 28–31, 2003. As indicated in Figures 1a–1c, two superstorms occurred during this period. Their respective recovery phases started at 0100 UT on 30 October (identified throughout the text as R1) and at ~2300 UT on 30 October (identified throughout the text as R2), facilitated by northward Bz. Lei et al. [2011] reported that the relaxation times, defined by the e-folding time of the post-storm recovery of thermosphere density during this period, are about 6–8 h; these time scales are much shorter than those predicted by empirical and theoretical models. Figure 1d depicts the variations of the orbit-averaged neutral mass densities from CHAMP during October 28–31, 2003. In this orbit-averaged plot, the measured densities at CHAMP altitudes are normalized to a constant altitude of 390 using MSISE00 [Picone et al., 2002] since the altitudes of the satellite changed significantly from pole to pole. Obviously, thermospheric mass densities on both the dayside and nightside were enhanced significantly in response to the storms, whereas they recovered rapidly during both R1 and R2. Thermosphere densities recovered to the pre-storm level during R1 and they became even lower during R2 compared

Figure 1. Variations of (a) interplanetary magnetic field Bz, (b) Kp, (c) Dst, (d) dayside (red) and nightside (green) orbit-averaged neutral densities at 390 km from CHAMP, orbit-averaged NO cooling rate from TIMED/SABER (e) in the altitudinal range between 100 and 200 km, and (f) dayside (red) and nightside (blue) averaged NO cooling rate at 125 km during 28–31 October 2003. Note that the dashed lines in Figure 1d stand for mass density in units of 10⁻¹² kg/m³ during the quietest period on October 28; NO cooling rate in (e–f) is in units of 10⁻³ erg/cm³/s.

2. Data Source

[5] Thermosphere mass densities obtained from CHAMP and GRACE accelerometer measurements using standard

[4] The purpose of this study is to explore whether this overcooling effect does occur in the post-storm thermosphere. Thermosphere mass densities retrieved from the accelerometer measurements on the CHAMP (Challenging Minisatellite Payload [Reigber et al., 2002]) and GRACE (Gravity Recovery and Climate Experiment [Tapley et al., 2004]) satellites along with the NO cooling rates estimated from TIMED/SABER (Thermosphere Ionosphere Mesosphere Energetics and Dynamics / Sounding of the Atmosphere using Broadband Emission Radiometry) instrument are used to examine the variations of both thermospheric mass density and the NO cooling rates during the recovery phases of the October 2003 geomagnetic storms. This interval is selected as our study period because two of the strongest storms in the solar cycle 23 occurred during this time and we would expect overcooling to be strongest in such events.
with their quiet time level. These dominant features in response to the October 2003 storms were also observed in the GRACE thermospheric mass density (see the details given by Lei et al. [2011]).

[8] Roble et al. [1987] demonstrated that under high solar activity conditions the dominant cooling process in the lower thermosphere is NO radiative cooling, which results in enhanced downward heat conduction in the upper thermosphere. In order to better understand the thermospheric mass density response during the recovery phase, the orbit-averaged NO cooling rate from TIMED/SABER in the altitudinal range between 100 and 200 km during October 2003 is presented in Figure 1e. It can be seen that the NO cooling rate peaks at around 125 km and decreases greatly above this altitude. The NO cooling rate started to increase in response to the arrival of the CME shocks and then increased profoundly during the main phase associated with the enhancement of auroral particle precipitation and thermosphere temperature [Barth et al., 2009; Lu et al., 2010]; this is similar to the behavior of the thermospheric mass density during this event. The strongest values of the NO cooling rate occurred during the late main or recovery phase of the storms, reaching levels as great as a factor of 6–8 larger than the pre-storm cooling rate, which should contribute to the rapid recovery of thermospheric mass density. Interestingly, the NO cooling rate during the later part of the recovery phase of the two storms was still much higher than that during the pre-storm period. This feature becomes immediately obvious in the line plot of the NO cooling rate at 125 km on both the dayside and nightside (Figure 1f). Additionally, the recovery times for the NO cooling rate are longer than those of thermosphere density.

[9] In order to examine the latitudinal structure of the thermospheric overcooling during these recovery phases, we analyzed the CHAMP and GRACE neutral mass densities near 390 and 485 km, respectively, from pole to pole. Figures 2a and 2b shows the relative depletion of the CHAMP mass density as a function of latitude during the October 2003 storms with respect to the reference values at the corresponding UT of October 27. Gray color area indicates the regions where the neutral densities are not lower than the reference values. Note the first two hours of data on November 1 are also included in these contours for better visualization. The local times represent the median values of the dayside or nightside orbits. See the details in the text.

Figure 2. (a–b) Relative depletion of CHAMP and (c–d) GRACE neutral mass densities near 390 and 485 km, respectively, during the recovery phases of the October 2003 storms (with respect to the reference values at the corresponding UT of October 27). Gray color area indicates the regions where the neutral densities are not lower than the reference values. Note the first two hours of data on November 1 are also included in these contours for better visualization. The local times represent the median values of the dayside or nightside orbits. $B_z$ is shown in the bottom panels for reference. See the details in the text.
The overcooling effect in recovery thermosphere mass density is clearly seen on both the dayside (Figure 2a) and nightside (Figure 2b). On the dayside, the density depletion tends to be a little bit stronger at around noon and nightside (Figure 2b). On the dayside, the density depletion during R1 is not as strong as that during R2. The maximum depletion of thermosphere density is 15.6% during R1 versus 23.3% during R2. On the nightside, the density depletion during R1 is also larger at higher middle latitudes due to larger distance between the cooling source and the satellite altitude. Note that during the recovery periods the differences between the cases with and without the F107 normalization are around 5%, except on November 1 when F107 value (F107 = 207.2) was relatively lower with respect to that on the reference day.

The overcooling effect in recovery thermosphere mass density is clearly seen on both the dayside (Figure 2a) and nightside (Figure 2b). On the dayside, the density depletion tends to be a little bit stronger at around 30° during both R1 and R2, whereas the depletion during R1 is not as strong as that during R2. The maximum depletion of thermosphere density is 15.6% during R1 versus 23.3% during R2. On the nightside, the density depletion during R1 with respect to the quiet time reference is only seen in the Southern Hemisphere with a maximum decrease of 35.7%, while the corresponding depletion during R2 occurred at most of latitudes. During R2 there was a maximum decrease of 26% at low latitudes. Additionally, the overcooling effects during R2 happened later on the nightside than on the dayside. The reason for this observation will be discussed in the following section. It should be pointed out that the orbit averaged density in Figure 1d obscures the overcooling effect significantly since the overcooling occurs mainly at low and middle latitudes, whereas at high latitudes thermosphere density is generally higher than that on the reference day (i.e., the overcooling effect is weak or absent there).

Next, we can estimate the corresponding overcooling effect on thermosphere temperature from the density observations with the aid of the MSISE00 model, if we assume that changes of thermospheric mass density are purely attributable to changes in thermosphere temperature. The MSISE00 uses the Bates-Walker temperature profile in which the exospheric temperature is a crucial parameter. In this study, the exospheric temperature in the MSISE00 is adjusted to change the neutral temperature profile and results in the calculated mass density to match the observed mass densities from the satellites. It is found that the maximum cooling in neutral temperature during R2 is about 110 and 70 K on the dayside and nightside, respectively.

Figures 2c and 2d shows the changes of the GRACE densities near 485 km during the October 2003 storms with the same format as Figures 2a and 2b. The observed features in the GRACE density depletion during the recovery phase are consistent with those from the CHAMP satellite, although the relative changes of the recovery density at GRACE altitudes are generally larger than at CHAMP altitudes except on the nightside during R1 when the density depletion is negligible. The maximum decreases of the GRACE mass density, with respect to quiet time values, are 22.6% and 28.7% on the dayside during R1 and R2 and 20.6% and 31.8% on the nightside. Lei et al. [2010] illustrated that the magnitude of the mass density increase at a fixed altitude during storm-time depends on the number of scale height increments separating the heat source from the satellite altitude. In the same way, the relative depletion in thermospheric density associated with the overcooling is also larger at higher altitudes due to larger distance between the cooling source and the satellite altitude. Note that the resultant decrease in neutral temperature at the altitudes of CHAMP and GRACE satellites would be similar because of the thermal equilibrium in the upper thermosphere.

4. Discussion

As shown in Figure 3, the maximum NO cooling rate in both hemispheres was observed at higher middle latitudes. It is expected that the overcooling would be observed in mass density at these latitudes during the recovery phase because a larger NO cooling rate leads to enhanced downward heat conduction. At high latitudes, NO density or cooling rate is high but neutral temperature also has a larger increase in response to the storm effect. Consequently, neutral gas at high latitudes takes a much longer time to cool down to the pre-storm level. The density depletion is also seen at low latitudes where the NO cooling rate during the recovery phase is not substantially higher than observed at middle latitudes (Figure 2). What likely occurs at middle and low latitudes is that the increase of Joule heating at high latitudes changes global neutral circulation and consequently its convergence and divergence pattern, which, in turn, increases the compressional heating and decreases the cooling by expansion at middle and low latitudes [Burns et al., 1992, 1995]. The net effect is to produce a fairly...
smooth pattern with greater increases in temperature at night and smaller ones during the day. These temperature increases, however, are weaker than those at high latitudes where the thermosphere is subjected to direct Joule heating. Thus, the interplay between energy sources and sinks of disparate time evolution determines where the overcooling effect occurs.

[14] Figure 2 also showed that during R2 the density depletion on the nightside occurred later than that on the dayside. This is probably associated with the temporal variation of the storm-time NO cooling rate. As can be seen in Figure 3, profound enhancements in the NO cooling rate were seen at middle latitudes during both main phases of the 29–30 and 30–31 October superstorms. The NO cooling rate along the nightside orbits (Figure 3b) during R2 remains a level as high as that during the main phase, which might be the contributor to the timing differences of the density depletion between the dayside and nightside. The density depletion is larger for R2 compared with R1 in Figure 2, while the cooling rates are not higher in R2 compared with R1 in Figure 3. This might be due to the limited amount of time for the thermosphere to experience overcooling before the second storm hit.

[15] One may notice that the local time sampling of the TIMED satellite during this study period is different from that of CHAMP and GRACE satellites, and this situation makes it difficult to claim a definite cause-effect relationship between the thermospheric overcooling and the slow recovery of NO cooling. In order to address this issue, we analyzed the NO cooling rate simulated from the NCAR-TIEGCM, which is driven by the AMIE high latitude convection and auroral particle precipitation [see Lei et al., 2011]. The simulated NO cooling rates at 125 km are sampled at fixed local times of TIMED, GRACE and CHAMP satellites. As illustrated in Figure S1, the NO cooling rates at all local times on either dayside (left) and nightside (right) are stronger during the storm time.1 The expected longitudinal dependence of the NO cooling rate due to the different local time sampling is significantly smeared out by its temporal variation with the development/decay of the storm driven by Joule heating. Therefore, the SABER NO cooling rates are still valuable to be used to explore the cause of the overcooling features seen in the CHAMP and GRACE observations, albeit the sampled local times of TIMED are different than those of CHAMP and GRACE.

[16] Finally, it would be expected that the overcooling effect should be seen in most storms if the storm-time NO density or cooling rate takes a longer time to recover to its quiet time level than the thermosphere temperature and mass density. As shown in Figure 1, thermosphere densities during the October 2003 storms recovered rapidly to the pre-storm level as a result of the strong enhancement of NO cooling during the late main or recovery phases of the storm. In this case, the elevated NO cooling rate during the post-storm period becomes effective at cooling the thermosphere to temperatures below the pre-storm level. For non-severe storms, the thermosphere recovers slowly in a relative sense; thus the overcooling effect is weak and this effect can be smeared out by other processes like advection and adiabatic heating/cooling. Additionally, in situ measurements from satellites usually provide information about the thermosphere at limited local times rather than a snapshot over the whole globe. Therefore, overcooling effects in the thermosphere may not necessarily be sampled by the satellite measurements even during severe geomagnetic storm events.

5. Concluding Remarks

[17] In this paper thermosphere densities from CHAMP and GRACE are used to further investigate thermosphere recovery during the October 2003 geomagnetic storms, as a follow-up study to Lei et al. [2011]. We found that not only did thermosphere densities recover rapidly, but they were lower during the post-storm period than the pre-storm, quiet time densities. Specifically, compared with the quiet values, the maximum depletion in the CHAMP and GRACE densities is about 23–36%, and the corresponding changes in thermospheric temperature reach 70–110 K. The TIMED/SABER measurements showed that the NO cooling rate remained at a high level during the recovery phases. Moreover, the recovery time of the NO cooling rate is longer than that of the observed thermosphere densities. Our observations suggest that the elevated NO cooling rate results in a cooler thermosphere during the post-storm period. This result indicates that the thermosphere energy balance is an intricate interplay between energy sources and sinks of disparate time evolution, which requires both processes to be evaluated to understand the observed response.

[18] We emphasize here that the purpose of this study is to report the observed post-storm overcooling (or density depletion) in the upper thermosphere from satellite data. Although thermosphere post-storm overcooling is expected theoretically, evidence of this phenomenon has not been reported before, to our best knowledge. The present study does not aim to examine all aspects of the overcooling phenomenon with very limited data sets available at the present, and there is definitely a need for more work in the future, especially direct observations that can simultaneously measure thermospheric density and temperature and NO density, as well as physics-based modeling.

[19] Acknowledgments. This work was supported by the National Natural Science Foundation of China (41174139, 41121003, 41025016, 40890165), Thousand Young Talents Program of China (USTC), the AFOSR MURI Award FA9550-07-1-0565 (CU), the Center for Integrated Space Weather Modeling (CISM), which is funded by the STC program under agreement ATM-0129050, and NASA grants NNX10A049G and NNH08AH37I from the NASA Guest Investigator Program (NCAR). The National Center for Atmospheric Research is sponsored by the National Science Foundation. Jiu Hou Lei thanks referees’ comments to improve the manuscript greatly and Gang Lu for providing the TIEGCM-AMIE simulations of the 2003 Halloween geomagnetic storms.

[20] Robert Lysak thanks the reviewers for their assistance in evaluating this paper.

References


