A simulation study of the thermosphere mass density response to substorm using GITM model

Xianjing Liu[1] and Aaron Ridley[1]

[1]University of Michigan, Atmospheric Oceanic and Space Sciences, Ann Arbor, MI

Xianjing Liu, University of Michigan, Atmospheric Oceanic and Space Sciences, Ann Arbor, MI 48105

(Email:xianjing@umich.edu)

Abstract.

The temporal and spatial variations of the thermospheric mass density during a series of idealized substorms were investigated using the Global Ionosphere Thermosphere Model (GITM). The maximum mass density perturbation of an idealized substorm with a peak variation of hemispheric power (HP) of 50 GigaWatts (GW) and interplanetary magnetic field (IMF) Bz of -2 nT was ~14% about 50 min after the substorm onset in the nightside sector of the auroral zone. The mass density response to different types of energy input has a strong local time dependence, with the mass density perturbation due to only an IMF Bz variation peaking in the dusk sector and the density perturbation due to only HP variations peaks in the nightside sector. Simulations with IMF Bz changes only, and HP changes only showed that the system behaves slightly nonlinearly when both IMF and HP variations are included (a maximum of 6% of the non-

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/2014JA020962

linearity), and that the non-linearity grows with energy input. The neutral gas heating rate due to Joule heating was of same magnitude as the heating rate due to precipitation, but the majority of the temperature enhancement due to the heating due to precipitation occurs at lower altitude as compared to the auroral heating. About 110 min after onset, a negative mass density perturbation (~-5%) occurred in the night sector, which was consistent with the mass density measurement of the CHAMP satellite.

1. Introduction

The mass density of the thermosphere is linearly proportional to the drag force that is felt by low-Earth orbiting objects. Uncertainties in thermospheric mass density variations are the major limiting factor for precise low-earth orbit determination/prediction at altitudes below about 700 km [*Marcos et al.,* 2010]. The perturbation of the thermospheric mass density is strongly controlled by the energy deposited into the upper atmosphere. The primary heating sources include solar radiation, Joule heating and particle precipitation, whereas the major cooling sources of the upper atmosphere include infrared radiative emissions by nitric oxide and heat conduction [*Roble et al.*, 1987]. Geomagnetic energy, which includes both the Joule heating and particle precipitation, contributes about 20% of the total energy input to the upper atmosphere during quiet conditions, but can increase to 67% of the total energy during geomagnetic storms [*Knipp et al*., 2004].

Changes in the thermospheric mass density include long-term, seasonal and storm time variations. The long-term thermospheric mass density follows the 11-year solar cycle, which has

been investigated in detail by *Keating et al.* [2000]; *Emmert et al.* [2004, 2010]; *Marcos et al.* [2005], *Solomon et al.* [2011] and references therein. The thermospheric mass density exhibits a strong seasonal variation, with maxima near the equinoxes, a primary minimum during northern hemisphere summer, and a secondary minimum during southern hemisphere summer [e.g. *Qian et al.*, 2009 and *A et al.*, 2012]. The time scale of the storm time variation of the thermospheric density is much smaller compared to the long-term and seasonal variations, and has a large impact on satellite orbital determination, but predicting geomagnetic storms is extremely challenging [e.g. *Valdivia et al.*, 1996; *Wu and Lundstedt*, 1996].

During geomagnetic storms, the thermospheric temperature increases due to heating due to precipitation and Joule heating (e.g., *Volland*, [1979]; *Fuller-Rowell and Rees*, [1981]; *Roble et al.*, [1982]). The behavior of the thermospheric composition during a storm is more complicated: Heavier species such as Ar and N_2 increase during the storm, whereas lighter species, such as helium, decrease [*Prölss* 1981]. *Liu et al.* [2014] investigated the altitude variation of the mass density perturbation during a geomagnetic storm using the Naval Research Laboratory Mass Spectrometer Incoherent Scatter Radar (NRL-MSIS) [*Picone et al.*, 2002] model along with satellite measurements, and found that the mass density perturbation is not only affected by the temperature, and therefore scale height enhancement, but is also strongly modified by the species ratios in the thermospheric composition transition region. *Thayer et al.* [2012] showed that the mass density response to a geomagnetic storm during solar minimum is modified by the ratio of oxygen and helium at the altitude of the Gravity Recovery and Climate Experiment (GRACE) satellite [*Tapley et al.*, 2004].

Clausen et al. [2014] used 5 years of Challenging Minisatellite Payload (CHAMP) satellite [*Reigber et al.*, 2002] data to study the mass density perturbations during substorms. Through a superposed epoch analysis of 2306 substorms, they found that the mass density perturbation peaks at about 6% ~90 min after the substorm onset and about 3 hours of local time east of onset region. *Ritter et al.* [2010] used data from the CHAMP satellite to estimate the mass density response to substorms at 400 km. They report that the mass density enhancement is about 4% to 15% in the polar region. The statistical studies of CHAMP density measurements by *Clausen et al.* [2014] and *Ritter et al.* [2010] provides a sense of how much thermosphere density perturbation is expected during a substorm. However, due to the limited parameters that the satellite can measure, it is hard to picture the whole physical process of how the upper atmosphere responds to a substorm. Simulating substorms in a physics-based model, such as the Global Ionosphere Thermosphere Model (GITM) model, can help us to better understand where the energy is deposited in the upper atmosphere and how the temperature, composition and density changes, given the energy input.

The motivation of this study was to investigate the spatial and temporal variation of the thermospheric mass density during different substorms and to investigate the difference in the mass density response to different sources of energy input using GITM simulations. This work is similar to *Clausen et al.* [2014], who used CHAMP data to investigate substorms, while this study used a global model.

2. GITM and model inputs

GITM uses a three-dimensional altitude-based spherical grid and does not assume a hydrostatic solution, which enables the model to capture physics in the high-latitude region with a more complete momentum equation [*Ridley et al.*, 2006]. The ion momentum equation is solved assuming steady-state, taking into account the pressure gradient, gravity, neutral winds, and external electric fields. GITM allows different models of high-latitude electric fields and auroral particle precipitation, but for this study, the *Weimer* [2005] electric potentials and *Fuller-Rowell and Evans* [1987] auroral precipitation patterns were used. The interplanetary magnetic field (IMF), solar wind and hemisphere power (HP) were used to drive these models. The resolution of the GITM simulation was set to 5° in longitude and 2.5° in latitude for this study. GITM was run for two days before the time period discussed here to allow a roughly diurnally repeatable pattern to form in the thermosphere. The start of the simulation presented here (i.e., at -2:00 epoch time) was at 00:00 UT on March 21.

GITM is a model of the ionosphere and thermosphere, and does not include self-consistent magnetospheric dynamics. Therefore, to simulate the thermospheric and ionospheric reaction to a substorm, the high-latitude drivers have to be altered in a non-self-consistent way, which is obviously an approximation to an actual substorm. In order to do this so that the results can be compared with the *Clausen et al.* [2014] study, similar solar and geomagnetic drivers as their superposed epoch results were used in this study. These drivers were derived from the superposed epoch analysis of all the geomagnetic conditions during 2306 substorms between

January 2001 and December 2005. Figure 1 shows the time history of the interplanetary magnetic field (IMF) z-component (i.e., Bz), auroral activity (i.e., hemispheric power, or HP) and F10.7. The x-axis of Figure 1 indicates the substorm epoch time, where 00:00 epoch time represents the onset of the substorm expansion phase. Before 00:00 epoch time, the IMF Bz was slightly negative, representing the growth phase of the substorm. At approximately 00:20, the substorm peaked in intensity, as evidenced by the maximum hemispheric power, after which, the substorm lessened in intensity and entered the recovery phase.

Based on the superposed epoch variations of the IMF Bz and the HP during the substorms used in the study of *Clausen et al.* [2014], five prototypical substorms with different combinations of IMF Bz and HP were simulated in this study. For Substorm 1, the IMF Bz shown in Figure 1a started to decrease an hour before the substorm onset, reaching a minimum value of -0.5 nT at -00:25 epoch time, and then recovered back to the pre-substorm condition at 01:20 epoch time. The HP in Figure 1b started to increase at the substorm onset and reached its maximum of 30 GW at 00:20 epoch time, and then recovered back to the pre-substorm condition at 02:00 epoch time. For the other 4 substorms, the IMF Bz and HP indices had the same temporal characteristics as that of Substorm 1 but with different peak values. The IMF Bz was set to 0 nT and HP was 20 GW for the pre- and post- substorm conditions. The IMF Bx and By were set to -2 nT and 2 nT (i.e., Parker Spiral conditions) for all the simulations. The F10.7, shown in Figure 1c, was set to 130 solar flux units (SFU) during the entire substorm interval, which represented a low to moderate level of solar radiation at the top of the atmosphere. Further, a sixth simulation was conducted with constant values equal to the pre-substorm conditions. This

simulation was used as a baseline case, so the model results could be compared and perturbations from the substorm only could be calculated.

The geomagnetic parameters of Substorm 3 were close to the median values of the superposed epoch results in *Clausen et al.* [2014], so for the following description, the simulation of Substorm 3 is presented in detail to describe the spatial and temporal variation of the thermospheric mass density response to the substorm.

3 Model simulations and observations

To represent the intensity of the mass density response to each substorm, the thermospheric mass density perturbation during the substorm was calculated as $\delta\rho = \frac{\rho_{SS}-\rho_{NoSS}}{\rho_{NoSS}}$ x100%. Where ρ_{SS} was the thermospheric mass density during the substorm, while ρ_{NoSS} was the mass density of the run with no substorm. Figure 2 shows the spatial variation of the thermospheric mass density perturbation between 40°N and 90°N at ~400km during Substorm 3 at a 10 minute cadence from 70 min before the substorm onset to 220 min after the substorm onset.

As illustrated in Figure 2, the mass density perturbation is 0% up until -60 min, which indicates that there was no mass density disturbance before the IMF Bz changed. An enhancement of the mass density, responding to the decrease of the IMF Bz during the pre-substorm time period, was observed after -60 min. The enhancement of the mass density caused by the IMF Bz occurred over all local time sectors, but the enhancement was larger on the day side than on the night side. At -10 min and 0 min, a second peak showed up in the dusk sector,

which also was associated with the IMF change. As the substorm onset began (i.e., after 0 min), the mass density responded quickly to the HP enhancement, and the perturbation in mass density caused by the increase in aurora occurred over the entire polar region with the maximum enhancement (14%) located on the nightside about 50 min after substorm onset. The perturbation then started weakening an hour after the substorm onset, and the density perturbation propagated to lower latitudes as a traveling atmospheric disturbance (TAD). During the recovery phase of the substorm from 110 min, a negative mass density perturbation occurred on the nightside (~00:00-03:00 LT), which indicates that the mass density was lower during the recovery phase than it would have been if no substorm had occurred.

In order to investigate the local time dependence of the mass density perturbation during the substorm, Figure 3 shows the average mass density perturbation as a function of epoch time during Substorm 3 in the auroral zone between 65°N and 75°N in four isolated local time sectors: (a) 03:00-09:00 LT, (b) 09:00-15:00 LT, (c) 15:00-21:00 LT, (d) 21:00 -03:00 LT. Figure 3e shows the average mass density perturbation over the entire polar area between 65°N and 75°N. The vertical dashed line indicates the onset of the substorm expansion phase. As illustrated in Figure 3, the mass density response to the IMF Bz change and the HP change had different features in their local time dependence. The perturbations before the substorm onset were caused by the IMF Bz change while the perturbations after zero epoch time were mainly due to the auroral enhancement. The mass density perturbation due to the IMF Bz variation was largest in the 15:00-21:00 LT (dusk) sector and minimum in the 03:00-09:00 LT (dawn) sector, while the mass density perturbation due to the HP enhancement maximized in the 21:00-03:00 LT

(nightside) sector, with a secondary maximum in the 03:00-09:00 LT (dawn) sector and a minimum in the 15:00-21:00 LT (dusk) sector. A significant depression of the thermospheric mass density during the recovery phase was observed in the 21:00-03:00 LT (nightside) sector about 2 hours after the substorm onset. The negative phase lasted for about 1.5 hours before it recovered back to pre-substorm conditions. The thermospheric mass density perturbation had the largest peak-to-peak oscillation on the nightside during the substorm as shown in Figure 3d.

As observed in Figure 2, a traveling atmospheric disturbance, or in situ generated large scale gravity wave was created as a result of this energy input. Figure 4 shows the mass density perturbation of Substorm 3 as a function of latitude and epoch time at 03:00 local time, similar to what an orbiting satellite would observe. As illustrated in Figure 4, a density enhancement of ~5%, caused by the IMF Bz variation, occurred in the high latitude region about 30 min before the storm onset. This was followed by a stronger mass density perturbation (>10%), which maximized in the auroral zone, and was caused by the auroral increase after the substorm onset. The density enhancement propagated from the auroral zone in both polar regions towards the equatorial regions as time elapsed. In addition, in the northern hemisphere, a TAD was observed to propagate poleward, away from the auroral oval, which is consistent with the results reported by *Bruinsma and Forbes* [2009] using the CHAMP satellite observations. The density enhancement arrived in the equator region about 3 hours after the substorm onset, which corresponded to a TAD propagation speed of about 600±120 m/s, which is also consistent with the result of *Bruinsma and Forbes* [2009]. As the density perturbation propagated to lower latitudes, the density perturbation in the auroral zone become negative between 01:30 and 03:20

epoch time, then recovered back to a positive mass density perturbation. The strong negative perturbation was only observed in the ~03:00 local time sector.

4. Discussion

4.1 Thermosphere mass density depression during the post-substorm period

In the statistical results of *Clausen et al.* [2014], there was no negative density perturbation observed on the nightside, as was observed in the GITM results. In order to tell whether this collapsing of the atmosphere ever happens after a substorm, the thermospheric mass density measured by the CHAMP satellite during two substorms during October 2003 was investigated. This time period was studied because the CHAMP satellite was at 03:00 LT, and the F10.7 index (~ 130 SFU) was close to the F10.7 input used in the GITM simulations. Besides these considerations, the substorms were chosen at random. Figures 5a and 5b show the latitudinal variation of the CHAMP mass density normalized to 400km during four satellite orbit periods (~ 4x92 min). Figures 5c and 5d show the AE index as a function of universal time during these periods. The thick black lines show the variation of the density (top plots) and AE index (bottom plots) before the substorm onset, the thick purple and red lines are times during the substorm and the thick yellow lines indicate times after the substorm. During the 04:00-10:00 UT, Oct 4, 2003 substorm (left plots), the AE index shown in Figure 5c was about ~100 nT before the substorm onset. The corresponding mass density is shown as the black line in Figure 5a. The AE index increased to about 450 nT during the substorm expansion phase and the mass density marked in purple indicates that there was no response to the substorm yet (at high

latitudes), since there may not have been time for the heating to have occurred and the atmosphere to have expanded. During the next pass (red line), an hour after the peak in the substorm, the mass density had increased at high-latitudes, but had decreased at mid-latitudes, similar to the GITM results. The AE index decreased to ~100 nT after the substorm (yellow line), and mass density indicated by the yellow line decreased below the density before the storm onset at mid-latitudes, which suggests a negative mass density perturbation during the recovery phase at 03:00 LT. This result from the satellite measurement is consistent with the GITM simulation shown in Figure 2. The right panels of Figure 5 display another substorm during 10:00-15:00 UT, Oct 4, 2003. Figure 5b also suggests a negative mass density perturbation in the mid-latitude region just after the substorm. Note that these two substorms occurred consecutively, so the pre-substorm density of the second substorm may not have totally recovered back to quiet condition from the previous substorm.

Compared to the model simulations, the latitudinal variation of the mass density from the satellite measurements shows a much more complicated structure at low latitudes, which requires further investigation in future studies. It is unclear why the density at low latitudes decreased below the density observed before the substorm, and whether the density behavior in the CHAMP measurements at low latitudes was even tied to the substorm. The main point that is suggested here is that at high- and mid-latitudes, the CHAMP-observed density can decrease below the pre-substorm value, similar to what was observed in the simulation. It is clear from the statistical results of *Clausen et al.* [2014], that, on average, this does not happen. It points to the need to understand why sometimes the atmosphere collapses after energy input, but

sometimes it does not. It should be noted, though, that the mean substorm in the *Clausen et al.* [2014] study was quite small compared to the substorm simulations that have negative perturbations. The depth of the negative perturbation appears to be related to the strength of the positive perturbation.

Lei et al. [2012] reported on the behavior of the thermospheric mass density during the Oct 30, 2003 geomagnetic storm as observed by the CHAMP and GRACE satellites. The measurements indicated that the mass density recovered rapidly and eventually decreased below the quiet time density during the pre-storm period. Their explanation of this 'overcooling effect' was that the time for the nitric oxide (NO) density to recover to a quiet time level was longer than the response time of the rest of the thermospheric densities. Since NO is one of the main coolers of the thermosphere, if NO created during the storm lingers, the thermosphere would reach a different, cooler energy equilibrium than the equilibrium that existed before the storm. In the case of the reduced mass density in the thermosphere during the recovery phase of the substorm in this simulation, the negative mass density perturbation at ~03:00 LT occurred at a substorm epoch time of 110 min, and the mass density perturbation at this location recovered back to positive values approximately 200 min after the start of the substorm. For such a short-term oscillation of mass density perturbation in this study, the driver is most likely not NO. It should also be noted that, statistically, *Clausen et al.* [2014] did not observe a negative density perturbation in the 03:00 LT sector. It is unclear why some substorms would show this negative perturbation while others would not.

4.2 Dependence of the mass density perturbation on different types of energy input

As documented in the previous section, the mass density perturbation had a strong local time dependence because the thermospheric response to different types of energy was different. Figure 6 shows the mean of the mass density perturbation in the auroral zone (i.e. between 65°N and 75°N) as a function of substorm epoch time for the 5 different substorms. The maximum average mass density perturbations in auroral zone for Substorm 1 to 5 were 2.1%, 3.5%, 6%, 8% and 10%. There were two "peaks" of the mass density perturbation as a function of epoch time: the "peak" before the substorm onset, which was associated with the IMF Bz enhancement, and the larger peak at ~ 40 min epoch time, which was associated with the auroral energy input increase. The density peak occurred 20 minutes after the peak hemispheric power energy input during the substorm.

In order to study the relationship of the mass density response to different sources of energy input, the amount of the density perturbation due to each source of energy input was calculated. In the case of the IMF variation, the high latitude electric field (and therefore ion flow) was altered. Because ion flows alter the electron density through advection, changing the electric field also altered the electron density. When the hemispheric power was varied, the electron density in the ionosphere was altered through the increased ionization rate. Because the electron density was altered, the gradients in pressure were also altered, which could have changed the ion flow velocities also, but this would be an extremely small effect. For the substorms described so far, the electric field changed first, then the aurora and ionization changed. Simplistically, one can think of these two changes as altering different terms in the ion-neutral frictional heating (i.e., the velocity difference and the electron density). Because this was a highly idealized numerical

study, it was possible to completely separate the two energy sources. Five additional simulations were run with simply the IMF change and no auroral (or HP) change. Five more simulations, beyond those, were run with the auroral change, but no IMF change. The thermospheric mass density response to each of these was then investigated. Figure 7 shows the zonal and latitudinal average mass density perturbation over the entire auroral zone for Substorms 1-5. The red line in each plot in Figure 7 indicates the density perturbation when both the HP and Bz variation were included (the same as in Figure 6). The blue line indicates the density perturbation that resulted from running the simulations with only the HP variation, with the variation in Bz not included. The black line indicates the density perturbation that resulted from running the simulation with only the Bz variation, with the variation in HP not included. The green line indicates the density perturbation that equals the sum of blue and black lines. In Figure 7a and 7b, the red line almost completely overlaps the green line, which indicates that the mass density perturbation due to both the IMF Bz and HP drivers together was quite similar to the sum of the density perturbations with HPI and Bz driven separately. This suggests the mass density response to the IMF Bz and HP inputs are nearly a linear system, during Substorm 1 and 2. If these changes were completely uncorrelated with each other, with heating from one type of event occurring in a different location than the heating from the other type of event, one would expect perfect "linearity", meaning that the average density change from one type of heating added to the average density change from the other would be the same as the average density change if both types of heating occurred at the same time. As the energy deposited into the upper atmosphere increased in Substorms 3 to 5, the mass density response to the different types of energy inputs

remained nearly linear, except near the peak of the hemispheric power input between 00:00-01:00 epoch time. The "non-linearity" of the mass density response to the energy was calculated by exploring the percentage difference $\frac{\rho_{(HP+IMF)}-(\rho_{HP}+\rho_{IMF})}{\max(\rho_{HP}+\rho_{IMF})} * 100\%$. Figure 7e shows the non-linearity of the mass density response to the different energy inputs for Substorms 1-5. The non-linearity of the mass density response for Substorm 1 and 2 was less than 2%, and the non-linearity increased as the driving energy increased. The non-linearity of the mass density response was about 6% for Substorm 5. As described above, the auroral variations tended to cause heating more on the nightside, while the IMF variations tended to cause more heating on the dayside and around dawn and dusk. Therefore, one would expect mostly a "linear" relationship, with little correlation between the two, as is observed. In the overlap region, though, where both the IMF and auroral variations caused heating, the IMF heating altered the state of the thermosphere and ionosphere, which altered the heating that resulted from the auroral inputs, making a non-linear relationship between the processes, where having both IMF and auroral changes caused a larger average heating than the two processes independent of each other.

Figure 8 illustrates that when the hemispheric power (only) was increased in the simulation, the Joule heating was also increased. This is because the HP increased the ionization and, therefore, the electron density. Since there was a pre-existing, stationary, electric field structure with-in the high latitude region, the Joule heating increased. In the specific cases described here, the resultant Joule heating was less than the hemispheric power, but this is most likely due to the weak driving in the background conditions (i.e., the constant zero IMF Bz drove a weak electric field). If the high latitude electric field were significantly larger, increasing the

hemispheric power would result in a significantly larger Joule heating increase. For substorm 3, for example, when 30 giga-watt hours (GWh) of energy was added to the system in the form of auroral precipitation, approximately 15 GWh of additional energy was added in the form Joule heating, resulting in 45 GWh of total energy added. In the real magnetosphere-ionosphere system, this relationship is not so simple, since increased precipitation sometimes results in reduced electric fields in the precipitation regions and faster flows outside of this region [e.g. *Lu et al.,* 1995, *Paschmann et al.,* 2002, *Thayer and Semeter,* 2004, *Lotko,* 2007]. The Joule heating, a mixture of the conductance and electric field is then quite complicated.

Figure 9 compares the hemispherically integrated perturbed energy (i.e. Joule heating plus hemisphere power) in GWh deposited into the upper atmosphere during Substorm 1 to Substorm 5 with the associated peak value of the mean mass density perturbation over the substorm period in the auroral zone at 150, 200, 300 and 400 km altitude. Only the hemispheric power variations were included in the simulations in Figure 9a, while only the variations in the IMF were included in the simulation in Figure 9b. Finally, Figure 9c shows the results of the simulations with both HP and IMF Bz variations included. As illustrated in Figure 9, the slope of the mass density perturbation versus energy input was larger at high altitudes than at low altitudes, which suggests that the thermosphere mass density at higher altitudes is more sensitive to the energy input. The mass density response to a single type of energy input (either by the HP or IMF Bz input) was closer to linear as shown in Figure 9a and 9b compared to the figure 9c. The mass density response to the combination of the HP and IMF Bz input is not perfectly linear, especially at 300 and 400 km.

Comparing Figures 9a and 9b, for the same energy amount of input, but in different form, the mass density perturbation caused by the HP variation (9a) was more significant than that caused by the IMF Bz enhancement (9b). Furthermore, the change of the mass density perturbation with altitude is larger for the same the Joule heating energy input caused by HP enhancement (Figure 9a) than that caused by the IMF Bz enhancement (Figure 9b). For example, for the same 60 GWh energy deposited into the upper atmosphere, the mass density perturbations were ~1.8% at 150 km and ~5.5% at 400 km for the HPI only simulation, whereas the mass density perturbation was ~0.9% at 150 km, ~3.2% at 400 km for the Bz only simulation. *Clausen et al.* [2014] reported that in order to produce a mass density increase of about 4% at the satellite altitude of 400 km, an energy deposition rate of 30 GW should be applied for 1.5 hours, which is 45 GWh total energy input. As shown in Figure 9c from the GITM simulations, when approximately 60 GWh of total energy was deposited into the upper atmosphere, the thermospheric mass density perturbation was increased by about 4% at 400 km. This shows that the energies are roughly consistent, but GITM needed about 33% more energy than the *Clausen et al.* [2014] estimate. It should be noted that the *Clausen et al.* [2014] study did not include chemistry or horizontal advection, since it was a 1D simulation, while GITM included thermodynamic terms such as advection and adiabatic cooling, which could account for the difference.

The estimation by *Ahn et al.* [1983] was that 20% of the total energy input was due to the particle precipitation, leaving about 80% to Joule heating. From Figure 9c, during Substorm 3, the total energy input to the upper atmosphere was about 102 GWh, with particle precipitation

about 30 GWh and Joule heating about 72 GWh. From the estimation results, the particle precipitation was about 30% of the total energy input. The ratio changed as a function of the strength of the electric field, since with a strong electric field, with a small increase in hemispheric power, the Joule heating can increase dramatically.

From Figure 9, the thermospheric mass density response was dependent on the type of energy input. In order to understand why this might be the case, the temporal response of the upper atmosphere during Substorm 3 was investigated further. The altitude variation of the temperature enhancement, the mass density perturbation, the enhancement of the Joule heating energy, and neutral gas heating rates due to Joule heating, auroral heating and chemical heating were contoured as a function of altitude and substorm epoch time in Figure 10. The white dashed vertical line indicates the substorm onset. The changes of the parameters before the white dashed line were caused by the IMF Bz variation and the changes of the parameters after the white line were mainly due to the HP variations. As illustrated in Figure 10a, the temperature perturbations were almost uniform with height above about 250 km. The maximum temperature enhancement caused by the IMF Bz was about 12 K. The maximum temperature enhancement after the HP increase was about 18 K. The corresponding density perturbations increased with altitude shown in Figure 10b, which is consistent with the studies by *Thayer et al.* [2012] and *Liu et al.* [2014], who concluded that the mass density increased with altitudes below the oxygen/helium transition because the mass density is an integral of all the density scale height change below, and the density scale height below the oxygen/helium composition transition region is mainly determined by the temperature change during a heating event. As shown in Figure 10b, there was

363 a density enhancement before the substorm onset due to the IMF Bz variation. During the
364 substorm, a more significant density perturbation was caused by the auroral variations, which
365 corresponded to a larger temperature enhancement after the substorm onset shown in Figure 10a.
366 The maximum density perturbation was about 8% at 600 km. Note that there was a slight
367 negative mass density perturbation during the substorm recovery phase, especially below 200 km.
368 As shown in Figure 10c, the Joule heating energy perturbation due to the HP increase
369 (after 00:00 epoch) was larger than that due to the IMF Bz enhancement (before 00:00 epoch). A
370 lower Joule heating energy compared to the no substorm simulation occurred during substorm
371 recovery phase (i.e., the Joule heating energy was negative). The possible explanation for this
372 phenomenon was that the O/N_2 ratio decreased during the substorm due to the atmospheric
373 expansion, which caused the total electron density to decrease compared to the no substorm case,
374 which subsequently led to the Joule heating decrease. Most of the Joule heating energy was
375 deposited between 100-150 km, but, as shown in Figure 10d, the neutral gas heating rate due to
376 Joule heating was larger at higher altitudes than at lower altitudes because the mass density
377 decreased exponentially with height [e.g. *Bank and Cockarts*, 1973]. If the transport of energy is
378 excluded from consideration, the main sources and sinks for the thermospheric heating
379 enhancement during the substorm include the Joule heating (Figure 10d), auroral heating (10e),
380 chemical heating (10f), heat conduction (not shown), and radiative cooling (not shown). It should
381 be noted that each of these were shown with different scales, with the Joule heating being largest
382 by almost a factor of two. The auroral heating being next largest, and the chemical heating scale
383 being smaller than the Joule heating scale by almost a factor of 10. As illustrated in Figure 10d,

384 the neutral gas heating rate due to Joule heating increased during both the IMF Bz and HP
385 enhancement. The enhancement of the IMF Bz altered the electric field, while the increase of the
386 HP caused the enhancement of the electron density in the upper atmosphere as described above.
387 However, the heating rate enhancement was larger during the IMF Bz variation than that during
388 the HP change. During the substorm, as the aurora increased, the electron density was most
389 strongly perturbed in the E-region, which resulted in significantly increased Joule heating energy
390 deposition in this region. The F-region electron density was not strongly affected, so the neutral
391 gas heating rate due to Joule heating, were not greatly affected in F-region. So, during the
392 substorm, the majority of the temperature enhancement due to the Joule heating occurred low in
393 the thermosphere. After the substorm, the neutral gas heating rate became negative, indicating
394 that it was lower than the neutral gas heating rate in the background simulation. Since the
395 electric fields were the same between the background and substorm simulations after the
396 substorm, the main reason that the neutral gas heating rate would be different in the substorm run
397 would be because of a change in the ion/electron density. As described above, large electric
398 fields on the nightside can drive downward flows of the ions, which reduce the density. This
399 would reduce the neutral gas heating rate, causing a lower temperature after the substorm.
400 The auroral heating rate, shown in Figure 10e, increased significantly after the substorm
401 onset because the enhancement of the HP increased the particle precipitation. The enhancement
402 of the auroral neutral gas heating rate decreased as altitude decreased because of the mass
403 density dependence on the heating rate. The majority of the energy was deposited in the E-region,
404 but the temperature increase was largest at higher altitudes. As shown in Figure 10f, before the

substorm onset, the chemical heating rate increased slightly due to the IMF Bz increase, and the chemical heating rate increased significantly after the substorm onset when the HP increased. The chemical heating peaked at about 300 km. The recombination rate in the E region was higher than in the F region, hence more heat was release at the low altitude, but due to the lower heat capacity at higher altitude, the temperature enhancement was larger at higher altitude. The enhancement of the chemical heating was about an order of magnitude smaller than the (peak) enhancement of the neutral gas heating rate due to Joule and auroral heating. The radiative cooling rate variation, which was not shown in this figure, was about an order magnitude smaller than the chemical heating rate variation. The main cooling in the high latitude region was caused by three things: (1) advection of the heating out of the region to lower latitudes; (2) adiabatic cooling due to the divergence of wind away from the high-latitudes; and (3) conduction of the heat from the upper thermosphere to the cooler lower thermosphere.

In order to further understand how the changes of the energy heating rates affect the temperature variation during the substorm, Figure 11a shows the temperature enhancement at 150 km and 500 km as a function of the substorm epoch time. The temperature increased due to IMF Bz variations about 1 hour ahead of the substorm onset and the temperature enhancement peaked around 1 hour epoch time, which lagged behind the HP peak by about half an hour. The temperature enhancement during the substorm at 500 km was about two times the temperature enhancement at 150 km, and the temperature enhancement had more fluctuation at 500 km than at 150 km due to the variation of the heat conduction (not shown here). Figure 11b shows the sum of the main energy heating rate (JouleHeatingRate+ AuroraHeatingRate +

426 ChemicalHeatingRate – RadiativeCooling) in the unit of K/s. There were two peaks in the upper
427 atmospheric heating rate corresponding to the two types of energy input (IMF Bz and HP). The
428 values of these two peaks were more comparable at 500 km than at 150 km, which indicates that
429 the heating of the atmosphere due to the IMF Bz variation was much smaller than that due to the
430 HP at 150 km, while the heating of the atmosphere due to IMF Bz and HP variations were closer
431 to each other at 500 km. Note that if the heat conduction, horizontal and vertical advection, and
432 adiabatic cooling were included, the heating rate calculation in the Figure 11b would be
433 significantly less than what was shown in Figure 11b and would become negative after 01:00
434 epoch time, indicating that cooling was overwhelming the heating after 01:00, which caused the
435 temperature enhancement to decrease as shown in Figure 11a.

436 ## 5. Summary

437 The spatial and temporal variations of the thermospheric mass density during a series of
438 different idealized substorms were studied using the Global Ionosphere Thermosphere Model.
439 From these simulations, the following conclusions were made:

440 • For the substorm with a peak hemisphere power enhancement of 50 GW and peak IMF
441 Bz variation of -2 nT (close to the median value of the results in *Clausen et al.* [2014]),
442 the corresponding peak mass density perturbation from GITM simulation was ~14%
443 about 50 min after the substorm onset. The maximum peak density occurred in the
444 nightside sector of the auroral zone. With similar hemisphere power and IMF Bz

445 enhancement, the peak mass density perturbation from CHAMP measurement was about
446 6% by the study of *Clausen et al.* [2014].

447 • The mass density response to the IMF Bz and auroral inputs had a strong local time
448 dependence. The mass density perturbation due the IMF Bz variation peaked in the
449 15:00-21:00 sector while the density perturbation due to HP input peaked at the 21:00-
450 03:00 sector.

451 • During the substorm recovery phase, a negative mass density perturbation(~ -5%)
452 occurred on the night in an isolated region. The lower mass density in post-substorm
453 period was shown to exist in at least two substorms measured by the CHAMP satellite
454 during October 2013. These were the only substorms explored, and both showed the
455 reduced density, although the *Clausen et al.* [2014] study did not statistically show a
456 decreased density in this region. It is unknown why this discrepancy exists.

457 • The mass density perturbation due to both the IMF Bz and HP variations together was
458 similar to the sum of the density perturbations with HP and Bz variations considered
459 separately, which suggests the mass density response to the IMF Bz and HP energy
460 inputs were almost linear, or, were not correlated during the substorm. The non-linearity
461 of the mass density response to different energy input for these 5 substorm was less than
462 6%, but grew with the amount of energy. It may very well be that the system becomes
463 highly nonlinear during extended energy input periods, such as during storms.

- The neutral gas heating rate due to Joule heating and the auroral heating rate were similar magnitudes but had different altitude distributions, with the auroral heating occurring higher in the atmosphere than the Joule heating. The temperature enhancement started to decrease about 20 minutes after the peak in the substorm due to the combination effect of all the heating rates changes and the heat conduction, horizontal and vertical advection, and adiabatic cooling process.

Acknowledgments. This work is supported by the NSF grants CNS1035236 and AGS1010812, and AFOSR grant FA9550-12-1-0401. The data from GITM simulation for this study are available upon request from the authors. All geomagnetic indices can be downloaded at http://cdaweb.gsfc. nasa.gov/. Dr. Eric Sutton provides the CHAMP data used in this study at http://sisko.colorado.edu/sutton/data.html.

475 **References**

476 A, E., A. J. Ridley, D. Zhang, and Z. Xiao (2012), Analyzing the hemispheric asymmetry in the
477 thermospheric density response to geomagnetic storms, Journal of Geophysical Research:
478 Space Physics, 117(A8), A08317, doi:10.1029/2011JA017259.
479 Ahn, B. H., S. I. Akasofu, and Y. Kamide (1983), The Joule heat production rate and the particle
480 energy injection rate as a function of the geomagnetic indices AE and AL, Journal of
481 Geophysical Research: Space Physics, 88(A8), 6275-6287, doi:10.1029/JA088iA08p06275.
482 Banks, P. M., and G. Kockarts (1973), Aeronomy, part B, Academic, San Diego, Calif, 259.
483 Bruinsma, S. L., and J. M. Forbes (2009), Properties of traveling atmospheric disturbances
484 (TADs) inferred from CHAMP accelerometer observations, *Advances in Space Research*,
485 *43*(3), 369-376, doi:http://dx.doi.org/10.1016/j.asr.2008.10.031.
486 Clausen, L. B. N., S. E. Milan, and A. Grocott (2014), Thermospheric density perturbations in
487 response to substorms, Journal of Geophysical Research: Space Physics, 119(6), 4441-4455,
488 doi:10.1002/2014JA019837.
489 Deng, Y and A. J. Ridley (2006), Role of vertical ion convection in the high-latitude ionospheric
490 plasma distribution, *J. Geophys. Res., 111*, A09314, doi:10.1029/2006JA011637.
491 Deng, Y., T. J. Fuller-Rowell, R. A. Akmaev, and A. J. Ridley (2011), Impact of the altitudinal
492 Joule heating distribution on the thermosphere, Journal of Geophysical Research: Space
493 Physics, 116(A5), A05313, doi:10.1029/2010JA016019.
494 Emmert, J., J. Lean, and J. Picone (2010), Record-low thermospheric density during the 2008
495 solar minimum, Geophysical Research Letters, 37(12), doi:10.1029/2010GL043671..

496 Emmert, J. T., J. M. Picone, J. L. Lean, and S. H. Knowles (2004), Global change in the
497 thermosphere: Compelling evidence of a secular decrease in density, *Journal of Geophysical*
498 *Research: Space Physics*, *109*(A2), A02301, doi:10.1029/2003JA010176.
499 Fuller-Rowell, T. J., and D. S. Evans (1987), Height-integrated Pedersen and Hall conductivity
500 patterns inferred from the TIROS-NOAA satellite data, Journal of Geophysical Research:
501 Space Physics, 92(A7), 7606-7618, doi:10.1029/JA092iA07p07606.
502 Fuller-Rowell, T. J., and D. Rees (1981), A three-dimensional, time-dependent simulation of the
503 global dynamical response of the thermosphere to a geomagnetic substorm, Journal of
504 Atmospheric and Terrestrial Physics, 43(7), 701-721, doi:http://dx.doi.org/10.1016/0021-
505 9169(81)90142-2.
506 Keating, G. M., R. H. Tolson, and M. S. Bradford (2000), Evidence of long term global decline
507 in the Earth's thermospheric densities apparently related to anthropogenic effects, *Geophysical*
508 *Research Letters*, *27*(10), 1523-1526, doi:10.1029/2000GL003771.
509 Knipp, D. J., W. K. Tobiska, and B. A. Emery (2004), Direct and indirect thermospheric heating
510 sources for solar cycles 21–23, Sol Phys, 224(1-2), 495-505, doi:10.1007/s11207-005-6393-4..
511 Lei, J., A. G. Burns, J. P. Thayer, W. Wang, M. G. Mlynczak, L. A. Hunt, X. Dou, and E. Sutton
512 (2012), Overcooling in the upper thermosphere during the recovery phase of the 2003 October
513 storms, Journal of Geophysical Research: Space Physics, 117(A3), A03314,
514 doi:10.1029/2011JA016994.

515 Liu, X., J. P. Thayer, A. Burns, W. Wang, and E. Sutton (2014), Altitude variations in the
516 thermosphere mass density response to geomagnetic activity during the recent solar minimum,
517 Journal of Geophysical Research: Space Physics, 2013JA019453, doi:10.1002/2013ja019453.
518 Lotko, W. (2007), The magnetosphere–ionosphere system from the perspective of plasma
519 circulation: A tutorial, Journal of Atmospheric and Solar-Terrestrial Physics, 69(3), 191-211,
520 doi:http://dx.doi.org/10.1016/j.jastp.2006.08.011.
521 Lu, G., A. D. Richmond, B. A. Emery, and R. G. Roble (1995), Magnetosphere-ionosphere-
522 thermosphere coupling: Effect of neutral winds on energy transfer and field-aligned current,
523 Journal of Geophysical Research: Space Physics, 100(A10), 19643-19659,
524 doi:10.1029/95JA00766.
525 Marcos, F. A., S. T. Lai, C. Y. Huang, C. S. Lin, J. M. Retterer, S. H. Delay, and E. K. Sutton
526 (2010), Towards next level satellite drag modeling, AIAA 2010-7840, paper presented at the
527 AIAA Atmospheric and Space Environments Conference, Toronto, Ontario, Canada, 2–5
528 August.
529 Marcos, F. A., J. O. Wise, M. J. Kendra, N. J. Grossbard, and B. R. Bowman (2005), Detection
530 of a long-term decrease in thermospheric neutral density, *Geophysical Research Letters*, *32*(4),
531 L04103, doi:10.1029/2004GL021269.
532 Paschmann, G., S. Haaland, and R. Treumann (Eds.) (2002), Auroral Plasma Physics, Springer,
533 Newyork.

534 Prölss, G. (1981), Latitudinal structure and extension of the polar atmospheric disturbance,
535 Journal of Geophysical Research: Space Physics (1978-2012), 86(A4), 2385-2396, doi:
536 10.1029/JA086iA04p02385.
537 Qian, L., S. C. Solomon, and T. J. Kane (2009), Seasonal variation of thermospheric density and
538 composition, Journal of Geophysical Research: Space Physics (1978–2012), 114(A1), doi:
539 10.1029/2008JA013643.
540 Reigber, C., H. Lühr, and P. Schwintzer (2002), CHAMP mission status and perspectives, Eos
541 Trans. AGU, 81(48), 48.
542 Ridley, A. J., Y. Deng, and G. Toth (2006), The global ionosphere-thermosphere model, Journal
543 of Atmospheric and Solar-Terrestrial Physics, 68(8), 839-864,
544 doi:10.1016/j.jastp.2006.01.008.
545 Ritter, P., H. Lühr, and E. Doornbos (2010), Substorm-related thermospheric density and wind
546 disturbances derived from CHAMP observations, Ann. Geophys., 28(6), 1207-1220,
547 doi:10.5194/angeo-28-1207-2010.
548 Roble, R. G., R. E. Dickinson, and E. C. Ridley (1982), Global circulation and temperature
549 structure of thermosphere with high-latitude plasma convection, Journal of Geophysical
550 Research: Space Physics, 87(A3), 1599-1614, doi:10.1029/JA087iA03p01599.
551 Roble, R. G., E. C. Ridley, and R. E. Dickinson (1987), On the global mean structure of the
552 thermosphere, Journal of Geophysical Research: Space Physics (1978–2012), 92(A8), 8745-
553 8758, doi: 10.1029/JA092iA08p08745.

554 Solomon, S. C., L. Qian, L. V. Didkovsky, R. A. Viereck, and T. N. Woods (2011), Causes of
555 low thermospheric density during the 2007- 2009 solar minimum, Journal of Geophysical
556 Research: Space Physics (1978-2012), 116(A2), doi: 10.1029/2011JA016508.
557 Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and
558 climate experiment: Mission overview and early results, Geophysical Research Letters, 31(9),
559 doi: 10.1029/2004GL019920.
560 Thayer, J., X. Liu, J. Lei, M. Pilinski, and A. Burns (2012), The impact of helium on
561 thermosphere mass density response to geomagnetic activity during the recent solar minimum,
562 Journal of Geophysical Research, 117(A7), A07315, doi: 10.1029/2012JA017832.
563 Thayer, J. P., and J. Semeter (2004), The convergence of magnetospheric energy flux in the polar
564 atmosphere, Journal of Atmospheric and Solar-Terrestrial Physics, 66(10), 807-824,
565 doi:http://dx.doi.org/10.1016/j.jastp.2004.01.035.
566 Valdivia, J. A., A. S. Sharma, and K. Papadopoulos (1996), Prediction of magnetic storms by
567 nonlinear models, Geophysical Research Letters, 23(21), 2899-2902,
568 doi:10.1029/96GL02828.
569 Volland, H. (1979), Magnetospheric electric fields and currents and their influence on large scale
570 thermospheric circulation and composition, Journal of Atmospheric and Terrestrial Physics,
571 41(7–8), 853-866, doi:http://dx.doi.org/10.1016/0021-9169(79)90128-4.
572 Weimer, D. R. (2005), Improved ionospheric electrodynamic models and application to
573 calculating Joule heating rates, Journal of Geophysical Research: Space Physics, 110(A5),
574 A05306, doi:10.1029/2004JA010884.

575 Wu, J.-G., and H. Lundstedt (1996), Prediction of geomagnetic storms from solar wind data
576 using Elman Recurrent Neural Networks, Geophysical Research Letters, 23(4), 319-322,
577 doi:10.1029/96GL00259.

578

579

580 Xianjing Liu, University of Michigan, Atmospheric Oceanic and Space Sciences, Ann Arbor, MI
581 48105
582 (e-mail:xianjing@umich.edu)

Figure Captions:

Figure 1. The variations of (a) IMF Bz (nT), (b) hemispheric power (GW) and (c) $F_{10.7}$ (SFU) as a function of substorm epoch time during Substorm 1 to 5. The azure dashed line indicates the baseline case.

Figure 2. The spatial variation of the mass density perturbation from 40°N to 90°N at ~400km during Substorm 3 by GITM simulation. The grey rings indicate the latitudes at 40°N, 60°N 80°N. The radial grey lines show each third hour of local time. The substorm epoch time is at a 10 min cadence from -70 min to 220 min. The sun (12:00 LT) is to the top of each panel. Dawn (06:00 LT) is to the right and dusk (18:00 LT) is to the left. The contour of the geomagnetic latitudes is shown in the last subplot.

Figure 3. The longitudinal and latitudinal average mass density perturbation during Substorm 3 by GITM simulation in the auroral zone for four local time sectors: (a) 03:00-09:00 LT, (b) 09:00-15:00 LT, (c) 15:00-21:00 LT, (d) 21:00 -03:00 LT. (e) shows the mass density perturbation average over the entire the auroral zone and local time sector from 00:00 to 24:00 LT. The zero vertical dash line indicates the onset of the substorm expansion phase.

Figure 4. The latitude variations of thermospheric mass density perturbation as a function of substorm epoch time during Substorm 3 at 03:00 LT.

Figure 5. (a, b) The latitude variation of the mass density from the measurement of CHAMP satellite normalized to 400 km for one satellite orbit period, and (c, d) The AE index variation as a function of UT during two substorms in Oct 4, 2003. The black lines show the variation of the density (top pannels) and AE index (bottom pannels) before the substorm onset. The purple and red lines are times during the substorm and the yellow lines indicate times after the substorm.

Figure 6. The variation of the zonal and latitudinal average mass density perturbation over the entire auroral zone during Substorm 1 (black), Substorm 2 (blue), Substorm 3 (red), Substorm 4 (purple) and Substorm 5 (green) as a function of substorm epoch time. The zero vertical dash line indicates the substorm expansion phase onset.

Figure 7. (a-e) The zonal and latitudinal average mass density perturbation over the entire auroral zone for Substorm 1-5 with four runs: Run 1 (red dotted): density perturbation includes both HP and Bz variation, Run 2 (blue): density perturbation due to HP variation only, Run 3 (black): density perturbation due to Bz variation only, Run 4 (green): the sum of the density

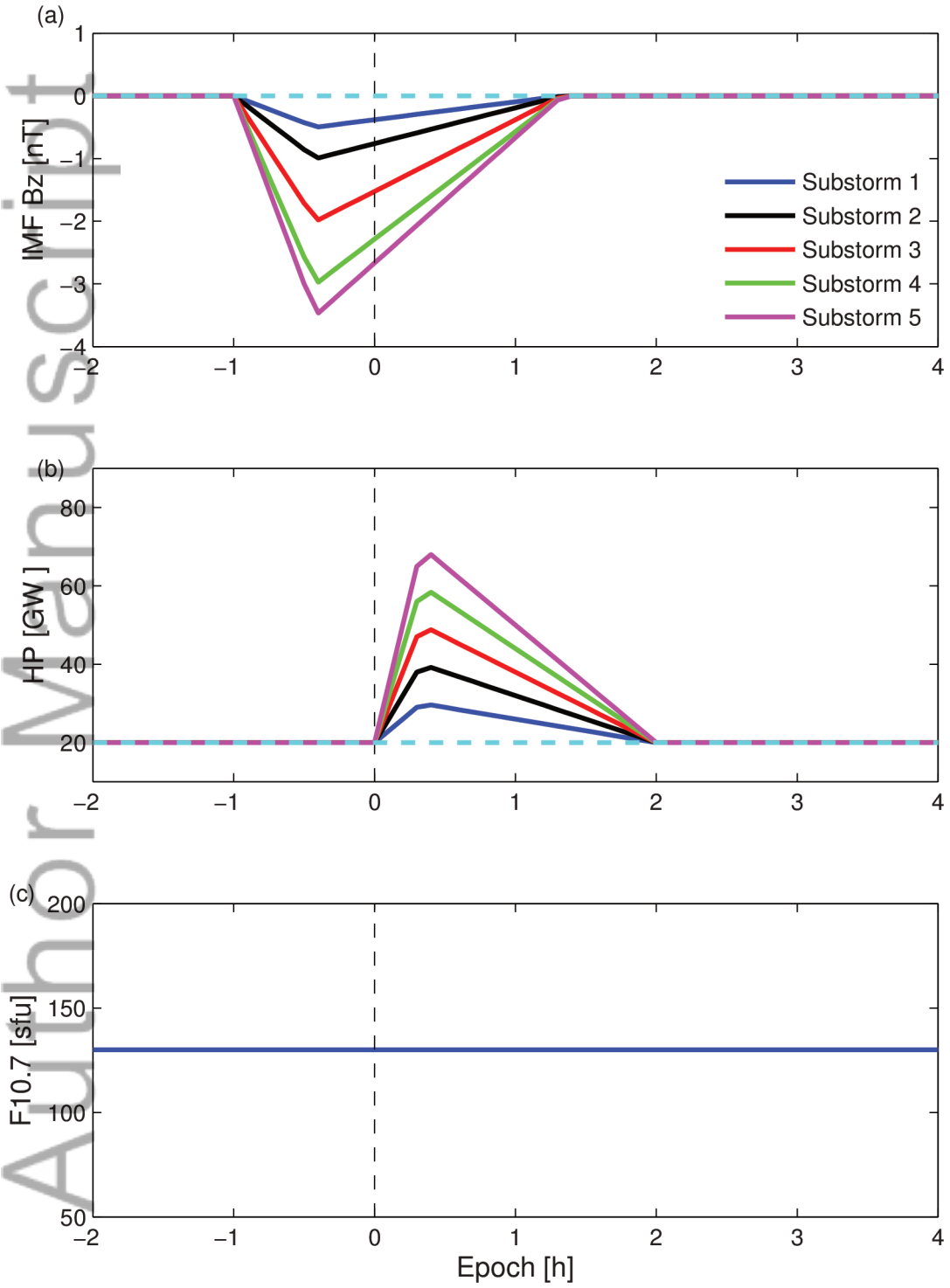
perturbation of Run 2 and Run 3. (f) the non-linear of mass density response to different energy input for substorm 1-5.

Figure 8. The hemispheric total energy and the globally integrated Joule heating total energy (GWh) during Substorm 3 at 400 km.

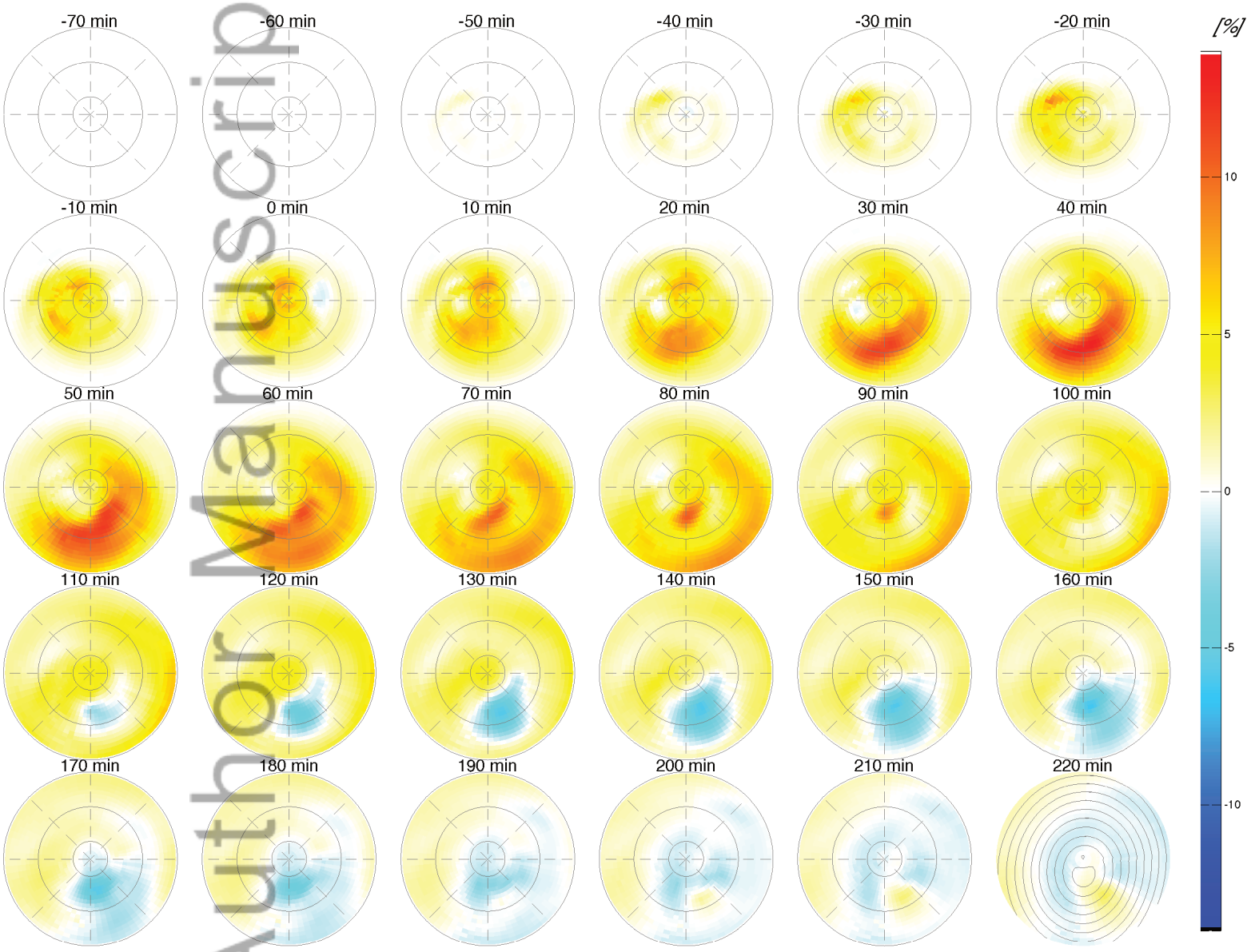
Figure 9. The globally integrated Joule heating energy and hemisphere power enhancement (GWh) with the associated maximum zonal and meridional mean mass density perturbation in the auroral zone (65° N to 75° N) at 150 km (black), 200 km (blue), 300 km (red) and 400 km (green) during Substorm 1-5. The five dots at each altitude indicate Substorm 1 to Substorm 5. Subplot (a) is for the substorms with HP variation only, (b) is for the substorms with IMF Bz variation only, (c) is for the substorms with both HP and IMF Bz variation included.

Figure 10. The altitude variations of (a) the temperature enhancement (K), (b) mass density perturbation (%), (c) Joule heating energy enhancement (W/m^3), (d) neutral gas heating rate enhancement (K/s), (e) auroral heating rate enhancement (K/s) and (f) chemical heating rate enhancement (K/s) averaged over the entire auroral zone as a function of the substorm epoch time during Substorm 3.

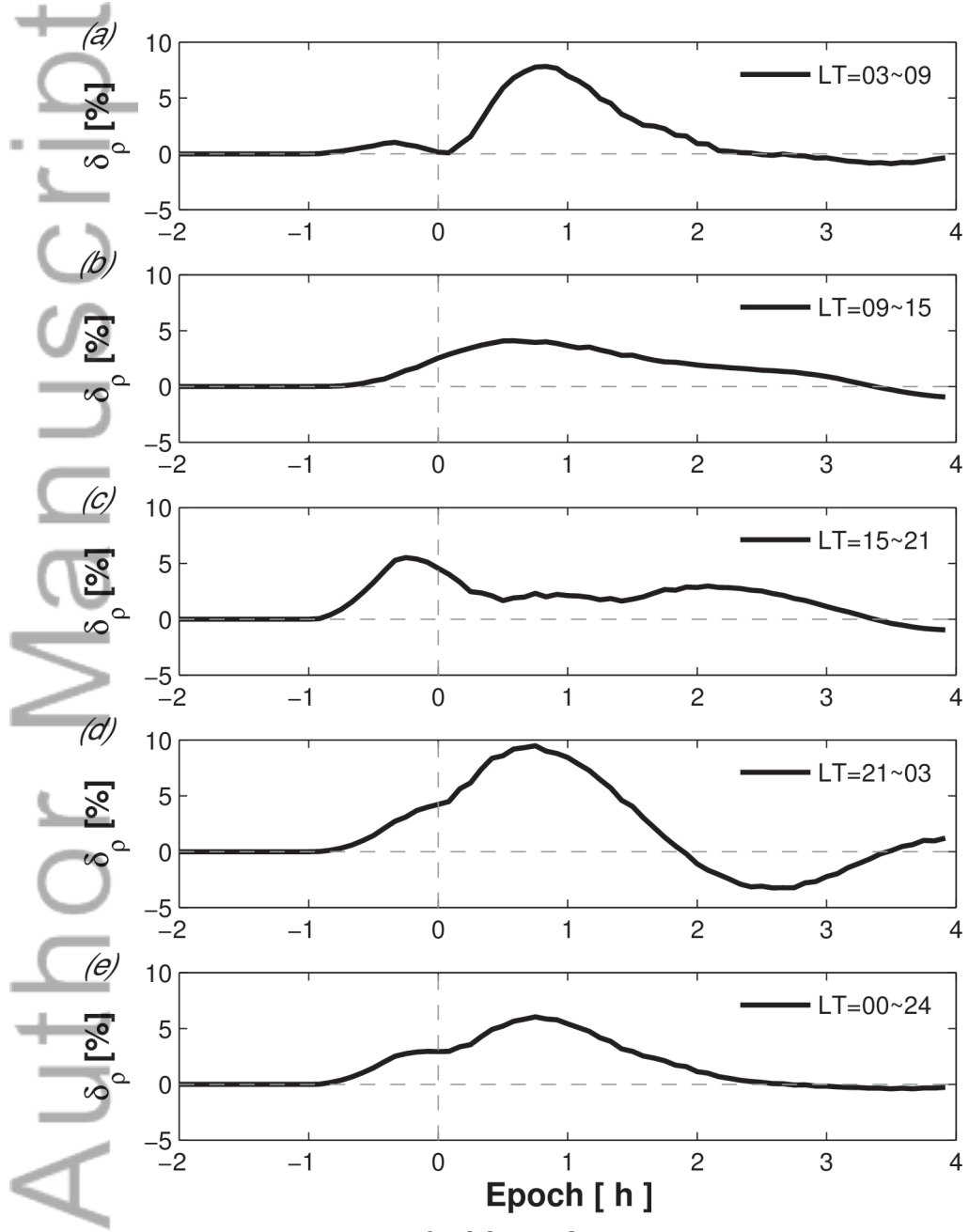
Figure 11. (a) The mean temperature enhancement, (b) the upper atmosphere total heating and cooling rate variations over the entire auroral zone at 150 (black) and 500 (red) km as a function of the substorm epoch time during Substorm 3.



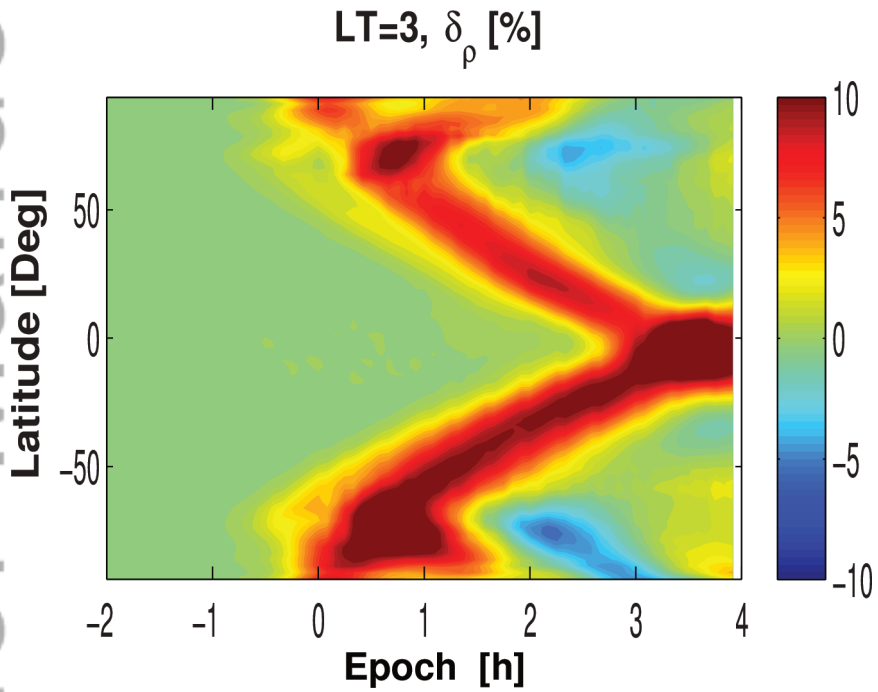
f1_solar_geomagnetic_input_ai



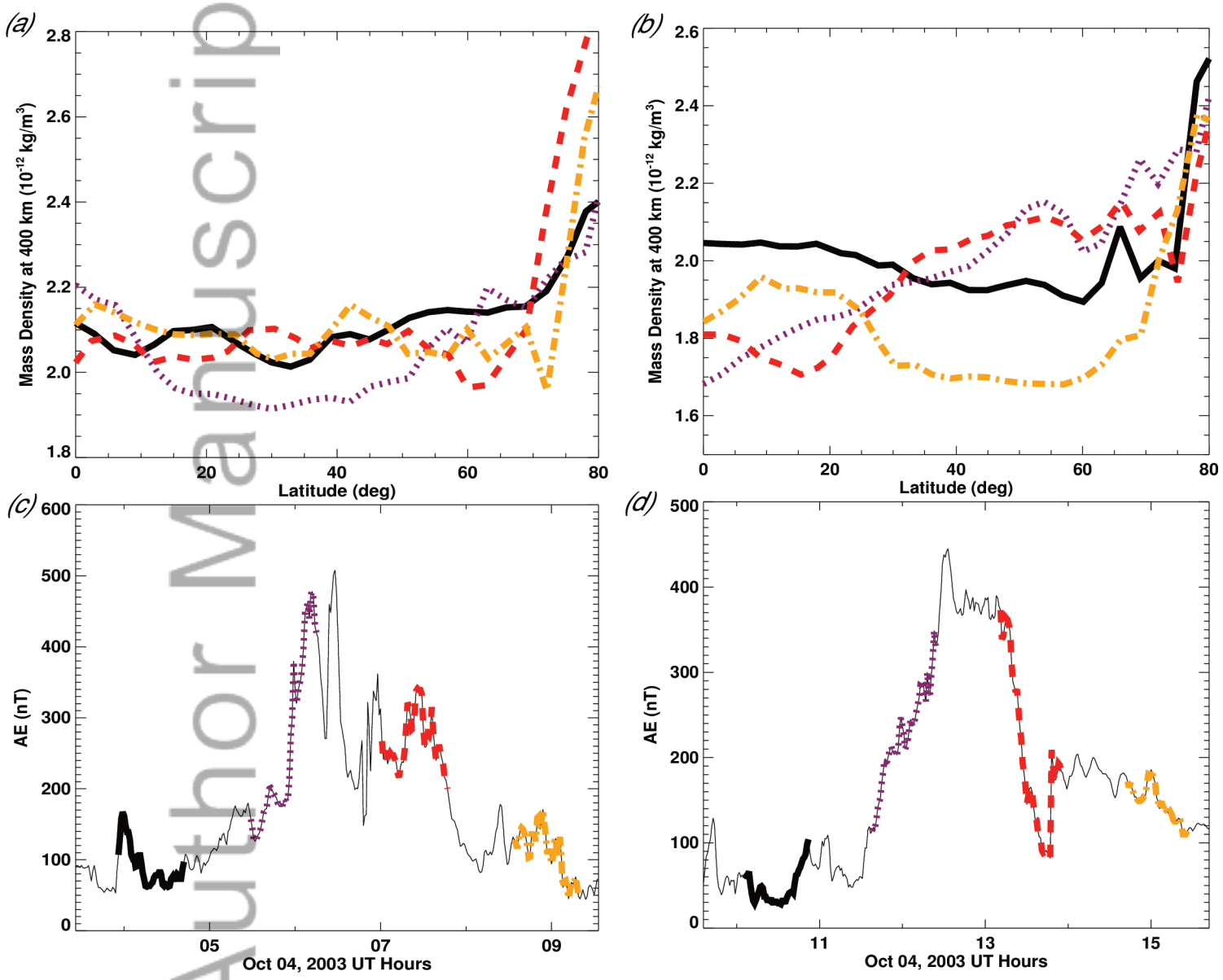
f2_drho_polar_v3_ai



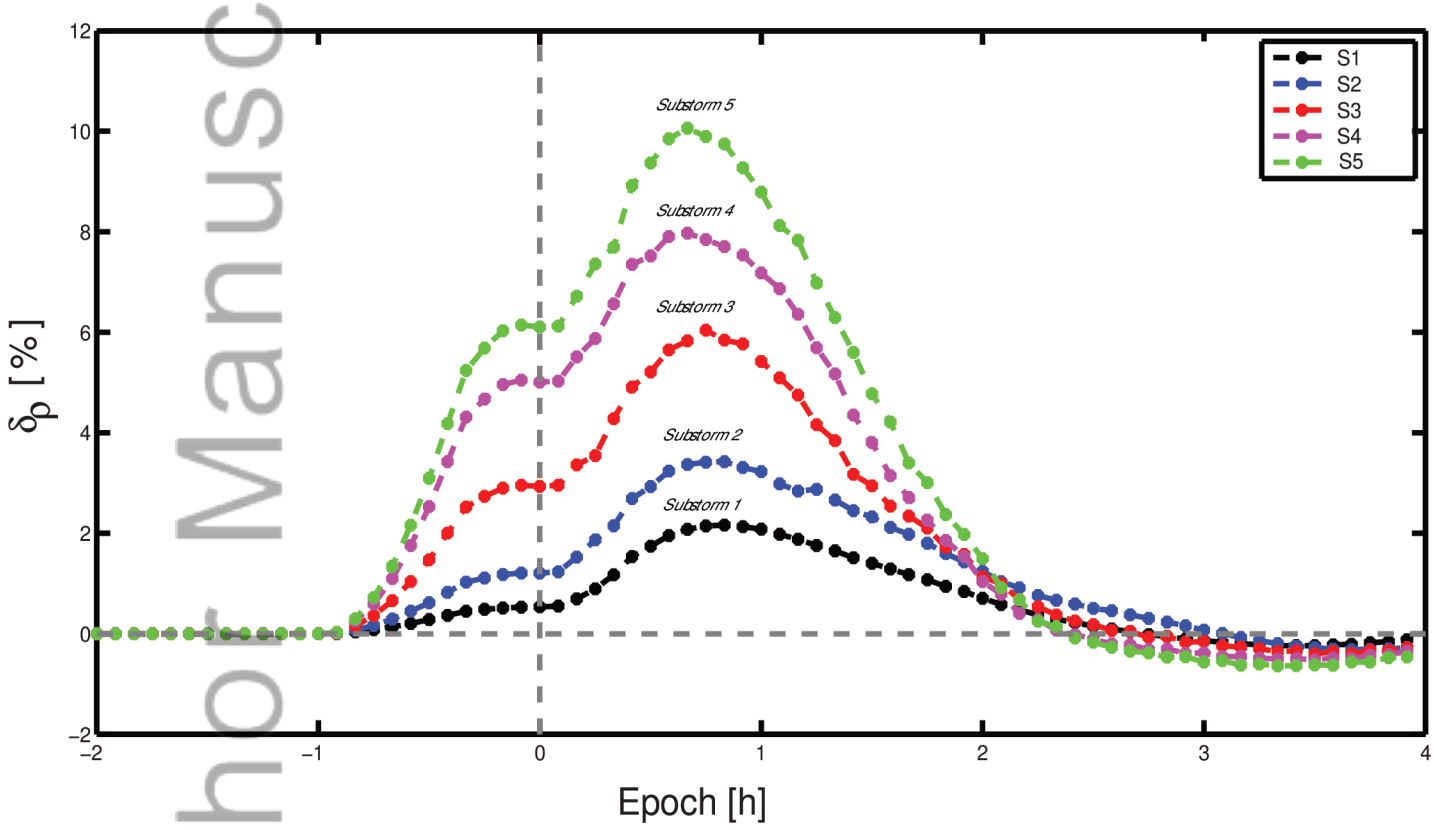
f3_drho_ut_lt_ai



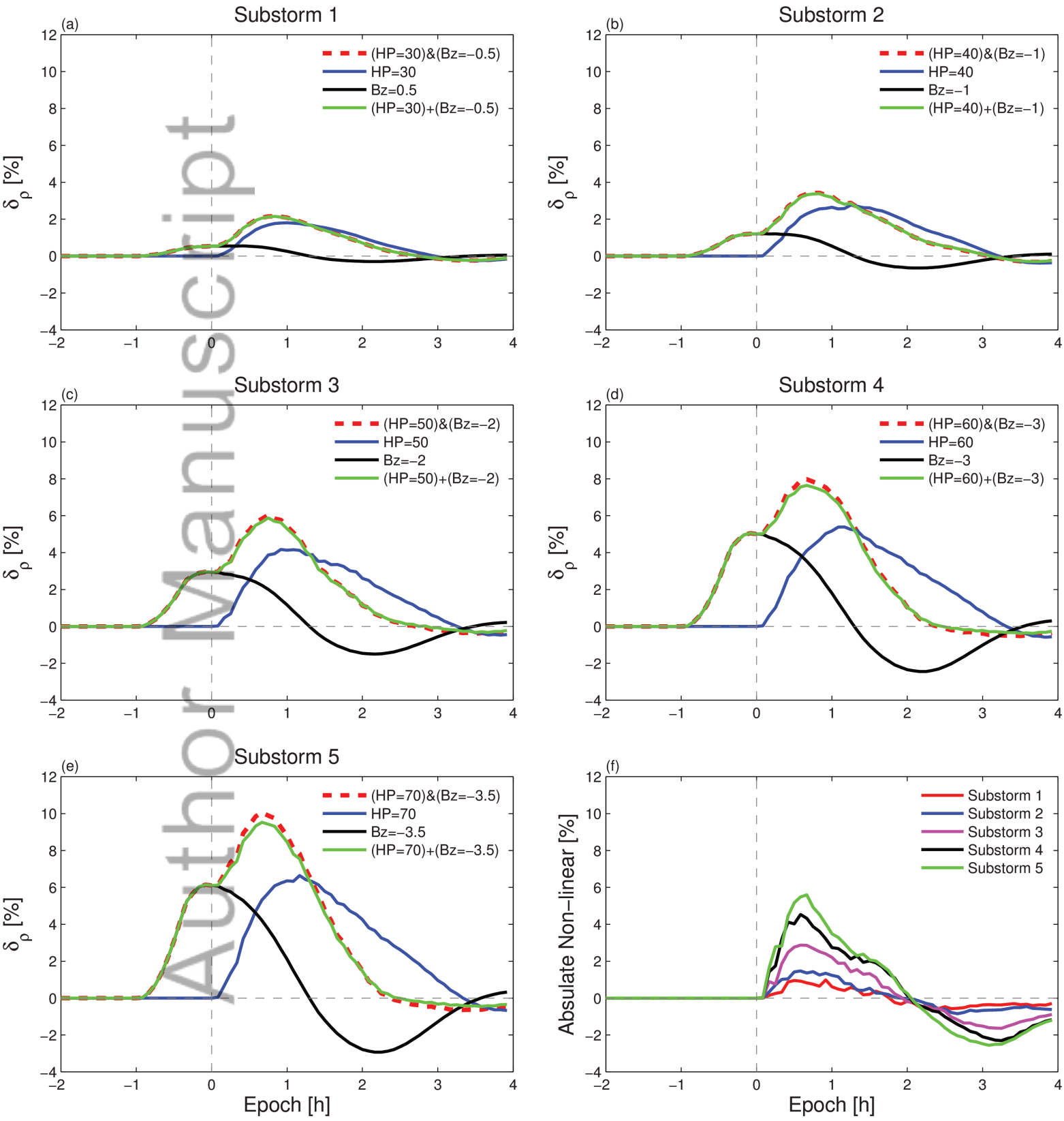
f4_drho_ut_lt_ai



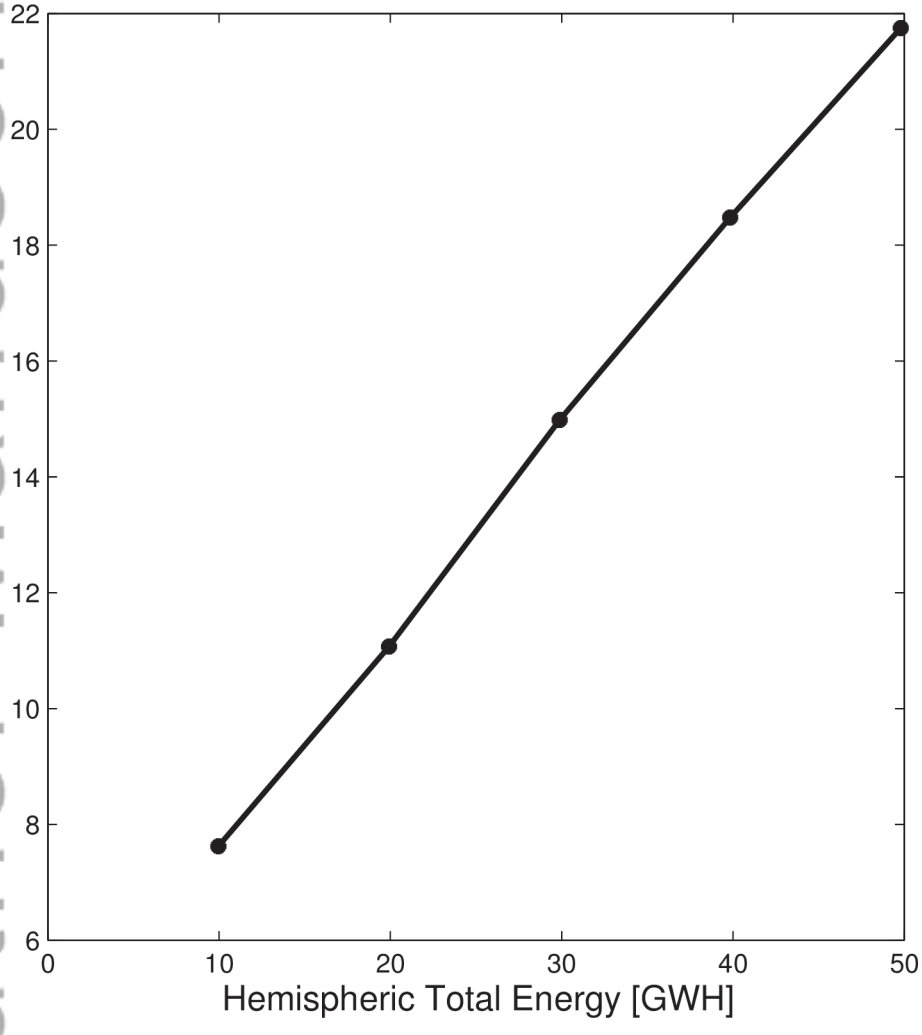
f5_champ_ae_combine_v4



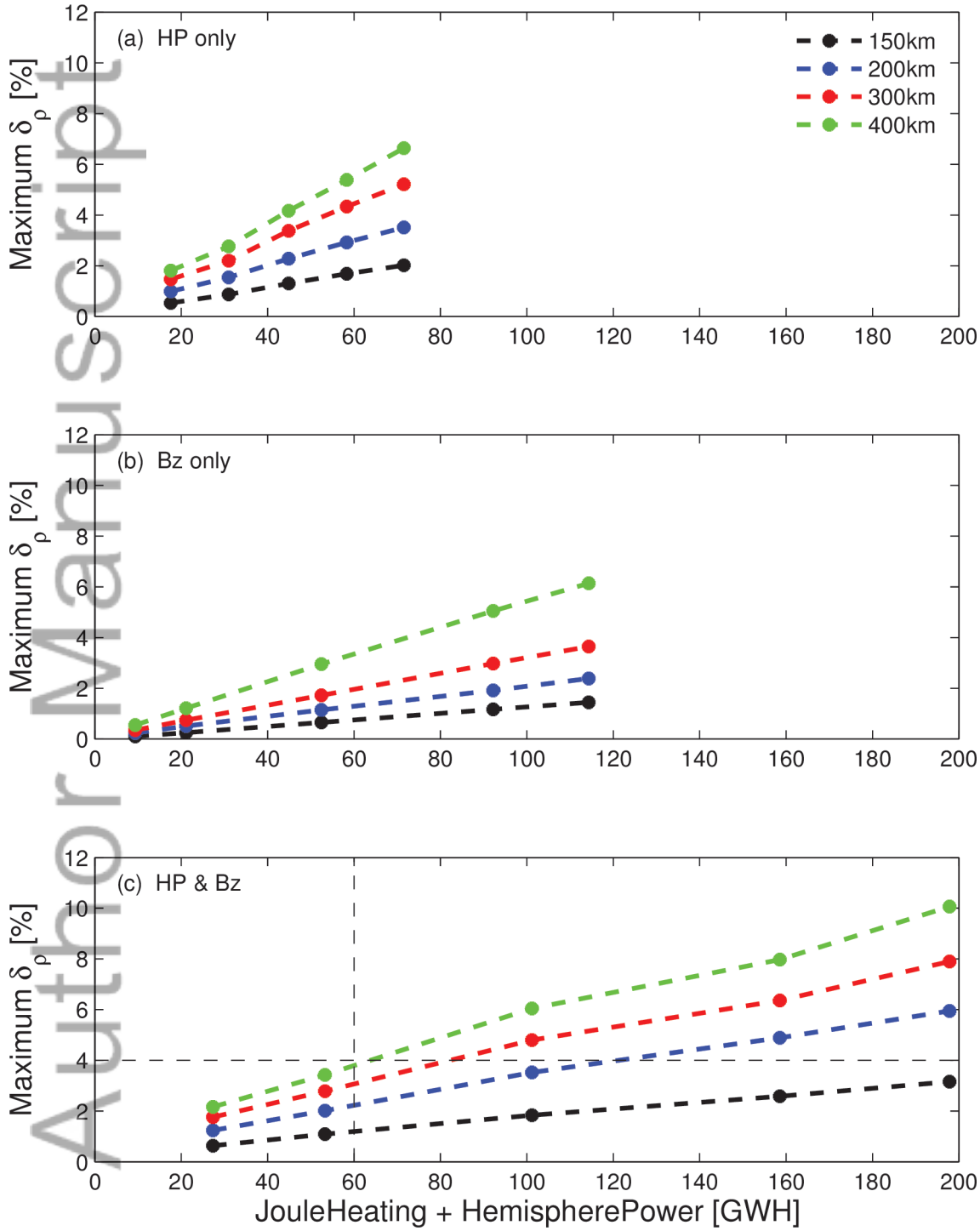
f6_deltarho_vs_ut_ai



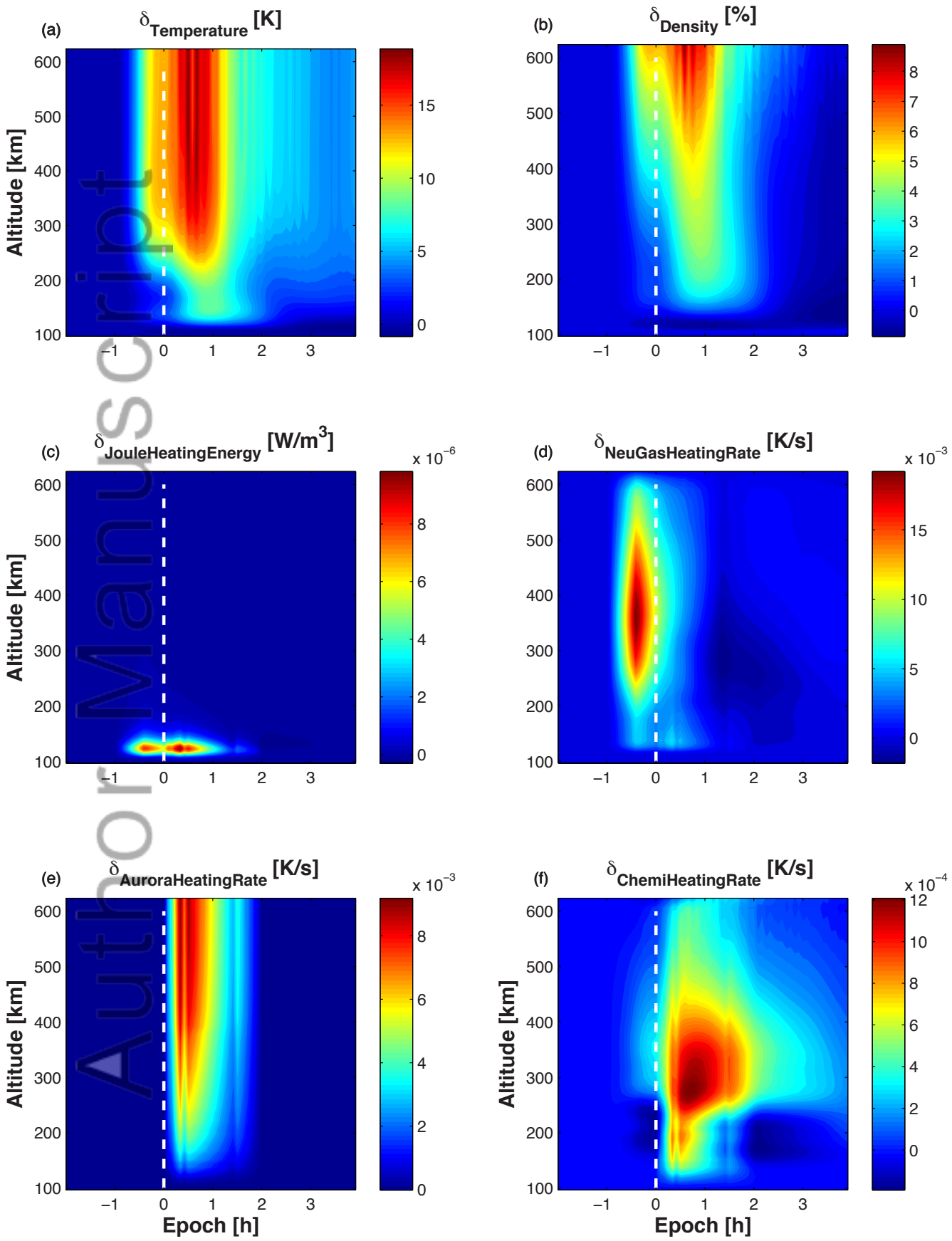
f7_5cases_all_vs_hpi_vs_bz_v5

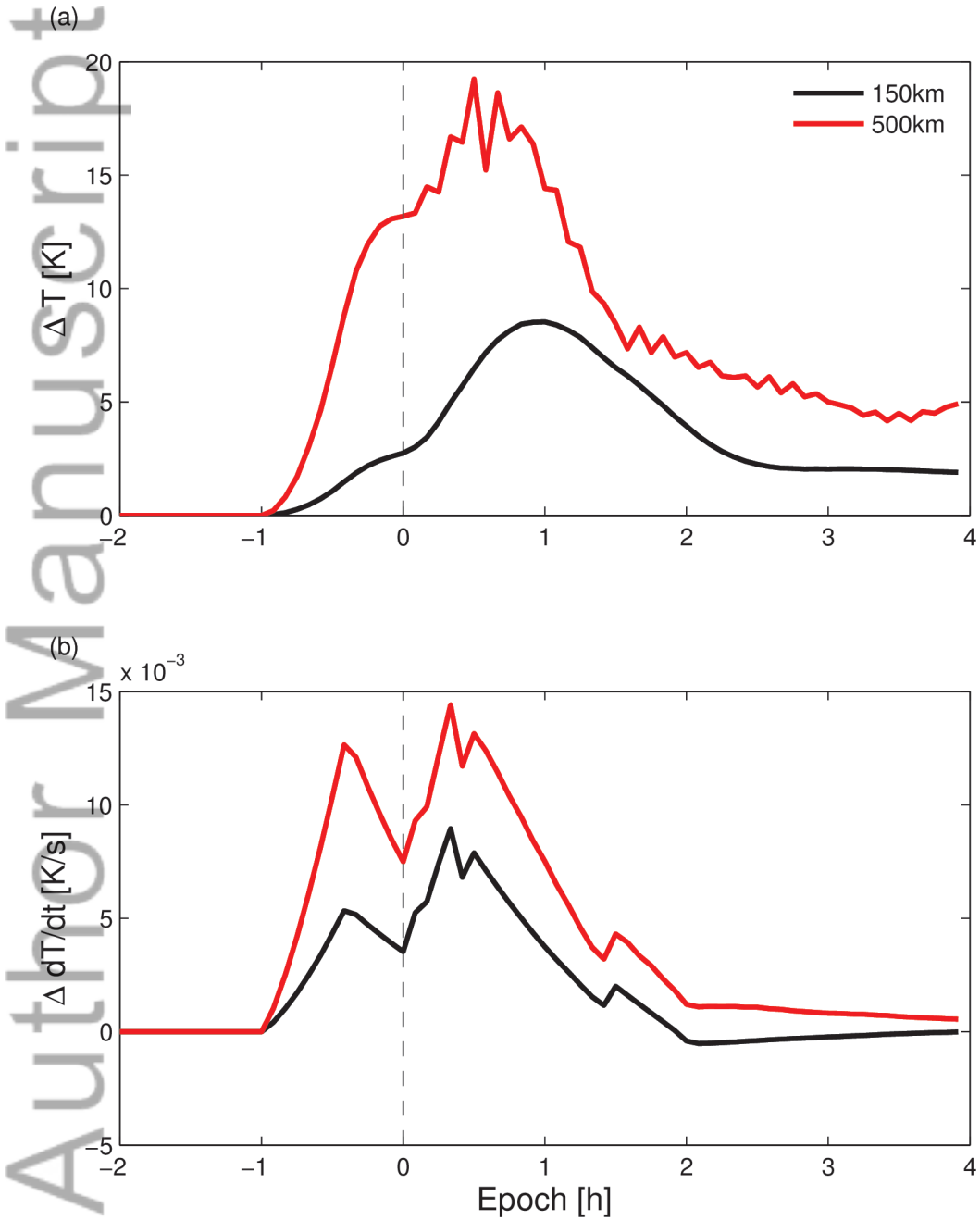


f8_jouleheating_vs_hp_v4



f9_combined_v4





f11_dt_dtdt