5-2023

Studies of the Ionosphere-Thermosphere Responses to Multi-Scale Energy Deposition Processes

Haonan Wu
haonanw@clemson.edu

Studies of the Ionosphere-Thermosphere Responses to Multi-Scale Energy Deposition Processes

---

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Physics

---

by
Haonan Wu
May 2023

Accepted by:
Dr. Xian Lu, Committee Chair
Dr. Xinzhao Chu
Dr. Stephen Kaeppler
Dr. Gerald Lehmacher
Dr. Marco Ajello

Abstract

The Ionosphere-Thermosphere (I-T) system is greatly affected by the magnetospheric energy deposition from above and subjected to forcing from the lower atmosphere simultaneously. A central problem in studying the coupled I-T system is to analyze the multi-scale processes naturally embedded in both ways. Magnetospheric energy input such as auroral precipitation and electric field demonstrates multi-scale structures during magnetic storms, resulting in multi-scale I-T responses when deposited into the I-T system. To better quantify the multi-scale aurora and electric field, we developed a new data assimilation model based on a multi-resolution Gaussian process model to synthesize empirical models and observational data from various sources and provide estimates in regions without observations. The new method mitigates the discrepancy between empirical models and observations by successfully capturing the dynamic evolutions of large-scale and mesoscale auroral and electric field structures. By further incorporating the assimilated aurora and electric fields into Thermosphere Ionosphere Electrodynamics General Circulation Model (TIEGCM) during the 2015 St. Patrick's Day storm, we significantly elevate Joule heating and largely reproduce the global and local I-T responses as observed, including the enhanced electron density and vertical wind. Data assimilation also helps introduce more spatial and temporal variabilities in TIEGCM, which propagate to low-latitude regions through Traveling Atmospheric Disturbance (TAD). In the other

direction, to study the atmospheric wave forcing from below and how it impacts the I-T system, we develop a nested-grid extension to TIEGCM to study the Gravity Wave (GW) propagation process and its ionospheric effect during the 2022 Tonga volcano eruption. Such a hybrid-grid design helps to better simulate the variations of a smaller scale than the standard model resolution while reducing computation costs at the same time. With proper seeding at the lower boundary, GW propagation in the thermosphere is successfully reproduced. The resulting Traveling Ionospheric Disturbance (TID) in the ionosphere has a similar speed to observations. The wave spectrum at different altitudes also indicates that the dominant GW has a shorter period and horizontal wavelength at higher altitudes. This dissertation discusses the detailed tool development, including data assimilation and TIEGCM-NG, which enables a better understanding of the influences of multi-scale magnetospheric forcing and lower-atmosphere wave forcing on the I-T system. This work provides a powerful set of tools for a better simulation of space weather.

Acknowledgments

I sincerely thank my adviser, Prof. Xian Lu, for her invaluable supervision and support of my PhD studies. She has provided me with tremendous help and encouragement for my graduate career at Clemson University. She made every effort to enable enlightening discussions, which stimulated great ideas about science and helped me to develop the new data assimilation model and TIEGCM-NG. She always encouraged me to be innovative, incisive, and extensive. I would also like to thank other committee members, Prof. Xinzhao Chu, Prof. Stephen Kaeppler, Prof. Gerald Lehmacher, and Prof. Marco Ajello, for their insightful discussions and suggestions on my research. Prof. Chu has also provided precious observational data, which triggered a motivation for the first topic of my model studies. I also want to thank Prof. Jens Oberheide for his support and helpful discussions about my research. Many thanks go to other faculty members in the Department of Physics and Astronomy at Clemson University for their excellent teaching. I also appreciate the help I received from the scientists at the High Altitude Observatory, National Center for Atmospheric Research. It would not be possible for me to finish my PhD degree without their selfless supervision and support.

Table of Contents

List of Tables

List of Figures

xi

xii

xiii

xiv

Chapter 1

Introduction

The thermosphere is the topmost layer of the atmosphere where the concentration of charged particles is high enough to have a significant impact. Due to different collision frequencies of electrons and ions with neutral particles and the resultant different mobilities, charge separation occurs, and electric fields form. Electric fields then react back on charged particles leading to the differential motion of neutral and charged particles. The collisional forces between neutral and charged particles, which depend on the relative velocities between two species, accelerate one at slow speeds. At the same time, deaccelerate the other at high speeds resulting in momentum and energy exchanges. The changes in the spatial distribution of charged particles feed back into the corresponding altering of electric fields. Due to the intimate nature of neutral and charged particles and the entangled relation between charges and fields, the ionosphere and thermosphere are often studied together and referred to as the Ionosphere-Thermosphere (I-T) system (Kelley, 2009).

Since ions are greatly affected by magnetic fields through Lorentz forces, the ionosphere is sensitive to geomagnetic field configurations. In high latitudes (polar regions) where geomagnetic field lines are connected to the magnetosphere, the

ionosphere is highly influenced by the interaction between the solar wind and magnetosphere through magnetic coupling processes. When the dayside Interplanetary Magnetic Field (IMF) points southward, the magnetic fields reconnect from solar wind to earth. Plasmas travel along geomagnetic field lines to the ionosphere, resulting in auroral activities (Dungey, 1961). During disturbed geomagnetic conditions, energy deposition from the magnetosphere in the form of electromagnetic waves and particle precipitation can reach 3 to 5 times than quiet time, far exceeding the energy input from upward propagating atmospheric waves (Knipp et al., 2011). Massive magnetospheric energy heats the atmosphere, changes the local circulation, and leads to atmospheric disturbances like temperature enhancement, density variation, and traveling waves (Lei et al., 2010; Bruinsma and Forbes, 2010).

In the other direction, the upward-propagating atmospheric wave is a persistent process that transports momentum and energy into the I-T region, leading to perturbations of various scales. From large-scale waves like tides and planetary waves to regional-scale Gravity Waves (GWs), atmospheric waves show different spectra and propagation properties, leading to distinct behaviors in the I-T system. For example, the interaction between planetary waves with tides generally causes large-scale features like the wave-4 structures in the ionospheric electron densities (Kil et al., 2008; Oberheide et al., 2011). At the same time, GWs have an imprint on the ionospheric E and F regions in the form of Traveling Ionospheric Disturbances (TIDs) (Hines, 1960; Fritts and Lund, 2011) and or sporadic E layers (van Eyken et al., 1982). The nonlinear wave-wave and wave-mean coupling processes further complicate the picture through cross-scale interactions, a significant source for the I-T variabilities. Such wave-induced I-T variabilities are studied extensively in literature (Holton, 1983; Fritts, 1984).

1.1 Observations

A series of satellite and ground-based observations have observed significant variations of the I-T system during geomagnetic storms. Instruments onboard low-earth orbit satellites provide sustained high-quality measurements of the near-earth environment. Special Sensor Ultraviolet Spectrographic Imager (SSUSI) (Paxton et al., 1992) and Special Sensor for Ion and Electron Scintillation (SSIES) (Rich, 1994) onboard Defense Meteorological Satellites Program (DMSP) provide precious auroral image data and ion drift measurements. Auroral particle precipitation is found to heat the atmosphere directly by collision and indirectly by changing conductivities (Banks, 1977). Downward Poynting flux on the order of $50\,\mathrm{mW\,m^{-2}}$ to $100\,\mathrm{mW\,m^{-2}}$ associated with sizeable electric field perturbations ($100\,\mathrm{mV\,m^{-1}}$) is also detected and further converted into Joule heating (Wygant et al., 2000). The ionospheric current system also varies. For example, Field-Aligned Current (FAC) derived from AMPERE can increase by over $100\,\%$ during storm time (Anderson et al., 2002, 2014).

Figure 1.1 shows a daily summary plot of auroral irradiances by DMSP/SSUSI during the 2015 St. Patrick's Day storm (to be discussed in Chapter 5). During this one-day interval, the satellite passes both poles multiple times and records auroral airglow at 5 different bands: $121.6\,\mathrm{nm}$, $130.4\,\mathrm{nm}$, $135.6\,\mathrm{nm}$, $140\,\mathrm{nm}$ to $150\,\mathrm{nm}$ (Lyman-Birge-Hopfield (LBH) short), and $165\,\mathrm{nm}$ to $180\,\mathrm{nm}$ (LBH long). The figure shows optical brightening at the auroral oval, indicating enhanced auroral activities.

Ground-based auroral observation network like Time History of Events and Macroscale Interactions during Substorms (THEMIS) All-Sky Imager (ASI) has also expanded and updated to cover more regions and reach higher precisions. THEMIS ASI array monitors the auroral activities in northern America since 2006 (Gabrielse
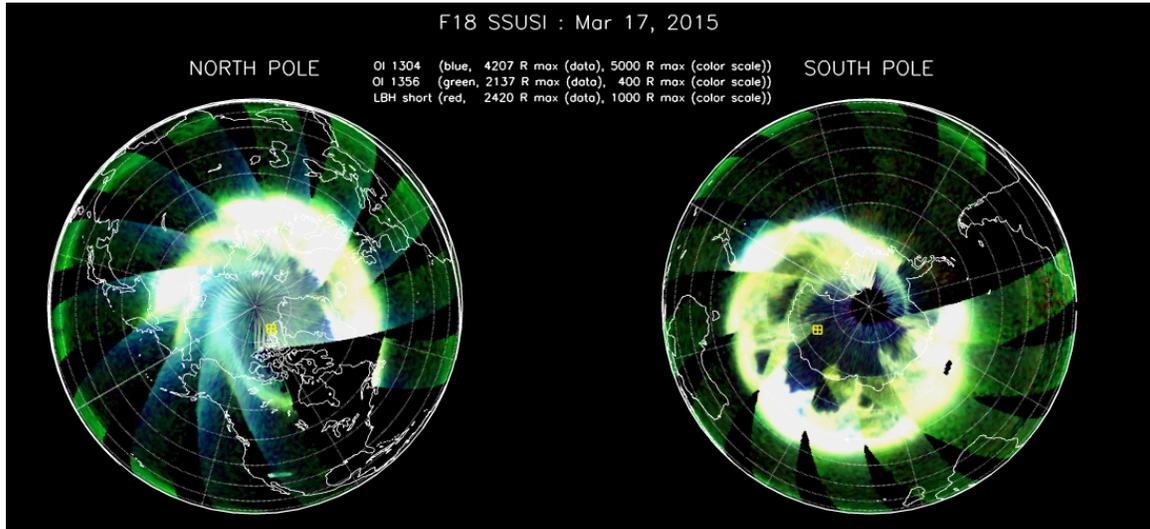
Figure 1.1: SSUSI auroral irradiance daily summary plot on 2015-03-17. This figure is reconstructed from optical measurements at 130.4 nm, 135.6 nm and LBH-short bands. Courtesy SSUSI website.

et al., 2021). Figure 1.2 depicts a typical picture of the ground-based auroral measurements of energy fluxes. Clear enhancements are present in the middle of the graph. At the same time, the aurora is not intense at a similar latitude but a different longitude, showing spatial dependencies. The continuous observations of the spatial distributions provide another important data set for analyzing the long-term and short-term variations of the aurora.

Besides auroral observations like DMSP/SSUSI and THEMIS ASI, high-latitude ion convections, or electric fields, have also been monitored since the 1990s (Greenwald et al., 1995). Super Dual Auroral Radar Network (SuperDARN), a widely used high-frequency radar network, provides Line-Of-Sight (LOS) Doppler ion velocity, which can infer large to small-scale plasma motions (Nishitani et al., 2019). Figure 1.3 shows SuperDARN ion drift measurements during the 2015 St. Patrick's Day storm (Chapter 5). Large ion velocities on the order of $1000 \, \mathrm{m \, s^{-1}}$ are present, indicating strong ion flows. The concurrent large ion velocity with enhanced aurora leads
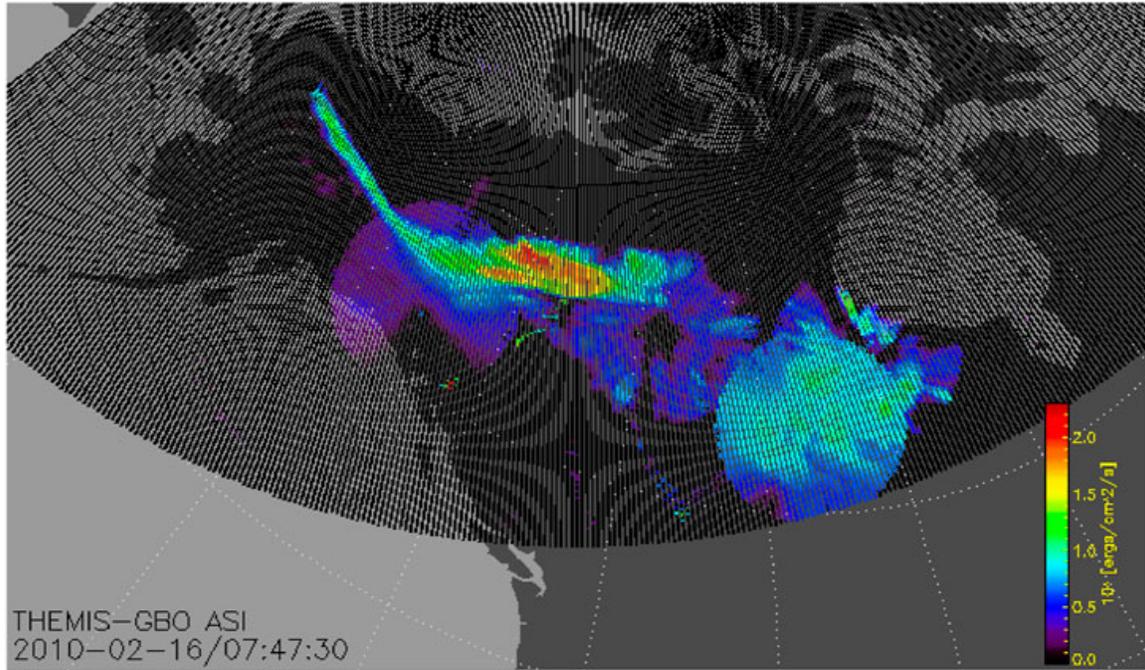
Figure 1.2: THEMIS ASI observations on 2010-02-16. Adapted from Gabrielse et al. (2021).

to a disturbed neutral atmosphere and increased variabilities (Chapter 5).

Incoherent Scatter Radar (ISR) also monitors the ionosphere with a range-resolved picture. ISR provides essential local electron density measurement and drift velocity derived from the scattered power spectrum. Figure 1.4 shows the Field-Of-View (FOV) of Poker Flat ISR (PFISR) with altitude dependence. PFISR continuously measures electron density and ion velocity since 2007, providing an essential data source for both statistical and case studies of the ionosphere.

I-T system reacts to magnetospheric energy deposition and tends to "balance" the input energy through dynamical and (photo-)chemical adjustments (Tian et al., 2008; Bharti et al., 2018). Li et al. (2019b) points out that NO cooling has a good correspondence with the magnetic storm. During the 2005-05-15 magnetic storm, NO cooling derived from Thermosphere Ionosphere Mesosphere Energetics and Dynam-
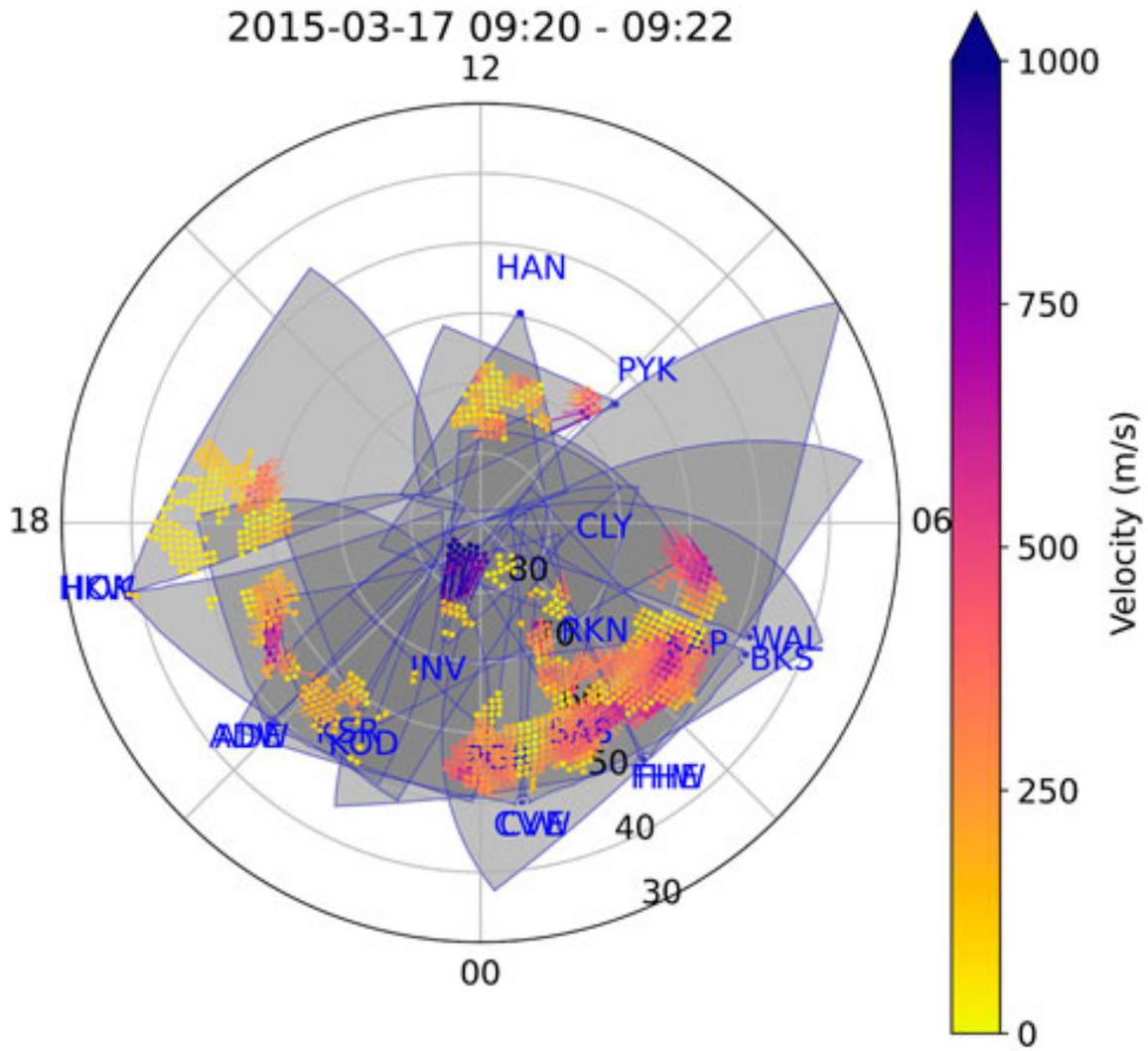
Figure 1.3: SuperDARN LOS ion drifts on 2015-03-17. Adapted from Shi et al. (2022).
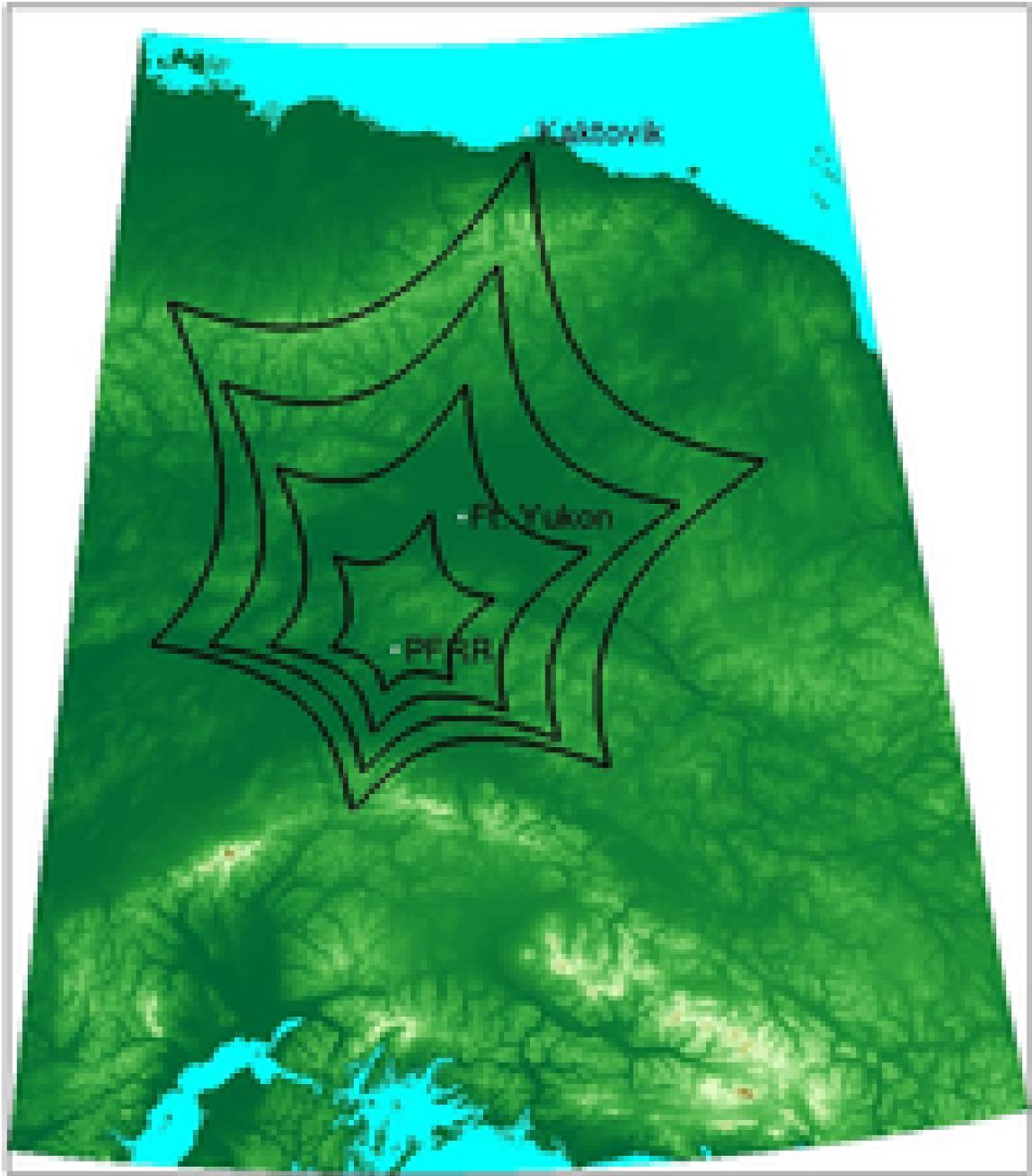
Figure 1.4: FOV of PFISR. The contours mark the radar's FOV at 100 km, 200 km, 300 km, and 400 km altitude. Courtesy Advanced Modular ISR website.

ics (TIMED)/Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) responds globally to the geomagnetic storm within 2 h after the onset of the main phase, with the most significant NO cooling increase occurring at the middle and low latitudes. This study indicates that the energy gained from the enhanced magnetic activity at high latitudes can be released via radiative cooling processes globally (Qian et al., 2010). Li et al. (2018) states that dynamic cooling (adiabatic cooling and horizontal/vertical advection) has the quickest responses and redistributes heat over the globe. The adiabatic processes are essential in redistributing the energy within the I-T system and changing the wind and circulation patterns.

On the other hand, GWs manifest their signals in various observations as spatial and temporal oscillations. Yue et al. (2022) shows satellite-based imaging observations from Geostationary Operational Environmental Satellites (GOES)/Advanced Baseline Imager (ABI) after the La Soufriere volcano eruption. The volcanic plume peaked at around 20 km, reaching the stratosphere. A close correlation between convective overshooting of the tropopause by 1 km to 3 km and concentric GWs in nightglow at 85 km has been established. Global Navigation Satellite System (GNSS) Total Electron Content (TEC) maps also demonstrate the ionospheric disturbances resulting from the volcano eruption. Figure 1.5 shows observations from ABI (left) and GNSS (right), where small-scale disturbances of the I-T system are discernable.

1.2 Numerical Models

Numerical simulation is performed to assist the understanding and explanation of observations. Numerical models are generally divided into physics-based models and statistical/empirical models. In addition, data assimilation models are developed to bridge the gap between models and observations. The principles of data assimila-
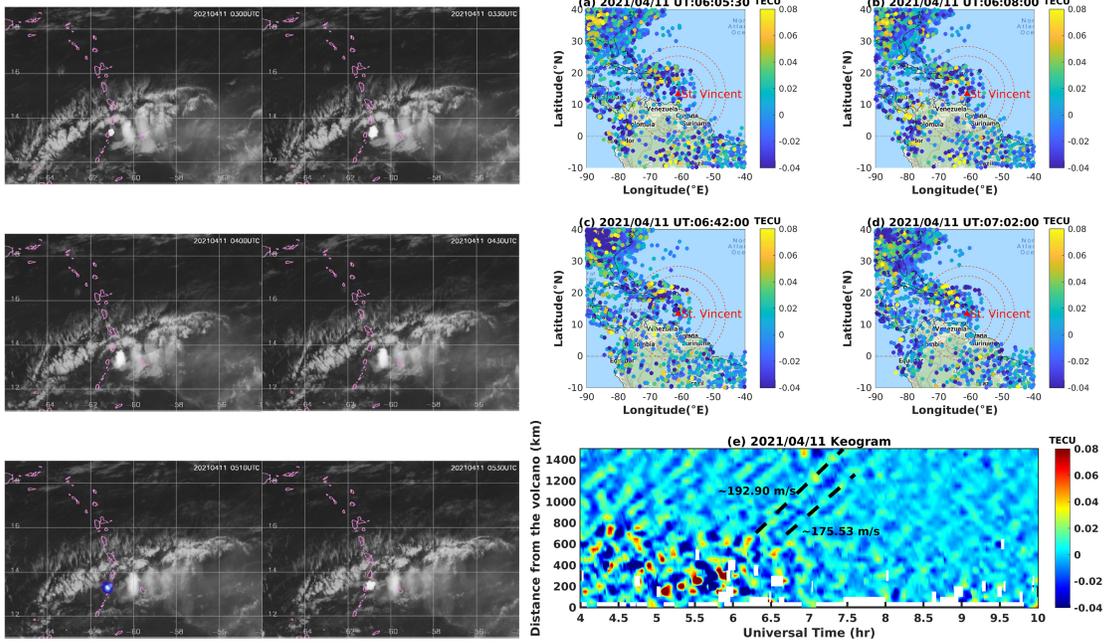
Figure 1.5: (Left) GOES-16 ABI infrared observations of the volcanic plume on 2021-04-11. (Right) Two-dimensional filtered GNSS TEC maps with the Butterworth band-pass filtering (10 min to 30 min) after the La Soufriere volcanic eruption. Adapted from Yue et al. (2022).

tion models are usually certain types of Gaussian statistics, and they are characterized as statistical models in this dissertation.

Physics-based models which self-consistently solve the I-T system can be further divided into two sub-groups: whole atmosphere models and I-T models. Whole atmosphere models refer to the models which solve the first-principle equations of neutral and charged particles in the whole altitude range from the ground to the thermosphere. The physical processes solved in whole atmosphere models usually include continuity, momentum, and energy equations of both ions and neutral particles. Chemical and photochemical reactions are also important parts of the whole atmosphere system. Some famous examples include Whole Atmosphere Community Climate Model with thermosphere and ionosphere extension (WACCM-X) (Liu et al., 2010), Whole Atmosphere Model with Ionosphere, Plasmasphere, Electrodynamics (WAM-IPE) (Akmaev et al., 2008), and Ground-to-topside model of Atmosphere and Ionosphere for Aeronomy (GAIA) (Jin et al., 2011). The most significant advantage of whole atmosphere models is that all the atmospheric waves are resolved self-consistently in the model, and the atmospheric wave propagation can be fully captured from the lower atmosphere to the upper atmosphere. However, due to the entire altitude range being solved, whole atmosphere models usually require large amounts of computing resources.

On the other hand, I-T models focus more on the I-T region ($100\,\mathrm{km}$ to $600\,\mathrm{km}$), and they are relatively computationally efficient compared to whole atmosphere models. I-T models also solve the motions of ions and charged particles and the energy transfer. They usually use a simplified set of chemical and photochemical reactions dominant in the I-T region. Performing control simulations in I-T models is relatively more straightforward than in whole atmosphere models since imposing variations at the lower boundary is straightforward. With the imposed variations, we

can then study the model evolution by performing such "forced" simulations. The shortage of I-T models is that for GW studies, manually placing GW seeding at the lower boundary is usually necessary since GWs cannot be generated spontaneously in the I-T region. The GW propagation in the I-T region is affected by the GW source. Therefore, the similarity to observations is determined mainly by the form of artificial GW seeding placed at the lower boundary. Some most widely used I-T models are Thermosphere Ionosphere Electrodynamics General Circulation Model (TIEGCM) (Richmond et al., 1992; Qian et al., 2014) and Global Ionosphere Thermosphere Model (GITM) (Ridley et al., 2006; Deng et al., 2008).

TIEGCM is a global 3-D numerical model that simulates the coupled I-T system from $97\,\mathrm{km}$ to $600\,\mathrm{km}$ using a hydrostatic assumption. It self-consistently solves the fully coupled, nonlinear, hydrodynamic, thermodynamic, and continuity equations of the neutral gas, the ion and electron energy and momentum equations, the ion continuity equation, and neutral wind dynamo (Richmond et al., 1992; Qian et al., 2014). It has been used to study various I-T phenomena, including geomagnetic storms (Lei et al., 2011), pre-reversal enhancements (Fesen et al., 2000), and annual ionospheric variations (Zeng et al., 2008). Developments of TIE-GCM since Richmond et al. (1992) include the extension of the upper boundary to higher altitude, variable tidal inputs (Hagan et al., 2001), improvements to the electro dynamo calculations (Richmond and Maute, 2014), revision of solar irradiance and photoelectron inputs (Solomon and Qian, 2005), modification of the upper boundary heat flux (Lei et al., 2007), application of variable eddy diffusivity at the lower boundary (Qian et al., 2009), revisions to auroral precipitation inputs (Emery et al., 2008) and background (nighttime) ionization rates, and inclusion of the Weimer potential model (Solomon et al., 2012). The current version of TIEGCM is 2.0, which has two standard resolutions: $5° × 5° × 1/2$ scale height and $2.5° × 2.5° × 1/4$ scale height.

Dang et al. (2018) incorporates a high-resolution configuration of TIEGCM up to $0.625° \times 0.625°$ horizontally and Cai et al. (2022) further extended the model upward to higher altitudes (over $1000\,\text{km}$) to study the magnetic storm impacts on the I-T system.

GITM is a non-hydrostatic model which solves the I-T system with adjustable grids in latitudes and altitudes. GITM explicitly solves for the neutral densities of O, O_2, N(2D), N($2\,$P), N($4\,$S), N_2, NO, H, and He; and ion species $O^+(4\,\text{S})$, $O^+(2\,\text{D})$, $O^+(2\,\text{P})$, $O_2{}^+$, N^+, $N_2{}^+$, NO^+, H^+, and He^+. A second-order flux-limited transport is used in the horizontal advection. The explicit advection solver enforces the numerical stability with a small model time step of $2\,\text{s}$. The neutral-neutral friction term, the heat conduction, and the viscosity terms are solved using an implicit scheme in the vertical direction. The chemistry in GITM is solved using a sub-cycling technique (Ridley et al., 2006). The non-hydrostatic effect is studied in Deng et al. (2008) by increasing Joule heating within a short period, and the results show that the vertical pressure gradient force can locally be $25\,\%$ larger than the gravity force, resulting in a significant disturbance away from hydrostatic equilibrium.

Physics-based models sometimes use empirical formulae where the detailed physical processes are not clearly understood or the actual processes are too computationally expansive to be implemented in global models. For example, the Weimer (2005) model is an empirical electric field model derived using two years of Dynamics Explorer (DE) 2 electric field measurements to provide statistical patterns of high-latitude electric fields. It estimates large-scale electric fields that successfully capture the variations of electric field magnitudes with magnetic activity levels. However, although the Weimer model gives reasonable estimates of the electric field in general, it tends to give overly smoothed electric field patterns. The missing variations of mesoscale to small-scale structures often lead to deviation from real-time observa-

tions.

To reduce the discrepancies between empirical models and actual observations, data assimilation techniques are developed like Assimilative Mapping of Ionospheric Electrodynamics (AMIE) (Richmond and Kamide, 1988; Richmond, 1992). AMIE is an optimally constrained, weighted least-squares fitting procedure of electric potential distribution to diverse types of atmospheric observations. The data fitting is carried out in two steps. In the first step, AMIE assimilates the height-integrated Pedersen and Hall conductances by modifying statistical auroral conductance models (Fuller-Rowell and Evans, 1987) with direct and indirect observations. In the second step, AMIE combines all available data related to electric fields (such as radar and satellite measurements of ion drifts) and currents (such as magnetic field perturbations observed on the ground and in space) to obtain an optimal linear estimation of electric potential over the high-latitude ionosphere (Lu, 2017). Future developments of AMIE are pursued by reformulating the Best Linear Unbiased Prediction (BLUP) problem presented in Richmond and Kamide (1988) into a Bayesian estimation problem (Matsuo et al., 2005) under the Gaussian distribution assumption of electrodynamic variables. A Bayesian perspective clarifies the critical role of the covariance function in modeling Gaussian processes, thus a better estimation of covariance functions from SuperDARN ion drift (Cousins et al., 2013), DMSP particle precipitation (McGranaghan et al., 2015), and Iridium magnetic perturbation (Cousins et al., 2015) lie in the center of modeling high-latitude electric fields (Matsuo, 2020). An end-user product of such AMIE extensions is Assimilative Mapping of Geospace Observations (AMGeO), which provides estimates of electric potential, Hall and Pedersen conductances, and FACs using concurrent SuperDARN, SuperMAG, and AMPERE measurements.

However, both empirical and physics-based models are far from reproducing

all observations. For example, Stasiewicz and Potemra (1998) found that FAC varies from several µA to hundred µA when horizontal sizes vary from $500\,\mathrm{km}$ down to sub km scale. Such large deviations at small scales from large scales are never reproduced in model studies. Localized strong FACs might also lead to neutral dynamical responses like neutral density variations (Luhr et al., 2004), further impacting the mid-to-low latitude atmosphere through Traveling Atmospheric Disturbances (TADs).

The reasons for the discrepancies come from two aspects. One is because empirical models could only give large-scale climatological structures, which tend to underestimate the actual observations for specific cases. For example, the Weimer model predicts a Joule heating rate of $15\,\mathrm{mW\,m^{-2}}$ during storm time, which is below the typical observations of $100\,\mathrm{mW\,m^{-2}}$ (Huang and Burke, 2004), and an extreme case has been detected to exceed $170\,\mathrm{mW\,m^{-2}}$ (Knipp et al., 2011). The essential point for resolving such underestimations is incorporating observations into the model as the high-latitude drivers.

The other aspect is due to the scale dependency of the I-T process. Since the atmosphere is a highly nonlinear system, small-scale structures are not simply superimposed onto large-scale structures. The interaction between small- and large-scale processes leads to energy transfer between small and large scales (Charney, 1971). Joule heating rate is larger when small-scale features are considered (Lu et al., 1998a). While mean electric fields are usually used for calculating Joule heating in models, electric field variabilities also have comparable contributions to the Joule heating as mean electric fields (Codrescu et al., 2000). Missing electric field variabilities will lead to underestimations of Joule heating and energy budget (Deng et al., 2009). Moreover, the intrinsic multi-scale nature of GWs further complicates the cross-coupling picture by introducing ways of wave-wave and wave-mean interactions (Liu et al., 2014). Therefore, increasing the model resolution to simulate space weather more

realistically has equal importance.

1.3   From Large to Small Scales

Satellite and ground-based observations demonstrate multi-scale structures in high-latitude drivers (i.e., auroral precipitation, electric field, and FAC). Luhr et al. (2015) studied the scale dependency of FACs and found that FACs on the order of 1.5° are stable for 1 min while those with sizes below 0.1° have a persistent time of no more than 10 s. A similar conclusion for auroral particle precipitation is also obtained by Humberset et al. (2017). Rocket measurements suggest that the effects of electric field variabilities of small-scale contributions to the heating can be more than a factor of 2 greater than the mean field contribution (Hurd and Larsen, 2016).

Figure 1.6 shows the multi-scale magnetospheric energy deposition processes from large scales over 1000 km down to kinetic scales of 1 m (Nishimura et al., 2021a). In the leftmost figure, statistical models give smooth large-scale (1000 km) averages of auroral precipitations. In real-time, the THEMIS ASI network depicts rich mesoscale (10 km to 100 km) precipitation structures during a substorm. Narrow FOV images show even smaller scale (10 km) structures and conductance gradients. In the rightmost figure, kinetic simulations predict turbulent structures of the ionospheric densities at a sub-meter scale, even though it is not observed so far. The aurora and corresponding change of ionospheric conductances and densities manifest in scale-dependent spatial structures, with each view demonstrating a completely different physical picture. The mesoscale, small-scale, and kinetic-scale structures can have comparable or larger magnitudes than those at larger scales, and thus those potentially affect larger-scale dynamics. From the perspective of modeling the multi-scale I-T processes, while the large-scale enhancement in the magnitude of aurora can be

captured in most models, the local structures of auroral activities are generally not reproduced.

In the other direction, atmospheric waves continuously transport energy and momentum upward into the I-T region, leading to disturbances of various scales. Figure 1.7 shows the vertical wind from a Whole Atmosphere Community Climate Model (WACCM) simulation indicative of GW activities at four representative heights: tropopause, stratosphere, mesosphere, and lower thermosphere (Liu et al., 2014). GWs resulting from a tropical cyclone appear as concentric ring patterns at low latitudes. With increasing altitudes, the horizontal extent and amplitude of the GWs grow, leading to more apparent wave structures at higher altitudes. Apart from the perturbations associated with the tropical cyclone, many other concentric GWs arise, likely due to deep convection. The growth of GWs from small to large scales is one of the critical cross-scale coupling processes. Such cross-scale processes are essential sources of variabilities in the I-T region.

Techniques assisting high-resolution model simulations are developed for atmospheric modeling studies. Wang et al. (1999) developed Thermosphere Ionosphere Nested Grid (TING) model, a nested-grid extension to Thermosphere Ionosphere General Circulation Model (TIGCM), to study the mesoscale and cross-scale I-T processes. GITM with local refinement (GITM-R) adopted a similar approach, which solves regional high-resolution dynamic evolutions in GITM (Zhao et al., 2020). Recently, a new dynamical core, Spectral Element (SE), was implemented in WACCM-X (Liu et al., 2022a), which employs a quasi-uniform cubed sphere grid (Lauritzen et al., 2018). Unlike orthogonal structured meshes used in other models with decreasing areas near the polar region, which places a singularity issue including TIEGCM and GITM, the cubed sphere grid addresses the polar singularity issue. The SE dynamical core further enables simulations at much higher horizontal resolutions. Other

16
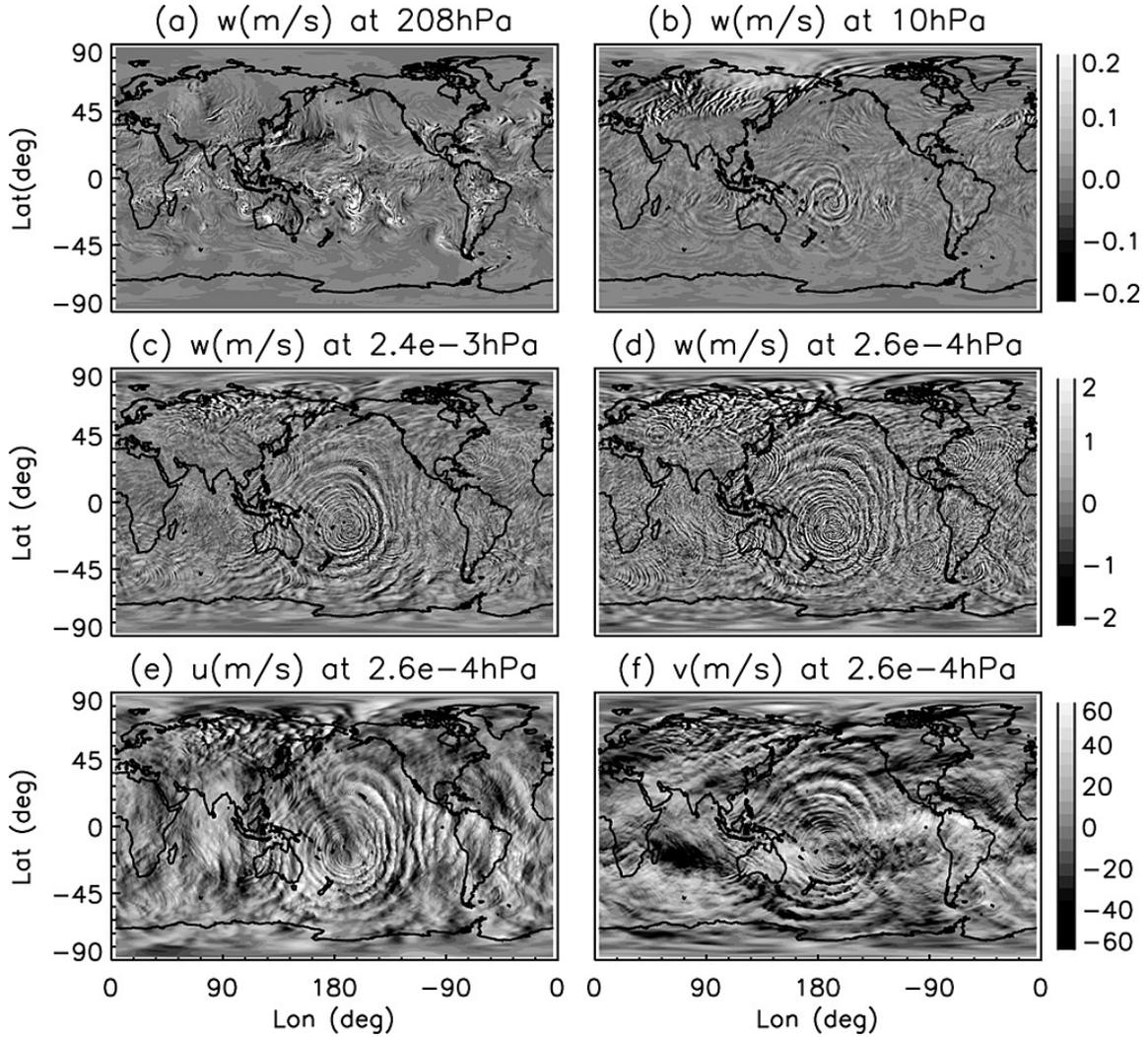
Figure 1.6: Illustration of multiscale processes. From left to right: (1000 km) large-scale statistical precipitating electron energy flux distribution from Oval Variation, Assessment, Tracking, Intensity, and Online Nowcasting (OVATION)-SuperMAG; (10 km to 100 km) instantaneous precipitating electron energy flux distribution in the nightside using THEMIS ASIs; (10 km) instantaneous small-scale Pedersen conductance distribution derived from narrow FOV imaging; (1 m) kinetic simulation of ionosphere density. Adapted from Nishimura et al. (2021a).
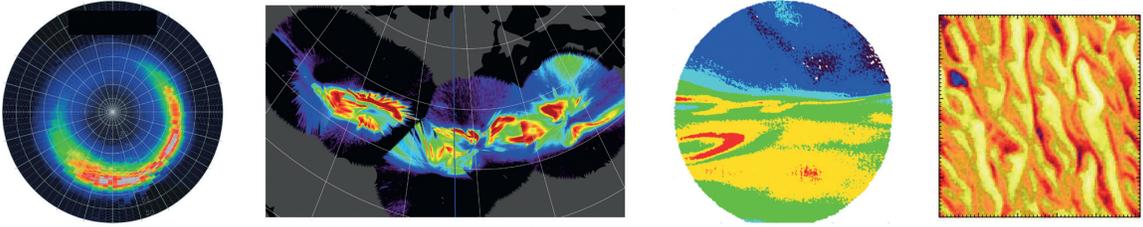
Figure 1.7: WACCM simulation results on February 4 at 21:00:00. (a–d) Vertical winds at $208\,\mathrm{hPa}$ ($11\,\mathrm{km}$), $10\,\mathrm{hPa}$ ($30\,\mathrm{km}$), $2.4\times10^{-3}\,\mathrm{hPa}$ ($87\,\mathrm{km}$), and $2.6\times10^{-4}\,\mathrm{hPa}$ ($100\,\mathrm{km}$), respectively. (e-f) Zonal and meridional winds at $2.6\times10^{-4}\,\mathrm{hPa}$. Adapted from Liu et al. (2014).

whole atmosphere models like WAM-IPE and GAIA use a spectral dynamical core to solve the atmosphere in different wavelengths (Akmaev et al., 2008; Jin et al., 2011). Instead of solving equations in a geographic grid, the spectral method reformulates the problem in the frequency domain and calculates the time evolutions of each wavelength.

1.4 Science Questions

Based on the discussions above, we raise the following science questions in terms of the downward and upward coupling processes:

1. How does the I-T system respond to multi-scale magnetospheric energy deposition?

2. How does the I-T system respond to lower-atmosphere forcing?

1.5 Objectives

We try to answer the above science questions using numerical models. Specifically, we perform the following efforts to untackle the science questions:

1. Develop a multi-resolution data assimilation model to quantify magnetospheric energy deposition. Then use this new data assimilation method to assist I-T simulations using TIEGCM.

2. Develop nested-grid TIEGCM to study GW propagations in the I-T system.

Regarding data assimilation methods that digest real-time observations with historical information, several existing models provide dynamic evolutions of high-latitude electrodynamic processes like AMIE. They usually provide realistic variations due to the data fusion procedure that uses actual observations to constrain the model. However, the fitting resolution is still a big restriction to the model in the smallest resolvable scale. AMIE has a resolution of 10° in longitude and 2° in latitude, which can resolve large-scale auroral structures and ion flows. However, the Magnetosphere-Ionosphere-Thermosphere (M-I-T) coupled system embraces a variety of critical medium-to-small-scale electrodynamic processes below the resolution resolvable by AMIE. For example, during an expansion phase of a substorm, the auroral structure with scales more minor than $500\,\mathrm{km}$ contributes to $50\,\%$ of the total energy flux. Mesoscale auroral processes such as poleward moving auroral forms, polar cap patches, auroral arcs, and streamers can give feedback to the large-scale dynamics and impose net effects on the global distribution of electron densities (Gabrielse et al., 2021). The electric fields at the polar cap and auroral region exhibit cross-scale power spectra from planetary scales down to $0.5\,\mathrm{km}$ (Golovchanskaya and Kozelov, 2010a; Kozelov and Golovchanskaya, 2006), which lead to the deviation from the global large-scale two-cell ion convection pattern (Cousins and Shepherd, 2012a,b). Small-scale electric fields have often been observed and found to impact the energy budget during magnetic storms (Codrescu et al., 1995; Cosgrove and Codrescu, 2009). All these processes will change the evolution of the I-T system and lead to enhanced

spatial and temporal variabilities.

A straightforward idea to extend AMIE to high resolutions is by increasing the number of basis functions used in AMIE. But the overfitting problem takes over the benefit of high resolutions at some point. From the point of inverse problems, AMIE retrieves global electric fields out of scattered electrodynamic measurements through matrix calculations (Matsuo, 2020). Due to the relatively complex covariance function used in AMIE to describe the spatial correlation, the internal mathematical structure of AMIE has a very high dimension, and the size of the inverse problem grows as the square of the number of basis functions. To increase the resolution of AMIE to $2°$ in longitude by $1°$ in latitude, the number of basis functions grows by 10 times, and the size of the problem grows by 100 times. The fitting procedure will likely fail when directly applied to a problem 100 times larger. The mathematical reason for this vulnerability is that the algorithm for inverting a matrix of relatively small sizes is different from the algorithm for inverting a huge matrix. Therefore, new algorithms must be introduced to extend data assimilation methods to high resolutions.

The current physics-based models are also insufficient in studying mesoscale to small-scale phenomena. Nyquist theorem states that the minimum resolved scale of the model is twice the grid size. Thus, for standard TIEGCM configuration of a global uniform $2.5°$ grid, the resolvable atmospheric waves are at least $5° \approx 500\,\mathrm{km}$. Wang et al. (1999) pointed out that a fine grid could provide more small-scale structures of the auroral precipitation and resolve more electro- and neutral dynamic structures than a coarse grid. High resolution is even more critical for atmospheric GW studies since small-scale waves will be damped in coarse grids (Fritts and Alexander, 2003). Extending current models to high resolutions is a critical requirement for GW-related studies.

But there are challenges regarding the costs and outcomes of extending physics-

based models to high resolutions. The computation grows as the number of grid points increases to perform global model simulations. For example, a global TIEGCM simulation with uniform 2.5° horizontal grid and 57 vertical grids takes about 5 CPU hours to finish 1 model day at a time step of 30 s. For a high-resolution simulation at a uniform 0.625° grid and double the number of vertical grids, the computation increases to 160 CPU hours. What further increases the computation is the requirement for numerical stability that the Courant number must be smaller than 1 (Durran, 2010). The time step of high-resolution simulations must decrease to seconds and sub-seconds to stay stable. For a typical TIEGCM simulation with a horizontal resolution of 0.625°, the time step has a maximum of 5 s. This leads to the total time cost going up to nearly 1000 CPU hours. The drastic increase in computational cost prohibits the general application of global high-resolution simulations.

However, considering practical uses, global high-resolution simulations provide much more information that can be used directly. For example, observations usually prefer spatial distributions: high-resolution observations often occur in a restricted region, while low-resolution observations exist globally. This uneven spatial distribution suggests we design a model simulation domain according to observations. It would save a lot of computation if we could increase the resolution in certain regions while keeping the global simulation in low resolutions. Such a hybrid-grid approach to physics-based modeling is achieved in other models like Weather Research and Forecasting model (WRF); the remaining problem is adapting this approach to I-T models.

1.6 Outline

This dissertation starts with a mechanism study of a significant temperature enhancement at the lower thermosphere observed by an Fe Boltzmann lidar at Mc-Murdo, Antarctica (Chu et al., 2011). The repeated observations on 2011-05-02 and 2011-05-28 indicate standard underlying physical processes, which turned out to be associated with localized auroral precipitation and electric field activities. Local aurora and electric fields lead to significant Joule heating deposited at the E region altitude. The thermodynamic term analysis indicates that the combined effect of Joule heating, adiabatic cooling, and vertical advective heating leads to the vertical temperature structure as observed (Chapter 2).

However, our method in the data processing step on aurora and electric fields (Section 2.2) is imperfect. Neglecting spatial correlation causes strong discontinuity and artificial perturbations to the model (Section 2.6). This shortage leads us to consider further a "better" tool to analyze the multi-scale processes coherently. In Chapters 3-4, we introduce a new data assimilation method based on a Gaussian process model to synthesize auroral data and ion drift measurements from different sources. This method incorporates the spatial covariance structure and the measurement error to estimate the actual field and provide predictions and quantified uncertainties, eliminating the spatial inconsistency in Chapter 2. With high-resolution modeling, the large-scale and mesoscale structures remain in the assimilation result. As a by-product, scale analysis of aurora and electric fields can also be achieved using this method. Chapter 5 talks about the local and global I-T impacts of the 2015 St. Patrick's Day storm (2015-03-17) using TIEGCM and the newly developed data assimilation method. In this study, the data assimilation method also significantly impacts spatial and temporal variabilities in the I-T region.

A nested-grid extension of TIEGCM is developed to extend the capability of physics-based models to study atmospheric processes from large to medium and small scales. Specifically, the GW activities in the I-T region during the 2022-01-15 Tonga volcano eruption are studied in detail, including their impacts on the ionosphere (Chapter 6). The analysis indicates that the resolution is one of the most critical factors affecting GW propagations in the I-T system.

Based on the science questions and methodologies discussed above, we summarize the main points of this dissertation:

1. Study the mechanism of the temperature enhancement at the lower thermosphere and identify the local auroral precipitation and sub-grid scale electric field as the dominant factor (Chapter 2).

2. In retrospect, the shortage of Chapter 2 is discussed in Section 2.6, which leads to the motivation of developing a new data assimilation method to study aurora and electric field (Chapters 3-4) and apply it to a geomagnetic storm study (Chapter 5). The local and global impact resulting from the enhanced aurora and electric fields are examined quantitatively.

3. Develop a nested-grid extension of TIEGCM to study GW propagation in the I-T system and identify one of the essential factors affecting GW propagations as the model resolution (Chapter 6).

4. Chapter 7 summarizes the conclusions and briefly describes possible future scientific projects.

Chapter 2

Studying the Importance of Mesoscale Aurora and Electrical Fields during Geomagnetic Storms

A dramatic Thermospheric Temperature Enhancement and Inversion Layer (TTEIL) was observed by the Fe Boltzmann lidar at McMurdo, Antarctica, during a geomagnetic storm (Chu et al., 2011). TIEGCM driven by empirical auroral precipitation and background electric fields cannot adequately reproduce TTEIL. We incorporate DMSP/SSUSI auroral precipitation maps, which capture the regional-scale features into TIEGCM and add subgrid electric field variability in the regions with intense auroral activity. These modifications enable the simulation of neutral temperatures closer to lidar observations and neutral densities closer to Gravity Recovery and Climate Experiment (GRACE) satellite observations (475 km). The regional scale auroral precipitation and electric field variabilities are both needed to generate intense Joule heating that peaks around 120 km. The resulting temperature increase leads to the change of pressure gradients, thus inducing a horizontal divergence of

airflow and large upward winds that increase with altitude. The adiabatic cooling is associated with the upwelling wind, gradually increasing with altitude and peaking at 200 km. The intense Joule heating around 120 km and strong cooling above result in differential heating that produces a sharp TTEIL. However, vertical heat advection broadens TTEIL and raises the temperature peak from 120 km to 150 km, causing simulations to deviate from observations. Intense local Joule heating also excites TADs that carry TTEIL signatures to other regions. Our study suggests the importance of including fine-structure auroral precipitation and subgrid electric field variability in modeling storm-time I-T responses.

2.1   Introduction

Magnetospheric energy is dissipated in the I-T region through Joule heating and particle precipitation (Cole, 1975). Most of the energy from the convergence of Poynting flux results in Joule heating (Knipp et al., 2004; Richmond and Thayer, 2000; Thayer et al., 1995; Thayer and Semeter, 2004), while auroral particle precipitation heats the atmosphere directly through collision and indirectly by increasing conductivity and thus Joule heating. The frictional collisions between ions and neutrals cause temperature enhancement and subsequent pressure divergence, leading to upward neutral motions. Furthermore, Lorentz force accelerates neutrals by transferring momentum from fast-moving ions (Mikkelsen et al., 1981). During periods of intense geomagnetic activity, both auroral precipitation and electric fields are elevated, leading to enhanced energy deposition in the I-T region (Deng et al., 2011; Fuller-Rowell et al., 1987; Heppner et al., 1993; Wang et al., 2005). For instance, the localized extreme Poynting flux has been detected to exceed $170\,\mathrm{mW\,m^{-2}}$ by DMSP spacecraft with substantial temporal and spatial variability in the auroral zone (Knipp

et al., 2011). Although the Weimer statistical model predicts a peak Joule heating rate of $15\,\mathrm{mW\,m^{-2}}$ (Weimer, 2005), satellite observations have revealed that the localized Earth-directed Poynting fluxes can be several times larger during magnetic storms (Huang et al., 2016; Huang and Burke, 2004).

However, such localized energy deposition and its effects are difficult to capture in the I-T models driven by empirical high-latitude inputs because those drivers are usually obtained from statistical auroral maps (Hardy et al., 1985; Newell et al., 2009) and electric convection patterns (Heelis et al., 1982; Weimer, 2005). The statistical auroral maps usually miss localized features in the resolved grids, and convection patterns usually miss subgrid temporal and spatial variabilities (Cosgrove et al., 2011; Matsuo and Richmond, 2008). To reduce the gap between observations and statistical models, data assimilation methods such as AMIE (Richmond, 1992; Richmond and Kamide, 1988) have been developed to synthesize available observations into coherent patterns to provide more realistic electric fields. However, even with data assimilation techniques, comparisons of AMIE outputs with the direct observations still show significant discrepancies (Cosgrove et al., 2009; Cosgrove and Codrescu, 2009).

It has been shown that electric field variability has an equal contribution to Joule heating as the mean electric field (Codrescu et al., 2000), and therefore, omitting electric field variability causes a significant underestimation of Joule heating in the models (Codrescu et al., 1995, 2008; Deng et al., 2009; Matsuo and Richmond, 2008). Since Joule heating is the dominant heating source in the polar region during storm times, the underestimation of Joule heating leads to an underestimated global heating budget, further influencing the calculation of global circulation. Emery et al. (1999) introduced a parameterized way of accounting for the missing contribution from subgrid electric field variabilities in TIEGCM by multiplying Joule heating by a factor of 1.5 globally, which effectively increases the global Joule heating budget to

a reasonable level but neglects the spatial distribution of electric field variabilities. Satellite and ground-based observations have found that large electric field variabilities occur mainly in the auroral zone (Cousins and Shepherd, 2012a,b; Matsuo et al., 2003). By analyzing plasma drift measurements obtained from the DE 2 mission, Matsuo et al. (2003) found that the electric field variability exceeds the magnitude of the mean electric field in the polar area. Cousins and Shepherd (2012a) statistically studied small-scale electric field variabilities on spatial scales between 45 km and 450 km and temporal scales between 2 min and 20 min using 4 years of SuperDARN LOS ion drift measurements. They found that the small-scale variability shows spatial distribution correlated with the gradient in the background plasma drift, likely originating from small-scale magnetospheric turbulences. Note that the mean spatial scale they studied (225 km) is comparable to the typical resolution of I-T models. Cousins and Shepherd (2012b) further examined the distribution of electric field temporal and spatial variabilities, which peak in the auroral region under negative IMF B_z and in winter times.

The motivation of this work originates from the large neutral thermospheric temperature elevation in the E region altitude detected by an Fe Boltzmann lidar (Chu et al., 2002) at McMurdo (77.8°S, 166.7°E), Antarctica around 15:00:00 during the 2011-05-28 storm event (Figure 2c in Chu et al. (2011), replotted as Figure 2.1a in this paper). The storm resulted from the passage of two Interplanetary Coronal Mass Ejections (ICMEs), both with clear magnetic cloud signatures (Chi et al., 2018), and was accompanied by strong auroral activity as shown in the Auroral Electrojet (AE) index. The IMF B_z southward fields from ICME joined in an interaction region that passed Earth from 13:00:00 to 14:30:00. The IMF B_z was southward from 06:00:00 to 15:00:00 during which the Kp index was 6 and Symmetric Disturbance Field in H component (SYM-H) index dropped to about -100 nT (Figures 2.1b-e). Compared

to NRLMSISE-00 model (Picone et al., 2002), the storm-time neutral temperature profile was dramatically elevated with a maximum increase of 550 K at 134 km (Chu et al., 2011). This leads to a local temperature peak in altitude and the formation of a neutral TTEIL at E region altitudes, defined as the structure showing temperature increase and then decrease with altitude.

Most storm studies have been dedicated to investigating the responses of ionospheric parameters such as ion drifts and electron densities (Lu et al., 2008, 1998b) or neutral density and wind variations in the upper thermosphere (Forbes et al., 2005; Lei et al., 2010; Sutton et al., 2005). A few studies have reported the storm effects on neutral temperatures in the mesosphere and lower thermosphere, which cause perturbation of tens of K (Fagundes et al., 1996; Lastovicka, 1996; Liu et al., 2018c; Li et al., 2018, 2019a; Yuan et al., 2015). Intriguingly, such tremendous neutral temperature elevation is present at E region altitudes (110 km to 150 km) with an associated inversion layer. Furthermore, another event with a similar temperature enhancement and inversion layer, but somewhat a smaller magnitude, was observed by the same Fe Boltzmann lidar at McMurdo on 2011-05-02 (Chu et al., 2011). Therefore, such thermospheric enhancement and inversion at E region altitudes may not be rare in the neutral atmospheric responses to storms. On the other hand, even though the I-T models with statistical high-latitude drivers provide fair predictions of the medium- and large-scale storm-time dynamics in the upper atmosphere, none of them can simulate TTEIL seen in the local lidar observations, which challenges our understanding and modeling of the M-I-T coupling processes. To investigate these features, we implement the auroral precipitation maps from SSUSI observations onboard DMSP satellites and incorporate the electric field variability into TIEGCM (Section 2.2). The model outputs are compared with neutral temperatures observed by the GRACE satellite's lidar and neutral density measurements. By pushing the

simulation closer to observations, we aim to investigate the generation mechanism of TTEIL in the model, the sources of local heating and cooling, and the neutral dynamics effects (Section 2.3). Additional modeling efforts to simulate TTEIL and the limitations of this work are discussed in Section 2.4.
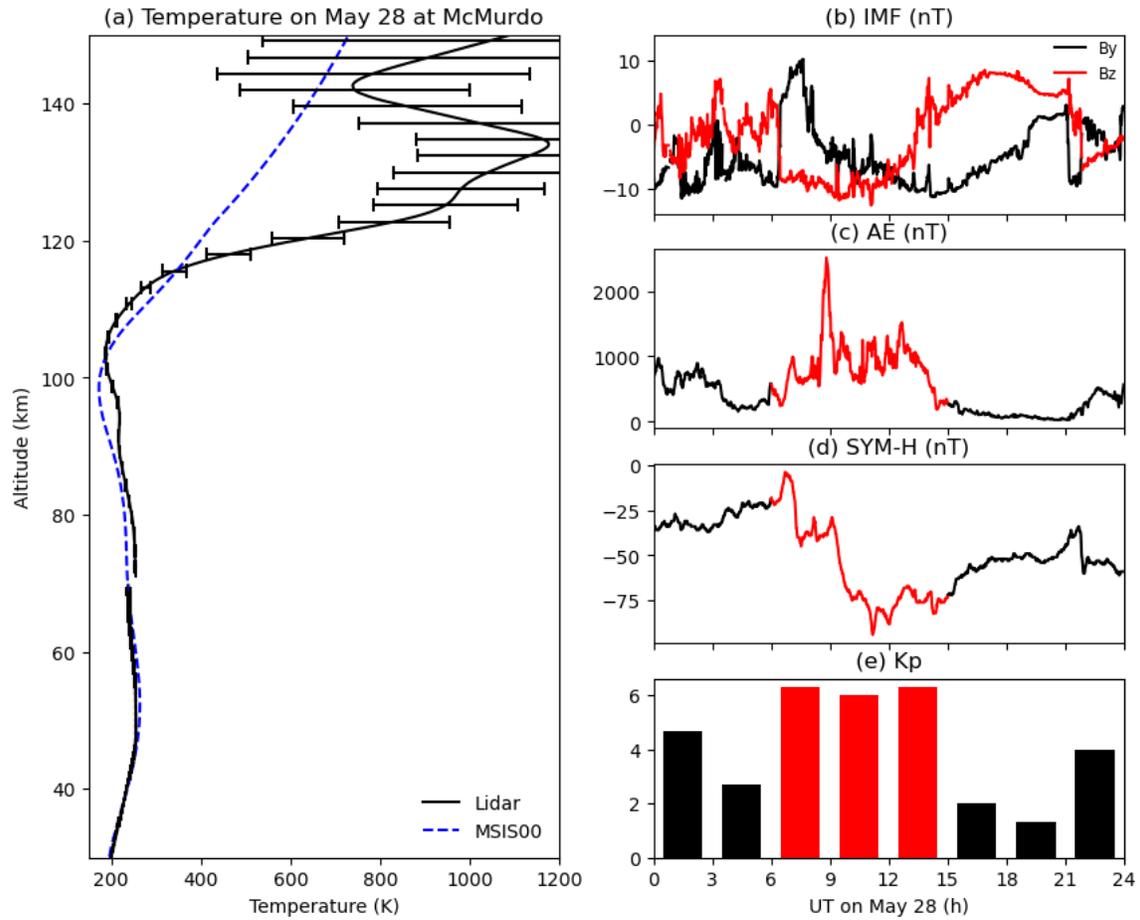
Figure 2.1: (a) The vertical profile of temperatures for 1 h integration around 15:00:00 measured by the lidar (replotted from Figure 2c in Chu et al. (2011)). (b-e) Geomagnetic indices showing IMF, AE, SYM-H, and Kp, respectively, on 2011-05-28.

2.2 Model, Data, and Methodology

2.2.1 Model Setup

This study uses TIEGCM version 2.0, which has a horizontal resolution of 2.5° in latitude and longitude and a vertical resolution of 1/4 scale height. TIEGCM is a global 3-D numerical model that simulates the coupled thermosphere/ionosphere system from 97 km to 600 km altitude. It self-consistently solves the fully coupled, nonlinear, hydrodynamic, thermodynamic, and continuity equations of the neutral gas, the ion and electron energy equations, the O^+ continuity equation and ion chemistry, and the neutral wind dynamo (Qian et al., 2014; Richmond, 1995; Richmond et al., 1992; Roble et al., 1987, 1988). The lower boundary tides are derived from a linear atmospheric model assimilated with TIMED SABER and TIMED Doppler Interferometer (TIDI) observations (Wu et al., 2012; Yamazaki et al., 2014). The time step of TIEGCM simulation is 30 s. Diagnostic outputs are saved every minute.

One of the essential high-latitude drivers for the I-T model is the electric field (or electric potential) mapped down from the magnetosphere. AMIE is designed to assimilate observational data from various sources to provide a more realistic high-latitude electric field convection pattern (Knipp et al., 1993; Lu et al., 1995; Ridley et al., 1998), which is used in this study to drive TIEGCM. The data assimilated in AMIE for this event include magnetic field perturbations from 265 ground-based magnetometers (197 in Northern Hemisphere, 68 in Southern Hemisphere) through SuperMAG network (Gjerloev, 2009, 2012) and from Iridium satellite magnetometers through AMPERE project (Anderson et al., 2008, 2014). LOS ion drifts from SuperDARN (Chisham et al., 2007; Greenwald et al., 1995) and auroral particle precipitation measured by Special Sensor J (SSJ) instrument onboard DMSP F16, F17, and F18 satellites (Kadinsky-Cade et al., 2004) are also used.

2.2.2 Incorporating DMSP/SSUSI Observations Into TIEGCM Auroral Precipitation Maps

Another high-latitude driver for TIEGCM is auroral precipitation, obtained through AMIE fitting procedure. Assuming Maxwellian distributions, auroral precipitation can be described by two key parameters: mean energy (E_m) and energy flux (Φ_E) (Rees and Luckey, 1974; Robinson and Vondrak, 1985; Roble and Ridley, 1987). Φ_E and E_m in TIEGCM/AMIE describe the enhancement of aurora during storm times, but underestimation of auroral activity still exists. Figure 2.2 shows an example of the comparisons of Φ_E and E_m between TIEGCM/AMIE and DMSP/SSUSI observations around 11:55:00. Comparing Figures 2.2a1-a2 with 2.2b1-b2, the position of the auroral oval is reasonably captured, but its magnitude is significantly underestimated, and regional structures are not resolved.

To mitigate the underestimation of auroral precipitation in TIEGCM for this study, we modify the auroral maps in TIEGCM according to SSUSI observations. We use auroral EDR data (Paxton et al., 1992), which contain Φ_E and E_m derived from the observed spectral intensities. For the 2011-05-28 event, three DMSP satellites (F16, F17, and F18) were taking measurements, and therefore, they are all used. The raw Φ_E and E_m data of each swath are provided in 2-D Magnetic Latitude (MLAT)/Magnetic Local Time (MLT) grids with excellent resolution (0.1°MLAT $\times$ 0.01 hMLT). The number of data samplings from all three satellites over each MLAT/MLT grid for the period of 2011-05-27 to 2011-05-29 is shown in Figure 2.3a. This polar map indicates the coverage of satellite observations. The satellites passed most of the region on the night side for more than 70 times while missing the dayside sector to a large extent. To take advantage of the available observations and fill up the observational gaps, we combine observations and the model as follows:
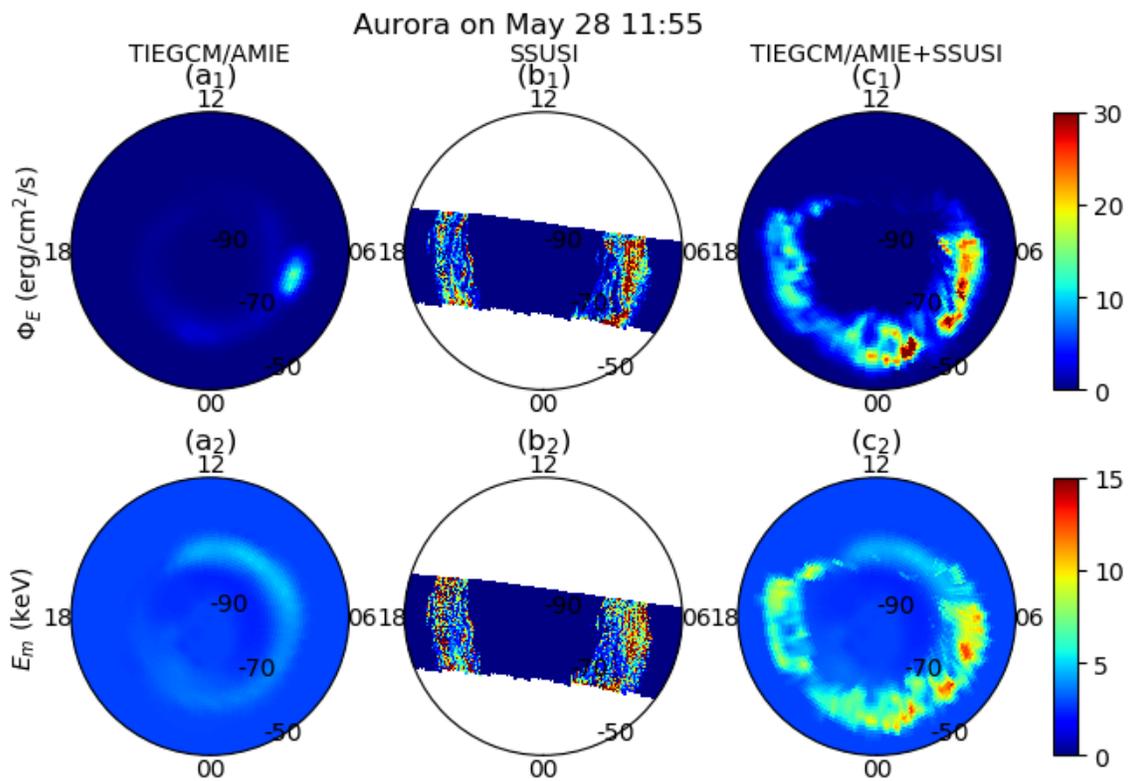
Figure 2.2: (a1-c1) auroral maps of Φ_E from TIEGCM/AMIE (a1), SSUSI (b1), and a combined map of TIEGCM/AMIE and SSUSI (c1). The plotting time is 11:55:00 for (a1) and (c1) and 11:42:00 to 12:05:00 for (b1). (a2-c2) are the same except for E_m. All plots are in MLAT/MLT coordinates.

1. We collect SSUSI observations from all three satellites falling into each MLAT/MLT grid to form a time series in UT, which is unevenly distributed. The red dots in Figure 2.3b show Φ_E measured by SSUSI on 2011-05-28 at 70°S MLAT and 3 MLT. We apply linear temporal interpolation in UT to project the data into an evenly distributed time series with a time interval of 30 s to be used by the model (black line in Figure 2.3b). Figure 2.3c shows the original Φ_E observations from one satellite pass between 11:21:00 and 11:43:00, and Figure 2.3d shows the temporally interpolated Φ_E at 11:30:00, which is obtained through interpolation from multiple satellite passes. Most polar areas are covered except for the dayside sector, where the satellites have no passes. When temporal interpolation happens at a time (e.g., 11:30:00 as in Figure 2.3d) close to the time when the observation is taken (e.g., 11:21:00 to 11:43:00 shown in Figure 2.3c), the interpolated aurora (Figure 2.3d) tends to approach the observation (Figure 2.3c) due to the proximity in time.

2. After linear interpolation, the auroral maps are converted from MLAT/MLT coordinates to Geographic Latitude (GLAT)/Geographic Longitude (GLON) coordinates (Figure 2.3e), which are nonuniform spatially after the conversion and have a finer spatial resolution than TIEGCM. To obtain auroral maps on the regular TIEGCM grids, spatial binning is performed on the geographic coordinate by taking the average of the 20 to 30 points falling into each model grid (Figure 2.3f), which includes both the real SSUSI observations and the interpolated results. The Standard Deviations (SDs) of these points are also calculated and retained for mechanism studies (Section 2.3.1). This binning operation is performed at every interpolating time step (30 s).

3. We replace TIEGCM/AMIE auroral maps with the temporally interpolated

and spatially binned SSUSI maps obtained after Step 2 wherever available. For the regions where satellites do not pass (such as the dayside sector in the polar map), we use the auroral maps from the model (TIEGCM/AMIE), which usually occupies 20 % of the overall auroral region. Using the model maps is a reasonable approximation considering that auroral activity in the dayside sector is generally weak (Newell et al., 2009), thus, the differences between observations and the model are relatively small.
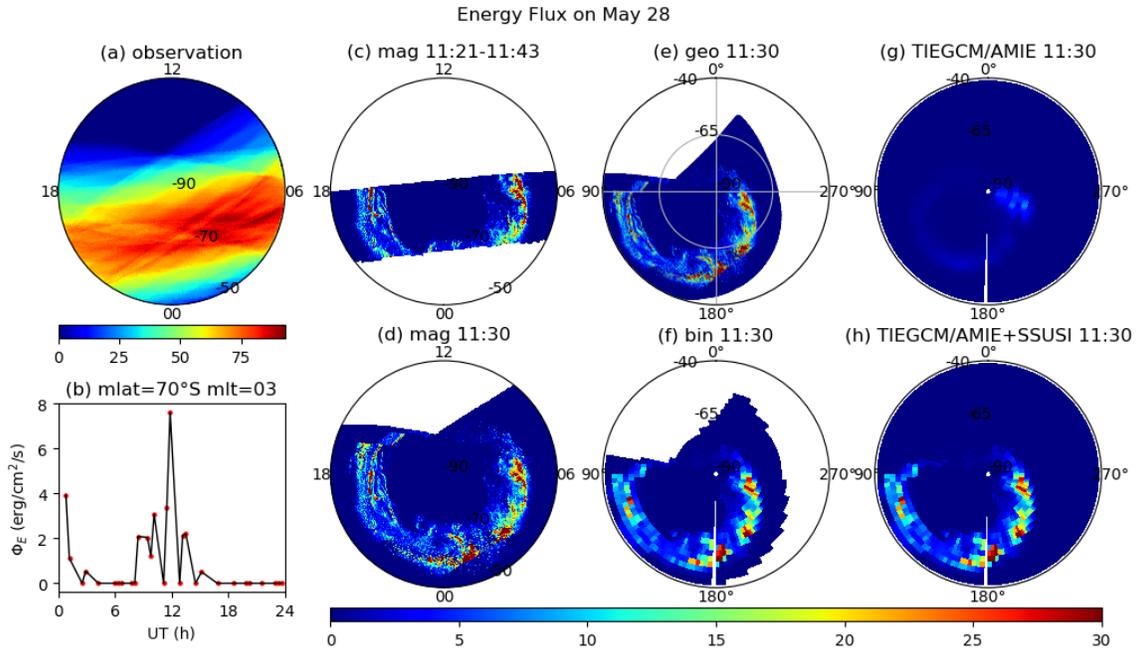
Figure 2.3: (a) Total sampling numbers for the satellite auroral observations measured by SSUSI from DMSP F16, F17, and F18 from 2011-05-27 to 2011-05-29, in MLAT/MLT coordinate. (b) Linear interpolation of energy flux (Φ_E) at 70°S MLAT and 3 MLT. Red dots are observations, and the black line corresponds to the interpolated results with a time interval of 30 s. (c) Raw Φ_E from 11:21:00 to 11:43:00 in MLAT/MLT coordinate. (d) Φ_E after temporal interpolation and projected to 11:30:00. (e) Φ_E in geographic coordinates after coordinate transformation. (f) Φ_E in geographic coordinates after binning to TIEGCM grids (2.5° in latitude and longitude). (g) Original Φ_E maps of TIEGCM/AMIE. (h) Φ_E in TIEGCM after combining AMIE and SSUSI.

The combined TIEGCM/AMIE and SSUSI maps are shown in Figures 2.3h,

2.2c1-c2, which are referred to as "SSUSI-modified" auroral maps. These maps are generated every 30 s and used as inputs to drive TIEGCM. Comparing Figures 2.3g with 2.3h, 2.2a1 with 2.2c1, and 2.2a2 with 2.2c2, the overall position of the auroral oval is well represented in the default model maps, while the magnitudes of Φ_E and E_m are enhanced significantly and show regional-scale structures in the modified auroral precipitation. The combination of the observed and default auroral maps is susceptible to discontinuity at their boundaries (Figure 2.3h). We have performed sensitivity tests on the smooth boundary by taking the running mean of the adjacent nine grid points. The results do not show noticeable differences from the runs we present here.

It should be noted that for any particular geophysical location where the SSUSI-interpolated-binned map (Step 3) is available, real-time observations are still sparse, so most data points are from the temporal interpolation (Figure 2.3b). Even so, since the real-time continuous observations covering the whole auroral region are unavailable, temporal interpolation is one of the simplest realizations of the auroral variations that takes advantage of actual data information for this storm event. It also helps maintain the regional structure of the aurora. As shown in Section 2.3, the auroral modification improves the simulation of neutral temperatures and densities in general.

2.2.3 Implementing Subgrid Scale Electric Field Variability

To account for the effects of subgrid-scale electric field variabilities and their nonuniform spatial distribution, we change the Joule heating factor from 1.5 used in the default TIEGCM runs (Emery et al., 1999) to 1 in all of our runs and introduce the variability by adding a random number (E') to the resolved electric field

35

(E_0) obtained from the AMIE procedure at each time step. This is similar to the procedure conducted in Matsuo and Richmond (2008), which treated the subgrid variability as a stochastic process and simulated it as random numbers in numerical modeling. Cousins and Shepherd (2012a) derived the statistical distribution of small-scale electric field variabilities, which follow a two-sided exponential distribution $(\exp(-|E'|/\mu))$ instead of Gaussian (also seen in Golovchanskaya et al. (2006)), where E' is the electric field fluctuation, and μ is the characteristic electric field variability. Cousins and Shepherd (2012b) found that the ratios of μ to the background electric field are significant in the auroral zone (often close to 1) and minor in other regions. Following their work, we choose the random number (E') from an exponential distribution $\exp(-|E'|/\mu)$ with μ equal to the local resolved electric field (E_0) as the proxy of electric field variability. To correlate the variabilities to the spatial distribution of the auroral zone, they are added only in the strong auroral precipitation regions where Φ_E is over a threshold of $\Phi_{E0} = 5\,\mathrm{mW\,m^{-2}}$, that is,

$$\frac{\mu(t)}{E_0(t)} = \begin{cases} 0, \Phi_E < \Phi_{E0} \\[2mm] 1, \Phi_E > \Phi_{E0} \end{cases}$$

And the overall electric field adopted in the model is $E(t) = E_0(t) + E'(t)$. The ratio μ/E_0 and threshold Φ_{E0} are also adjusted by checking the temperature responses. Generally, a larger μ/E_0 and a smaller Φ_{E0} lead to a more significant temperature enhancement. A smaller threshold Φ_{E0} value leads to a larger area where the implementation of adding electric field variability is applied; a larger μ/E_0 ratio means that for the regions where electric field variabilities are added, the magnitudes of the variabilities are larger. We choose the numbers that best reproduce the neutral density and temperature observations while maintaining numerical stability.

2.3   Mechanism Studies of the Strong TTEIL

Mechanism studies are conducted in two major steps. The first step covers auroral regions (Sections 2.3.1-2.3.4), and the second step focuses on a local region around McMurdo (Section 2.3.5).

2.3.1   Neutral Temperature Structures and Simulated TTEIL

In compliance with the modifications introduced in Sections 2.2.2-2.2.3, four sets of TIEGCM runs with different configurations of the two high-latitude drivers are designed and named in Table 2.1.

Run 1 (AMIE)	use AMIE electric potential and auroral maps (control run)
Run 2 (AMIE_SSUSI)	use AMIE electric potential and SSUSI-modified auroral maps
Run 3 (AMIE_varE)	use AMIE electric potential and auroral maps, but add electric field variabilities in the auroral region according to SSUSI-modified auroral maps
Run 4 (AMIE_SSUSI_varE)	use AMIE electric potential, SSUSI-modified auroral maps, and add electric field variabilities according to SSUSI-modified auroral maps

Table 2.1: Name convention for each run and details

Figure 2.4a reveals the vertical profiles of neutral temperatures when and where the most significant temperature increase and a strong TTEIL are located (73.75°S, 117.5°W). We refer to this as the "primary TTEIL" to differentiate from the relatively weak ones generated by different mechanisms (to be discussed in Sections 2.3.3 and 2.3.5). Compared with Run 1 (black line), only Run 4 (red line) produces a significant temperature increase ($600\,\mathrm{K}$ at $150\,\mathrm{km}$) and a prominent TTEIL. In contrast, the magnitudes of the temperature increase from Runs 2 and 3 are much smaller. The comparisons of these neutral temperature profiles imply that the en-

hanced auroral precipitation (which increases ionization rates) and the electric field variabilities play essential roles in elevating the temperatures at 150 km and forming TTEIL.
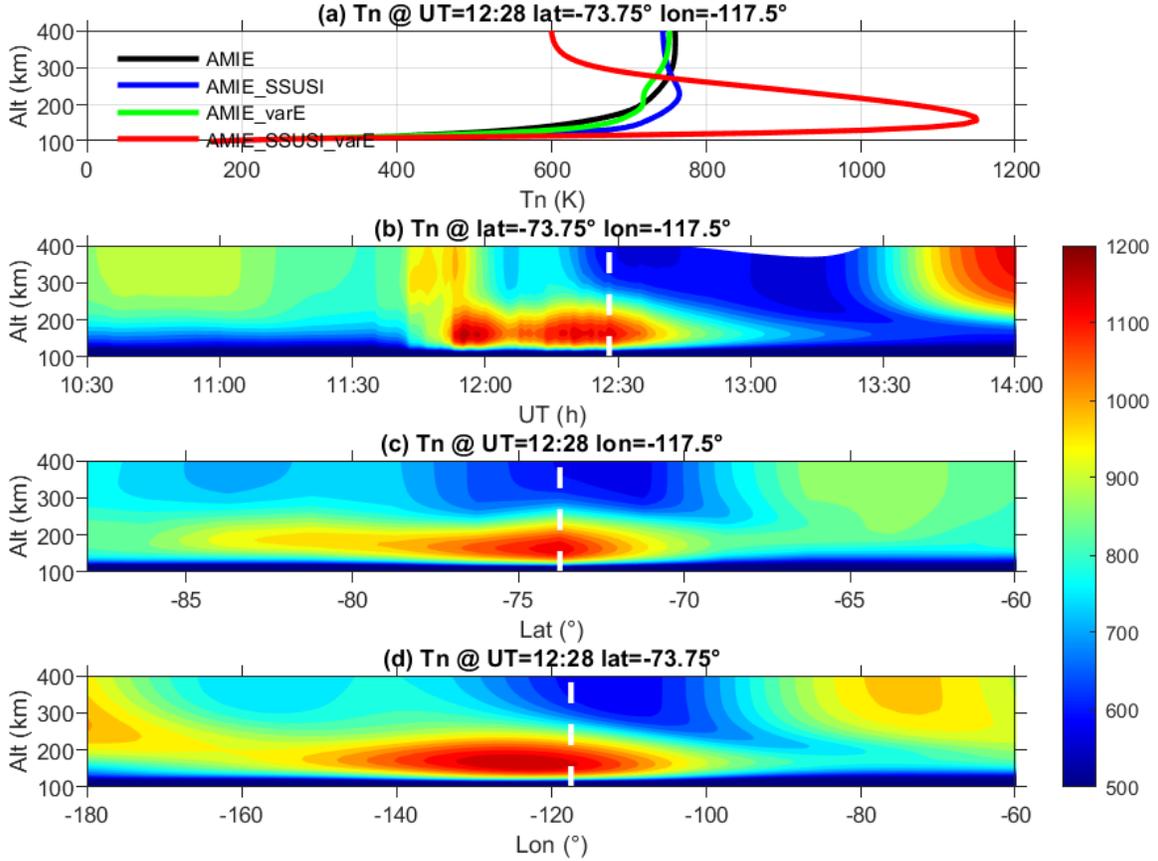
Figure 2.4: (a) Vertical temperature profiles at (73.75°S, 117.5°W) and 12:28:00 from four different TIEGCM runs indicated by different colors. (b-d) TTEIL structures from Run 4 in GLAT and GLON. White dashed lines highlight the time and location for the vertical profile shown in (a).

As described in Section 2.2.2, the SDs of auroral precipitation in terms of Φ_E and E_m are calculated and stored. We incorporate the variability of auroral precipitation into the model using the same method as that for the electric field, except that the random numbers follow Gaussian distribution. The model tests (not shown here) of adding subgrid auroral variability show similar vertical structures of

neutral temperatures as the runs without adding it, which suggests that the impact of auroral precipitation variability is minor compared with the precipitation itself. This is expected since Joule heating is linearly related to Pedersen conductivity determined by precipitation. The precipitation variability will likely be smoothed out with time and space leading to a minimal net effect on the Joule heating budget (Lu et al., 1998a).

We focus on Run 4 to demonstrate the temporal and spatial structures of TTEIL (Figures 2.4b-d). The horizontal red stripe around 100 km to 200 km from 11:50:00 to 12:50:00 indicates that the simulated TTEIL lasts for 1 h. Starting from 11:40:00, neutral temperature enhancement is present above 120 km. After 12:00:00, neutral temperature above 200 km starts to decrease but the enhancement between 120 km and 200 km lasts, which increases the contrast of temperatures between 150 km and above, strengthening the magnitude of TTEIL. TTEIL magnitude is defined as the temperature difference between the maximum temperature identified in the vertical range of 100 km to 200 km and the minimum temperature above this maximum temperature peak. We choose the time of 12:28:00 when TTEIL reaches its maximum to examine its spatial extension (Figures 2.4c-d). The latitudinal and longitudinal spans of TTEIL are 20° and 80°, indicating that TTEIL is a large-scale phenomenon.

2.3.2 Thermodynamic Term Analysis and TTEIL Formation Mechanism

TIEGCM calculates neutral temperature by solving the thermodynamic equation in the dimensionless log-pressure coordinate $z(= \ln p_0/p)$ with the form

$$
\begin{aligned}
\frac{\partial T}{\partial t} =& \frac{g\,\mathrm{e}^z}{p_0 C_p} \frac{\partial}{\partial z} \left[\frac{K_T}{H} \frac{\partial T}{\partial z} + K_E H^2 C_p \rho \left(\frac{g}{C_p} + \frac{1}{H} \frac{\partial T}{\partial z} \right) \right] \\
& - \boldsymbol{v} \cdot \nabla T - w \left(\frac{\partial T}{\partial z} + \frac{R^* T}{C_p m} \right) + \frac{Q - \mathrm{e}^z\, L_e}{C_p} - L_i T
\end{aligned}
\tag{2.1}
$$

Table 2.2 defines each letter.

T	neutral temperature
g	gravity constant
p_0	reference pressure
C_p	heat capacity in constant pressure
K_T	heat conduction coefficient
H	scale height
K_E	eddy diffusion coefficient
ρ	neutral density
v	horizontal neutral wind vector
w	vertical neutral wind $= \mathrm{d}z/\mathrm{d}t$
$R*$	universal gas constant
m	mean molar mass
Q	other diabatic heating except for heat conduction and eddy diffusion
L_e, L_i	diabatic cooling coefficients

Table 2.2: Definitions of each letter in the thermodynamic equation

The heating terms on the Right Hand Side (RHS) of Equation (2.1) are grouped into six categories, summarized in Table 2.3.

All parameters on the RHS of Equation (2.1) are outputs from the model simulation and retrieved every minute, as are their contributions to the time rate of temperature change $\partial T/\partial t$. Figure 2.5 reveals $\partial T/\partial t$ induced by Joule heating (Q_J), adiabatic cooling (Q_A), vertical heat advection (Q_V), horizontal heat advection (Q_H),

$\dfrac{g\,e^z}{p_0 C_p}\dfrac{\partial}{\partial z}\left(\dfrac{K_T}{H}\dfrac{\partial T}{\partial z}\right)$	heat conduction (Q_C)
$\dfrac{g\,e^z}{p_0 C_p}\dfrac{\partial}{\partial z}\left[K_E H^2 C_p \rho\left(\dfrac{g}{C_p}+\dfrac{1}{H}\dfrac{\partial T}{\partial z}\right)\right]$	vertical heat transfer by eddy diffusion
$-\boldsymbol{v}\cdot\nabla T$	horizontal heat advection (Q_H)
$-w\,\partial T/\partial z$	vertical heat advection (Q_V)
$-wR^* T/C_p m$	adiabatic cooling (Q_A)
Q/C_p	diabatic heating, including Joule heating (Q_J), particle precipitation, solar, and chemical heating
$-\,e^z L_e/C_p - L_i T$	diabatic cooling

Table 2.3: Explanation of heating terms in the thermodynamic equation

and heat conduction (Q_C) at (73.75°S, 117.5°W). Joule heating is the most important term at this location among all the diabatic heating terms when TTEIL is generated. Other diabatic heating/cooling terms in TIEGCM, including solar radiative heating and heating due to oxygen recombination, heat transfer by molecular diffusion, and diabatic cooling (such as NO and $O_3{}^+$ radiative cooling) are at least 1 order of magnitude smaller than the major terms in the altitude of 100 km to 200 km, which are thus ignored in the term analysis. In the long-term time scale, sometimes extending to the recovery phase of storms, the radiative cooling such as the NO cooling is large and important to the overall energy budget (Knipp et al., 2017; Lei et al., 2011; Li et al., 2019b; Lu et al., 2010).

Since Joule heating results from external driving (magnetospheric perturbation) and is related to electrodynamic processes, it appears much more sporadically (Figure 2.5a) than other heating/cooling terms which involve neutral motion and requires more inertia to change. Joule heating penetrates down to 120 km with several intense episodes lasting for a few minutes extending from 11:40:00 to 12:40:00. The maximum heating rate reaches over $3\,\mathrm{K\,s^{-1}}$ at 11:50:00. This strong heating changes the pressure gradient and leads to a horizontal divergence of airflow above
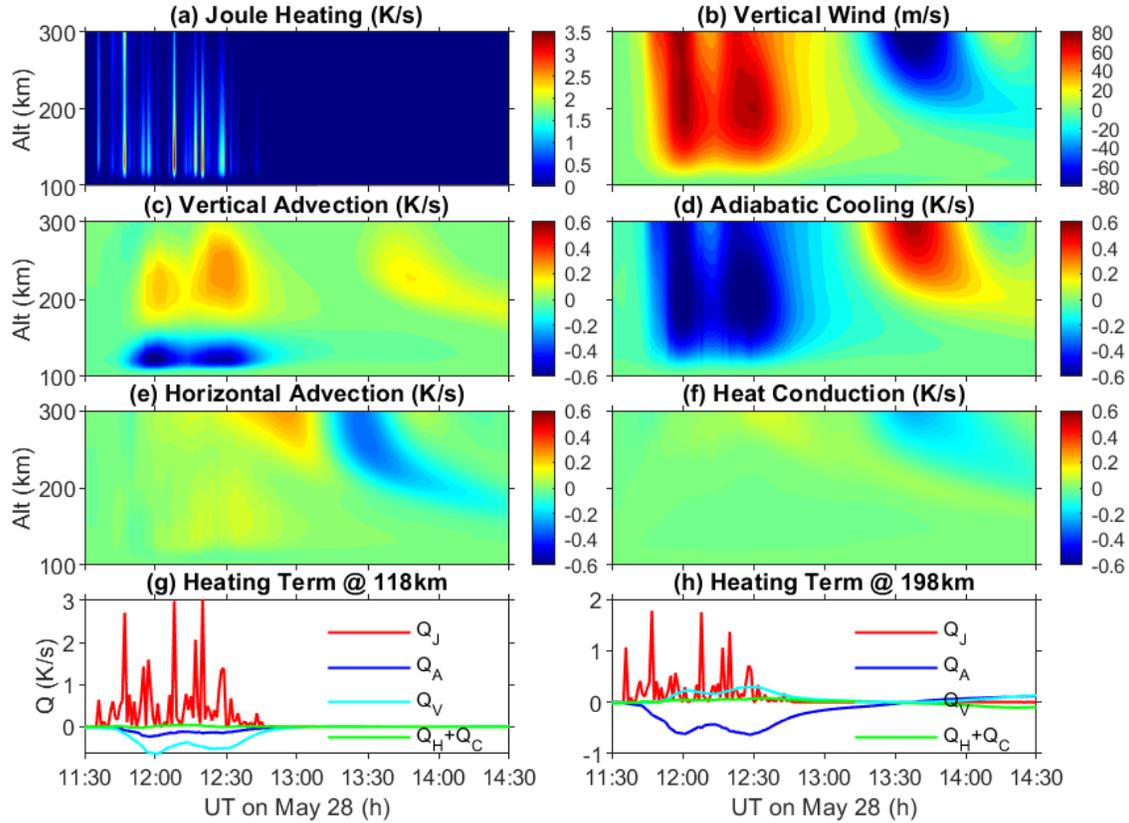
Figure 2.5: Temperature tendency ($\partial T/\partial t$) induced by (a) Joule heating, (c) adiabatic cooling, (e) vertical advection, (f) horizontal advection, and (g) heat conduction at (73.75°S, 117.5°W) where the maximum TTEIL is found. (b) Vertical winds at the same location. (d and h) $\partial T/\partial t$ induced by Joule heating, adiabatic cooling, vertical advection, and horizontal advection+conduction, at 120 km and 200 km, respectively.

120 km. The induced large upward vertical motion reaches $80 \, \mathrm{m \, s^{-1}}$ around $200 \, \mathrm{km}$ at 12:00:00, with a time delay of $10 \, \mathrm{min}$ after Joule heating peaks (Figure 2.5b). Accompanying this upward motion and the resultant atmosphere expansion is the adiabatic cooling reaching a negative maximum of $-0.6 \, \mathrm{K \, s^{-1}}$ at $200 \, \mathrm{km}$ (Figure 2.5c), consistent with the altitude where the vertical upward motion is the strongest. Since temperature peaks around $150 \, \mathrm{km}$, such an upward motion also transports cold air upward below the peak, leading to cooling (up to $-0.6 \, \mathrm{K \, s^{-1}}$) below $150 \, \mathrm{km}$. Above the peak, this upward motion brings warm air at $150 \, \mathrm{km}$ upward, leading to heating (up to $0.3 \, \mathrm{K \, s^{-1}}$) at higher altitudes (Figure 2.5e). Horizontal heat advection and heat conduction tend to be moderate most of the time and play minor roles below $250 \, \mathrm{km}$ (Figures 2.5f-g). The time series of these heating/cooling terms at $120 \, \mathrm{km}$ and $200 \, \mathrm{km}$ are shown in Figures 2.5d and 2.5h, respectively. Joule heating (red line) is dominant at $120 \, \mathrm{km}$, while adiabatic cooling (blue line) becomes strong at $200 \, \mathrm{km}$. Vertical advection (black line) induces cooling at $120 \, \mathrm{km}$, while warming at $200 \, \mathrm{km}$. The effects of horizontal heat advection + heat conduction (green line) are small at both altitudes compared to the other terms.

To further identify which heating/cooling terms dominate TTEIL, we perform a post-processing diagnostic analysis of the temperature evolution induced by each term on the RHS of Equation (2.1) and by various combinations of these terms. Considering the impact of an individual term, we start from zero and add up the time integration of $\partial T / \partial t$ on fixed pressure levels only caused by this term to examine the net temperature changes. Figures 2.6a-d show the temperature profiles by only integrating Joule heating, adiabatic cooling, vertical heat advection, horizontal heat advection, and heat conduction, respectively. For the different combinations of these terms, we start from an initial temperature profile and perform similar temperature integrations (Figures 2.6e-g). Since the intense Joule heating mainly occurs after

43

11:30:00, after which the neutral atmosphere starts to respond and change dramatically, we choose the temperature profile at 11:30:00 as an initial state. TTEIL is a time-integrated result of the thermodynamic balance of heating/cooling terms; thus, an integrated view is suitable for diagnosing their relative contribution to temperature changes.

Figure 2.6a shows Joule heating alone is capable of producing a strong inversion layer with temperature increases by over $2000\,\mathrm{K}$ within $40\,\mathrm{min}$ (from 11:50:00 to 12:30:00) and peaks around $120\,\mathrm{km}$. Along with the reduced temperature induced by adiabatic cooling peaking at $180\,\mathrm{km}$ (Figure 2.6b), a sharp inversion layer with a peak around $120\,\mathrm{km}$ and temperature drop reaching over $1500\,\mathrm{K}$ above the peak forms after 12:30:00 (Figure 2.6e). After accounting for the vertical heat advection (Figure 2.6c), however, the peak altitude where the temperature reaches maximum moves up to $150\,\mathrm{km}$, and the inversion layer becomes much broader and weaker (Figures 2.6f, after 12:00:00, and 2.6i). This indicates that vertical heat advection, which cools the layer below $150\,\mathrm{km}$ and warms it above, cancels out the sharp temperature gradient formed by Joule heating and adiabatic cooling in the lower thermosphere, thereby hindering the formation of a sharp TTEIL with the peak at a very low altitude. Horizontal advection and heat conduction play a minor role in shaping the temperature profile below $200\,\mathrm{km}$ by comparing Figures 2.6f and 2.6g, which is consistent with their minor heating/cooling effects (Figures 2.5f-g, and 2.6d). By comparing Figures 2.6g and 2.6h, the overall temperature evolution is similar. The differences mainly originate from the numerical smoothing effect, which is not included in the thermodynamic equation (Shapiro, 1970) but is implemented at every time step in the model to maintain numerical stability.

From Figures 2.5 and 2.6, the physical processes in generating the primary TTEIL in the model are summarized as follows:
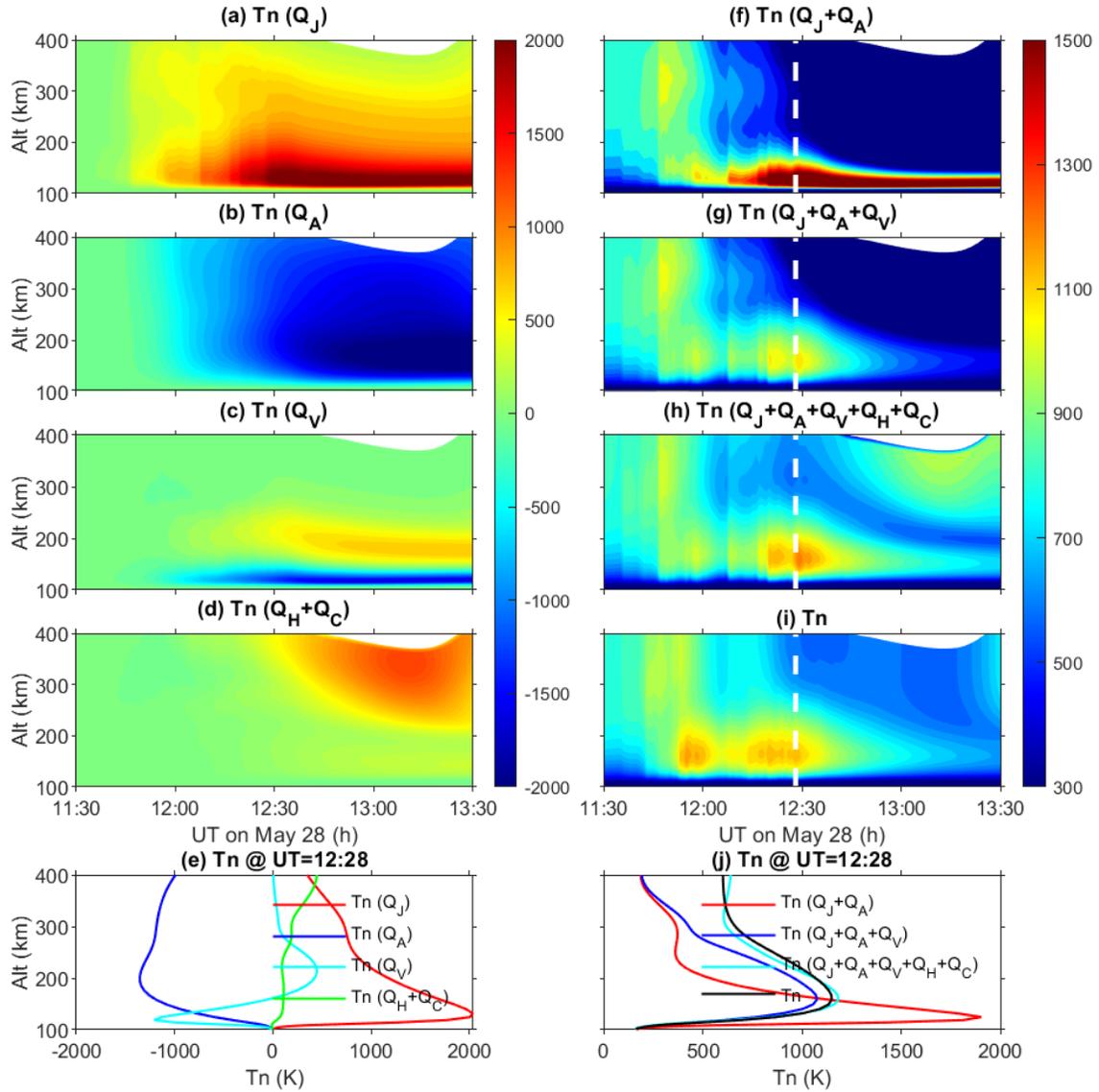
Figure 2.6: Temperature profiles at (73.75°S, 117.5°W) forced solely by (a) Joule heating (Q_J), (b) adiabatic cooling (Q_A), (c) vertical advection (Q_V), and (d) horizontal advection (Q_H) + heat conduction (Q_C), integrated from the initial state of zero temperatures at 11:30:00. Temperature profiles forced by the combination of (e) $Q_J + Q_A$, (f) $Q_J + Q_A + Q_V$, and (g) $Q_J + Q_A + Q_V + Q_H + Q_C$. (h) The model results (Run 4) are given as a reference. (i) The vertical temperature profiles of (e-h) at 12:28:00 (dashed black lines in a-h), respectively. Units are K.

1. strong Joule heating penetrating down to 120 km in a relatively short period changes the pressure gradient, and horizontal divergence of airflow induces upward air motion;

2. upward wind induces strong adiabatic cooling reaching a maximum at 200 km;

3. strong heating at 120 km and cooling at 200 km lead to strong vertical differential heating (Figures 2.5d and 2.5h) and a sharp TTEIL;

4. vertical heat advection, on the other hand, acts as a strong cooling term below 150 km and a moderate heating term above (Figure 2.5e), therefore playing a negative effect in forming the sharp TTEIL in general (Figure 2.6i).

Therefore, the intense Joule heating is a trigger. At the same time, the neutral dynamic terms, including adiabatic cooling and vertical advection, contribute to the formation and the ultimate structure of the TTEIL in the model simulation.

2.3.3 Electrodynamics and Neutral Dynamics Associated With Joule Heating

Joule heating $Q_J = \sigma_P(\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B})^2$ is determined mainly by Pedersen conductivity and electric fields. Figure 2.7 shows the time series of Pedersen conductivity, the magnitude of the total electric field $\sqrt{E_x^2 + E_y^2}$ and Joule heating interpolated to 150 km obtained from four different model runs. Considering that TTEIL is a large-scale phenomenon and the local electric fields are produced by adding random variabilities, the parameters shown in Figure 2.7 are averaged within the ambient nine grids around the location of primary TTEIL to form a regional and statistical picture. Comparing black and green lines with blue and red lines in Figure 2.7a, high-latitude conductivity calculated from the default Φ_E and E_m maps is elevated by 3 times on

46

average after incorporating SSUSI auroral maps from 11:30:00 to 12:30:00. The magnitudes of electric fields after adding variabilities demonstrate more variations with time than the original AMIE fields, while their means remain at a comparable level (Figure 2.7b). Such significant electric field variations are also reported in the literature (Heppner, 1972); in real-time observations from SSIES onboard DMSP, large spikes in ion drifts are also identified, which indicates significant variations in electric fields (to be discussed in Section 2.3.5). The combined effects of the enhanced conductivity and electric field variability dramatically enhance Joule heating after 11:30:00. The comparison of the green and blue lines with the red line in Figure 2.7c illustrates that increasing conductivity or adding electric field variability alone cannot enhance Joule heating significantly. Therefore, realistic auroral precipitation and electric field (including its mean and variability), as the two critical high-latitude drivers of I-T models, are essential for calculating Joule heating and the resultant neutral atmosphere responses (Zhu et al., 2020).

Using the same SSUSI-modified auroral maps, Figure 2.7a shows that Pederson conductivity from Run 4 (red) is more significant than that from Run 2 (blue). Adding the electric field variability, which alters local Joule heating, and then the neutral temperature and number density (converted from neutral mass density using mean molar mass calculated from mass mixing ratios of O, O_2 and N_2), tends to impose feedback on the local conductivity which is determined by both neutral and plasma densities (Kelley, 2009). Neutral number density shown in Figure 2.7d is higher in Run 4 than in Run 3 while plasma density barely changes at 150 km (not shown here), which suggests that neutral number density is mainly responsible for the conductivity change. The increase of local neutral number density in Run 4 is likely associated with the thermal expansion of the air below this altitude triggered by the enhanced Joule heating, and the denser air supplied from below leads to the increase of neutral
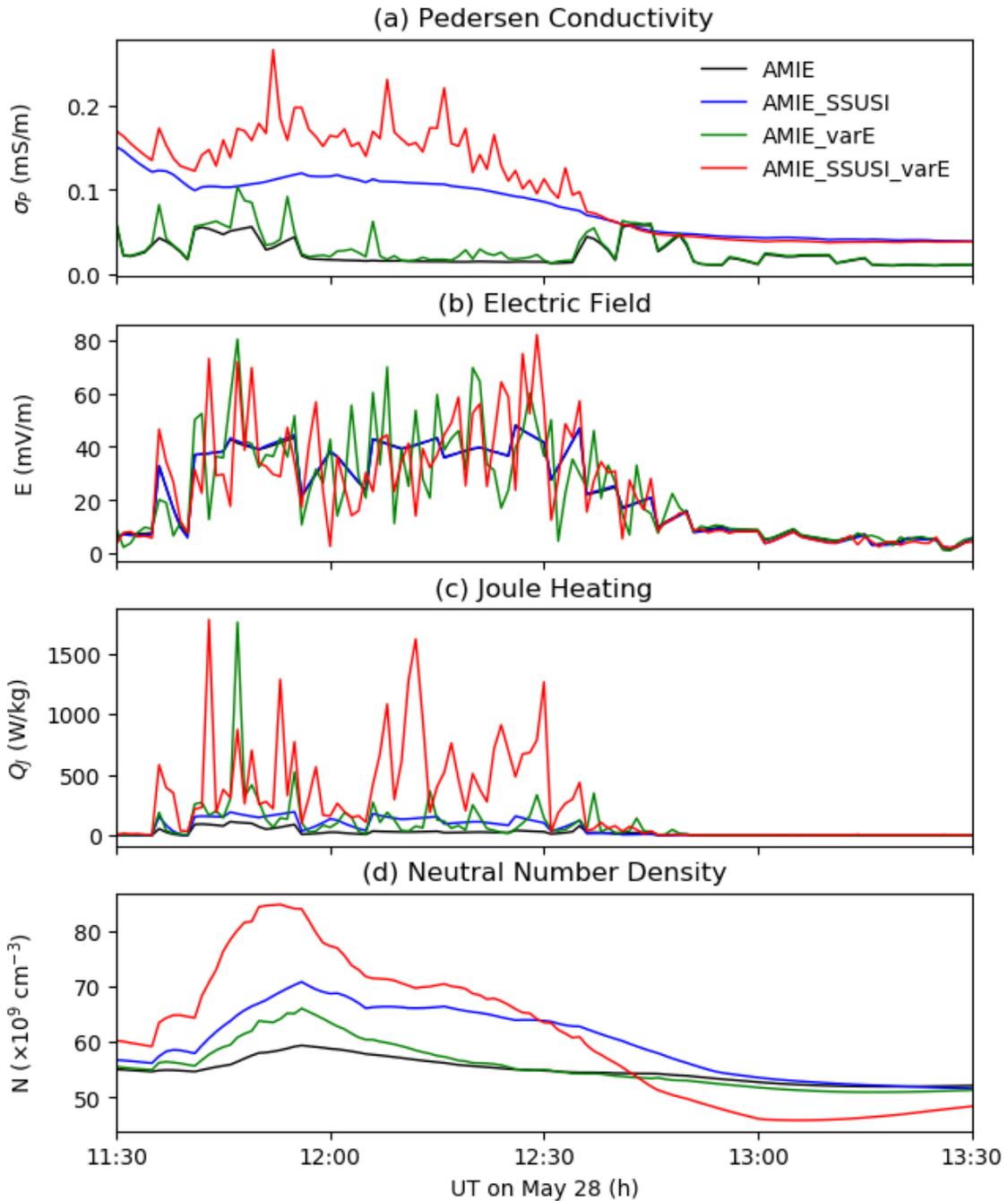
Figure 2.7: TIEGCM simulation of (a) Pedersen conductivity, (b) electric field magnitude, (c) Joule heating, and (d) neutral number density averaged within nearest nine grid points around (73.75°S, 117.5°W) at 150 km. The model results from Runs 1 to 4 are indicated in black, blue, green, and red, respectively.

number density at this altitude ($150\,\text{km}$).

The global effect of Joule heating enhancement on neutral dynamics is also examined. Figure 2.8 shows the polar view of Joule heating, TTEIL magnitude (defined in Section 2.3.1), neutral temperatures and vertical winds at two different altitudes in geographic coordinates. As shown in Figures 2.8(1a-b), TTEIL emerges where Joule heating is intense (the maximum integrated Joule heating reaches $1\,\text{W}\,\text{m}^{-2}$). The large vertical wind is present at both $165\,\text{km}$ and $222\,\text{km}$ (Figures 2.8(1d1-d2)). At the same time, neutral temperatures increase at lower altitudes but decrease at higher altitudes (Figures 2.8(1c1-c2)) due to the strong differential heating. This picture is consistent with what is shown in Figures 2.5 and 2.6. An intriguing feature is the wave-like perturbations in temperatures and vertical winds, which propagate outward from the primary TTEIL region with time, indicating a large-scale TAD structure. By perturbing the neutral dynamics in the surrounding environment, TTEIL itself also propagates outward spirally and forms TTEIL at other locations, as shown in Figure 2.8(2b). At 12:53:00, the front of TTEIL approaches McMurdo (Figure 2.8(2b)), but with a magnitude significantly reduced (compared to Figure 2.8(1b)). The magnitude of TTEIL at McMurdo remains similar and then gradually diminishes afterward.

2.3.4 Comparison of Neutral Mass Densities With Satellite Observations

Figure 2.7d shows that the neutral number density increases at $150\,\text{km}$ in Run 4 compared with the other runs. To examine whether the neutral mass densities from our modified simulations are in a reasonable scope, we compare the neutral mass densities from the four different runs with GRACE measurements in Figure 2.9a. All simulation results are projected along the satellite orbit and interpolated vertically
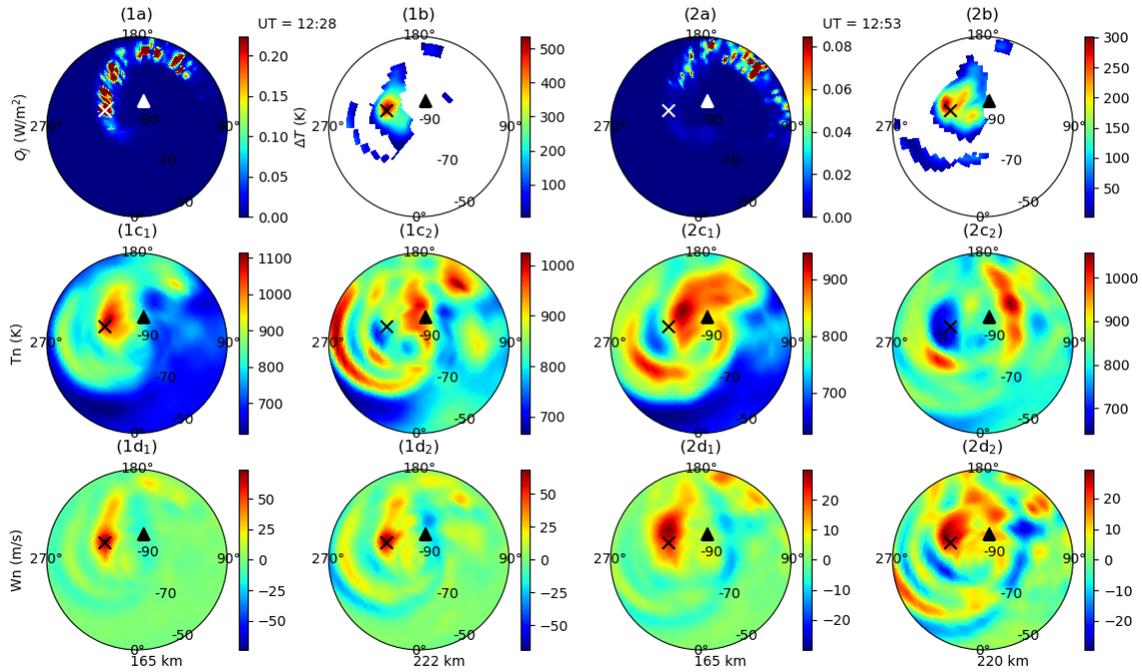
Figure 2.8: Spatial distribution of (1a) height integrated Joule heating, (1b) TTEIL magnitude, neutral temperatures at (1c1) 165 km and (1c2) 222 km, and vertical winds at (1d1) 165 km and (1d2) 222 km, at 12:28:00 in geographic coordinates. The 180° longitude is plotted at the top to visualize TTEIL propagation better. The right two columns are the same except for 12:53:00 and the two examined altitudes are 165 km and 220 km. Cross and triangle represent the location of the strongest TTEIL (73.75°S, 117.5°W) and McMurdo (77.8°S, 166.7°E), respectively.

to the satellite altitude (475 km). The corresponding GRACE orbit as a function of MLAT, MLT, GLAT, and Solar Local Time (SLT) is indicated in Figure 2.9b. The neutral mass density from the control run underestimates GRACE observations by 40 % during the geomagnetically active period. This underestimation is consistent with the previous studies showing that the default configurations of I-T models tend to underestimate the neutral mass densities along satellite orbits during geomagnetically active times, likely due to the underestimation of Joule heating (Deng et al., 2013; Shim et al., 2012). After SSUSI auroral maps and electrical field variabilities are implemented in the model, the simulated neutral mass density is elevated overall and best matches the observations (red and gray lines in Figure 2.9a). The better agreements with observations are found not only in high-latitude nightside regions where the modification of the model drivers is made but also in high-latitude dayside and low-latitude regions, which imply that the neutral densities around 475 km can be affected by neutral dynamics (such as TADs) induced by remote heating sources. Overall, the model-GRACE comparison suggests that our modification in both aurora and electric field, which increases Joule heating, generally improves the simulation of neutral mass density response to the storm globally.

A closer examination of Figure 2.9 shows at some locations (e.g., time around 09:30:00, 11:00:00, 14:00:00, and 15:30:00), model simulations are still smaller than observations. Some of the underestimations occur in dayside high-latitude regions (such as 14:00:00 and 15:30:00), which might be related to localized FACs in the cusp region (Luhr et al., 2004). Since the current TIEGCM runs do not treat FACs in the cusp region as a direct input, the reasons for the discrepancy deserve more detailed investigation in future studies. The underestimations in nighttime high-latitude regions such as those around 13:00:00 and 14:30:00 may be more directly related to Joule heating still being underestimated locally. This is not unexpected
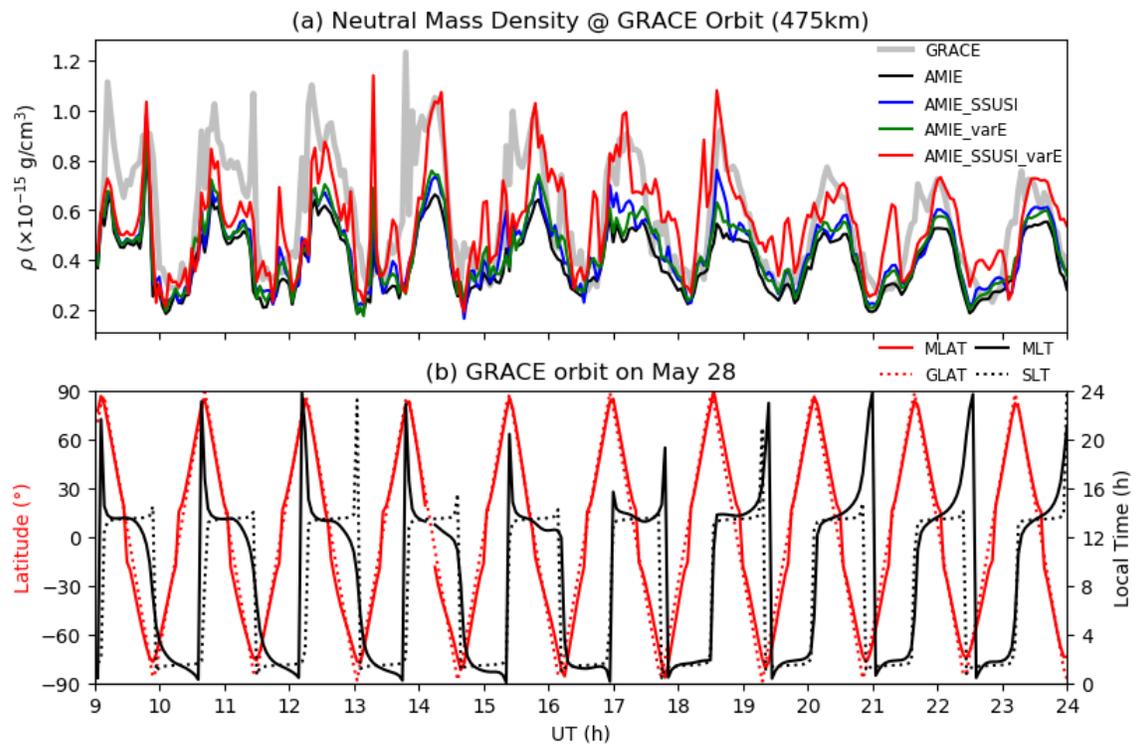
Figure 2.9: (a) Neutral mass densities projected along GRACE orbit from the four different TIEGCM runs (colored lines) and the comparison with GRACE observations (gray line). (b) GRACE orbit during the time: Red solid line shows MLAT and the black solid line shows MLT, the red dotted line shows GLAT and the black dotted line shows SLT.

since the interpolation and binning method for auroral maps, and AMIE may miss the localized real-time auroral precipitation and electric field, respectively. For instance, at any specific point, such as McMurdo, the satellite (SSUSI) observations cover a tiny portion of the time series. Hence, most of the auroral precipitation at a particular location used in TIEGCM has to be interpreted from the scarce observations plus the empirical model (Section 2.2.2). In addition, electric field variability in the model is added as a random number, which may also miss the realistic features locally.

2.3.5 TTEIL Structure at McMurdo With Local Joule Heating Enhancement

Indeed, the underestimation of electric field fluctuation is seen at McMurdo. During the time window of 14:46:00 to 14:56:00 when TTEIL was detected by lidar (15:00:00), DMSP F16 has conjunction measurements (Figure 2.10a). The satellite orbit is marked as the black line from right to left. Strong spectral radiances in LBH short band observed by SSUSI appear very close to McMurdo at 14:50:00 (Figure 2.10a). During the same time window, ion drift measurements from Ion Drift Meter (IDM) in SSIES on DMSP show large spikes (up to $3\,\mathrm{km\,s^{-1}}$) and fluctuations indicating strong local electric fields near McMurdo (Figures 2.10b-c). Red vertical lines mark five consecutive time ticks with the corresponding orbit locations marked as red dots in Figure 2.10a. The significant ion drift ($3\,\mathrm{km\,s^{-1}}$) is identified within the auroral region. Such auroral enhancement and rapid-varying large electric field near McMurdo are not captured in Run 4. Therefore, after the TTEIL signal propagates to McMurdo at 12:53:00 in Run 4 with instead a small magnitude (Figure 2.11a1, red line), it does not last for long, and TTEIL structure is missing at 15:00:00 (Figure 2.11a2, black line). The discrepancies in timing and TTEIL magnitude at McMurdo

prompt us to further consider the localized high-latitude drivers by incorporating DMSP observations around McMurdo and examining the temperature responses.
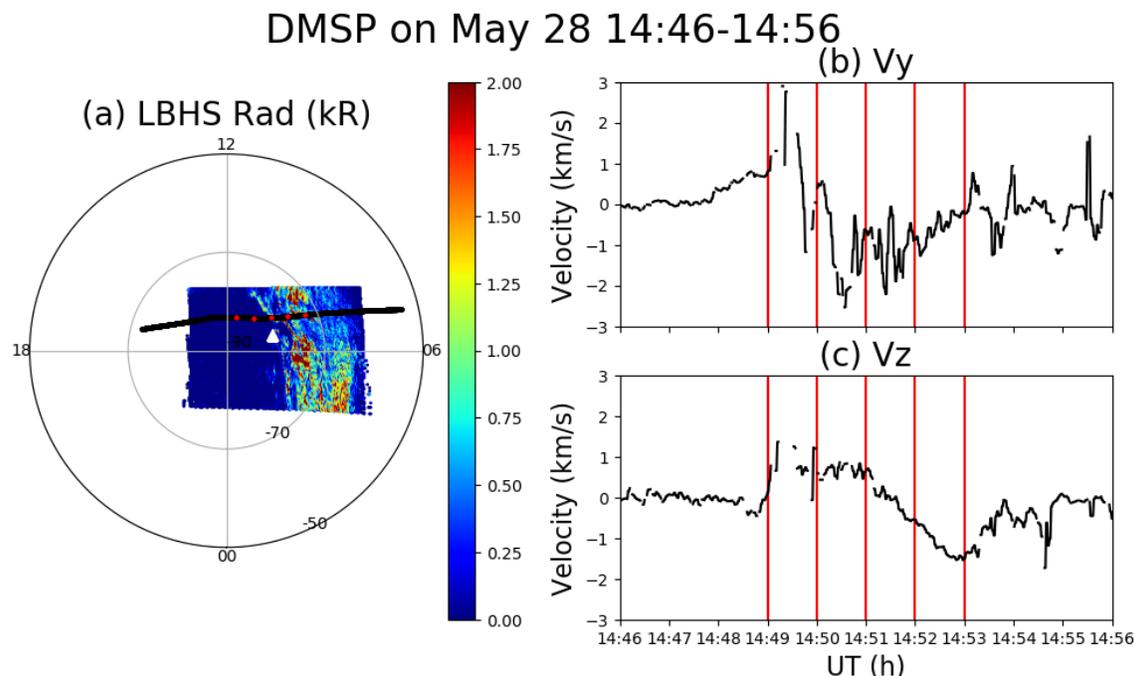
Figure 2.10: (a) DMSP F16 SSUSI observation of spectral radiance at LBH short band from 14:46:00 to 14:56:00 in MLAT/MLT coordinate; black line marks satellite orbit which travels from right to left; red dots mark the spatial locations of five consecutive time ticks; white triangle denotes McMurdo. (b-c) DMSP F16 SSIES measurements of cross-track ion drift (V_y) and vertical ion drift (V_z) from IDM during the same period. Red vertical lines mark five consecutive time ticks in (a).

Run 5 (named AMIE_SSUSI_varE_McM) is performed based on the settings of Run 4 but with enhanced auroral precipitation and electric field added around Mc-Murdo. We set auroral mean energy to be $10\,\mathrm{keV}$ and energy flux to be $30\,\mathrm{mW\,m^{-2}}$ from 14:45:00 to 14:55:00, during which McMurdo is located at the poleward edge of the auroral oval (Figure 2.10a). Within this period, we impose an electric field of $150\,\mathrm{mV\,m^{-1}}$ in the north/south direction (E_y) with flipping direction every $2\,\mathrm{min}$ similar to Deng et al. (2009) and Zhu et al. (2018). Note that the evolution of DMSP ion drifts shown in Figures 2.10b-c contains temporal and spatial variations. They

provide a reference for the possible magnitudes but not exact values of the electric fields over McMurdo. The direction and magnitude of electric fields are chosen to reproduce ion drifts comparable to the maximum IDM observation (Figure 2.10b), in which the cross-track ion drifts are primarily oriented along eastward/westward direction. Auroral energy and electric field are tuned to reproduce a similar magnitude of temperature peak close to the lidar observation. Such local adjustment is implemented uniformly in a region centered at (78.75°S, 167.5°E) with a spatial expansion of three grid points in latitude and seven grids in longitude. Imposing an additional electric field in the east/west direction (E_x) is also done, but the difference between these two tests is slight, thus not shown here.

The simulation results are shown in Figures 2.11a2-f2. Compared with the results from Run 4 at the exact location and time (black line in Figure 2.11a2), the temperature inversion structure in Run 5 (red line) is more prominent. The temperature peak is more significant and closer to the lidar observation (over 1000 K). This temperature structure is comparable to the primary TTEIL studied in Section 2.3.2. The heating term analysis shows that Joule heating is the most substantial thermal term and penetrates to 120 km. The second and third dominant terms are the strong adiabatic cooling accompanying the upward motion and vertical advection. Again, similar to the mechanism for the primary TTEIL formation studied in Section 2.3.2, Joule heating leads to temperature increase below 150 km, adiabatic cooling leads to temperature decrease above 120 km altitude, and along with vertical advection, these three terms determine the vertical structure of neutral temperature.

To show the unique features associated with the primary TTEIL generated at McMurdo in Run 5, the same term analysis is carried out for the weak TTEIL identified around 12:53:00 in Run 4 for comparison (Figures 2.11b1-f1). In Run 5, the intense Joule heating penetrates down to 120 km and produces a secondary and also
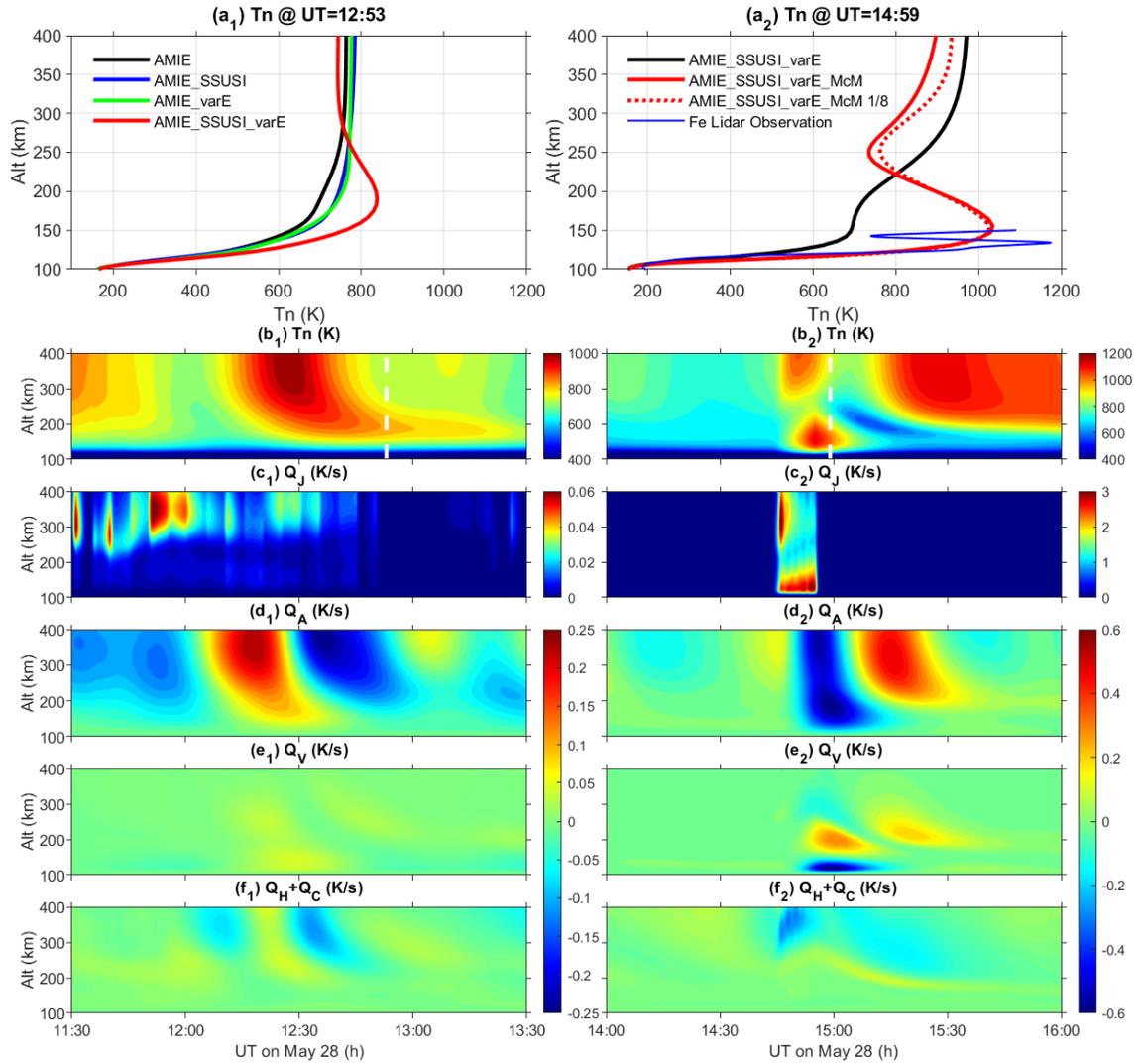
Figure 2.11: (a1) Vertical temperature profiles from Runs 1 to 4 at 12:53:00, McMurdo. (a2) Vertical temperature profiles from Runs 4 and 5 at 14:59:00, McMurdo; the temperature profile from high-resolution configuration of Run 5 (AMIE_SSUSI_varE_McM 1/8) at 14:55:00 are also plotted in red dashed line. (b1-f1) Temporal variations of neutral temperature, Joule heating, adiabatic heating/cooling, vertical advection, and horizontal advection + heat conduction at McMurdo, obtained from Run 4. (b2-f2) Same as (b1-f1) except that the results are obtained from Run 5 during a different period.

56

more substantial peak in the E region than the F region peak around 300 km (Figure 2.11c2). In Run 4, the F region peak is more prominent, and the penetration to the E region is insignificant (Figure 2.11c1). Distinct from Run 5, adiabatic heating/cooling is the most substantial thermal term in Run 4, contributing the most to the morphology of the weak TTEIL according to the analysis of integrated temperature profiles (not shown here). The vertical wind perturbation likely induces alternating adiabatic heating/cooling due to the propagation of TADs. The comparison between Runs 4 and 5 indicates that for TTEIL formation at different times/locations, the relative importance of Joule heating, adiabatic cooling, vertical advection, horizontal advection, and heat conduction can change. For the formation of primary TTEIL (characterized by huge temperature enhancement and large TTEIL magnitude), the localized intense Joule heating is a critical trigger, which likely happened at McMurdo. For the weak TTEIL, the dynamics associated with wave-induced transport and disturbances may play dominant roles.

2.4   Discussion

By refining the high-latitude drivers globally and further fine-tuning their localized characteristics around McMurdo, Run 5 captures a more realistic temperature elevation at McMurdo and compares better with the lidar observations in neutral temperature than the other runs. Despite numerous improvements, there are still noticeable discrepancies between the model simulation and lidar observations by comparing Figure 2.11a2 and Figure 2.1a. The peak altitude is about 20 km higher than the observation (130 km), and the modeled layer width is much broader than the observation. We have performed a few other tests to sort out possible causes for the discrepancies, even though they are insufficient to resolve them. To give a reference

for future modeling, we list them as follows:

1. Test of the model lower boundary: The same modifications of high-latitude drivers are implemented into TIMEGCM (Roble and Ridley, 1994), which is governed by similar dynamic and thermodynamic equations in the upper atmosphere as TIEGCM but with a lower boundary at 30 km. The simulated TTEIL still has the same peak altitude at 150 km, which suggests that the lower boundary of TIEGCM (97 km) close to the observed TTEIL is likely not the main reason for the high peak altitude in the simulations.

2. To examine whether the broad vertical span is due to the "strong" diffusion coefficients, we decrease the eddy diffusion coefficient (K_E) and heat conduction coefficient (K_T) in the model to half. The test results show that the vertical temperature structure does not differ much from what is shown in Figure 2.4, which is consistent with the minor contribution from the diffusion terms in the thermodynamic term analysis (Section 2.3.2).

3. The possibility of high altitude for temperature peak and broad span due to low vertical resolution is also examined by performing the high-resolution (1.25° in latitude and longitude, 1/8 scale height) TIEGCM run. The vertical temperature profile from the high-resolution model run using the settings for Run 5 (named AMIE_SSUSI_varE_McM 1/8) is illustrated as the red dashed line in Figure 2.11a2. Intensifying aurora and electric field around McMurdo produce nearly the same temperatures as the standard resolution runs below 220 km, but the maximum magnitude of TTEIL occurs 4 min earlier. The altitude for the peak temperature enhancement (150 km) and the mechanism for producing TTEIL at McMurdo are the same.

4. Liu et al. (2016) found that the electron temperature can increase by $2000\,\mathrm{K}$ at $100\,\mathrm{km}$ to $130\,\mathrm{km}$ due to Farley-Buneman Instability (FBI). Still, it is not known whether such instability can also lead to a significant neutral temperature increase. We test this possibility by implementing the electron anomalous heating module resulting from FBI and find that neutral temperature is only elevated by several K. Such an insignificant neutral temperature change may be caused by the electron-to-neutral density ratio at this altitude so small that the energy exchange from electrons to neutrals is insufficient to heat the neutrals, leaving neutral temperatures almost unchanged.

It should be noted that the above test results and their implication rely on physics, chemistry, and electrodynamics considered in TIEGCM. The effects from the processes not considered in the model, such as non-hydrostatic processes and local heating induced by wave breaking, cannot be tested for now. In addition, gravity wave activities in the lower boundary of TIEGCM have not been included at a level comparable to observations (Chen et al., 2016; Chu et al., 2011), so their effects are not considered in the current model. The narrow width of the observed TTEIL, likely associated with a vertically propagating wave structure (Figure 2.1a), could be a combined effect from the wave perturbations in electrodynamic or neutral dynamic processes (Chu and Yu, 2017) superimposed on the elevated temperature background. Testing all these possibilities is out of the scope of this study and deserves further investigation.

2.5 Conclusion

To understand the underlying physics for the significant thermospheric temperature enhancement ($550\,\mathrm{K}$ at $130\,\mathrm{km}$) and inversion layer (TTEIL) structure ob-

served by lidar at McMurdo during a geomagnetic storm on 2011-05-28, we modify auroral maps and electric field variabilities in TIEGCM with guidance from DMSP observations and previous statistical results from SuperDARN. Such modifications are made because the default model runs with empirical high-latitude drivers that cannot adequately reproduce storm responses. We first incorporate auroral precipitation observed by DMSP/SSUSI into the auroral maps, which leads to increased conductivity. On average, Pedersen conductivity increases by 3 times at $150\,\mathrm{km}$ in the auroral region during the active period. We then consider the subgrid-scale electric field variabilities by adding a randomly generated number in the auroral region to the original electric field obtained from AMIE. Two modifications are implemented to cover high-latitude regions first and then focus on the local region around McMurdo. The modifications bring the simulation of neutral densities close to GRACE observations and generate TTEIL, sharing similarities with lidar observations.

The most important effects of the modified auroral precipitation and electric field variability are significant enhancements in Joule heating and its deeper penetration. To generate the primary TTEIL, including the one observed at McMurdo, the local intense Joule heating is an essential factor that not only provides the heating source down to $120\,\mathrm{km}$ but also triggers the following dynamical processes that work together with Joule heating to shape the ultimate temperature structure. Associated with the intense Joule heating, pressure gradient changes and large upward winds of over $80\,\mathrm{m\,s^{-1}}$ resulting from a horizontal divergence of airflow lead to strong adiabatic cooling at high altitudes ($-0.6\,\mathrm{K\,s^{-1}}$ maximizing at $200\,\mathrm{km}$). Intense Joule heating at $120\,\mathrm{km}$ and adiabatic cooling higher up result in intense differential heating vertically and generate a sharp temperature inversion. On the contrary, vertical heat advection induces strong cooling ($-0.6\,\mathrm{K\,s^{-1}}$) below $150\,\mathrm{km}$ and moderate heating above ($0.3\,\mathrm{K\,s^{-1}}$) mainly due to the upward vertical motion and the temperature peak at

150 km. Therefore, vertical advection weakens TTEIL and renders a higher temperature peak altitude and a broader layer than the observations. The vertical structure of the primary TTEIL is mainly determined by the combined effects of these three thermodynamic terms. Simulation results also show that TADs are generated from where the maximum Joule heating is deposited, and the primary TTEIL is generated. TADs propagate outward to other locations, forming TTEILs with much smaller magnitudes than the primary ones in distant regions. For the weak TTEIL, the local Joule heating is minimal, and temperature structure can be mainly determined by other terms such as adiabatic heating/cooling effects.

The improved agreement with the observed neutral density and temperature in our modified simulations indicates the importance of regional-scale auroral precipitation and electric fields as critical high-latitude drivers in the I-T models. Our results also imply that the local fine-scale characteristics of high-latitude drivers, often missing in empirical drivers, are essential for simulating local storm-time dynamics. Run 5, which incorporates the high-latitude modification of auroral maps and electric field variability, along with additional local enhancement in Joule heating around Mc-Murdo, reasonably reproduces the dramatic temperature enhancement in the lower thermosphere observed by the lidar. However, the modeled peak altitude is still higher (150 km) than the observed one (130 km). Note that the amount of heating needed in the modeling is large (maximum value is on the order of $1\,\mathrm{W\,m^{-2}}$), but it is not implausible. We suspect that this much heating may require, in reality, some physics not included in the current TIEGCM, for example, strong anomalous heating associated with plasma irregularities, in addition to the missing physics mentioned earlier. The unsolved discrepancies and the possible modeling efforts to tackle this problem remain open questions.

2.6 Revisit Auroral and Electric Field Modeling

Before we end this chapter, we must revisit how we incorporate aurora and electric fields into TIEGCM.

2.6.1 Evaluation of the Interpolation of Aurora

The essential step is that we interpolated SSUSI measurements in time (Section 2.2.2) to expand the coverage of auroral observations. We compared the interpolated aurora and concurrent SSUSI observations and found a similarity. However, one question remains: how does the interpolated aurora approach the "truth", which is the aurora in the real world? Of course, the truth is unknown and one-to-one comparison is not achievable. In other words, can we find a way to quantify the deviation of the processed data from the truth? For example, for each observation, an associated error (uncertainty) estimates the confidence of this measurement. Then can we find a way to estimate the uncertainty of the interpolated aurora?

More generally, there might be multiple data sources for other cases, such as satellite-borne instruments and ground-based observation sites. Each data source provides a distinct spatial and temporal coverage and a different measurement error. Is there a way to synthesize them together and give an estimate of the uncertainty for the final data product? This problem will be answered in Chapter 3.

2.6.2 Shortage of Implementing Subgrid Electric Field

In implementing subgrid electric fields, we assume that AMIE captures the resolved scale electric field well, which is generally a reasonable assumption (Knipp, 1989). However, the treatment of the subgrid electric field is debatable. We add a random perturbation to every spatial location (quantified by auroral distribution)

and treat that as the effect of the subgrid electric field on the total electric field. However, a simple hypothetical experiment as in Figure 2.12 will demonstrate the shortage. The spatial structure of the mean field (above row) is wholly disrupted when adding up the perturbation field (bottom row). However, such strong spatial discontinuity is never supported by observations.
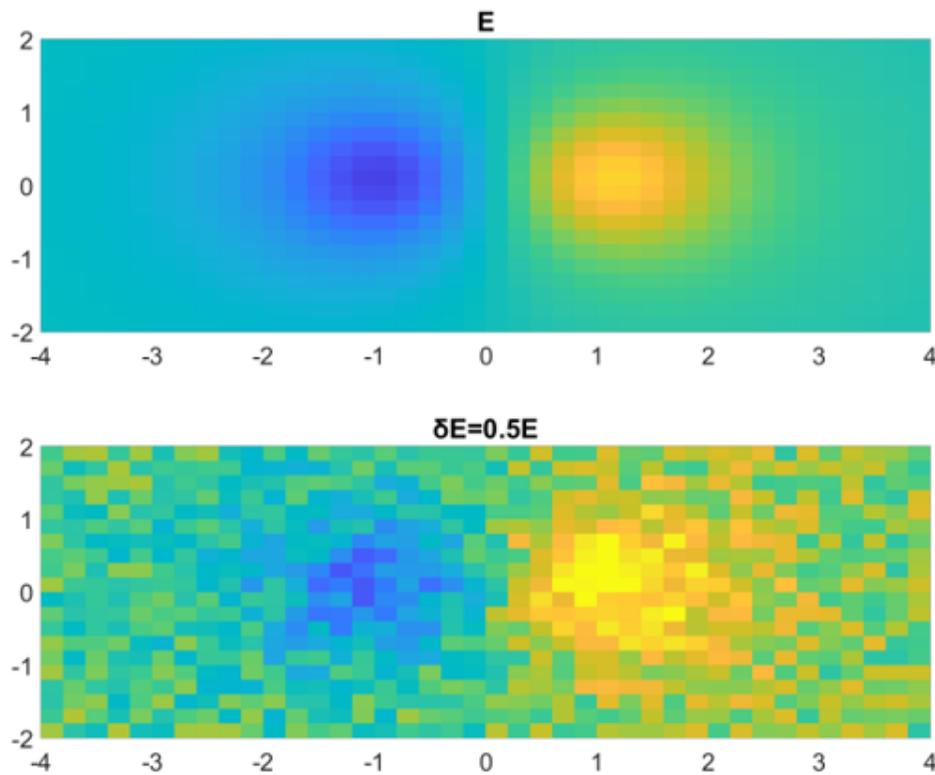
Figure 2.12: An arbitrary field with perturbations at half of the mean.

The spatial correlation is missing in the treatment of electric field variability. Without considering the spatial correlation, it is usually the case that two adjacent locations behave drastically differently. Such discontinuity is generally harmful in numerical models since this introduces unrealistic perturbations and changes the physical processes we are studying. So a natural question arises: how can spatial

variations be modeled with spatial correlations considered? This question will be answered in Chapter 4.

Chapter 3

Developing a Multi-Resolution Data Assimilation Model for Auroral Observations and Empirical Models

We apply a multiresolution Gaussian process model (Lattice Kriging) to combine satellite observations, ground-based observations, and an empirical auroral model to produce the assimilation of auroral energy flux and mean energy over high-latitude regions. Compared to simple padding, the assimilation coherently combines various data inputs leading to continuous transitions between different datasets. The multiresolution modeling capability is achieved by allocating multiple layers of basis functions with different resolutions. Higher-resolution fitting results capture more mesoscale (10 km to 100 km) structures, such as auroral arcs, than the low-resolution ones and the empirical model. Two preprocessing steps, temporal interpolation of satellite data and spatial down-sampling of low-fidelity data, are implemented to rec-

oncile different datasets better. The inherent smoothing effect of the fitting, which causes an unrealistic spreading of the aurora, is mitigated by a post-processing step: the K-Nearest Neighbors (KNN) algorithm. KNN identifies the probability of a region with significant aurora, thereby eliminating regions with low values. This methodology can maintain realistic and mesoscale auroral structures without boundary issues. We then run TIEGCM driven by the high- and low-resolution auroral assimilations and compare TECs. TIEGCM driven by data assimilation produces enhanced TECs by a factor of 2 than the one driven by the empirical aurora, and high-resolution results show mesoscale structures. Our study shows the value of incorporating realistic auroral inputs via assimilation to drive I-T models to better understand the consequences of mesoscale phenomena.

3.1 Introduction

The dynamics and electrodynamics of the I-T system are closely related to the magnetosphere's coupling of magnetically conjugate regions and its interaction with the solar wind (Wang et al., 2004; Wiltberger et al., 2004). In particular, during geomagnetically active periods (magnetic storms or substorms), the I-T system manifests a series of ionospheric phenomena like enhanced aurora, electric fields, and FACs in response to the different phases of day and night-side reconnection (Nishimura et al., 2021a). These ionospheric processes further affect the neutral atmosphere by exchanging and transporting momentum, energy, and composition in the coupled M-I-T system (Thayer and Semeter, 2004). Eventually, the change of the I-T system will lead to the change of conductance, current systems, and ion outflows, which in turn poses a feedback effect to the magnetosphere (Merkin and Lyon, 2010; Merkine et al., 2003; Tanaka, 2007; Yau and Andre, 1997).

Due to the dynamic and turbulent nature of magnetospheric processes, the high-latitude forcing of the I-T system, such as aurora, electric fields, and FACs, is highly multiscale, and mesoscale structures that play an essential role in the M-I-T coupling (Nishimura et al., 2021a). For example, empirical models usually give the large-scale ($> 1000\,\mathrm{km}$) morphology of the auroral oval, which can differ significantly from the ground-based ASI observations such as those from THEMIS ASIs (Donovan et al., 2006). THEMIS ASIs depict rich mesoscale ($10\,\mathrm{km}$ to $100\,\mathrm{km}$) structures, while a narrow FOV imaging can even resolve small-scale ($< 10\,\mathrm{km}$) structures. During an expansion phase of a substorm, the auroral structure with scales smaller than $500\,\mathrm{km}$ contributes to $50\,\%$ of the total energy flux. Mesoscale auroral processes such as poleward moving auroral forms, polar cap patches, auroral arcs, and streamers can give feedback to the large-scale dynamics and impose net effects on the global distribution of electron densities (Gabrielse et al., 2021). Similar to the aurora, electric fields also show multiscale features. Using SuperDARN measurements, Cousins and Shepherd (2012b) found a large ratio ($75\,\%$) of mesoscale to large-scale electric fields in terms of magnitude under a southward IMF condition. The scale analyses by Cousins et al. (2015) and Shi et al. (2020) show that mesoscale FACs contribute to nearly $60\,\%$ of the spatial variability of FACs. These magnetosphere-originated processes are highly correlated. Mesoscale auroral structures such as auroral arcs are often associated with enhanced FACs and electric fields (Nishimura et al., 2021a), which profoundly affect the I-T system.

For I-T models, aurora and electric fields are the two most important drivers at high latitudes. Thus, it is critical to capture these two drivers realistically. This study focuses on the assimilation of auroral particle precipitation, specifically energy flux and mean energy. Even though the empirical auroral models derived from historical data can capture large scales reasonably well (Newell et al., 2009; Wu et al., 2021;

Zhang and Paxton, 2008; Zhu et al., 2021), they still miss the essential mesoscale features. Wu et al. (2020) showed that only when the empirical auroral model is replaced by auroral observations from SSUSI (Paxton and Meng, 1999; Paxton et al., 2002) onboard the DMSP satellites to drive the I-T model, TTEIL observed by the Fe Boltzmann lidar at McMurdo, Antarctica can be reproduced. Neutral densities in the F region match GRACE observations. Similarly, Sheng et al. (2020) implemented THEMIS ASI auroral observations into GITM and compared them with the simulations driven by the empirical model. The authors found that the magnitude of TIDs in GITM is almost doubled when driven by realistic THEMIS ASI observations and more consistent with observations. These previous studies indicate the necessity of developing data-driven auroral maps for high-latitude drivers, primarily when we focus on specific storms. Such efforts have rarely been made, and the current work aims to address this challenge.

The existing techniques for auroral measurements include satellite and ground-based imagers, which have distinct spatial coverage and temporal samplings. SSUSI/DMSP measures global auroral emissions with a high spatial resolution and a revisit time of 30 min (three satellites) to the same MLAT and MLT. Ground-based instruments such as THEMIS ASIs provide both high temporal and spatial resolution (3 s, 0.1°) observations in North America. Empirical auroral models are built upon the statistics of many historical observations and provide global auroral maps with highly smoothed patterns (Hardy et al., 1985; Roble and Ridley, 1987). They usually deviate from real-time observations, especially for mesoscale features. These deviations can often lead to systematic biases in estimating the general auroral activity level. The empirical model can still provide sensible information for large-scale features such as auroral boundaries. These data sources provide complementary information on auroral activities but are rarely used synergistically. One way to combine all data sources is

by simply padding different types of auroral observations. Still, this method usually leads to discontinuous boundaries among different data sources and introduces unphysical gradients, which could lead to artificial perturbations in I-T models. Another approach to synthesizing various data sources is AMIE (Lu, 2017; Richmond, 1992; Richmond and Kamide, 1988), but its resolution is limited by the order of spherical cap harmonics (Matsuo, 2020).

In this paper, we apply a novel multiresolution spatial Gaussian process model (Lattice Kriging, Nychka et al. (2015)) to incorporate auroral observations from satellite and ground-based data, as well as an empirical model where observations are unavailable. It uses a range-limited basis function that better serves localized auroral assimilation. The mesoscale features in the satellite and ground-based observations are primarily kept in the assimilation results. In addition, the multiresolution modeling capability is fulfilled by locating multiple layers of basis functions with different resolutions. This method, therefore, provides a valuable tool to study the multiscale processes and the corresponding impacts. Lattice Kriging has already been used in lower atmospheric studies like surface temperature analysis (Heaton et al., 2019; Wiens et al., 2020). Wu and Lu (2022) have extended this model to vector fields and assimilated high-latitude electric fields using SuperDARN and PFISR data, demonstrating its effectiveness in space weather studies. It is the first time that this model has been applied to auroral assimilation.

The manuscript is organized as follows. Section 3.2 introduces the data sources. Section 3.3 describes the Lattice Kriging model, including the principles and mathematical formula. Section 3.4 provides the detailed procedures to apply this model for auroral assimilation. Section 3.5 presents TIEGCM simulations driven by the empirical auroral model and the two scales of auroral assimilation maps. Section 3.6 gives the conclusions and discussion.

3.2 Data Sources

The data sources used for the auroral assimilation include SSUSI onboard three DMSP satellites (F16, F17, F18), THEMIS ASIs, and a Kp-based empirical auroral model (Zhang and Paxton, 2008). The choice of empirical models is relatively flexible as long as the model provides full MLAT and MLT coverage. SSUSI is a remote-sensing instrument that measures ultraviolet emissions from the Earth's upper atmosphere in five different wavelength bands. The spatial resolution of SSUSI data product is $0.15°$, sufficient to analyze the mesoscale structures of aurora in this study. The derived data products include the precipitating electron mean energy and energy flux. The three satellites sweep through the polar cap alternatively every 30 min, sampling through the auroral region (a swath) across the pole.

THEMIS ASIs observe the white light aurora over the North American continent from Canada to Alaska at a sampling rate of every 3 s, which provides high-resolution information about the rapid evolution of the aurora. The white light data are converted to red-green-blue colors by comparing them with the nearest Northern Solar Terrestrial Array (NORSTAR) meridian scanning photometers and multispectral ASIs. Then the color ratios are converted to energy fluxes and mean energies using the Strickland et al. (1983) formula (Mende et al., 2008). In this study, the electron mean energy and energy flux maps of spatial resolution $0.1°$ are used, and the data are temporally down-sampled to a 1 min basis.

The Zhang and Paxton (2008) model is built upon 4 years of Global Ultraviolet Imager (GUVI) data onboard TIMED satellite from 2002 to 2005. The model provides auroral predictions (mean energy and energy flux) covering all MLAT/MLT sectors overall Kp ranges (0 to 9) using Epstein function fitting. Comparing to other Kp-based models like Hardy et al. (1987), this model provides a more physical spec-

ification of the geo-effective energy flux and mean energy. Such information is also helpful for the assessment of the statistical mean needed in the auroral assimilation (Equation (3.1) in Section 3.3.1) and for the regions where observations are not available. The empirical model can be generated at an arbitrary resolution, that is, on the satellite grids in the present study.

Owing to the noticeable auroral activity and decent data coverage on 2014-02-20, we use the auroral observations on this day as an example to demonstrate the methodology. The geomagnetic indices are shown in Figure 3.1. After 03:00:00, a negative turning of IMF B_z marks the start of geomagnetic disturbances. The Kp index reaches 6, and the SYM-H index reaches $-100\,\mathrm{nT}$, indicative of a moderate-intense storm. Significant variations in the AE indices reach $1200\,\mathrm{nT}$, suggesting considerable auroral activities during this period.

Figure 3.2 displays the auroral energy fluxes from the three data sources in the northern hemisphere at 11:50:00, plotted in MLAT and MLT coordinates. This UT is chosen due to the clear auroral structures in SSUSI and THEMIS observations. The instantaneous SSUSI observations are limited: to assimilate the auroral maps (mean energy and energy flux) for a particular time, SSUSI data falling into a $20\,\mathrm{min}$ time window ($10\,\mathrm{min}$ before and $10\,\mathrm{min}$ after) are gathered. For example, SSUSI data from 11:40:00 to 12:00:00 are binned for the auroral assimilation at 11:50:00 (Figure 3.2a, more details in Section 3.4.1). SSUSI observations after the binning mainly cover the dawn and dusk sectors and show scattered auroral arc features. THEMIS ASIs provide night-time observations with several auroral enhancements spreading between $60°$ and $70°$ MLAT around midnight. The empirical model has a locally much smaller magnitude and smoother structure than the actual observations but provides reasonable large-scale patterns and auroral boundaries for $\mathrm{Kp} = 6$ geomagnetic condition.
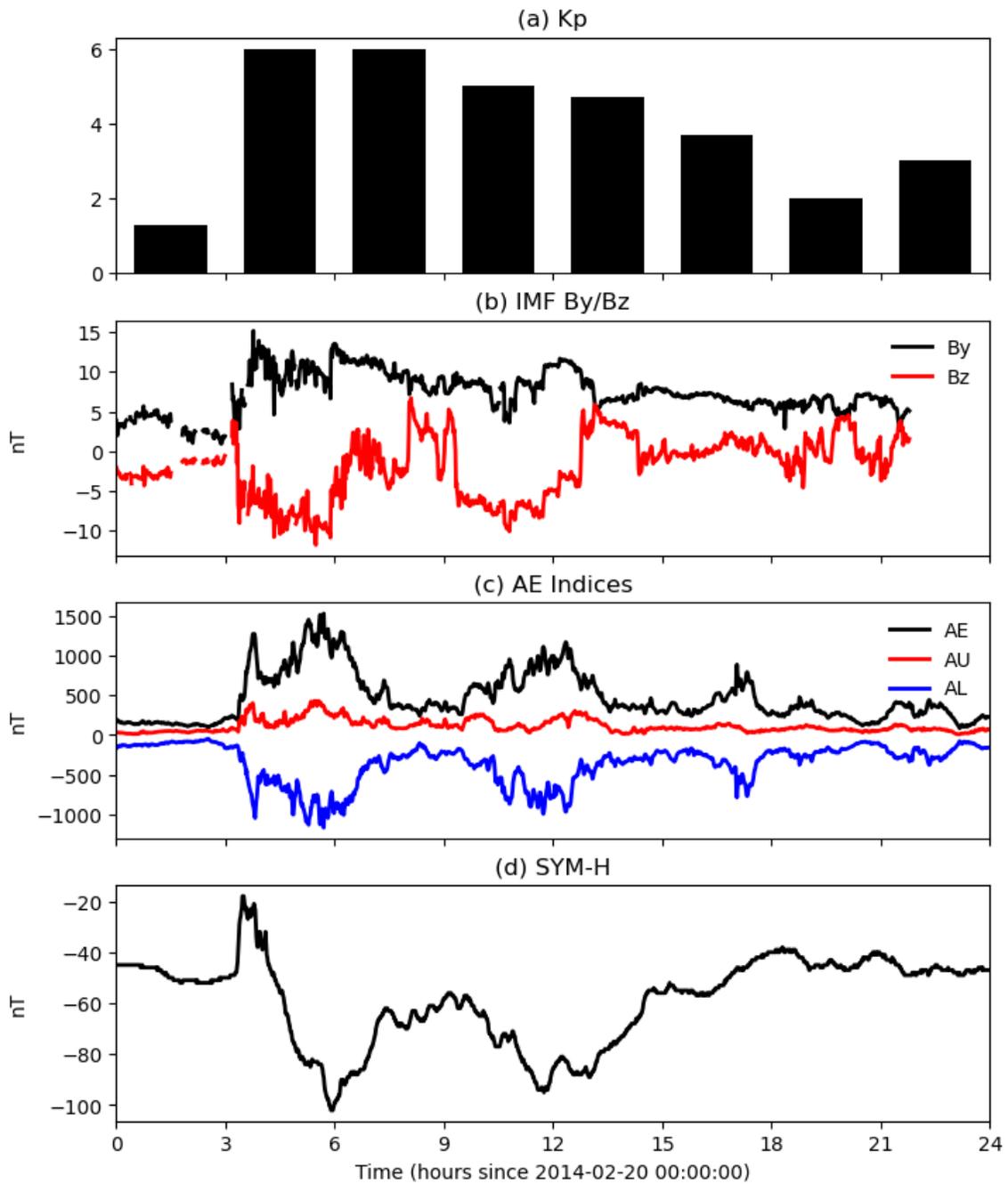
Figure 3.1: Geomagnetic indices on 2014-02-20 (a) Kp, (b) IMF B_y (black) and B_z (red), (c) AE indices: AL (blue), AU (red) and AE (black), and (d) SYM-H.
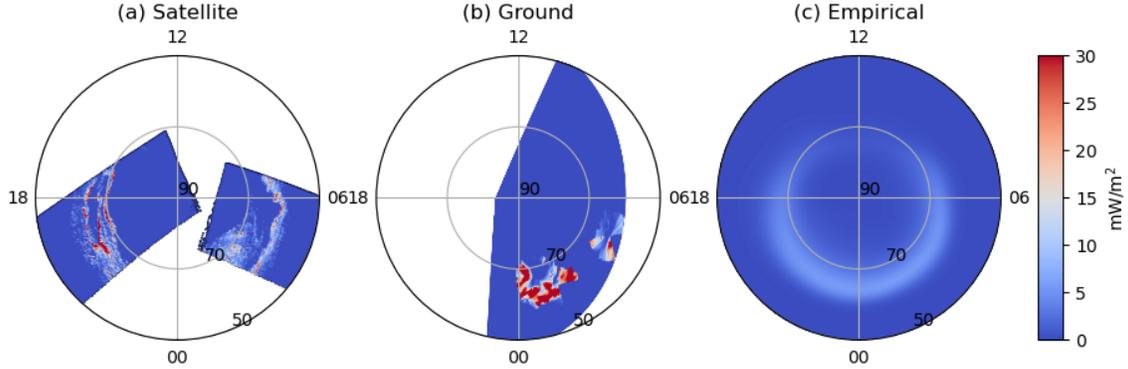
Figure 3.2: Auroral energy fluxes from (a) SSUSI, (b) THEMIS ASI, and (c) empirical model at 11:50:00. All are plotted in MLAT and MLT coordinates. Unit is $\mathrm{mW\,m^{-2}}$.

3.3 Lattice Kriging Model

In this section, we introduce the principles of the Gaussian process model adopted for Lattice Kriging (Section 3.3.1) and the implementation of the multiresolution data assimilation (Section 3.3.2).

3.3.1 Principles of Gaussian Process Model

We consider a spatial field y, whose values at all locations in a spatial domain $\mathbb{X}$ are assumed to follow a Gaussian process. Hence, the value at every finite location $\{x_i, 1 \leq i \leq n\}$ follows a multivariate normal distribution. We will use the observations $\{y(x_i), 1 \leq i \leq n\}$, where n is the total number of observations to predict the values for a set of new locations $\{x_i', 1 \leq i \leq n'\}$ without observations, $\{y'(x_i'), 1 \leq i \leq n'\}$. Krige (1951) states that the best prediction, in terms of minimizing the prediction variance, for y at any unobserved location x_i' can be expressed as a linear superposition of the observed values, that is, $\hat{y}(x_i') = \sum_{i=1}^{n} a_i y(x_i) + a_0$, and these optimal coefficients, $\{a_i, 0 \leq i \leq n\}$ can be estimated from the observed data.

The spatial field of interest (the auroral map in this study) can be decomposed into a combination of a spatially varying mean $\mu(x)$, a spatially correlated field, $g(x)$, and a spatially uncorrelated error term $\epsilon(x)$, which represents measurement uncertainties:

$$y(x) = \mu(x) + g(x) + \epsilon(x), x \in \mathbb{X} \tag{3.1}$$

The spatial mean function $\mu(x)$ can be retrieved from our auroral application's empirical model $z(x)$. As indicated earlier, there tends to be a systematic discrepancy between the values from the empirical model and observation. Here, we assume a scaling factor d to account for such a multiplicative bias, that is, $\mu(x) = z(x)d$.

The majority of the spatial predictability is achieved through the spatial random field, $g(x)$, which characterizes the detailed spatial variations of aurora. In this work, we take a spatial basis function approach by decomposing $g(x)$ onto a series of predefined basis functions $\{\phi_j(x), 1 \leq j \leq m\}$, that is, $g(x) = \sum c_j \phi_j(x)$, where c_j is the coefficient of jth basis function and m is the total number of basis functions (see Section 3.3.2 for further details about the basis functions). The coefficient vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_m)$ jointly follow a multivariate normal distribution with mean zero and covariance matrix $\boldsymbol{Q}^{-1}$ (therefore, $\boldsymbol{Q}$ represents the distribution's precision matrix). As a result, $g(x)$ is a zero mean Gaussian process, and the covariance function takes the following form:

$$\text{cov}\left(g(x), g\left(x'\right)\right) = \sum_{1 \leq j, j' \leq m} \rho \phi_j(x) \boldsymbol{Q}_{j,j'}^{-1} \phi_{j'}(x'), x, x' \in \mathbb{X}$$

where ρ is the spatial marginal variance of the process of interest. The detailed description for $\boldsymbol{Q}^{-1}$ is given in Appendix A.

We will use matrix notation to simplify the presentation regarding parameter

estimation and spatial prediction. First, we write the basis functions evaluated at the observed locations into an $n \times m$ matrix $\boldsymbol{\phi}$ such that $\boldsymbol{\phi}_{ij} = \phi_j(x_i)$, the value of the jth basis function at x_i. We use a vector x to denote the observed spatial locations, that is, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. Then, we have $g(\boldsymbol{x}) = \boldsymbol{\phi}\boldsymbol{c}$ and the covariance matrix of $\text{cov}(g(x), g(x')) = \rho\boldsymbol{\phi}\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T$.

Second, we stack all auroral observations $\{y(x_i), 1 \leq i \leq n\}$ and errors $\{\epsilon(x_i), 1 \leq i \leq n\}$ into vectors $\boldsymbol{y}$ and $\boldsymbol{\epsilon}$, respectively. Since we assume the errors are spatially uncorrelated, the covariance matrix of $\boldsymbol{\epsilon}$ is $\sigma^2\boldsymbol{W}^{-1}$, where $\boldsymbol{W}^{-1}$ is a diagonal error covariance matrix, and σ^2 is a scaling factor of the error term. We also stack the empirical model at each location $\{z(x_i), 1 \leq i \leq n\}$ as vector $\boldsymbol{Z}$. We can now write the model (Equation (3.1)) in the following matrix form:

$$\boldsymbol{y} = \boldsymbol{Z}d + \boldsymbol{\phi}\boldsymbol{c} + \boldsymbol{\epsilon} \tag{3.2}$$

Here, $\boldsymbol{y}$ follows a MVN distribution with a mean of $\boldsymbol{Z}d$ and a covariance matrix $\rho\boldsymbol{\phi}\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T + \sigma^2\boldsymbol{W}^{-1}$

$$\boldsymbol{y} \sim \text{MVN}\left(\boldsymbol{Z}d, \rho\boldsymbol{\phi}\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T + \sigma^2\boldsymbol{W}^{-1}\right)$$

In terms of parameter estimation, we will need to estimate the fixed scaling constant d and the spatially varying effects at the observed locations $\boldsymbol{c}$ based on observations $\boldsymbol{y}$ and their spatial locations $\boldsymbol{x}$. The best estimates of d and the conditional distribution of $\boldsymbol{c}$ can be obtained via the standard results of generalized least squares

(Cressie, 1993), which are

$$\hat{d} = \left(\boldsymbol{Z}^T \boldsymbol{M}_\lambda^{-1} \boldsymbol{Z} \right)^{-1} \boldsymbol{Z}^T \boldsymbol{M}_\lambda^{-1} \boldsymbol{y}$$

$$\left[\boldsymbol{c} | \boldsymbol{y}, d, \sigma^2, \rho, \boldsymbol{Q}^{-1} \right] \sim \mathrm{MVN} \left(\boldsymbol{Q}^{-1} \boldsymbol{\phi}^T \boldsymbol{M}_\lambda^{-1} (\boldsymbol{y} - \boldsymbol{Z}d), \rho \boldsymbol{Q}^{-1} - \rho \boldsymbol{Q}^{-1} \boldsymbol{\phi}^T \boldsymbol{M}_\lambda^{-1} \boldsymbol{\phi} \boldsymbol{Q}^{-1} \right)$$

where $\boldsymbol{M}_\lambda = \boldsymbol{\phi} \boldsymbol{Q}^{-1} \boldsymbol{\phi}^T + \lambda \boldsymbol{W}^{-1}$ and $\lambda = \sigma^2/\rho$. Then, the estimate of $\boldsymbol{c}$ is set to the conditional mean

$$\hat{\boldsymbol{c}} = \boldsymbol{Q}^{-1} \boldsymbol{\phi}^T \boldsymbol{M}_\lambda^{-1} (\boldsymbol{y} - \boldsymbol{Z}d)$$

and the variance of $\hat{\boldsymbol{c}}$ is

$$\mathrm{var} \left(\hat{\boldsymbol{c}} \right) = \rho \boldsymbol{Q}^{-1} - \rho \boldsymbol{Q}^{-1} \boldsymbol{\phi}^T \boldsymbol{M}_\lambda^{-1} \boldsymbol{\phi} \boldsymbol{Q}^{-1}$$

Therefore, the predictions (conditional mean and variances) of $\hat{\boldsymbol{y}}'$ at new locations are

$$\hat{\boldsymbol{y}}' = \boldsymbol{Z}'d + \boldsymbol{\phi}' \hat{\boldsymbol{c}} \tag{3.3}$$

$$\mathrm{var} \left(\hat{\boldsymbol{y}}' \right) = \boldsymbol{\phi}' \left(\rho \boldsymbol{Q}^{-1} - \rho \boldsymbol{Q}^{-1} \boldsymbol{\phi}^T \boldsymbol{M}_\lambda^{-1} \boldsymbol{\phi} \boldsymbol{Q}^{-1} \right) \boldsymbol{\phi}' \tag{3.4}$$

where the primes on $\hat{\boldsymbol{y}}'$, $\boldsymbol{Z}'$, and $\boldsymbol{\phi}'$ indicate that the prediction can be taken at different locations from the input data. A more detailed derivation of $\hat{d}$ and $\hat{\boldsymbol{c}}$ can be found in Appendix A.

In summary, our goal is to predict the values $\boldsymbol{y}'$ at unobserved locations $\boldsymbol{x}'$ (with corresponding empirical model output $\boldsymbol{Z}$ as a predictor) and to quantify the prediction uncertainty. Equation 3.3 gives the prediction of conditional mean $\hat{\boldsymbol{y}}'$, and the associated prediction uncertainty is the square root of the diagonal terms in Equation 3.4. In real applications, the calculations of variances are relatively computationally

expensive. Therefore, the variances at each spatial location are usually approximated by the sample variance of independent draws from the conditional distribution of $\hat{\boldsymbol{y}}'$, given available observations (Monte Carlo method).

3.3.2 Multiresolution Capability and Implementation

The auroral data in MLAT (ψ) and MLT (t) coordinates are mapped to the modeling coordinates, which are stretched spherical surface coordinates, using the following equations:

$$
\begin{aligned}
x &= \left(\frac{\pi}{2} - \frac{\pi\psi}{180}\right) \cos\frac{\pi t}{12} \\
y &= \left(\frac{\pi}{2} - \frac{\pi\psi}{180}\right) \sin\frac{\pi t}{12}
\end{aligned}
\tag{3.5}
$$

The setup of the basis functions is on the $x - y$ plane. Following Nychka et al. (2015), the basis functions $\phi_j(x)$ are chosen as compactly supported Radial Basis Functions (RBF) φ, which are bell-shaped curves with a common width θ

$$
\phi_j(x) = \varphi(\|x - u_j\|/\theta)
$$

where $u_j(1 \leq j \leq m)$ is the center of RBFs. Typically, u_j is equidistant, that is, $\Delta u = u_j - u_{j-1}$ is a constant that represents the grid size (also referred to as the model resolution), and all $\{u_j, 1 \leq j \leq m\}$ of the same θ form a regular grid map covering the whole domain, which consists of one level of RBFs. The number of RBFs at this level m is approximately related to the grid size Δu by the reciprocal rule $m\Delta u^2 = $ domain size. Since the latitudinal direction is directly mapped, but the longitudinal direction is scaled in this coordinate, we refer the resolution of $1°$ (without differentiating between the latitudinal and longitudinal directions) to the

spacing of RBFs by $\pi/180$. Take the auroral modeling, for example, in the high-latitude region over 50° MLAT, the domain size is $(40° \times 2)^2 = 6400$. In terms of 1° modeling resolution ($\Delta u = 1°$, basis functions separated by 1°), approximately 6400 basis functions are used ($m = 6400$).

For the multiresolution fitting, RBFs of different Δu and θ can be combined into a large basis set (see Figure 3.3 for a three-level setup of RBFs; Figure 3.3a shows a 1-D case, and Figure 3.3b shows a 2-D case). In this sense, we relabel m with m_l, Δu with Δu_l, and θ with θ_l with l representing the number of levels. These parameters can take different values across different levels, which leads to different resolutions. The multilevel reconstruction of the spatial variation field is then written as

$$g(x) = \sum_{l=1}^{L} \sum_{j=1}^{m_l} c_{j,l} \varphi(\|x - u_j\|/\theta_l)$$

where $c_{j,l}$ is the coefficient of the jth RBF at l^{th} level. L is the total number of levels, a critical parameter in describing the multiresolution properties of the basis functions. For typical usage, θ_l is set as a fixed multiple of Δu_l (greater than one) to allow for overlapping RBFs at every point. Both L and m_l (equivalently, Δu_l) can be adjusted to obtain basis function maps of different scales. Higher-resolution RBFs have more free parameters (c_j) to simulate the details of the input data. They are expected to provide more small-scale structures of aurora than the lower-resolution RBFs.

In our auroral modeling setup, the modeling domain is a 2-D square over the high-latitude region. To have m_l basis functions for the 2-D map in the lth level, we distribute $N_l = \sqrt{m_l}$ basis functions on each side. We double N_l every time as we go from a coarse to a fine level, so the overall number of basis functions (m_l) approximately increases by a factor of 4. By recalling that the grid size and the number of basis functions are related by the reciprocal rule, the grid size is approximately
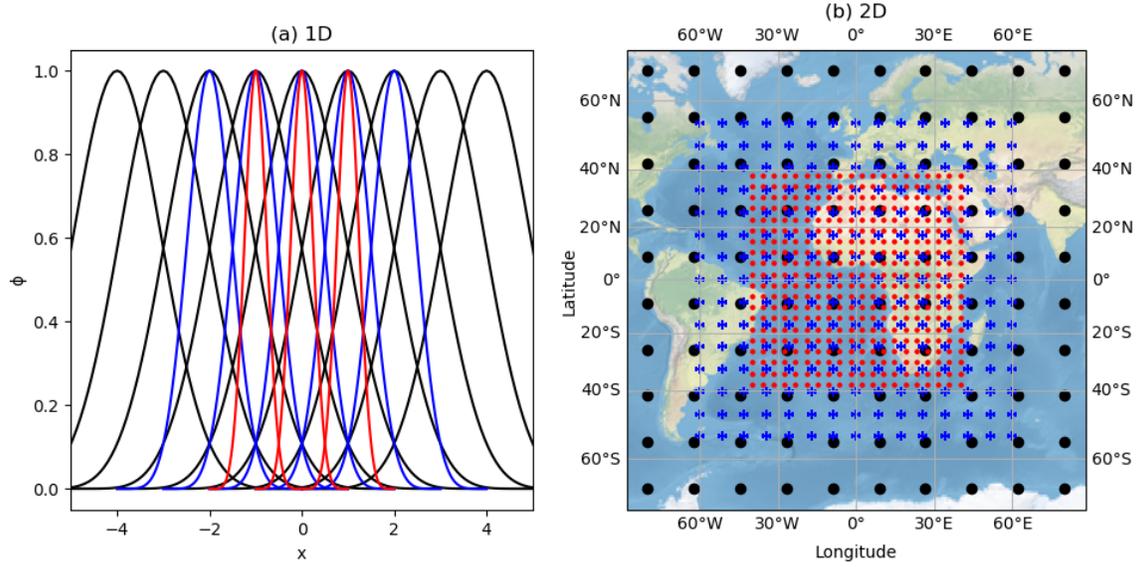
Figure 3.3: Demonstration of the multilevel setup of basis functions. (a) 1-D setup of basis functions with three levels (b) 2-D setup of basis functions with three levels, black, blue, and red colors represent three levels from the coarsest to the finest grids.

halved with increasing L, and the fitting resolution is doubled. From Nyquist's theorem, we can take the smallest resolvable scale of our model approximately as double the grid size ($2\Delta u_l$). This study selects the number of levels (L) and the number of basis functions on each side at the coarsest level (N_c) as the fundamental parameters to control the modeling resolution. In this sense, we define low-resolution modeling as $L = 1, N_c = 13(N_1 = 13)$, which means the fitting resolution is 6.2° (modeling domain is 80°), and the resolvable scale is 12.3°. Similarly, the medium resolution is defined as $L = 2, N_c = 15$ ($N_{1,2} = (15, 30)$, 30 basis functions on each side at the finest level, the fitting resolution is 2.6°, the resolvable scale is 5.3°) and the high resolution is $L = 3, N_c = 25$ ($N_{1,2,3} = (25, 50, 100)$, 100 basis functions on each side at the finest level, the fitting resolution is 0.8°, the resolvable scale is 1.6°).

Even though the number of RBFs roughly quadruples if we increase L, the overall computations do not grow exponentially with the number of levels. Since we

formulate the precision matrix $\boldsymbol{Q}$ at each level to be sparse, the overall precision matrix is still sparse (see Appendix A for details). Therefore, the matrix calculation does not increase cubically with the total number of the matrix elements (e.g., Gaussian elimination) but only linearly with the nonzero elements. Thus, the increase in levels of basis functions leads to a moderate increase in the whole computation.

3.4 Procedures and Results of Auroral Data Assimilation

Before feeding SSUSI and THEMIS observations and the empirical model into Lattice Kriging, two preprocessing steps are implemented. First, even after we collect 20 min of SSUSI data to form the satellite binned map (Section 3.2), its spatial coverage is still limited. Therefore, an interpolated satellite map using the data 1 h before and 1 h after the modeling time is generated to enlarge the spatial coverage and used as the fourth data source (details in Section 3.4.1.1). Second, we assign larger weights to higher-fidelity data (i.e., SSUSI and THEMIS observations) and smaller ones to lower-fidelity data (the interpolated satellite data and empirical model) such that the former two data sources dominate the fitting results. At the same time, the latter two only play roles in the regions where observations are missing. The weighting in Lattice Kriging is realized by attributing different sampling ratios to different data sources. The low-fidelity data are downsampled to decrease their sampling rates and, equivalently, the weights in the fitting (details in Section 3.4.1.2). After the preprocessing, we use Lattice Kriging to synthesize all four data sources to generate auroral maps at all locations (Section 3.4.2) and produce the intermediate result. Due to the smoothing effect inherent in the fitting procedure, Lattice Kriging causes spreading.

It introduces nonzero values in regions with no aurora, such as the polar cap, and tends to smear out the auroral boundaries. This solicits a postprocessing weighting method (KNN) for mitigation (details in Section 3.4.3). Figure 3.4 provides a flow chart of these procedures.
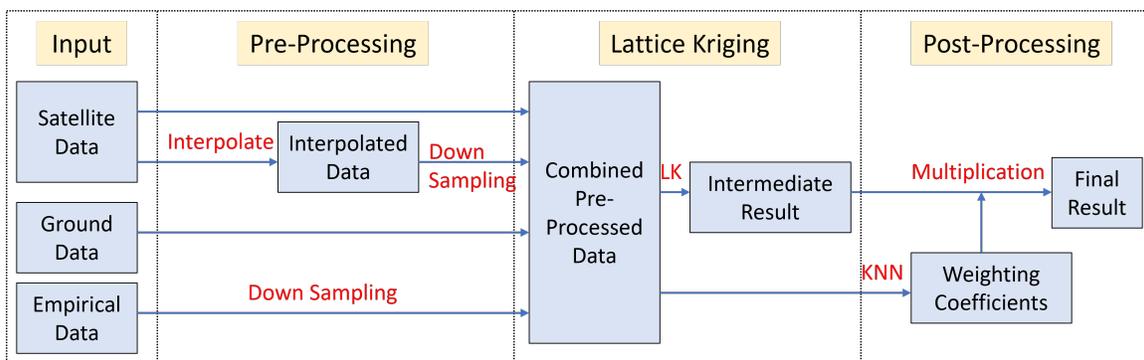
Figure 3.4: Flow chart of the auroral data assimilation model.

3.4.1 Data Preprocessing

3.4.1.1 Temporal Interpolation of Satellite Data

Figure 3.5 illustrates an example of the satellite data by 20 min binning centered around 11:40:00 (Figure 3.5d), and the satellite map after the linear interpolation with time (Figure 3.5c). The data used for the interpolation are collected within a 2 h window from 10:40:00 to 11:40:00 shown as the "1 h before" data in Figure 3.5a and from 11:40:00 to 12:40:00 shown as the "1 h after" data in Figure 3.5b, respectively. Compared to Figure 3.5d, the interpolated map (Figure 3.5c) shows similar results if the data being interpolated are within the 20 min window, such as the region around the dawn (6 MLT). The similarity originates from the proximity in time for the temporal interpolation. For the cases where the satellite data are available within the 2 h but not the 20 min window, the binning method would not show anything. At

81

the same time, the interpolation can fill up the aurora, such as in a significant portion of the dusk region where a few auroral arcs are seen (18 MLT in Figure 3.5c). Even though the assumption that the aurora should change linearly during this period does not necessarily represent the truth, the interpolated results (such as their magnitude) are still closer to reality than the empirical model. Note that compared with the relatively instantaneous observations (e.g., binned satellite and ground-based data), the interpolated data are downsampled (Section 3.4.1.2) to ensure that they do not override the 20 min binned data when they both exist in the same regions. A similar interpolation method is used in Wu et al. (2020), simulating TTEIL during the storm time better than using the empirical auroral drivers. In this study, the similarity and correlation between the auroral activities separated by over 2 h are thought to be weak, so the linear interpolation is conducted within the 2 h window.

3.4.1.2 Down-Sampling and Weight Adjustment

As discussed earlier, due to the different fidelities of the data sources (observation > satellite interpolation > empirical model), we attribute different weights to them by controlling the data sampling ratios. For simplicity, we refer to the satellite 20 min binned data as "satellite data" and the interpolated results as "satellite interpolation" or "satellite interpolated data". Sampling ratios of 1 meaning no downsampling are assigned to the satellite and ground-based data ($r_{sat} = r_{grd} = 1$). The ratios for the satellite interpolation and empirical model whose original spatial grids are the same as the satellite data are $r_{int} = 1/3$ and $r_{emp} = 1/20$, respectively. Considering that the spatial resolution of ground-based data is higher than that of the satellite (Section 3.2), the absolute sampling ratios are satellite : ground-based : satellite interpolation : empirical model $= 1 : 2 : 1/3 : 1/20$. If they overlap in the same region, their weights follow the sequence of ground-based data > satellite data > satellite interpolation >
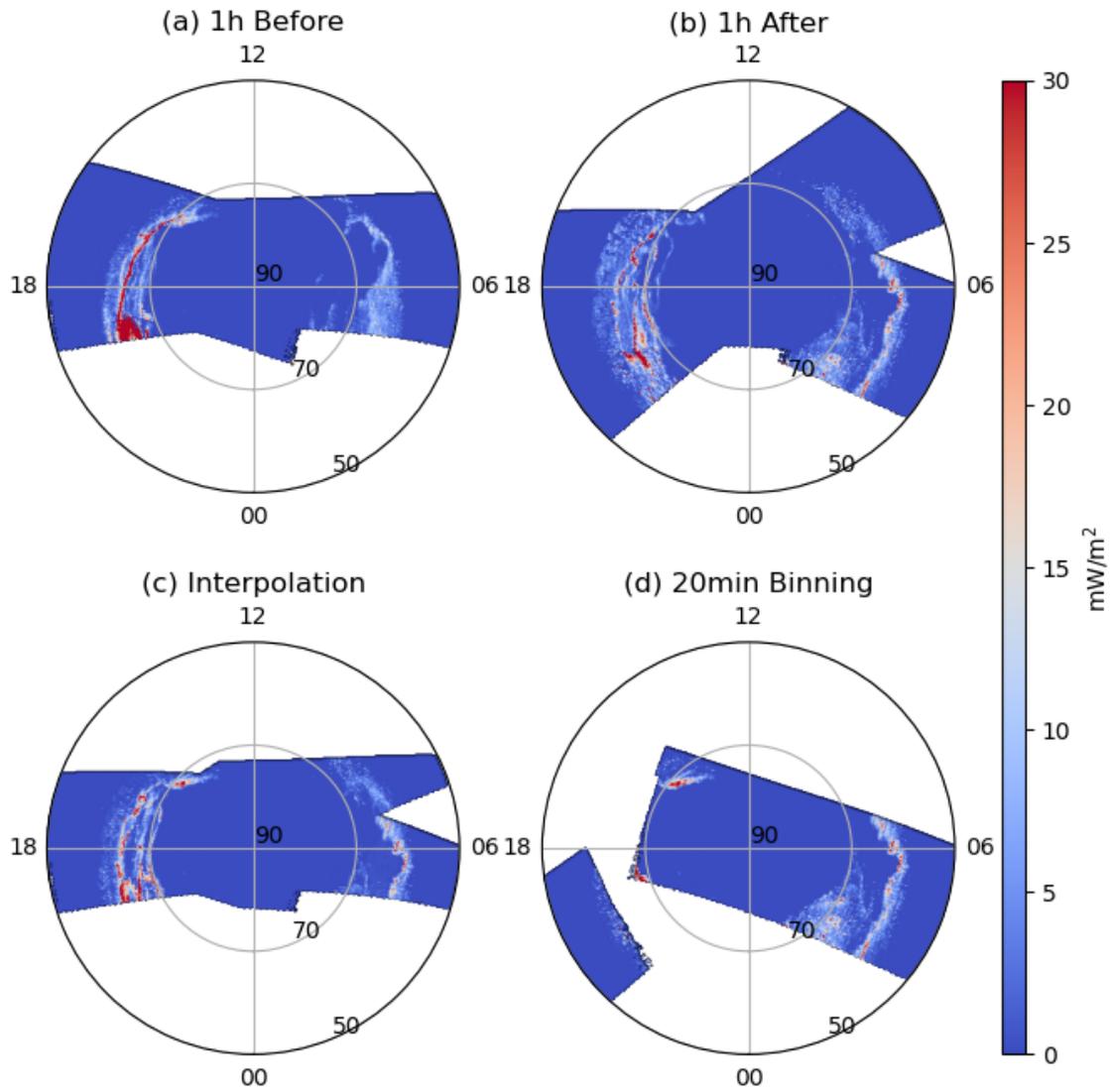
Figure 3.5: Temporal interpolation of satellite observations (a-b) SSUSI energy flux combining 1 h period of data before and after 11:40:00, respectively, into a snapshot, (c) temporal interpolated energy flux at 11:40:00, (d) 20 min binned energy fluxes around 11:40:00. Unit is $\mathrm{mW\,m^{-2}}$.

empirical model. The ratios are adjustable depending on the data quality and application purposes.

Figure 3.6 demonstrates the downsampling procedure at 11:50:00. Figures 3.6a-d are the original data, and Figures 3.6e-h are the data after downsampling. The satellite and ground observations are fully kept in this case following their original resolutions. The interpolation data are down-sampled to 1/3 of the original satellite grids (Figures 3.6b and 3.6f). Due to the lowest fidelity of the empirical model, it has the lowest data sampling density (Figure 3.6h). Therefore, its information is assimilated mainly in the regions where the other three data sources are unavailable.
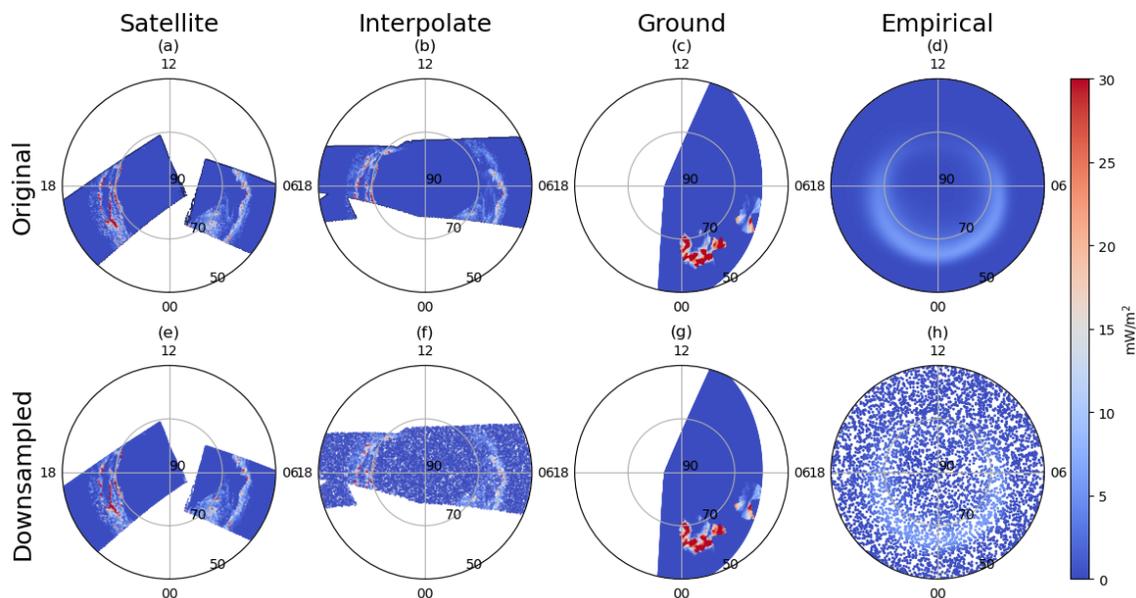
Figure 3.6: Downsampling and weight adjustment (a-d) auroral energy fluxes from different sources with original resolutions (e-h) scattered plots after the downsampling showing different sampling ratios and weightings. Unit is $mW\,m^{-2}$.

3.4.2 Data Assimilation Using Lattice Kriging

Feeding the preprocessed data (Figures 3.6e-h) as the inputs to Lattice Kriging ($\boldsymbol{y}$ in the formula of Section 3.3.1), we obtain the assimilation results in Figure 3.7

for 11:50:00, which corresponds to the "Intermediate Results" in Figure 3.4 (before the KNN postprocessing method being applied). We adopt three levels ($L = 3$), and the numbers of basis functions from coarse to fine grids are $N_{1,2,3} = (25, 50, 100)$, respectively. For the dawn and dusk sectors, the assimilation result mainly resembles the SSUSI observations (Figure 3.6e); THEMIS data (Figure 3.6g) contribute to the midnight sector. For the pre-midnight sector (21 MLT to 24 MLT), where no observations are available, the assimilated aurora follows the empirical model. Figure 3.7a shows the prediction of the conditional mean (Equation (3.3) in Section 3.3.1). The mesoscale structures, including the auroral arcs in SSUSI data and the hot spots spreading in the midnight sector measured by THEMIS ASIs, are largely maintained.
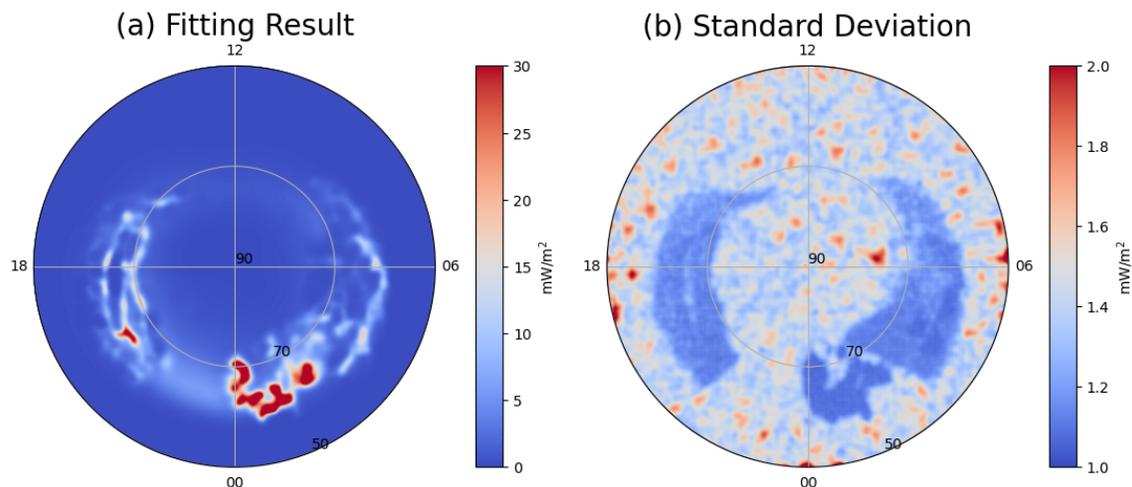
Figure 3.7: Lattice Kriging fitted map of auroral energy fluxes (a) prediction of mean and (b) prediction of SD. Unit is $\mathrm{mW\,m^{-2}}$.

The assimilated energy fluxes' uncertainty depends on the data sources' uncertainties. Based on the error assessment of the historical data, the uncertainty of SSUSI data can be taken as 15 % of the measurement, and the uncertainty of THEMIS observation can be taken as 20 % of the data itself (Gabrielse et al., 2021). Since the interpolated SSUSI data has lower fidelity, an uncertainty of 30 % is assigned to the

interpolated result. The uncertainty of the empirical model is chosen to be 100 % of its value as a proxy since no related information is available yet. These uncertainty terms are used as inputs to ϵ in Equation (3.2) (Section 3.3.1). The SD/fitting uncertainty is then calculated following Equation (3.4) and shown in Figure 3.7b. The uncertainties are considerably smaller than the predictions of the means and smaller in the regions with observations than those without observations, reflecting the data constraints.

Despite the similarity between the assimilated auroral map and the input data (Figures 3.6 and 3.7a), the former appears blurry, and small energy fluxes spread into the polar cap and sub auroral regions where the observations show no aurora in the input data. Even though delicate structures such as auroral arcs are retained, the peak values in the assimilated map are also lower than the actual observations. In other words, the Lattice Kriging model introduces a smoothing effect, which causes leakage to the regions without aurora and reduced auroral peaks. A possible explanation is that Gaussian process models (including Lattice Kriging) use a distance-weighted mean strategy to attribute contributions from input data at different locations. In addition, the specific covariance structure used in the model predicts the variances at two nearby locations with similar magnitudes. Therefore, Gaussian process models rarely predict extraordinarily high or low values, making overall spatial predictions smoother than the data. Also, the smooth spatial structures of the basis functions tend to create a smooth representation of the spatial process.

3.4.3 Data Postprocessing

To suppress such smoothing effects, a postprocessing step relying on the KNN algorithm is applied. The same algorithm is used in Syrjäsuo and Donovan (2002)

and Syrjäsuo and Donovan (2004) to classify different types of aurora from ASIs automatically. We use this algorithm to identify the likelihood of having aurora for each location and eliminate the low-likelihood points. KNN is a common classification method widely used in machine learning. It relies on the assumption that the points in the same category share similar features and lies close in the feature space. Therefore, a straightforward way to divide points into different categories is by grouping them in the feature space justified by distance. Given a set of labeled training data, we calculate the distance of a new point to all training data points and pick up the k nearest points. We identify the category of these k points in the feature space, and the new point belongs to the category with the most members.

Based on our data set, we define the auroral activity with energy flux higher than $2\,\mathrm{mW\,m^{-2}}$ as significant and set it as 1 in the feature space. Otherwise, it is insignificant (0). The threshold of $2\,\mathrm{mW\,m^{-2}}$ is chosen based on trial and error. It is suitable to identify substantial auroral activity, avoid contamination from low fidelity data (typically on the order of $0.2\,\mathrm{mW\,m^{-2}}$), and effectively maintain the fine structure. The preprocessed data (Figures 3.6e-h) with their significant/insignificant labels (1/0) are used as the training data for KNN. The k-nearest points to the training data are identified for each location on the fitted map. Assuming the number of data labeled as 1 is n_1, then a ratio of n_1/k is calculated, which represents the percentage of the k-nearest points falling into the category of significant aurora. This ratio is used as the weighting coefficient for this location. Doing so forms a coefficient matrix with the same dimension as the intermediate result. Their multiplication leads to a weighting process producing the final results of the auroral assimilation (Figure 3.4).

In Figure 3.8, we display the post-processing results with $k = 10$ at 11:50:00. The weighting coefficients from KNN are shown in Figure 3.8a. The intermediate

87

results from Lattice Kriging are shown in Figure 3.8b (same as Figure 3.7a). The final assimilation by multiplying Figures 3.8a and 3.8b is given in Figure 3.8c, where we see the smearing of energy fluxes into the polar cap is largely suppressed. In the polar cap region where the preprocessed data indicate that there are no auroral activities, the k-nearest points all fall into the feature space of 0, thus the spreading values in the polar cap region are effectively removed by multiplying a KNN weighting coefficient of 0. This can also help remove the isolated points (the ambient areas show no aurora) that may be due to measurement noise. In the auroral region around midnight, THEMIS ASI observations indicate strong auroral activity, and KNN labels are mostly 1 so as the weighting coefficients, therefore, the fitting results in the auroral region are kept. In the dawn and dusk regions, the strengths of auroral activities vary, so the feature space consists of both 0 and 1, and the resulting weighting coefficients are between 0 and 1. The multiplication of the weighting coefficients and the intermediate results helps decrease the aurora if the ambient region does not show enough significant auroral activities. This process sharpens the auroral boundary and to some extent, corrects the smoothing effect caused by Lattice Kriging. The overall auroral structures become more comparable to real observations since the training data set in KNN relies on real observations. Using a larger k involves more points in a larger area to be weighted and introduces a smoother structure than a smaller k.

In Figure 3.9, we compare simple padding of the satellite and ground-based observations with the final auroral assimilation at 11:50:00. The padding results show an obvious discontinuity and sharp cut-off at the boundaries among the satellite data, ground-based data, and the regions without observations (Figure 3.9a), which largely disappear in Figure 3.9b. The data assimilation effectively removes the boundary discontinuity and combines different data sources more coherently than the padding. A trade-off for such coherence and continuity is the reduced peak magnitude of the au-
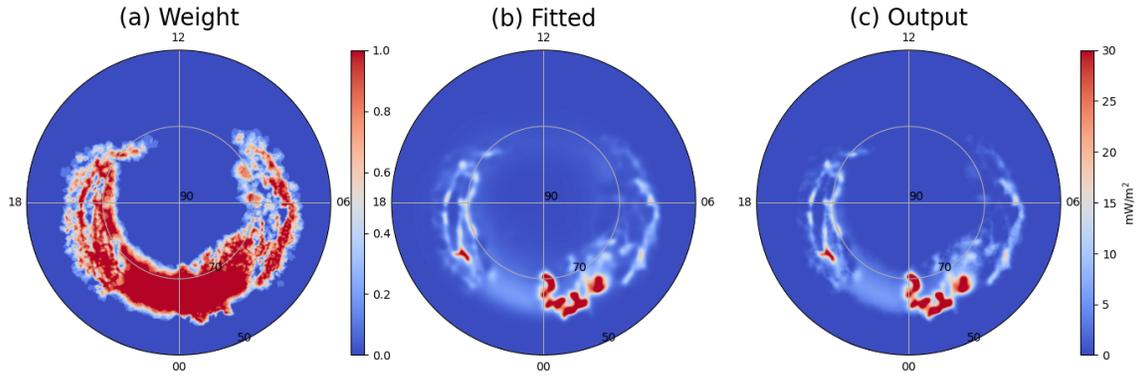
Figure 3.8: Postprocessing with KNN (a) the map of weighting coefficients from KNN, (b) the intermediate Lattice Kriging fitting results, (c) the final assimilation outputs by multiplying (a) and (b). (b) and (c) have unit $\mathrm{mW\,m^{-2}}$.

rora, which the KNN postprocessing step cannot correct. Nevertheless, the mesoscale aurora is largely retained including the auroral arcs, which significantly improves the reproduction of the real-time behavior of aurora compared to the empirical model.
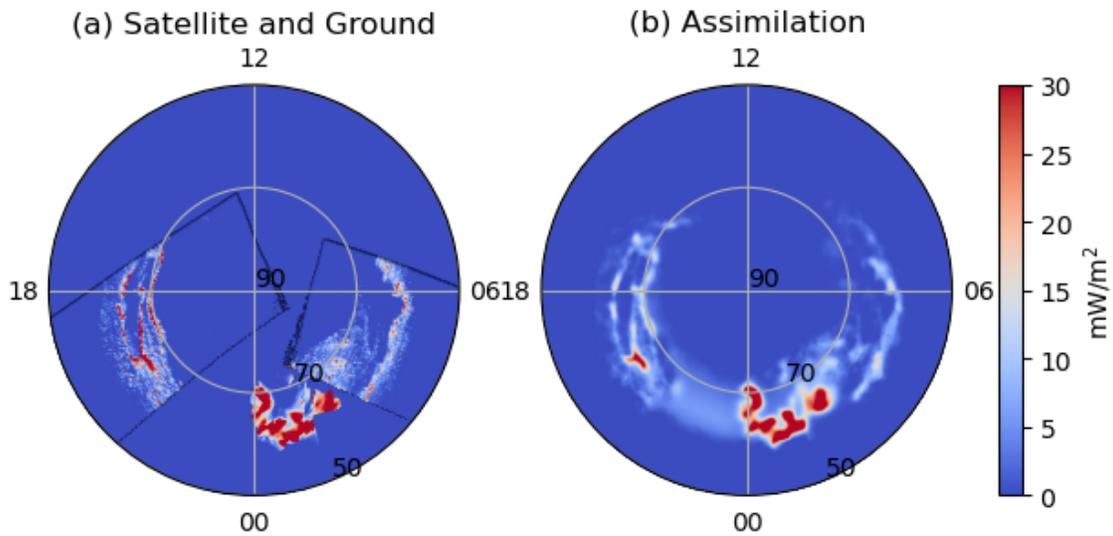
Figure 3.9: (a) The padding results combining satellite data and ground-based observations (b) auroral data assimilation (same as Figure 3.8c). Unit is $\mathrm{mW\,m^{-2}}$.

3.4.4   Auroral Assimilation With Different Scales

As mentioned in Section 3.3.2, the generated auroral maps of different scales can be obtained by tuning the number of fitting levels and the number of basis functions in each level (L and N). The resolution increases and the resolvable scale becomes smaller when we increase L and N. Figure 3.10 shows the assimilated auroral maps at three scales at 11:50:00. The parameters to generate these three auroral maps are $L = 1, N_1 = 13, k = 30$ for large scale; $L = 2, N_{1,2} = (15, 30), k = 20$ for medium scale; and $L = 3, N_{1,2,3} = (25, 50, 100), k = 10$ for small scale. Figures 3.10a-c show the auroral energy fluxes while Figures 3.10d-f show the mean energy maps from large to small scales and, equivalently, low to high resolutions. The assimilated aurora becomes more fine-structured from low to high resolutions, and the peak values increase. The auroral arcs in the dusk sector are distinct in the high-resolution results but absent in the low-resolution ones.

3.5   TIEGCM Simulations Driven by Auroral Assimilation Maps

To study how the data-assimilated drivers improve the simulation of I-T models and how different scales of aurora impact the I-T system, we run TIEGCM with different auroral maps. TIEGCM is a global 3-D numerical model that simulates the coupled thermosphere/ionosphere system from $97\,\mathrm{km}$ to $600\,\mathrm{km}$ altitude. It self-consistently solves the fully coupled nonlinear, hydrodynamic, thermodynamic, and continuity equations of the neutral gas, the ion and electron energy equations, the O^+ continuity equation and ion chemistry, and the neutral wind dynamo (Qian et al., 2014; Richmond et al., 1992). In the default setup, the high-latitude drivers such
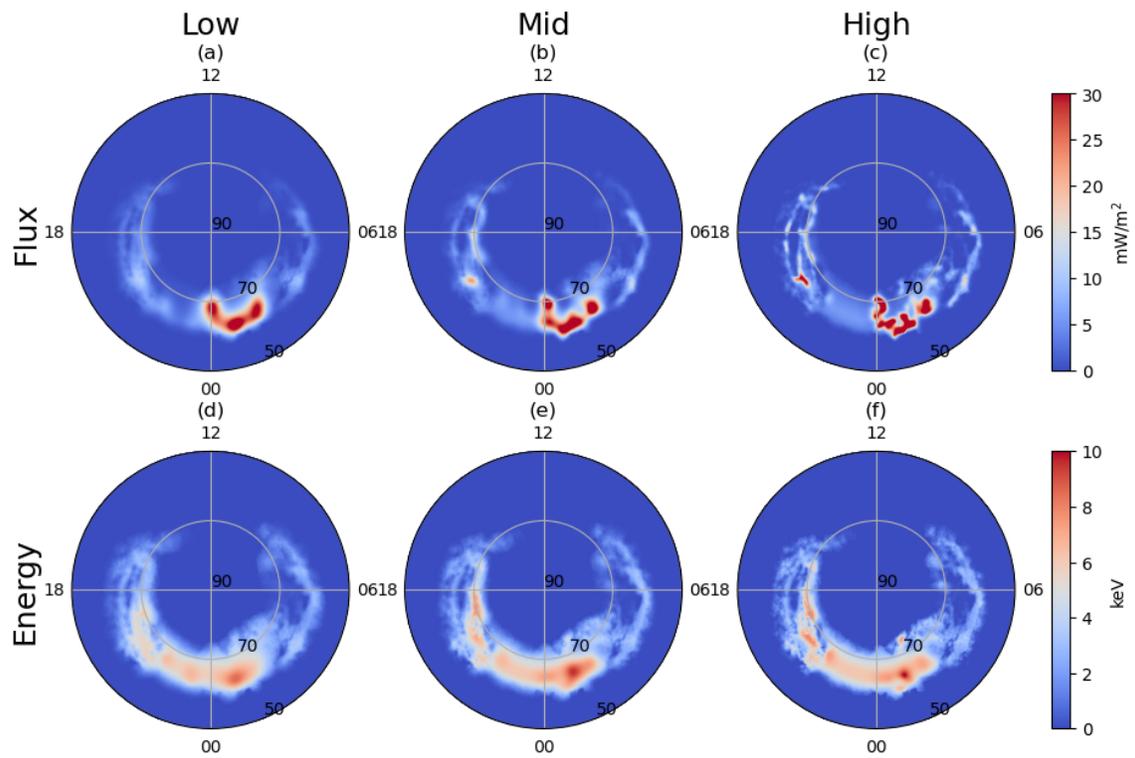
Figure 3.10: (a-c) Data assimilation of energy fluxes ($\mathrm{mW\,m^{-2}}$) from low to high resolutions. (d-f) The same for the mean energy (keV).

as aurora and electric fields (or potentials) are specified as empirical models (Heelis et al., 1982; Roble and Ridley, 1987; Weimer, 2005). In our TIEGCM runs, the time-varying SuperDARN electric potential pattern is derived from a Spherical Harmonics Fitting (SHF) of LOS ion velocities (Ruohoniemi and Baker, 1998) is used as a driver for electric fields. The assimilated electron precipitation pattern and the Zhang and Paxton (2008) model are used to specify auroral particle precipitation in TIEGCM. The spatial resolution of TIEGCM is $1.25° \times 1.25° \times 1/4$scale height in latitude $\times$ longitude $\times$ altitude (Dang et al., 2018, 2021). Realistic Kp and F10.7 are used in all simulations. The time step of the TIEGCM simulation is $10\,$s. Diagnostic outputs are saved every $5\,$min.

We perform three model runs; the only differences are the auroral energy flux and mean energy inputs. These three drivers are the empirical auroral model and the assimilated aurora at low and high resolutions (Figures 3.11a-c). In Run 1, the empirical model of the Zhang and Paxton (2008) is used as the auroral input to TIEGCM; In Runs 2 and 3, low- and high-resolution auroral patterns created in this study are used. The high-latitude electric field input in all runs is the SuperDARN potential pattern. Since the auroral particle precipitation affects the ionization rate and, therefore, the electron density, we show TECs from these three runs at 11:50:00 in Figures 3.11d-f and compare with GNSS observations (Figure 3.11g). GNSS TEC is measured by the trans-ionospheric propagation time difference between two radio frequencies from the GNSS satellite to the dual-frequency GNSS receiver. This propagation delay difference is directly proportional to the line integral of the electron density (Vierinen et al., 2016).

Compared with the TEC results driven by the empirical model (Figure 3.11d), the significant changes after we apply the auroral assimilation maps to drive the TIEGCM are the TEC enhancement (by a factor of 2) in the midnight sector where
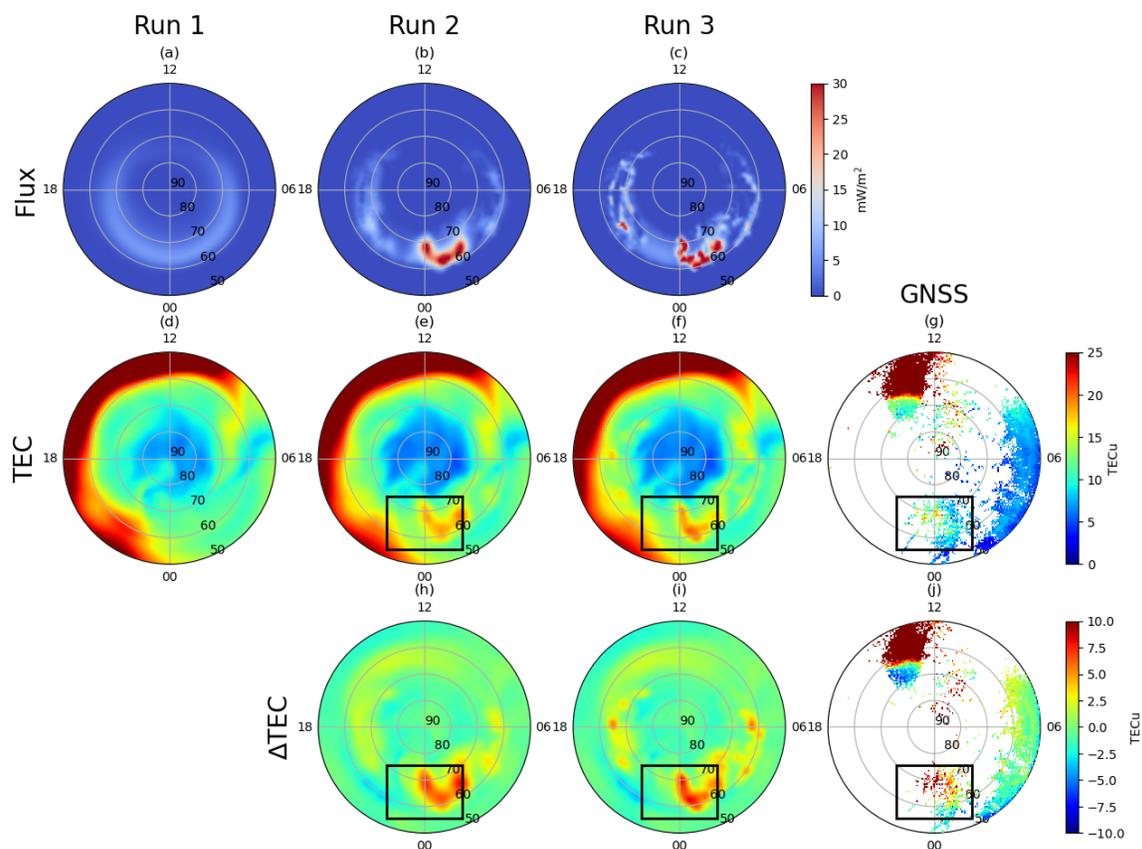
Figure 3.11: (a-c) Auroral energy fluxes (mW m^{-2}) from Zhang and Paxton (2008) empirical model, and assimilation maps at low and high resolutions at 11:50:00. All are projected into MLAT and MLT coordinates. (d-f) TIEGCM simulations of TECs (TECu) using (a-c) as the drivers for aurora. (g) Simultaneous GNSS TEC observations. (h-i) Differential TECs (TECu) by subtracting (d) from (e-f), respectively. (j) Differential TEC (TECu) from GNSS observations (details given in the text). Black rectangles highlight the regions with enhanced TECs due to the data assimilation and comparison with observations.

93

the SSUSI and THEMIS observations weight in (black rectangles in Figures 3.11e-f). The changes from low to high resolutions are noticeable in TEC as more mesoscale structures are seen in the high resolution. We also compare the storm-quiet time TEC differences in Figures 3.11h-j. From an observational perspective, the differential TEC is obtained by subtracting the TEC 24 h before the targeted storm-time (Figure 3.11j), which corresponds to 11:50:00 on 2014-02-19. From the modeling perspective, we use the Run 1 result, which does not involve data assimilation and only shows the large-scale pattern as a proxy for the quiet-time response. Figures 3.11h-i demonstrate the differential TECs from TIEGCM simulations by subtracting Figure 3.11d from Figures 3.11e and 3.11f. The regions and magnitudes of TEC enhancements from data assimilation agree with the observations, which means that the data assimilation can better simulate the mesoscale ionospheric responses to auroral precipitation. The model simulations with data assimilation capture the locations of strong TEC responses more precisely than the ones driven by the empirical model. The differential TECs also show comparable enhancements, which indicates the robustness of our auroral assimilation method and the resulting improvement.

It is noted that the original data resolutions of both DMSP SSUSI and THEMIS ASI data are much higher than the TIEGCM. Data assimilation can match the observation to a large extent but would still be limited by the I-T model, which incorporates it as an input. To further simulate small-scale processes and better use the data assimilation, the resolutions of the I-T models need to be improved. Moreover, the corresponding physics down to small scales must also be considered. Nevertheless, this work highlights the substantial changes from using the empirical model to data-assimilated aurora as drivers to simulate the responses of the I-T system, which is essential to understand better and predict the impacts of realistic and localized magnetospheric energy deposition.

3.6   Conclusions and Discussion

We introduce a multiresolution Gaussian process model (Lattice Kriging) to self-consistently synthesize various data sources (satellite, ground-based, and empirical models) for the auroral assimilation for the first time. This model assumes that the auroral activity follows a Gaussian process. It uses the available data to estimate the fitting coefficients of the basis functions within the Kriging theory framework. Then it uses these coefficients to project the estimation to the whole high-latitude region. The multilevel (or multiresolution) capability is fulfilled by distributing different levels of basis functions with different resolutions. Different scales of aurora can be assimilated, facilitating the study of multiscale processes like the aurora.

We introduce two preprocessing steps and one postprocessing step to customize the Lattice Kriging model to auroral assimilation. First, we interpolate the satellite data temporally to expand the spatial coverage at a particular time. The interpolated satellite data and empirical model (low-fidelity data) are then downsampled to decrease their weightings and ensure that the assimilation results are dominated by the satellite and ground-based observations (high-fidelity data) where the low- and high-fidelity data overlap. These data sources (satellite and ground-based observations, satellite interpolation, and empirical model) are fed into Lattice Kriging to obtain the intermediate results. Due to an inherent smoothing effect of the fitting procedure, which smears out auroral boundaries and introduces nonzero values in the regions with no aurora (such as polar cap), we generate a postprocessing weighting map using KNN trained by observations to mitigate these issues. The KNN weighting coefficient indicates how likely one location has significant auroral activity. These coefficients are multiplied by the intermediate fitting results to eliminate the isolated points likely caused by measurement noises and unrealistic spreading values produced

in the intermediate Lattice Kriging modeling. However, the reduced peak values of aurora due to the smoothing effect are difficult to compensate for. Compared with the simple padding of satellite and ground-based observations, the auroral assimilation model can effectively remove the discontinuity at the boundaries of different datasets.

We use the 2014-02-20 case (a moderate geomagnetic condition) as an example to demonstrate the assimilation procedures and generate the energy flux and mean energy maps with three different scales. The large-scale maps corresponding to the low-resolution fitting miss mesoscale structures such as auroral arcs, while the small-scale maps corresponding to the high-resolution fitting show mesoscale structures that more closely resemble observations. We then apply the assimilation maps of low and high resolutions to drive TIEGCM to study the impacts of different scales on TEC. TEC in the auroral region (especially the midnight sector) generally shows substantial enhancement that better matches observations after data assimilation due to increased auroral particle precipitation and ionization. High-resolution auroral precipitation maps also produce mesoscale structures of TEC. Overall, TIEGCM simulations highlight the importance of implementing realistic aurora as one of the magnetospheric drivers to model the mesoscale electrodynamics at high latitudes.

Despite the noticeable advantages of fusing actual data to simulate the mesoscale auroral structures, the current auroral assimilation model has the following limitations, which may need further improvements. In the data preprocessing step, we combine SSUSI data over 20 min to form a snapshot, then we interpolate over a 2 h period to expand the data coverage. One limitation of these steps is that the information on the development of aurora within that time interval is lost. This may lead to the distortion of the auroral oval if the aurora changes rapidly during the 20 min interval. For example, if a substorm onset occurred between the time when the dawn and dusk-side oval was observed, the dawn-side oval would appear expanded, while

96

the dusk side would appear contracted. While each side of the oval might appear as narrow features, they would come from entirely different auroral ovals. Combining observations from such a situation might lead to a double-edge structure, which is purely due to the binning of SSUSI data. It is difficult to mitigate this issue by the technique itself, and more data are needed to solve it fundamentally.

In spatial modeling of the aurora, simplified assumptions may not represent actual observations when specifying the covariance structure. First, the covariance matrix used here is derived from a Gaussian Markov random field (which assumes two locations are correlated only if adjacent, Nychka et al. (2015)). In the real world, however, even distant auroral regions can be correlated if the aurora in these regions is generated from a closely connected region in the magnetotail (Nishimura et al., 2020a). The covariance matrix must include an additional term indicating the medium-to-large range correlation to describe the realistic auroral characteristics (Cousins et al., 2013; Matsuo, 2020). Second, the auroral activity may not follow the Gaussian distribution as assumed in this study. Since the different high-latitude regions connect to different regions in the magnetosphere, the auroral distributions may not be the same. They may deviate from Gaussian distribution due to the pitch angle diffusion and other wave-particle interactions (Nishimura et al., 2020b). Therefore, the mathematical formulation may need to be modified based on a non-Gaussian process model. Still, Gaussian statistics have good properties for fast computation, such as the sparse matrix calculation as aforementioned, which satisfies as a starting point. The improvements in the covariance matrix and distribution type solicit statistical studies of the aurora, which is beyond the scope of this study. Third, the current methodology can efficiently combine various data sources and coherently conduct spatial fitting; thus, the boundary issue disappears. However, it is not an auroral prediction model and cannot be used to predict auroral activity for the next

97

time steps. The prediction of aurora may be achieved by the machine learning technique training a large amount of historical data. For our case, real-time observational data are still the key to driving models to produce realistic I-T responses. It is worth pointing out that there may be discrepancies between satellite and ground observations. For this event, the magnitudes from these two types of observations match to a large extent despite discrepancies in some small-scale structures. However, if these two data sources deviate, examining the data quality and downsampling to the one with lower fidelity is necessary.

It is worth mentioning that the Lattice Kriging modeling is not limited to scalar field assimilation. Wu and Lu (2022) have extended it to assimilate vector fields such as electric fields under the curl-free condition. They obtained results with much smaller errors than the global SHF using SuperDARN data. The fundamental principles are the same except for the assimilation of electric fields. We need to project the basis functions of electrical potential (scalar) to electric fields (vector) and then project them onto the LOS direction, along which the observations are made (SuperDARN measures LOS ion drifts). Such extended capability makes the Lattice Kriging modeling appealing for scalar assimilation, such as GNSS TEC measurements, and wind measurements, such as those from Ionospheric Connection Explorer (ICON) in the future.

Chapter 4

Developing a Multi-Resolution Data Assimilation Model for Electric Fields

We develop a new methodology for the multi-resolution assimilation of electric fields by extending a Gaussian process model (Lattice Kriging) used for scalar fields initially to vector fields. This method takes the background empirical model as "a priori" knowledge and fuses actual observations under the Gaussian process framework. The comparison of assimilated results under two different background models and three different resolutions suggests that (a) the new method significantly reduces fitting errors compared with the global SHF because it uses range-limited basis functions ideal for the local fitting and (b) the fitting resolution, determined by the number of basis functions, is adjustable, and higher resolution leads to smaller errors, indicating that more structures in the data are captured. We also test the sensitivity of the fitting results to the total amount of input data: (a) as the data amount increases, the fitting results deviate from the background model and become

more determined by data, and (b) the impacts of data can reach remote regions with no data available. The assimilation also better captures short-period variations in local PFISR measurements than the SHF and maintains a coherent pattern with the surrounding. The multi-resolution Lattice Kriging is examined via attributing basis functions into multiple levels with different resolutions (the fine level is located in the region with observations). Such multi-resolution fitting has the least error and shortest computation time, making the regional high-resolution modeling efficient. Our method can be modified to achieve the multi-resolution assimilation for other vector fields from unevenly distributed observations.

4.1   Introduction

Ionospheric plasma convection is primarily driven by the electrodynamic processes in the magnetosphere, which is controlled by the interaction between the magnetosphere and solar wind. Thus, ion convection is a crucial indicator of the ionospheric responses to geomagnetic variations. Ion motion can enhance, recede, and even reverse in reacting to different IMF and solar wind conditions. During disturbed periods, enhanced ion convection transports mid-latitude plasma into the polar cap, leading to tongues of ionization and patches, which can disrupt communication and navigation in the polar region (Buchau et al., 1983; Nishimura et al., 2021a; Weber et al., 1984). The radio backscatter technique has been widely used to measure ion motions. For instance, SuperDARN scans over azimuth sectors regularly (typically $2\,\mathrm{min}$) and measures LOS ion drifts therein. The LOS ion drift measurements provide ion convection information over high-latitude and mid-latitude regions. When different radars receive signals from different directions at the exact location, the vector ion drift can be directly retrieved (Bristow et al., 2016; Hanuise et al., 1993). However,

due to the limited coverage of SuperDARN radars, the rate of an overlapping FOV is relatively low; thus, the direct derivation of vector drifts from the LOS measurements is limited (Bristow et al., 1995; Sanchez et al., 1996).

The retrieval of the global convection pattern from LOS measurements using other techniques has long been investigated. Best-known techniques include Super-DARN SHF (Ruohoniemi and Baker, 1998) and AMIE (Richmond and Kamide, 1988). SuperDARN SHF derives vector ion drifts by minimizing the weighted squared errors between the LOS ion drifts and spherical harmonic expansions. The fitted patterns are widely used in quiet- and storm-time studies (Maimaiti et al., 2018; Zhang et al., 2020). AMIE uses spherical cap harmonics as the basis function to fit the LOS ion drifts. It provides more realistic high-latitude electric fields in the storm time than the empirical models (Hsu et al., 2021; Lu et al., 2020; Richmond, 1992). Both methods use a known background model to provide constraints in the fitting process. The degree of the fitting precision, referred to as "resolution", is controlled by the order of basis functions, which describes the number of harmonics along the longitudinal circle. The background model is weighted for both methods where no observations are available. More model points are sampled for higher orders, so the patterns are more heavily weighted by the background model (Bristow et al., 2016). Thus, for those methods using global basis functions (e.g., spherical harmonics and spherical cap harmonics), the fitting is constrained by limiting the sampling of the background model; thus, the fitting resolution cannot be too high. A typical choice is on the order of $10°$ in longitude and $2°$ in latitude (Lu, 2017; Matsuo, 2020).

Nevertheless, the M-I-T coupled system embraces a variety of critical medium-to-small-scale electrodynamic processes, which are below the resolution resolvable by SuperDARN SHF or AMIE methods. The electric fields at the polar cap and auroral region exhibit cross-scale power spectra from planetary scales down to 0.5 km

(Golovchanskaya and Kozelov, 2010a; Kozelov and Golovchanskaya, 2006), which lead to the deviation from the global large-scale two-cell ion convection pattern (Cousins and Shepherd, 2012a,b). Small-scale electric fields have often been observed and found to impact the energy budget during magnetic storms (Codrescu et al., 1995; Cosgrove and Codrescu, 2009). Wu et al. (2020) found that an accurate specification of the local electric fields varying in short temporal scales and satellite-observed auroras showing small-scale spatial variations is essential to reproduce the observed local temperature enhancement (500 K) and inversion layer in the E region (130 km). Sheng et al. (2020) found that using ground-based auroral imager observations characterized by mesoscale features better resolves the large-scale TADs than empirical auroral maps. These studies illustrate that better quantifying energy inputs fusing data information with regional scales significantly improves the simulation of ionospheric/thermospheric responses to geomagnetic storms.

Several methods have been proposed to accommodate the high-resolution data to use better the LOS ion drift measurements, typical of 1° (e.g., SuperDARN). For example, Amm et al. (2010) used Spherical Elementary Current Systems (SECS) as the basis functions and solved the coefficients of SECS using the divergence-free condition of ion drifts. This method does not rely on "a priori" information provided by background models. Bristow et al. (2016) proposed a local divergence-free fitting technique, which also relies on the divergence-free assumption of ion drifts. The relation between the vector field and its LOS component is imposed as an additional constraint of the system from which the vector ion drift is derived during the fitting process. Large-scale SHF results are imposed as boundary constraints and as "a priori" knowledge to the framework. Both methods produce vector ion drifts at higher resolutions than the typical global basis function fitting technique.

Here, we propose an alternative method of retrieving vector ion drifts or equiv-

alently electric fields, out of the LOS measurements using a multi-resolution Gaussian process model called Lattice Kriging (Nychka et al., 2015). This methodology has been used to analyze surface temperatures and make the prediction at regions without observations (Heaton et al., 2019; Wiens et al., 2020). However, the previous applications of the Lattice Kriging method are largely limited to the assimilation of scalar fields. In this paper, we develop an extension of this methodology and apply it to the assimilation of vector fields, that is, electric fields in our case. Such an extension assumes that the high-latitude electric field is curl-free (equivalent to the divergence-free constraint of ion drifts). The electric potential can be retrieved over the whole domain using LOS information on electric fields and certain background models. By adjusting the sparseness/fineness of the basis functions in multiple levels, the model can be used to study the multi-resolution structures of the electric field. Our testing results show that the fitted results follow the inputs locally and over the high-latitude region where SuperDARN observations are available. The method reduces the error of LOS electric fields, and the improvement is more significant when a set of higher-resolution basis functions is used.

The mathematical formulation is presented in Section 4.2, including a synthetic test to verify the formulation. Then, we apply this method to actual observations (i.e., SuperDARN and PFISR) to explore its applicational performance (Section 4.3). Conclusions are given in Section 4.4.

4.2   Methodology

4.2.1   Fundamental Formulation of Lattice Kriging for Data Assimilation

The full derivation is skipped since it has been discussed in Section 3.3.1.

Conventionally, electrical potentials are denoted as ϕ. But in Section 3.3.1, the basis functions also share the same symbol. In distinguishing between these two, the basis functions are re-denoted as R in this chapter. The specific form of RBF used in this study is

$$R(s) = \begin{cases} (1-s)^6 \left(35s^2 + 18s + 3\right)/3, 0 < s < 1 \\ 0, \text{otherwise} \end{cases}$$

where s is the normalized distance between observations and basis functions. Even though Nychka et al. (2015) use RBF as the basis function, the choices of basis functions are flexible as far as the function is range limited, which enables localized fitting. The basis functions can be easily modified to accommodate applicational needs.

4.2.2   Extension of Lattice Kriging to Assimilate Electric Fields

A straightforward approach to vector field modeling is to perform data assimilation separately for its components. However, as mentioned earlier, the measurements are only LOS components for electric fields, so such independent fitting is not feasible. Therefore, to derive the electric field from its LOS component, an additional equation that relates the two orthogonal components of the vector must be used as a

constraint. For electric fields, we can use the curl-free condition, which is a reasonable approximation in the I-T system (Eccles, 1998; Mayr and Harris, 1978)

$$\frac{\partial E_y}{\partial x} = \frac{\partial E_x}{\partial y}$$

This differential equation indicates that these two variables are not independent. Using this relation, only one unknown parameter must be derived despite the electric field having two components. By using this relation and projecting the vector electrical fields to the LOS direction along which the LOS drifts have measurements, the retrieval of E_x and E_y is possible and described as follows.

With the curl-free condition for electric fields, a common practice is to define a scalar potential ϕ satisfying

$$\boldsymbol{E} = -\nabla\phi$$

Then, the curl-free condition is automatically satisfied.

Since electric fields and potentials are related by partial derivatives, a natural choice to obtain the basis functions of electric fields is by taking the directional derivatives of potential fields. We choose the basis function of the potential field ϕ to be the RBF $R(\boldsymbol{x})$ following Nychka et al. (2015). We further define two functions $R_x(\boldsymbol{x})$ and $R_y(\boldsymbol{x})$, which are related to $R(\boldsymbol{x})$ by

$$R_x(\boldsymbol{x}) = -\frac{\partial R(\boldsymbol{x})}{\partial x}$$
$$R_y(\boldsymbol{x}) = -\frac{\partial R(\boldsymbol{x})}{\partial y}$$

Therefore, if the potential field is decomposed onto a set of $R_j(\boldsymbol{x})$ with coeffi-

cients c_j satisfying that

$$\phi(\boldsymbol{x}) = \sum_{j=1}^{m} c_j R_j(\boldsymbol{x})$$

Then, the components of the electric fields will follow the relation

$$
\begin{aligned}
E_x(\boldsymbol{x}) &= -\frac{\partial \phi(\boldsymbol{x})}{\partial x} = -\sum_{j=1}^{m} c_j \frac{\partial R_j(\boldsymbol{x})}{\partial x} = \sum_{j=1}^{m} c_j R_{x,j}(\boldsymbol{x}) \\
E_y(\boldsymbol{x}) &= -\frac{\partial \phi(\boldsymbol{x})}{\partial y} = -\sum_{j=1}^{m} c_j \frac{\partial R_j(\boldsymbol{x})}{\partial y} = \sum_{j=1}^{m} c_j R_{y,j}(\boldsymbol{x})
\end{aligned}
\tag{4.1}
$$

This suggests that $R_x(\boldsymbol{x})$ and $R_y(\boldsymbol{x})$ are basis functions of $E_x(\boldsymbol{x})$ and $E_y(\boldsymbol{x})$, respectively. Equation (4.1) can be rewritten into a vector form as

$$\boldsymbol{E}(\boldsymbol{x}) = -\nabla \phi(\boldsymbol{x}) = -\sum_{j=1}^{m} c_j \nabla R_j(\boldsymbol{x})$$

To further relate the electric field $\boldsymbol{E}(\boldsymbol{x})$ with its LOS component, we project every electric field observation onto its corresponding LOS direction $\boldsymbol{k}(\boldsymbol{x}_i)$

$$E_{\text{LOS}}(\boldsymbol{x}_i) = \boldsymbol{E}(\boldsymbol{x}_i) \cdot \boldsymbol{k}(\boldsymbol{x}_i) = -\sum_{j=1}^{m} c_j \nabla R_j(\boldsymbol{x}) \cdot \boldsymbol{k}(\boldsymbol{x}_i) = \sum_{j=1}^{m} c_j (-\nabla R_j(\boldsymbol{x}) \cdot \boldsymbol{k}(\boldsymbol{x}_i))$$

This is equivalent to defining a new set of observation-dependent basis functions

$$R_{\text{LOS},j}(\boldsymbol{x}_i) = -\nabla R_j(\boldsymbol{x}_i) \cdot \boldsymbol{k}(\boldsymbol{x}_i)$$

and projecting the LOS electric field onto the new basis set

$$E_{\text{LOS}}(\boldsymbol{x}_i) = \sum_{j=1}^{m} c_j R_{\text{LOS},j}(\boldsymbol{x}_i)$$

Gaussian process model for LOS electric field then becomes

$$\boldsymbol{y}_{\text{LOS}} = \boldsymbol{Z}_{\text{LOS}}d + \boldsymbol{R}_{\text{LOS}}\boldsymbol{c} + \boldsymbol{\epsilon}_{\text{LOS}} \tag{4.2}$$

$\boldsymbol{Z}_{\text{LOS}}$ is the projection of background model values onto the LOS direction. $\boldsymbol{\epsilon}_{\text{LOS}}$ is the measurement error of the LOS electric field. After obtaining BLUPs of $\hat{d}$ and $\hat{\boldsymbol{c}}$ using the observations of LOS electric fields ($\boldsymbol{y}_{\text{LOS}}$), the potential field can be reconstructed using Equation (3.3).

In summary, we reduce the electric field (vector) modeling problem into the fitting of its LOS component (scalar) assuming the curl-free condition and use the fitting information of $\hat{d}$ and $\hat{\boldsymbol{c}}$ from the LOS measurements to simulate electric potentials for all locations over the domain. This approach is a simplified formulation of Fan et al. (2018), which uses Helmholtz-Hodge decomposition to study a broader range of vector fields on a sphere.

4.2.3 Validation Test Using Synthetic Inputs

The electric field Lattice Kriging is validated using an artificial two-cell potential map with an arbitrary statistical background model in a 2-D plane. The test domain is $-4 < x < 4$ and $-2 < y < 2$. In order to apply Equation (4.2) to estimate the BLUPs of $\hat{d}$ and $\hat{\boldsymbol{c}}$, we need to input $\boldsymbol{y}_{\text{LOS}}$ (LOS electric fields), $\boldsymbol{Z}_{\text{LOS}}$ (LOS projection of the background model), and $\boldsymbol{\epsilon}_{\text{LOS}}$.

$\boldsymbol{y}_{\text{LOS}}$ is obtained from a reference potential ϕ, which is a combination of two cells with equal magnitudes centered symmetrically around the origin

$$\phi(x, y) = \frac{1}{1 + (x - 1)^2 + y^2} - \frac{1}{1 + (x + 1)^2 + y^2}$$

Then, the two electric field components are

$$E_x(x,y) = -\frac{\partial \phi(x,y)}{\partial x} = 2\left(\frac{x-1}{[1+(x-1)^2+y^2]^2} - \frac{x+1}{[1+(x+1)^2+y^2]^2}\right)$$
$$E_y(x,y) = -\frac{\partial \phi(x,y)}{\partial y} = 2\left(\frac{y}{[1+(x-1)^2+y^2]^2} - \frac{y}{[1+(x+1)^2+y^2]^2}\right)$$

The locations of inputs into the fitting model are taken randomly within the whole test domain. The azimuth angle θ_i at each location is randomly chosen from 0 to 2π. Then, the LOS direction is

$$\boldsymbol{k}(x_i, y_i) = \hat{\boldsymbol{i}} \cos \theta_i + \hat{\boldsymbol{j}} \cos \theta_i$$

The projection of the electric field in the LOS direction is

$$E_{\mathrm{LOS}}(x_i, y_i) = \boldsymbol{E}(x_i, y_i) \cdot \boldsymbol{k}(x_i, y_i)$$

Such LOS electric fields are fed into the model as $\boldsymbol{y}_{\mathrm{LOS}}$. These chosen LOS electric fields in real applications correspond to the spatially scattered LOS observations. All error terms ($\boldsymbol{\epsilon}_{\mathrm{LOS}}$) are taken as identities for simplicity.

In this validation test, the background potential model is taken as a linear function in the x direction

$$Z_\phi(x,y) = x$$

Then, the background LOS electric field model ($\boldsymbol{Z}_{\mathrm{LOS}}$) is

$$Z_{\mathrm{LOS}} = -\nabla Z_\phi(x_i, y_i) \cdot \boldsymbol{k}(x_i, y_i) = -\hat{\boldsymbol{i}} \cdot (\hat{\boldsymbol{i}} \cos \theta_i + \hat{\boldsymbol{j}} \cos \theta_i) = -\cos \theta_i$$

Following Equation (4.2), we obtain BLUPs of the model parameters $\hat{d}$ and $\hat{\boldsymbol{c}}$. Then we

simulate potential over the whole domain out of the obtained $\hat{d}$ and $\hat{c}$ using Equation (3.3). The simulated potential ϕ' is then compared with the reference potential ϕ (Figure 4.1).
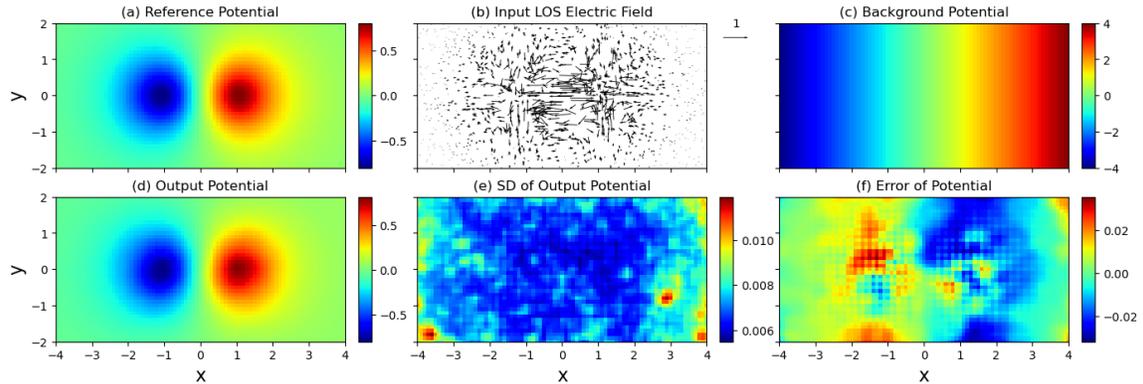
Figure 4.1: (a) Reference electric potential; (b) model inputs of LOS electric fields at selected locations; (c) background potential model; (d) fitted potential; (e) SD of fitted potential; (f) errors of fitted potential, (d-a). Units are arbitrary.

Figure 4.1a shows the reference potential field over the whole domain. Figure 4.1b shows LOS electric fields chosen for the model fitting. The background potential model is shown in Figure 4.1c. The fitted potential is depicted in Figure 4.1d, with its uncertainty/SD shown in Figure 4.1e. Figure 4.1f is the fitting error, which is defined as the difference between output (Figure 4.1d) and reference (Figure 4.1a). Note that SD and fitting error are magnitudes smaller than the field. The minor errors and the agreement between the original (reference) and fitted potentials confirm the validity of extending Lattice Kriging to assimilate vector fields.

4.2.4 Multi-Resolution Modeling by Using Multiple Levels

The math derivation is skipped since it has been discussed in Section 3.3.2. Figure 4.2 gives another example of a multi-level setup of the basis functions.

Throughout the fitting process, the coefficients of every basis function for each
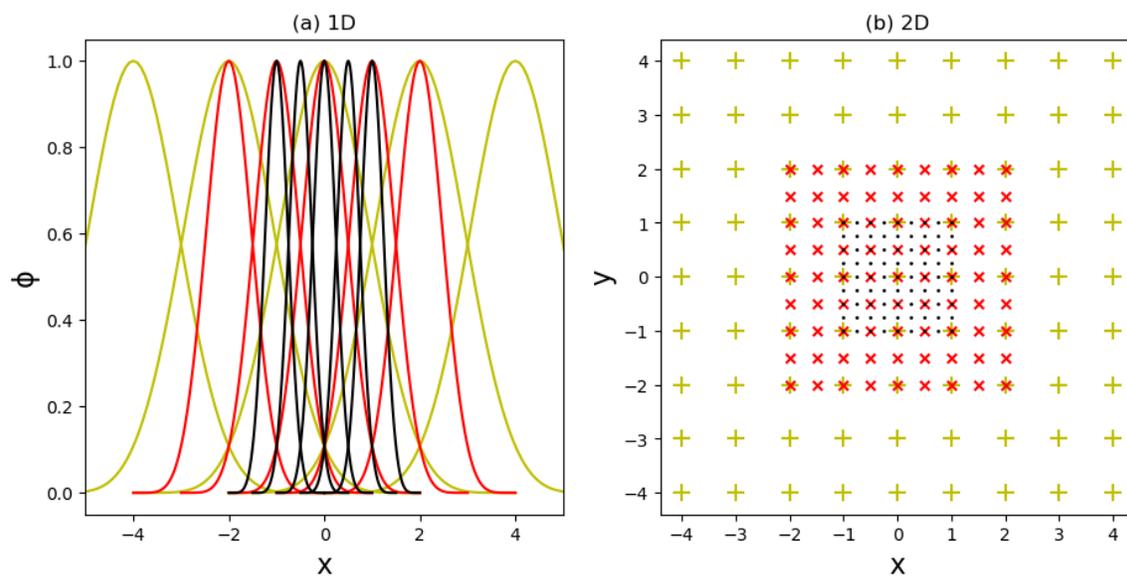
Figure 4.2: Examples of the configuration of multi-level basis functions in (a) 1-D and (b) 2-D. Units are arbitrary.

level are obtained, then used to reconstruct the field at this level. The final fitting result is a weighted mean of the reconstructed fields at all levels. The weighting factors depend on the application and can be adjusted toward small scales by attributing large weights on fine levels or toward large scales by putting large weights on coarse levels. In real applications, the finer grids are suggested to be located in the regions with more observations.

4.2.5   Data Preparation and Boundary Treatment for Electric Field Assimilation

The model setup for assimilating actual observations is similar to the validation test (Section 4.2.3) except that the coordinate system is different. The coordinate system used in simulating electric fields is a scaled plane coordinate system on the sphere's surface centered at the magnetic north pole. The coordinate transform from

MLAT and MLT to the model coordinates in the northern hemisphere is provided in Equation (3.5). After the data assimilation, the inverse coordinate transform is performed to obtain magnetic coordinates.

In the high-latitude ionosphere, the plasma motion is dominated by the drift motion

$$\boldsymbol{v} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} \tag{4.3}$$

So, all ion drift measurements used in this study are transformed to equivalent electric fields using

$$\boldsymbol{E} = -\boldsymbol{v} \times \boldsymbol{B} \tag{4.4}$$

For current development, the lower latitude boundary is set at 30° MLAT, where we assume the electric field vanishes. Since typical electric field observations are on the order of $10\,\mathrm{mV\,m^{-1}}$ and the errors are about 1/5-1/3 of the observations, the setup of the boundary condition consists of small values ($10^{-3}\,\mathrm{mV\,m^{-1}}$) with significant errors ($100\,\mathrm{mV\,m^{-1}}$). The small value term is used to force the fitting results to approach 0 at the boundary, while the significant error term is to minimize the boundary impact on the fitting of the internal field (poleward of 50° MLAT), which is the focus of this study.

4.3 Application of Lattice Kriging to Real Observations

The validation test demonstrates that Lattice Kriging can recover the synthetic electric fields (Section 4.2.3); now, we apply it to actual observations and examine its performance. We choose the St. Patrick's Day storm (2015-03-17) as an example

because it is the strongest geomagnetic storm in the solar cycle 24 and has reasonable SuperDARN data coverage. Figure 4.3 shows the geomagnetic indices during the two-day storm period. IMF B_z turns south at around 05:00:00 on 2015-03-17 (Figure 4.3a), marking the start of the geomagnetic storm. The southward B_z lasts for almost one day before it returns to neutral at around 05:00:00 on 2015-03-18, after which there are still minor IMF variations. IMF B_y is quite variable during the time. Figure 4.3b shows solar wind velocities and densities. The enhancement in the solar wind is evident during the southward B_z period. Figure 4.3c shows AE indices. Substantial AE variations indicate that the auroral activity is high during the time. The SYM-H index is shown in Figure 4.3d, from which we can tell that the significant substorm activity lasts until midnight in 2015-03-17 and the storm is still in the recovery phase till midnight in 2015-03-18.

During this time, there are 19 northern hemisphere SuperDARN radars operating. For reference, the names and locations are given in Table 4.1. The gridded ion velocity data are used, which have a spatial resolution of 0.5° and temporal resolution of 2 min. We will first use SuperDARN LOS ion drift data to perform data assimilation and analyze the errors using different statistical background models and resolutions in Section 4.3.1. Multi-level fitting is discussed in Section 4.3.2. Then, we add the measurements from PFISR into the data assimilation and compare the assimilated results locally (Section 4.3.3).

4.3.1 Assimilation of Electric Potential Using SuperDARN Data

We carry out six different settings for the fitting, which comes from the combination of two different background models (SuperDARN SHF and Weimer (2005)

112
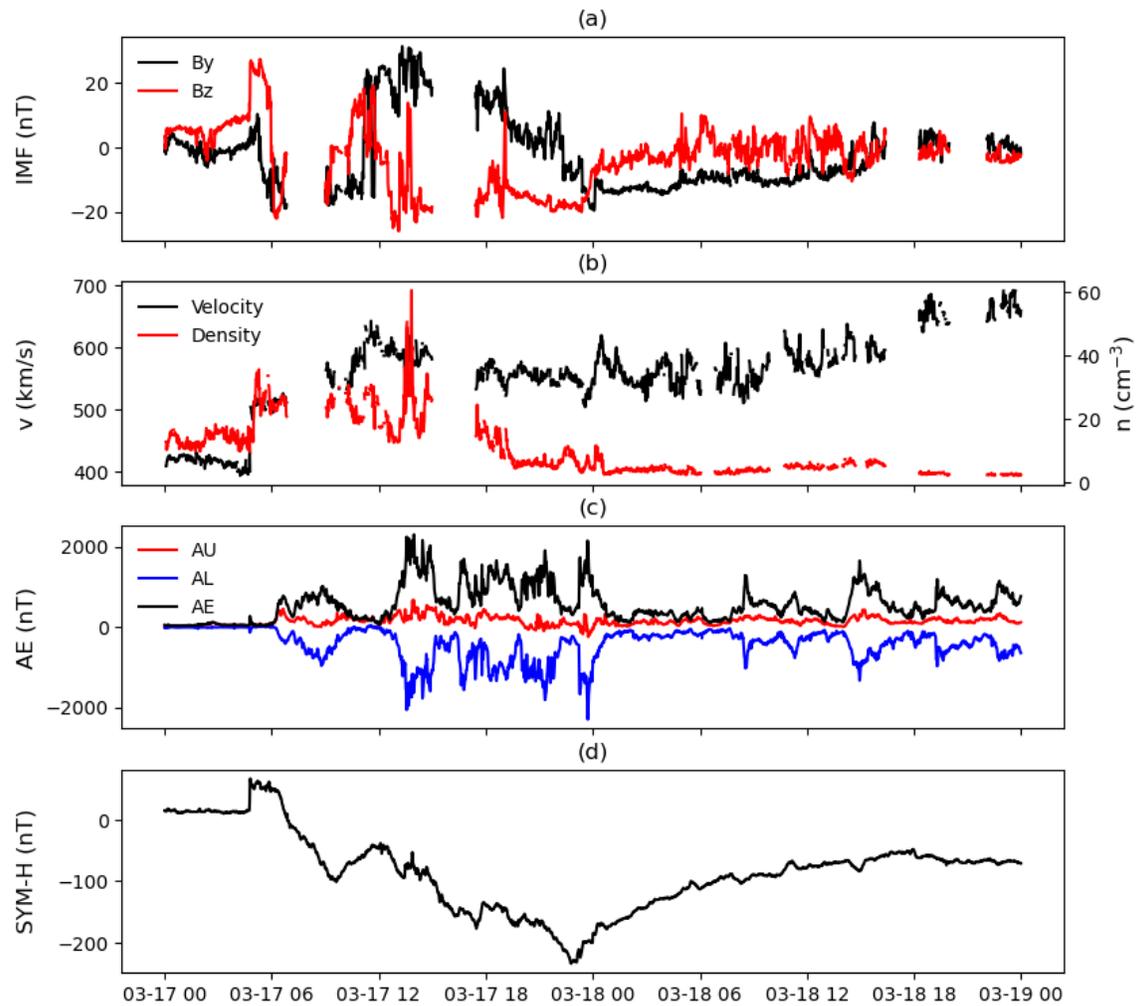
Figure 4.3: Geomagnetic indices on 2015-03-17 and 2015-03-18: (a) IMF B_y and B_z, (b) solar wind velocity and density, (c) AE indices, and (d) SYM-H.

Radar Name	GLAT	GLON	MLAT	MLON	Code	ID
Kapuskasing	49.39	−82.32	60.2	−8.3	kap	3
Saskatoon	52.16	−106.53	60.9	−43.8	sas	5
Prince George	53.98	−122.59	59.6	−64.3	pgr	6
Kodiak	57.62	−152.19	57.2	−94.9	kod	7
Pykkvibaer	63.77	−20.54	64.6	67.3	pyk	9
Hankasalmi	62.32	26.61	59.1	104.5	han	10
King Salmon	58.68	−156.65	57.5	−99.1	ksr	16
Wallops Island	37.93	−75.47	48.7	0.8	wai	32
Blackstone	37.10	−77.95	48.2	−2.7	bks	33
Hokkaido East	43.53	143.61	37.3	−144.9	hok	40
Hokkaido West	43.54	143.61	37.3	−144.9	hkw	41
Inuvik	68.414	−133.772	71.5	−85.1	inv	64
Rankin Inlet	62.82	−92.113	72.6	−26.4	rkn	65
Fort Hays West	38.86	−99.39	48.9	−32.2	fhw	204
Fort Hays East	38.86	−99.39	48.9	−32.2	fhe	205
Christmas Valley West	43.27	−120.36	49.5	−58.3	cvw	206
Christmas Valley East	43.27	−120.36	49.5	−58.3	cve	207
Adak Island West	51.88	−176.63	47.6	−113.0	adw	208
Adak Island East	51.88	−176.63	47.6	−113.0	ade	209

Table 4.1: A list of the radars that provide LOS ion drift observations on 2015-03-17. From left to right, the columns show the name, GLAT, GLON, MLAT, Magnetic Longitude (MLON), code, and station ID of the radar.

model) and three different resolutions (2°, 5°, and 8° in longitude and latitude). In all cases, the basis functions fill the whole domain with equal distances, a single-resolution assimilation for each setting. Since the fitting domain is set as a square with edge length $60° \times 2 = 120°$, and the basis functions are equally spaced, the number of basis functions used is then $(120/r + 1)^2$, where r is the resolution. For the three cases mentioned in this section, 256 basis functions are used for the 8° case, 625 basis functions are used for the 5° case, and 3721 basis functions are used for the 2° case. Figure 4.4a shows SuperDARN LOS ion drift measurements at 09:37:00 on 2015-03-17. In this study, we examine the impact of the background model on the data assimilation results by using SuperDARN SHF potential maps versus the Weimer model, shown in Figures 4.4b-c, respectively. Both SuperDARN SHF and Weimer potentials give two-cell patterns of similar magnitudes. In SuperDARN SHF potential, the peak magnitude is about 30 kV, while in the Weimer model, the peak magnitude is slightly larger at about 40 kV. The positive cell of SuperDARN SHF potential is located at a lower MLAT than the Weimer model. SuperDARN SHF potential shows more spatial variations, while the Weimer model is more uniform. The background potential maps are used to derive LOS electric fields and used as $\boldsymbol{Z}_{\text{LOS}}$ in Equation (4.2) for the assimilation procedure.

Figures 4.4b1-b3 and 4.4c1-c3 show assimilated electric potentials for the six different settings with SuperDARN SHF and Weimer as background models, respectively. Using the same background model, the fitted results of different resolutions are generally similar except with slightly different magnitudes. A cross-comparison of different background models (e.g., Figure 4.4b1 vs. 4.4c1, 4.4b2 vs. 4.4c2, and 4.4b3 vs. 4.4c3) shows that the fitted potential is smoother. The negative cell has a larger amplitude using the Weimer model than SuperDARN SHF, even though the general two-cell structure is similar.
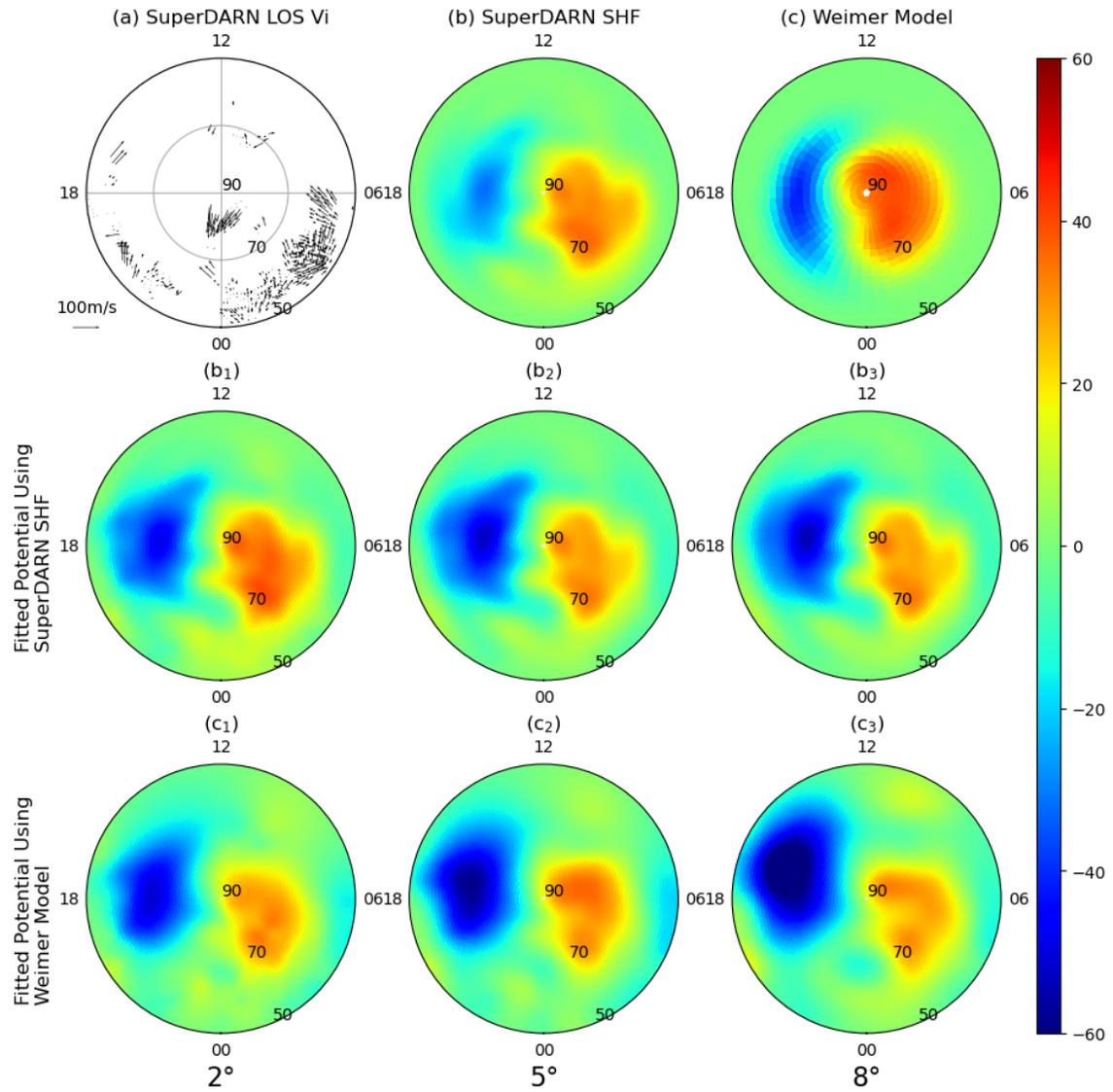
Figure 4.4: (a) SuperDARN LOS ion drift measurements, (b) SuperDARN SHF potential, (c) Weimer potential model, (b1-b3) fitted potentials using SuperDARN SHF potential as the background model, (c1-c3) fitted potentials using the Weimer model as the background model. Units are $mV\,m^{-1}$ for electric fields and kV for potentials.

To examine how the fitting results are related to the background model, we compare Figures 4.4b1-b3 with 4.4b (SuperDARN SHF) and Figures 4.4c1-c3 with 4.4c (Weimer model). Figures 4.4b1-b3 are similar to Figure 4.4b to a large extent because SuperDARN SHF is primarily fitted upon SuperDARN LOS ion drift measurements, which has already contained a large bulk of the observational information. One noticeable change is that the negative cell becomes more negative after applying Lattice Kriging. On the other hand, from the Weimer model to the fitting result using it as a background model (Figures 4.4c vs. 4.4c1-c3), significant differences are seen: (a) the comparable magnitudes of the positive and negatives cells are modified to a very negative cell (below $-80\,\mathrm{kV}$) and a mild positive cell ($30\,\mathrm{kV}$), (b) the unequal areas of the two cells are modified to be approximately equal, and (c) the negative cell is moved to a lower MLAT. Such modifications result from fusing the real-time observed LOS electric fields, which are not captured in the Weimer empirical model. The fitting process weighs more on the observations than the background model; therefore, the fitting results are dominated by observations whenever they are available.

From Figure 4.4c (original Weimer model) to 4.4c3 (fitted results using Weimer as background model), the potential pattern changes significantly even in regions where data coverage is sparse, for example, the post-noon sector from $12\,\mathrm{MLT}$ to $18\,\mathrm{MLT}$. We perform several tests to examine how the input data impact the fitting process by varying the total amount of data inputs. The fitting results in $8°$ using the Weimer background model are shown in Figure 4.5. In the extreme case that no data are inputted into the model, the fitting result (Figure 4.5c) is the same as the background model. In this case, no data are incorporated into the fitting process. The result converges to the background model with $d = 1$ and $\boldsymbol{c} = 0$ in Equation (3.2), which further validates our method. In Figure 4.5d where we only include

sparse observations from 12 MLT to 24 MLT, the fitting result still appears close to the background model. As more data are used (from Figure 4.5d to 4.5f), the fitting results gradually deviate from the background model and become more determined by the data. This test suggests that data can significantly influence the background patterns in the region away from the data. In other words, the impacts of data tend to be remote and global.
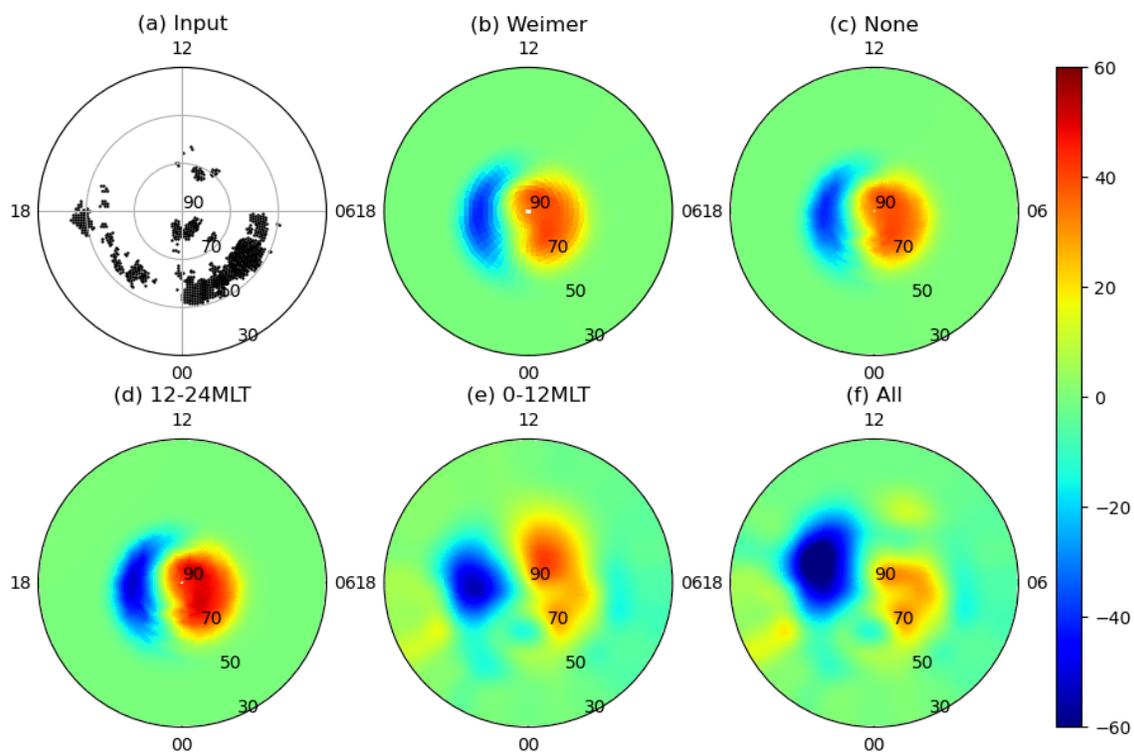
Figure 4.5: (a) Input data coverage at 09:37:00, (b) electric potentials from Weimer model, (c) fitting results with no input data, (d) fitting results with data only from 12 MLT to 24 MLT, (d) fitting results with data only from 0 MLT to 12 MLT, (e) fitting results with all available data.

To demonstrate how the fitting process introduces mesoscale variations in electric fields and ion drifts (converted from electric fields using Equation (4.3)), the fitting results in 2° using Weimer as the background model at 09:37:00 are shown in Figures 4.6a1-a2. Electric field and ion drift solely from the Weimer model are

shown in Figures 4.6b1-b2 for comparison. Apparent differences exist, and the fitted maps show more regional structures than the empirical Weimer model. For instance, around 70° MLAT and 0 MLT, the local divergence and convergence in electric fields (Figure 4.6a1) and local ion drift vortices (Figure 4.6a2) are only seen in the fitted maps. Mesoscale electric fields, which are missing in the empirical model, start to emerge when SuperDARN observations are included in the fitting process.

To better evaluate the fitting outcome and performance, we analyze the fitting error defined as Root-Mean-Square Error (RMSE) regarding the LOS electric field differences between fitted results and observations.

$$\epsilon = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(-\nabla\phi(\boldsymbol{x}_i)\cdot\boldsymbol{k}(\boldsymbol{x}_i) - E_{\mathrm{LOS}}(\boldsymbol{x}_i))^2}$$

where n is the total points of observations, $\phi(\boldsymbol{x}_i)$ is the fitted potential, and $-\nabla\phi(\boldsymbol{x}_i)\cdot\boldsymbol{k}(\boldsymbol{x}_i)$ gives its projection onto the fitted LOS electric field. $E_{\mathrm{LOS}}(\boldsymbol{x}_i)$ is the observed LOS electric field.

Figure 4.7 compares the fitting errors from all cases, and the errors solely from background models are also shown as a reference. Figures 4.7a1-b1 show fitting errors at 09:37:00. In Figure 4.7a1, the fitting errors decrease as resolutions increase. The errors are smaller than those from SuperDARN SHF, which means that the fitting process successfully reduces the errors and captures more information in the data as resolution increases. The fitting error of using 2° resolution decreases by 30 % compared with the background model. In Figure 4.7b1, the fitting errors decrease by more than half compared with the Weimer model, and more significant improvements are seen in the higher resolution assimilation. Cross-comparing Figure 4.7a1 with 4.7b1 in each resolution, the fitting errors are of a similar magnitude no matter which background model is invoked. This confirms that the fitting results are not sensitive

119
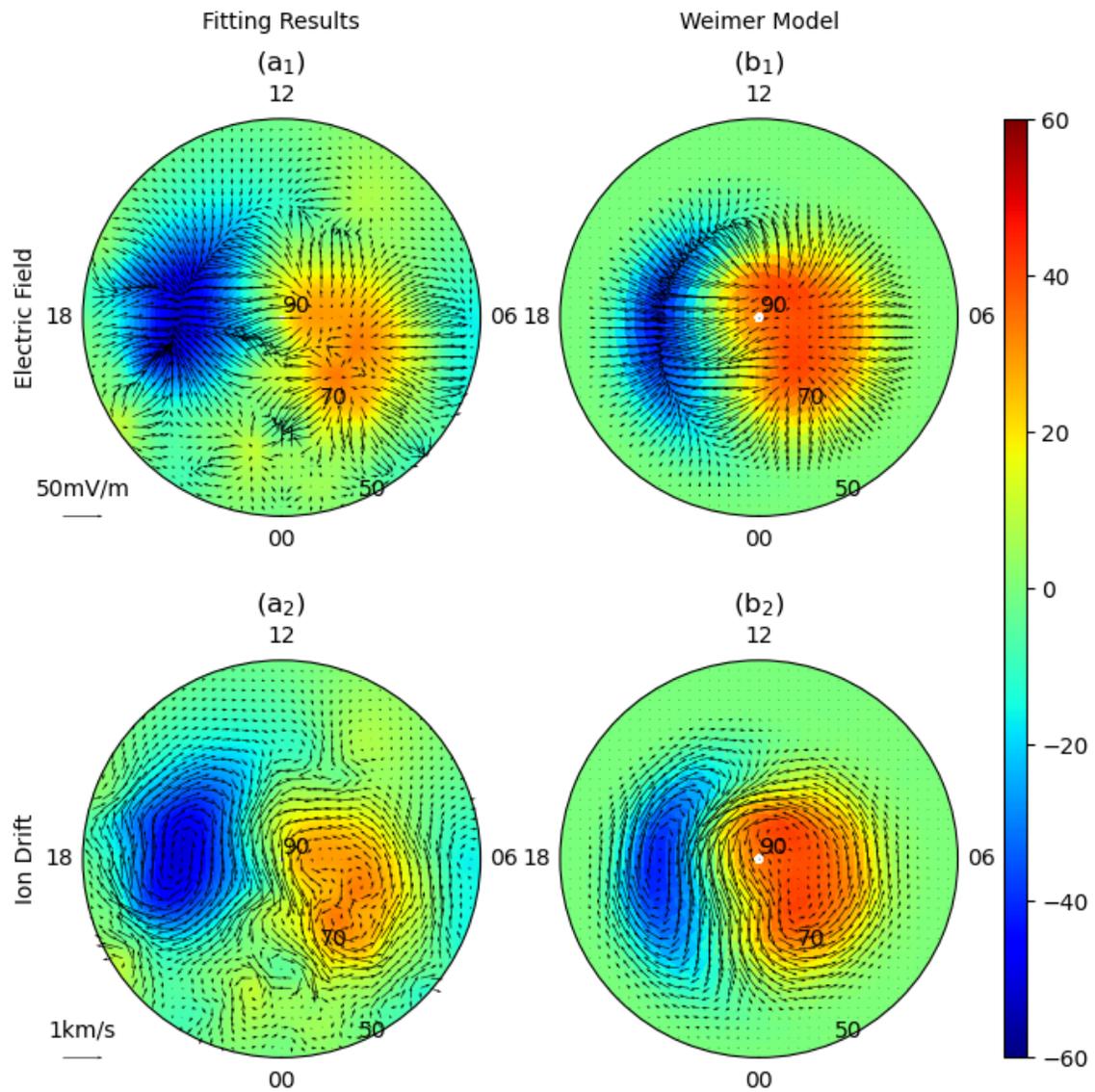
Figure 4.6: (a1) Fitted electric field, (b1) Weimer electric field, (a2) fitted ion drift, (b2) Weimer ion drift. The potential is overplotted in all subfigures. Arrows represent electric fields or ion drifts; the color contour represents the fitted potential. Units are mV m^{-1} for electric fields, km s^{-1} for ion drifts, and kV for potentials.

to the background model when sufficient observations are available.
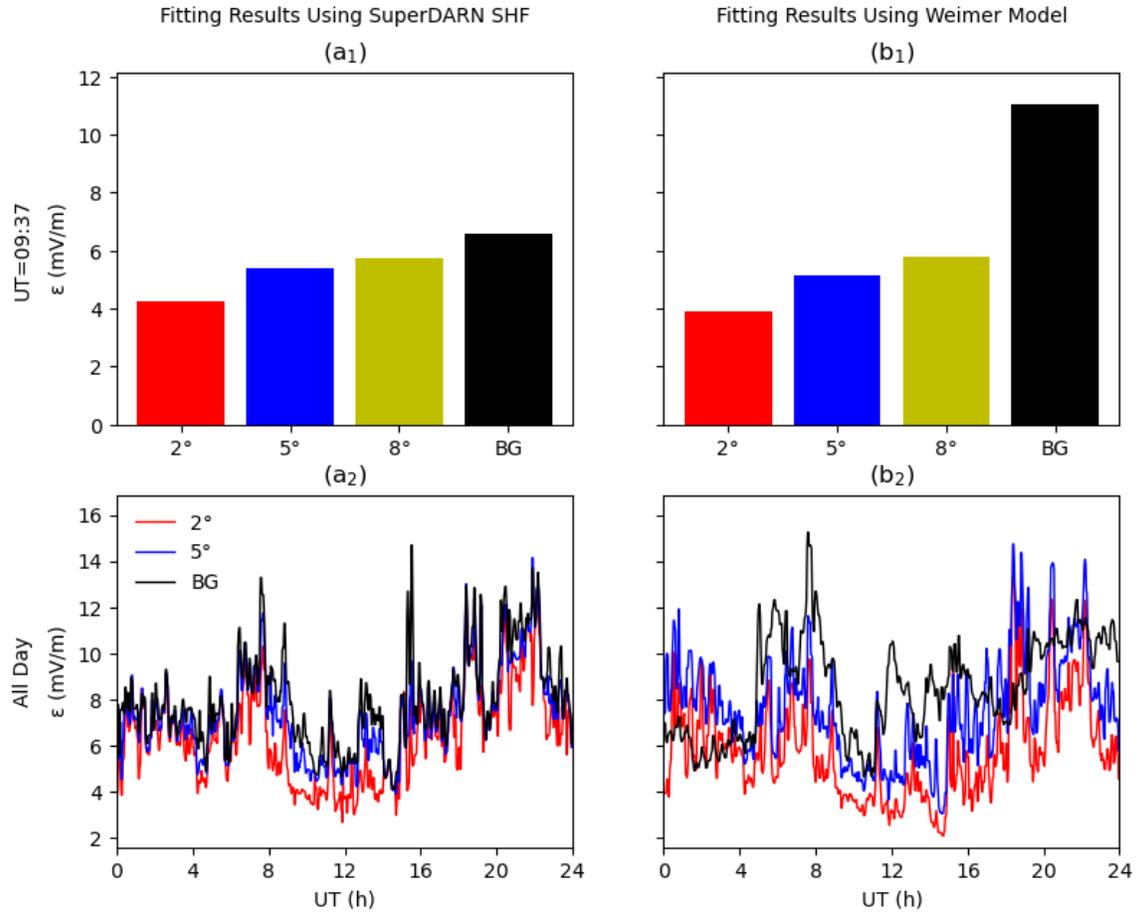
Figure 4.7: RMSE comparison of the fitting results: (a1) RMSE of the LOS electric fields using SuperDARN SHF as background model at 09:37:00; (b1) same as (a1) except for using Weimer model as the background model. (a2-b2) RMSE during the day using SuperDARN SHF and Weimer as background models, respectively. The errors of the Weimer model are divided by 2 in (b2) for display purposes. Red, blue, and yellow colors are for the fitting results under the resolutions of 2°, 5°, and 8°, respectively, while black is for the background model.

Figures 4.7a2-b2 show the fitting errors during the whole day (the results of 2° and 5° are given as examples). RMSE in the Weimer model is divided by 2 for displaying purposes in Figure 4.7b2. Lattice Kriging generally reduces the LOS electric field error when 2° and 5° of fitting resolutions are used. In Figure 4.7a2, the

121

fitting results using SuperDARN SHF decrease the error by more than 30 % during 07:00:00 and 15:00:00, while in Figure 4.7b2, the fitting results using Weimer background model decrease by more than half during most of the time. The fitting errors decrease with increasing fitting resolutions in both cases following Figures 4.7a1-b1. This implies that more data structures are captured using higher resolutions, and Lattice Kriging generally performs better than the fitting using global harmonics.

4.3.2 Multi-Resolution Assimilation of Electric Fields Using SuperDARN Data

To demonstrate the capability of multi-resolution data assimilation (Section 3.3.2), we set up two-level basis functions to perform the fitting and analyze the results (Figure 4.8a). The basis functions of the coarse level are separated by 5° and cover the whole domain; those of the fine level are separated by 2° and only cover half of the domain. Note that for the pure 2° and 5° cases, a single resolution (either 2° or 5°) is used throughout the whole domain. In Figure 4.8a, the outer circle marks the low-latitude boundary at 30° MLAT, and the middle circle marks the 50° MLAT circle, which is of our concern and outputted. The multi-level fitting region (inner square in Figure 4.8a) largely overlaps with the output domain (middle circle in Figure 4.8a). Such configuration (fixed-hybrid fitting) is designed to take advantage of the most available SuperDARN observations, primarily poleward of 50° MLAT.

Figure 4.8b shows the fitted potential and electric field at 09:37:00. Compared with Figures 4.4b1-b3, the large and mesoscale structures are similar. The fitting error shown in Figure 4.8c is slightly smaller than the 2° fitting error, which indicates that the fixed-hybrid fitting can better simulate the mesoscale structures of the electric
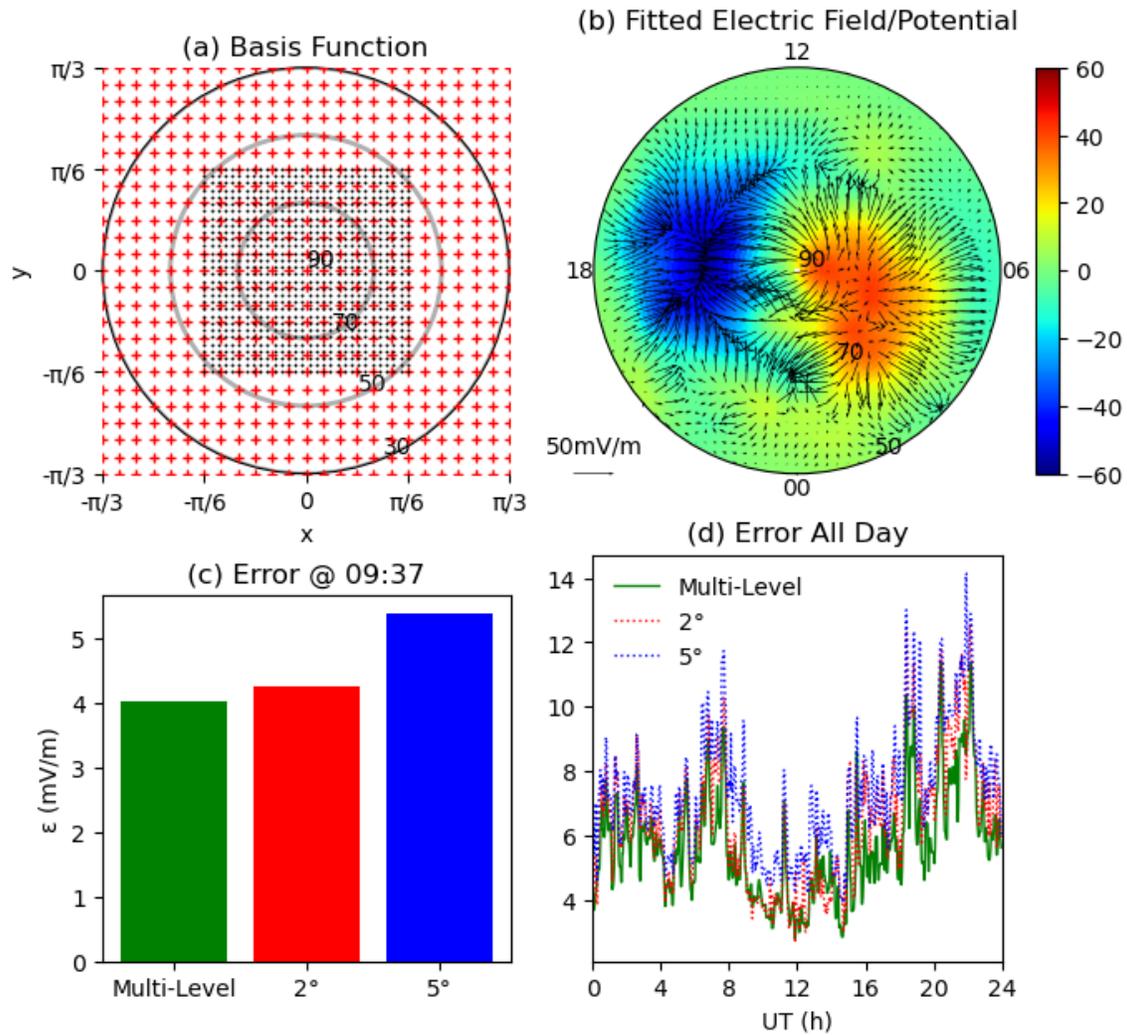
Figure 4.8: Multi-resolution fitting case: (a) Two-level basis function setup, (b) fitted potential and electric field, (c) comparison of fitting errors with 2° and 5° cases, and (d) fitting errors for all day. Units are $mV\,m^{-1}$ for electric fields and kV for potentials. The correspondence of x and y in (a) with MLT and MLAT can be found in Section 4.2.5.

field. Figure 4.8d shows the fitting errors for the whole day. Again, the fixed-hybrid fitting has the lowest errors most of the time. We list the averaged RMSEs of LOS electric fields throughout the day for all cases in Table 4.2. The top row indicates the resolution used for each case, and the first left column indicates the selection of the background model. The daily mean error from the background model itself is listed in the last column. Using SuperDARN SHF as the background model, the fixed-hybrid fitting decreases the RMSE by 29 %, and the pure 2° case decreases by 19 %, compared with the background model itself. Using Weimer model, these two decreases are 68 % and 66 %, respectively.

	Fixed-hybrid	Auto-hybrid	2°	5°	8°	BG
SuperDARN	5.71	6.51	6.44	7.37	7.64	7.99
Weimer	5.46	6.45	5.80	7.67	8.81	17.11

Table 4.2: Daily Means of RMSEs of LOS Electric Fields for All Cases (Unit is $\mathrm{mV\,m^{-1}}$)

Here, the multi-level case applies the high-resolution grids to a fixed region, but the actual measurements and data coverage usually change with time. Applying high-resolution grids to the region without data coverage causes a waste of computing time, and no substantial improvement is guaranteed. We perform a new test case with an auto-hybrid fitting approach (distinguished from the aforementioned fixed-hybrid fitting). The new setup of the auto-hybrid fitting consists of two levels: the low-resolution (5°) level covers the whole domain, while the high-resolution (2°) level is implemented in the region with observations, that is, determined by the data. We apply an auto adjustment to the high-resolution level every time to collocate the basis functions with observations. From the fixed-hybrid fitting to this auto-hybrid fitting, since the high-resolution basis functions in regions where observations are unavailable are removed, it has fewer basis functions than the fixed-hybrid fitting. The fitting

error of the auto-hybrid fitting is shown in Figure 4.9 as the blue line. The fitting errors from the uniform 2° basis setup (black line) and the fixed-hybrid fitting (red line) are also plotted in the figure. The errors from the auto-hybrid fitting are only slightly larger than the fixed-hybrid fitting and of similar magnitude as the 2° case.
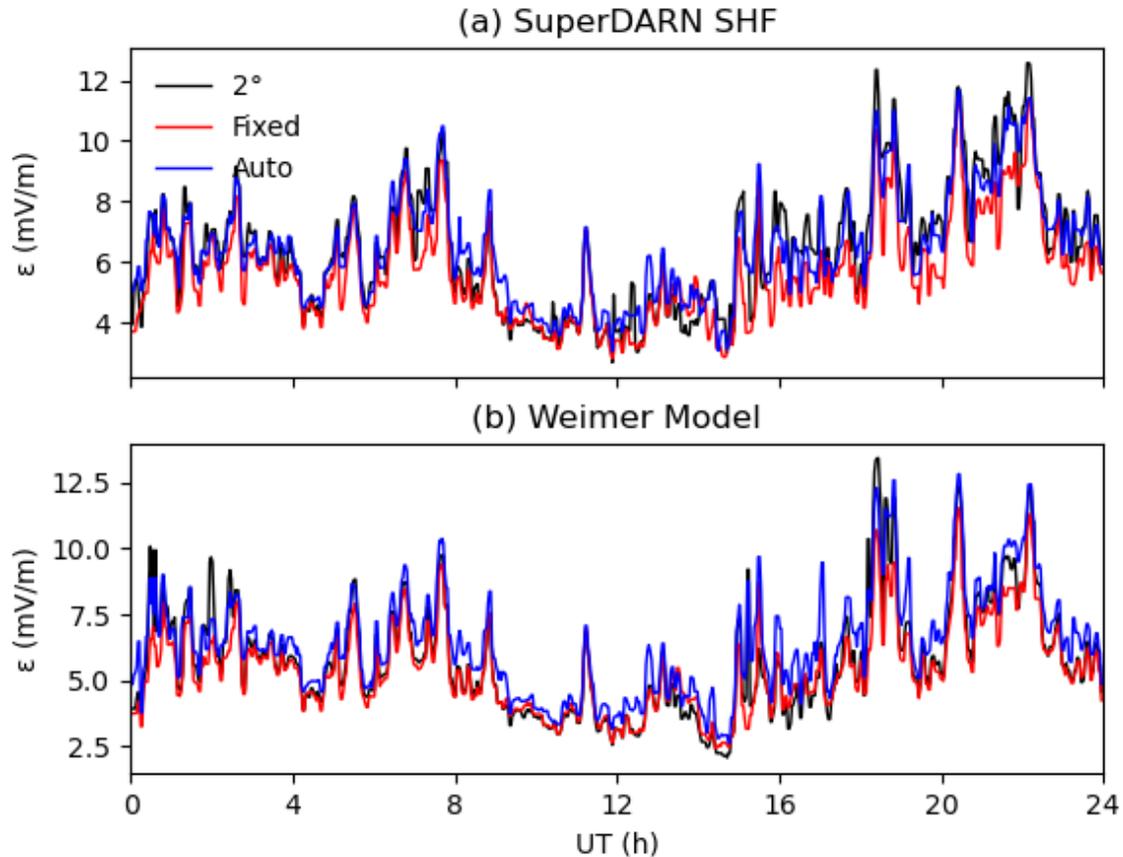
Figure 4.9: Fitting errors (RMSE) using (a) SuperDARN SHF as background model, (b) Weimer model as a background model. Red lines show fitting errors from the fixed multi-level setup. Blue lines show fitting errors from the auto-adjusted multi-level setup, and black lines show fitting errors from the uniform 2° setup for comparison.

It is worth mentioning that the computation time for the fixed-hybrid fitting is 40 % shorter than the pure 2° case. The auto-hybrid fitting enables a time decrease of 80 %, which suggests that with a proper setup of basis functions, the auto-hybrid fitting is likely the most affordable and efficient choice for regional high-resolution

assimilation.

4.3.3 Local Electric Field Modeling Using PFISR Data

On 2015-03-17, PFISR was operating in several different modes (IPY27_Tracking, Freg732, LTCS35, and WorldDay35) during the day, and ion drifts from all modes over Poker Flat as available are used to derive electric fields (Equation (4.4)) for the entire storm period. The derived PFISR electric fields are then fed into the data assimilation model with SuperDARN measurements to obtain the electric fields. The spatial resolution of PFISR measurements is $0.25°$, and the temporal resolution is typically every $1\,\mathrm{min}$ to $2\,\mathrm{min}$.

Figure 4.10 shows the $2°$ fitted electric fields at Poker Flat when using Super-DARN SHF as the background model. PFISR measurements and the electric field from SuperDARN SHF are also plotted for comparison. Roughly speaking, the background model agrees well with PFISR measurements though some large fluctuations are missing (e.g., E_x at 08:30:00, 12:00:00 and E_y at 06:30:00, after 11:00:00). Instead, the fitted eastward and northward electric fields from Lattice Kriging capture such large fluctuations and follow the actual PFISR observations more closely. For example, the positive E_y peak at 06:30:00 underestimated in SuperDARN SHF is largely elevated to a comparable level to PFISR measurements, while at 07:30:00 the peak only present in the background model is attenuated and becomes more comparable with PFISR by the fitting process. Nevertheless, there are periods when the fitting results deviate from the observations, such as E_x after 12:00:00. The ambient SuperDARN LOS electric field measurements might influence the fitting results during these periods, which show differences from PFISR.

In summary, Lattice Kriging can capture the short-period variations shown in

126
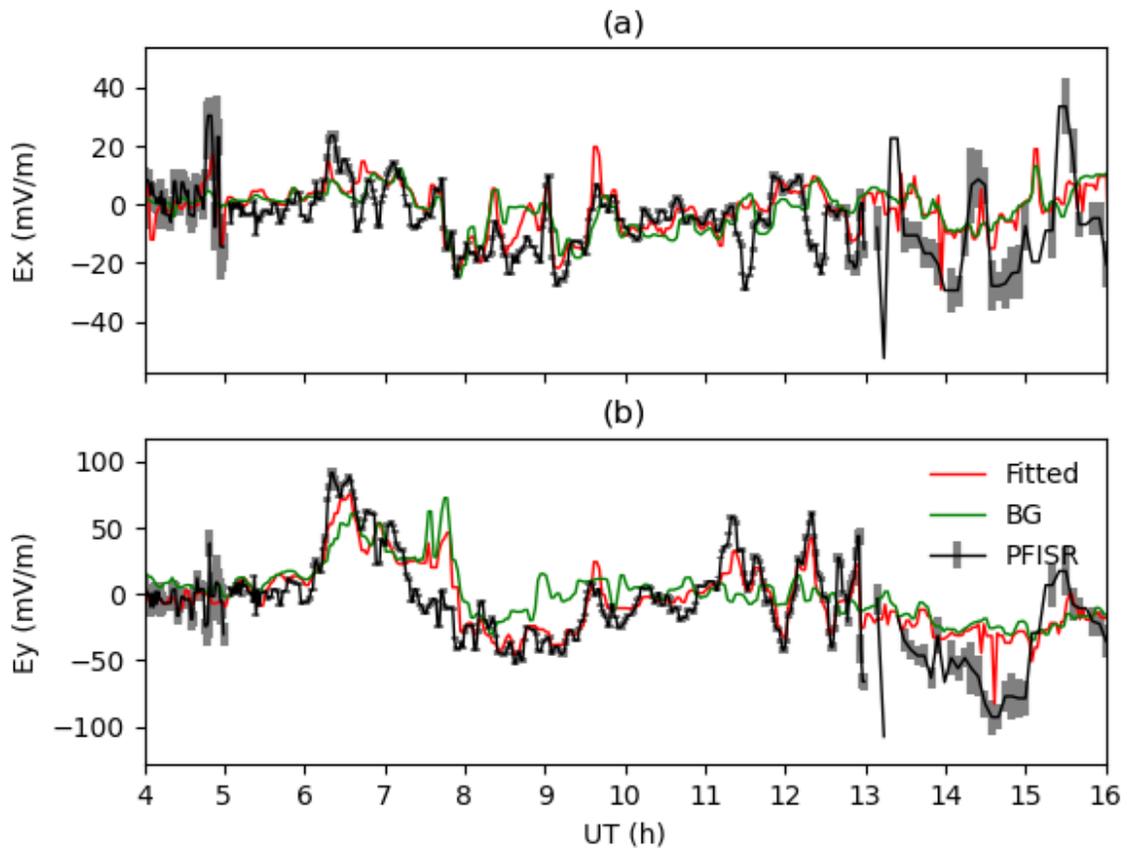
Figure 4.10: (a) Eastward and (b) northward electric fields at Poker Flat. The dotted black lines are PFISR observations, blue lines are SuperDARN SHF, and red lines are Lattice Kriging fitted results.

the local data and maintain the coherence to the ambient electric fields to a large extent. Compared with just padding the local observations into the background model, Lattice Kriging avoids the discontinuity problem and largely fuses the information from real, local observations.

4.4 Conclusions and Outlook

This paper develops a new methodology to assimilate high-latitude electric fields by extending the Lattice Kriging framework to vector fields. This modeling assumes that the fitted field is a Gaussian process. By combining the background model (SuperDARN SHF or Weimer model in our case) used as "a priori" knowledge and the available observations (SuperDARN and PFISR), the means, variances, and covariances of the to-be-assimilated variables (electric fields and ion drifts) are calculated and further used to reconstruct the fields where no observations are available. This methodology assimilates the observational data and provides the fitted results for the whole domain of interest.

We systematically evaluate the performance of Lattice Kriging using different resolutions and background models. We first assimilate the electric fields for the St. Patrick's Day storm (2015-03-17) with SuperDARN data. When the same amount of data is provided, the fitting results are similar to whichever background model is used. This suggests that data is more important than the background model in fitting the electric fields. By varying the amount of input data, we find that the impacts of data tend to be remote and reach the regions without observations. Compared with the background model, data assimilation leads to considerable decreases in the RMSE of the LOS electric field. Such improvement is more significant against the Weimer model than SuperDARN SHF, likely because the former is empirical, while

the latter has already incorporated some of the data information. Comparing across the different fitting resolutions (2°, 5°, and 8° in our case), higher resolution always leads to smaller RMSE, suggesting that more details in the observations are captured with more basis functions (higher resolutions) used in the fitting model.

We demonstrate the capability and advantages of multi-resolution modeling using multi-level basis functions. 2° in the fine level and 5° in the coarse level are adopted to form a two-level framework as an example. Two types of configuration of the high-resolution (2° in this case) grids are tried. The fixed-hybrid grid covers the most prominent region where all the data appear and uses it as a fixed region for high-resolution fitting. The auto-hybrid grid adjusts the high-resolution region according to the real-time data coverage. The fixed-hybrid fitting errors further decrease compared with the pure 2° case, and the computation time shortens by 40%. The auto-hybrid fitting has a similar fitting error as the pure 2° case. The decrease of computation time reaches 80%, suggesting that (a) the multi-level basis function can further improve the fitting, (b) using the relatively coarse grid in the region without observations does not degrade the performance and saves computational cost, and (c) the auto-hybrid fitting provides an efficient way to perform the regionally high-resolution data assimilation. Using SuperDARN SHF as the background model, the multi-level assimilation decreases RMSE by 29%, and 2° case decreases by 19%, compared with the background model itself. High-resolution observations can be better assimilated using the multi-level basis function setup (especially auto-hybrid fitting) with affordable computational resources.

Even though the multi-level basis function setup can effectively reduce errors compared with other setups, the errors are still substantial (daily mean of $5.71\,\mathrm{mV\,m^{-1}}$ for the auto-hybrid fitting and $6.51\,\mathrm{mV\,m^{-1}}$ for the fixed-hybrid fitting). There are two possibilities to reduce the errors further: (a) The covariance matrix is

to be improved. The covariance matrix used here is derived from a Gaussian Markov random field (which assumes two locations are correlated only if adjacent, Nychka et al. (2015)). However, electric potentials/fields are not necessarily uncorrelated, even if they are apart for some distance. Strictly, an additional term indicating the medium-range electric field correlation needs to be included in the covariance matrix to describe the real-world electric field characteristics, and (b) the mathematical formulation is to be modified based on non-Gaussian process models. The current development assumes the electric field is a Gaussian process, while in the real world, the distribution of electric fields deviates from Gaussian (Golovchanskaya and Kozelov, 2010a,b). Still, Gaussian statistics have good properties for fast computation, such as the sparse matrix calculation as aforementioned, which satisfies as a starting point.

Further, we include PFISR observations in the model. The fitting results at Poker Flat better capture the short-period electric field variations shown in the data than those from the global SHF (i.e., SuperDARN SHF). This indicates that our method can efficiently fuse and then recover the local measurements and, importantly, maintain the coherence of the patterns with the ambient electric fields.

The decreases in RMSE, the flexibility of incorporating various data sources, and the benefits of the multi-level setup embedded in Lattice Kriging show that it is a powerful tool in the data assimilation application. The application of such a method is not limited to an electric field and ion drift. Still, it can also be applied to other physical quantities, such as FAC as a scalar field and wind as a vector field. We use the curl-free assumption for electric fields to provide an additional constraint for the modeling. For neutral winds in relatively large scales (e.g., planetary and synoptic scales), the vertical gradient of the vertical wind is negligible, and the horizontal winds are approximately divergence-free. A stream function is well-defined with the divergence-free assumption. Defining basis functions similarly as mentioned in Section

130

4.2, the modeling of the stream function using horizontal wind measurements can be similarly formulated. The new Michelson Interferometer for Global High-resolution Thermospheric Imaging (MIGHTI) instrument onboard ICON provides neutral wind measurements over the mid-latitude and low-latitude regions, which may provide an optimal data set to assimilate the multi-resolution structures of neutral dynamics using Lattice Kriging.

Chapter 5

Applying Data-Driven Aurora and Electric Fields to TIEGCM in Studying the I-T Responses during the 2015 St. Patrick's Day Storm

Through Chapters 3-4, we describe the new data assimilation method for aurora and electric fields in detail. A natural question arises: how does this method perform in real applications? To answer this question, we choose the 2015 St. Patrick's Day storm when substantial vertical wind measurements (over $50\,\mathrm{m\,s^{-1}}$) occur at several stations in Alaska. This chapter uses TIEGCM and the new data assimilation method to study the global and local storm impacts.

As one of the strongest geomagnetic storms in Solar Cycle 24, the 2015 St. Patrick's Day storm has attracted significant attention. We revisit this event by taking advantage of simultaneous observations of high-latitude forcings (aurora and electric fields) and I-T responses. The forcing terms are assimilated to drive TIEGCM

using a newly adopted Lattice Kriging method (Wu and Lu, 2022; Wu et al., 2022). Compared to the default run, the TIEGCM simulation with assimilation captures (a) secondary E region electron density peak due to aurora intensification; (b) strongly elevated ion temperatures (up to $3000\,\mathrm{K}$) accompanied by a strong northward electric field ($80\,\mathrm{mV\,m^{-1}}$) and associated ion frictional heating; (c) elevation of electron temperatures; and (d) substantially enhanced neutral vertical winds (order of $50\,\mathrm{m\,s^{-1}}$). RMSEs decrease by $30\,\%$ to $50\,\%$. The strong neutral upwelling is caused by large Joule heating down to $120\,\mathrm{km}$ resulting from enhanced aurora and electric field. Data assimilation increases the height-integrated Joule heating at Poker Flat to a level of $50\,\mathrm{mW\,m^{-2}}$ to $100\,\mathrm{mW\,m^{-2}}$ while globally, its maximum value is comparable with the default run: the location of energy deposition becomes guided by data. TADs in the assimilation run show stronger magnitudes and larger extensions, leading to increased vertical wind variability by a factor of 1.5 to 3. Our work demonstrates that data assimilation of model drivers helps produce realistic storm-time I-T responses, which show richer dynamic range, scales, and variability than what has been simulated before.

5.1 Introduction

Joule heating and particle heating induced by particle precipitation are the two most crucial thermospheric heating sources in the auroral zones and compete with solar irradiance during geomagnetic storm times (Knipp et al., 2004). Most of the energy deposited in the magnetosphere by the solar wind is ultimately dissipated in the I-T system as a result of the convergence of Poynting flux and auroral particle precipitation that heats the atmosphere directly through collision and indirectly by increasing Joule heating (Lu et al., 1998a; Richmond and Thayer, 2000; Thayer and

133

Semeter, 2004; Thayer et al., 1995). During periods of geomagnetic activity, both auroral precipitation and electric fields are elevated, leading to significant energy deposition in the I-T region and localized Poynting fluxes (Deng et al., 2011; Fuller-Rowell et al., 1987; Heppner et al., 1993; Wang et al., 2005). For instance, DMSP observed a local Poynting flux exceeding $170\,\mathrm{mW\,m^{-2}}$ associated with merging at the magnetopause flank and lobe during an east-west IMF dominant event (Knipp et al., 2011). Satellite observations have also revealed that localized Earth-directed Poynting fluxes can be several times larger than the prediction from statistical models (order of $15\,\mathrm{mW\,m^{-2}}$) during magnetic storms (Huang and Burke, 2004; Huang et al., 2016).

However, such localized energy deposition and its effects are difficult to capture in I-T models driven by empirical high-latitude inputs obtained from statistical auroral models (Hardy et al., 1985; Newell et al., 2009) and electric field models (Heelis et al., 1982; Roble and Ridley, 1987; Weimer, 2005). Empirical models usually provide the large-scale ($> 1000\,\mathrm{km}$) morphology of the auroral oval, which differs from the auroral precipitation seen in the observations, such as that observed by THEMIS ground-based ASIs. THEMIS ASIs generally depict rich mesoscale ($10\,\mathrm{km}$ to $100\,\mathrm{km}$) structures, while narrow FOV imaging reveals even small-scale ($< 10\,\mathrm{km}$) auroral patterns (Nishimura and Lyons, 2021; Nishimura et al., 2021a,b). Electric fields in the polar cap and auroral region also exhibit cross-scale spectra ranging from planetary scales down to a few kilometers (Golovchanskaya and Kozelov, 2010a; Kozelov and Golovchanskaya, 2006), which deviate from the global large-scale two-cell ion convection pattern (Cousins and Shepherd, 2012a,b). Cousins et al. (2013) performed a scale analysis of electric fields using SuperDARN, showing that the first several EOF modes contribute to global-scale features. In contrast, higher-order EOFs (> 8) contribute to mesoscales ($< 1000\,\mathrm{km}$), which have a shorter decorrelation time indicating

134

more variability. Small-scale electric fields have often been observed and found to impact the energy budget during magnetic storms (Codrescu et al., 1995; Cosgrove and Codrescu, 2009). Wu et al. (2020) found that implementing electric fields varying on short temporal scales and the aurora observed by satellites into TIEGCM is essential to reproducing TTEIL observed by the Fe Boltzmann lidar at McMurdo, Antarctica. Sheng et al. (2020) found that GITM driven by the THEMIS/ASI auroral observations better resolves the magnitude of TIDs (doubled) than that by empirical auroral inputs. Both studies require the fusion of aurora and electric field observations to constrain the model drivers for a better understanding of storm-time I-T responses.

The St. Patrick's Day storm (2015-03-17) is the strongest geomagnetic storm during Solar Cycle 24. A special section in the Journal of Geophysical Research Space Physics entitled "Geospace system responses to the St. Patrick's Day storms in 2013 and 2015" has highlighted a series of storm impacts (Zhang et al. (2017) and references therein), ranging from the ionospheric tongue of ionization and storm enhanced density at high latitudes, sub auroral polarization stream and ion up flow at sub auroral and midlatitude regions, to significant Prompt Penetration Electric Field and Disturbance Dynamo Electric Field effects on low-latitude electrodynamics. Neutral dynamics involve composition changes (mainly O/N_2) and the resultant TEC depletion during the storm's recovery phase. However, studies covering magnetospheric forcings and I-T responses and using well-constrained models to connect them were relatively rare.

This investigation aims to marry the improved data assimilation technique with the I-T model (TIEGCM in this case) to understand better what physics drove the significant I-T responses observed locally by the three Fabry-Perot Interferometers (FPIs) and PFISR. Three FPIs located at Eagle (64.8°N, 141.2°W), Toolik (68.6°N, 149.6°W), and Kaktovik (70.1°N, 143.6°W), Alaska, all showed strong upward winds

with maxima reaching an order of $50\,\mathrm{m\,s^{-1}}$ to $100\,\mathrm{m\,s^{-1}}$ during the period of 08:30:00 to 09:30:00 at $250\,\mathrm{km}$. Larsen and Meriwether (2012) reported that vertical winds with a magnitude of $10\,\mathrm{m\,s^{-1}}$ to $20\,\mathrm{m\,s^{-1}}$ are common at high latitudes. In contrast, significant vertical winds (up to $50\,\mathrm{m\,s^{-1}}$) are often accompanied by substorm events at high latitudes or the generation of irregularities or sporadic layers at mid and low latitudes. Since the atmosphere is highly structured vertically, large vertical wind can cause significant displacement and mixing, thus playing an essential role in the dynamics. The observations showing vertical winds larger than $50\,\mathrm{m\,s^{-1}}$ were relatively rare and deserved an investigation. In addition to the FPI measurements, the PFISR, which is located at Poker Flat (65.1°N, 147.5°W), Alaska, had continuous measurements of electron densities (N_E), LOS ion drifts (V_{LOS}) which can be used to derive electric fields, and ion and electron temperatures (T_I and T_E, respectively). The THEMIS/ASIs also had decent coverage. The geometry of the above four stations along with an ASI site at Yakutat (59.5°N, 139.7°W), Alaska, is shown in Figure 5.1. Such simultaneous observations of magnetospheric drivers (aurora and electric fields) and responses (neutral winds and ionospheric parameters) provide an excellent opportunity to constrain the I-T model at high latitudes, which can be further used to study the connections.

The paper is organized as follows. Data sources, assimilation method, and TIEGCM runs are introduced in Section 5.2. Model results, along with the comparisons with data and critical physical processes responsible for the strong upward winds and salient changes in the I-T system, are discussed in Section 5.3. Discussion and conclusions are given in Section 5.4.
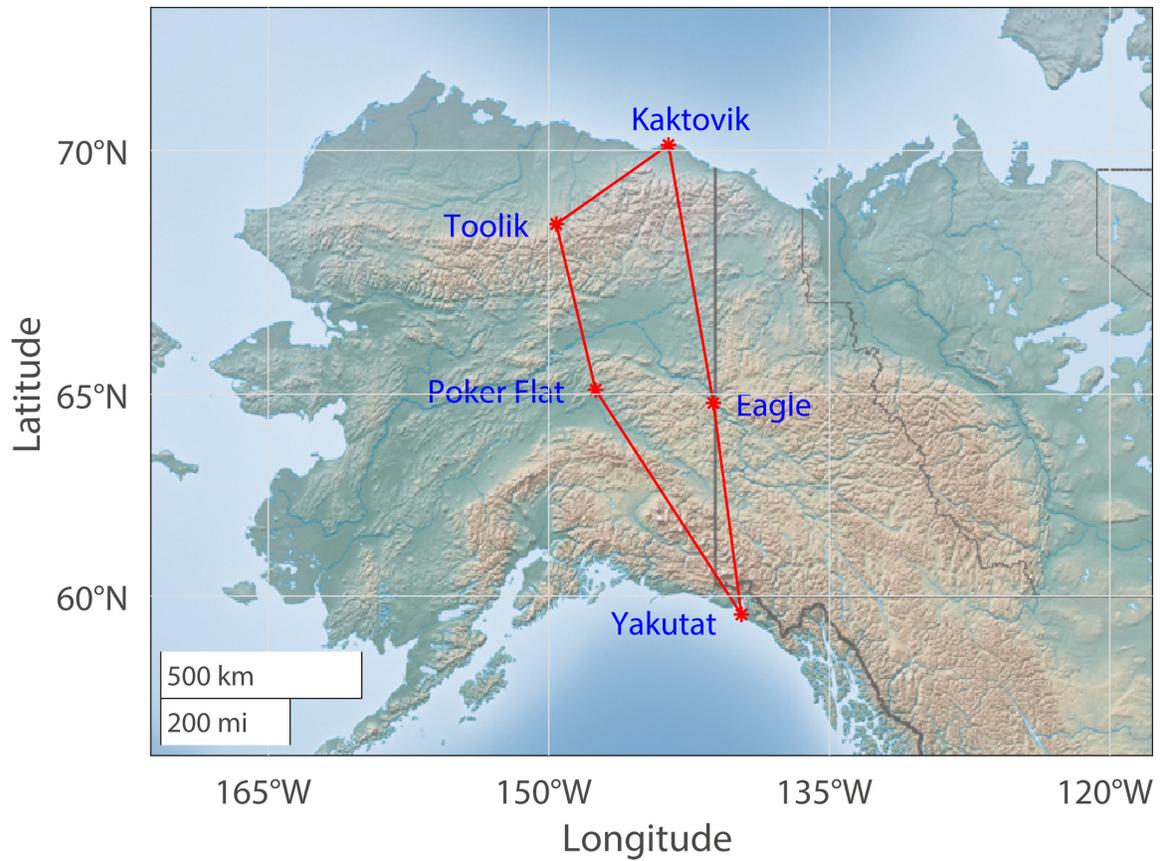
Figure 5.1: Geometry of the five observational stations in Alaska, including Poker Flat (65.1°N, 147.5°W), Eagle (64.8°N, 141.2°W), Toolik (68.6°N, 149.6°W), Kaktovik (70.1°N, 143.6°W), and Yakutat (59.5°N, 139.7°W).

5.2 Ground-Based Observations, Data Assimilation, and TIEGCM Runs

5.2.1 St. Patrick's Day Storm and Ground-Based Observations

Figure 5.2 shows the geomagnetic indices on 2015-03-17 and 2015-03-18. The geomagnetic storm starts at 06:00:00 on 2015-03-17 as the IMF B_z turns south. The period of southward B_z lasts for almost a whole day before this parameter returns to zero at 05:00:00 on 2015-03-18, after which there are some minor fluctuations. IMF B_y turns to negative around 05:00:00, 2015-03-17, and lasts for 6 h before becoming positive and lasting for another 12 h. Solar wind velocities and densities show obvious enhancements during the southward B_z period (Figure 5.2b). Figure 5.2c shows AE indices, which illustrate strong auroral activities after 06:00:00 and reach the first peak around 09:00:00 on 2015-03-17. The SYM-H index reaches a negative maximum ($-230\,\mathrm{nT}$) around midnight 2015-03-17, and the storm remains in the recovery phase till midnight 2015-03-18. Both IMF and AE indices show fast oscillations with periods of tens of minutes to an hour.

The simultaneous PFISR observations of ion drifts and electric fields at Poker Flat, Alaska, and neutral vertical wind measurements observed by the three FPIs are shown in Figure 5.3. During the period of 05:00:00 to 13:00:00, PFISR was running at an experiment containing two long-pulse modes, which was suitable for the derivation of F region parameters including N_E, T_I, T_E, and V_{LOS}. The F region measurements use a long pulse with a 480 μs uncoded pulse and are gated to have a spacing of 36 km with a range resolution of 72 km, respectively. The E region data presented are smeared due to the ambiguity function associated with the long

138
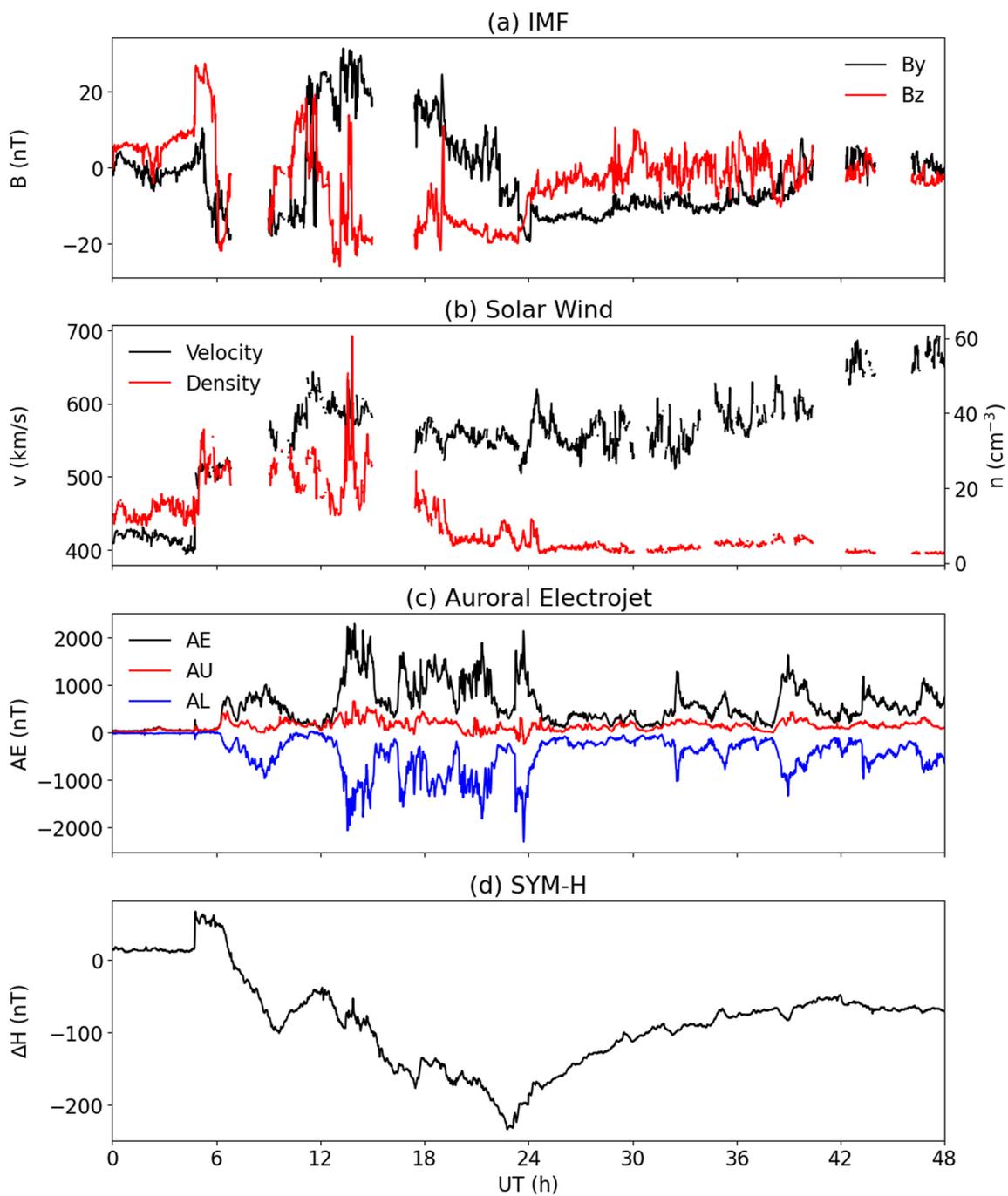
Figure 5.2: Geomagnetic indices on 2015-03-17 and 2015-03-18: (a) IMF B_y and B_z; (b) solar wind velocity and density; (c) AE indices; and (d) SYM-H.

pulse, although significant enhancements are captured. The methodology used to take the F region V_{LOS} observations and determine the plasma drifts is described in Heinselman and Nicolls (2008). Right after the storm onset, a strong westward ion drifts with a magnitude of $-1800\,\mathrm{m\,s^{-1}}$ is detected around 06:30:00 corresponding to a northward E_y of $85\,\mathrm{mV\,m^{-1}}$. After 08:00:00, E_y switches to southward and maintains a significant magnitude of $-50\,\mathrm{mV\,m^{-1}}$ which corresponds to an eastward ion drift of $1000\,\mathrm{m\,s^{-1}}$. Such southward E_y and eastward ion drift values last for nearly 2 h until 10:00:00 during this period (08:00:00 to 10:00:00), the westward E_x increases to a magnitude of $-30\,\mathrm{mV\,m^{-1}}$ (blue line in Figure 5.3b) along with strong southward ion drift (blue line in Figure 5.3a).

At the same time, the three FPIs all show sustained strong upward winds for a considerable period (from 08:30:00 to 10:00:00). The peaks of the vertical winds (W_N) reach $50\,\mathrm{m\,s^{-1}}$ to $100\,\mathrm{m\,s^{-1}}$ at Toolik and Kaktovik, and even exceed $100\,\mathrm{m\,s^{-1}}$ at Eagle. Around 06:00:00, the FPIs observe strong downward neutral winds and reach about $-50\,\mathrm{m\,s^{-1}}$. Even though the temporal evolutions of the three observations follow each other in general, the differences in the magnitude and direction at a particular time still exist. For instance, after 10:00:00, the vertical wind keeps positive at Kaktovik, fluctuates around zero at Poker Flat, while turns into negative at Toolik, until 11:00:00 when they become both negative at Toolik and Kaktovik and turn back to positive at Poker Flat. Considering that the distances among these three stations are only about a few hundred kilometers (Figure 5.1), such local differences suggest that vertical winds are highly structured and characterized by mesoscale patterns.

In addition to electric fields and neutral vertical winds, the discrete auroral activity is also intensified as captured by the THEMIS/ASIs (Mende et al., 2008). A substorm onset occurs at 06:45:00 (a snapshot is shown in Figure 5.4a), followed by a bulge expanding poleward and azimuthally toward the Alaska-Canada border.
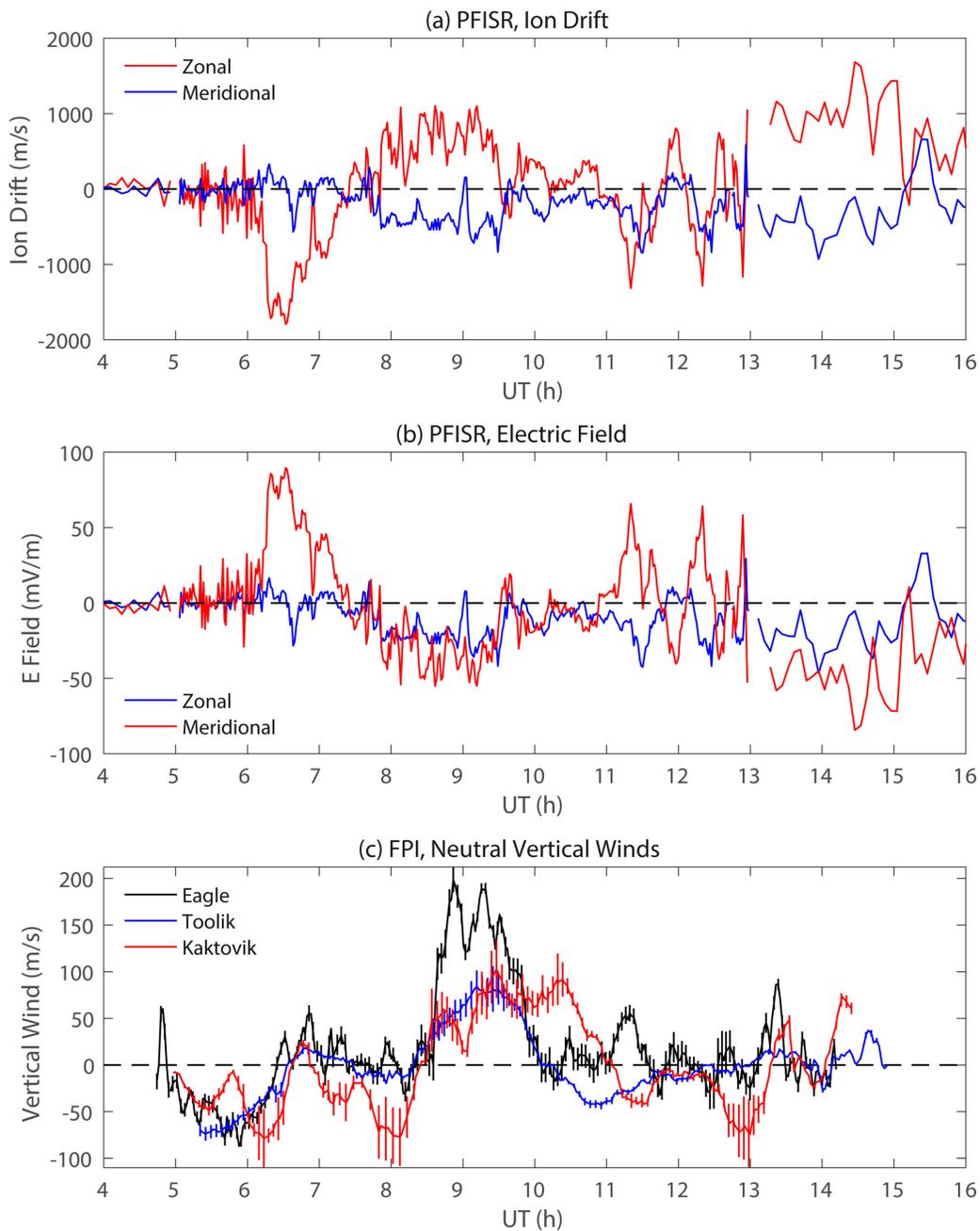
Figure 5.3: (a) Ion drifts, (b) electric fields (zonal E_x, meridional E_y) measured by PFISR, and (c) neutral vertical winds measured by FPI on 2015-03-17 at Eagle (black), Toolik (blue), and Kaktovik (red), Alaska, respectively. Vertical error bars denote measurement uncertainties.

Another substorm onset occurs at 08:33:00 near Yakutat, Alaska, with magnificent auroral emission and brightening lasting until 09:40:00 (Figures 5.4c-d), during which multiple intensifications of the bulge and auroral streamers appeared (Nishimura and Lyons, 2021) near Poker Flat and Eagle. Even though the sky was somewhat cloudy over Poker Flat, which disabled the reliable derivation of energy flux and mean energy from the auroral emission, the raw images still show considerable auroral activity at Poker Flat during this substorm (Figures 5.4c-d), which lends confidence of padding aurora for Poker Flat in the data assimilation process (more details in Section 5.2.2). The precipitating energy flux and characteristic energy were obtained from the THEMIS ASI data (Nishimura and Lyons, 2021) and were incorporated into the data assimilation.

5.2.2 Data Assimilation for Aurora and Electric Fields

The Lattice Kriging modeling has been recently adopted for the data assimilation of aurora (Wu et al., 2022) and extended for the assimilation of electric fields (Wu and Lu, 2022). It has largely captured actual data's temporal and spatial variability. The field to be estimated is assumed to consist of a statistical mean, a variability term (spatially), and an error term. The empirical model can be used as the statistical mean with a scaling factor. We choose the (Zhang and Paxton, 2008) model for aurora, built upon the historical SSUSI observations and Kp driven, and the Weimer model (Weimer, 2005) for electric fields. The variability term is projected into the superposition of a series of RBFs more suitable to assimilate local features than the global spherical harmonic fitting. The scaling factor and coefficients of RBFs are estimated from existing observations, which are then applied to reconstruct the field globally. The detailed mathematical formulation and principles can be found in Wu
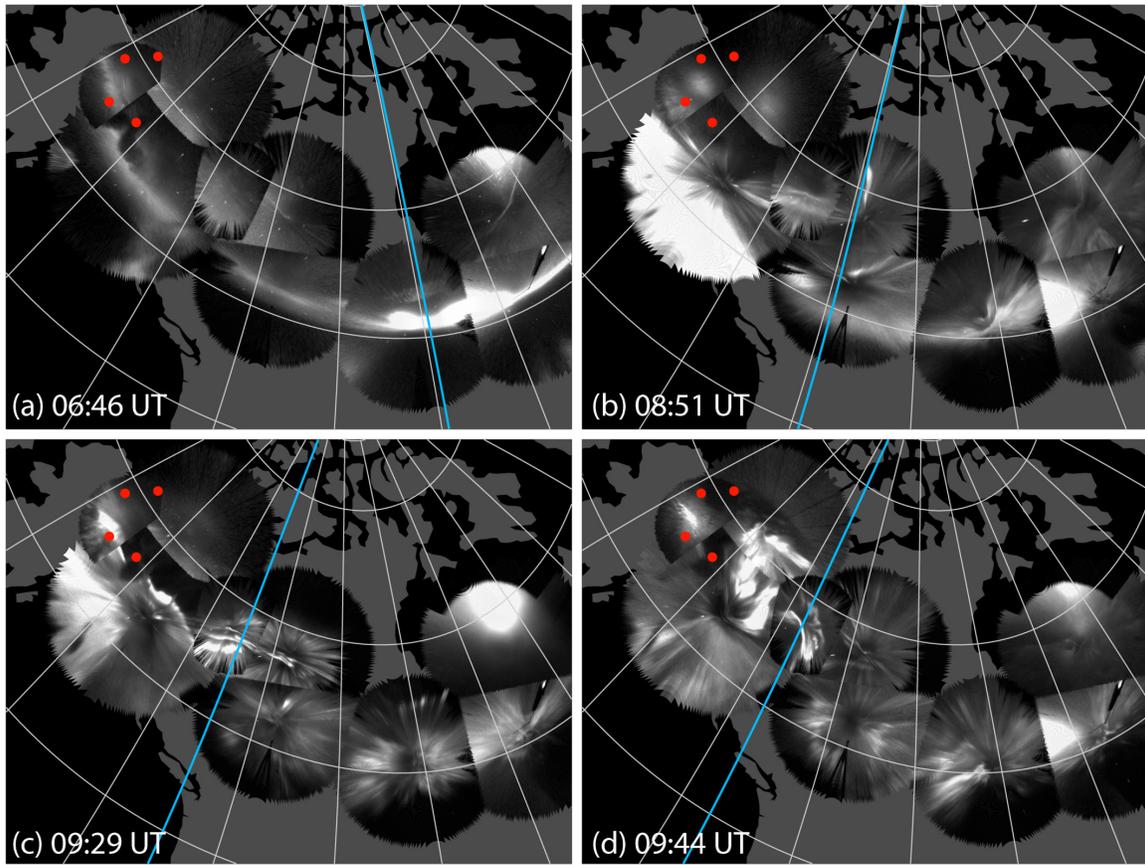
Figure 5.4: Evolution of aurora: white line emission (count/s) from THEMIS/ASI observations. Four red spots denote Poker Flat, Eagle, Toolik, and Kaktovik, Alaska.

and Lu (2022) and Wu et al. (2022).

SSUSI (Paxton and Meng, 1999; Paxton et al., 2002) onboard DMSP and THEMIS/ASI data are used for auroral assimilation. Figure 9 of Wu et al. (2022) shows that the Lattice Kriging model can coherently combine the three data sets (empirical auroral model, SSUSI, and THEMIS/ASIs) with a smooth boundary transition and largely keep mesoscale features such as auroral arcs shown in the observations. The auroral image from Yakutat, Alaska (57°-61°N, 130°-140°W), is padded to a box of (65°-71°N, 140°-160°W) near Poker Flat, partially guided by the raw auroral images which illustrate that the temporal variations of aurora at these two locations are decently correlated (Figure 5.4). Such an implementation is reasonable by comparing TECs with GNSS observations (Section 5.3.2). Without padding, the TEC at the northwest corner of Alaska, including Poker Flat, is too small compared with the GNSS observations. This implies that significant auroral activity is expected during the time of interest. Figure 5.5 compares the mean (characteristic) energy and number flux between the empirical model and data assimilation. The assimilated maps are more structured, and the magnitude of number flux calculated as energy flux divided by mean energy is significantly elevated due to the incorporation of the SSUSI and THEMIS/ASI observations which have much more significant energy fluxes than the empirical model (Figures 5.5c-d). It is worth mentioning that the energy fluxes of THEMIS/ASIs need to be halved before feeding into the assimilation model to produce electron densities and TECs in the TIEGCM that match observations. The measurements' uncertainties and ionization rate calculation in the model may cause this adjustment. Further work is needed to pinpoint the exact source.

The V_{LOS} (converted to LOS electric fields) from the SuperDARN measurements and the PFISR observations of electric fields (Figure 5.3a) at Poker Flat (PFISR data are repeated in the ambient $4° \times 4°$ region to ensure its weighting
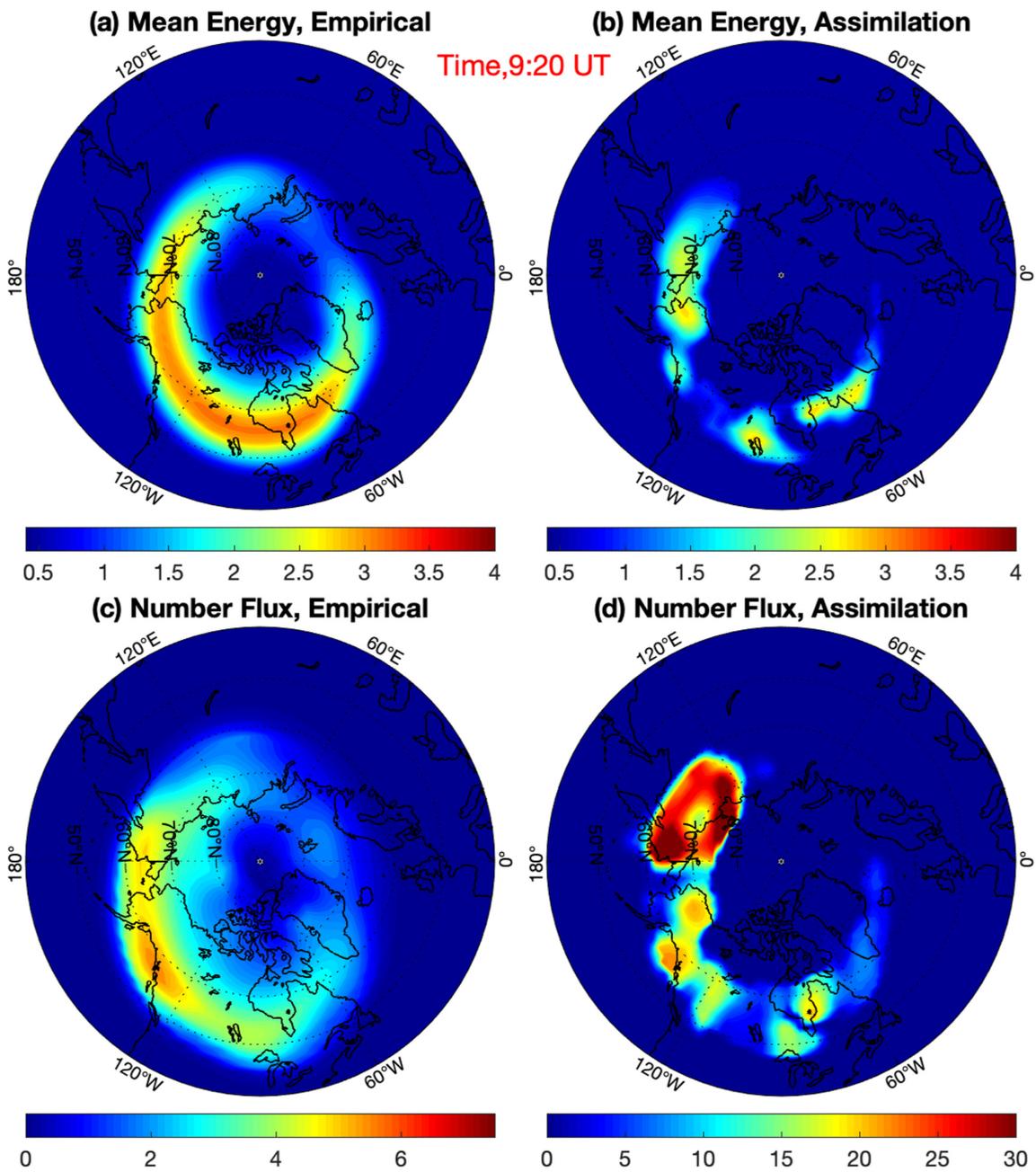
Figure 5.5: (a, b) Auroral mean energy (keV) maps at 09:20:00 from the empirical model (Zhang and Paxton, 2008) and assimilation, respectively. (c) and (d) are the same except for the number flux (10^8 cm^{-2} s^{-1}). Note that panels (c) and (d) color bars differ. All these maps are the ones used to drive different TIEGCM runs.

in the fitting) are used as data sources for the electric field assimilation. Wu and Lu (2022) have successfully extended the Lattice Kriging model to vector fields, and they presented results from the same event. Figures 7 and 10 from Wu and Lu (2022) show that the assimilation results largely follow the PFISR observations locally. RMSEs (determined by the differences between assimilation and observations) decrease by more than half compared with the Weimer model. We compare the empirical and assimilated electric fields at 06:28:00, when E_y reaches the maximum at Poker Flat, in Figure 5.6. The empirical model (Weimer, 2005) shows the large-scale patterns reasonably well, while the assimilation results resolve much finer structures. The empirical model shows positive E_y near Poker Flat with a much smaller magnitude versus the observation (Figure 5.6c vs. 5.3b). After the data assimilation, the strong northward E_y is enforced near Poker Flat (Figure 5.6d), which helps to constrain the local electrodynamics. Since the SuperDARN data used for assimilation cover the high-latitude northern hemisphere reasonably well for this case (see Figure 4a in Wu and Lu (2022)), the assimilated results should, in general, impose more realistic electric fields than the empirical model, especially for the regions where the data are available. Figure 10 of Wu and Lu (2022) shows that the data assimilation also captures the short-term temporal variability (within an hour) of electric fields from the PFISR observations.

5.2.3 Four TIEGCM Runs

TIEGCM is a global 3-D numerical model that simulates the coupled thermosphere/ionosphere system from $97\,\mathrm{km}$ to $600\,\mathrm{km}$ altitude. It self-consistently solves the fully coupled nonlinear, hydrodynamic, thermodynamic, and continuity equations of the neutral gas, the ion and electron energy equations, the O^+ continuity equation
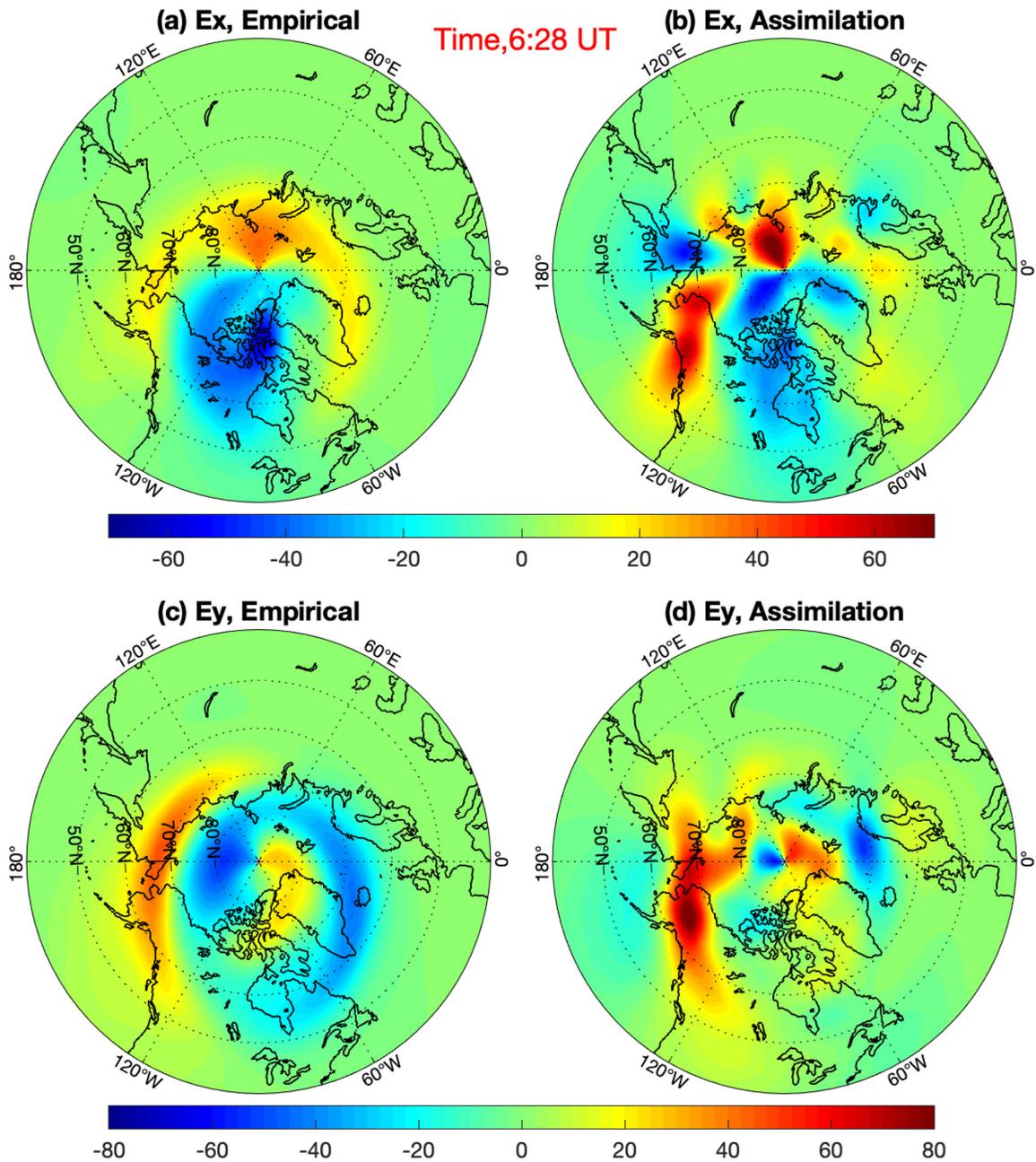
Figure 5.6: (a, b) E_x and (c, d) E_y at 06:28:00, respectively. (a) and (c) are from the empirical Weimer model, while (b) and (d) are from the data assimilation. The unit is $\mathrm{mV\,m^{-1}}$.

and ion chemistry, and the neutral wind dynamo (Qian et al., 2014; Richmond et al., 1992). In this study, the resolution of TIEGCM is $1.25° \times 1.25° \times 1/8$ scale height in latitude $\times$ longitude $\times$ altitude (Dang et al., 2018, 2021). Realistic F10.7 is used in all simulations. Realistic IMF conditions drive the Weimer model for the empirical high-latitude electric field run. The time step of the TIEGCM simulation is $10\,$s. The output frequency of the diagnostic terms is $1\,$min.

In the current study, both aurora and electric fields have two types of maps, one from the empirical model and the other from the data assimilation, giving rise to a combination of four options for high-latitude drivers. These four options are used to drive the TIEGCM. We list the names (R1-R4: $R_{a\_emp\_e\_emp}$, $_{a\_assi\_e\_emp}$, $R_{a\_emp\_e\_assi}$, and $R_{a\_assi\_e\_assi}$) and the corresponding setups for the auroral and electric field drivers of these four runs in Table 5.1.

	R1: $R_{a\_emp\_e\_emp}$	R2: $R_{a\_assi\_e\_emp}$	R3: $R_{a\_emp\_e\_assi}$	R4: $R_{a\_assi\_e\_assi}$
Aurora	Empirical	Assimilated	Empirical	Assimilated
Electric fields	Empirical	Empirical	Assimilated	Assimilated

Table 5.1: Names and setups for the auroral and electric field drivers for the four different TIEGCM runs. "a" and "e" are the short names for "aurora" and "electric field", respectively; "emp" and "assi" are the short names for "empirical model" and "assimilated results", respectively

5.3 Model Results, Model-Data Comparisons, and Mechanism Studies

5.3.1 Local Neutral and Ionospheric Responses

Before we compare the modeled neutral and ionospheric responses among the four different TIEGCM runs, we show the PFISR measurements of electron density

(N_E), electron temperature (T_E), and ion temperature (T_I) to provide an observational baseline (Figure 5.7). The most prominent feature in N_E compared with the quiet-time (not shown here) is the secondary peak formation (around 150 km to 200 km) during the period of 09:00:00 to 10:30:00 (black rectangular in Figure 5.7a), most likely produced by auroral particle precipitation. The timing is consistent with the THEMIS/ASI observations of auroral surge and brightening (Figure 5.4, Section 5.2.1). Ion temperatures show significant elevation and reach 3500 K around 06:30:00 (Figure 5.7c), concurrent with the occurrence of strong E_y (Figure 5.3b) and indicative of strong frictional heating. Intriguingly, the strongest enhancement of T_I occurs in the altitude range below 250 km. Electron temperatures also increase compared to the quiet time, and the two most significant heating periods occur around 06:30:00 and 09:00:00 when either the electric field or aurora is intensified. Different from the enhancement of ion temperatures, which maximizes around 150 km to 200 km near 06:30:00, the enhancement of electron temperatures generally increases with altitude. Around 08:00:00, 08:30:00, 09:00:00, 09:45:00, 11:30:00, 12:20:00, and 12:40:00, both ion and electron temperatures show significant increases especially above 250 km. All the three quantities (N_E, T_E, and T_I) show oscillations with periods of tens of minutes to an hour, which comply with the periodicity of the magnetospheric forcing (Figures 5.2 and 5.3).

We illustrate T_I, T_E, and N_E from the four TIEGCM runs at Poker Flat in Figures 5.8-5.10, respectively. Figures 5.8a, 5.9a, and 5.10a show the results from the default TIEGCM run (R1) driven by empirical aurora and empirical electric fields. The R1 simulation fails to reproduce the observed features for all three quantities: T_I and T_E do not show significant increases, and N_E does not capture the secondary peak compared to the PFISR observations. From Figure 5.8, the differences between R1 and R2, and R3 and R4, are trivial, which suggests that electric fields are the most
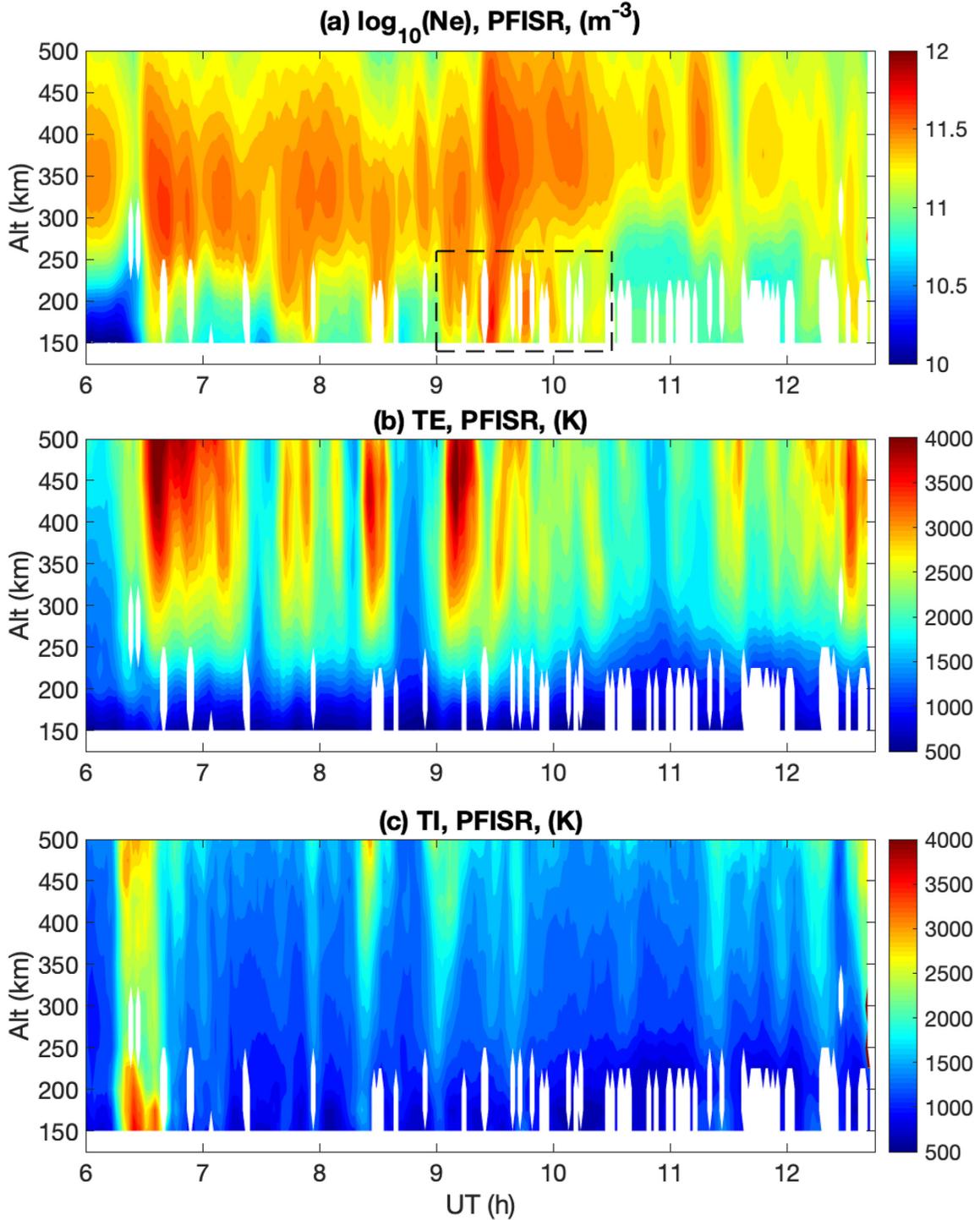
Figure 5.7: (a) Electron density (N_E), (b) electron temperature, and (c) ion temperature observed by PFISR on 2015-03-17. The dashed black rectangular in panel (a) highlights the occurrence of the secondary peak in N_E.

important factor for enhanced ion temperature during the storm time. This is sensible since ion frictional heating, which dictates the ion temperature, is proportional to the square of the electric field magnitude. Both auroral particle precipitation and the electric field affect T_E (Figure 5.9). Even though auroral assimilation enhances T_E during the period of 08:30:00 to 11:00:00 around 250 km (Figure 5.9b), the most significant heating effect originates from the assimilation of the electric field, by comparing Figures 5.9a, 5.9c-d.
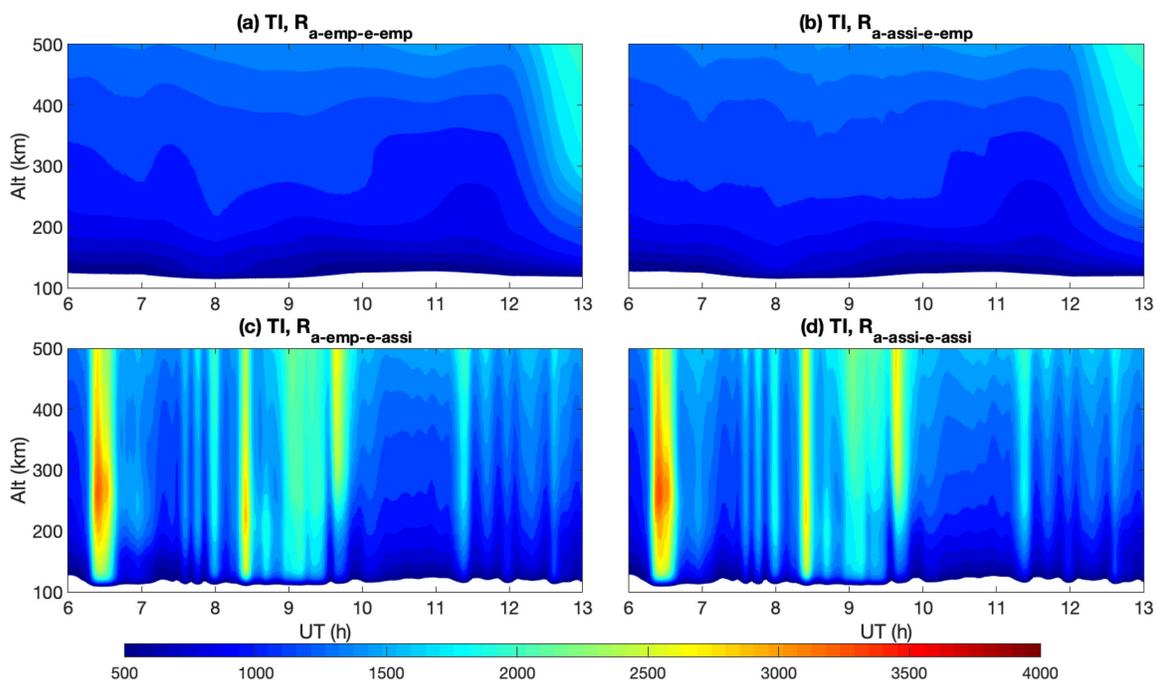
Figure 5.8: TIEGCM simulations of ion temperatures at Poker Flat from the four different runs. Unit is K.

N_E can be affected by both aurora and electric field (Figure 5.10), even though aurora seems to be the key for the formation of the secondary peak in the upper E and lower F regions (Figures 5.10b and 5.10d). The auroral assimilation leads to more substantial particle precipitation locally, which increases the ionization rate and thus causes the enhancement in electron densities. By comparing Figures 5.10a with 5.10c and 5.10b with 5.10d, which differ solely by electric fields, the empirical electric field
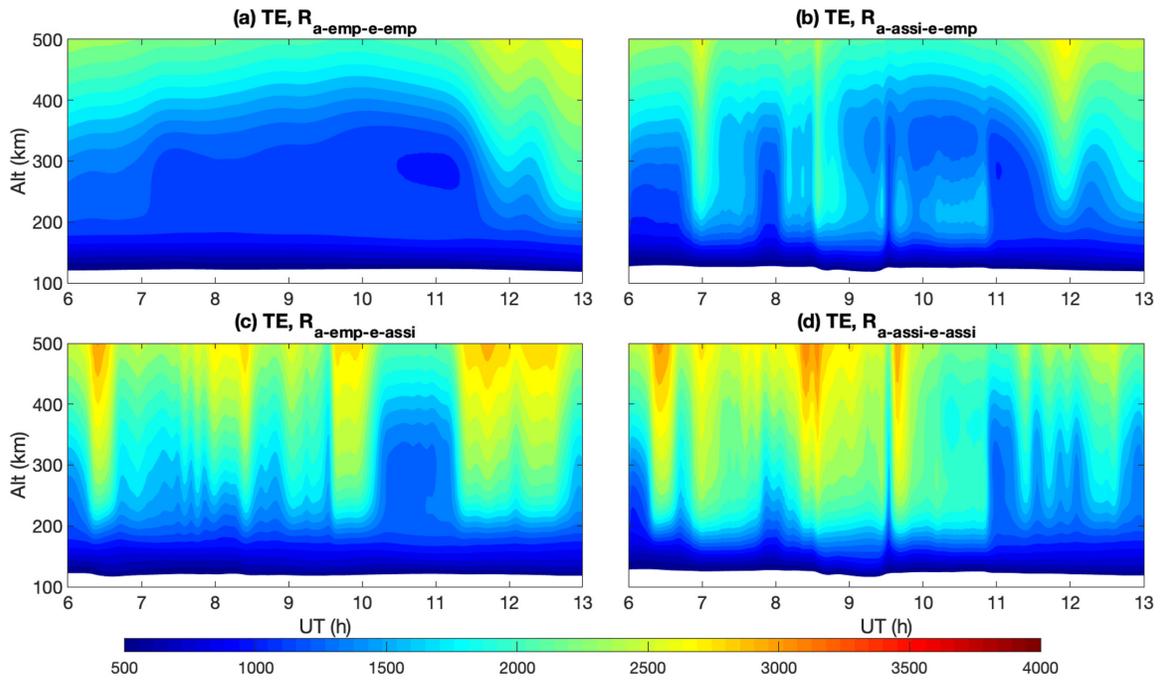
Figure 5.9: Similar to Figure 5.8 except for electron temperatures (K).
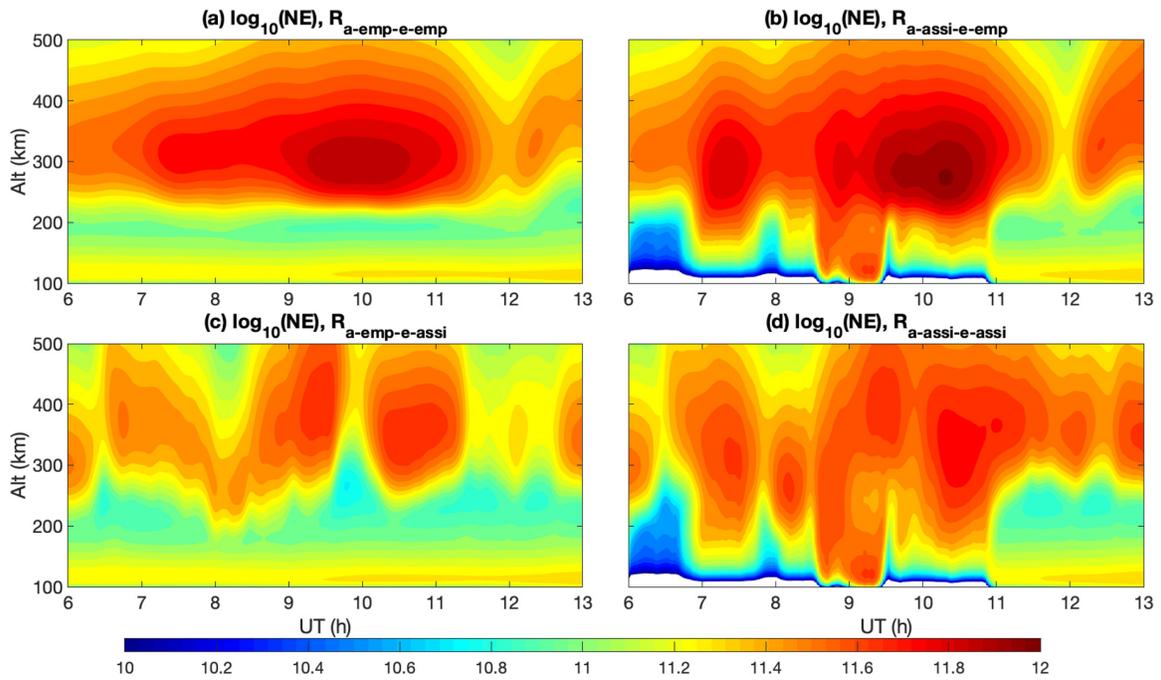
Figure 5.10: Similar to Figure 5.8 except for electron densities (m^{-3}).

model (Weimer in this case) tends to overestimate N_E. There are two processes potentially affecting local electron densities at Poker Flat: (a) electric fields lead to Joule heating, causing an upwelling in the thermosphere, changing composition with a depleted O/N_2, and (b) electric fields can transport the plasma in/out of the region to introduce temporal variations and therefore change local electron densities.

We compare the neutral vertical winds in Figure 5.11. We first compare the simulations at Poker Flat from the four different runs (Figures 5.11a1-a4). From the runs driven by the empirical electric fields (R1 and R2), vertical winds are relatively small below 300 km. The assimilation of electric fields introduces stronger vertical winds and more temporal variability (Figures 5.11a3-a4). R4 assimilates aurora and electric field and generates the strongest vertical winds below 300 km. In particular, the enhancement of vertical winds of the order of $40 \, \mathrm{m \, s^{-1}}$ during the period of 08:30:00 to 09:30:00 below 300 km is only captured by R4, implying that both aurora and electric field are responsible for the strong upward vertical winds.

We compare the neutral vertical winds from R1 and R4 for the three FPI stations (Eagle, Toolik, and Kaktovik) in Figures 5.11b1-b2, 5.11c1-c2, 5.11d1-d2, respectively. Like Poker Flat, the vertical winds are significantly strengthened and show temporal variabilities. During the period of 08:30:00 to 10:00:00, strong upwelling winds are found, especially for Toolik, which reaches $60 \, \mathrm{m \, s^{-1}}$ and are pretty comparable with observations (Figure 5.3c). Note that the local measurements of aurora and electric fields were unavailable to assimilate the FPI locations. Considering that magnetospheric energy input can be structured, the exact matching with the FPI observations has not required in this study. The magnitude and variability in neutral vertical winds are improved overall.

According to Figure 5 in Wu et al. (2020), the intense Joule heating in the 100 km to 150 km region causes a significant increase in the neutral temperature
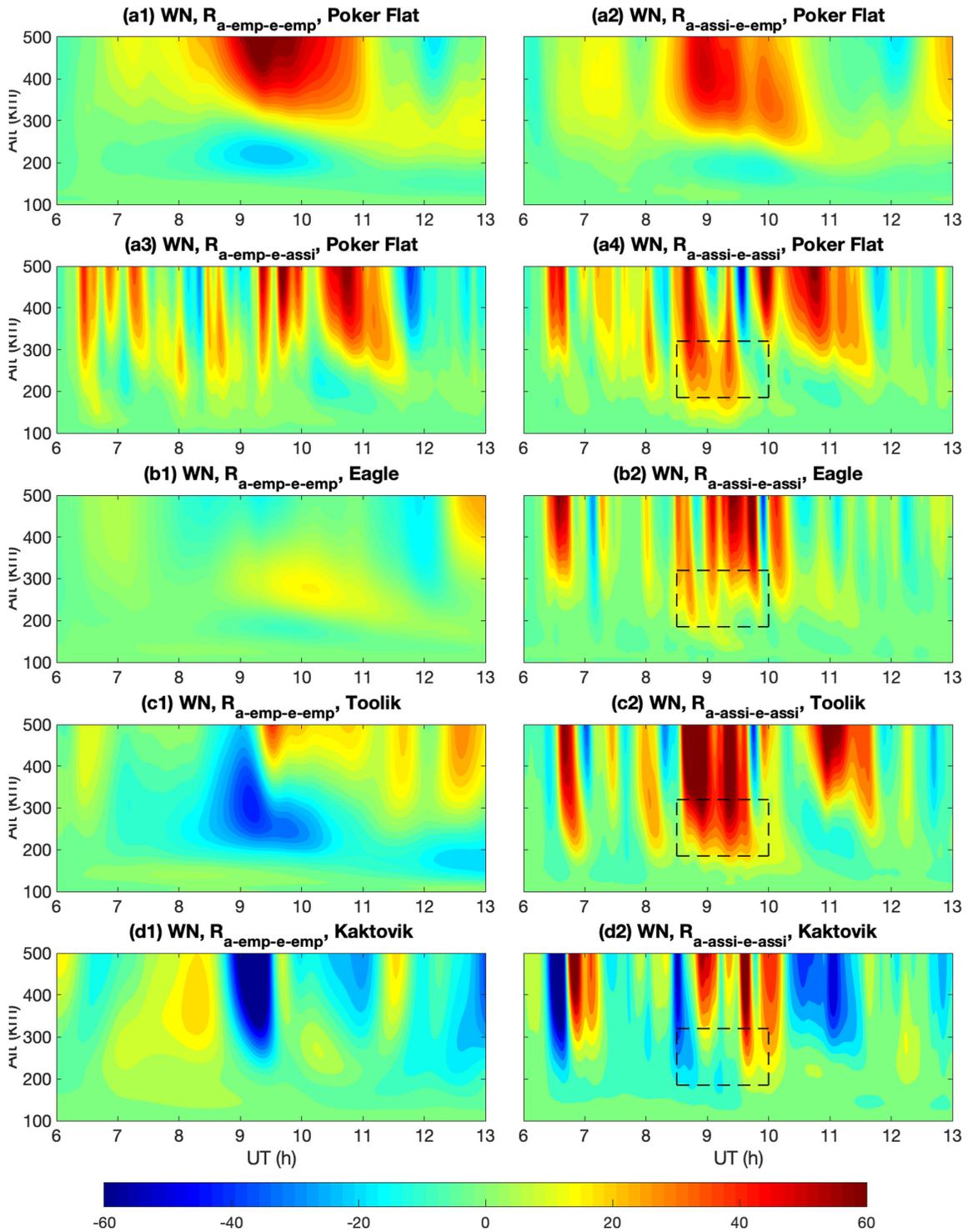
Figure 5.11: Similar to Figure 5.8 except for vertical winds $(\mathrm{m\,s^{-1}})$. The black dashed rectangle highlights the period with enhanced vertical winds around $250\,\mathrm{km}$.

above this altitude range. This thermal enhancement is responsible for TTEIL and the generation of strong vertical winds $(100\,\mathrm{m\,s}^{-1})$ above. To simulate the TTEIL and produce a better agreement with the lidar observations, the authors had to incorporate the realistic SSUSI auroral precipitation and sub-grid electric field variability to gain the necessary magnitude of Joule heating. In the current study, we implement the data assimilation of aurora and electric fields for the same purpose: to reproduce more realistic Joule heating locally and the effects on neutral temperatures and winds.

Figure 5.12 shows the comparisons of Joule heating rate per unit mass $(\mathrm{W\,kg}^{-1})$ at Poker Flat from the four runs and for the three FPI stations from R1 and R4. The differences are introduced mainly by electric fields. The Joule heating rate also shows short-term temporal variations (order of minutes). Significant Joule heating around 06:30:00 (Figures 5.12a3-a4) associated with the strong E_y (red line in Figure 5.3b) explains the sharp increases of T_I at the same time observed by PFISR (Figure 5.7c) and simulated by the run with assimilation (R4, Figure 5.8d). Compared with the Joule heating around 06:30:00, which mainly concentrates in the F region, the strong auroral activity around 09:00:00 leads to the relatively strong Joule heating in the lower E region (red rectangles in Figure 5.12), which corresponds reasonably well with the strong upwelling above (black rectangle in Figure 5.11). By cross-comparing Figures 5.12a1-a4, both aurora and electric field, which become strong around 09:00:00, are responsible for the increase of the magnetospheric energy deposition, inducing the strong upward vertical wind. Such a heating process is regional and partially responsible for the strong upward winds in the other three stations.

We compare the height-integrated Joule heating rate at Poker Flat and the maximum integrated Joule heating rate at high northern latitudes ($> 50°$N) in Figure 5.13. The significant local enhancement associated with the data assimilation at Poker Flat is salient, which indicates that the default model run driven by empirical

155
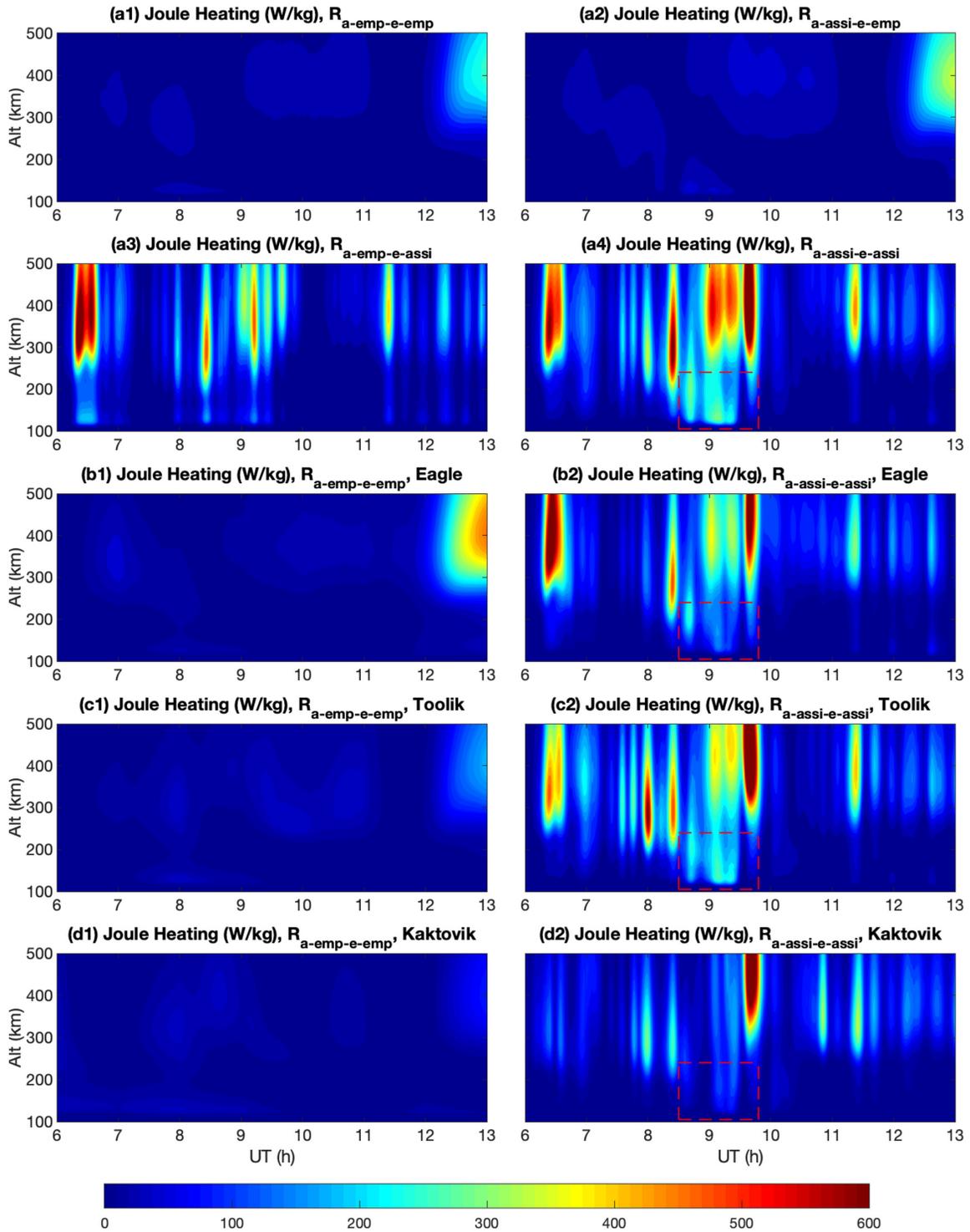
Figure 5.12: Same as Figure 5.11 except for altitude-resolved Joule heating. The red dashed rectangle highlights the period with the significant Joule heating to the E region.

magnetospheric drivers misses the characterization of localized energy deposition. The magnitudes of the maximum height-integrated Joule heating rate over the globe ($> 50°$N), however, are more comparable among the four different runs, except that the assimilation runs show stronger peaks and more temporal variability driven by observational data. Such comparisons suggest that the data assimilation has guided the spatial distribution of energy deposition and led to different locations that peak in Joule heating. The large Poynting flux exceeding $100\,\mathrm{mW\,m^{-2}}$ has been observed by the DMSP F13 during a strong magnetic storm (Huang and Burke, 2004). Knipp et al. (2011) has also reported a day-side Poynting flux as large as $170\,\mathrm{mW\,m^{-2}}$ occurring during the period of high-speed solar winds. It is noted that the majority of Poynting flux is transferred to Joule heating in the upper atmosphere (Lu et al., 1995; Thayer and Semeter, 2004). Thus the magnitude of the integrated Joule heating rate simulated in this study is within a reasonable scope.

To further quantify the improvements introduced by the assimilated drivers compared to the empirical ones, we provide two metrics, maximum values, and RMSEs, for N_E, T_E, T_I, and W_N (Table 5.3.1). For N_E, T_E, and T_I, the RMSEs are calculated from the differences between each model run and PFISR observation. All the time-altitude points with valid observations are considered. For the RMSE of W_N, the model data at $250\,\mathrm{km}$ is extracted first, and the difference from FPI measurements is used for calculation. We use Toolik as an example in Table 5.3.1. There is a clear trend that maximum values increase and become the closest to the observations as both aurora and electric fields are assimilated. At the same time, RMSEs decrease by $30\,\%$ to $50\,\%$ depending on physical quantity, indicating that the model better matches observations, primarily as the electric field assimilation is implemented.
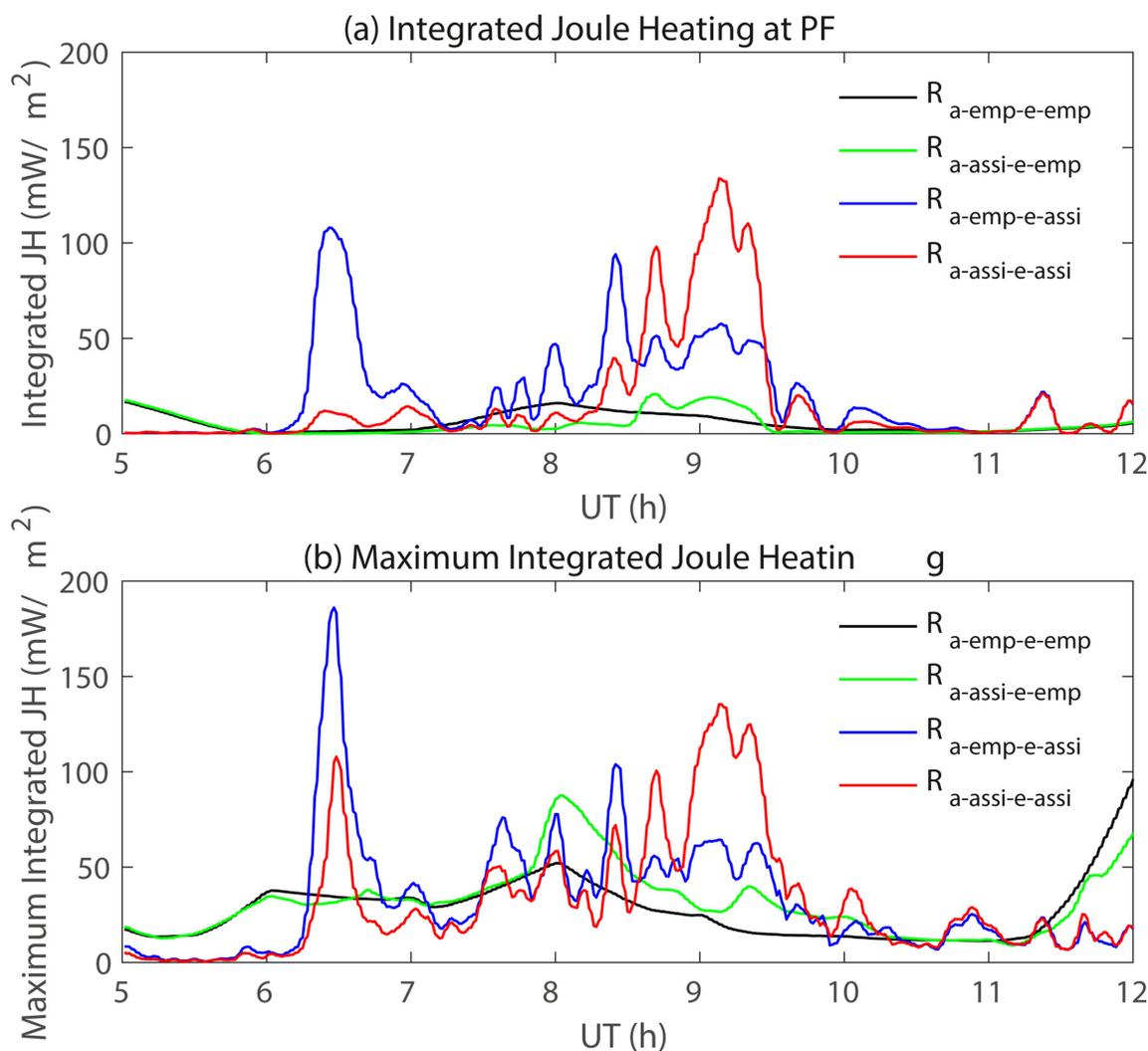
Figure 5.13: (a) Comparison of height-integrated Joule heating $(\mathrm{mW\,m^{-2}})$ at Poker Flat from the four TIEGCM runs. (b) Same as (a) except for the maximum height-integrated Joule heating at high latitudes $(> 50°\mathrm{N})$ of the northern hemisphere.

		Run 1	Run 2	Run 3	Run 4	Observation
N_E $(\mathrm{m^{-3}})$	Max	2.1×10^{11}	4.7×10^{11}	2.0×10^{11}	4.7×10^{11}	5.2×10^{11}
	RMSE	2.0×10^{11}	2.5×10^{11}	1.0×10^{11}	1.5×10^{11}	
T_E (K)	Max	2760	2710	3080	3260	4440
	RMSE	763	681	486	505	
T_I (K)	Max	1980	2010	3350	3400	3310
	RMSE	458	445	342	330	
W_N $(\mathrm{m\,s^{-1}})$	Max	3.6	10.0	48.5	58.8	84.5
	RMSE	47.7	38.4	30.0	23.3	

5.3.2 Regional to Global Neutral and Ionospheric Responses

Since the data assimilation considers SSUSI and THEMIS for aurora and SuperDARN for electric fields with substantial spatial coverage, it should improve regional or global simulations. Figure 5.14 shows the TEC distributions from the four runs compared with the GNSS observations. With the realistic aurora (R2 and R4), the TECs from Alaska circling to the Great Lake regions are elevated, which follows the enhancement of auroral energy flux shown in Figure 5.4. R2 tends to overestimate the TEC in Alaska. At the same time, R4 matches the GNSS observations the best, especially in the magnitude, consistent with the local simulation of TEC that empirical electric field run leads to a larger N_E (Figure 5.10). R4 mostly captures the temporal evolution of TECs as observations, and such improvement lasts for the whole period of the storm.

A better representation of the time, location, and strength of energy deposition with data assimilation leads to the differences in the simulated neutral dynamics and winds from different runs. To explore this, we examine the signal of TADs from neutral vertical winds. According to the dispersion relation of gravity waves, vertical wind can better manifest higher-frequency and smaller-horizontal scale waves than horizontal winds and temperature (Lu et al., 2015, 2017; Vadas, 2013). Figure 5.15 shows the vertical winds from R1 and R4 at 250 km and 550 km altitudes at 09:45:00. The TADs are always stronger and can reach lower latitudes in R4 than R1, especially at 550 km. This means that the wave source for TADs is stronger, and neutral responses to the storm are more dynamic in R4, characterized by richer scales and broader extension, than those captured by the default model run. To quantify such effects, we compute the longitudinal variability (SD along longitude) of vertical winds as a function of latitude and UT and temporal variability as a function of
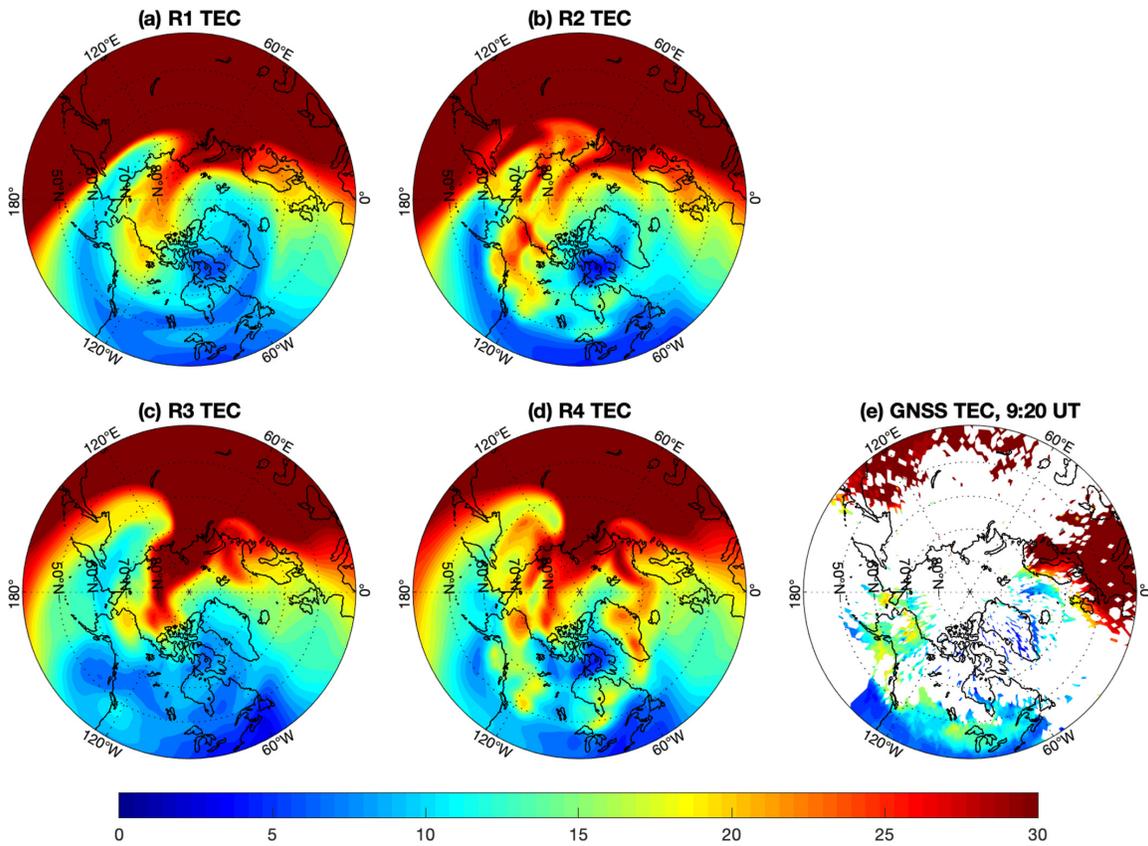
Figure 5.14: Model simulations of TEC (unit: TECu) compared with GNSS TEC observations at 09:20:00. (a-d) TEC simulation from R1, R2, R3, and R4, respectively. (e) GNSS TEC observations at 09:20:00.

latitude and longitude. Figure 5.16a1-d1 show longitudinal variabilities for the four runs. Figure 5.16e1 shows the time-averaged means of the longitudinal variabilities from the four runs. Figure 5.16a2-d2 show temporal variabilities of vertical winds, and Figure 5.16e2 shows their zonal means. The assimilated electric fields are more important than the aurora in increasing spatial and temporal variabilities. For the time-averaged spatial variability (Figure 5.16e1), R4 shows a larger magnitude by a factor of 1.5 to 3 compared to R1 (default run), depending on latitude. And for the zonal-mean temporal variability (Figure 5.16e2), R4 is also larger than R1 by a factor of 1.5 to 3. Such ratios suggest that the wave activities (TADs) are stronger and embrace more short-term temporal and small-scale spatial scales in the data-assimilation run than in the default one.

5.4 Discussion and Conclusions

We apply the newly implemented Lattice Kriging model to assimilate aurora and electric fields for the 2015 St. Patrick's Day storm. We use them to drive TIEGCM and produce more realistic I-T system responses than the default run driven with empirical high-latitude inputs. The improvement of model simulation enabled by data assimilation includes (a) higher T_I and T_E (up to $3000\,\text{K}$) due to the enhancement of local heating; (b) larger N_E below $200\,\text{km}$ and the emergence of the secondary peak in the upper E and lower F region around 09:00:00 when auroral surges occur; (c) stronger neutral vertical winds around $250\,\text{km}$ due to significant Joule heating below that caused by both the intensified aurora and enhanced electric fields; (d) enhanced TECs around auroral oval because of the auroral particle precipitation; and (e) a larger amplitude of TADs and a further extension to low latitudes likely due to the elevation of local heating and ion drag as a wave source; and (f) short-term
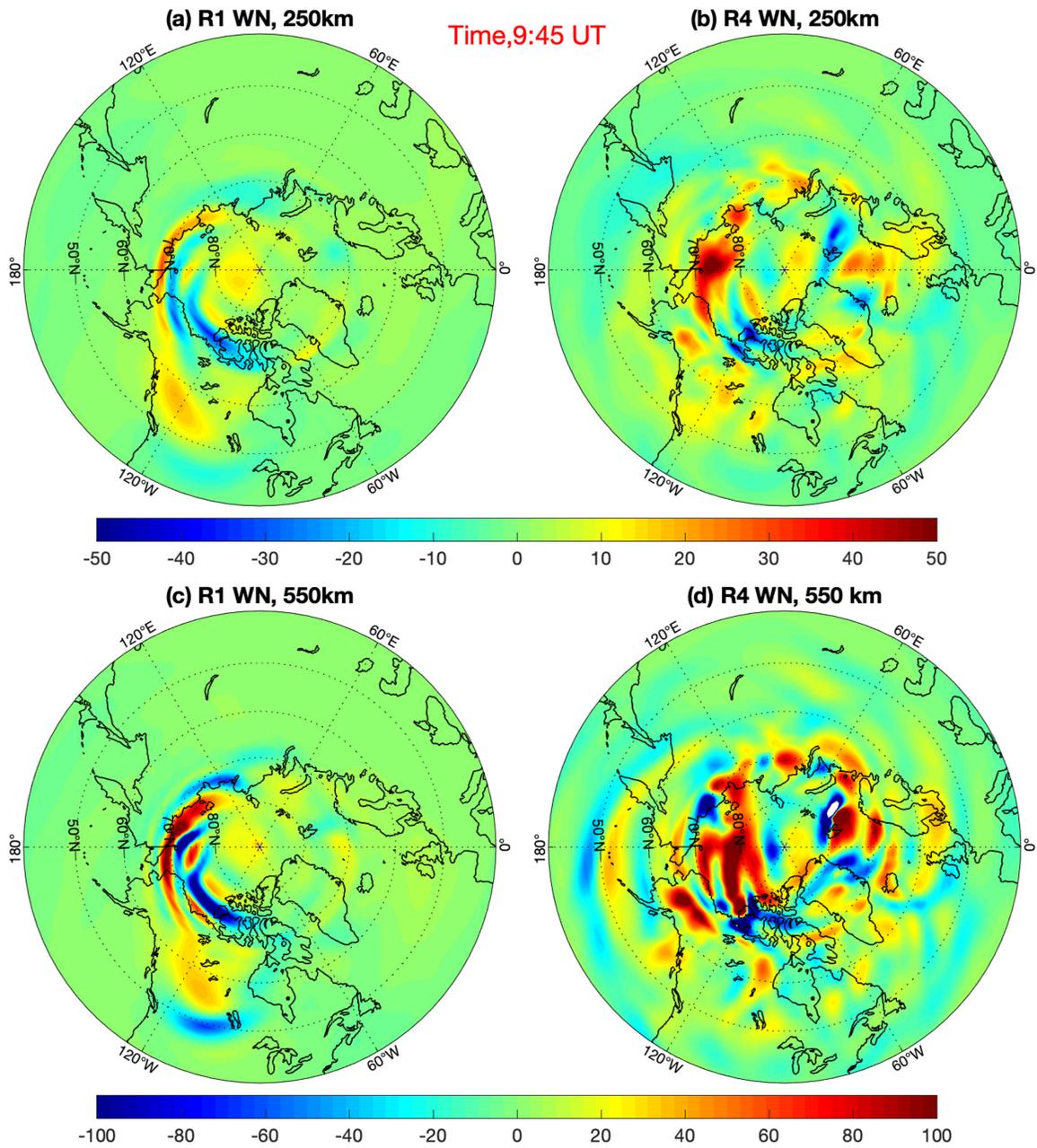
Figure 5.15: (a, b) Neutral vertical winds at $250\,\mathrm{km}$ from R1 and R4, respectively. (c) and (d) are the same except at $550\,\mathrm{km}$. Unit is $\mathrm{m\,s^{-1}}$.
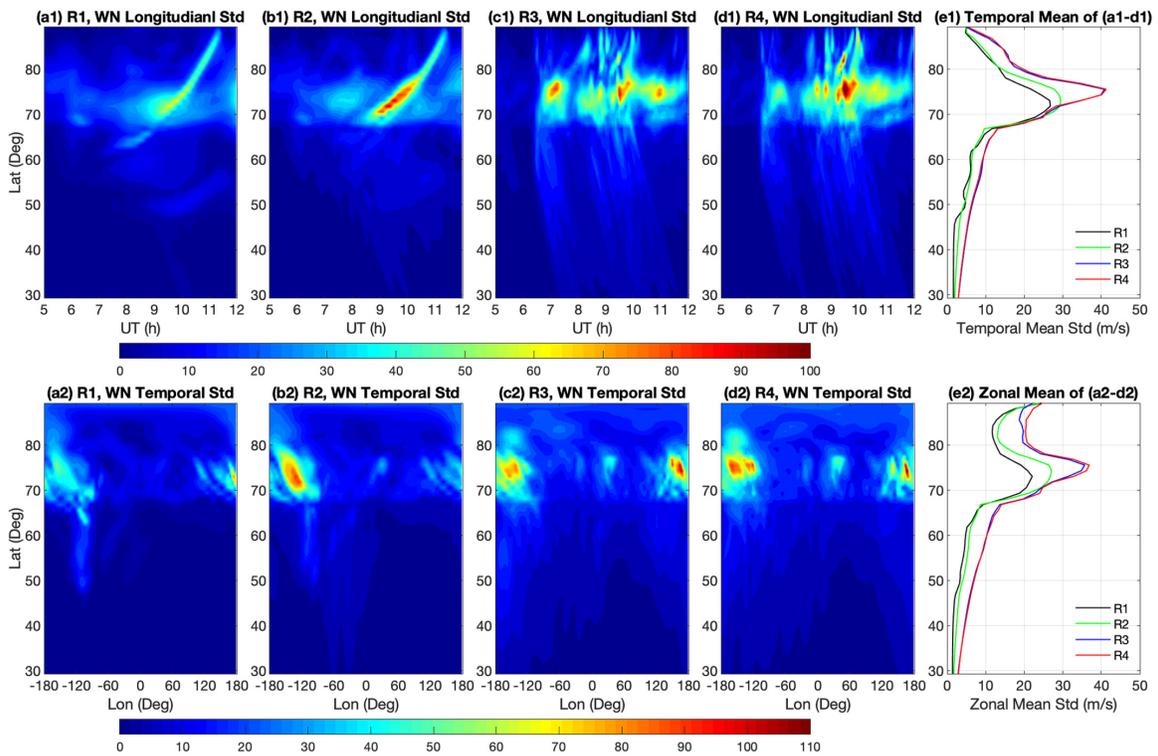
Figure 5.16: (a1-d1) Longitudinal variability (calculated as SD along longitude) of neutral vertical wind as a function of latitude and UT at 550 km. (e1) Temporal means of panels (a1-d1). (a2-d2) Temporal variability of neutral vertical wind as a function of latitude and longitude at 550 km. (e2) Zonal (longitudinal) mean of panels (a2-d2). The unit is m s^{-1}.

163

temporal variability more comparable with the PFISR and FPI observations. In addition, the maximum values of N_E, T_E, T_I, and W_N responses are better captured in assimilated model runs. The RMSEs calculated from the differences between model results and observations decrease by 30% to 50% compared to the default run. All these improvements show that the I-T responses to the St. Patrick's Day storm, in reality, have a more extensive dynamic range and more variability than those simulated by the default TIEGCM run with empirical driving conditions. Thus, we have shown that data assimilation of high-latitude drivers can help us better understand the storm's impacts locally and globally. Even though the data assimilation run is a significant improvement, the remaining discrepancies with observations include that the modeled T_E is about $1000\,\mathrm{K}$ colder than the PFISR observation in the F region, which needs further investigation.

Once the model drivers are refined with aurora and electric field data, and the outputs are systematically evaluated and found to be in better agreement with observations, this data-constrained model run is then used to diagnose the changes in the system responses, for which we use the longitudinal and temporal variability of neutral vertical winds as a proxy. Compared to the default run, the data assimilation run has shown stronger longitudinal (spatial) and temporal variabilities than the default run by a factor of 1.5 to 3. Since the wave's kinetic energy is proportional to the square of wave perturbations, this would increase the kinetic energy received by the upper atmosphere by an order of magnitude. Such enhancement is seen in the auroral regions and prominent at middle and low latitudes, likely via the generation of TADs at high latitudes and then equatorward wave propagation. In addition, the refinement of high-latitude drivers can change disturbance dynamo and penetration electric fields, influencing low-latitude and equatorial electrodynamics. Detailed analysis of the model performance at middle and low latitudes and comparison with

observations there are worthwhile and deserve a future investigation.

Our work highlights the importance of observing the storm-time manifestation of magnetospheric drivers and I-T responses simultaneously to understand better neutral-ion coupling in the upper atmosphere, a prerequisite for the predictability of space weather. Such observations can also help constrain data assimilation models and evaluate their capabilities and limitations. The quality of data assimilation still highly depends on observations. Multi-point and simultaneous observations of ion drifts, auroral precipitation, electron densities, neutral responses such as density and wind, and temperature, such as those solicited by the Geospace Dynamics Constellation (GDC) mission, are highly desirable.

Chapter 6

Simulating GW Propagations in the I-T System Using Nested-Grid TIEGCM

In reality, data coverage is uneven, and what usually happens is that high-resolution observations occur in a restricted region while globally, low-resolution observations dominate. Then in terms of global modeling, when comparing model results with observations, low-resolution modeling cannot take advantage of high-resolution observations, and high-resolution modeling consumes too many unnecessary computation resources in the region where there are only low-resolution observations. Is there a way to distribute more computing resources into the region where high-resolution observations are available while only necessary resources are in other regions? This leads to developing a hybrid-grid approach to the global models where a low-resolution grid is used globally, and regionally, a high-resolution grid is deployed (nested grid). The design and application of the nested-grid TIEGCM (TIEGCM-NG) will be discussed in this chapter.

The Hunga Tonga-Hunga Ha'apai volcano eruption on 2022-01-15 triggered intense atmospheric GW activity in the upper atmosphere, which was later detected by various observations. We perform one of the first I-T model studies of the GW effects from the Tonga eruption in the I-T system using the TIEGCM. We implement a high-resolution mesh inside a regional domain (nested grid) in addition to the global low-resolution mesh, which differs from the standard global uniform resolution setup. TIEGCM-NG successfully simulates the observed wave propagation and effects in the I-T system by further nudging GW fields at TIEGCM lower boundaries (97 km) using output from the high-resolution WACCM-X simulations. The simulation results indicate that horizontal resolution is the critical parameter to simulate GW propagation. Inside the high-resolution nested region, GWs with horizontal wavelengths of 400 km and periods of 10 min to 30 min can propagate outward and upward and produce significant ionospheric disturbances close to observations. Outside the nested region, only long-wavelength, low-frequency waves survive. Another test indicates that GWs can be better resolved when geopotential height is nudged at TIEGCM lower boundaries. With the capability of simultaneously simulating local, small to mesoscale I-T processes, TIEGCM-NG is superior to global high-resolution simulations due to its primarily reduced computation cost and may find its application in the study of I-T system regional dynamics.

6.1   Introduction

Atmospheric GWs play an essential role in energy and momentum coupling from the lower to the upper atmosphere (Fritts, 1984; Fritts and Alexander, 2003). GW sources are manifold, including topography, convection, wind shear, geostrophic adjustment, body force, and wave-wave interactions (Nastrom and Fritts, 1992; Fritts

and Nastrom, 1992). GWs can propagate to the mesosphere and thermosphere, which leads to changes in the energy and momentum budget (Vadas and Fritts, 2005; Alexander et al., 2010). GW impacts manifest in the ionospheric E and F regions as TIDs (Hines, 1960; Fritts and Lund, 2011) or sporadic E layers (van Eyken et al., 1982). The upward propagating GWs may dissipate or break and transfer momentum and energy to the mean background winds, making it an efficient way to connect the lower atmosphere source region to the upper atmosphere (Lindzen, 1981; Holton, 1982; Fritts and Alexander, 2003). It is also crucial in numerical models to correctly account for the GW effects to produce temperature and wind profiles comparable to observations (Holton, 1983; Palmer et al., 1986; McFarlane, 1987; Vadas and Liu, 2009).

On 2022-01-15, a volcano erupted at Hunga Tonga-Hunga Ha'apai (20.5°S, 175.4°W) and produced a vertical plume over 30 km tall with a top above 55 km (Carr et al., 2022). The huge energy release in the initial explosion (10 EJ to 28 EJ) and the subsequent plume triggered atmospheric waves across a wide range of wavelengths and frequencies. The eruption also ejects a large amount of water vapor (150 Tg) into the stratosphere, which acts as a secondary wave source in the form of latent heat release as reported by Millan et al. (2022) and Zhu et al. (2022). The generated waves propagated over long distances and were reported worldwide from the surface to the ionosphere by satellite and ground-based observations (Wright et al., 2022). Large neutral wind variations were captured by MIGHTI onboard ICON, causing the equatorial electrojet to switch directions from eastward to westward (Harding et al., 2022). SABER recorded the GW activity due to the eruption as temperature variations were enhanced by a factor of 1.5 compared to the days before and after the eruption (Liu et al., 2022b). Ern et al. (2022) analyzed the temperature profiles from the Atmospheric Infrared Sounder (AIRS) and the Microwave Limb Sounder (MLS)

and found strong mesoscale GW signals. They performed a ray tracing experiment using the Modern-Era Retrospective analysis for Research and Applications (MERRA) 2 as the background atmosphere. The ray tracing results showed the source was near the volcano, which suggests that the GW activities were strongly correlated to the volcano eruption.

There was a moderate geomagnetic storm (Kp = 5.7) the day before the eruption (2022-01-14), and the geomagnetic field impacts from the volcano eruption were examined by Schnepf et al. (2022); Yamazaki et al. (2022). Concentric ring structures of differential TEC indicating the impact from the eruption-induced outward propagating GWs were reported by Themens et al. (2022) and Zhang et al. (2022b). The geomagnetic conjugate effects of the ionospheric perturbations caused by the volcano-triggered GWs were reported by Lin et al. (2022). Aa et al. (2022) pointed out that the depletion of local TEC measurements could reach as high as 10 TECU to 15 TECU. Volcano effects manifested in the range-time-intensity plots of the Super-DARN observations as rapid oscillations in the LOS Doppler velocities (Zhang et al., 2022a). In-situ ion density and drift measurements from Ion Velocity Meter (IVM) onboard ICON also showed clear signatures of GW impacts (Gasque et al., 2022).

Modeling efforts have been conducted to study the interaction between atmospheric acoustic GWs and tsunamis at the ocean surface (Omira et al., 2022; Lynett et al., 2022). Inchin et al. (2022) solved the 3-D nonlinear compressible Navier-Stokes equations using Model for Acoustic-Gravity wave Interactions and Coupling (MAGIC) and confirmed that tsunami-generated acoustic GWs could directly influence the mesopause airglow. Amores et al. (2022) utilized Semi-implicit Cross-scale Hydroscience Integrated System Model (SCHISM) to study the fast-moving Lamb waves excited by the Tonga volcano eruption, and the simulation results agreed well with the arrival time measured by satellite observations.

However, the modeling efforts of GW propagation in the upper atmosphere and its impact on the ionosphere remain sparse. We present a modeling study of the volcano-triggered GWs and their effects in the I-T system using TIEGCM with a nested grid extension. TIEGCM self-consistently solves the continuity, momentum, thermodynamic, and ion transport equations from 97 km to 600 km (Qian et al., 2014). TIEGCM-NG is an extension of TIEGCM version 2.0, which solves the same complete set of first-principle equations inside the regional domain. It provides a high-resolution regional description of the I-T system and extends the capability of studying neutral and electro-dynamical processes to mesoscales. The multi-level nesting capability provides an even finer view of the I-T processes and enables comparative studies with the local high-resolution observations. TIEGCM-NG is more computationally efficient for studying mesoscale to small-scale I-T processes than the globally high-resolution model. In this study, we add the nudging capability, an efficient approach to imposing subscribing waves propagating from below.

TIEGCM-NG is inspired by TING (Wang et al., 1999), a nested-grid extension built upon TIGCM (Roble et al., 1988). TIGCM is the predecessor of TIEGCM, in which electrodynamics were missing. TIEGCM self-consistently solves electrodynamic coupling processes (Richmond et al., 1992), can better simulate the I-T system than TIGCM, and produces output with better agreements with observations. That leads to TIEGCM-NG, an updated version of TING with a complete set of physical processes of self-consistently resolved wind dynamo. TIEGCM-NG also extends the nesting capability to allow for arbitrary nesting levels compared to a maximum of 2 levels in TING.

This paper is organized as follows: Section 6.2 describes the design of TIEGCM-NG in detail and the setup of control experiments. Section 6.3 discusses the model results. We conclude this study and briefly discuss the possible application of this

new model in Section 6.4.

6.2 Model Description and Experiment Design

The design of TIEGCM-NG largely follows the original TIEGCM version 2.0 (Richmond et al., 1992; Qian et al., 2014). The physical equations, the same as those in the global domain, are solved separately in the nested domain. The vertical coordinate system is a log-pressure coordinate,

$$z = \ln \frac{p_0}{p}$$

where $p_0 = 5 \times 10^{-5}\,\mathrm{Pa}$. The lower and upper boundaries are at $z = -7(p = 5\,\mathrm{e}^7 \times 10^{-5} = 0.55\,\mathrm{Pa})$ and $z = 7(p = 5\,\mathrm{e}^{-7} \times 10^{-5} = 4.55 \times 10^{-8}\,\mathrm{Pa})$, respectively. The temperature (T), zonal and meridional winds (U and V), and geopotential height (Z) consist of the complete set of lower boundary conditions and are chosen to specify the lower boundary fields. The differences and connections between the global and nested grid domains in TIEGCM-NG are highlighted in Section 6.2.1. The additional functionality (nudging) is described in Section 6.2.2.

6.2.1 Design of TIEGCM-NG

The spatial design of TIEGCM-NG is illustrated in Figure 6.1a, which shows a 3-level nesting scheme labeled by g (global), n1 (1st level nested grid) and n2 (2nd level nested grid). The nested domain is designed as a rectangular region in latitude and longitude that can cover an arbitrary portion of the globe except for the poles. The field mapping between the global and nested domains is achieved by a 2-D spatial interpolation which allows for the maximum flexibility of the position of the nested

grid. When TIEGCM-NG starts, the nested fields are initialized with the global fields. Then, they iterate without re-initialization from the global fields, which means that the field mapping from the global to the whole nested domain happens only once per model run at the model's first timestep. The lateral boundary conditions of the nested grid domain are chosen to be the time-dependent Dirichlet condition, i.e., the fields at the boundary are pre-obtained from the global domain at every nested grid time step. The boundary condition's time dependence guarantees the maximum resemblance of the global and nested fields. The information exchange between the global and nested domains follows different procedures from outside inward (global to nested) and from inside outward (nested to global):

1. To pass information from the global into the nested domain (inward), boundary mapping is performed. The global fields are extracted and interpolated to the nested grid boundaries at every nested grid time step to be used as the lateral boundary conditions of the nested domain.

2. To pass information from the nested to the global domain (outward), direct field mapping is performed. At every global time step, the global fields inside the nested domain are replaced by the nested fields interpolated onto the global grids. Outside the nested domain, the global fields remain untouched.

In terms of the horizontal finite difference scheme, instead of using a 4-point-centered finite difference (4th order accuracy) uniformly in the global domain, a 2-point-centered difference (2nd order accuracy) is used at the lateral boundary while a 4-point centered difference is used internally in the nested grid domain. Numerical tests show minimal resulting numerical noise (not shown here). The boundary reflection problem, sometimes present in the GW vertical propagation near the upper boundary (Klemp and Durran, 1983), does not show up in the GW horizontal

172
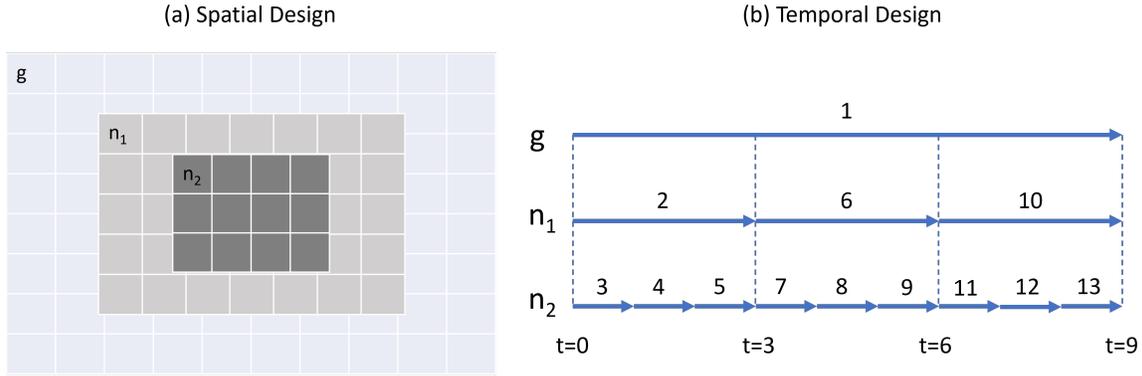
Figure 6.1: Schematic diagrams showing the spatial (a) and temporal (b) design of a 3-level nested grid. g stands for global, n1 stands for the 1st level nested grid, and n2 stands for the 2nd level nested grid. Numbers in (b) represent the workflow of TIEGCM-NG.

propagation close to the nested grid lateral boundary in this study.

Due to the decreased grid sizes in the nested domains compared to the global domain, it is desirable to decrease the time step to retain numerical stability correspondingly. Sub-cycling in time is used to achieve numerical stability in the nested domain. The time integration of TIEGCM-NG is illustrated in Figure 6.1b. A 3-level nesting scheme with a 3-step sub-cycling demonstrates the advancing procedure. The numbers in Figure 6.1b marks the proceeding sequence in each domain. The details in each step are as follows:

1. The time iteration starts from the global domain. Within one iteration (time step), the global domain advances from $t = 0$ to $t = 9$. The 1st level boundaries are extracted from the global field at these two timestamps ($t = 0$ and $t = 9$) and temporally interpolated to every sub-cycle ($t = 3$ and $t = 6$) for iteration.

2. Then the 1st level is advanced from $t = 0$ to $t = 3$. Repeating the inward boundary mapping procedure from the global domain to the 1st level, the 2nd level nested grid boundaries are extracted from the 1st level at $t = 0$ and $t = 3$.

173

Again, the 2nd level lateral boundary conditions are interpolated in time at the 2nd level sub-cycle ($t = 1$ and $t = 2$).

3. After that, the 2nd level is advanced from $t = 0$ to $t = 1$.

4. Since there are no higher-level nested domains, the 2nd level is then advanced from $t = 1$ to $t = 2$.

5. Again, the 2nd level is advanced from $t = 2$ to $t = 3$. Now it is the end of the 2nd level iteration, and an outward field mapping happens from the 2nd to the 1st level. At $t = 3$, the 1st-level fields inside the 2nd-level nested domain are replaced by the 2nd-level fields.

6. Then, the 1st level iterates from $t = 3$ to $t = 6$ (steps 6 to 9) and from $t = 6$ to $t = 9$ (steps 10 to 13), which are repetitions of what has happened from $t = 0$ to $t = 3$. Then at $t = 9$, after the 2nd level updates the 1st-level fields, they are further mapped back to the global domain to update the global fields. This finishes the iteration.

Such time iteration is repeated for the next cycle ($t = 9$ to $t = 18$), and the model is advanced.

Besides the differences in the spatial setup and time integration highlighted above, another difference between the global and the nested domain is the low latitude electric fields, which are directly obtained from the interpolation of the global electric fields from the global grid instead of solving the electro-dynamo equation in the nested domain. In other words, the electric fields in the global domain, derived from the global electro-dynamo equation, are used as an additional input to the nested domain.

From the perspective of computation efficiency, TING is generally slow due to its serial programming design, which is unsuitable for current multi-processor computers. TIEGCM-NG is fully parallelized to better use the available computation resources and runs much faster than TING. The efficiency of TIEGCM-NG is tested through a series of different numbers of nested grid points and core processors. The profiling of the parallel implementation indicates that the time cost grows linearly with the nested grid number, and the additional time spent on parallelism is minimal.

TIEGCM-NG also leads to a significant improvement in computation efficiency as compared to the globally uniform high-resolution TIEGCM. To perform a run with the highest resolution of 1°, TIEGCM-NG can be faster as much as 3 times based on a 20-core machine. For higher-resolution runs, more time is saved by utilizing multi-level nesting. The reduction of computation cost is substantial, and it becomes affordable to perform high-resolution TIEGCM-NG runs on workstations.

6.2.2   Nudging of the WACCM-X Fields

Specific wave sources need to be subscribed at the lower boundary to study the propagation of GWs originating from the lower atmosphere in TIEGCM-NG. However, due to the sparse observations near the lower boundary altitude (pressure level $z = -7$, $97\,\mathrm{km}$), it is nearly impossible to provide self-consistent high-resolution lower boundary conditions from current observations. Here, we resort to numerical results from other models that can provide fields at $z = -7$. The simulation results from WACCM-X are used as the lower boundary input for TIEGCM-NG. WACCM-X self-consistently resolves the atmospheric processes from the ground to the upper thermosphere (Liu et al., 2010, 2018a,b). In this study, a high-resolution WACCM-X

SE run at 0.25° horizontal resolution and 0.1 scale height vertical resolution is used (Liu et al., 2022a). The new SE dynamical core eliminates the polar singularity of the finite volume core used in previous versions of WACCM-X. It enables much higher resolutions than it would achieve in the past. High-resolution model simulations fully embrace resolved GW generations, typically parameterized in low-resolution runs (Gettelman et al., 2019). For the 2022-01-15 volcano eruption at Hunga Tonga-Hunga Ha'apai, a large surface pressure perturbation is introduced at 04:35:00 (the reported time of eruption) to approximate the effect of the volcano eruption in the model. The magnitude of the perturbation is adjusted to match available pressure observations in other locations. The pressure disturbance-induced GWs are soon reflected in other fields and grow in amplitude vertically, leading to large variations at the lower thermosphere. It should be noted that a large amount of water vapor was emitted into the stratosphere during the volcano eruption. The corresponding release of latent heat acts as another source of GWs (Millan et al., 2022; Zhu et al., 2022). However, this secondary GW (generated by latent heat release) is not considered in this study. The only type of GWs to be studied in this paper is the one excited by the initial pressure pulse. WACCM-X outputs diagnostic fields at a frequency of one minute.

The WACCM-X fields T, U, V and Z are extracted at $z = -7 (p = 0.55\,\mathrm{Pa}, 97\,\mathrm{km})$ to be used as the lower boundary of TIEGCM-NG. In addition, to better resolve the wave propagation near the lower boundary of TIEGCM-NG, another 2 scale heights of the WACCM-X fields are nudged into TIEGCM-NG, up to $z = -5 (p = 5\,\mathrm{e}^5 \times 10^{-5} = 0.07\,\mathrm{Pa}, 110\,\mathrm{km})$. To apply WACCM-X simulation results in TIEGCM-NG, a spatial 3-D interpolation is applied to convert from WACCM-X longitude-latitude-z grids to TIEGCM-NG grids. Temporal interpolation is applied to accommodate the difference in TIEGCM-NG time steps and the WACCM-X output frequency so that WACCM-X

176

fields force TIEGCM-NG at the lower pressure levels at each time step.

The coupling function used to nudge TIEGCM-NG with WACCM-X fields is

$$f_{\text{TIEGCM-NG}} \leftarrow f_{\text{TIEGCM-NG}} + r(f_{\text{WACCM-X}} - f_{\text{TIEGCM-NG}})$$

where f is either T, U, V, or Z. As mentioned above, such a replacement is applied at every time step at TIEGCM-NG lower pressure level grids to reflect the driven state by WACCM-X. r is a vertical relaxation factor, which is chosen to decrease exponentially from $r = 1$ at $z = -7$ to $r = 10^{-4}$ at $z = -5$ to allow for a smooth transition from strong constraints near the lower boundary to weak constraints above. The nudging procedure follows the same idea as Maute et al. (2015) and Jones Jr. et al. (2018) except for a different altitude range.

In summary, we developed TIEGCM-NG with a nudging capability to study the GW propagation in the I-T system. We applied the TIEGCM-NG to simulate the I-T response to the Tonga eruption. In this study, a global 2.5° resolution is used with a single level nested domain of 60°S-20°N, 140°E-230°E(130°W) at 1° resolution. This domain is chosen to include Tonga (21°S, 175°W) at the nested grid center and have a sufficient horizontal span (40° on each side) to simulate the wave propagation. The vertical resolution in the global and the nested domains are all set as 1/4 scale height. The global domain has a time step of 30 s, and 5 sub-cycles are iterated in the nested domain, indicating a nested grid time step of 6 s. Diagnostic outputs are saved every minute in the global and nested domains. Nudging of WACCM-X fields is switched on at 04:35:00 when the eruption takes place, and the nudging altitude ranges in both the global and the nested domain are 2 scale heights. Two testing groups are performed to study the sensitivities of model resolutions and nudging fields:

1. The nesting functionality is switched on and off to study the impact of the

177

model resolution on the simulation of GW propagation.

2. To analyze the sensitivity of nudging different fields in wave propagation, runs with nudging different combinations of T, U, V, and Z, all with the nesting functionality on, are performed.

To investigate GW impacts, a different type of runs (control runs) without nudging lower-boundary GWs is performed to present the undisturbed states. The simulation results in all testing runs are subtracted from their corresponding control runs to represent the effects caused by the nudged GWs. Unless otherwise notified, all the figures shown in this paper are deviations from the undisturbed states (difference fields).

6.3 Model Results in GW Propagation

Two groups of runs are performed separately. The first group nudges all 4 fields (T, U, V, and Z) and differs by whether the nesting functionality is on. This run group mainly studies the importance of resolution in simulating the propagation of waves (Section 6.3.1). The model runs in the second group are all nested, but the nudged fields are changed among different ones. This group focuses on the sensitivity of TIEGCM-NG to the nudged fields (Section 6.3.2).

6.3.1 Importance of Model Resolution on Wave Propagation

To simplify the notation for the different runs in this section, the run without nesting is denoted as Run 1, and the one with nesting is denoted as Run 2. Run 2 has two sets of outputs from global and nested domains. Therefore, they are denoted as Run 2g and Run 2n, respectively.

Figure 6.2 compares the time-altitude profiles of vertical wind and temperature perturbations between Runs 1 and 2n. Figures 6.2a-b show the difference fields at 21.25°S, 175°W (close to the epicenter) in Run 1, and Figures 6.2c-d show those at 21°S, 175°W in Run 2n. Due to interpolation, the fields in Run 2g are nearly the same as in Run 2n and are therefore omitted. The wave signatures are present at high altitudes in both runs. Still, the wave amplitude of the vertical wind is significantly larger in Run 2n than in Run 1, over $100\,\mathrm{m\,s^{-1}}$ compared to less than $30\,\mathrm{m\,s^{-1}}$ (Figures 6.2a and 6.2c, color scales are different in Figures 6.2a and 6.2c). Both Figures 6.2b and 6.2d show downward phase progression below $z = 0$ of similar amplitude in temperature perturbations $(100\,\mathrm{K})$, and the periods at these altitudes are generally long $(> 8\,\mathrm{h})$, indicating upward propagating tides or inertial GWs of similar magnitude. Above $z = 0$, the vertical wavelengths become large due to increased diffusivity and ion drag, while nested results tend to show higher-frequency perturbations $(< 1\,\mathrm{h})$ than Run 1.

As shown in Figures 6.2a and 6.2c, the vertical propagating wave shifts towards shorter periods at higher altitudes, and wave periods are shorter at these altitudes in Figure 6.2c compared to Figure 6.2a. Figure 6.3 shows the vertical wind wave spectra from 05:00:00 to 08:00:00 at different altitudes in Runs 1 and 2n. Here, the vertical wind is plotted (not the difference field). The direct component is marked as the infinite period in the figure. At the lowest altitude $(z = -7)$, there are no significant wave components shorter than $15\,\mathrm{min}$, and long-period components dominate, especially for the temporal mean (direct component). At high altitudes in Run 1, the wave periods range from $15\,\mathrm{min}$ to $45\,\mathrm{min}$ centered around $40\,\mathrm{min}$, and no shorter period signals than $12\,\mathrm{min}$ are observed. However, in Run 2n, the period components below $40\,\mathrm{min}$ grow in altitude, and eventually, the wave spectra are excited in a broad band from $5\,\mathrm{min}$ to $45\,\mathrm{min}$ at high altitudes. Also, different
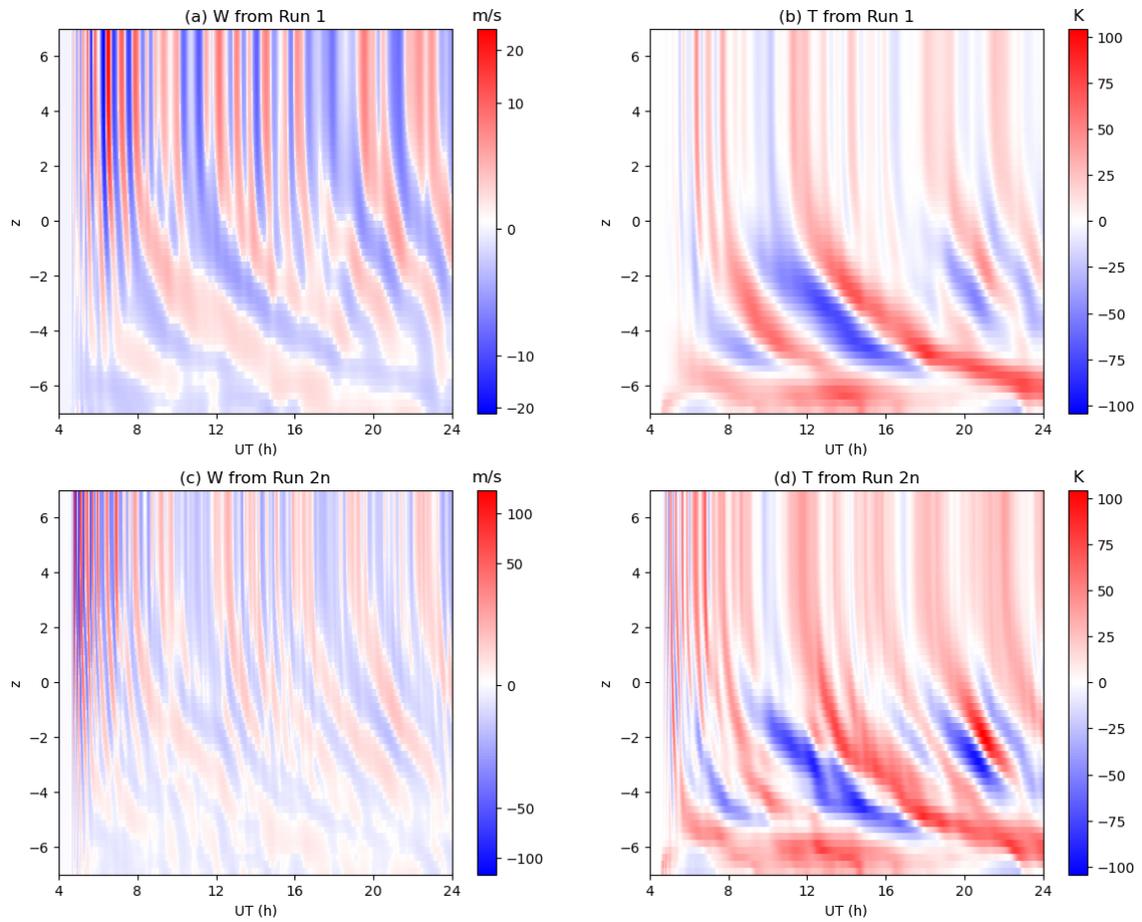
Figure 6.2: Time-altitude cross sections of vertical wind and temperature perturbations at (21.25°S, 175°W) in Run 1 (a-b) and at (21°S, 175°W) in Run 2n (c-d). The color scales are different for vertical winds in (a) and (c).

180

from the dominant 40 min period in Run 1, three spectral peaks at 40 min, 20 min and 12 min arise with a similar magnitude. The change of the dominant wave period with altitude is likely associated with the filtering effect of background winds (Fritts and Alexander, 2003; Alexander et al., 2010).

The amplitude growth of waves with altitudes can cause significant perturbations in the upper thermosphere. Figure 6.4 shows the horizontal structures of vertical wind and temperature perturbations at 08:00:00, $z = 2(320\,\mathrm{km})$ in Runs 1 and 2. Figures 6.4a-b are from Run 1; Figures 6.4c-d are from Run 2g overlaid by Run 2n inside the boxed region. The geopotential height and temperature (not the perturbation field) from WACCM-X are extracted at the TIEGCM lower boundary ($z = -7$) and plotted in Figures 6.4e-f for comparison. Comparing Figures 6.4d and 6.4f, the concentric ring patterns are clear at both the lower boundary and upper thermosphere (also seen in Figures 6.4c and 6.4e) indicating the upward propagation of GWs. However, the differences between Figures 6.4a and 6.4c (Figures 6.4b and 6.4d) are remarkable. The wave components of short wavelengths ($4° \approx 400\,\mathrm{km}$) dominate the nested domain, while in the global run with no nesting, only waves with much longer wavelengths ($20° \approx 2000\,\mathrm{km}$) survive. Also, wave amplitudes in Run 2n are significantly larger than those in Run 1, which indicates that the shorter wavelengths ($400\,\mathrm{km}$) carry the most wave energy in the wave propagation, but in Run 1, due to the insufficient horizontal resolution ($2.5° \approx 250\,\mathrm{km}$), the wave components below $500\,\mathrm{km}$ are dissipative and eventually removed from the wave spectra (Nyquist theorem).

In terms of traveling TAD propagation, time-latitude profiles of vertical wind and temperature perturbations are shown at a constant altitude ($z = 2$) and longitude ($175°\mathrm{W}$) in Figure 6.5 to demonstrate the horizontal wave propagation. This longitude is chosen to cross the epicenter, representing the wave propagation path. Both
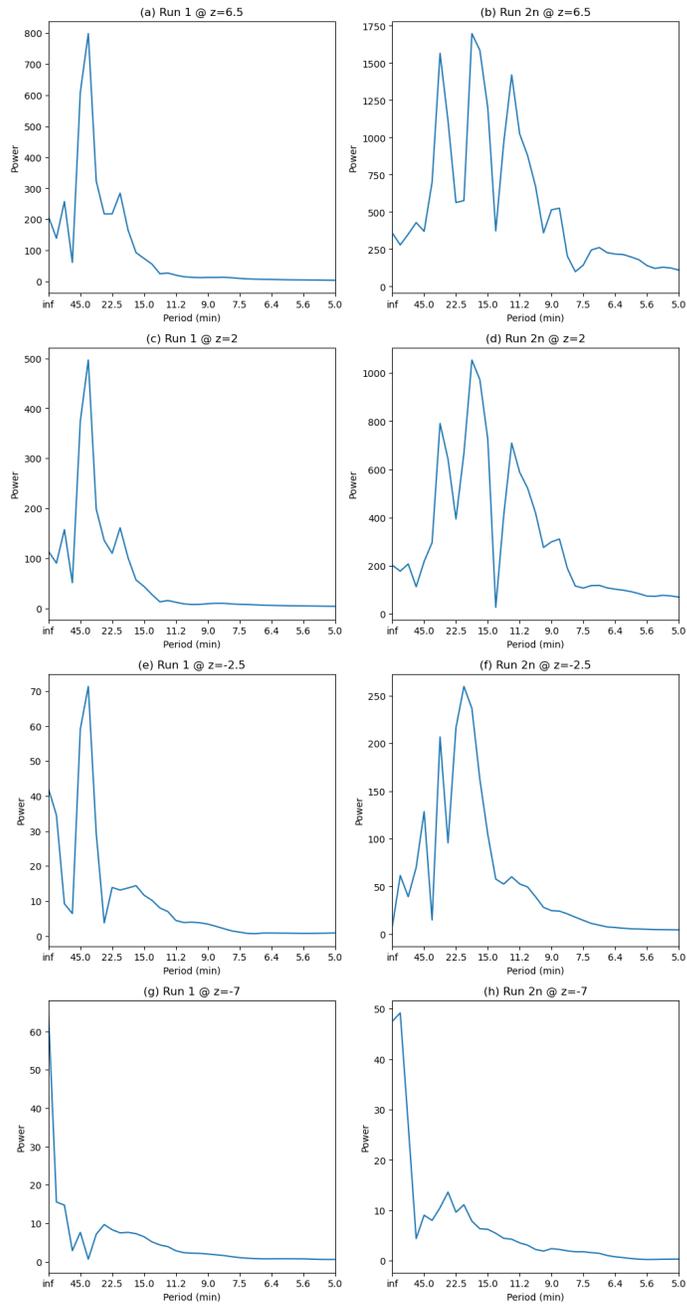
Figure 6.3: Unnormalized wave spectra of vertical winds (not perturbations) at different altitudes ($z = -7, -2.5, 2, 6.5$) in Runs 1 (a, c, e, g) and 2n (b, d, f, h), both at the same locations as in Figure 6.2.
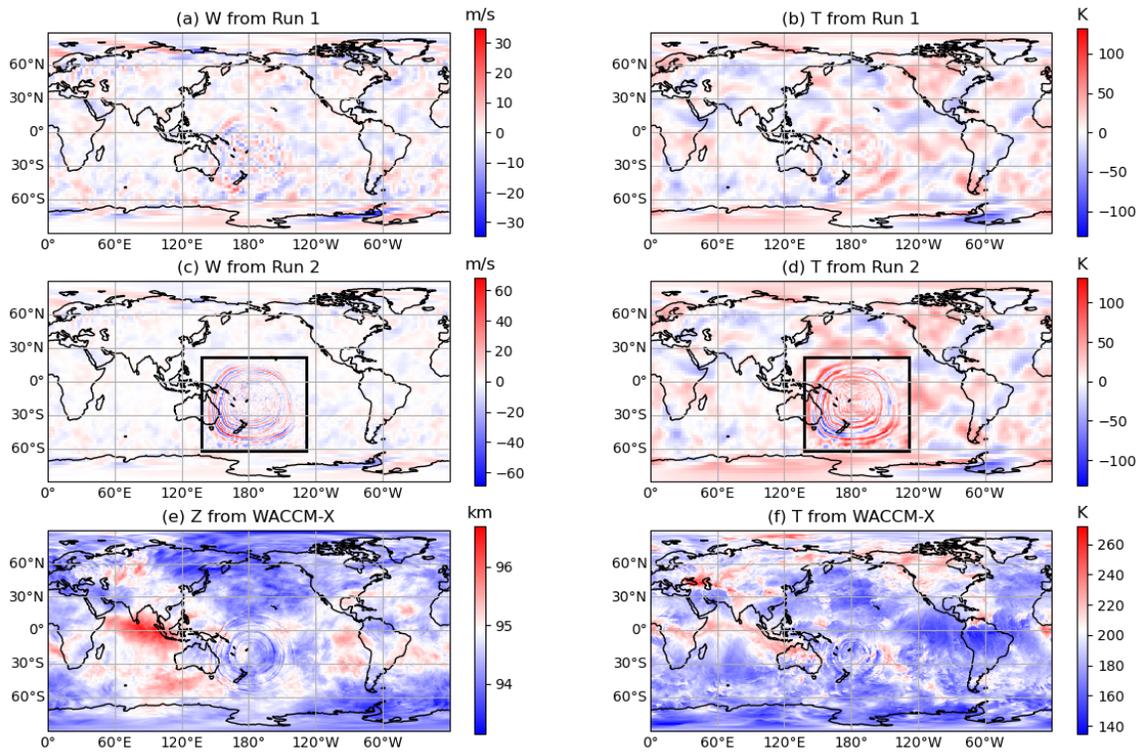
Figure 6.4: (a-d) Horizontal cross sections of vertical wind and temperature pertur-bations at $z = 2(320\,\mathrm{km})$, 08:00:00 in Runs 1 (a-b) and 2 (c-d). Figures 4c-d are from Run 2g overlaid by Run 2n inside the boxed region. (e-f) WACCM-X geopotential height and temperature extracted at TIEGCM lower boundary ($z = -7$). These two are not the perturbation fields.

runs show clear latitudinal propagations, but the wave amplitude is significantly larger in Run 2n compared to Run 1; the wave speed is estimated to be $10\,^\circ\mathrm{h}^{-1} \approx 300\,\mathrm{m\,s}^{-1}$ in both runs, much slower than the local sound speed (typically $800\,\mathrm{m\,s}^{-1}$), indicating that the propagating wave is indeed a GW. No significant difference is found in the wave amplitude or speed along northward and southward directions, which indicates that the wave propagation does not have a preferential direction.
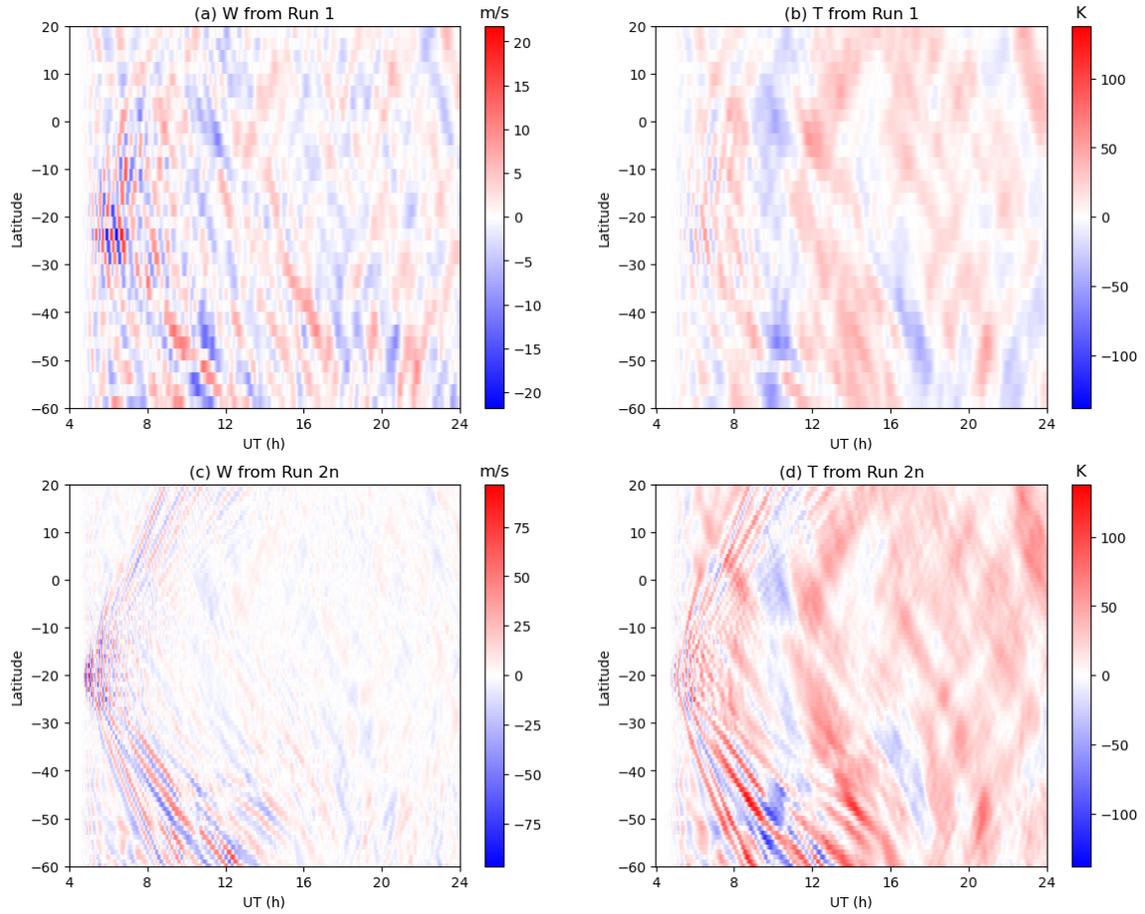
Figure 6.5: Time-latitude cross sections of vertical wind and temperature perturbations at $z = 2$ and 175°W in Runs 1 (a-b) and 2n (c-d).

The large GW amplitudes cause significant ionospheric perturbations. Figure 6.6 shows electron density perturbations at $z = 2$ and differential TEC in both runs.

184

This altitude is close to the ionospheric F2 peak and is therefore chosen to represent the ionospheric variations. Again, the concentric ring structure indicative of outward propagating waves is evident in Figure 6.6c-d, on the order of $10^5\,\mathrm{cm}^{-3}$ and $5\,\mathrm{TECU}$, at a similar level as GNSS TEC observations ($5\,\mathrm{TECU}$ as shown in Figure 2 of Aa et al. (2022)). In contrast, though electron density perturbations are present in Run 1, the wave pattern is much larger, and its magnitude is smaller. In both runs, the wave patterns are more evident in the north direction than the south because the electron density peaks at a northern equatorial latitude.
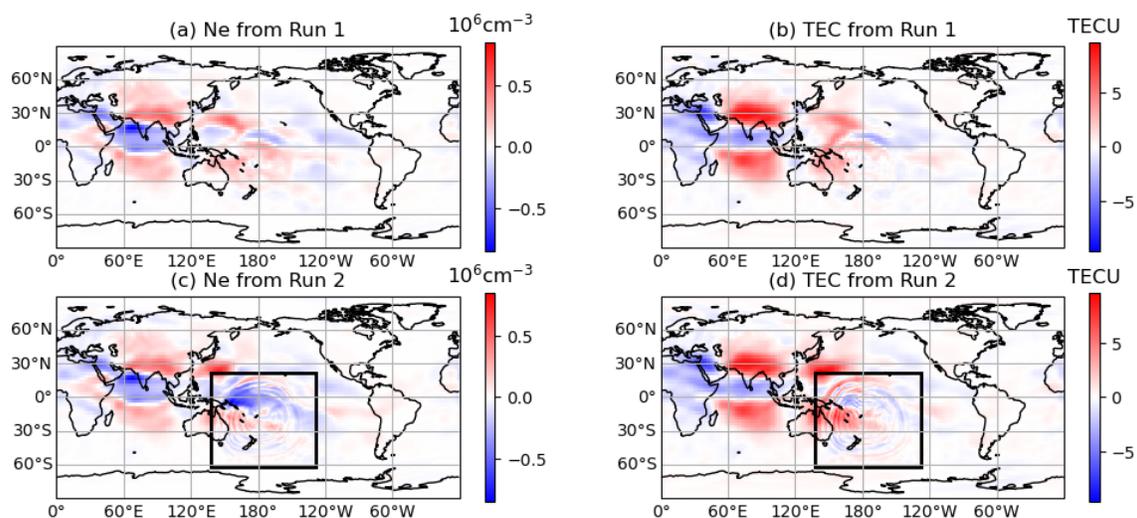
Figure 6.6: Horizontal cross sections of electron density perturbations at $z = 2$ and TEC perturbations at 08:00:00 in Runs 1 (a-b) and 2 (c-d).

The horizontal propagations of ionospheric perturbations are shown in Figure 6.7. The time-latitude cross sections are shown at the same altitude ($z = 2$) and longitude (175°W) as in Figure 6.5. The TID speed is estimated to be $300\,\mathrm{m\,s}^{-1}$, of a magnitude similar to TAD, indicating that the source of the TIDs is most likely the neutral variation resulting from the volcano eruption. The estimated TID speed follows the slow mode of the near-field variations in Figure 2 of Themens et al. (2022) and Figure 3 of Zhang et al. (2022b). Just as shown in Figure 6.6, the northward

propagation of TID ($2\,\text{TECU}$) in Figure 6.5 is significantly stronger than that in the southward direction ($< 1\,\text{TECU}$).
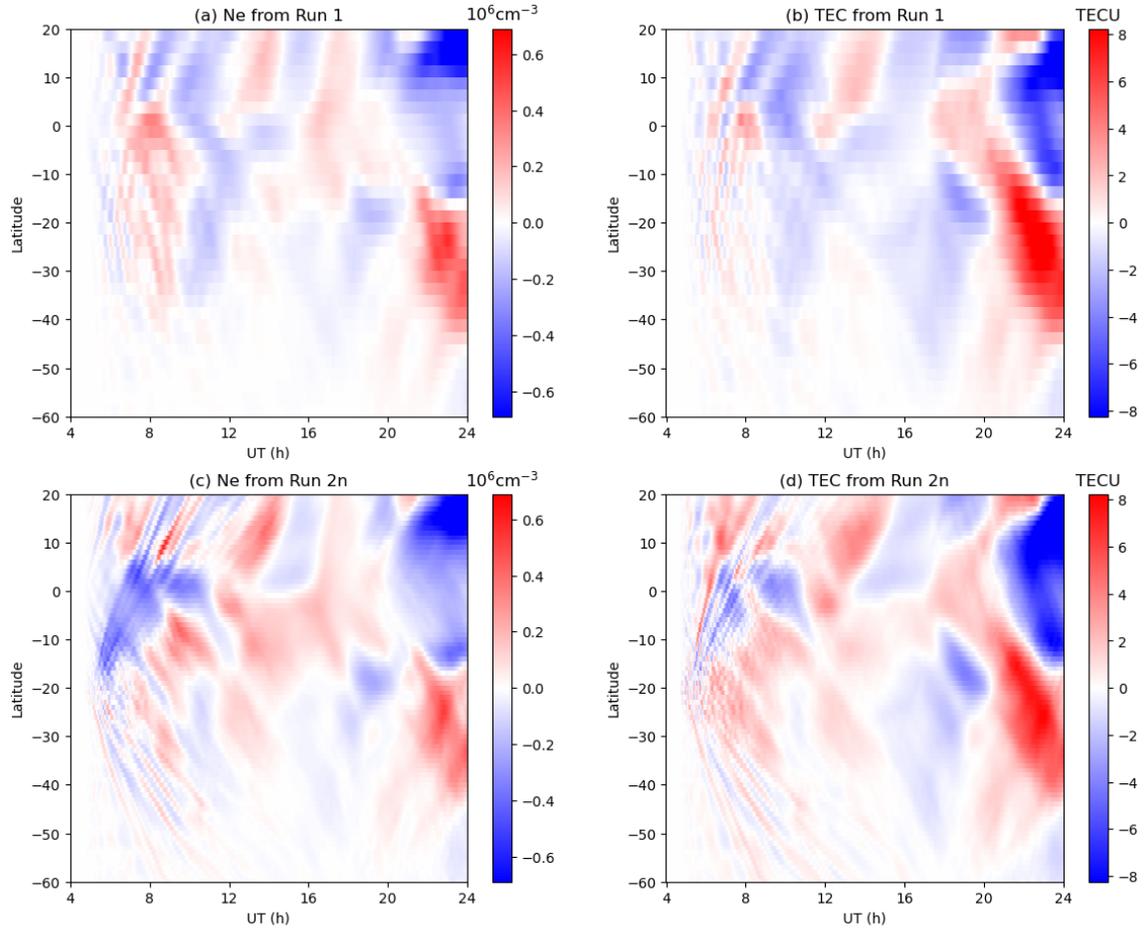
Figure 6.7: Time-latitude cross sections of electron density perturbations at $z = 2$ and TEC perturbations at 175°W in Runs 1 (a-b) and 2n (c-d).

6.3.2 Model Sensitivity to Nudging Fields

To examine the capability of incorporating waves from below at the TIEGCM-NG lower boundary by nudging different wave fields, the model runs belonging to the second testing group in Section 6.2.2 are performed. Three different runs are performed: Run TUV, indicating T, U, and V are nudged near the lower boundary;

Run Z, indicating only Z is nudged; and Run TUVZ, indicating all fields are nudged at the model low-pressure levels (Run 2 in Section 6.3.1).

The vertical and horizontal propagations in the nested domain are shown in Figure 6.8 for Runs TUV, Z, and TUVZ. Different cross-sections are drawn in the figure for different runs, including the time-altitude cross-section at the epicenter (21°S, 175°W), the time-latitude cross-section at $z = 2$, 175°W, and the longitude-latitude cross-section at 08:00:00, $z = 2$. Although the vertical and horizontal structures of the GW are similar, comparing Figures 6.8a and 6.8g (6.8b and 6.8h, 6.8c and 6.8i), the magnitude is significantly smaller in Run TUV than in Run TUVZ, $40\,\mathrm{m\,s}^{-1}$ versus $100\,\mathrm{m\,s}^{-1}$. On the other hand, comparing Runs Z and TUVZ, no distinguishable differences are identified from Figures 6.8d and 6.8g (6.8e and 6.8h, 6.8f and 6.8i). The comparison indicates that Z is the most important field to be nudged in GW propagation. In other words, high-frequency GW is most sensitive to Z. The possible reason for this sensitivity can be summarized in two points. First, in this study, high-frequency waves have larger amplitudes and dominate the region of interest. The vertical wind spectrum is sensitive to high-frequency waves, while zonal and meridional wind spectra are sensitive to low-frequency waves (Geller and Gong, 2010). This leads to the dominant high-frequency waves in vertical winds. Second, the vertical wind is proportional to the displacement of Z, but nudging T, U, and V only helps improve the vertical gradient of vertical winds under hydrostatic conditions. So nudging Z directly improves the vertical wind and supports high-frequency waves.
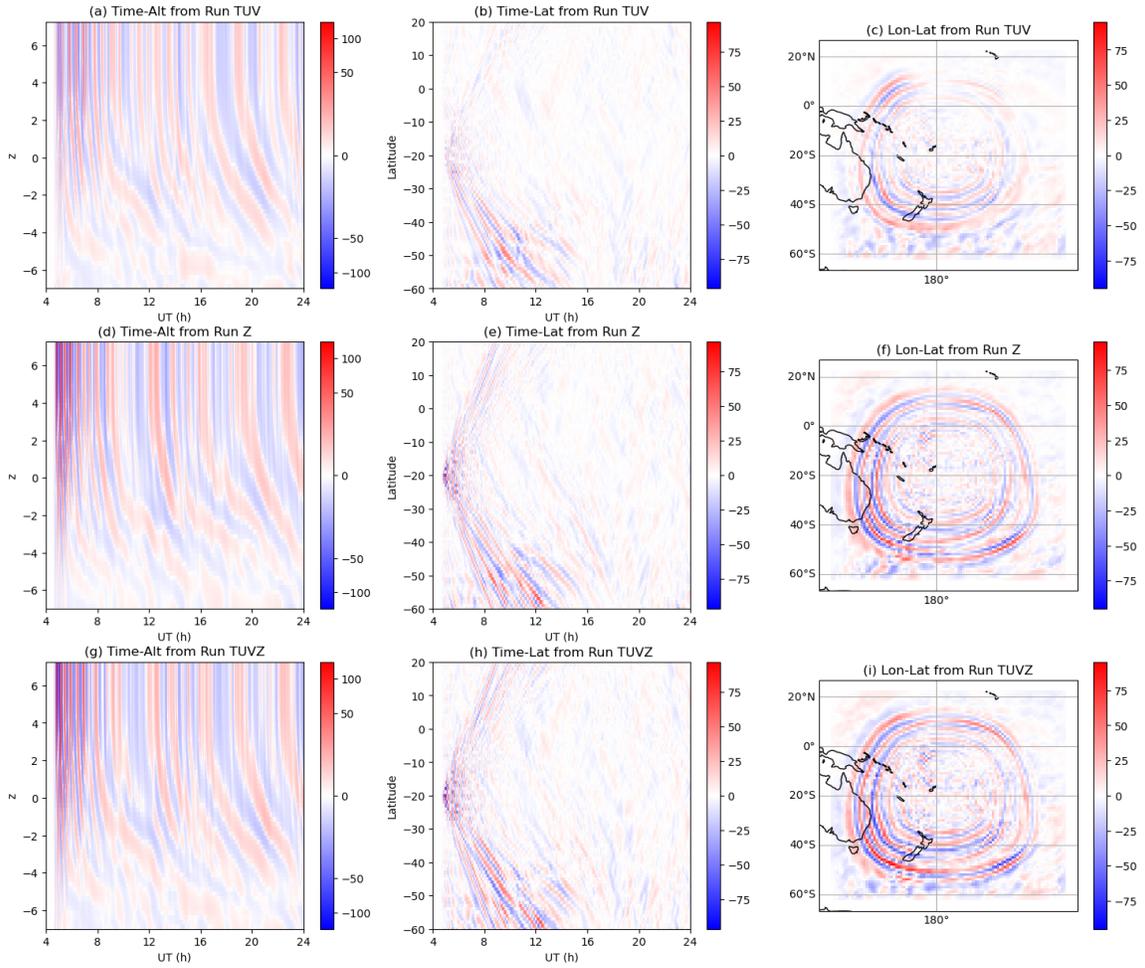
Figure 6.8: Different cross sections of vertical wind perturbations in Runs TUV (a-c), Z (d-f), and TUVZ (g-i). Time-altitude cross sections at 21°S, 175°W are shown in (a, d, g); time-latitude cross sections at $z = 2$, 175°W are shown in (b, e, h); longitude-latitude cross sections at 08:00:00, $z = 2$ are shown in (c, f, i). All units are $\mathrm{m\,s^{-1}}$.

6.4   Conclusion and Discussion

We develop a new TIEGCM-NG model, a regional high-resolution extension of the original TIEGCM, to study I-T processes on regional scales. It solves the same equations as the original TIEGCM and uses time-dependent Dirichlet lateral boundary conditions obtained from the global low-resolution run. Near the nested grid boundaries, the original 4-point stencil in the horizontal finite difference is substituted by a 2-point stencil. The sub-cycling in time guarantees the model's stability in the nested grid domain. We further add the capability of nudging the model at low pressure levels to better simulate the propagation of GWs from the lower atmosphere.

We then nudge WACCM-X simulation results to subscribe GWs into TIEGCM-NG to the impacts of 2022-01-15 volcano eruption at Hunga Tonga-Hunga Ha'apai on the ionospheres via the vertical coupling of GWs. The TIEGCM-NG simulations indicate that high resolution is critical to simulate wave propagation. The wave period shifts from long to short as it propagates from low to high altitudes, likely due to the filtering effect of background winds. The wave can further propagate to the ionosphere, leading to significant ionospheric disturbances such as large electron density perturbations. The TIEGCM-NG simulated eruption-induced ionospheric perturbations have a similar magnitude to the observations. The speeds of TAD and TID at $320\,\mathrm{km}$ are estimated to be $300\,\mathrm{m\,s^{-1}}$, following those derived from the TEC observations. In addition, we perform a sensitivity test to determine the essential nudging fields to fuse in GWs at the lower boundary. The test shows that geopotential height is the essential field to be nudged into TIEGCM-NG, and geopotential height variations are the easiest to propagate through the I-T system.

However, it is necessary to point out that the simulated GW in this study only covers the one excited by the initial pulse of the volcano eruption. Due to a large

amount of water vapor emitted into an intense column throughout the stratosphere, the release of latent heat at later times also contributes to the excitation of GWs in the authentic atmosphere. The secondary GW excited by such a mechanism is not studied.

It is worth emphasizing that the computation efficiency is much higher in TIEGCM-NG than in global high-resolution TIEGCM. Thus TIEGCM-NG can be an ideal choice in applications where computing time is a critical factor, like near-real-time forecasting on regional scales.

One advantage of TIEGCM-NG compared to local models is that the nested region's background atmosphere (winds, temperatures, etc.) can be more interactive and realistic, representing the time dependence on different external conditions. For example, TIEGCM-NG can assist in investigating under what conditions GWs can survive after long-distance propagations and impact the ionosphere (e.g., in the form of TID). Although high-resolution WACCM-X is involved in this study, it is necessary to mention that TIEGCM-NG is not tied to global high-resolution simulations. The purpose of subscribing external field from WACCM-X is to provide GW seeding near the lower boundary. The choice of GW seeding can vary from comprehensive physical models like MAGIC (Heale et al., 2022; Inchin et al., 2022) or theoretical GW models to point-like sources in the lower atmosphere. The computation-efficient nature of TIEGCM-NG is only partially appreciated in the current coupled WACCM-X/TIEGCM-NG study due to the high cost of high-resolution WACCM-X simulations. However, TIEGCM-NG will most likely find its advantage in real applications by coupling to other regional high-resolution models like MAGIC. MAGIC can simulate small-scale GWs while saving computing time (compared to global models) by focusing on a regional area. While MAGIC itself cannot simulate the change of the lateral boundary, a coupled MAGIC/TIEGCM-NG in the future will likely bring

new insight into the GW simulation by combining the advantages of global models (no lateral boundaries) and regional models (high resolution).

It is also arguable that a time-dependent boundary flux (Neumann condition) instead of the field (Dirichlet condition) might better serve the wave propagation across the boundary. Still, as a starting point and for easy implementation, the field is prescribed as the lateral boundary condition in the current development. Among all simulations we perform, there are no significant boundary distortions or reflections; thus, we choose to use a Dirichlet condition. The current design of the nested grid extension is a separate module of the original TIEGCM subroutines, making it convenient to modify the physics in the nested grid domain in the future. It is possible to incorporate different sets of physical equations (such as those describing the non-hydrostatic processes) in the nested domain and combine them into global TIEGCM simulations. In this way, even localized high-frequency compressible atmospheric waves can be simulated in the framework of TIEGCM. Further extensions of TIEGCM-NG are left as a future task in studying different atmospheric processes.

6.5   Comparing to High-Resolution GW Simulations in GITM

Recently, GITM-R has implemented the nesting capability using a similar design discussed in Section 6.2.1. Zhao et al. (2020) performs an ionospheric GW study resulting from a tropical cyclone using GITM-R. A theoretical monochromatic wave is imposed on the horizontal winds at the lower boundary as

$$\delta u = A \cos(2\pi r/\lambda - \omega t)$$

191

where A is the amplitude, λ and ω are the horizontal wavelength and frequency, r is the distance from the epicenter, and t is the time. λ and ω are chosen to match the observations best.

Figure 6.9 displays the GITM-simulated TEC perturbations (δTEC). The maximum δTEC is 0.019, 0.064, 0.086, and 0.093 at 5 min, 15 min, 35 min, and 1 h. The magnitude increases with time, and the TEC perturbations expand radially outward from the wave center. The magnitude of the TEC perturbations is not circularly symmetric, especially at 1 h, likely caused by the variation of background winds and the effect of the magnetic field.

Compared to our approach of imposing realistic perturbations in the lower boundary (Section 6.2.2), imposing idealized GWs in the lower boundary produces oversimplified spatial structures, making the simulation deviate from real-time TEC observations.
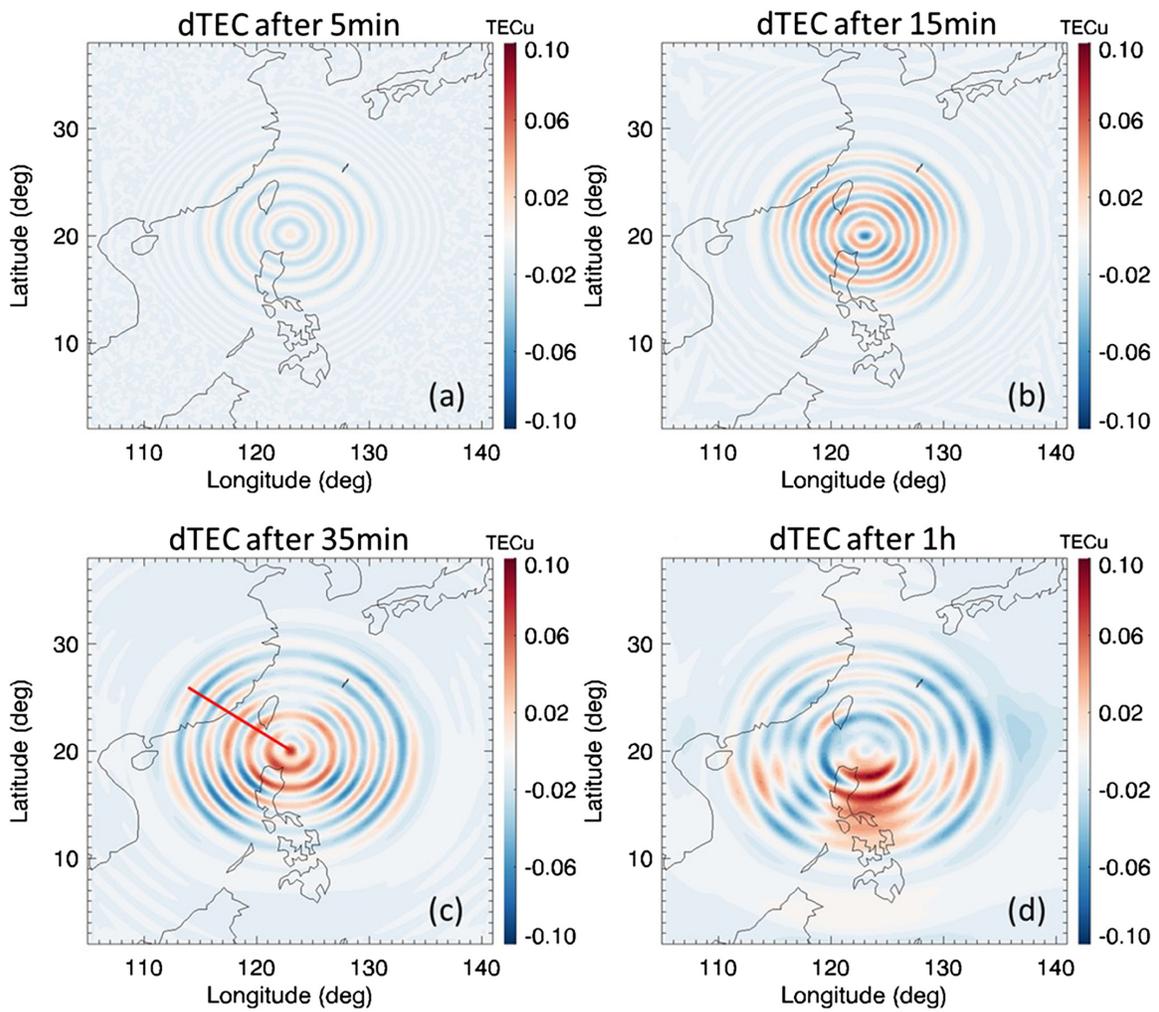
Figure 6.9: Snapshots of simulated δTEC at four different times after forcing imposed at the lower boundary. Adapted from Zhao et al. (2020).

Chapter 7

Conclusions and Outlook

We conclude by summarizing the results of the individual projects comprising this dissertation and by briefly discussing possible future work based on the current stage of model development.

7.1 Engineering Achievements

Development of Auroral Data Assimilation   We introduce a multi-resolution Gaussian process model to synthesize various data sources self-consistently (empirical models, satellite-based and ground-based data) for the auroral assimilation for the first time. The assimilation results reproduce the large-scale and mesoscale auroral structures with quantified uncertainties over the assimilation domain. To produce assimilation results that agree better with observations, we use data downsampling to decrease the weightings of low-fidelity data (interpolated satellite data and empirical model) and ensure the dominance of high-fidelity data (satellite and ground-based observations) in the assimilation results. Further, we generate a postprocessing weighting map using KNN trained by auroral observations to mitigate the inherent

smoothing effect of the fitting procedure. The combined effect of the data assimilation model and the postprocessing weighting map helps reproduce mesoscale auroral arcs to a large extent. We use assimilated aurora to drive TIEGCM, and the simulated TEC structures show more spatial structures and agree with GNSS observations.

Development of Electric Field Data Assimilation We introduce a multi-resolution Gaussian process model to assimilate high-latitude electric fields by extending the Lattice Kriging framework to vector fields. Several sensitivity tests suggest that data takes a leading role over the background model. Comparing fitting errors using different resolutions, higher-resolution assimilation leads to more minor fitting errors, which suggests that higher-resolution assimilation can capture more medium-to-small-scale structures. We further implement the multi-resolution assimilation capability by locating multiple levels of basis functions in the same region. By setting up multi-level basis functions, we significantly decrease the computation cost while achieving better error performance than the single-resolution assimilation. The capability of tracing data coverage automatically and locating high-resolution basis functions according to data is further developed to reduce the computation cost to a level similar to low-resolution assimilation while retaining good error performance as the high-resolution assimilation. We compare local assimilation results with PFISR measurements and confirm the new data assimilation model can also reproduce local observations to a large extent.

Development of Nested-Grid TIEGCM We extend the original TIEGCM with the regional high-resolution refinement capability using a time-dependent Dirichlet lateral boundary condition at each level. The numerical stability is reinforced by time sub-cycling in the nested region. We design and implement a highly parallel

scheme to enable real-time high-resolution modeling. A nudging capability is developed to support atmospheric wave propagation in the model. GW propagations are therefore reproduced in the I-T region by multi-level simulations and nudging WACCM-X results near the lower boundary. The simulated GWs in TIEGCM-NG have a comparable phase speed to observations.

7.2 Scientific Findings

Mechanism Study of TTEIL during the 2011-05-28 Storm Regional enhancement in aurora leads to a considerable increase in Pedersen conductivity and deeper penetration of Joule heating to the lower thermosphere. The implementation of electric field variabilities leads to intense Joule heating, which triggers dynamic responses in the thermosphere. Enhancements in aurorae and electric fields lead to intense Joule heating in the lower thermosphere, which causes the temperature to increase drastically. Joule heating-induced pressure gradient changes lead to the horizontal divergence of airflow, which results in large vertical velocities. Strong vertical motions lead to strong adiabatic cooling at higher altitudes which assists the formation of TTEIL. Vertical advective heating acts as a cooling term below and a heating term above, hampering the formation of inverse temperature profiles. Summing up the heating effects from three terms together, Joule heating and adiabatic cooling support the formation of TTEIL, and vertical advective heating hinders TTEIL, and their joint effect leads to the observed TTEIL. Dynamic disturbances resulting from Joule heating propagate to regions without local Joule heating in the form of TADs, which lead to secondary TTEIL with smaller amplitude.

Model Study of Strong Vertical Winds during the St. Patrick's Day Storm
With enhanced aurora, TIEGCM generates a secondary electron density peak at E
region altitudes. After further implementing realistic aurora, intense Joule heating
appears, which leads to significant vertical winds at 250 km. Data-driven aurora
and electric fields cause a significant increase in spatial and temporal variabilities in
the model. TADs with large amplitudes, resulting from the intense Joule heating,
propagate globally, which leads to elevated variabilities in low latitudes.

Model Study of GW Propagation during Tonga Volcano Eruption Tonga
volcano eruption-induced GWs have a prominent wave period shorter than 45 min.
GWs of short periods at 15 min and short wavelengths below 500 km, present in the
high-resolution simulation, carry the most wave energy and cause the most significant
I-T responses. To produce high-frequency GW signals in vertical winds, it is essential
to include geopotential height perturbations in the simulation due to direct physical
relations between geopotential heights and vertical winds.

In summary, during magnetic storms, magnetospheric energy deposition can
significantly impact the I-T system in the form of elevated Joule heating resulting
from enhanced aurora and electric fields. The elevated Joule heating induces dynamic
and thermodynamic I-T responses and causes small to large-scale TADs and TIDs
which leads to the global propagation of I-T variations. During both storm and quiet
time, wave activities of various wavelengths and periods from the lower atmosphere
transport momentum and energy into the I-T system and lead to variabilities across
a wide range of spatial and temporal scales.

7.3 Outlook

The newly developed tools (the data assimilation model, TIEGCM-NG, and TIEGCM-nudging) described in this dissertation have shown their great potential to better study the coupled I-T system by refining the forcing both from above and from below towards reality. With these tools, several major science questions remain to be better answered:

1. How does the I-T system respond on different scales to magnetic storms? What is the relative contribution from each scale to the I-T responses?

2. How does the I-T system respond to downward magnetospheric energy depositions and upward-propagating atmospheric waves simultaneously?

3. How can these tools leverage the existing observations/models, and future missions such as GDC and Dynamical Neutral Atmosphere-Ionosphere Coupling (DYNAMIC) to investigate fundamental physical processes in the I-T system?

We are nearly ready to begin work on the first topic. All that we need are high-resolution observations of aurora and electric fields. We can analyze the scale dependency using the new data assimilation model and produce auroral and electric field assimilation results at different scales. Then we will use these assimilation results in TIEGCM and study the simulated I-T responses. The runs driven by different scales of magnetospheric energy inputs (aurora and electric field) can be quantitatively analyzed through comparisons among different runs.

It is known that geomagnetic storms can change circulations and compositions. Such impacts reach the whole globe which potentially alters the wave propagation from below. On the other hand, changes in wave momentum and energy deposition modulate the I-T states, and then the I-T responses to storms. Again, if we can better

quantify each source, timing, and spatial extension of the storm and wave impacts by combining data and model, we are in a better position to realistically quantify the underlying processes, and unravel the mysteries of observations with a large dynamic range.

GDC mission will enable the observations of high-latitude magnetospheric forcing including aurora and ion drifts, which are the perfect inputs to the data assimilation model we developed. The DYNAMIC mission will observe atmospheric waves in an unprecedented cadence, which delineates more details of both large and small-scale waves. And their impacts can be better resolved in the nudging procedure and higher-resolution modules. So the overall set of tools developed in this dissertation is highly relevant and holds great potential to contribute.

Appendices

Appendix A Lattice Kriging

Assume that the field $y(x)$ satisfies an additive model with a statistical mean $\mu(x)$ and deviation $g(x)$

$$y(x) = \mu(x) + g(x) + \epsilon(x) \tag{1}$$

$\mu(x)$ is directly taken from an empirical model $z(x)$ but multiplied by a global linear scaling factor d (to be estimated) to account for the bias

$$\mu(x) = z(x)d$$

Decompose $g(x)$ onto a set of m basis functions $\phi_j(x)(1 \leq j \leq m)$

$$g(x) = \sum_{j=1}^{m} c_j \phi_j(x) \tag{2}$$

where c_j is the coefficient of the jth basis function. The covariance function of $g(x)$ is written as

$$\text{cov}\left(g(x), g(x')\right) = \sum_{1 \leq j, j' \leq m} \rho \phi_j(x) \boldsymbol{Q}_{j,j'}^{-1} \phi_{j'}(x') \tag{3}$$

where $\boldsymbol{Q}^{-1}$ is the covariance matrix in the representation of $\phi_j(x)$ and ρ serves as a scaling factor of the covariance matrix. To guarantee the positive definiteness of the covariance matrix, it is desired to write the precision matrix $\boldsymbol{Q}$ in an inner product form

$$\boldsymbol{Q} = \boldsymbol{B}^T \boldsymbol{B} / \rho$$

Following Lindgren et al. (2011), $\boldsymbol{B}$ is formulated as

$$\boldsymbol{B}_{j,j'} = \begin{cases} 4 + \kappa^2, j' = j \\ -1, j' = j \pm 1 \\ 0, \text{otherwise} \end{cases}$$

where $\kappa \neq 0$ for the determinant to be non-zero. In this formulation, $\boldsymbol{B}$ is a sparse matrix, and $\boldsymbol{Q}$ is also sparse. This precision matrix $\boldsymbol{Q}$ is derived from a Gaussian Markov random field in which two locations are correlated only if they are adjacent.

Given a set of observations y_i at $x_i (1 \leq i \leq n)$, write the basis functions into an $n \times m$ matrix $\boldsymbol{\phi}$ with its elements satisfying

$$\boldsymbol{\phi}_{i,j} = \phi_j(x_i)$$

Then expand Equations (2) and (3) to all points and obtain their corresponding matrix forms

$$g(\boldsymbol{x}) = \boldsymbol{\phi c}$$

$$\text{cov}\left(g(\boldsymbol{x}), g\left(\boldsymbol{x'}\right)\right) = \rho \boldsymbol{\phi Q}^{-1} \boldsymbol{\phi}^T$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, and $\boldsymbol{c} = (c_1, c_2, \ldots, c_m)$. The error covariance matrix $\boldsymbol{W}^{-1}$ is diagonal

$$\boldsymbol{W}^{-1} = \text{diag}(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$$

When combining the covariance and the error covariance into the model, another scaling parameter σ^2 is added to the error covariance matrix $\sigma^2 \boldsymbol{W}^{-1}$. Stack all observations y_i into a vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, and empirical model at input locations

z_i into $\boldsymbol{Z} = (z_1, z_2, \ldots, z_n)$, then the matrix form of Equation (1) can be written as

$$\boldsymbol{y} = \boldsymbol{Z}d + \boldsymbol{\phi}\boldsymbol{c} + \boldsymbol{\epsilon}$$

$\boldsymbol{y}$ satisfies a multi-variate normal distribution with a mean of $\boldsymbol{Z}d$ and a SD of $\rho\boldsymbol{\phi}\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T + \sigma^2\boldsymbol{W}^{-1}$

$$\boldsymbol{y} \sim \mathrm{MVN}\left(\boldsymbol{Z}d, \rho\boldsymbol{\phi}\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T + \sigma^2\boldsymbol{W}^{-1}\right)$$

Define a new parameter $\lambda = \sigma^2/\rho$ and an auxiliary matrix

$$\boldsymbol{M}_\lambda = \boldsymbol{\phi}\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T + \lambda\boldsymbol{W}^{-1}$$

The subscript indicates that $\boldsymbol{M}$ is a function of λ but not ρ. Re-parameterize in terms of λ and ρ

$$\boldsymbol{y} \sim \mathrm{MVN}(\boldsymbol{Z}d, \rho\boldsymbol{M}_\lambda)$$

The log-likelihood of the model is

$$l(d, \rho, \lambda) = -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{Z}d)^T(\rho\boldsymbol{M}_\lambda)^{-1}(\boldsymbol{y} - \boldsymbol{Z}d) - \frac{1}{2}\log|\rho\boldsymbol{M}_\lambda| - \frac{n}{2}\log 2\pi \qquad (4)$$

To maximize l, it is desired to maximize over d. The generalized least squares estimate of d is

$$\hat{d} = \left(\boldsymbol{Z}^T\boldsymbol{M}_\lambda^{-1}\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}^T\boldsymbol{M}_\lambda^{-1}\boldsymbol{y}$$

Set $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{Z}\hat{d}$ and substitute back into Equation (4)

$$l(\rho, \lambda) = -\frac{1}{2}\boldsymbol{r}^T(\rho\boldsymbol{M}_\lambda)^{-1}\boldsymbol{r} - \frac{1}{2}\log|\rho\boldsymbol{M}_\lambda| - \frac{n}{2}\log 2\pi \qquad (5)$$

203

Maximize l over ρ gives

$$\hat{\rho} = \boldsymbol{r}^T \boldsymbol{M}_\lambda^{-1} \boldsymbol{r}/n$$

Equation (5) then becomes

$$l(\lambda) = -\frac{1}{2}\boldsymbol{r}^T(\hat{\rho}\boldsymbol{M}_\lambda)^{-1}\boldsymbol{r} - \frac{1}{2}\log|\hat{\rho}\boldsymbol{M}_\lambda| - \frac{n}{2}\log 2\pi \qquad (6)$$

Finally, the best estimate of λ is given by maximizing Equation (6) over the whole possible value range $(0, \infty)$.

The coefficient of the basis functions $\boldsymbol{c}$ satisfies an MVN distribution given all other parameters

$$\left[\boldsymbol{c}|\boldsymbol{y}, d, \sigma^2, \rho, \boldsymbol{Q}^{-1}\right] \sim \mathrm{MVN}\left(\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T\boldsymbol{M}_\lambda^{-1}\boldsymbol{r}, \rho\boldsymbol{Q}^{-1} - \rho\boldsymbol{Q}^{-1}\boldsymbol{\phi}^T\boldsymbol{M}_\lambda^{-1}\boldsymbol{\phi}\boldsymbol{Q}^{-1}\right)$$

Appendix B   Acronyms

This table list all the acronyms used in this dissertation and their full names.

Acronym	Full Name
ABI	Advanced Baseline Imager
AE	Auroral Electrojet
AIRS	Atmospheric Infrared Sounder
AMGeO	Assimilative Mapping of Geospace Observations
AMIE	Assimilative Mapping of Ionospheric Electrodynamics
AMPERE	Active Magnetosphere and Planetary Electrodynamic Response Experiment
ASI	All-Sky Imager
BLUP	Best Linear Unbiased Prediction
DE	Dynamics Explorer
DMSP	Defense Meteorological Satellites Program
DYNAMIC	Dynamical Neutral Atmosphere-Ionosphere Coupling
FAC	Field-Aligned Current
FBI	Farley-Buneman Instability
FOV	Field-Of-View
FPI	Fabry-Perot Interferometer
GAIA	Ground-to-topside model of Atmosphere and Ionosphere for Aeronomy
GDC	Geospace Dynamics Constellation
GITM	Global Ionosphere Thermosphere Model
GITM-R	GITM with local refinement
GLAT	Geographic Latitude
GLON	Geographic Longitude

GNSS	Global Navigation Satellite System
GOES	Geostationary Operational Environmental Satellites
GRACE	Gravity Recovery and Climate Experiment
GUVI	Global Ultraviolet Imager
GW	Gravity Wave
I-T	Ionosphere-Thermosphere
ICME	Interplanetary Coronal Mass Ejection
ICON	Ionospheric Connection Explorer
IDM	Ion Drift Meter
IMF	Interplanetary Magnetic Field
ISR	Incoherent Scatter Radar
IVM	Ion Velocity Meter
KNN	K-Nearest Neighbors
LBH	Lyman-Birge-Hopfield
LOS	Line-Of-Sight
M-I-T	Magnetosphere-Ionosphere-Thermosphere
MAGIC	Model for Acoustic-Gravity wave Interactions and Coupling
MERRA	Modern-Era Retrospective analysis for Research and Applications
MIGHTI	Michelson Interferometer for Global High-resolution Thermospheric Imaging
MLAT	Magnetic Latitude
MLON	Magnetic Longitude
MLS	Microwave Limb Sounder
MLT	Magnetic Local Time
NG	Nested Grid

NORSTAR	Northern Solar Terrestrial Array
OVATION	Oval Variation, Assessment, Tracking, Intensity, and Online Nowcasting
PFISR	Poker Flat ISR
RBF	Radial Basis Function
RHS	Right Hand Side
RMSE	Root-Mean-Square Error
SABER	Sounding of the Atmosphere using Broadband Emission Radiometry
SCHISM	Semi-implicit Cross-scale Hydroscience Integrated System Model
SD	Standard Deviation
SE	Spectral Element
SECS	Spherical Elementary Current Systems
SHF	Spherical Harmonics Fitting
SLT	Solar Local Time
SSIES	Special Sensor for Ion and Electron Scintillation
SSJ	Special Sensor J
SSUSI	Special Sensor Ultraviolet Spectrographic Imager
SuperDARN	Super Dual Auroral Radar Network
SYM-H	Symmetric Disturbance Field in H component
TAD	Traveling Atmospheric Disturbance
TEC	Total Electron Content
THEMIS	Time History of Events and Macroscale Interactions during Substorms
TID	Traveling Ionospheric Disturbance
TIDI	TIMED Doppler Interferometer
TIEGCM	Thermosphere Ionosphere Electrodynamics General Circulation Model

TIGCM	Thermosphere Ionosphere General Circulation Model
TIMED	Thermosphere Ionosphere Mesosphere Energetics and Dynamics
TING	Thermosphere Ionosphere Nested Grid
TTEIL	Thermospheric Temperature Enhancement and Inversion Layer
WACCM	Whole Atmosphere Community Climate Model
WACCM-X	WACCM with thermosphere and ionosphere extension
WAM-IPE	Whole Atmosphere Model with Ionosphere, Plasmasphere, Electrodynamics
WRF	Weather Research and Forecasting model

Appendix C   External Links

This section lists links to the publications and code repositories.

Chapter 2	10.1029/2020JA028224
Chapter 3	10.1029/2022SW003146
Chapter 4	10.1029/2021SW002880
Chapter 5	10.1029/2022SW003308
Chapter 6	10.1029/2023JA031354
Code Repository	`https://github.com/hzfywhn`

Bibliography

Aa, E., Zhang, S.-R., Erickson, P. J., Vierinen, J., Coster, A. J., Goncharenko, L. P., Spicher, A., and Rideout, W. (2022). Significant ionospheric hole and equatorial plasma bubbles after the 2022 tonga volcano eruption. *Space Weather*, 20(7):e2022SW003101. e2022SW003101 2022SW003101.

Akmaev, R. A., Fuller-Rowell, T. J., Wu, F., Forbes, J. M., Zhang, X., Anghel, A. F., Iredell, M. D., Moorthi, S., and Juang, H.-M. (2008). Tidal variability in the lower thermosphere: Comparison of whole atmosphere model (wam) simulations with observations from timed. *Geophysical Research Letters*, 35(3).

Alexander, M. J., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., Sato, K., Eckermann, S., Ern, M., Hertzog, A., Kawatani, Y., Pulido, M., Shaw, T. A., Sigmond, M., Vincent, R., and Watanabe, S. (2010). Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Quarterly Journal of the Royal Meteorological Society*, 136(650):1103–1124.

Amm, O., Grocott, A., Lester, M., and Yeoman, T. K. (2010). Local determination of ionospheric plasma convection from coherent scatter radar data using the secs technique. *Journal of Geophysical Research: Space Physics*, 115(A3).

Amores, A., Monserrat, S., Marcos, M., Argueso, D., Villalonga, J., Jorda, G., and Gomis, D. (2022). Numerical simulation of atmospheric lamb waves generated by the 2022 hunga-tonga volcanic eruption. *Geophysical Research Letters*, 49(6):e2022GL098240. e2022GL098240 2022GL098240.

Anderson, B. J., Korth, H., Waters, C. L., Barnes, R. M., and Weiss, M. B. (2008). The active magnetosphere and planetary electrodynamics response experiment (ampere): A new facility for real-time magnetosphere-ionosphere monitoring. In *Fall Meeting Abstracts*. American Geophysical Union.

Anderson, B. J., Korth, H., Waters, C. L., Green, D. L., Merkin, V. G., Barnes, R. J., and Dyrud, L. P. (2014). Development of large-scale birkeland currents determined from the active magnetosphere and planetary electrodynamics response experiment. *Geophysical Research Letters*, 41(9):3017–3025.

Anderson, B. J., Takahashi, K., Kamei, T., Waters, C. L., and Toth, B. A. (2002). Birkeland current system key parameters derived from iridium observations: Method and initial validation results. *Journal of Geophysical Research: Space Physics*, 107(A6):SMP 11–1–SMP 11–13.

Banks, P. M. (1977). Observations of joule and particle heating in the auroral zone. *Journal of Atmospheric and Terrestrial Physics*, 39(2):179–193.

Bharti, G., Sunil Krishna, M. V., Bag, T., and Jain, P. (2018). Storm time variation of radiative cooling by nitric oxide as observed by timed-saber and guvi. *Journal of Geophysical Research: Space Physics*, 123(2):1500–1514.

Bristow, W. A., Hampton, D. L., and Otto, A. (2016). High-spatial-resolution velocity measurements derived using local divergence-free fitting of superdarn observations. *Journal of Geophysical Research: Space Physics*, 121(2):1349–1361.

Bristow, W. A., Sibeck, D. G., Jacquey, C., Greenwald, R. A., Sofko, G. J., Mukai, T., Yamamoto, T., Kokubun, S., Hughes, T. J., Hughes, W. J., and Engebretson, M. J. (1995). Observations of convection vortices in the afternoon sector using the superdarn hf radars. *Journal of Geophysical Research: Space Physics*, 100(A10):19743–19756.

Bruinsma, S. L. and Forbes, J. M. (2010). Large-scale traveling atmospheric disturbances (lstads) in the thermosphere inferred from champ, grace, and seta accelerometer data. *Journal of Atmospheric and Solar-Terrestrial Physics*, 72(13):1057–1066.

Buchau, J., Reinisch, B. W., Weber, E. J., and Moore, J. G. (1983). Structure and dynamics of the winter polar cap f region. *Radio Science*, 18(6):995–1010.

Cai, Y., Yue, X., Wang, W., Zhang, S.-R., Liu, H., Lin, D., Wu, H., Yue, J., Bruinsma, S. L., Ding, F., Ren, Z., and Liu, L. (2022). Altitude extension of the ncar-tiegcm (tiegcm-x) and evaluation. *Space Weather*, 20(11):e2022SW003227. e2022SW003227 2022SW003227.

Carr, J. L., Horvath, A., Wu, D. L., and Friberg, M. D. (2022). Stereo plume height and motion retrievals for the record-setting hunga tonga-hunga ha'apai eruption of 15 january 2022. *Geophysical Research Letters*, 49(9):e2022GL098131. e2022GL098131 2022GL098131.

Charney, J. G. (1971). Geostrophic turbulence. *Journal of Atmospheric Sciences*, 28(6):1087 – 1095.

Chen, C., Chu, X., Zhao, J., Roberts, B. R., Yu, Z., Fong, W., Lu, X., and Smith, J. A. (2016). Lidar observations of persistent gravity waves with periods of 3-10 h in the antarctic middle and upper atmosphere at mcmurdo (77.83s, 166.67e). *Journal of Geophysical Research: Space Physics*, 121(2):1483–1502.

Chi, Y., Zhang, J., Shen, C., Hess, P., Liu, L., Mishra, W., and Wang, Y. (2018). Observational study of an earth-affecting problematic icme from stereo. *The Astrophysical Journal*, 863(1):108.

Chisham, G., Lester, M., Milan, S. E., Freeman, M. P., Bristow, W. A., Grocott, A., McWilliams, K. A., Ruohoniemi, J. M., Yeoman, T. K., Dyson, P. L., Greenwald, R. A., Kikuchi, T., Pinnock, M., Rash, J. P. S., Sato, N., Sofko, G. J., Villain, J.-P., and Walker, A. D. M. (2007). A decade of the super dual auroral radar network (superdarn): scientific achievements, new techniques and future directions. *Surveys in Geophysics*, 28(1):33–109.

Chu, X., Pan, W., Papen, G. C., Gardner, C. S., and Gelbwachs, J. A. (2002). Fe boltzmann temperature lidar: design, error analysis, and initial results at the north and south poles. *Appl. Opt.*, 41(21):4400–4410.

Chu, X. and Yu, Z. (2017). Formation mechanisms of neutral fe layers in the thermosphere at antarctica studied with a thermosphere-ionosphere fe/fe+ (tife) model. *Journal of Geophysical Research: Space Physics*, 122(6):6812–6848.

Chu, X., Yu, Z., Gardner, C. S., Chen, C., and Fong, W. (2011). Lidar observations of neutral fe layers and fast gravity waves in the thermosphere (110-155 km) at mcmurdo (77.8°s, 166.7°e), antarctica. *Geophysical Research Letters*, 38(23).

Codrescu, M. V., Fuller-Rowell, T. J., and Foster, J. C. (1995). On the importance of e-field variability for joule heating in the high-latitude thermosphere. *Geophysical Research Letters*, 22(17):2393–2396.

Codrescu, M. V., Fuller-Rowell, T. J., Foster, J. C., Holt, J. M., and Cariglia, S. J. (2000). Electric field variability associated with the millstone hill electric field model. *Journal of Geophysical Research: Space Physics*, 105(A3):5265–5273.

Codrescu, M. V., Fuller-Rowell, T. J., Munteanu, V., Minter, C. F., and Millward, G. H. (2008). Validation of the coupled thermosphere ionosphere plasmasphere electrodynamics model: Ctipe-mass spectrometer incoherent scatter temperature comparison. *Space Weather*, 6(9).

Cole, K. D. (1975). Energy deposition in the thermosphere caused by the solar wind. *Journal of Atmospheric and Terrestrial Physics*, 37(6):939–949.

Cosgrove, R., McCready, M., Tsunoda, R., and Stromme, A. (2011). The bias on the joule heating estimate: Small-scale variability versus resolved-scale model uncertainty and the correlation of electric field and conductance. *Journal of Geophysical Research: Space Physics*, 116(A9).

Cosgrove, R. B. and Codrescu, M. (2009). Electric field variability and model uncertainty: A classification of source terms in estimating the squared electric field from an electric field model. *Journal of Geophysical Research: Space Physics*, 114(A6).

Cosgrove, R. B., Lu, G., Bahcivan, H., Matsuo, T., Heinselman, C. J., and McCready, M. A. (2009). Comparison of amie-modeled and sondrestrom-measured joule heating: A study in model resolution and electric field-conductivity correlation. *Journal of Geophysical Research: Space Physics*, 114(A4).

Cousins, E. D. P., Matsuo, T., and Richmond, A. D. (2013). Mesoscale and large-scale variability in high-latitude ionospheric convection: Dominant modes and spatial/temporal coherence. *Journal of Geophysical Research: Space Physics*, 118(12):7895–7904.

Cousins, E. D. P., Matsuo, T., Richmond, A. D., and Anderson, B. J. (2015). Dominant modes of variability in large-scale birkeland currents. *Journal of Geophysical Research: Space Physics*, 120(8):6722–6735.

Cousins, E. D. P. and Shepherd, S. G. (2012a). Statistical characteristics of small-scale spatial and temporal electric field variability in the high-latitude ionosphere. *Journal of Geophysical Research: Space Physics*, 117(A3).

Cousins, E. D. P. and Shepherd, S. G. (2012b). Statistical maps of small-scale electric field variability in the high-latitude ionosphere. *Journal of Geophysical Research: Space Physics*, 117(A12).

Cressie, N. A. C. (1993). *Statistics for Spatial Data*, chapter 1, pages 1–26. John Wiley and Sons, Ltd.

Dang, T., Lei, J., Wang, W., Zhang, B., Burns, A., Le, H., Wu, Q., Ruan, H., Dou, X., and Wan, W. (2018). Global responses of the coupled thermosphere and ionosphere system to the august 2017 great american solar eclipse. *Journal of Geophysical Research: Space Physics*, 123(8):7040–7050.

Dang, T., Zhang, B., Lei, J., Wang, W., Burns, A., Liu, H., Pham, K., and Sorathia, K. A. (2021). Azimuthal averaging–reconstruction filtering techniques for finite-difference general circulation models in spherical geometry. *Geoscientific Model Development*, 14(2):859–873.

Deng, Y., Fuller-Rowell, T. J., Ridley, A. J., Knipp, D., and Lopez, R. E. (2013). Theoretical study: Influence of different energy sources on the cusp neutral density enhancement. *Journal of Geophysical Research: Space Physics*, 118(5):2340–2349.

Deng, Y., Huang, Y., Lei, J., Ridley, A. J., Lopez, R., and Thayer, J. (2011). Energy input into the upper atmosphere associated with high-speed solar wind streams in 2005. *Journal of Geophysical Research: Space Physics*, 116(A5).

Deng, Y., Maute, A., Richmond, A. D., and Roble, R. G. (2009). Impact of electric field variability on joule heating and thermospheric temperature and density. *Geophysical Research Letters*, 36(8).

Deng, Y., Richmond, A. D., Ridley, A. J., and Liu, H.-L. (2008). Assessment of the non-hydrostatic effect on the upper atmosphere using a general circulation model (gcm). *Geophysical Research Letters*, 35(1).

Donovan, E., Mende, S., Jackel, B., Frey, H., Syrjasuo, M., Voronkov, I., Trondsen, T., Peticolas, L., Angelopoulos, V., Harris, S., Greffen, M., and Connors, M. (2006). The themis all-sky imaging array-system design and initial results from the prototype imager. *Journal of Atmospheric and Solar-Terrestrial Physics*, 68(13):1472–1487. Passive Optics Aeronomy.

Dungey, J. W. (1961). Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett.*, 6:47–48.

Durran, D. R. (2010). *Numerical methods for fluid dynamics: With applications to geophysics*, volume 32. Springer Science and Business Media.

Eccles, J. V. (1998). Modeling investigation of the evening prereversal enhancement of the zonal electric field in the equatorial ionosphere. *Journal of Geophysical Research: Space Physics*, 103(A11):26709–26719.

Emery, B. A., Coumans, V., Evans, D. S., Germany, G. A., Greer, M. S., Holeman, E., Kadinsky-Cade, K., Rich, F. J., and Xu, W. (2008). Seasonal, kp, solar wind, and solar flux variations in long-term single-pass satellite estimates of electron and ion auroral hemispheric power. *Journal of Geophysical Research: Space Physics*, 113(A6).

Emery, B. A., Lathuillere, C., Richards, P. G., Roble, R. G., Buonsanto, M. J., Knipp, D. J., Wilkinson, P., Sipler, D. P., and Niciejewski, R. (1999). Time dependent thermospheric neutral response tothe 2-11 november 1993 storm period. *Journal of Atmospheric and Solar-Terrestrial Physics*, 61(3):329–350.

Ern, M., Hoffmann, L., Rhode, S., and Preusse, P. (2022). The mesoscale gravity wave response to the 2022 tonga volcanic eruption: Airs and mls satellite observations and source backtracing. *Geophysical Research Letters*, 49(10):e2022GL098626. e2022GL098626 2022GL098626.

Fagundes, P. R., Sahai, Y., Takahashi, H., Gobbi, D., and Bittencourt, J. A. (1996). Thermospheric and mesospheric temperatures during geomagnetic storms at 23°s. *Journal of Atmospheric and Terrestrial Physics*, 58(16):1963–1972. Atmospheric Studies by Optical Methods.

Fan, M., Paul, D., Lee, T. C. M., and Matsuo, T. (2018). Modeling tangential vector fields on a sphere. *Journal of the American Statistical Association*, 113(524):1625–1636.

Fesen, C. G., Crowley, G., Roble, R. G., Richmond, A. D., and Fejer, B. G. (2000). Simulation of the pre-reversal enhancement in the low latitude vertical ion drifts. *Geophysical Research Letters*, 27(13):1851–1854.

Forbes, J. M., Lu, G., Bruinsma, S., Nerem, S., and Zhang, X. (2005). Thermosphere density variations due to the 15-24 april 2002 solar events from champ/star accelerometer measurements. *Journal of Geophysical Research: Space Physics*, 110(A12).

Fritts, D. C. (1984). Gravity wave saturation in the middle atmosphere: A review of theory and observations. *Reviews of Geophysics*, 22(3):275–308.

Fritts, D. C. and Alexander, M. J. (2003). Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, 41(1).

Fritts, D. C. and Lund, T. S. (2011). *Gravity Wave Influences in the Thermosphere and Ionosphere: Observations and Recent Modeling*, pages 109–130. Springer Netherlands, Dordrecht.

Fritts, D. C. and Nastrom, G. D. (1992). Sources of mesoscale variability of gravity waves. part ii: Frontal, convective, and jet stream excitation. *Journal of Atmospheric Sciences*, 49(2):111 – 127.

Fuller-Rowell, T. J. and Evans, D. S. (1987). Height-integrated pedersen and hall conductivity patterns inferred from the tiros-noaa satellite data. *Journal of Geophysical Research: Space Physics*, 92(A7):7606–7618.

Fuller-Rowell, T. J., Rees, D., Quegan, S., Moffett, R. J., and Bailey, G. J. (1987). Interactions between neutral thermospheric composition and the polar ionosphere using a coupled ionosphere-thermosphere model. *Journal of Geophysical Research: Space Physics*, 92(A7):7744–7748.

Gabrielse, C., Nishimura, T., Chen, M., Hecht, J. H., Kaeppler, S. R., Gillies, D. M., Reimer, A. S., Lyons, L. R., Deng, Y., Donovan, E., and Evans, J. S. (2021). Estimating precipitating energy flux, average energy, and hall auroral conductance from themis all-sky-imagers with focus on mesoscales. *Frontiers in Physics*, 9.

Gasque, L. C., Wu, Y.-J., Harding, B. J., Immel, T. J., and Triplett, C. C. (2022). Rapid volcanic modification of the e-region dynamo: Icon's first glimpse of the tonga eruption. *Geophysical Research Letters*, 49(18):e2022GL100825. e2022GL100825 2022GL100825.

Geller, M. A. and Gong, J. (2010). Gravity wave kinetic, potential, and vertical fluctuation energies as indicators of different frequency gravity waves. *Journal of Geophysical Research: Atmospheres*, 115(D11).

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., Tilmes, S., Vitt, F., Bardeen, C. G., McInerny, J., Liu, H.-L., Solomon, S. C., Polvani, L. M., Emmons, L. K., Lamarque, J.-F., Richter, J. H., Glanville, A. S., Bacmeister, J. T., Phillips, A. S., Neale, R. B., Simpson, I. R., DuVivier, A. K., Hodzic, A., and Randel, W. J. (2019). The whole atmosphere community climate model version 6 (waccm6). *Journal of Geophysical Research: Atmospheres*, 124(23):12380–12403.

Gjerloev, J. W. (2009). A global ground-based magnetometer initiative. *Eos, Transactions American Geophysical Union*, 90(27):230–231.

Gjerloev, J. W. (2012). The supermag data processing technique. *Journal of Geophysical Research: Space Physics*, 117(A9).

Golovchanskaya, I. V. and Kozelov, B. V. (2010a). On the origin of electric turbulence in the polar cap ionosphere. *Journal of Geophysical Research: Space Physics*, 115(A9).

Golovchanskaya, I. V. and Kozelov, B. V. (2010b). Properties of electric turbulence in the polar cap ionosphere. *Geomagnetism and Aeronomy*, 50(5):576–587.

Golovchanskaya, I. V., Ostapenko, A. A., and Kozelov, B. V. (2006). Relationship between the high-latitude electric and magnetic turbulence and the birkeland field-aligned currents. *Journal of Geophysical Research: Space Physics*, 111(A12).

Greenwald, R. A., Baker, K. B., Dudeney, J. R., Pinnock, M., Jones, T. B., Thomas, E. C., Villain, J. P., Cerisier, J. C., Senior, C., Hanuise, C., Hunsucker, R. D., Sofko, G., Koehler, J., Nielsen, E., Pellinen, R., Walker, A. D. M., Sato, N., and Yamagishi, H. (1995). Darn/superdarn. *Space Science Reviews*, 71(1):761–796.

Hagan, M. E., Roble, R. G., and Hackney, J. (2001). Migrating thermospheric tides. *Journal of Geophysical Research: Space Physics*, 106(A7):12739–12752.

Hanuise, C., Senior, C., Cerisier, J. C., Villain, J. P., Greenwald, R. A., Ruohoniemi, J. M., and Baker, K. B. (1993). Instantaneous mapping of high-latitude convection with coherent hf radars. *Journal of Geophysical Research: Space Physics*, 98(A10):17387–17400.

Harding, B. J., Wu, Y.-J. J., Alken, P., Yamazaki, Y., Triplett, C. C., Immel, T. J., Gasque, L. C., Mende, S. B., and Xiong, C. (2022). Impacts of the january 2022

tonga volcanic eruption on the ionospheric dynamo: Icon-mighti and swarm observations of extreme neutral winds and currents. *Geophysical Research Letters*, 49(9):e2022GL098577. e2022GL098577 2022GL098577.

Hardy, D. A., Gussenhoven, M. S., and Holeman, E. (1985). A statistical model of auroral electron precipitation. *Journal of Geophysical Research: Space Physics*, 90(A5):4229–4248.

Hardy, D. A., Gussenhoven, M. S., Raistrick, R., and McNeil, W. J. (1987). Statistical and functional representations of the pattern of auroral energy flux, number flux, and conductivity. *Journal of Geophysical Research: Space Physics*, 92(A11):12275–12294.

Heale, C. J., Bossert, K., and Vadas, S. L. (2022). 3d numerical simulation of secondary wave generation from mountain wave breaking over europe. *Journal of Geophysical Research: Atmospheres*, 127(5):e2021JD035413. e2021JD035413 2021JD035413.

Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., Gerber, F., Gramacy, R. B., Hammerling, D., Katzfuss, M., Lindgren, F., Nychka, D. W., Sun, F., and Zammit-Mangion, A. (2019). A case study competition among methods for analyzing large spatial data. *Journal of Agricultural, Biological and Environmental Statistics*, 24(3):398–425.

Heelis, R. A., Lowell, J. K., and Spiro, R. W. (1982). A model of the high-latitude ionospheric convection pattern. *Journal of Geophysical Research: Space Physics*, 87(A8):6339–6345.

Heinselman, C. J. and Nicolls, M. J. (2008). A bayesian approach to electric field and e-region neutral wind estimation with the poker flat advanced modular incoherent scatter radar. *Radio Science*, 43(5).

Heppner, J. P. (1972). Electric field variations during substorms: Ogo-6 measurements. *Planetary and Space Science*, 20(9):1475–1498.

Heppner, J. P., Liebrecht, M. C., Maynard, N. C., and Pfaff, R. F. (1993). High-latitude distributions of plasma waves and spatial irregularities from de 2 alternating current electric field observations. *Journal of Geophysical Research: Space Physics*, 98(A2):1629–1652.

Hines, C. O. (1960). Internal atmospheric gravity waves at ionospheric heights. *Canadian Journal of Physics*, 38(11):1441–1481.

Holton, J. R. (1982). The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere. *Journal of Atmospheric Sciences*, 39(4):791 – 799.

Holton, J. R. (1983). The influence of gravity wave breaking on the general circulation of the middle atmosphere. *Journal of Atmospheric Sciences*, 40(10):2497 – 2507.

Hsu, C.-T., Matsuo, T., Maute, A., Stoneback, R., and Lien, C.-P. (2021). Data-driven ensemble modeling of equatorial ionospheric electrodynamics: A case study during a minor storm period under solar minimum conditions. *Journal of Geophysical Research: Space Physics*, 126(2):e2020JA028539. e2020JA028539 2020JA028539.

Huang, C. Y. and Burke, W. J. (2004). Transient sheets of field-aligned current observed by dmsp during the main phase of a magnetic superstorm. *Journal of Geophysical Research: Space Physics*, 109(A6).

Huang, Y., Wu, Q., Huang, C. Y., and Su, Y.-J. (2016). Thermosphere variation at different altitudes over the northern polar cap during magnetic storms. *Journal of Atmospheric and Solar-Terrestrial Physics*, 146:140–148.

Humberset, B. K., Gjerloev, J. W., Samara, M., and Michell, R. G. (2017). Scale size-dependent characteristics of the nightside aurora. *Journal of Geophysical Research: Space Physics*, 122(2):2455–2466.

Hurd, L. D. and Larsen, M. F. (2016). Small-scale fluctuations in barium drifts at high latitudes and associated joule heating effects. *Journal of Geophysical Research: Space Physics*, 121(1):779–789.

Inchin, P. A., Heale, C. J., Snively, J. B., and Zettergren, M. D. (2022). Numerical modeling of tsunami-generated acoustic-gravity waves in mesopause airglow. *Journal of Geophysical Research: Space Physics*, 127(8):e2022JA030301. e2022JA030301 2022JA030301.

Jin, H., Miyoshi, Y., Fujiwara, H., Shinagawa, H., Terada, K., Terada, N., Ishii, M., Otsuka, Y., and Saito, A. (2011). Vertical connection from the tropospheric activities to the ionospheric longitudinal structure simulated by a new earth's whole atmosphere-ionosphere coupled model. *Journal of Geophysical Research: Space Physics*, 116(A1).

Jones Jr., M., Drob, D. P., Siskind, D. E., McCormack, J. P., Maute, A., McDonald, S. E., and Dymond, K. F. (2018). Evaluating different techniques for constraining lower atmospheric variability in an upper atmosphere general circulation model: A case study during the 2010 sudden stratospheric warming. *Journal of Advances in Modeling Earth Systems*, 10(12):3076–3102.

Kadinsky-Cade, K., Holeman, E., McGarity, J., Rich, F., Denig, W., Burke, W., and Hardy, D. (2004). First results from the ssj5 precipitating particle sensor on dmsp f16: simultaneous observation of kev and mev particles during the 2003 halloween storms. In *Spring Meeting Abstracts*, SH53A-03. American Geophysical Union.

Kelley, M. C. (2009). Edited by. In Kelley, M. C., editor, *The Earth's Ionosphere*, volume 96 of *International Geophysics*, page iii. Academic Press.

Kil, H., Talaat, E. R., Oh, S.-J., Paxton, L. J., England, S. L., and Su, S.-Y. (2008). Wave structures of the plasma density and vertical e × b drift in low-latitude f region. *Journal of Geophysical Research: Space Physics*, 113(A9).

Klemp, J. B. and Durran, D. R. (1983). An upper boundary condition permitting internal gravity wave radiation in numerical mesoscale models. *Monthly Weather Review*, 111(3):430 – 444.

Knipp, D., Eriksson, S., Kilcommons, L., Crowley, G., Lei, J., Hairston, M., and Drake, K. (2011). Extreme poynting flux in the dayside thermosphere: Examples and statistics. *Geophysical Research Letters*, 38(16).

Knipp, D. J. (1989). Quantifying and reducing uncertainty in the assimilative mapping of ionospheric electrodynamics. Technical report, AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH.

Knipp, D. J., Emery, B. A., Richmond, A. D., Crooker, N. U., Hairston, M. R., Cumnock, J. A., Denig, W. F., Rich, F. J., de la Beaujardiere, O., Ruohoniemi, J. M., Rodger, A. S., Crowley, G., Ahn, B. H., Evans, D. S., Fuller Rowell, T. J., Friis Christensen, E., Lockwood, M., Kroehl, H. W., Maclennan, C. G., McEwin, A., Pellinen, R. J., Morris, R. J., Burns, G. B., Papitashvili, V., Zaitzev, A., Troshichev, O., Sato, N., Sutcliffe, P., and Tomlinson, L. (1993). Ionospheric convection response to slow, strong variations in a northward interplanetary magnetic field: A case study for january 14, 1988. *Journal of Geophysical Research: Space Physics*, 98(A11):19273–19292.

Knipp, D. J., Pette, D. V., Kilcommons, L. M., Isaacs, T. L., Cruz, A. A., Mlynczak, M. G., Hunt, L. A., and Lin, C. Y. (2017). Thermospheric nitric oxide response to shock-led storms. *Space Weather*, 15(2):325–342.

Knipp, D. J., Tobiska, W. K., and Emery, B. A. (2004). Direct and indirect thermospheric heating sources for solar cycles 21-23. *Solar Physics*, 224(1):495–505.

Kozelov, B. V. and Golovchanskaya, I. V. (2006). Scaling of electric field fluctuations associated with the aurora during northward imf. *Geophysical Research Letters*, 33(20).

Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119–139.

Larsen, M. F. and Meriwether, J. W. (2012). Vertical winds in the thermosphere. *Journal of Geophysical Research: Space Physics*, 117(A9).

Lastovicka, J. (1996). Effects of geomagnetic storms in the lower ionosphere, middle atmosphere and troposphere. *Journal of Atmospheric and Terrestrial Physics*, 58(7):831–843. Geomagnetic Storms: Their Origin, Mechanism and Ionospheric/Atmospheric Effects.

Lauritzen, P. H., Nair, R. D., Herrington, A. R., Callaghan, P., Goldhaber, S., Dennis, J. M., Bacmeister, J. T., Eaton, B. E., Zarzycki, C. M., Taylor, M. A., Ullrich, P. A., Dubos, T., Gettelman, A., Neale, R. B., Dobbins, B., Reed, K. A., Hannay, C., Medeiros, B., Benedict, J. J., and Tribbia, J. J. (2018). Ncar release of cam-se in cesm2.0: A reformulation of the spectral element dynamical core in dry-mass vertical coordinates with comprehensive treatment of condensates and energy. *Journal of Advances in Modeling Earth Systems*, 10(7):1537–1570.

Lei, J., Roble, R. G., Wang, W., Emery, B. A., and Zhang, S.-R. (2007). Electron temperature climatology at millstone hill and arecibo. *Journal of Geophysical Research: Space Physics*, 112(A2).

Lei, J., Thayer, J. P., Burns, A. G., Lu, G., and Deng, Y. (2010). Wind and temperature effects on thermosphere mass density response to the november 2004 geomagnetic storm. *Journal of Geophysical Research: Space Physics*, 115(A5).

Lei, J., Thayer, J. P., Lu, G., Burns, A. G., Wang, W., Sutton, E. K., and Emery, B. A. (2011). Rapid recovery of thermosphere density during the october 2003 geomagnetic storms. *Journal of Geophysical Research: Space Physics*, 116(A3).

Li, J., Wang, W., Lu, J., Yuan, T., Yue, J., Liu, X., Zhang, K., Burns, A. G., Zhang, Y., and Li, Z. (2018). On the responses of mesosphere and lower thermosphere temperatures to geomagnetic storms at low and middle latitudes. *Geophysical Research Letters*, 45(19):10,128–10,137.

Li, J., Wang, W., Lu, J., Yue, J., Burns, A. G., Yuan, T., Chen, X., and Dong, W. (2019a). A modeling study of the responses of mesosphere and lower thermosphere winds to geomagnetic storms at middle latitudes. *Journal of Geophysical Research: Space Physics*, 124(5):3666–3680.

Li, Z., Knipp, D., and Wang, W. (2019b). Understanding the behaviors of thermospheric nitric oxide cooling during the 15 may 2005 geomagnetic storm. *Journal of Geophysical Research: Space Physics*, 124(3):2113–2126.

Lin, J.-T., Rajesh, P. K., Lin, C. C. H., Chou, M.-Y., Liu, J.-Y., Yue, J., Hsiao, T.-Y., Tsai, H.-F., Chao, H.-M., and Kung, M.-M. (2022). Rapid conjugate appearance of the giant ionospheric lamb wave signatures in the northern hemisphere after hunga-tonga volcano eruptions. *Geophysical Research Letters*, 49(8):e2022GL098222. e2022GL098222 2022GL098222.

Lindgren, F., Rue, H., and Lindstrom, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Lindzen, R. S. (1981). Turbulence and stress owing to gravity wave and tidal breakdown. *Journal of Geophysical Research: Oceans*, 86(C10):9707–9714.

Liu, H.-L., Bardeen, C. G., Foster, B. T., Lauritzen, P., Liu, J., Lu, G., Marsh, D. R., Maute, A., McInerney, J. M., Pedatella, N. M., Qian, L., Richmond, A. D., Roble, R. G., Solomon, S. C., Vitt, F. M., and Wang, W. (2018a). Development and validation of the whole atmosphere community climate model with thermosphere and ionosphere extension (waccm-x 2.0). *Journal of Advances in Modeling Earth Systems*, 10(2):381–402.

Liu, H.-L., Foster, B. T., Hagan, M. E., McInerney, J. M., Maute, A., Qian, L., Richmond, A. D., Roble, R. G., Solomon, S. C., Garcia, R. R., Kinnison, D., Marsh, D. R., Smith, A. K., Richter, J., Sassi, F., and Oberheide, J. (2010). Thermosphere extension of the whole atmosphere community climate model. *Journal of Geophysical Research: Space Physics*, 115(A12).

Liu, H.-L., Lauritzen, P. H., Vitt, F., and Goldhaber, S. (2022a). Thermospheric and ionospheric effects by gravity waves from the lower atmosphere. *Earth and Space Science Open Archive*.

Liu, H.-L., McInerney, J. M., Santos, S., Lauritzen, P. H., Taylor, M. A., and Pedatella, N. M. (2014). Gravity waves simulated by high-resolution whole atmosphere community climate model. *Geophysical Research Letters*, 41(24):9106–9112.

Liu, J., Liu, H., Wang, W., Burns, A. G., Wu, Q., Gan, Q., Solomon, S. C., Marsh, D. R., Qian, L., Lu, G., Pedatella, N. M., McInerney, J. M., Russell III, J. M., and Schreiner, W. S. (2018b). First results from the ionospheric extension of waccm-x during the deep solar minimum year of 2008. *Journal of Geophysical Research: Space Physics*, 123(2):1534–1553.

Liu, J., Wang, W., Oppenheim, M., Dimant, Y., Wiltberger, M., and Merkin, S. (2016). Anomalous electron heating effects on the e region ionosphere in tiegcm. *Geophysical Research Letters*, 43(6):2351–2358.

Liu, X., Xu, J., Yue, J., and Kogure, M. (2022b). Strong gravity waves associated with tonga volcano eruption revealed by saber observations. *Geophysical Research Letters*, 49(10):e2022GL098339. e2022GL098339 2022GL098339.

Liu, X., Yue, J., Wang, W., Xu, J., Zhang, Y., Li, J., Russell III, J. M., Hervig, M. E., Bailey, S., and Nakamura, T. (2018c). Responses of lower thermospheric

temperature to the 2013 st. patrick's day geomagnetic storm. *Geophysical Research Letters*, 45(10):4656–4664.

Lu, G. (2017). Large scale high-latitude ionospheric electrodynamic fields and currents. *Space Science Reviews*, 206(1):431–450.

Lu, G., Baker, D. N., McPherron, R. L., Farrugia, C. J., Lummerzheim, D., Ruohoniemi, J. M., Rich, F. J., Evans, D. S., Lepping, R. P., Brittnacher, M., Li, X., Greenwald, R., Sofko, G., Villain, J., Lester, M., Thayer, J., Moretto, T., Milling, D., Troshichev, O., Zaitzev, A., Odintzov, V., Makarov, G., and Hayashi, K. (1998a). Global energy deposition during the january 1997 magnetic cloud event. *Journal of Geophysical Research: Space Physics*, 103(A6):11685–11694.

Lu, G., Goncharenko, L. P., Richmond, A. D., Roble, R. G., and Aponte, N. (2008). A dayside ionospheric positive storm phase driven by neutral winds. *Journal of Geophysical Research: Space Physics*, 113(A8).

Lu, G., Mlynczak, M. G., Hunt, L. A., Woods, T. N., and Roble, R. G. (2010). On the relationship of joule heating and nitric oxide radiative cooling in the thermosphere. *Journal of Geophysical Research: Space Physics*, 115(A5).

Lu, G., Pi, X., Richmond, A. D., and Roble, R. G. (1998b). Variations of total electron content during geomagnetic disturbances: A model/observation comparison. *Geophysical Research Letters*, 25(3):253–256.

Lu, G., Richmond, A. D., Emery, B. A., and Roble, R. G. (1995). Magnetosphere-ionosphere-thermosphere coupling: Effect of neutral winds on energy transfer and field-aligned current. *Journal of Geophysical Research: Space Physics*, 100(A10):19643–19659.

Lu, G., Zakharenkova, I., Cherniak, I., and Dang, T. (2020). Large-scale ionospheric disturbances during the 17 march 2015 storm: A model-data comparative study. *Journal of Geophysical Research: Space Physics*, 125(5):e2019JA027726. e2019JA027726 2019JA027726.

Lu, X., Chen, C., Huang, W., Smith, J. A., Chu, X., Yuan, T., Pautet, P.-D., Taylor, M. J., Gong, J., and Cullens, C. Y. (2015). A coordinated study of 1 h mesoscale gravity waves propagating from logan to boulder with crrl na doppler lidars and temperature mapper. *Journal of Geophysical Research: Atmospheres*, 120(19):10,006–10,021.

Lu, X., Chu, X., Li, H., Chen, C., Smith, J. A., and Vadas, S. L. (2017). Statistical characterization of high-to-medium frequency mesoscale gravity waves by lidar-measured vertical winds and temperatures in the mlt. *Journal of Atmospheric and Solar-Terrestrial Physics*, 162:3–15. Layered Phenomena in the Mesopause Region.

Luhr, H., Park, J., Gjerloev, J. W., Rauberg, J., Michaelis, I., Merayo, J. M. G., and Brauer, P. (2015). Field-aligned currents' scale analysis performed with the swarm constellation. *Geophysical Research Letters*, 42(1):1–8.

Luhr, H., Rother, M., Kohler, W., Ritter, P., and Grunwaldt, L. (2004). Thermospheric up-welling in the cusp region: Evidence from champ observations. *Geophysical Research Letters*, 31(6).

Lynett, P., McCann, M., Zhou, Z., Renteria, W., Borrero, J., Greer, D., Fa'anunu, O., Bosserelle, C., Jaffe, B., La Selle, S., Ritchie, A., Snyder, A., Nasr, B., Bott, J., Graehl, N., Synolakis, C., Ebrahimi, B., and Cinar, G. E. (2022). Diverse tsunamigenesis triggered by the hunga tonga-hunga ha'apai eruption. *Nature*, 609(7928):728–733.

Maimaiti, M., Ruohoniemi, J. M., Baker, J. B. H., and Ribeiro, A. J. (2018). Statistical study of nightside quiet time midlatitude ionospheric convection. *Journal of Geophysical Research: Space Physics*, 123(3):2228–2240.

Matsuo, T. (2020). *Recent Progress on Inverse and Data Assimilation Procedure for High-Latitude Ionospheric Electrodynamics*, pages 219–232. Springer International Publishing, Cham.

Matsuo, T. and Richmond, A. D. (2008). Effects of high-latitude ionospheric electric field variability on global thermospheric joule heating and mechanical energy transfer rate. *Journal of Geophysical Research: Space Physics*, 113(A7).

Matsuo, T., Richmond, A. D., and Hensel, K. (2003). High-latitude ionospheric electric field variability and electric potential derived from de-2 plasma drift measurements: Dependence on imf and dipole tilt. *Journal of Geophysical Research: Space Physics*, 108(A1):SIA 1–1–SIA 1–15.

Matsuo, T., Richmond, A. D., and Lu, G. (2005). Optimal interpolation analysis of high-latitude ionospheric electrodynamics using empirical orthogonal functions: Estimation of dominant modes of variability and temporal scales of large-scale electric fields. *Journal of Geophysical Research: Space Physics*, 110(A6).

Maute, A., Hagan, M. E., Yudin, V., Liu, H.-L., and Yizengaw, E. (2015). Causes of the longitudinal differences in the equatorial vertical e x b drift during the 2013 ssw period as simulated by the time-gcm. *Journal of Geophysical Research: Space Physics*, 120(6):5117–5136.

Mayr, H. G. and Harris, I. (1978). Some characteristics of electric field momentum coupling with the neutral atmosphere. *Journal of Geophysical Research: Space Physics*, 83(A7):3327–3336.

McFarlane, N. A. (1987). The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *Journal of Atmospheric Sciences*, 44(14):1775 – 1800.

McGranaghan, R., Knipp, D. J., Matsuo, T., Godinez, H., Redmon, R. J., Solomon, S. C., and Morley, S. K. (2015). Modes of high-latitude auroral conductance variability derived from dmsp energetic electron precipitation observations: Empirical orthogonal function analysis. *Journal of Geophysical Research: Space Physics*, 120(12):11,013–11,031.

Mende, S. B., Harris, S. E., Frey, H. U., Angelopoulos, V., Russell, C. T., Donovan, E., Jackel, B., Greffen, M., and Peticolas, L. M. (2008). The themis array of ground-based observatories for the study of auroral substorms. *Space Science Reviews*, 141(1):357–387.

Merkin, V. G. and Lyon, J. G. (2010). Effects of the low-latitude ionospheric boundary condition on the global magnetosphere. *Journal of Geophysical Research: Space Physics*, 115(A10).

Merkine, V. G., Papadopoulos, K., Milikh, G., Sharma, A. S., Shao, X., Lyon, J., and Goodrich, C. (2003). Effects of the solar wind electric field and ionospheric conductance on the cross polar cap potential: Results of global mhd modeling. *Geophysical Research Letters*, 30(23).

Mikkelsen, I. S., Jorgensen, T. S., Kelley, M. C., Larsen, M. F., Pereira, E., and Vickrey, J. (1981). Neutral winds and electric fields in the dusk auroral oval 1. measurements. *Journal of Geophysical Research: Space Physics*, 86(A3):1513–1524.

Millan, L., Santee, M. L., Lambert, A., Livesey, N. J., Werner, F., Schwartz, M. J., Pumphrey, H. C., Manney, G. L., Wang, Y., Su, H., Wu, L., Read, W. G., and Froidevaux, L. (2022). The hunga tonga-hunga ha'apai hydration of the stratosphere. *Geophysical Research Letters*, 49(13):e2022GL099381. e2022GL099381 2022GL099381.

Nastrom, G. D. and Fritts, D. C. (1992). Sources of mesoscale variability of gravity waves. part i: Topographic excitation. *Journal of Atmospheric Sciences*, 49(2):101 – 110.

Newell, P. T., Sotirelis, T., and Wing, S. (2009). Diffuse, monoenergetic, and broadband aurora: The global precipitation budget. *Journal of Geophysical Research: Space Physics*, 114(A9).

Nishimura, Y., Deng, Y., Lyons, L. R., McGranaghan, R. M., and Zettergren, M. D. (2021a). *Multiscale Dynamics in the High-Latitude Ionosphere*, chapter 3, pages 49–65. American Geophysical Union (AGU).

Nishimura, Y., Lessard, M. R., Katoh, Y., Miyoshi, Y., Grono, E., Partamies, N., Sivadas, N., Hosokawa, K., Fukizawa, M., Samara, M., Michell, R. G., Kataoka, R., Sakanoi, T., Whiter, D. K., Oyama, S.-i., Ogawa, Y., and Kurita, S. (2020a). Diffuse and pulsating aurora. *Space Science Reviews*, 216(1):4.

Nishimura, Y. and Lyons, L. R. (2021). *The Active Magnetosphere*, chapter 18, pages 277–291. American Geophysical Union (AGU).

Nishimura, Y., Lyons, L. R., Gabrielse, C., Weygand, J. M., Donovan, E. F., and Angelopoulos, V. (2020b). Relative contributions of large-scale and wedgelet currents in the substorm current wedge. *Earth, Planets and Space*, 72(1):106.

Nishimura, Y., Verkhoglyadova, O., Deng, Y., and Zhang, S.-R. (2021b). *Cross-scale Coupling and Energy Transfer in the Magnetosphere-ionosphere-thermosphere System*. Elsevier.

Nishitani, N., Ruohoniemi, J. M., Lester, M., Baker, J. B. H., Koustov, A. V., Shepherd, S. G., Chisham, G., Hori, T., Thomas, E. G., Makarevich, R. A., Marchaudon, A., Ponomarenko, P., Wild, J. A., Milan, S. E., Bristow, W. A., Devlin, J., Miller, E., Greenwald, R. A., Ogawa, T., and Kikuchi, T. (2019). Review of the accomplishments of mid-latitude super dual auroral radar network (superdarn) hf radars. *Progress in Earth and Planetary Science*, 6(1):27.

Nychka, D., Bandyopadhyay, S., Hammerling, D., Lindgren, F., and Sain, S. (2015). A multiresolution gaussian process model for the analysis of large spatial datasets. *Journal of Computational and Graphical Statistics*, 24(2):579–599.

Oberheide, J., Forbes, J. M., Zhang, X., and Bruinsma, S. L. (2011). Wave-driven variability in the ionosphere-thermosphere-mesosphere system from timed observations: What contributes to the "wave 4"? *Journal of Geophysical Research: Space Physics*, 116(A1).

Omira, R., Ramalho, R. S., Kim, J., Gonzalez, P. J., Kadri, U., Miranda, J. M., Carrilho, F., and Baptista, M. A. (2022). Global tonga tsunami explained by a fast-moving atmospheric source. *Nature*, 609(7928):734–740.

Palmer, T. N., Shutts, G. J., and Swinbank, R. (1986). Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. *Quarterly Journal of the Royal Meteorological Society*, 112(474):1001–1039.

Paxton, L. J. and Meng, C.-I. (1999). Auroral imaging and space-based optical remote sensing. *Johns Hopkins APL technical digest*, 20(4):556–569.

Paxton, L. J., Meng, C.-I., Fountain, G. H., Ogorzalek, B. S., Darlington, E. H., Gary, S. A., Goldsten, J. O., Kusnierkiewicz, D. Y., Lee, S. C., Linstrom, L. A., Maynard, J. J., Peacock, K., Persons, D. F., and Smith, B. E. (1992). Special sensor ultraviolet spectrographic imager: an instrument description. In Chakrabarti, S. and Christensen, A. B., editors, *Instrumentation for Planetary and Terrestrial Atmospheric Remote Sensing*, volume 1745, pages 2 – 15. International Society for Optics and Photonics, SPIE.

Paxton, L. J., Morrison, D., Zhang, Y., Kil, H., Wolven, B., Ogorzalek, B. S., Humm, D. C., and Meng, C.-I. (2002). Validation of remote sensing products produced by the Special Sensor Ultraviolet Scanning Imager (SSUSI): a far UV-imaging spectrograph on DMSP F-16. In Larar, A. M. and Mlynczak, M. G., editors, *Optical Spectroscopic Techniques, Remote Sensing, and Instrumentation for Atmospheric and Space Research IV*, volume 4485, pages 338 – 348. International Society for Optics and Photonics, SPIE.

Picone, J. M., Hedin, A. E., Drob, D. P., and Aikin, A. C. (2002). Nrlmsise-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. *Journal of Geophysical Research: Space Physics*, 107(A12):SIA 15–1–SIA 15–16.

Qian, L., Burns, A. G., Emery, B. A., Foster, B., Lu, G., Maute, A., Richmond, A. D., Roble, R. G., Solomon, S. C., and Wang, W. (2014). *The NCAR TIE-GCM*, chapter 7, pages 73–83. American Geophysical Union (AGU).

Qian, L., Solomon, S. C., and Kane, T. J. (2009). Seasonal variation of thermospheric density and composition. *Journal of Geophysical Research: Space Physics*, 114(A1).

Qian, L., Solomon, S. C., and Mlynczak, M. G. (2010). Model simulation of thermospheric response to recurrent geomagnetic forcing. *Journal of Geophysical Research: Space Physics*, 115(A10).

Rees, M. H. and Luckey, D. (1974). Auroral electron energy derived from ratio of spectroscopic emissions 1. model computations. *Journal of Geophysical Research (1896-1977)*, 79(34):5181–5186.

Rich, F. (1994). Users guide for the topside ionospheric plasma monitor (ssies, ssies-2 and ssies-3) on spacecraft of the defense meteorological satellite program, volume 1: Technical description. Technical report, PHILLIPS LAB HANSCOM AFB MA.

Richmond, A. D. (1992). Assimilative mapping of ionospheric electrodynamics. *Advances in Space Research*, 12(6):59–68.

Richmond, A. D. (1995). Ionospheric electrodynamics using magnetic apex coordinates. *Journal of geomagnetism and geoelectricity*, 47(2):191–212.

Richmond, A. D. and Kamide, Y. (1988). Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique. *Journal of Geophysical Research: Space Physics*, 93(A6):5741–5759.

Richmond, A. D. and Maute, A. (2014). *Ionospheric Electrodynamics Modeling*, chapter 6, pages 57–71. American Geophysical Union (AGU).

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

Richmond, A. D. and Thayer, J. P. (2000). *Ionospheric Electrodynamics: A Tutorial*, pages 131–146. American Geophysical Union (AGU).

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

Ridley, A. J., Lu, G., Clauer, C. R., and Papitashvili, V. O. (1998). A statistical study of the ionospheric convection response to changing interplanetary magnetic field conditions using the assimilative mapping of ionospheric electrodynamics technique. *Journal of Geophysical Research: Space Physics*, 103(A3):4023–4039.

Robinson, R. M. and Vondrak, R. R. (1985). Characteristics and sources of ionization in the continuous aurora. *Radio Science*, 20(3):447–455.

Roble, R. G. and Ridley, E. C. (1987). An auroral model for the ncar thermospheric general circulation model (tgcm). *Annales Geophysicae Series A-Upper Atmosphere And Space Sciences*, 5:369–382.

Roble, R. G. and Ridley, E. C. (1994). A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (time-gcm): Equinox solar cycle minimum simulations (30500 km). *Geophysical Research Letters*, 21(6):417–420.

Roble, R. G., Ridley, E. C., and Dickinson, R. E. (1987). On the global mean structure of the thermosphere. *Journal of Geophysical Research: Space Physics*, 92(A8):8745–8758.

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

Ruohoniemi, J. M. and Baker, K. B. (1998). Large-scale imaging of high-latitude convection with super dual auroral radar network hf radar observations. *Journal of Geophysical Research: Space Physics*, 103(A9):20797–20811.

Sanchez, E. R., Ruohoniemi, J. M., Meng, C.-I., and Friis-Christensen, E. (1996). Toward an observational synthesis of substorm models: Precipitation regions and high-latitude convection reversals observed in the nightside auroral oval by dmsp satellites and hf radars. *Journal of Geophysical Research: Space Physics*, 101(A9):19801–19837.

Schnepf, N. R., Minami, T., Toh, H., and Nair, M. C. (2022). Magnetic signatures of the 15 january 2022 hunga tonga-hunga ha'apai volcanic eruption. *Geophysical Research Letters*, 49(10):e2022GL098454. e2022GL098454 2022GL098454.

Shapiro, R. (1970). Smoothing, filtering, and boundary effects. *Reviews of Geophysics*, 8(2):359–387.

Sheng, C., Deng, Y., Zhang, S.-R., Nishimura, Y., and Lyons, L. R. (2020). Relative contributions of ion convection and particle precipitation to exciting large-scale traveling atmospheric and ionospheric disturbances. *Journal of Geophysical Research: Space Physics*, 125(2):e2019JA027342. e2019JA027342 2019JA027342.

Shi, X., Schmidt, M., Martin, C. J., Billett, D. D., Bland, E., Tholley, F. H., Frissell, N. A., Khanal, K., Coyle, S., Chakraborty, S., Detwiller, M., Kunduri, B., and McWilliams, K. (2022). pydarn: A python software for visualizing superdarn radar data. *Frontiers in Astronomy and Space Sciences*, 9.

Shi, Y., Knipp, D. J., Matsuo, T., Kilcommons, L., and Anderson, B. (2020). Modes of (facs) variability and their hemispheric asymmetry revealed by inverse and assimilative analysis of iridium magnetometer data. *Journal of Geophysical Research: Space Physics*, 125(2):e2019JA027265. e2019JA027265 2019JA027265.

Shim, J. S., Kuznetsova, M., Rastatter, L., Bilitza, D., Butala, M., Codrescu, M., Emery, B. A., Foster, B., Fuller-Rowell, T. J., Huba, J., Mannucci, A. J., Pi, X., Ridley, A., Scherliess, L., Schunk, R. W., Sojka, J. J., Stephens, P., Thompson, D. C., Weimer, D., Zhu, L., and Sutton, E. (2012). Cedar electrodynamics thermosphere ionosphere (eti) challenge for systematic assessment of ionosphere/thermosphere models: Electron density, neutral density, nmf2, and hmf2 using space based observations. *Space Weather*, 10(10).

Solomon, S. C., Burns, A. G., Emery, B. A., Mlynczak, M. G., Qian, L., Wang, W., Weimer, D. R., and Wiltberger, M. (2012). Modeling studies of the impact of high-speed streams and co-rotating interaction regions on the thermosphere-ionosphere. *Journal of Geophysical Research: Space Physics*, 117(A9).

Solomon, S. C. and Qian, L. (2005). Solar extreme-ultraviolet irradiance for general circulation models. *Journal of Geophysical Research: Space Physics*, 110(A10).

Stasiewicz, K. and Potemra, T. (1998). Multiscale current structures observed by freja. *Journal of Geophysical Research: Space Physics*, 103(A3):4315–4325.

Strickland, D. J., Jasperse, J. R., and Whalen, J. A. (1983). Dependence of auroral fuv emissions on the incident electron spectrum and neutral atmosphere. *Journal of Geophysical Research: Space Physics*, 88(A10):8051–8062.

Sutton, E. K., Forbes, J. M., and Nerem, R. S. (2005). Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from champ accelerometer data. *Journal of Geophysical Research: Space Physics*, 110(A9).

Syrjäsuo, M. and Donovan, E. (2002). Analysis of auroral images: Detection and tracking. *Geophysica*, 38(1-2):3–14.

Syrjäsuo, M. T. and Donovan, E. F. (2004). Diurnal auroral occurrence statistics obtained via machine vision. *Annales Geophysicae*, 22(4):1103–1113.

Tanaka, T. (2007). Magnetosphere-ionosphere convection as a compound system. *Space Science Reviews*, 133(1):1–72.

Thayer, J. P. and Semeter, J. (2004). The convergence of magnetospheric energy flux in the polar atmosphere. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(10):807–824. Upper Atmosphere Tutorials from the 2001 Joint CEDAR SCOSTEP Meeting.

Thayer, J. P., Vickrey, J. F., Heelis, R. A., and Gary, J. B. (1995). Interpretation and modeling of the high-latitude electromagnetic energy flux. *Journal of Geophysical Research: Space Physics*, 100(A10):19715–19728.

Themens, D. R., Watson, C., Zagar, N., Vasylkevych, S., Elvidge, S., McCaffrey, A., Prikryl, P., Reid, B., Wood, A., and Jayachandran, P. T. (2022). Global propagation of ionospheric disturbances associated with the 2022 tonga volcanic eruption. *Geophysical Research Letters*, 49(7):e2022GL098158. e2022GL098158 2022GL098158.

Tian, F., Solomon, S. C., Qian, L., Lei, J., and Roble, R. G. (2008). Hydrodynamic planetary thermosphere model: 2. coupling of an electron transport/energy deposition model. *Journal of Geophysical Research: Planets*, 113(E7).

Vadas, S. L. (2013). Compressible f-plane solutions to body forces, heatings, and coolings, and application to the primary and secondary gravity waves generated by a deep convective plume. *Journal of Geophysical Research: Space Physics*, 118(5):2377–2397.

Vadas, S. L. and Fritts, D. C. (2005). Thermospheric responses to gravity waves: Influences of increasing viscosity and thermal diffusivity. *Journal of Geophysical Research: Atmospheres*, 110(D15).

Vadas, S. L. and Liu, H.-l. (2009). Generation of large-scale gravity waves and neutral winds in the thermosphere from the dissipation of convectively generated gravity waves. *Journal of Geophysical Research: Space Physics*, 114(A10).

van Eyken, A. P., Williams, P. J. S., Maude, A. D., and G., M. (1982). Atmospheric gravity waves and sporadic-e. *Journal of Atmospheric and Terrestrial Physics*, 44(1):25–29.

Vierinen, J., Coster, A. J., Rideout, W. C., Erickson, P. J., and Norberg, J. (2016). Statistical framework for estimating gnss bias. *Atmospheric Measurement Techniques*, 9(3):1303–1312.

Wang, W., Burns, A. G., Solomon, S., and Killeen, T. L. (2005). High-resolution, coupled thermosphere-ionosphere models for space weather applications. *Advances in Space Research*, 36(12):2486–2491. Space Weather.

Wang, W., L., K. T., G., B. A., and G., R. R. (1999). A high-resolution, three-dimensional, time dependent, nested grid model of the coupled thermosphere-ionosphere. *Journal of Atmospheric and Solar-Terrestrial Physics*, 61(5):385–397.

Wang, W., Wiltberger, M., Burns, A. G., Solomon, S. C., Killeen, T. L., Maruyama, N., and Lyon, J. G. (2004). Initial results from the coupled magnetosphere-ionosphere-thermosphere model: thermosphere-ionosphere responses. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(15):1425–1441. Towards an Integrated Model of the Space Weather System.

Weber, E. J., Buchau, J., Moore, J. G., Sharber, J. R., Livingston, R. C., Winningham, J. D., and Reinisch, B. W. (1984). F layer ionization patches in the polar cap. *Journal of Geophysical Research: Space Physics*, 89(A3):1683–1694.

Weimer, D. R. (2005). Improved ionospheric electrodynamic models and application to calculating joule heating rates. *Journal of Geophysical Research: Space Physics*, 110(A5).

Wiens, A., Nychka, D., and Kleiber, W. (2020). Modeling spatial data using local likelihood estimation and a matern to spatial autoregressive translation. *Environmetrics*, 31(6):e2652.

Wiltberger, M., Wang, W., Burns, A. G., Solomon, S. C., Lyon, J. G., and Goodrich, C. C. (2004). Initial results from the coupled magnetosphere ionosphere thermosphere model: magnetospheric and ionospheric responses. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(15):1411–1423. Towards an Integrated Model of the Space Weather System.

Wright, C. J., Hindley, N. P., Alexander, M. J., Barlow, M., Hoffmann, L., Mitchell, C. N., Prata, F., Bouillon, M., Carstens, J., Clerbaux, C., Osprey, S. M., Powell, N., Randall, C. E., and Yue, J. (2022). Surface-to-space atmospheric waves from hunga tonga-hunga ha'apai eruption. *Nature*, 609(7928):741–746.

Wu, C., Ridley, A. J., DeJong, A. D., and Paxton, L. J. (2021). Fta: A feature tracking empirical model of auroral precipitation. *Space Weather*, 19(5):e2020SW002629. e2020SW002629 2020SW002629.

Wu, H. and Lu, X. (2022). Data assimilation of high-latitude electric fields: Extension of a multi-resolution gaussian process model (lattice kriging) to vector fields. *Space Weather*, 20(1):e2021SW002880. e2021SW002880 2021SW002880.

Wu, H., Lu, X., Lu, G., Chu, X., Wang, W., Yu, Z., Kilcommons, L. M., Knipp, D. J., Wang, B., and Nishimura, Y. (2020). Importance of regional-scale auroral precipitation and electrical field variability to the storm-time thermospheric temperature enhancement and inversion layer (tteil) in the antarctic e region. *Journal of Geophysical Research: Space Physics*, 125(9):e2020JA028224. e2020JA028224 2020JA028224.

Wu, H., Tan, X., Zhang, Q., Huang, W., Lu, X., Nishimura, Y., and Zhang, Y. (2022). Multiresolution data assimilation for auroral energy flux and mean energy using dmsp ssusi, themis asi, and an empirical model. *Space Weather*, 20(9):e2022SW003146. e2022SW003146 2022SW003146.

Wu, Q., Ortland, D. A., Foster, B., and Roble, R. G. (2012). Simulation of nonmigrating tide influences on the thermosphere and ionosphere with a timed data driven tiegcm. *Journal of Atmospheric and Solar-Terrestrial Physics*, 90-91:61–67. Recent Progress in the Vertical Coupling in the Atmosphere-Ionosphere System.

Wygant, J. R., Keiling, A., Cattell, C. A., Johnson, M., Lysak, R. L., Temerin, M., Mozer, F. S., Kletzing, C. A., Scudder, J. D., Peterson, W., Russell, C. T., Parks, G., Brittnacher, M., Germany, G., and Spann, J. (2000). Polar spacecraft based comparisons of intense electric fields and poynting flux near and within the plasma sheet-tail lobe boundary to uvi images: An energy source for the aurora. *Journal of Geophysical Research: Space Physics*, 105(A8):18675–18692.

Yamazaki, Y., Richmond, A. D., Maute, A., Wu, Q., Ortland, D. A., Yoshikawa, A., Adimula, I. A., Rabiu, B., Kunitake, M., and Tsugawa, T. (2014). Ground magnetic effects of the equatorial electrojet simulated by the tie-gcm driven by timed satellite data. *Journal of Geophysical Research: Space Physics*, 119(4):3150–3161.

Yamazaki, Y., Soares, G., and Matzka, J. (2022). Geomagnetic detection of the atmospheric acoustic resonance at 3.8 mhz during the hunga tonga eruption

event on 15 january 2022. *Journal of Geophysical Research: Space Physics*, 127(7):e2022JA030540. e2022JA030540 2022JA030540.

Yau, A. W. and Andre, M. (1997). Sources of ion outflow in the high latitude ionosphere. *Space Science Reviews*, 80(1):1–25.

Yuan, T., Zhang, Y., Cai, X., She, C.-Y., and Paxton, L. J. (2015). Impacts of cme-induced geomagnetic storms on the midlatitude mesosphere and lower thermosphere observed by a sodium lidar and timed/guvi. *Geophysical Research Letters*, 42(18):7295–7302.

Yue, J., Miller, S. D., Straka III, W. C., Noh, Y.-J., Chou, M.-Y., Kahn, R., and Flower, V. (2022). La soufriere volcanic eruptions launched gravity waves into space. *Geophysical Research Letters*, 49(8):e2022GL097952. e2022GL097952 2022GL097952.

Zeng, Z., Burns, A., Wang, W., Lei, J., Solomon, S., Syndergaard, S., Qian, L., and Kuo, Y.-H. (2008). Ionospheric annual asymmetry observed by the cosmic radio occultation measurements and simulated by the tiegcm. *Journal of Geophysical Research: Space Physics*, 113(A7).

Zhang, J., Xu, J., Wang, W., Wang, G., Ruohoniemi, J. M., Shinbori, A., Nishitani, N., Wang, C., Deng, X., Lan, A., and Yan, J. (2022a). Oscillations of the ionosphere caused by the 2022 tonga volcanic eruption observed with superdarn radars. *Geophysical Research Letters*, 49(20):e2022GL100555. e2022GL100555 2022GL100555.

Zhang, J. J., Wang, W., Wang, C., Lan, A. L., Yan, J. Y., Xiang, D., Zhang, Q. H., Ruohoniemi, J. M., Kunduri, B. S. R., Nishitani, N., Shi, X., and Qiu, H. B. (2020). First observation of ionospheric convection from the jiamusi hf radar during a strong geomagnetic storm. *Earth and Space Science*, 7(1):e2019EA000911. e2019EA000911 10.1029/2019EA000911.

Zhang, S.-R., Erickson, P. J., Zhang, Y., Wang, W., Huang, C., Coster, A. J., Holt, J. M., Foster, J. F., Sulzer, M., and Kerr, R. (2017). Observations of ionneutral coupling associated with strong electrodynamic disturbances during the 2015 st. patrick's day storm. *Journal of Geophysical Research: Space Physics*, 122(1):1314–1337.

Zhang, S.-R., Vierinen, J., Aa, E., Goncharenko, L. P., Erickson, P. J., Rideout, W., Coster, A. J., and Spicher, A. (2022b). 2022 tonga volcanic eruption induced global propagation of ionospheric disturbances via lamb waves. *Frontiers in Astronomy and Space Sciences*, 9.

Zhang, Y. and Paxton, L. J. (2008). An empirical kp-dependent global auroral model based on timed/guvi fuv data. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70(8):1231–1242.

Zhao, Y., Deng, Y., Wang, J.-S., Zhang, S.-R., and Lin, C. Y. (2020). Tropical cyclone-induced gravity wave perturbations in the upper atmosphere: Gitm-r simulations. *Journal of Geophysical Research: Space Physics*, 125(7):e2019JA027675. e2019JA027675 2019JA027675.

Zhu, Q., Deng, Y., Maute, A., Kilcommons, L. M., Knipp, D. J., and Hairston, M. (2021). Ashley: A new empirical model for the high-latitude electron precipitation and electric field. *Space Weather*, 19(5):e2020SW002671. e2020SW002671 2020SW002671.

Zhu, Q., Deng, Y., Richmond, A., and Maute, A. (2018). Small-scale and mesoscale variabilities in the electric field and particle precipitation and their impacts on joule heating. *Journal of Geophysical Research: Space Physics*, 123(11):9862–9872.

Zhu, Q., Deng, Y., Richmond, A., Maute, A., Chen, Y.-J., Hairston, M., Kilcommons, L., Knipp, D., Redmon, R., and Mitchell, E. (2020). Impacts of binning methods on high-latitude electrodynamic forcing: Static versus boundary-oriented binning methods. *Journal of Geophysical Research: Space Physics*, 125(1):e2019JA027270. e2019JA027270 2019JA027270.

Zhu, Y., Bardeen, C. G., Tilmes, S., Mills, M. J., Wang, X., Harvey, V. L., Taha, G., Kinnison, D., Portmann, R. W., Yu, P., Rosenlof, K. H., Avery, M., Kloss, C., Li, C., Glanville, A. S., Millan, L., Deshler, T., Krotkov, N., and Toon, O. B. (2022). Perturbations in stratospheric aerosol evolution due to the water-rich plume of the 2022 hunga-tonga eruption. *Communications Earth and Environment*, 3:248.