Understanding Processes Occurring in the Upper Atmosphere of Mars Using NGIMS Data Analysis

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment
of the requirements for the degree

Doctor of Philosophy

by

Hayley N Williamson

May 2019

APPROVAL SHEET

This Dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Author Signature: Hayley Williamson

This Dissertation has been read and approved by the examining committee:

Advisor: Robert Johnson

Committee Member: Petra Reinke

Committee Member: Julia Spencer

Committee Member: Anne Verbiscer

Committee Member: Meredith Elrod

Committee Member: ______________

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

May 2019

Declaration of Authorship

I, Hayley N. WILLIAMSON, declare that this thesis titled, "Understanding Processes Occurring in the Upper Atmosphere of Mars Using NGIMS Data Analysis" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Hayley Williamson

Date: 04/23/2019

UNIVERSITY OF VIRGINIA

Abstract

Department of Engineering Physics

Doctor of Philosophy

Understanding Processes Occurring in the Upper Atmosphere of Mars Using NGIMS Data Analysis

by Hayley N. WILLIAMSON

Two main processes in the present Martian atmosphere are examined using the data from the Neutral Gas and Ion Mass Spectrometer (NGIMS) on the Mars Atmosphere and Volatile Evolution (MAVEN) spacecraft, orbiting Mars since late 2014. Because of the prediction that Mars atmosphere might have been lost due to its interaction with the solar wind, maps of average atmospheric densities at altitudes of 180-220 km are created using the NGIMS data for three species: O, Ar, and CO_2. The density data are averaged and then organized according to the direction of the solar wind convective electric field, which determines the average direction of flow of the solar wind ions. By mapping the average densities, I look for evidence that solar wind ions that penetrate and collide with the neutral atmosphere affect the neutral densities. However, while the data examined suggest there might be a small effect at present, the evidence is not statistically significant. Since the Martian atmosphere is very thin it is also highly perturbed and the effect of these perturbations are debated. Therefore, on 252 trajectories through the Martian atmosphere large amplitude, high altitude perturbations seen in the NGIMS database are examined. When the perturbations are organized by column density rather than altitude, the perturbations both peak and dissipate at column densities roughly independent of the time of day. Additionally, these perturbations increase the O/CO_2 ratio above that measured for orbits without a significant perturbation. To understand this effect, the perturbations are subsequently categorized by location and found to be roughly consistent with wave activity seen lower in the atmosphere. Because the NGIMS data for each perturbation cannot measure the temperature or long term behavior, model simulations of wave propagation are described based on the Direct Simulation Monte Carlo (DSMC) model. The results from such simulations suggest that these perturbations are most likely large amplitude acoustic gravity waves, whose high frequency and fast phase speed allow them to propagate into the Martian exosphere, affecting the diffusive separation of species and depositing heat.

Acknowledgements

I would like to acknowledge the support of the Jefferson Scholars Foundation, who supported me as a Graduate Fellowship for the majority of this work. The data used in this study are available from the NASA Planetary Data Systems Atmospheres Node and the MAVEN Science Data Center at the University of Colorado Boulder Laboratory for Atmospheric and Space Physics. Additional support was provided by the UVA School of Engineering and Applied Science Teaching Fellowship. I also thank the entire MAVEN team for their gracious assistance with the MAVEN data. Finally, I would like to acknowledge my advisor, Robert Johnson, for his years of help and encouragement.

Support at the University of Virginia comes from NASA's Planetary Data Systems Program via grant NNX15AN38G to create and publish a database on the NASA Planetary Data System Atmospheres Node. This summarized published collision cross sections for use by us and other atmospheric research programs. As part of my graduate program I was also supported for a semester to work with Dr. J. Spencer to develop new techniques for teaching Differential Equations in the Applied Math Department in the School of Engineering and Applied Science. This not only involved creating new lecture notes and in class worksheets, but also giving lectures, holding office hours, writing new tests and then evaluating the changes in performance. We created a collaborative, peer-group focused active learning class to determine if this made a difference in both performance and the willingness of the students to learn. We additionally created multiple short cumulative tests for the class throughout the semester, rather than have midterm exams. At the end of the semester, we found that both the in-class group work and frequent cumulative tests had given our students the ability to have high final grades, with all students but one achieving a final grade above 70%.

Contents

List of Figures

List of Tables

List of Abbreviations

DSMC	**D**irect **S**imulation **M**onte **C**arlo Model
EUV	**E**xtreme **U**ltra**V**iolet Light
IMF	**I**nterplanetary **M**agnetic **F**ield
IUVS	**I**maging **U**ltra**V**iolet **S**pectrograph
KP	**K**ey **P**arameters Data
LST	**L**ocal **S**olar **T**ime
MAVEN	**M**ars **A**tmosphere and **V**olatile **E**volutio**N** Mission
MGS	**M**ars **G**lobal **S**urveyor
MRO	**M**ars **R**econnaisance **O**rbiter
MSE	**M**ars **S**olar **E**lectric Coordinates
MSO	**M**ars **S**olar **O**rbital Coordinates
NGIMS	**N**eutral **G**as and **I**on **M**ass **S**pectrometer
OIMS	**O**rbiter **I**on **M**ass **S**pectrometer
OMAG	**O**rbiter **Mag**netometer
PVO	**P**ioneer **V**enus **O**rbiter
SZA	**S**olar **Z**enith **A**ngle
VSO	**V**enus **S**olar **O**rbital Coordinates

Physical Constants

Boltzmann constant	$k_B = 1.380\,65 \times 10^{-23}\,\mathrm{J\,K^{-1}}$
Molar gas constant	$R = 8.314\,46\,\mathrm{J\,mol^{-1}\,K^{-1}}$
Electron volt	$\mathrm{eV} = 1.602\,176\,6 \times 10^{-19}\,\mathrm{J}$
Dalton	$\mathrm{Da} = 1.660\,539 \times 10^{-27}\,\mathrm{kg}$

List of Symbols

A	amplitude	%
b	impact parameter	m
BV	Brunt-Väisälä frequency	s^{-1}
c_p	heat capacity at constant P	$J\,K^{-1}$
c_v	heat capacity at constant V	$J\,K^{-1}$
F	force	N
g	gravitational acceleration	$m\,s^{-2}$
H	scale height	m
Kn	Knudsen number	unitless
k_x	horizontal wavenumber	m^{-1}
k_z	vertical wavenumber	m^{-1}
m	mass	kg
n	number density	m^{-3}
p	pressure	Pa
Q	heat	J
Ri	Richardson number	unitless
T	temperature	K
t	time	s
$\vec{u}$	flow velocity	$m\,s^{-1}$
V	volume	m^3
v	velocity	$m\,s^{-1}$
W	work	J
z	altitude	m
β	plasma pressure/magnetic pressure	unitless
Γ_{ad}	adiabatic lapse rate	$K\,m^{-1}$
γ	c_p/c_v	unitless
Θ	potential temperature	K
ρ	mass density	$kg\,m^{-3}$
σ	collision cross section	m^2
σ_{ij}	stress tensor	$N\,m^{-2}$
τ_{if}	stress tensor force	N
ϕ	phase angle	rad
ω	angular frequency	rad

Dedicated to my long-suffering husband Alex, who has supported me through a difficult journey, and my doctors, for keeping me intact.

Chapter 1

Introduction

The atmospheres of terrestrial planets are well-studied phenomena. However, for the atmospheres of unmagnetized planets such as Venus and Mars, special consideration must be given to the effect of solar weather on the upper atmosphere. Such effects include erosion of the atmosphere through atmospheric escape, the addition of energy via solar extreme ultraviolet radiation or ions carried in the solar wind which heat the atmosphere, and stripping of the ionosphere by the interplanetary magnetic field, which is created via photoionization of the upper atmosphere. In addition to solar weather, a significant source of vertical energy transfer is wave activity, specifically propagating gravity waves, which refer to perturbations that propagate through the atmosphere with gravity as the restoring force. These waves are ubiquitous in all atmospheres, but are best understood on Earth due to the relative ease of observation. However, with the Mars Atmosphere and Volatile EvolutioN (MAVEN) mission, *in situ* analysis of the Martian atmosphere is now possible, and both gravity waves and the effects of solar weather have been detected (e.g. Thiemann et al. (2015), Curry et al. (2015), England et al. (2017), and Terada et al. (2017), etc.). While there have been many previous Mars missions, including both orbiters and landers, the MAVEN mission provides the first true opportunity for ongoing, long term *in situ* study of the Martian atmosphere, as previous orbiters such as Mars Global Surveyor and Mars Reconnaissance Orbiter have only obtained relevant upper atmospheric density data during their aerobraking phases (Bougher et al., 1999; Fritts, Wang, and Tolson, 2006). As a result, the scientific community continues to develop an increasing understanding of atmospheres on the terrestrial planets and how atmospheric escape and wavelike processes can cause them to change over time. In this thesis I focus on processes in the Martian atmosphere, with an examination of the atmosphere of Venus found in Appendix B.

1.1 Motivations and Background

1.1.1 Atmospheric escape and evolution

It is evident when studying large-scale geology that at one time Mars must have had significantly more water. Surface features provide evidence of past glaciation in Hellas Planitia, tributary fluvial networks in the cratered southern highlands, and outflow channels are indicative of immense flows of liquid water (Baker, 2001). At present, briny water is only stable for short periods of time on the surface due to the low atmospheric pressure. This suggests that much of the atmosphere must therefore have been lost to space over the past 4.5 billion years (Johnson and Liu, 1996). Studies of argon isotopes seem to confirm that the majority of the missing atmosphere has been lost to space, or escaped, rather than become chemically locked into rocks (Jakosky et al., 1994; Jakosky et al., 2017). Therefore, to understand the atmospheric evolution of an unmagnetized terrestrial planet, which has implications for habitability, one must understand both the past and current rates of atmospheric escape.

Atmospheric escape occurs when a particular molecule is given enough energy to escape the gravitational well of a planet. This generally occurs in the exosphere, the uppermost layer of an atmosphere where particles follow ballistic trajectories, rather than collide with each other. The very rough boundary of the exosphere, called the exobase, is defined as the region where the scale height of the atmosphere (i.e. the vertical distance by which atmospheric pressure and density decrease by a factor of $1/e$) is equal to the mean free path of a particle (Johnson, Schnellenberger, and Wong, 2000; Jakosky et al., 2017). There are many ways a particle in the exosphere can attain enough energy to escape a planet's gravity well, including by interacting with the particles in the incident solar wind. The cause of atmospheric escape also differs depending on if the particle is neutral or ionized. Methods of atmospheric escape for the neutral portion of an atmosphere are shown in Figure 1.1 from Jakosky et al. (2015) and include:

- Jeans escape, when the thermal energy of an atom exceeds the gravitational escape energy. It is most effective for light atoms, such as hydrogen, on lower mass planets, such as Mars (Chassefière and Leblanc, 2004; Walterscheid, Hickey, and Schubert, 2013; Brain et al., 2016);

- Dissociative recombination, in which photochemical dissociation and recombination of an atmospheric molecular ion (such as O_2^+) give an end result of fast-moving neutral particles, which can then exceed the gravitational escape energy (Fox, 2004; Fox and Hać, 2009; Lillis et al., 2015); and

- Sputtering, when ions traveling along the interplanetary magnetic field (IMF) contained within the solar wind either directly precipitate or charge exchange producing precipitating pick-up O^+ that collide with neutral species. Such collisions can add enough energy to the neutral particle for it to escape (Luhmann, Johnson, and Zhang, 1992; Johnson, 1994; Leblanc and Johnson, 2002).

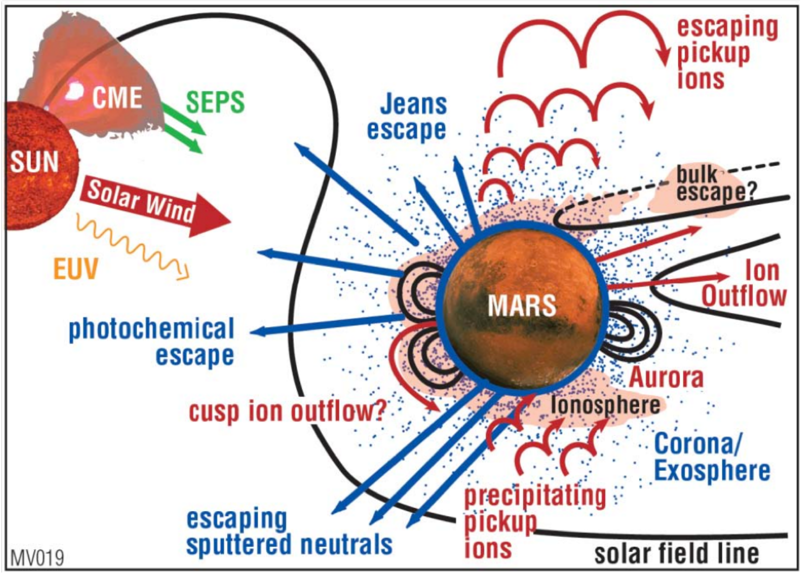


FIGURE 1.1: Different methods of escape possible at Mars, from Jakosky et al. (2015).

While much of the work contained in this dissertation does not directly affect the atmospheric escape rate, the atmospheric features studied here can do so indirectly by modifying the composition and the heating rate of the region of the atmosphere from which escape occurs. Since escape is a crucial and ubiquitous process in the Martian upper atmosphere, it continues to be an important area of research. The possibility of seeing the effects of ion precipitation on the neutral species is examined in Chapter 3, as well as the density changes induced in the upper atmosphere, which Walterscheid, Hickey, and Schubert (2013) show can induce Jeans escape of neutral species. I focus on two areas affecting the exobase region: the effect of an incoming plasma, an exogenic effect, and then the effect of perturbations that propagate into this regime, an endogenic effect.

Another important aspect of an atmosphere is its wave activity. While there are multiple types of waves possible in a planetary atmosphere, including tidal phenomena and Kelvin-Helmholtz waves, this work focuses on gravity waves, which occur when an instability propagates both vertically and horizontally with gravity (or, more technically, buoyancy) as the restoring force. While gravity waves are a non-linear process, they can often be reasonably well described using linear plane wave theory. In Appendix A, the wave equation for gravity waves is derived, as well as describing how gravity waves propagate and change the atmosphere. The appendix also contains information on atmospheric parameters such as the Brunt-Väisälä frequency, which is the frequency for which a parcel of air will oscillate in an atmosphere due to gravity and the buoyant force under adiabatic conditions.

Gravity waves are present in all planetary atmospheres, as they arise from instabilities, which can be caused by, for example, air flow over topography. Their most important role in a planetary atmosphere is the transfer of energy from the lower atmosphere to the upper atmosphere (Nappo, 2013; Walterscheid, Hickey, and Schubert, 2013; Hickey, Walterscheid, and Schubert, 2011; Midgley and Liemohn, 1966; Fritts, 1984; Fritts and Alexander, 2003; Tolstoy, 1963; Charney and Drazin, 1961; Hines, 1960). A wave oscillating under the influence of gravity can be more specifically categorized as either a *propagating gravity wave* or an *acoustic wave* (Hines, 1960; Midgley and Liemohn, 1966). The difference between the two lies in their phase speeds and oscillatory frequencies. For an acoustic wave, the phase speed is greater than the speed of sound (c) in the atmosphere, while the frequency is greater than the adiabatic frequency, i.e. the Brunt-Väisälä (BV) frequency. For a propagating (previously called internal) gravity wave, the frequency must be less than the atmospheric Brunt-Väisälä frequency and the phase speed less than the speed of sound. If the properties of a wave fall in between these two categories, e.g. the phase speed is greater than c but the frequency is less than the BV frequency, the wave does not propagate but decays and is known as an *evanescent wave*. Evanescent waves do not carry energy and so do not affect the temperature or composition of the atmosphere, thus it is important to ascertain the type of wave seen in observations. This has not previously been clarified in such works as England et al. (2017), Terada et al. (2017), and Yiğit et al. (2015) and so it is unknown what type of wave is most common in the Martian atmosphere. Tidal waves, with periods being an integer fraction value of the planetary rotation period, have also been detected in the Martian atmosphere (England et al., 2016), but as will be shown in Chapter 7, the waves of interest for this work do not have periods consistent with tidal phenomena. Additionally, the waves studied here are unlikely to be generated by topography, as waves propagating in the lower atmosphere generally dissipate in an unstable region of the atmosphere known as the turbopause (Slipski et al., 2018). While I do not discuss potential generation mechanisms in detail in the work, it is most likely at these altitudes that the generation mechanisms are temperature gradients which induce a disturbance.

Gravity waves (here referring to both propagating gravity waves and acoustic waves), however, can carry significant amounts of energy into the thermosphere and ionosphere. This energy transfer mainly occurs when the wave begins to dissipate due to turbulence, also known as saturation (Vincent, 2009). Acoustic waves generally propagate energy along the direction of phase speed, while the direction of energy propagation for a propagating gravity wave can vary more widely, including opposite to the phase depending on the mode of the wave (Hines, 1960). Walterscheid, Hickey, and Schubert (2013) and Hickey, Walterscheid, and Schubert (2011) show that both types of gravity waves can contribute significantly to the heat budget of the upper thermosphere, up to several hundred K/day, which can then in turn affect the Jeans escape rate. The amount of heat added to the thermosphere is dependent on the phase speed and intrinsic wave frequency, as some waves will actually create a net cooling effect as they dissipate (Hickey, Walterscheid, and Schubert, 2011). Wave saturation is also an important source of turbulence in the middle and upper atmospheres (Hodges, 1967). These facts indicate that gravity waves can be significant contributors to atmospheric escape and evolution that must be considered in order to gain a fuller understanding of the Martian atmosphere.

1.1.2 MAVEN Mission

The Mars Atmosphere and Volatile EvolutioN (MAVEN) mission was launched on November 18, 2013, with Mars orbital insertation occurring on September 22, 2014 (Jakosky et al., 2015). It was developed for the purpose of studying the Martian atmosphere and interaction with the solar wind. The mission's science objectives are to "measure the composition and structure of the upper atmosphere and ionosphere today, and determine the processes responsible for controlling them; measure the rate of loss of gas from the top of the atmosphere to space, and determine the processes responsible for controlling them; and determine properties and characteristics that will allow us to extrapolate backwards in time to determine the integrated loss to space over the four-billion-year history recorded in the geological record" (Jakosky et al., 2015). It began science operations in November 2014 and continues in extended missions. The spacecraft contains eight science instruments, several of which were instrumental to the work here. The following instruments' data were used for this dissertation:

- NGIMS, the Neutral Gas and Ion Mass Spectrometer, is the primary instrument of interest. NGIMS is a quadrupole mass spectrometer and will be discussed in more detail below (Mahaffy et al., 2015b).

- MAG, the dual vector fluxgate magnetometer, which measures magnetic field direction and magnitude in the vicinity of the spacecraft, is used when examining the effects of the solar wind on the neutral atmosphere. This instrument is part of the Particles and Fields Package (Connerney et al., 2015; Jakosky et al., 2015).

- EUV, the extreme ultraviolet monitor, is also used as a measure of solar activity. The EUV instrument is part of the Langmuir probe and is used at apoapsis to determine solar EUV output (Eparvier et al., 2015; Jakosky et al., 2015).

- SWIA, the Solar Wind Ion Analyzer, is used to find the velocity of the solar wind, used in calculations of the solar wind convective electric field (see Chapter 3). This instrument, also part of the Particles and Fields package, measures the energy and angular distribution of solar wind ions, including protons (Halekas et al., 2015; Jakosky et al., 2015).

The Key Parameters, or KP, dataset (*MAVEN Insitu Key Parameters Data Bundle* 2019) was used to obtain the data from instruments other than NGIMS, while the NGIMS Level 2 data products were used for neutral densities (Elrod, 2015).

The spacecraft has a nominal periapsis of 150 km and apoapsis of 6220 km, with a period of 4.5 hours. This highly elliptical orbit allows for study of both the Martian atmosphere and the solar wind environment, including the magnetosheath and bow shock, regions of interaction between the planet and the IMF. The orbit precesses to allow observations at periapsis during all local solar times (LST) and latitudes. Instruments such as EUV and SWIA predominately observe during the apoapse portion of the orbit, while NGIMS only operates on the periapse segment (Jakosky et al., 2015; Eparvier et al., 2015; Mahaffy et al., 2015b).

The NGIMS instrument was of particular interest for this investigation, as the long term goal is identifying the dominant sources of neutral escape in the Martian upper atmosphere. NGIMS is the descendant of many other spacecraft mass spectrometers, such as the Ion and Neutral Mass Spectrometer onboard the Cassini spacecraft at Saturn (Waite et al., 2005). The instrument has two modes for detecting neutral species, closed source (CSM) and open source neutral beaming (OSNB). In closed source mode, a hot ionizing filament is coupled to an antechamber with a small aperture, thus allowing neutrals to be ionized and flow into the quadrupole analyzer and detector portion of the instrument. In open source neutral beaming mode, the incoming particles are collimated into a beam rather than be allowed to collide in an antechamber; this prevents more reactive species such as O and N from undergoing collisions and, hence, possibly reacting. After being ionized by the filament, the particles pass into the quadrupole mass filter, which measures a mass-to-charge ratio ranging from 1.5 to 150 Da with a 0.1 Da resolution. Finally, the mass-filtered particles are beamed into the detector assembly, measuring particle counts per second, which are then converted to mass-specific densities (Mahaffy et al., 2015b). As some molecular species such as CO_2 will dissociate in the instrument, as well as some species having the same mass-to-charge ratio, the mass densities are calculated using isotopic mass fractionations found from laboratory calibrations (Benna and Elrod, 2017).

NGIMS primarily operates in science mode below 500 km in altitude, due to the extremely low atmospheric densities at high altitudes. In open source mode when the filament is on, the instrument cannot detect ions; therefore, the filament must be off in open source ion (OSION) mode so ions can flow into the quadrupole mass spectrometer portion of the instrument. This allows simultaneous detections of ions and neutrals by switching between CSN and OSION modes. In order to conduct observation of the reactive neutral species such as O and CO_2, NGIMS operated in two nominal orbits: switching between

CSN and OSION, known as ion observation, and CSN and OSNB, or neutral observation. These reactive species are able to stick to the sides of the antechamber and form more CO_2, thus increasing the background for O, CO_2, CO, and C. The effect increases throughout pass through the atmosphere and is quickly dissipated once out of the atmosphere. As a result, the outbound segments of the NGIMS density profiles have an artificially elevated background. Therefore for each orbit, the outbound densities can be higher than the inbound densities, so for all of the analysis in this work, only the inbound portion of each orbit is used. Background is subtracted from the densities for the higher-level (i.e. instrument counts have been converted to density) NGIMS data products below approximately 350 km. Above these altitudes, the densities are low enough for instrument background to outweigh the atmospheric signal. Including only the inbound densities at altitudes where the signal-to-noise ratio is high ensures that any features seen are due to actual atmospheric phenomena, rather than instrument background.

The MAVEN mission is invaluable for the study of the Martian atmosphere because it provides the first continuous set of *in situ* atmospheric data, in addition to providing context for changes in the atmosphere by collecting data on the solar wind environment. While this work does not directly study the main escape processes, the atmospheric density profiles found in the NGIMS dataset provide a crucial look at short-term atmospheric phenomena, such as the high-altitude gravity waves that are the focus of this dissertation. Previous missions such as the Mars Reconnaissance Orbiter and Mars Global Surveyor made observations of the atmosphere during their aerobraking phases, including observations of wave activity (Bougher et al., 1999; Fritts, Wang, and Tolson, 2006), but these observations were brief. MAVEN, however, orbits through the atmospheric region of interest continuously and so allows for long term observation of both the response to solar activity and wave-like processes.

1.2 Research Objectives

The aim of this dissertation was twofold: first, to ascertain if continuous ion precipitation has a visible, long term effect on the global neutral densities as seen by NGIMS (Chapter 3) and, second, to study the effect of multiple large amplitude density perturbations in the exosphere. The first goal was obtained by examining the average densities of O, Ar, and CO_2 in the altitude range of 180-220 km in a coordinate system dependent on the direction of the solar wind electric field, thus potentially indicating an effect due to ion precipitation. The second goal was accomplished using several objectives:

1. Determining criteria for identifying an exospheric perturbation in the NGIMS data based on altitude and amplitude;

2. Characterizing these perturbations in terms of location and solar wind conditions from both NGIMS and KP data (Chapter 6);

3. Calculating the column density of the altitude vs. density profiles using the Newton method of integration as a way to better categorize exospheric perturbations in relation to the exobase (Chapter 4);

4. Creating a profile of the O/CO_2 ratio for exospheric perturbation orbits and comparing it to that for orbits without a significant exospheric perturbation to see how exospheric perturbations change atmospheric composition on average over multiple orbits (Chapter 4);

5. Retrieving the frequency and wavelength of the exospheric perturbations and comparing to the BV frequency in order to understand the type of wave (Chapter 7);

6. Modeling a wave with similar parameters to those seen in the NGIMS data in a two-component (O and CO_2) Direct Simulation Monte Carlo (DSMC) simulation to understand how these exospheric perturbations affect the energy balance of the atmosphere (Chapters 5 and 7).

The modeling for this work was performed by Ludivine Leclercq, a postdoctoral fellow who created a one-dimensional multi-species DSMC model capable of simulating the vertical propagation of a disturbance, as well as Lucia Tian, an undergraduate engineering student. The model was used to create a density perturbation with an amplitude in the simulated atmosphere similar to that observed by MAVEN and then to vary the wave frequency of the perturbation to study its effect on propagation and energy transfer. Additionally, the Leclercq et al. (2019) paper included in this dissertation in Chapter 5 shows that the typical method of deriving temperature from a NGIMS density versus altitude profile, which calculates a pressure profile assuming hydrostatic equilibrium and then uses the ideal gas law to obtain temperature, can drastically differ from the actual kinetic temperature of a gas being perturbed as calculated in the model.

1.3 Research Value

While atmospheric escape and gravity waves are well-studied fields, the research in this dissertation contains several aspects for both that have not yet been examined. Firstly, while many attempts have been made to find evidence of sputtering, the escape of neutrals due to collisions with precipitating ions, in Martian atmospheric data (e.g. Leblanc et al. (2015)), these generally look for evidence of sputtering within selected density profiles. Here, rather than see how density changes in altitude may indicate sputtering, a single altitude range is selected to examine the global trends in density. There are many reasons the average density in the upper atmosphere of Mars can change; the most obvious reason being solar insolation, making the night side of the planet colder than the day side. However, to help indicate how solar wind factors may influence the density, the average density data are mapped in coordinates dependent on the direction of the solar wind electric field, which determines the motion of ions and hence ion precipitation. This creates a map looking for the potential average effects of ion precipitation on the neutral atmosphere rather than seeking individual orbits where precipitation may have occurred. Creating an average density map helps account for the stochastic nature of ion precipitation, which varies in location and frequency (Hara et al., 2017). I believe that this is a valuable approach, as sputtering was suggested to be, on average, a large contributor to Martian atmospheric escape in previous solar epochs (Luhmann, Johnson, and Zhang, 1992) and would have occurred globally. Taking this approach allowed for the examination potential sputtering in a similarly global context.

The population of orbits that contain the exospheric perturbations described in this dissertation have yet to be studied, as they differ from the more common lower amplitude thermospheric waves seen in, for example, England et al. (2017), Terada et al. (2017), and Yiğit et al. (2015). The perturbations examined here are outliers from those in other studies of gravity waves due to both their large amplitudes, which is required to be larger than 40% of the background density profile (see Chapter 4), and their appearance in the exosphere, where the atmosphere is no longer collisional. Indeed, even studies of gravity waves on Earth, where wave activity has been studied for decades (e.g. Hines (1960), Bretherton (1969), Midgley and Liemohn (1966), and Tolstoy (1963)), do not discuss exospheric waves, as the typical fluid physics used for theoretical examination is no longer valid. Thus the physics used at lower altitudes must be adapted to understand both how these waves propagate into the ballistic regime of the atmosphere and how they could potentially deposit energy. Similar data analysis methods have been used for other populations of gravity waves seen in NGIMS data (e.g. Terada et al. (2017), etc.), which I apply here to this unique subset of the data. Additionally, these waves can then be modeled with molecular kinetic simulations which, unlike the fluid models more typically used for gravity wave analysis (e.g. England et al. (2017) and Garcia et al. (2017)), is valid at the altitudes where the gas is rarefied and is the region from which escape is most likely to occur.

Previous studies of gravity waves found in the MAVEN NGIMS data also lack a discussion of potential compositional changes induced by vertical propagation, which could potentially arise as a result of heating or cooling by the wave. The NGIMS data profiles are first reconstructed in terms of column density rather than altitude, which helps remove the dependence of density on solar insolation and background temperature. The density at a given altitude may vary significantly with local time due to changes in scale height induced by the background temperature. Column density, however, is integrated vertically and so helps to account for scale height variations. Organizing by column density allows for the creation of an average density versus column density profile and calculate the ratio between O and CO_2, the main constituents of the upper atmosphere. The O/CO_2 ratio always increases with altitude due to the changing scale heights of the various species, but this ratio increases significantly more quickly for orbits with a significant exospheric perturbation, indicating that such perturbations are capable of changing the atmospheric composition, a likely indicator of energy deposition.

I also calculate the intrinsic wave frequency and the average BV frequency for each perturbation orbit so they can be categorized as propagating or acoustic gravity waves, which has previously only been done with Earth data or model results (Midgley and Liemohn, 1966; Hickey, Walterscheid, and Schubert, 2011; Walterscheid, Hickey, and Schubert, 2013). This presents another opportunity to further understand the energy deposition of these exospheric perturbations with implications for atmospheric escape. The temperature change is expected to be small, but previous work has shown that a small effect can integrate over time to be a significant contribution to the energy budget of the upper atmosphere (Walterscheid, Hickey, and Schubert, 2013). Determining the intrinsic wave and BV frequencies also allows us to more accurately simulate a wave in the DSMC simulation created by Ludivine Leclercq and run by Lucia Tian. Using model simulations gives further opportunities for understanding the physics occurring rather than using the data alone, which is limited by being a single track in time and, therefore, cannot show the evolution and eventual dissipation of a single wave, although it can be inferred. The model simulations, however, can track the progression of the wave in both density and temperature, therefore expanding the amount of information available for analysis of the effect of large amplitude perturbations in the exosphere of Mars.

1.4 Overview

In the following work, I first give in Chapter 2 a theoretical background for my research, including an overview of atmospheric structure, the full definition of the exobase, and a description of the DSMC model.

Williamson et al. (2019)a, Chapter 3, examines the average densities of three species, O, Ar, and CO_2, chosen for their prominence in Mars' upper atmosphere and, in the case of Ar, for its chemical stability, so that changes in Ar density can be inferred to be due to non-chemical processes, in an altitude corridor of 180-220 km. This altitude range is consistent with the most common definitions of the exobase. These average densities are mapped in a local time based physical coordinate system and then on a solar wind electric field based coordinate system, as well as separating the data by season. It is seen that mapping the data in solar wind electric field coordinates suggests that ion precipitation raises the average density of the neutral species, but it is not a strong conclusion.

Leclercq et al. (2019), Chapter 5, uses a 1D DSMC model to show that current methods of temperature extraction from atmospheric density data are not valid for a non-steady state atmosphere. In order to simulate an exospheric perturbation, the density at the lower boundary is varied by a chosen amplitude, with a period corresponding to the steady-steady atmospheric BV frequency. The results show the evolution in time of both density and temperature for a single-species O atmosphere and a two species mixed atmosphere. These data is then compared to NGIMS data for one of the exospheric perturbation profiles. It is seen that the derived temperature is out of phase from the actual temperature of the particles in the model no matter the amplitude of the induced perturbation. The model results also show the amplitude of the perturbation growing with amplitude, which would be expected from the linear perturbation theory of gravity waves; however, the amount of growth with altitude does differ from that described in, for example, Hines (1960), suggesting that horizontal transport is likely an important aspect of wave dissipation at some point above the exobase. My contribution to this paper was in the form of theory, giving the first author the theoretical framework for modeling the gravity wave and calculating the BV frequency. I additionally provided information on the choice of perturbation amplitude and frequency so that the model results were comparable to those seen on Mars, as well as contributing data analysis and comparison to the model data.

Chapter 4 contains objectives 1, 3, and 4 from the second research aim, categorizing large amplitude exospheric perturbations. I define an exospheric perturbation by creating a smooth background fit for each density profile, then subtracting it from the data to get the amplitude and show an example from a MAVEN orbit. This amplitude profile is then converted to in terms of total atomic column density, i.e. the column density of O plus three times the column density of CO_2. Doing so allows averaging all of the large amplitude profiles together, which shows that these exospheric perturbations both peak and dissipate at consistent column densities. Using this averaged profile, I also calculate O/CO_2 for both perturbation and non-perturbation orbits, which show that this ratio is substantially increased when an orbit contains an exospheric perturbation, suggesting these perturbations do deposit energy in the exosphere.

Next, Chapter 6 gives more information about the exospheric perturbations, fulfilling objective 2. I compare the distribution of the perturbations in local time/solar zenith angle, latitude, and solar wind conditions, including the solar wind magnetic field and proton velocity vector. These variables are also used as a comparison for the species-specific amplitude of the perturbations, with O generally having an amplitude roughly half that of Ar and CO_2 due to scale height differences. This chapter shows that overall, the exospheric perturbations are unsurprisingly similar in distribution to the thermospheric gravity waves seen in Terada et al. (2017), leading to the conclusion that these are most likely particularly large specimens of gravity waves that have been able to escape dissipating before reaching the exobase.

Finally, Chapter 7 examines the frequencies of the exospheric perturbations described in previous chapters as compared to the background atmospheric BV frequency. This is used to categorize the perturbations as either propagating or acoustic gravity waves. The retrieved frequencies of the data are used to motivate the simulations of waves with similar frequencies in the DSMC model from Chapter 5. The simulations show that changing the intrinsic frequency of a wave-like perturbation does not significantly change the accuracy of the temperature typically extracted from the perturbed density data. However, it has a large effect on both the amount of energy transferred vertically by the wave and the amount of time it takes for a given wave to dissipate at high altitudes. The simulation results are then used to infer how much energy these exospheric perturbations add to the exosphere.

The dissertation also includes two appendices. The first contains a full derivation of the physics of gravity waves, primarily focused on the vertical dimension. The derivation includes an explanation of the BV frequency and wave amplitude growth with increasing altitude. The second appendix summarizes work done before the current data analysis and examines a peculiar atmospheric phenomenon in the ionosphere of Venus as seen by the Pioneer Venus Orbiter spacecraft, which orbited the planet from 1978-1992. The data show that ionospheric "holes", where the plasma pressure of the ionosphere drops and the magnetic field carried in the ionosphere increases, occasionally appeared in the nightside ionosphere. In the appendix, I examine the location of these "holes", more formally known as low β regions, where β is defined as the ratio of plasma pressure to magnetic pressure, in coordinates rotated according to the solar wind magnetic field, as well as finding the convective electric field across these regions. This work was largely done with Dr. J. Grebowsky at NASA Goddard Spaceflight Center from 2011-2014.

Chapter 2

Understanding the Martian Atmosphere

2.1 Atmospheric composition and structure

The first *in situ* measurements of the Martian atmosphere were performed by the Viking 1 and 2 landers as they descended to the surface via onboard mass spectrometers. The mass spectrometers measured six neutral species, $CO_2, N_2, CO, O_2,$ NO and Ar, measuring the densities as a function of altitude (Nier and McElroy, 1977; Izakov, 1978). The Viking landers showed that CO_2 is the dominant species in the atmosphere, as well as a "sinuous shape" in the density profile that Nier and McElroy (1977) believed to be indicative of atmospheric tidal waves. Due to uncertainties about the adsorption of O in the Viking neutral mass spectrometers, an O profile was never calculated from the Viking data but its presence was inferred, as O is a typical byproduct of CO_2 and O_2 dissociation (Withers et al., 2015). Other missions such as Mars Global Surveyor and Mars Reconnaissance Orbiter also took atmospheric density measurements during their aerobraking phases in 1997-1998 and 2006, respectively, which showed similar composition to that seen by Viking. Additionally, both spacecraft saw wave-like features in the density data, inferred to be gravity waves (Bougher et al., 1999; Fritts, Wang, and Tolson, 2006). However, prior to the MAVEN mission, no spacecraft was able to continuously measure atmospheric properties for long periods of time, hence limiting understanding of both the atmospheric response to the solar wind and the effect of wave activity.

The NGIMS observations from the MAVEN mission also show that the bulk of the Martian atmosphere is composed of CO_2. The atmosphere additionally contains other species such as O, Ar, N_2, CO, O_2, NO, N, H_2O and He with their various isotopes, all able to be measured by NGIMS (Mahaffy et al., 2015a). While work is ongoing to understand the original CO_2 budget of the atmosphere using isotope ratios, it is clear that the atmosphere was once much more robust and able to support the presence of liquid water, as discussed in the introduction (Jakosky et al., 1994). It is likely that atmospheric loss to space began after the interior of the planet cooled sufficiently to stop the internal dynamo and hence protective planetary magnetic field, allowing the solar wind to interact directly with the atmosphere (Chassefière and Leblanc, 2004).

Much like Earth's, Mars' atmosphere can be roughly divided into layers based on the temperature profile as a function of altitude. On Earth, the lower and upper atmosphere are separate by the layer called the stratosphere, where the temperature increases with altitude due to the absorption of UV light by the atmospheric ozone layer. Because of the temperature profile the stratosphere is a stable regime with very little vertical transport (Holton, 2015). Mars, however, lacks a stratosphere, allowing for a strong coupling between the lower atmosphere, the troposphere, and the middle-to-upper atmosphere, consisting of the mesosphere, thermosphere, and exosphere, in ascending order (Chassefière and Leblanc, 2004).

The density and pressure of a single component, isothermal atmosphere follows the *law of atmospheres* or barometric law, assuming the atmosphere is in hydrostatic equilibrium. The derivation of this law is found in Appendix A, with the result being that density and pressure fall exponentially with altitude when T is a constant, i.e.

$$p(z) = p_0 e^{-\frac{mg}{k_b T} z} \tag{2.1}$$

where p is pressure, p_0 is the pressure at a chosen lower boundary altitude, k_b is the Boltzmann constant, T the atmospheric temperature, m the atomic mass, and g the gravitational acceleration. Using the ideal gas law a similar dependence applies for mass and number density as a function of altitude. The *scale height* of the atmosphere, then, is the distance for which density falls by a factor of $1/e$ and is given by the inverse of the z coefficient in the exponent of the above equation:

$$H = \frac{k_b T}{mg} \tag{2.2}$$

The steady-state temperature of the troposphere generally decreases with altitude due to vertical transport, while in the thermosphere temperature can rapidly increase due to absorption of UV light. However, in the upper thermosphere and exosphere, the atmosphere becomes roughly isothermal. Temperatures may vary slightly by species due to changes in scale height, but does not generally vary with altitude in the absence of perturbations.

In the troposphere and mesosphere, below the region of the atmosphere known as the homopause, the scale heights of each species are the same because the atmosphere is well-mixed and so a single scale height can be found using the average atmospheric mass. When densities are high, the atmosphere can be well-mixed because *eddy diffusion* due to turbulence dominates over *molecular diffusion* due to the thermal motions of individual particles, which is a function of temperature and particle mass (Green, 1999; Brown, 1991). The homopause is defined as the region where the molecular diffusion coefficient is equal to the eddy diffusion coefficient, and so thermal motion, dependent on species mass, begins to dominate over mixing due to turbulence (Green, 1999). As a result, the species no longer have the same scale heights. This can occur anywhere from 60-120km altitude in the Martian atmosphere, with recent estimates found by extrapolating the N_2/Ar ratio measured by the NGIMS instrument (Slipski et al., 2018).

While CO_2 dominates all other species even above the homopause, as the species continue to separate, O with its smaller mass has a larger scale height than CO_2 and so does not fall off as quickly with altitude. Additionally, CO_2 photodissociates at high altitudes, ensuring a large population of O, unlike species such as H or He, which are rarer in the atmosphere. At extreme altitudes of several Mars radii, the H in the atmosphere forms an extended corona because of its small mass (Halekas et al., 2017). As a result of the growing dominance of O over CO_2 with altitude, there is a point in the upper atmosphere where O becomes the dominant species, usually between approximately 230-270km depending on the background temperature of the atmosphere. The temperature is dependent on solar input and so is much colder on the nightside of the planet than near the subsolar region. This in turn affects the estimates of the scale height, the homopause and the altitude at which O becomes dominant (Mahaffy et al., 2015a; Slipski et al., 2018). At very high altitudes, the O density is sustained by the dissociative recombination of O_2^+ at various energies (Tully and Johnson, 2001):

$$
\begin{aligned}
O_2^+ + e^- &\rightarrow O(^3P) + O(^3P) && (6.98eV) && (2.3)\\
&\rightarrow O(^3P) + O(^1D) && (5.02eV) && (2.4)\\
&\rightarrow O(^3P) + O(^1S) && (2.79eV) && (2.5)\\
&\rightarrow O(^1D) + O(^1D) && (3.05eV) && (2.6)\\
&\rightarrow O(^1D) + O(^1S) && (0.83eV) && (2.7)
\end{aligned}
$$

S, D, and P here refer to the electron configuration of the resulting oxygen atoms, with $O(^3P)$ being the most energetic configuration and $O(^1S)$ being the least energetic configuration. Any collision that results in energies greater than 2eV will have sufficient energy to escape from the atmosphere; thus dissociative recombination is a significant source of atmospheric loss at Mars (Deighan, J et al., 2015; Lillis et al., 2017; Chassefière and Leblanc, 2004; Lee et al., 2015; Brain et al., 2016). Oxygen atoms that lack the energy to escape populate the *oxygen corona*, which extends out to several planetary radii and, while not detectable in NGIMS data, can be seen using the Imaging Ultraviolet Spectrograph (IUVS) instrument on MAVEN (Deighan, J et al., 2015). It is believed, although so far difficult to prove, that sputtering, the addition of energy to a neutral atom through collision with a precipitating ion, also contributes to the corona (Johnson, Schnellenberger, and Wong, 2000; Leblanc and Johnson, 2001).

While the corona is extensive, it is also low density, with the O number density being approximately 10^2 cm^-3 above 300 km altitude (Lee et al., 2015). Thus the corona is well below the densities at which the atmosphere transitions from collisional to dominated by ballistic transport, or the exosphere. The exosphere is the primary atmospheric region of importance for this work, as it is the portion of the atmosphere from which escape is most likely to occur. The exosphere and upper thermosphere are capable of more directly interacting with the solar wind and EUV at Mars; solar wind protons typically collide with ions in the exosphere (Rahmati et al., 2015), while precipitating ions will typically collide with neutrals to add heat and expand the atmosphere, in addition to potentially causing sputtering, when a neutral collision with a precipitating ion provides enough energy for the neutral to escape (Johnson, 1994).

The different methods of escape are only some of the ongoing processes in the Martian atmosphere. At low altitudes, there may be sublimation of surface and sub-surface ice (Dundas et al., 2018), as well as clouds and fogs, both seen by the Mars Science Laboratory Curiosity rover (Kloos et al., 2018). There is also large scale general circulation, which transfers energy into the middle and upper atmosphere. In the upper atmosphere, neutrals are photoionized by solar EUV, creating an ionosphere, which also presents a boundary to the solar wind, resulting in an induced magnetosphere from the draping of interplanetary magnetic field lines (Lillis et al., 2015). The ionosphere and neutral atmosphere also undergo charge exchange and other chemical reactions such as dissociative recombination described above (Chassefière and Leblanc, 2004). Gravity waves, discussed extensively in this work, are ubiquitous in all layers of the atmosphere, being generated whenever there is an instability. All of these processes serve as methods to transport energy throughout the atmosphere, creating a complex and dynamic system.

2.1.1 Collisions in the atmosphere

A key aspect of any gas is the collisions between constituent atoms. Collisions allow the atmosphere to remain well-mixed below the homopause and transfer energy both horizontally and vertically. For two atoms to collide, they must be within the *collision cross section*, the effective area surround a particle where scattering (i.e. a changing of the velocity of one or more of the particles) will occur (Griffiths, 1995). In its simplest form, a collision between two particles can be assumed to be between two spheres, known as the hard sphere approximation. For many situations, this cross section suffices, as specificity is not required. The hard sphere cross section is defined as $\sigma = \pi(2r)^2$, where r is the radius of the atomic hard sphere. The hard sphere approach assumes that an incident particle will interact with a target particle of the same radius in a circle with radius twice that of the target particle (Griffiths, 1995). For $O + O$, this number is generally taken to be $2 \times 10^{-15} cm^2$. However, the hard sphere approximation does not take into account the energies of either the incident or target species, which can drastically change the interaction size, nor does it consider the angle of incidence between the two particles, which can change the result of the scattering.

The cross section and the density of the gas determine, on average, the distance a particle can travel in the gas before undergoing a collision, called the *mean free path*. We can determine the mean free path by considering the hard sphere cross section and the distance a particle will travel during a length of time t with velocity v. We assume that, if the particle is traveling in a straight line, the target particle will sweep out a cylindrical interaction region with volume $V = v \cdot t \cdot \sigma$. The number of collisions can then be written as the number of particles in the interaction region, i.e. $v \cdot t \cdot \sigma \cdot n$ where n is the number density of the gas. The mean free path, then, will be the length the particle travels divided by the number of collisions:

$$mfp = \frac{v\,t}{v\,t \cdot \sigma \cdot n} = \frac{1}{\sigma n} \tag{2.8}$$

Therefore, if we know the cross section and number density of an atmosphere, we can easily calculate the average distance a particle will travel before it interacts with another. However, in a multispecies atmosphere such as Mars, choosing a cross section can become complicated, as each type of collision will have a differently sized interaction region.

One way to simplify the cross section calculations is to use an "atomic" cross section; that is, treat a molecule as multiple atoms, which therefore ignores internal modes of energy in the molecule due to rotation and vibration. This is possible for O and CO_2 at thermal energies of a few eV seen in the upper atmosphere (Luhmann, Johnson, and Zhang, 1992; Johnson, 1990b; Fox and Hać, 2009). Additionally, because the range of energies for O in the altitudes of interest is so low, the cross section for O in the Martian atmosphere varies very little with energy (Luhmann, Johnson, and Zhang, 1992; Fox and Hać, 2009; Lewkow and Kharchenko, 2014). In the Martian atmosphere, we can roughly treat CO_2 as three oxygen atoms for the purposes of collisions, since C has a similar mass and radius to O. Treating CO_2 as three O atoms allows us to use the O + O cross section as an approximate total collision cross section for the purposes of calculating a rough exobase. Using the O cross section to approximate the cross section for CO_2 is useful for the upper atmosphere of Mars, as it greatly simplifies calculation of column density and location of the exobase. We use this method when calculating the total atomic column density in Chapter 4 for simplification purposes, with $\sigma_O = 2 \times 10^{-15}$ cm^2.

While using a hard sphere approximation is sufficient for our data analysis, we choose to use more detailed cross sections in our model, to be discussed. Scattering theory examines collisions with three possible outcomes: elastic collisions, inelastic collisions, and absorption. Here, however, we will only focus on elastic collisions, as inelastic collisions and absorption are rare for collisions of the energies typical in the Martian atmosphere (Lewkow and Kharchenko, 2014; Leblanc et al., 2017).

For an elastic collision, we consider the number of particles n per unit time scattered into an element of solid angle $d\Omega$ in the spherical coordinates (θ, ϕ) given by $N\ d\Omega$. The number of particles will be dependent on the flux of particles into the solid angle element, which we can define as J, i.e. the number of particles per unit time crossing a unit area perpendicular to the direction of particle motion. We can then define the *differential cross section* as the ratio of the number of particles scattered in the (θ, ϕ) direction per unit time per unit solid angle to the incident flux:

$$\frac{d\sigma}{d\Omega} = \frac{n}{J} \tag{2.9}$$

From the differential cross section, we can obtain the *total cross section* by integrating across all solid angles Ω:

$$\sigma = \int \frac{d\sigma}{d\Omega}\, d\Omega = \int_0^{2\pi} \int_0^{\pi} \frac{d\sigma}{d\Omega}\, d\theta\, d\phi \tag{2.10}$$

where θ is the polar angle and ϕ is the azimuthal angle. In scattering theory, the calculation of $n\ d\Omega$ and subsequently $d\sigma$ is dependent on the *impact parameter* b, which determines the angle of scattering. If b is defined using classical theory, the resulting cross section will be the hard sphere cross section defined above. However, using quantum mechanics to provide a better atomic potential will change the impact parameter and provide a cross section as a function of collision energies (Griffiths, 1995). These energy-dependent cross sections may be calculated theoretically or using laboratory experiments. As mentioned in Chapter 1, this work was partially supported by creating a database of published cross section collisions as a function of energy for a variety of species and energies for the NASA Planetary Data System Atmospheres Node.

2.1.2 Exobase calculations

To study the effect of large amplitude perturbations in the exosphere as seen in NGIMS data, we first define the location where the exosphere begins in the data, the exobase. As the exosphere is defined as the region of the atmosphere where atomic motion is ballistic rather than collisional, the exobase is defined as the location where the mean free path of a particle equals the species scale height; that is, the density is low enough for that a particles on a ballistic trajectory has a 1/e chance of a collision. This indicates that the exobase location will vary by species; throughout this work, the exobase for O will be used. Using the definitions of mean free path and scale height, we can write this as:

$$mfp = \frac{1}{\sigma n} = H = \frac{k_b T}{mg} \tag{2.11}$$

$$\implies n = \frac{mg}{kT\sigma} \tag{2.12}$$

where n is then the number density at the exobase, mfp is the mean free path, σ is the collision cross section, H is the species scale height, k_b the Boltzmann constant, T the temperature, m the species mass, and g the gravitational acceleration. However, this method relies on knowing the background temperature or scale height of the atmosphere, which can be difficult when the atmosphere is not in hydrostatic equilibrium. One method of finding the mean atmospheric scale height is by fitting the altitude-density profile at two points z_1 and z_2, with z_1 being the lower altitude, i.e.

$$n(z_2) = n(z_1)\, e^{-(z_2-z_1)/H} \tag{2.13}$$

$$\implies \ln\left(\frac{n(z_2)}{n(z_1)}\right) = -(z_2-z_1)/H \tag{2.14}$$

$$\implies H = \frac{z_1 - z_2}{\ln\left(n(z_2)/n(z_1)\right)} \tag{2.15}$$

However, this is an average scale height and may not be representative of the perturbed state of the atmosphere.

In an atmosphere with multiple species, like the Mars atmosphere, the atomic column density can be used using the "atomic" cross section for O + O as described above. The column density N, or the number of atoms per unit area is defined by integrating the number density downwards:

$$N = \int_{z_0}^{\infty} n(z)\, dz \sim \sum_{z_0}^{\infty} n_i\, z_i \tag{2.16}$$

where z is the altitude. Thus, writing column density as atoms/area, we can define the exobase as the location where

$$N \sim \frac{1}{\sigma} \tag{2.17}$$

for our chosen σ. If σ is the O cross section and we treat Ar and CO_2 as having the same cross section as one and three O, respectively, we can calculate a total atomic column density N_{tot} profile and the corresponding exobase using

$$N_{tot} = N_O + N_{Ar} + 3 \cdot N_{CO2} = \frac{1}{\sigma} \tag{2.18}$$

Using our O cross section of 2×10^{-15} cm^2, our resulting total atomic column density of the exobase is 5×10^{14} cm^{-2} $= 5 \times 10^{18}$ m^{-2} (Johnson, 1990a; Luhmann, Johnson, and Zhang, 1992). This definition of the exobase is used throughout this work, specifically in Chapter 4, as a rough indicator that the atmosphere is transitioning to a ballistic regime.

Jakosky et al. (2017) defined the exobase similarly, but they instead chose to use the $Ar + CO_2$ cross section of 3×10^{-15} cm^2 for comparison to the mean free path. However, given that O is nearly an order of magnitude higher in density than Ar at the altitudes of interest, as well as being the species most likely to escape, we believe using the O + O cross section is better as a guideline for where the exosphere begins for our purposes (Fox and Hać, 2009). It is important to note here than any definition of the exobase is only a rough estimate. Additionally, a given column density integrated from the NGIMS data will only be an average at one particular time, whereas the upper atmosphere of Mars is highly variable. So we use this definition of the exobase as a guide only to determine if a perturbation is in the region of the atmosphere where escape is most likely to originate.

In order to use the NGIMS data to estimate the exobase, we must first compute the total atomic column density of each profile using the midpoint method of numerical integration. First we fit each density profile with a 5$^{\text{th}}$ order polynomial to create a smooth background profile. This method is similar to that used in, for example, Yiğit et al. (2015), and is also used to compute the amplitude of density perturbations in the profile. We then use this background density profile to compute the total atomic column density, integrating from high altitudes to low altitudes:

$$N(z_i) = \left[n_b \left(\frac{z_i + z_{i-1}}{2} \right) \right] \cdot (z_{i-1} - z_i) \tag{2.19}$$

$$N_{tot}(z) = N_O(z) + N_{Ar}(z) + 3 \cdot N_{CO_2}(z) \tag{2.20}$$

$N(z_i)$ is the total atomic column density at the ith point, $n_b\left(\frac{z_i+z_{i-1}}{2}\right)$ is the density at the midpoint of points i and i-1, and $(z_{i-1} - z_i)$ is the change in altitude between the two points. By finding where this column density is equal to 5×10^{18} m^{-2}, we can estimate a guideline exobase, as seen in Figure 4.1 in Chapter 4.

2.2 Gravity waves

Propagating waves in the upper atmosphere were first inferred on Earth from observations of perturbations in meteor trails at high altitudes (Hines, 1960). These waves were understood to be upper atmospheric gravity waves, a type of linear wave in a stratified, compressible fluid under the influence of gravity (Tolstoy, 1963). The derivation of the linear fluid dynamics of gravity waves is outlined in Appendix A. However, here we detail key concepts for our understanding of the perturbations seen in NGIMS data studied here: the Brunt-Väisälä (BV) frequency and the dispersion relations for gravity waves in the upper atmosphere, as well as the vertical transport of energy via wave dissipation or saturation.

Firstly, the BV frequency is that at which parcel of air will oscillate adiabatically in a background atmosphere, i.e. without the loss or addition of heat. The BV frequency ω_{bv}

in an isothermal atmosphere is given by the equation

$$\omega_{bv} = \sqrt{-\tfrac{g}{n}\tfrac{dn}{dz}} \tag{2.21}$$

where g is gravitational acceleration, z altitude, and n the number density (Nappo, 2013). From the barometric law of atmospheres, we know that

$$\frac{dn}{n} = -\frac{dz}{H} \implies -\frac{dn}{dz}\frac{1}{n} = \frac{1}{H} \tag{2.22}$$

and so the BV frequency can be calculated from the atmospheric scale height (Nappo, 2013).

2.2.1 Dispersion relation

In order to properly describe gravity waves, we must find a *dispersion relation*, the equation for a wave that relates the wavelength or wavenumber to the wave frequency. For gravity waves, a dispersion relation is particularly useful because it indicates the type of gravity wave, due to the dispersion equation having more than one branch of solution. Because of the complex non-linearity of an atmosphere, several assumptions are made to derive this equation: that the atmosphere is isothermal, which is roughly true for our region of interest; that the atmosphere is roughly quiescent; and that the thermal conductivity is nearly constant, which is also roughly true for the upper thermosphere and exosphere. These assumptions allow for a linear approximation of the gravity wave, although they are not valid for a dissipating wave, which, due to turbulence, is an inherently non-linear process and will be discussed further later. The following derivation primarily follows that found in Midgley and Liemohn (1966).

In Appendix A, we give the one dimensional governing fluid equations for a parcel of air. Here, we begin by using those equations for a two dimensional (vertical and horizontal) first order perturbation. First, we have the fluid equations of continuity, momentum, energy, and state. Bold with an arrow indicates a vector quantity and $\frac{D}{Dt}$ indicates the total derivative, i.e. $\frac{\partial}{\partial t} + \vec{\boldsymbol{V}} \cdot \nabla$ where $\vec{\boldsymbol{V}}$ is the perturbation flow velocity.

$$\frac{D\rho}{Dt} + \rho\nabla\cdot\vec{\boldsymbol{V}} = 0 \tag{2.23}$$

$$\rho\,\frac{D\vec{\boldsymbol{V}}}{Dt} = \rho\,\vec{\boldsymbol{g}} - \nabla p + \nabla\cdot\vec{\boldsymbol{S}} - 2\,\rho\,\vec{\boldsymbol{\Omega}}\times\vec{\boldsymbol{V}} \tag{2.24}$$

$$\frac{\rho k_b}{(\gamma-1)m}\frac{DT}{Dt} = Q + \nabla\cdot(\lambda\nabla T) - p\nabla\cdot\vec{\boldsymbol{V}} + \vec{\boldsymbol{S}}\cdot\nabla\vec{\boldsymbol{V}} \tag{2.25}$$

and then our state equation is the usual ideal gas law, $p = \frac{\rho k_b T}{m}$. Here ρ is density, p pressure, $\vec{V}$ the perturbation velocity, $\vec{g}$ gravitational acceleration, $\vec{S}$ the viscous stress tensor, $\vec{\Omega}$ the planetary rotation rate, k_b the Boltzmann constant, γ ratio of specific heats, Q is scalar heat, λ thermal conductivity, T temperature, and m mass.

We will make several assumptions to make solving the equations possible, as shown in Midgley and Liemohn (1966).

1. We can ignore planetary rotation, so $\Omega = 0$. This is possible because the frequency of the perturbation will be much larger than the Coriolis frequency.
2. The oscillations are small enough to make linearization possible.
3. The unperturbed quantities (e.g. density, temperature, pressure, etc.) will only vary in the vertical direction.
4. The unperturbed variables are at rest, so there is no background wind.
5. Composition and viscosity of the atmosphere do not vary, so $\frac{dm}{dt} = 0$ and the viscous stress tensor can be ignored.
6. Finally, we assume all perturbed variables vary harmonically in the x and z directions, i.e. horizontally and vertically.

Using the assumptions, we can create our first order perturbation variables assuming a wavelike perturbation, using a zeroth order term only varying in altitude and the first order term as the wave propagating in space and time:

$$\rho = \rho_0(z) \left[1 + A\, R(z)\, e^{i(\omega t - k_x x)}\right] \tag{2.26}$$

$$p = p_0(z) \left[1 + A\, P(z)\, e^{i(\omega t - k_x x)}\right] \tag{2.27}$$

$$T = T_0(z) \left[1 + A\, \hat{T}(z)\, e^{i(\omega t - k_x x)}\right] \tag{2.28}$$

$$\vec{V} = A\, U(z)\, e^{i(\omega t - k_x x)}\, \hat{x} + A\, W(z)\, e^{i(\omega t - k_x x)}\, \hat{z} \tag{2.29}$$

A is an arbitrary amplitude, small enough for only linear terms to be considered, and R, P, T, U, and W are vertically varying coefficients of the harmonic perturbation, then ω and k_x are the perturbation frequency and horizontal wavenumber, respectively. The first terms are the zeroth order, i.e. $\rho_0(z), p_0(z), T_0(z)$, that only vary with altitude and the second with exponential terms are the first order perturbation terms. We can easily plug in the zeroth order terms to find the zeroth order equations, since the unperturbed system is at rest. The continuity equation in 2.23 will be equal to zero on both sides of the equation since $\frac{D\rho_0(z)}{Dt} = 0$ and there is no zeroth order $\vec{V}$ term, so is not shown. Plugging the zeroth order terms for our perturbed variables into equations 2.24 and 2.25 gives

$$0 = -\rho_0(z)g(z) - \frac{dp}{dz} \tag{2.30}$$

$$0 = Q + \lambda\frac{d^2T}{dz^2} + \frac{d\lambda}{dz}\frac{dT}{dz} \tag{2.31}$$

$$p_0(z) = \frac{\rho_0(z)k_bT_0(z)}{m} \tag{2.32}$$

Now, we find the perturbation equations by plugging the first order, perturbed terms into our fluid equations. We will only show the full derivation for the continuity equation 2.23, as they all follow similarly. First, we find the total derivative of ρ. A prime indicates the partial z derivative, where z is altitude.

$$\frac{D\rho}{Dt} = \frac{\partial\rho}{\partial t} + \vec{V}\cdot\nabla\rho \tag{2.33}$$

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial t}\left(\rho_0(z) + \rho_0(z)\,A\,R(z)e^{i(\omega t-k_xx)}\right) \tag{2.34}$$

$$\frac{\partial\rho}{\partial t} = i\omega\rho_0\,A\,Re^{i(\omega t-k_xx)} \tag{2.35}$$

$$\begin{aligned}\nabla\rho &= \frac{\partial}{\partial x}\left(\rho_0(z) + \rho_0(z)\,A\,R(z)e^{i(\omega t-k_xx)}\right)\hat{x}\\ &\quad+\frac{\partial}{\partial z}\left(\rho_0(z) + \rho_0(z)\,A\,R(z)e^{i(\omega t-k_xx)}\right)\hat{z}\end{aligned} \tag{2.36}$$

$$\begin{aligned}\nabla\rho &= -ik_x\rho_0A\,R(z)\,e^{i(\omega t-k_xx)}\hat{x} + \rho_0'(z)\hat{z} + \rho_0'(z)\,A\,R(z)e^{i(\omega t-k_xx)}\hat{z}\\ &\quad+\rho_0(z)\,A\,R'(z)e^{i(\omega t-k_xx)}\hat{z}\end{aligned} \tag{2.37}$$

$$\begin{aligned}\vec{V}\cdot\nabla\rho &= \left[A\,U(z)\,e^{i(\omega t-k_xx)}\hat{x}\right] * \left[-ik_x\rho_0A\,R(z)\,e^{i(\omega t-k_xx)}\hat{x}\right]\\ &\quad+\left[A\,W(z)\,e^{i(\omega t-k_xx)}\hat{z}\right] * \left[\rho_0'(z) + \rho_0'(z)\,A\,R(z)e^{i(\omega t-k_xx)}\hat{z}\right]\end{aligned} \tag{2.38}$$

$$\begin{aligned}\vec{V}\cdot\nabla\rho &= A^2e^{2i(\omega t-k_xx)}\left[-ik_x\rho_0(z)\,R(z)\,U(z) + \rho_0'(z)\,R(z)\,W(z) + \rho_0(z)\,R'(z)\,W(z)\right]\\ &\quad+\rho_0'(z)\,A\,W(z)e^{i(\omega t-k_xx)}\end{aligned} \tag{2.39}$$

We can cancel the non-linear A terms to get from 2.35 and 2.39

$$\frac{D\rho}{Dt} = i\omega\rho_0(z)\,A\,R(z)e^{i(\omega t-k_xx)} + \rho_0'\,A\,We^{i(\omega t-k_xx)} \tag{2.40}$$

We then calculate the second term in the equation, $\rho\nabla\cdot\vec{V}$.

$$\nabla\cdot\vec{V} = \frac{\partial}{\partial x}\left(A\,U(z)e^{i(\omega t-k_x x)}\right) + \frac{\partial}{\partial z}\left(A\,W(z)e^{i(\omega t-k_x x)}\right) \tag{2.41}$$

$$\nabla\cdot\vec{V} = -ik_x A\,U(z)e^{i(\omega t-k_x x)} + A\,W'(z)e^{i(\omega t-k_x x)} \tag{2.42}$$

$$\begin{aligned}\rho\nabla\cdot\vec{V} = &\left(\rho_0(z) + \rho_0(z)\,A\,R(z)e^{i(\omega t-k_x x)}\right) * \\ &\left(-ik_x A\,U(z)e^{i(\omega t-k_x x)} + A\,W'(z)e^{i(\omega t-k_x x)}\right)\end{aligned} \tag{2.43}$$

$$\begin{aligned}\rho\nabla\cdot\vec{V} = &\, e^{i(\omega t-k_x x)}\left(-ik_x\rho_0(z)A\,U(z) + \rho_0(z)A\,W'(z)\right) \\ &+ e^{2i(\omega t-k_x x)}\left(-ik_x\rho_0(z)A^2\,R(z)\,U(z) + \rho_0(z)A^2\,R(z)\,W'(z)\right)\end{aligned} \tag{2.44}$$

Canceling the nonlinear terms with A^2 in 2.44 and adding it to 2.40, we have our first order perturbation continuity equation, which, unlike the zeroth order equation, is not equal to zero on the left hand side:

$$\begin{aligned}\frac{D\rho}{Dt} + \nabla\cdot\boldsymbol{V} = &\, A\,e^{i(\omega t-k_x x)}\left[i\omega\rho_0(z)\,R(z) + \rho_0'(z)W(z)\right] \\ &+ A\,e^{i(\omega t-k_x x)}\left[-ik_x\rho_0(z)\,U(z) + \rho_0(z)W'(z)\right] = 0\end{aligned} \tag{2.45}$$

$$\therefore\, i\omega\rho_0(z)\,R(z) + \rho_0'(z)\,W(z) + \rho_0(z)\left(W'(z) - ik_x\right) = 0 \tag{2.46}$$

where ω is the perturbation frequency, ρ_0 is the unperturbed density, R is the coefficient of the perturbed density term, W is the coefficient of the vertical perturbation velocity, and k_x is a horizontal wavenumber, while a prime indicates a partial derivative with respect to z.

Similarly, we can find a first order perturbation equation for the momentum, energy, and state equations. These derivations are not shown in full due to length. The equations below can be found in Midgley and Liemohn (1966).

$$\rho_0(z)\,i\omega\,U(z) = ik_x p_o(z)\,P(z) \tag{2.47}$$

$$\rho_0(z)\,i\omega\,W(z) = -\rho_0(z)g\,R(z) - p_0'(z)\,P(z) - p_0(z)\,P'(z) \tag{2.48}$$

$$\begin{aligned}\frac{\rho_0(z)\,k_b}{(\gamma-1)m}\left(i\omega\hat{T}(z) + T_0'(z)\,W(z)\right) = &\,\lambda\left(\hat{T}''(z) - k_x^2\hat{T}(z)\right) + \lambda'\hat{T}'(z) \\ &- p_0(z)\left(W'(z) - ik_x\,U(z)\right)\end{aligned} \tag{2.49}$$

where $\rho_0(z)$, $p_0(z)$, $T_0(z)$ are the unperturbed background density, pressure, and temperature respectively, ω is the perturbation frequency, k_x the perturbation horizontal wavenumber, k_b is the Boltzmann constant, γ the ratio of specific heats for the gas, λ

the thermal conductivity of the gas, W(z) and U(z) are the vertical and horizontal perturbation velocity coefficients, P(z) is the coefficient of the perturbation pressure, R(z) the coefficient of perturbation density, and $\hat{T}(z)$ the coefficient of the perturbation temperature. A prime indicates a partial derivative with respect to z and two primes indicates a second partial derivative with respect to z.

In order to find the useful dispersion equation derived in Hines (1960), further assumptions are made. As discussed previously, the region of interest in the atmosphere is nearly isothermal and we assume the wave is not dissipating. The coefficients of the perturbed variables (e.g. R, P, $\hat{T}$, U, W) in the above equations then become constants and we can find solutions of the form $e^{-i\kappa z}$ where κ is a complex vertical wavenumber if and only if the determinant of the coefficients vanishes:

$$\omega^4 - \omega^2 \left(C_s^2 k_x^2 + C_s^2 \kappa^2 - i\gamma g \kappa \right) + (\gamma - 1) g^2 k_x^2 = 0 \tag{2.50}$$

which is the dispersion relation in Hines (1960), with C_s being the speed of sound

$$\sqrt{\frac{\gamma p_0}{\rho_0}} = \sqrt{\gamma g H}$$

where γ is the ratio of specific heats C_p/C_v for the constituent gas. We can see that there are two branches of the solution by solving explicitly for κ, the complex vertical wavenumber.

$$\kappa = \frac{i\omega_a}{C_s} \pm \left[\frac{\omega^2 - \omega_a^2}{C_s^2} + k_x^2 \, \frac{\omega_{bv}^2 - \omega^2}{\omega^2} \right]^{1/2} \tag{2.51}$$

$$\omega_a = \gamma g / 2C_s \tag{2.52}$$

$$\omega_{bv} = \frac{(\gamma - 1)^{1/2} \, g}{C_s} \tag{2.53}$$

ω_a here is known as the *acoustic cutoff frequency*, by definition the lowest possible frequency of a sound wave that will produce a real wavenumber, while ω_{bv} is the natural adiabatic frequency, or the previously found BV frequency for an isothermal atmosphere in terms of sound speed and specific heats rather than density (Hines, 1960). To find a real wavenumber, we can set $\kappa = \frac{i\omega_a}{C_s} + k_z$, where k_z is the *real part* of the vertical wavenumber. The first imaginary term can also be written as $i/2H$ from the definition of ω_a and C_s and thus describes the amplitude growth of the perturbation as a function of scale height (Midgley and Liemohn, 1966). If k_z is imaginary, i.e. $k_z = ib$ where b is some constant, the wave solution will have the form

$$\sim e^{-i(\omega t + \kappa z - k_x x)} = e^{-i(iz/2H + \omega t - k_x x + ibz)} = e^{z/2H} e^{bz} e^{-i(\omega t - k_x x)} \tag{2.54}$$

and so be *evanescent* with only horizontal propagation. A real k_z will produce a horizontally and vertically propagating wave of the form

$$A\, e^{-i(iz/2H+\omega t-k_x x+k_z z)} = A\, e^{z/2H} e^{-i(\omega t-k_x x+k_z z)} \tag{2.55}$$

There are thus two possible cases for real k_z in equation 2.51, either $\omega > \omega_a$ or $\omega < \omega_{bv}$. If the wave frequency is greater than the acoustic cutoff frequency, then the wave is referred to as an *acoustic wave*, which a wave frequency lower than the BV frequency can be referred to as a *gravity wave* (often called internal gravity waves in older terminology, although it is largely outdated) (Hines, 1960; Midgley and Liemohn, 1966; Tolstoy, 1963). It is important to note here that in the technical sense of having gravity as a restoring force, an acoustic wave is also a gravity wave, but at acoustic frequencies and is therefore distinguished from the lower frequency propagating/internal gravity waves, which we will refer to as gravity waves for simplicity.

The ability of a gravity wave to propagate is also determined by its horizontal phase speed, $c = \omega/k_x$. For k_z to be real, $c > C_s$ in the high frequency limit, while $c < C_s$ in the low frequency limit. As a result, acoustic waves propagate much more quickly than gravity waves, above the local speed of sound. Therefore there are five possible categories of atmospheric wave governed by gravity:

1. $\omega > \omega_a$ and $c > C_s \implies k_z$ real. *Acoustic wave*
2. $\omega > \omega_a$ and $c < C_s \implies k_z$ imaginary. *Evanescent wave*
3. $\omega_{bv} < \omega < \omega_a \implies k_z$ imaginary. *Evanescent wave*
4. $\omega < \omega_{bv}$ and $c > C_s \implies k_z$ imaginary. *Evanescent wave*
5. $\omega < \omega_{bv}$ and $c < C_s \implies k_z$ real. *(Propagating/internal) Gravity wave*

This dispersion relation is used in nearly all gravity wave analysis papers, including Hickey, Walterscheid, and Schubert (2011), Terada et al. (2017), and Walterscheid, Hickey, and Schubert (2013), among others. England et al. (2017) uses a dispersion relation that includes the Coriolis frequency $f = 2\Omega \sin\phi$, ϕ latitude, but $\Omega_{Mars} \approx 1.13 \times 10^{-5}$ Hz, a full order of magnitude smaller than a typical BV frequency such as those found in Slipski et al. (2018), so ignoring Mars rotation is a reasonable assumption.

2.2.2 Wave dissipation

It is important to have at least a cursory understanding of wave dissipation, as the induction of turbulence due to wave dissipation or saturation is the primary method of energy transfer due to gravity waves (Fritts and Dunkerton, 1984; Fritts, 1984; Charney and Drazin, 1961; Fritts and Alexander, 2003; Geller, Tanaka, and Fritts, 1975; Hodges, 1967). However, because turbulence is an inherently non-linear process, it is difficult to describe analytically and must often be solved numerically unless the equations are linearized, as we have done above. There are a few ways that an atmospheric gravity wave can be dissipated; firstly, it can dissipate via viscous force, whether through molecular viscosity, eddy diffusion, or thermal conduction. Secondly, it can undergo "breaking", defined as the point where amplitude growth becomes unsustainable and the wave breaks down into turbulence, with energy then dissipating through convective instabilities. This is also called wave "saturation" (Fritts, 1984). Because this work focuses on perturbations in the exosphere, densities are sufficiently low for viscous forces to be nearly non-existent, as few particles undergo collisions. So the waves shown in this work most likely dissipate due to convective instabilities, i.e. saturation.

At altitudes below the exobase, the wave amplitude growth with altitude as shown above is tempered by the eddy diffusion coefficient (Hodges, 1969). Because waves with longer vertical wavelengths have corresponding higher group velocities, longer vertical wavelengths (on the order of 20km) require correspondingly higher eddy diffusion coefficients to prevent exponential amplitude growth (Fritts, 1984). In the exosphere, where eddy diffusion is non-existent, there is then no balancing force and the amplitude will grow unchecked. This creates a greater risk of saturation due to convective instabilities, which Terada et al. (2017) concluded was responsible for the thermospheric gravity wave amplitude dependence on background temperature below the exobase. This saturation then induces a drag on the atmosphere, which in term serves to transfer energy from the wave to the surrounding atmosphere. Further explanation of wave breaking and instabilities can be found in Appendix A. While eddy diffusion can be ignored in the exosphere, convective instabilities are dependent on temperature rather than viscosity and so can still occur where there are few collisions.

2.3 The DSMC model

To better understand the behavior of the exospheric perturbations, we simulate the upper atmosphere using a Direct Simulation Monte Carlo (DSMC) model developed by Ludivine Leclercq (Leclercq et al., submitted, see Chapter 5). The DSMC method allows for the study of a rarefied gas outside of the continuum regime assumed by equations 2.23, 2.24, and 2.25. For a low density gas, we must have a molecular-kinetic model to simulate the behavior of the particles. The DSMC method, developed by G.A. Bird, is one such method (Bird, 2013; Bird, 1994).

To determine if the DSMC method is appropriate for the desired simulated atmosphere, we can examine the dimensionless Knudsen number of the flow, defined as $Kn \equiv mfp/L$, where L is the characteristic length scale of the system. The Knudsen number determines if a flow is almost collisionless or totally collisionless. If $Kn \gtrsim 0.1$, indicating the flow is nearly collisionless, a continuum description is no longer valid and molecular motion is instead governed by the Boltzmann equation and kinetic theory. If $Kn \gtrsim 10$, on the other hand, the molecules in the gas essentially travel in ballistic trajectories. For Kn between 0.1 and 10, however, the flow is considered to be transitional, where a continuum model is no longer accurate but since the number of particles is too large to track every molecule individually, representative particles are tracked. Therefore the DSMC method was developed to allow for statistical molecular kinetics, reducing the amount of computing time while still accurately representing the physics of the rarefied gas. This method is then the most useful for our purposes, as we study the transitional region surrounding the exobase, where the flow goes from a continuum region to almost collisionless. Therefore a fluid model would be inaccurate, whereas a full molecular dynamics model would be too costly to run.

Instead of using a simulated particle with an exact potential for each real molecule, the DSMC uses simulated particles that statistically represent w_s real molecules, reducing the number of particles the model must track at any given time. The model then only considers binary, elastic collisions due to the relative rarity of collisions. The collision is treated as an instantaneous change in the particle velocity, as if the particles are hard spheres. The model tracks the position and velocity of each particle, possible due to the lower number of simulated particles. The velocities and coordinates of each particle are random and modified with time as they undergo motion due to gravity and collisions.

To simulate the rarefied gas using the statistically weighted particles and collisions, the DSMC model goes through several steps. First, the simulation domain must be set up and populated. This is done by choosing a top and bottom boundary location, then dividing the domain into cells. The height of the cells is usually determined to be slightly smaller than the mean free path of the particles, so that particles are able to cross cell boundaries. As our model is one dimensional, the horizontal dimensions of each cell are one unit length, with the vertical dimension chosen to be close to the molecular mean free path. In our model, this height is approximately 6-7km, depending on the cell. Each cell is then given a number of weighted particles with random coordinates and velocities chosen from a Maxwell-Boltzmann distribution for a given temperature.

After the initial conditions are set, the model undergoes a series of processes for each time step. The time step is usually chosen to be less than the mean collision time for the particles, so that particles do not travel non-physical distances before undergoing collisions. Thus, for each time step, the model:

1. Moves particles according to gravity and records the new position. Any boundary effects are also considered.

2. Re-indexes the particles into cells based on their new position

3. Selects collision pairs within each cell and assigns new velocities to the particles that undergo a collision
4. Samples properties such as density and temperature
5. Outputs the sampling results

This continues until the desired number of time steps has been reached.

For the first step, the particles undergo motion due to gravity, with first a new velocity v_{new} calculated, then a new position x_{new} using the kinematic equations of ballistic motion:

$$v_{new} = v_{old} + g_{old}\,dt \tag{2.56}$$

$$x_{new} = x_{old} + v_{old}\,dt + \tfrac{1}{2}\,g_{old}\,dt^2 \tag{2.57}$$

The gravitational acceleration g_{old} is also calculated at each time step depending on the altitude of the particle. After the new positions and velocities for each particle are found, the particles are re-indexed according to cell. This is necessary for the next step, as collision pairs are only chosen within a cell, so that particles are not crossing cell boundaries when colliding. At this stage any boundary conditions are also applied to the particles depending on their new positions. At the top of the simulation domain, a particle that crosses the upper boundary with energy greater than the escape energy is counted as an escaped particle, while a particle with energy less than the escape energy is treated as a ballistic particle, with its trajectory tracked until it reenters the simulation domain and can subsequently undergo collisions. At the lower boundary, the flux of particles entering the simulation is chosen to maintain the density of the lowest cell at roughly consistent levels. In our model, this upward flux is chosen to be $\Phi_0 = n_0\langle v_s\rangle/4$, where n_0 is the number density and $\langle v_s\rangle = \sqrt{\frac{8k_BT_s}{m_s\pi}}$ is the average thermal velocity with T_s is the species temperature and m_s is the species mass. The time step for our model is approximately 0.5s.

To calculate collisions, we must first find the number of collision pairs in a cell, n_{coll}. This number is proportional to the number of particles in the cell, the collision cross section, and the volume of the cell:

$$n_{coll} = \tfrac{1}{2}\,n_p(n_p-1)\,F_N\,(\sigma|v_i - v_j|)_{max}\,\Delta t/V_{cell} \tag{2.58}$$

Here n_p indicates the number of particles in the cell, F_N is a distribution function, $(\sigma|v_i - v_j|)_{max}$ is the maximum possible product of the collision cross section and the relative velocities of the ith and jth particles, Δt is the time step, and V_{cell} is the volume of the cell, which in our 1D case is equivalent to the difference between the top altitude and bottom altitude of the cell (Bird, 2013; Bird, 1994). Then random pairs of particles in the cell are chosen and undergo an acceptance-rejection procedure based on the probability of their collision relative to the maximum collision probability:

$$\frac{P}{P_{max}} = \frac{\sigma|v_i - v_j|}{(\sigma|v_i - v_j|)_{max}} \tag{2.59}$$

where P is the collision probability, P_{max} the maximum possible collision probability, and $\sigma|v_i - v_j|$ is the product of the collision cross section and the relative velocity of the ith and jth particles. $(\sigma|v_i - v_j|)_{max}$ indicates the maximum such product.

When two particles collide, their velocities are chosen to conserve momentum, energy, and angular momentum, with the new velocity vector chosen such that

$$v_1' = v_1 + [(v_2 - v_1) \cdot \hat{n}]\,\hat{n} \tag{2.60}$$

$$v_2' = v_2 - [(v_2 - v_1) \cdot \hat{n}]\,\hat{n} \tag{2.61}$$

where v_1' and v_2' are the new velocities of the first and second particle, v_1 and v_2 are the original velocities, and $\hat{n}$ is a random isotropic vector:

$$\begin{array}{lll} n_x = \cos\theta & n_y = \sin\theta\,\cos(2\pi r_1) & n_z = \sin\theta\,\sin(2\pi r_1) \\ \cos\theta = 1 - 2r_2 & & \sin\theta = \sqrt{1-\cos^2\theta} \end{array} \tag{2.62}$$

and $r_1,\ r_2$ are random numbers from 0 to 1. This creates the probabilistic collisions of the DSMC that statistically represent the outcomes of particle collisions without requiring a full atomic potential.

For our model, however, we use the Lewkow and Kharchenko (2014) cross section and collision method. This method is anisotropic, accounting for the angular part of a cross section, which is important for calculations of atmospheric escape. Using these cross sections also accounts for the energy-dependence of the differential cross section, as well as collisions between multiple species. Lewkow and Kharchenko (2014) calculates a variety of atom-atom differential and total cross sections for elastic collisions as a function of energy, since inelastic collisions are rare for the energy range of interest. These cross sections are then scaled such that differential cross sections for various species lie on a single curve, which is linear for small collision energies and scattering angles and quadratic for large energies and scattering angles. In our model, we use this curve to define the cross section for a given collision pair depending on their energies and species. For analysis of the NGIMS data, the constant cross section is sufficient as we do not have particle energies, unlike in the DSMC model.

After particles have undergone collisions, the density and temperature are sampled in each cell. For each cell, the density and temperature are calculated using

$$n_i = \frac{n_i w_s}{V_i} \tag{2.63}$$

$$T_i = \frac{m}{3k_B}\left(\frac{\sum_{p=1}^{n_i} v_p^2 w_s}{n_i V_i} - \langle v_p(i)\rangle^2\right) \tag{2.64}$$

where n_i is the number of particles in a given cell, w_s is the aforementioned statistical weight, V_i is the volume of the cell, m is the molecular mass, v_p is the velocity of particle p, and $\langle v_p(i)\rangle$ is the average particle velocity in the cell, separate from the average thermal velocity used above when choosing the upward flux at the bottom of the simulation domain. The model then outputs these results, which are subsequently used for data analysis such as that seen in Chapter 5. If the model contains more than one species, such as CO_2 and O, each species is given a statistical weight in proportion to their density ratio. This accounts for their different masses and densities and is the method used in Chapter 5.

Chapter 3

Examining NGIMS Neutral Data Response to Solar Wind Drivers

Paper submitted to the Journal of Geophysical Research: Planets in January 2019 as *Examining MAVEN NGIMS Neutral Data Response to Solar Wind Drivers*, authors Hayley N. Williamson (UVA), Meredith K. Elrod (NASA GSFC, University of Maryland College Park), Shannon M. Curry (UC Berkeley Space Sciences Laboratory), and Robert E. Johnson (UVA).

3.1 Abstract

The Martian upper atmosphere is known to vary diurnally and seasonally due to changing amounts of solar radiation. However, in the upper thermosphere and exosphere, the neutrals are also subject to ion precipitation. This can increase the temperature in the region of precipitation, resulting in density changes that might be seen in *in situ* data. Therefore, we examine neutral density data from the Mars Atmosphere and Volatile EvolutioN (MAVEN) Neutral Gas and Ion Mass Spectrometer (NGIMS) in Mars-Solar-Electric (MSE) coordinates, where location is determined by the direction of the solar wind convective electric field, resulting in a hemispherical asymmetry in the ion precipitation. By examining densities in MSE coordinates we are able to look for a detectable effect in the region where ion precipitation is more likely. Using the NGIMS neutral data and Key Parameters *in situ* solar wind data from February 2015 to August 2017 we look for asymmetries by constructing average density maps in Mars-Solar-Orbital (MSO) and MSE coordinates near the exobase. The NGIMS densities for O, Ar, and CO_2 from 180-220 km altitude for each orbit are averaged and then binned by location in MSO coordinates and transformed to MSE coordinates. The resulting MSE map exhibits a small density increase in the southern hemisphere, where one would expect to see enhanced precipitation. Although suggestive, the change is not statistically significant, so that the effect of ion precipitation, thought to be an important driver in the evolution of Mars' atmosphere remains elusive.

3.2 Introduction

The Mars Atmosphere and Volatile EvolutioN (MAVEN) mission seeks to both understand the structure of Mars' current atmosphere and ascertain how much of this atmosphere has been lost to space over time, known as atmospheric escape (Jakosky et al., 2015). Solar photons and the solar wind, a stream of charged particles that flows from the Sun, can affect the atmosphere and drive escape in several ways. For example, the interplanetary magnetic field (IMF) carried by the solar wind can penetrate Mars' upper atmosphere and thereby set ions in the ionosphere in motion. These are often referred to as pick-up ions. Such ions can then be lost by flowing down the planet's magnetospheric tail (Curry et al., 2015). However, while many of the ions picked up by the IMF are accelerated away from the planet, some reenter the atmosphere in a process known as pick-up ion precipitation (Johnson and Luhmann, 1998). The precipitating ions can then collide with and transfer energy to the neutral atmosphere. Non-thermal collisions can result in sputtering, a splashing out effect on the neutrals which has been suggested to be responsible for much of early neutral atmospheric escape (Leblanc and Johnson, 2002). Although it is still debated whether it was an important or dominant process early on, MAVEN data clearly indicate that it is a very small part of the present escape rate and as such is difficult to directly detect even in the extensive MAVEN data base (Leblanc et al., 2015). The energy transfer from precipitating ion and neutral collisions can raise the temperature of the atmosphere, and, hence, the scale height, leading to an increase in density of species in the region of the exobase. Therefore, here we examine the changes in the average density in the exobase region as a proxy for the pick-up ion heating. Precipitation is an ongoing process even during quiet solar conditions (Leblanc et al., 2015; Hara et al., 2017), so in this paper we look at average densities at a given altitude range to see if there is at present an effect from precipitation.

The neutral densities in this paper were obtained by the Neutral Gas and Ion Mass Spectrometer (NGIMS) on MAVEN, a quadrupole mass spectrometer that measures *in situ* densities every orbit (Mahaffy et al., 2015b). NGIMS provides a unique opportunity to study upper atmospheric composition, as it has now measured densities between the nominal altitudes of approximately 150-350 km for over a full Martian year. This provides a large dataset for analysis of the neutral densities in the altitude regime where the atmosphere transitions from collisional to ballistic and neutral escape becomes more likely. Since ion precipitation is correlated with the direction of the solar wind convective electric field, it is more likely when the field is directed towards the planet (Brain et al., 2015; Fang et al., 2013; Hara et al., 2017). Thus examining neutral densities in a frame of reference dependent on the direction of said electric field, the Mars-Solar-Electric Field coordinates (MSE), can help in trying to understand how and if ions driven by the solar wind can contribute to neutral escape. We also use Mars-Solar-Orbital (MSO) coordinates, which are useful for both atmospheric data and data taken farther from the planet, such as in the upstream solar wind, unlike geodetic coordinates, as a control comparison.

In MSE coordinates, the 'southern' latitudes indicate the electric field is pointed towards the planet, while 'northern' latitudes indicate the electric field is pointed away from the planet. Therefore precipitation predominately occurs in the southern MSE, or -E hemisphere, whereas ions flow out in the northern MSE latitudes or +E hemisphere. Transforming the neutral data to this coordinate system allows us to examine the effect on the neutrals of the potential presence of ion precipitation. We do this by comparing measured densities where precipitation is occurring with densities where precipitating ions are probably not providing additional energy to the neutral atmosphere.

We calculate average densities for oxygen, argon, and carbon dioxide between 180 and 220 km, then map these averages in MSO coordinates. We then use the upstream solar wind proton flow velocity and IMF vectors to calculate the solar wind convective electric field direction and use this to find MSE coordinates. As such, we will presume that changes in density in this altitude range between the northern and southern hemispheres are most likely due to the transfer of energy from precipitating ions in the southern -E hemisphere. The absence of a clear effect would suggest that ion precipitation is not affecting the neutrals. We also look at the data for different solar longitudes, to compare with global circulation models of the neutral atmosphere in order to ascertain if there is a seasonal effect. While we do not see any clear difference between seasonal densities in the data and previous model results (e.g. Bougher et al., 2015), we do see a suggestion that average neutral densities are slightly higher in areas where precipitation likely occurs. Although the evidence is not statistically significant, it is suggestive and could become clearer with additional data.

3.3 Methods

This study uses the publicly available NGIMS level 2, version 7, revision 3 data from February 2015 to August 2017, slightly more than a full Martian year, for a total of 3828 orbits. NGIMS is a quadrupole mass spectrometer that measures *in situ* ion and neutral counts for a range of 2-150 amu with 1 amu resolution every orbit. Both ion and neutrals counts are measured by NGIMS in channels for mass-to-charge ratio in a 2.6 s cadence (Benna and Elrod, 2017; Mahaffy et al., 2015b). The counts from each mass channel are converted to abundances in particles per cubic centimeter vs. altitude, time, latitude, and longitude of the measurement in the instruments level 2 data files.

MSO coordinates refer to a Mars-fixed solar-pointing coordinate system, with the X vector pointing towards the Sun, the Y vector anti-parallel to the direction of the orbit, and the Z vector completing the orthogonal system (Vignes et al., 2000). In the MSE coordinate system, the X unit vector maintains the same direction, but the XZ plane is defined by the

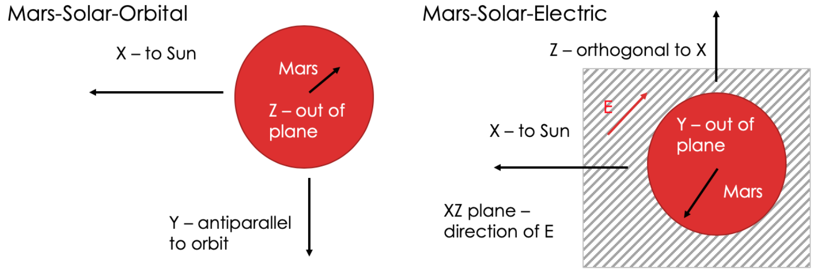


FIGURE 3.1: A cartoon depiction of MSO coordinates on the left and MSE coordinates on the right. In the right panel, the gray lines indicate the direction of the convective electric field.

direction of the positive solar electric field, with the Z unit vector chosen to be along the electric field direction and orthogonal to the X and Y vectors. A cartoon depiction of the two coordinate systems is shown in Figure 3.1. To define MSE coordinates, we use the Key Parameters (KP) version 12, revision 1 dataset, which includes data from the Particles and Fields instrument package onboard MAVEN (Dunn, 2015). The Solar Wind Ion Analyzer (SWIA) provides the proton flow velocity vector (Halekas et al., 2015), while the magnetometer provides the IMF vector (Connerney et al., 2015), giving us the background convection electric field from $\vec{E} = -\vec{v} \times \vec{B}$. To find these values, I average the velocities and magnetic field values above 4000 km in altitude for $|v| > 200$ km/s, a method used in Halekas et al. (2017) to determine if a measurement is taken in the solar wind. Due to the period of a MAVEN orbit, these measurements are thus taken approximately 2 hours prior to the NGIMS density measurements. However, the solar wind typically varies on the timescale of solar rotation, i.e. multiple days (Halekas et al., 2017) and so there should not be significant change in the solar wind electric field between the time of apoapsis and periapsis. Due to its precessing orbit, MAVEN does not always sample the solar wind directly when its apoapsis is on the night side of the planet, so we remove those orbits from our study by examining the proton flow velocities and magnetic field signatures in the KP data, as has been done previously (Halekas et al., 2017). Both proton flow velocities and magnetic field vectors are chosen to be the average values above 4000 km, to ensure that we are using solar wind values rather than those in the ionosphere.

To look at changes in average neutral density near the exobase for a variety of solar wind conditions, we compare densities for species O, Ar, and CO_2 in the 180-220 km altitude range. This is generally at or just above the region where the atmosphere transitions from being dominated by collisions to becoming ballistic. By looking at this altitude range, we can examine neutrals that might be heated by incident particle flux to sufficient temperatures to affect the escape rate from the planet's gravity well. Ar was chosen because it is chemically inert, meaning it does not undergo photochemical processes, while O and CO_2 were chosen due to their dominance at the altitudes of interest. For each orbit we average the measured densities for each species over altitudes 180-220 km. These averages are then separated into location bins of 5 degrees latitude and longitude in MSO and MSE coordinates, then the mean is found for each bin to produce the average density maps. This effectively normalizes the data by data density, reducing any bias due to multiple observations in the same location. Data is also split by solar longitude to examine seasonal neutral density variations outside of those expected from GCM models such as M-GITM, the Mars Global Ionosphere-Thermosphere Model (Bougher et al., 2014). While we do not directly compute temperature, neutral heating can be inferred from higher average densities, as an increase in temperature will increase the scale height and hence we will see higher densities at a given altitude.

3.4 Results

Figure 3.2 shows the binned densities as described in 3.3 in MSO coordinates. We observe, as expected, a distinct difference in density between dayside, the center of the figure, and nightside. For all three species, the density from subsolar point to antisolar point (or the data closest to those points) decreases by a significant amount. However, the highest density for O is not exactly in the noon region, as would be expected from atmospheric models in this altitude region (Bougher et al., 2014). For Ar and CO_2, the difference between nightside and dayside at these altitudes is quite clear and is close to two orders of magnitude. This, not surprisingly, is consistent with a significant difference in average temperature due primarily to the effect of UV heating. The density gradient across the terminators ($\pm 90^\circ$) is relatively steep, dropping approximately an order of magnitude for argon and carbon dioxide across 50° longitude. The density on the dayside and nightside is similar for the lighter species O due to ballistic transport and atmospheric winds. There are several places where the MSO paths cross, but do not have the same densities (for example, 0° longitude, -50° latitude). This is due to changes in solar longitude between the passes.

Figure 3.3 shows the same data as figure 3.2 but plotted in MSE coordinates. The MSE coordinates offer much higher spatial coverage than MSO, due to MAVEN retracing similar paths in MSO coordinates. While there are more filled bins, the number of orbits per bin on average is lower in figure 3.3. The MSE coordinates for a particular region in MSO coordinates change frequently due to the transient nature of the solar wind and the flapping of the IMF. The highly variable Z coordinate in MSE gives much more latitudinal coverage resulting in the near azimuthal symmetry. Because of the higher coverage, the general trends observed in MSO coordinates for Ar and CO_2 are easily seen: the dayside

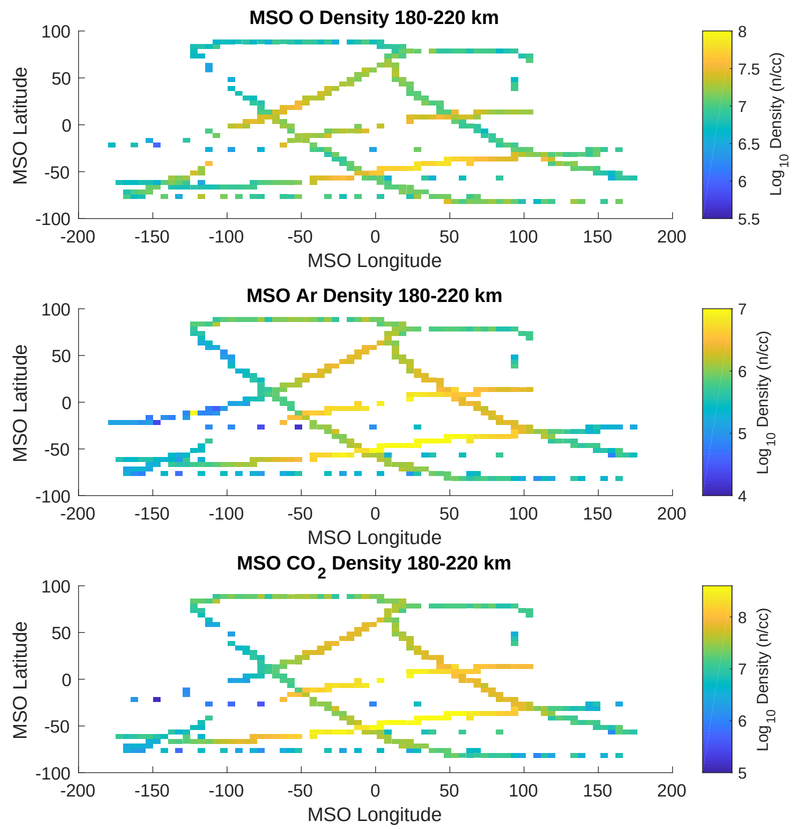


FIGURE 3.2: All available data averaged and binned in MSO coordinates with bin sizes of $5°x5°$ as described in the text: X pointing out from the plane towards the sun. Color indicates the log of the species density, with panels indicating density for O, Ar, and CO_2, respectively. Dayside, with noon at 0°, is uniformly higher density than the nightside (midnight at 180°), although the magnitude of the difference varies with species.

is higher density at this altitude than the nightside. However, any finer structure with respect to terminator or high latitude changes in density is lost because MSE coordinates are largely independent of geographic coordinates. The gap in the subsolar region is due to MAVEN not sampling the solar wind when its periapsis is at low solar zenith angles. At those times, its apoapsis is in the magnetotail or magnetosheath and so those data points are excluded, as the coordinate transformation will not be valid. The MSE coordinates give the densities an artificial oval or circular shape due to the rotation of the MSO coordinates with the solar wind electric field vector. So the circular spread represents rotation of similar locations in MSO for a variety of solar wind and IMF directions.

Previous papers have shown the presence of an ion polar plume at the north MSE pole (Dong et al., 2017) as a significant source of ion escape. Because ions escaping from below the exobase can also heat this region of the atmosphere, such a feature would appear as a density enhancement in the northern MSE hemisphere. Additionally, ion precipitation

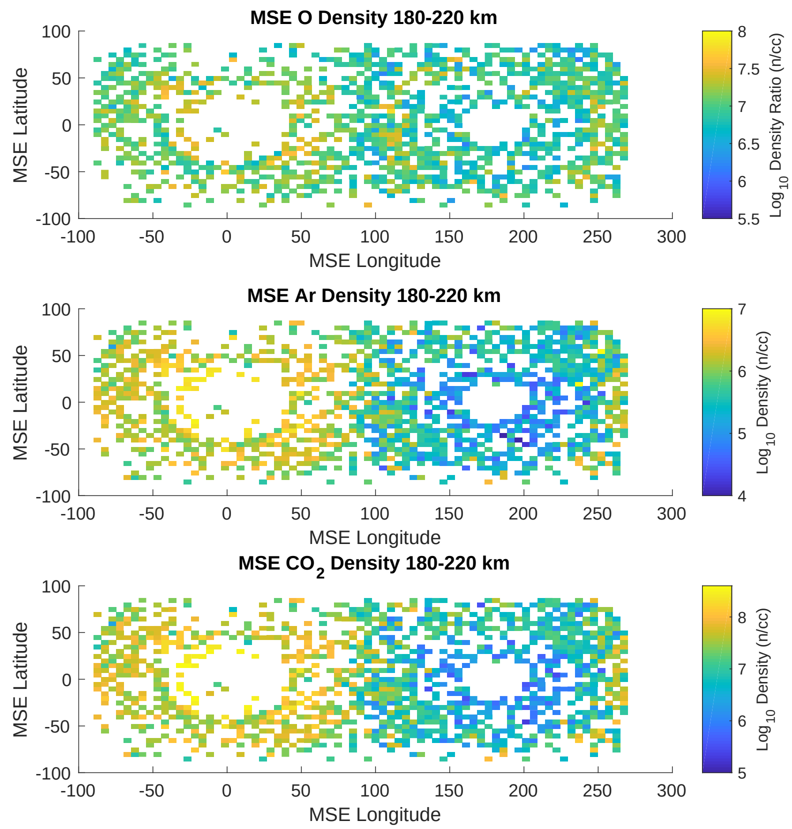


FIGURE 3.3: All available data binned and averaged with same resolution and color scale as in figure 3.2 but in MSE coordinates: X pointing out from the plane of each figure towards the sun and Z along projection of the solar wind electric field into the latitudinal plane, with Y perpendicular to the XZ plane and given as longitude. The solar wind electric field points away from the planet in the positive Z plane and towards the planet in the negative Z plane (positive and negative latitudes, respectively). Here, unlike in figure 3.2, dayside is on the left side of the plot and nightside is on the right sides.

into the atmosphere has been shown to be more common on the dayside hemisphere, specifically in the -E hemisphere (Brain et al., 2015; Hara et al., 2017), where the solar electric field points towards the planet. If this had a significant effect on the neutrals it would appear as an enhancement in the southern MSE hemisphere. Although the data density is not high, no such enhancements are obvious in figure 3.3. Although it is possible that this would be more evident with additional data, these data suggest the effect of the incident plasma either does not, on average, significantly heat the neutrals in this altitude region or that the effect is uniformly distributed. The effect of the ions might be clearer if it was possible to have useful solar wind data when MAVEN's periapsis is on the dayside, thus providing data in the subsolar MSE coordinates region. Currently, the near noon dayside coverage is poorer than that on the nightside.

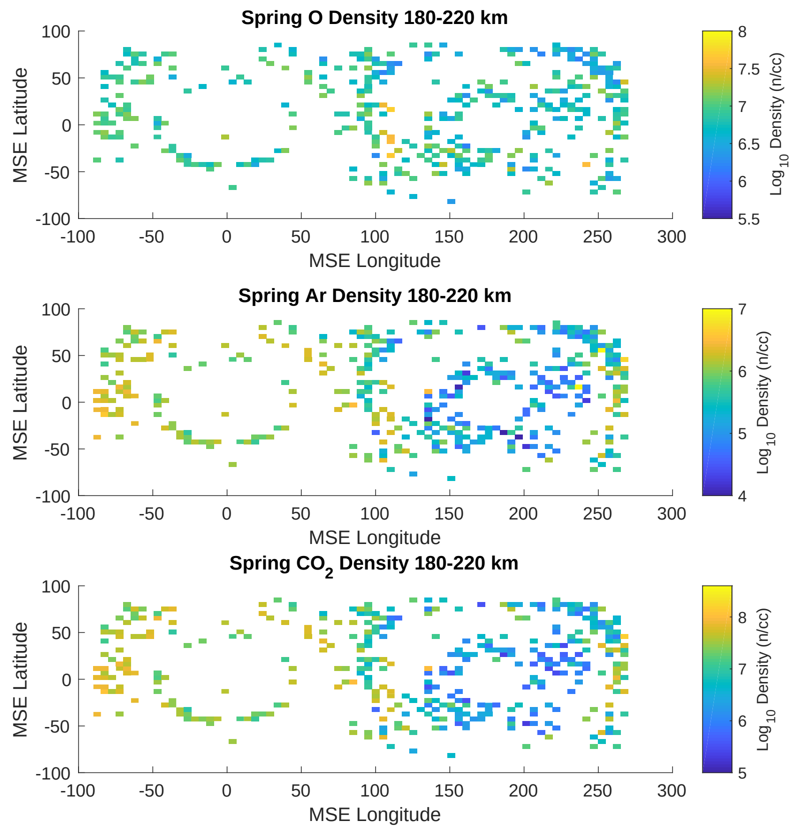


FIGURE 3.4: NGIMS data in MSE coordinates with a solar longitude less than 90°, i.e. northern spring, binned and averaged as described in the text. See figure 3.3 for a description of the figure.

To pursue this further, in figure 3.4 we bin the MSE density data for the Martian northern hemisphere spring and southern hemisphere fall, i.e. solar longitude less than 90°. Even though the data is even more sparse, for spring the coverage is best near the terminators and in the anti-solar region. We see the same general trends as in the entire dataset: dayside hemispheric densities are higher than the nightside densities for Ar and CO_2. However, there is a suggestion that O seems to be higher density near 100° longitude, which would be interesting as it matches M-GITM predictions quite well (e.g. Bougher et al., 2014). This suggests that longitudinal trends in the density at this altitude are visible for the different species in MSE coordinates and that the enhancement is not an artifact of the data processing. The same local enhancement is roughly visible in Ar and CO_2, which again would match global circulation models. However, any seasonal changes beyond those present in models for the neutral atmosphere near the exobase are not evident with the present data set.

Figure 3.5 shows the MSE binned data for northern summer and southern winter, which is also sparse. During this solar longitude, Mars is near its apoapsis. Again, overall

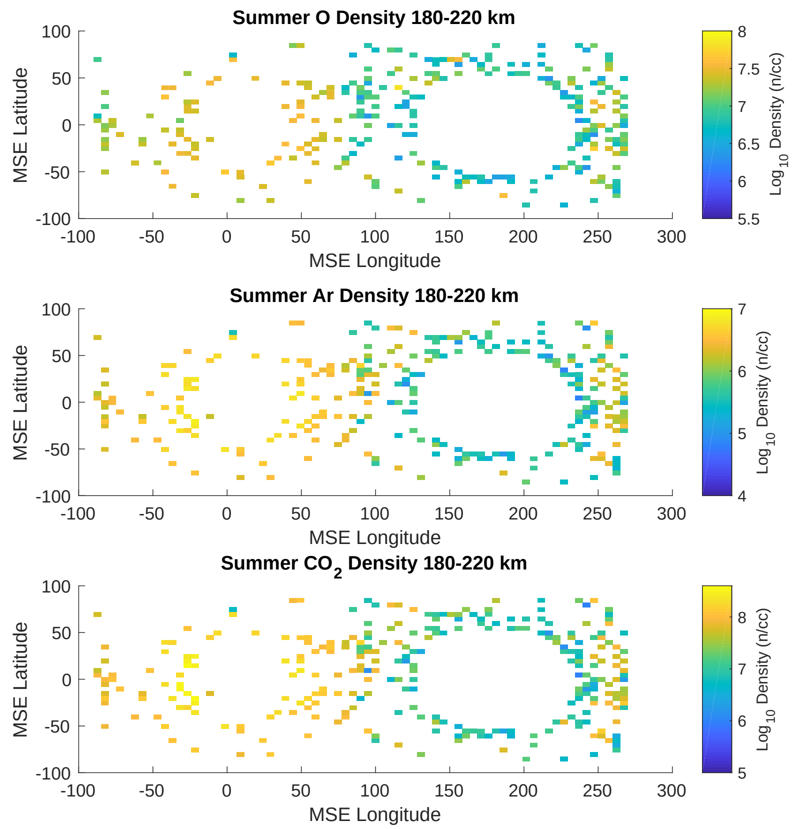


FIGURE 3.5: NGIMS data with a solar longitude between $90° \leq L_s \leq 180°$, i.e. northern summer, binned and averaged as described in the text. See figure 3.3 for a description of the figure.

longitudinal trends correspond roughly with those predicted by models. However, diurnal differences for O and CO_2 are higher than predicted in the M-GITM model (Bougher et al., 2014). Specifically, the dayside for O is about an order of magnitude higher in density than the nightside in this altitude regime, contrary to what is seen in Mars global circulation models at similar altitudes. It is unclear why this might be the case; since this seems to be in both the +E and -E hemispheres, it is unlikely the precipitation is warming the dayside enough to increase the density by an order of magnitude. Therefore, it may be due to the coordinate transformation changing what the normal seasonal trends would look like. Unsurprisingly, the diurnal change in density is highest for this season.

While data coverage is slightly different in figure 3.6 for the fall equinox, the overall density trends are similar to figure 3.4 at the spring equinox. The same high density region near 100° and 250° longitude for O is somewhat visible, although it is not as concentrated in a particular latitude region, indicating the variable solar wind conditions changing the MSE coordinates. CO_2 and Ar densities are again as expected by GCM models such as the previously mentioned M-GITM.

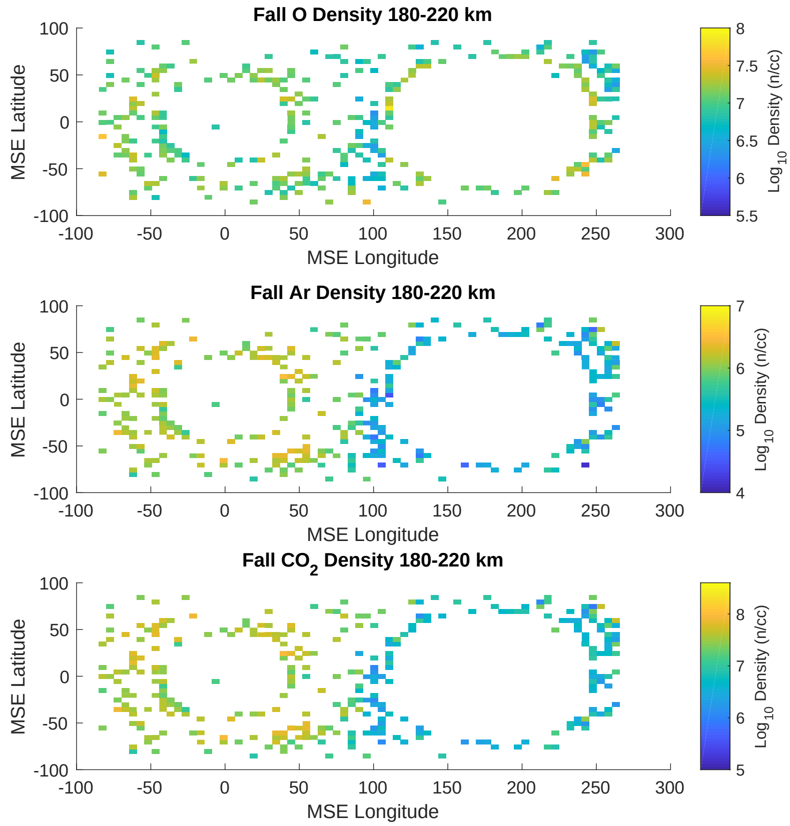


FIGURE 3.6: NGIMS data with a solar longitude between $180° \leq L_s \leq 270°$, i.e. northern fall, binned and averaged as described in the text. See figure 3.3 for a description of the figure.

Because the start of the MAVEN mission was in Martian winter and the dataset used in this paper is slightly over a full Martian year, figure 3.7 is the only one of the four that contains data from more than one Martian year. Figure 3.7 shows data from northern winter, which is also the season closest to solar periapse and is typically dusty, which can warm or "puff up" the atmosphere. This season differs most from that predicted by GCM models, which predict that for O the nightside should be an order of magnitude higher density than the dayside at these altitudes, which we do not see here. Instead, both hemispheres are nearly equal in density. Likewise, Ar and CO_2 are lower density on the nightside than would be expected. As mentioned above, the atmosphere close to periapsis can be quite dusty, which can change atmospheric temperatures and densities. GCM models typically use a dust average, usually for a weakly dusty season (Bougher et al., 2014). Thus is it likely that the difference between the data and model can be accounted for by dust, for while there has not been a global dust storm in some years, there are generally multiple large dust storms during this season every Martian year.

In table 3.1, we have split the MSE density data into dayside positive and negative E

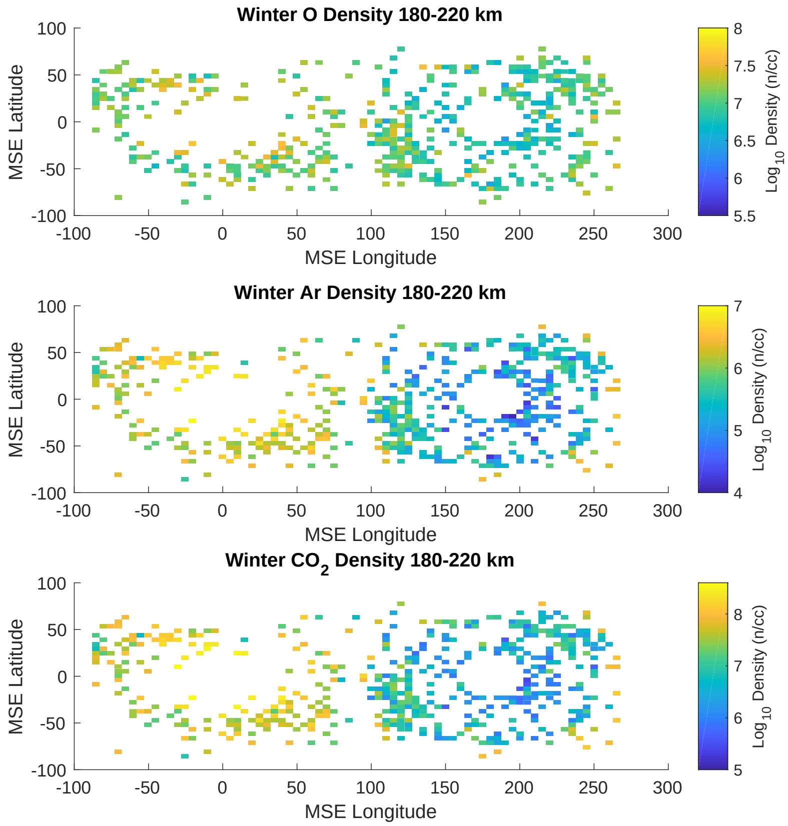


FIGURE 3.7: NGIMS data with a solar longitude greater than 270°, i.e. northern winter, binned and averaged as described in the text. See figure 3.3 for a description of the figure.

field direction (i.e. latitude) and nightside positive and negative E field direction, then taken the mean for each group. This shows that for all three species, regardless of solar zenith angle, the -E hemisphere is slightly higher in density than the +E hemisphere. Therefore it is possible that ion precipitation is adding energy to the neutral atmosphere in the -E hemisphere, hence raising the temperature and expanding the atmosphere. However, the standard deviation for these averages is large, up to 50 percent of the average density value, due to the data being from an entire hemisphere and thus varying widely in density. As a result, the differences between the ±E hemispheres are not individually statistically significant, but are consistent for all three species, as is the magnitude of the change, suggesting that precipitation might be affecting the neutrals on a global scale.

3.5 Discussion

Both diurnal and seasonal local time variations in density near the exobase are easily visible in the MSE coordinates and show the expected changes. Seasonal variations are

TABLE 3.1: Average densities for ±E day and night hemispheres

Hemisphere	O Density	Ar Density	CO_2 Density
+E Day	1.2	0.14	4.5
-E Day	1.4	0.15	4.6
+E Night	0.94	0.07	1.9
-E Night	1.1	0.08	2.4

[a]*1.0e7/cm^3

more complex, since they also depend on geographic latitude, not just solar longitude. However, the data shown in MSE coordinates is consistent with those studies in which the data is examined in geodetic coordinates. The orbital distance of Mars from the Sun, of course, also plays a role as it is highly eccentric affecting the solar insolation and, hence, the atmospheric structure. Comparing figure 3.5 and figure 3.7 shows densities at solar apoapse are lower than those near periapse as expected (Bougher et al., 2014; Bougher et al., 2015).

By definition, MSE coordinates provide information about the effect of ion flow in various regions. Because models have demonstrated that ion precipitation can transfer energy and increase neutral temperatures (Fang et al., 2013; Michael and Johnson, 2005), which in turn causes an increased scale height, comparing densities at the same altitude could be a proxy for the upper atmosphere heating rate. This is reflected in the seasonal effect and higher average dayside densities at fixed altitude seen in figures 3.2 and 3.3 for Ar and CO_2 (Valeille et al., 2009; Mahaffy et al., 2015a; Lillis et al., 2015). Therefore, the slightly higher densities seen in the -E hemisphere in table 3.1 suggests it is on average, slightly warmer than the +E hemisphere for both day and night.

Consistent with other studies using MAVEN data, atmospheric heating by incident plasma ions is found to be a small effect during the period examined. Our analysis suggests that heating of the neutrals on entering (-E hemisphere) slightly dominates heating on exiting (+E hemisphere). Since the change in density from MSE north to south in the altitude region studied is within one standard deviation of the mean, it is not statistically significant. However, the observed few percent difference is consistent with recent models of ion precipitation and sputtering (Fang et al., 2013; Leblanc et al., 2017). Due to the expected size of the density change and the lack of coverage on the dayside, it is not obvious in the global density maps (e.g. figure 3.3). To further evaluate its importance will require data at times of high solar activity, as well as data at lower solar zenith angles, which is not possible without independent solar wind measurements made simultaneously with the density measurements. There is currently work being done to estimate solar wind properties based on proxies while MAVEN has its apoapsis on the night side, but these results were not available at the time of this work. Although there is a suggestion of ion precipitation affecting the neutral densities, the role of pick-up ion precipitation, thought to be critical in the evolution of the Martian atmosphere, remains elusive.

Chapter 4

Large Amplitude Perturbations in the Martian Exosphere

Paper submitted to Icarus in January 2019 as *Large amplitude perturbations in the Martian exosphere seen in MAVEN NGIMS data*, authors Hayley N. Williamson (UVA), Robert E. Johnson (UVA), Ludivine Leclercq (UVA), and Meredith K. Elrod (NASA GSFC, University of Maryland College Park)

4.1 Abstract

We examine 252 Mars Atmosphere and Volatile EvolutioN (MAVEN) passes through the Martian atmosphere for which the Neutral Gas and Ion Mass Spectrometer (NGIMS) altitude-density profiles show perturbations with amplitudes larger than 40 percent of the background density that persist above the nominal exobase as defined here. The density profiles exhibiting such perturbations are plotted as a function of atmospheric column density rather than altitude. This roughly removes the dependence of the altitude-density profiles on composition, scale height, and local solar time. Such density structures are of interest as they can affect the local heating rate and, possibly, the escape rate. We find that the observed structures dissipate at roughly the same atomic column density and they affect the composition of the exosphere, raising the O/CO_2 ratio as compared with orbits that do not exhibit significant perturbations.

4.2 Introduction

Atmospheric waves with gravity as the restoring force, known generically as gravity waves, are pervasive and prevalent in Mars' atmosphere. *In situ* detections were made during the aerobraking phases of Mars Global Surveyor (MGS), Mars Odyssey, and Mars Reconnaissance Orbiter (MRO), as well during the Mars Atmosphere and Volatile EvolutioN (MAVEN) mission using the Neutral Gas and Ion Mass Spectrometer (NGIMS) (Mahaffy et al., 2015b). The NGIMS instrument measures the density and composition of neutrals as the spacecraft passes through the Martian atmosphere (Mahaffy et al., 2015a). Earth atmospheric studies show such perturbations are an important mechanism for the vertical transport of energy from the middle atmosphere to the upper atmosphere (Vincent, 2009). Linear fluid dynamics theory predicts that the amplitude can grow nearly exponentially with altitude and that once growth reaches a certain point, the amplitude begins to saturate, defined as the process by which the wave activity is dissipated by turbulence (Fritts, 1984; Fritts and Alexander, 2003). Multiple studies of wave activity at Mars based on the MAVEN *in situ* data have shown that the majority of these features saturate in the upper thermosphere, at densities where the atmosphere transitions from being collisional to ballistic. However, large wave-like features, or density perturbations, are seen here above the transition region and into the exosphere. This could be due to gravity waves with large amplitudes or long wavelengths penetrating beyond the nominal exobase into the exosphere (England et al., 2017; Terada et al., 2017; Yiğit et al., 2015) or, possibly, perturbations due to solar or plasma heating events directly produced in this region (Thiemann et al., 2015; Terada et al., 2017).

Because these features occur at such high altitudes, they can enhance neutral escape via energy transfer (Walterscheid, Hickey, and Schubert, 2013). However, it remains to be seen how much energy high altitude waves contribute to the ballistic regime, as wave theory is predominately focused on modeling the behavior in the collisional region of an atmosphere. Here we examine wave-like features that penetrate into the exosphere as seen in the NGIMS data to ascertain how these features behave and dissipate in this nearly collisionless regime. To do so, we first give our criteria for the exospheric features to be studied as seen in the NGIMS data from February 2015 to February 2018. We then examine these features in terms of atomic column density by integrating the CO_2, Ar and O densities along the altitude of the MAVEN spacecraft. The atomic column density was chosen rather than altitude as it directly correlates with the nominal exobase as defined below. Unlike density as a function of altitude, it is not as sensitive to the local temperature, so the effect of scale height changes with solar zenith angle are mitigated. Using plots of CO_2, O, and Ar densities versus atomic column density, we discuss the mechanisms for dissipation in the exosphere and how this might affect the local heating.

4.3 Data

4.3.1 Exospheric perturbations

To identify large exospheric perturbations, we searched the NGIMS neutral density data from February 2015 to November 2017. We examined O, Ar, and CO_2 profiles because O and CO_2 are the dominant species in the upper thermosphere and exosphere, while Ar is chemically inert and so unaffected by photochemical processes. For each orbit, we selected only the inbound data due to potential issues with the O density on the outbound leg of the orbit (Benna and Elrod, 2017). As discussed in Chapter 1, reactive species such as O and CO_2 build up on the wall of the instrument throughout the orbit, creating an artificial increase in the density on the outbound leg. Thus using only the inbound leg allows for more accurate density measurements. For each set of inbound orbit density data, we fit a 5^{th} degree polynomial to the density-altitude curve to obtain a smooth background density profile, as we found that a 5^{th} order polynomial captured long-wavelength waves with sufficient accuracy. Using a polynomial fit to obtain a background density for the purposes of calculating amplitude has been used previously in papers such as Cui et al. (2014) and England et al. (2017); and Yiğit et al. (2015), although the degree of the polynomial varies. This method is preferred to an exponential fit as it better accounts for a non-constant scale height; with an exponential profile, the scale height must be reevaluated throughout the profile to maintain the quality of the fit. Thus the polynomial fit is both simpler and smoother.

The amplitudes of the perturbations were calculated from the relative difference between the actual density and the background profile: i.e., $\frac{n_{data}-n_0}{n_0}$, where n_0 is the density of the 5^{th} degree polynomial fit to the NGIMS data and n_{data} is the NGIMS number density. This number, multiplied by 100, gives the percent amplitude relative to the background. We searched the NGIMS data to find passes such that the measured density profile exhibits perturbations that fit the following criteria:

1. There must be a perturbation with an Ar or CO_2 peak amplitude greater than 40% of the background density, where amplitude percentage is defined by $\frac{n_{data}-n_0}{n_0} \cdot 100$. The O amplitude is significantly smaller than that of Ar or CO_2 so is not used for this criterion.

2. The profile exhibits only one peak with amplitude larger than 40% of the background density. Most profiles also have small amplitude perturbations with amplitudes below this limit, assumed to be of the same wave train.

3. The peak amplitude must occur above 5×10^{18} atoms/m^2, a rough estimate of the onset of the transition into the exosphere as discussed below. This ensures the features examined dissipate in the nearly ballistic regime.

Using these criteria, we find 252 examples out of the 4259 orbits from February 2015 to November 2017. These profiles are subsequently compared to an average of the profiles without perturbations. The perturbations found span all local solar times, with slightly larger amplitudes on the nightside, and have a median apparent Ar wavelength, defined as the altitude difference between the two largest amplitude peaks, of 29.6 ± 9.5 km. Additionally, they are spread relatively evenly throughout all solar longitudes, indicating they are a global, if rare, phenomenon not affected by season. Both the spatial distribution and mean wavelength are consistent with the thermospheric waves analyzed in Terada et al. (2017), which shows waves appearing throughout the atmosphere and average vertical wavelengths of 20-40 km. The orbits of interest are slightly more common in the terminator regions and the nightside, although it is not a strong preference. This spatial distribution indicates that these large exospheric perturbations are, while uncommon, not unique to any particular region in the Martian atmosphere, much like gravity waves at lower altitudes, indicating that this could be a subset of the larger population of atmospheric gravity waves.

Figure 4.1 shows one of 252 examples that fit the above criteria, the inbound portion of orbit 2521, from January 16, 2016. Panels (a) and (b) show the density and amplitude profiles as a function of altitude, including in panel (a) the 5$^{\text{th}}$ degree polynomial fit used as a smooth background for the purposes of calculating amplitude. However, because the exobase region is dependent on the column density rather than altitude, we also show the density and amplitude as a function of total atomic column density with a nominal exobase as defined in the text. The total column density is calculated using Newton's method of integration from high altitudes to low altitudes for each species: $N = \int n\, dz = \frac{n_i + n_{i-1}}{2} * (z_{i-1} - z_i)$. The column densities for each species are then summed in proportion to the number of atoms for each species, i.e., $N_{tot} = N_O + N_{Ar} + 3\, N_{CO_2}$ giving a net atomic column for each MAVEN trajectory. At periapsis, the trajectory of the spacecraft becomes nearly horizontal, rendering the integration inaccurate, so the column density values are truncated at approximately 5 km above periapsis, consistent with England et al. (2017). Thus there are slightly fewer data points in the density versus column density profiles at high column and number densities, evident when comparing panel (a) to panel (c) in figure 4.1. Because the perturbations of interest penetrate into the exosphere, this does not affect the quality of the data examined. Plotting density and amplitude versus column density allows us to use a common reference level which in this paper we define below as a nominal atomic exobase. Because the principal escaping heavy neutral is atomic O (Lillis et al., 2017), in the following we use as a reference level an estimate of the exobase, the location where the mean free path of an escaping O is of the order of the atomic scale height, H: i.e., $1/(n\sigma) \sim H$. Here σ is an average *collision cross section* between atoms in the atmosphere. Since the column density $N \sim nH$, we use the equation $N_{ex} \sim 1/\sigma$ to define the reference level using an estimate of the O + O cross section at these temperatures (i.e. a few hundred K), $\sim 2 \times 10^{-15}$ cm^2 (Johnson, Schnellenberger, and Wong, 2000; Kharchenko et al., 2000; Tully and Johnson, 2001). This differs somewhat from other estimates, such as in Jakosky et al. (2017), where the cross section for Ar + CO_2 $= 3 \times 10^{-15}$ cm^2 is used to estimate a nominal exobase. Here we use the oxygen cross section as an "atomic" cross section, as we are interested in a rough guide as to where collisions of atoms with any of the gas atoms or molecules become much less likely.

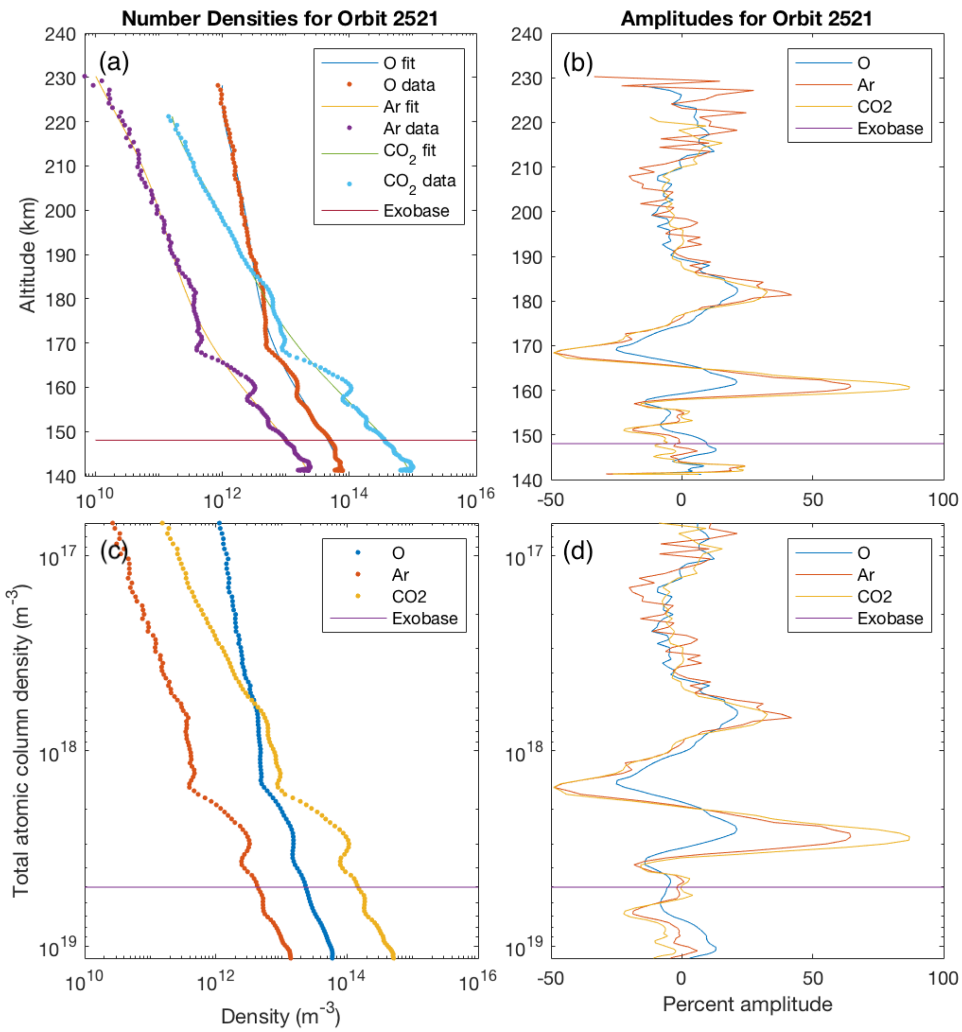


FIGURE 4.1: Example of a high altitude, large amplitude gravity wave from MAVEN orbit 2521 for species O, Ar, and CO_2. Panels (a) and (b) show the density plus background fit and amplitude plotted versus altitude, while panels (c) and (d) show the same plotted versus total atomic column density. The horizontal line is shown as a reference level of the order of our nominal atomic exobase, as defined in the text.

We use the O cross section here as the baseline rather than Ar because O is the primary escaping species of interest at these altitudes (Fox and Hać, 2009). Despite the slight difference in cross sections, the altitude range of the exobase is roughly the same as seen in Jakosky et al. (2017) and Slipski et al. (2018). As the transition region and exosphere are dominated by O and CO_2, treating CO_2 as 3 atoms, we use $N_{ex} \sim 5 \times 10^{18}$ atoms/m^2 as a nominal exobase (e.g. Jakosky et al. (2015), Jakosky et al. (2017), Lillis et al. (2017), and Slipski et al. (2018)). This line is marked on all plots and is particularly useful as we plot versus atomic column density.

In panels (a) and (b), the nominal exobase is located at approximately 150 km in altitude, which is on the lower end of the typical estimates of the exobase altitude, generally considered to be around 140-200 km (e.g., Jakosky et al. (2017) and Slipski et al. (2018)). However, orbit 2521 is located at 3 hours local solar time, on the nightside, where scale heights are depressed due to colder temperatures. As a result, the exobase is subsequently lower in altitude. However, the column density for the exobase remains the same as for all other orbits, independent of local time. Thus plotting density versus column density better organizes the data and clearly indicates when the perturbation is above the nominal exobase, which is not obvious in panel (a).

4.3.2 Average perturbation profiles

We show in Figure 4.2 an average profile of the absolute value of the amplitudes in terms of the total atomic column density as described above. Because the range of column densities in each orbit is slightly different, we standardize each amplitude profile by fitting it with a 15$^{\text{th}}$ degree polynomial between 10^{19} and 10^{16} m^{-2}, with the degree of the fit chosen to minimize residuals between the data and polynomial. We include for comparison a CO_2 mean amplitude profile constructed from all the orbits from February 2015 and February 2018 except the 252 case studies in order to show that the case studies examined here deviate significantly from the normal conditions.

The solid lines in Figure 4.2a show the result for each species when the mean is taken of all 252 absolute value amplitude fits, indicating that the maximum appears at a similar column density. In Figure 4.2b the solid line gives the averaged O/CO_2 ratio versus atomic column density for the orbits selected. As in Figure 4.1, the nominal exobase level is indicated by the horizontal line. These profiles are compared with the average of those orbits for which a perturbation fitting our criteria was not detected. The averaged profiles in Figure 4.2a show a growing amplitude until 2×10^{18} m^{-2}, after which it begins to decay. As can be seen in the sample orbit in Figure 4.1, the O amplitude is approximately half that of the Ar and CO_2 amplitudes due to differences in scale height. Gravity wave physics predicts that the amplitude is inversely proportional to scale height for long-wavelength waves even as the species amplitudes grow with altitude; therefore, lighter species should have smaller amplitudes, while species of similar masses such as Ar and CO_2 should have very similar amplitudes (England et al., 2017), which fits the data shown here. Additionally, we also see an increasing O/CO_2 ratio near 2×10^{18} m^{-2}, unlike the mean ratio for orbits without an exospheric perturbation. In Figure 4.2b, it is seen that the O/CO_2 ratio increases at similar column densities up to the point at which the mean amplitude begins to decrease. This suggests that the amplitude dissipation and the changing composition are related, as discussed in section 4.4.

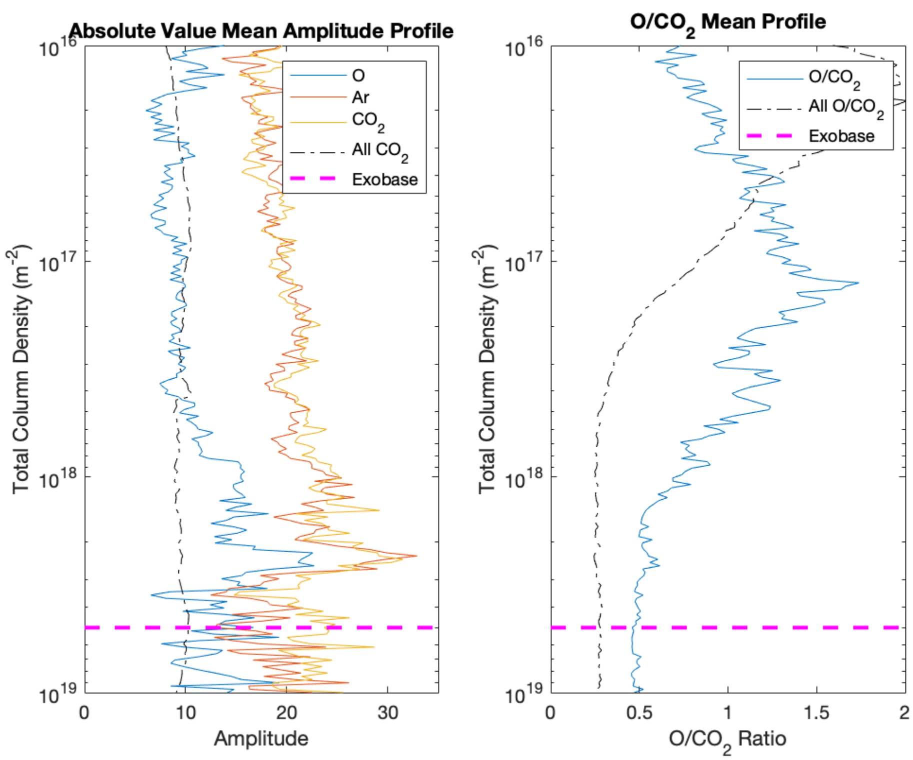


FIGURE 4.2: (a) Mean amplitude profile for exospheric waves of species O, Ar, and CO_2, chosen as described in the text. The black dashed line indicates the mean CO_2 amplitude for all orbits without a perturbation from Feb 2015-Nov 2017. (b) The mean ratio of O to CO_2 for all background density profiles plotted versus total atomic column density for the cases studied in blue and the mean ratio for all other orbits in black. Both panels are plotted versus total atomic column density in number/m^2. The reference level of nominal exobase as chosen in the text is marked with a horizontal dashed line.

4.4 Discussion

We examine the density structure in terms of atomic column density rather than altitude for several reasons. First and foremost is that it is much less sensitive than the density at a given altitude to conditions below the region being studied, as well as to the scale height and temperature. Since on the colder nightside of Mars, scale heights are smaller, a similar altitude to that on the dayside can have vastly different densities for each species. Therefore, averaging density profiles at a given altitude is not very useful in this region of the atmosphere. Additionally, the boundary between the predominantly collisional and the predominantly ballistic regions of an atmosphere can be characterized by column density for any chosen definition of the collisional exobase, as discussed in section 4.3. Therefore, a reference level, such as the nominal atomic exobase, is roughly common to all passes on a density versus column density profile. In this way, the density structure versus atomic column density is a significant help in interpreting the behavior of large perturbations in the exobase region of Mars' atmosphere.

Amplitude growth with decreasing column density is expected in continuum models of gravity waves (Fritts, 1984; Delisi and Orlanski, 1975). This, however, is likely limited by saturation for smaller amplitude gravity waves (England et al., 2017; Terada et al., 2017). England et al. (2017) examined average gravity wave amplitudes as a function of altitude for 116 orbits with results for CO_2 almost identical to the dashed line in Figure 4.2a, indicating that the roughly constant 10% amplitude for orbits without our defined exospheric perturbations is due to many smaller amplitude waves that saturate before their amplitudes grow significantly. However, we see in Figures 4.1 and 4.2 that amplitude growth persists even above our reference exobase, where the amplitude of the density perturbations grows with increasing altitude or decreasing column density. This is consistent with what would be expected for long wavelength gravity waves at high altitudes, even though the atmosphere is nearly collisionless. In our 1D simulations of the exobase region of a Mars-like atmosphere such growth can continue, resulting in relatively large amplitudes in the nearly collisionless regime (Leclercq et al., 2018). Here it is seen that in Mars' exosphere, as in the continuum regime, the wave amplitude growth does not continue indefinitely. The growth of the perturbations examined in Figure 4.2 begins to dissipate at an atomic column density of 2×10^{18} m^{-2}. Limits on wave amplitude growth in the continuum region, known as wave saturation, are due to instabilities and turbulence introduced into the atmosphere by the wave itself, dampening amplitude growth and often leading to total dissipation (Fritts, 1984). However, in the region of interest the continuum models do not apply, as the atmosphere is nearly ballistic. Below we consider the cause of the observed amplitude decay, as linear gravity wave theory is not applicable and, as usual, the temperature profile is not directly measured.

In Figure 4.2a, it is clear that the significant exospheric perturbations examined in this study appear at roughly consistent column densities as opposed to below our reference exobase, where there is significant variability, which appears like noise. This suggests that amplitude growth of these perturbations does vary somewhat depending on the conditions of the atmosphere at the time of being perturbed, including, for example, temperature. However, at lower atomic column densities (equivalent to higher altitudes), the noise in the mean amplitude profiles decreases and all three species experience a consistent decay in amplitude. The comparative lack of variability between passes in the region where the amplitude is decaying suggests that these exospheric perturbations begin to dissipate at a similar atomic column density, so the dissipation process is likely related to the atomic column density. Since these features are only present in a few percent of the total orbits, they are not a common feature. Indeed, they are not present in consecutive orbits, indicating they dissipate within the roughly 4 hours of MAVEN's orbital period and are not located in any particular geographic region. However, when the features are present, they are consistent in terms of peak amplitude and dissipation locations, unlike the other orbits with comparatively small average perturbations as seen in Figure 4.2a.

Wave activity has been shown to be an important mechanism for transferring energy from the middle atmosphere below the homopause, which varies from $\sim 60-130$ km in altitude (Slipski et al., 2018), to the upper atmosphere (the thermosphere and exosphere), increasing the velocity of molecules as induced turbulence (Charney and Drazin, 1961; Fritts and Alexander, 2003; Geller, Tanaka, and Fritts, 1975; Hodges, 1967; Midgley and Liemohn, 1966; Vincent, 2009; Walterscheid, Hickey, and Schubert, 2013). As the waves break in the collisional regime, their dissipation can lead to both heating and subsequent cooling (Hickey, Walterscheid, and Schubert, 2011). In the Martian middle atmosphere and lower thermosphere there can be cooling above the amplitude peak and heating below (Medvedev et al., 2015). It is possible here that the dissipation of the perturbations is also contributing energy to the atmosphere, increasing the diffusive separation of the different species and leading to the observed increase in the O/CO_2 ratio. The dissipation observed is likely due to the increased horizontal transport of O at these column densities as seen in various global climate models (Bougher et al., 2014).

In the transition region of the atmosphere up to the nominal exobase, it has typically been assumed that the dissipation of the wave energy produces local heating by increasing the atomic kinetic energy which in turn can affect the escape rate (Snowden et al., 2013; Walterscheid, Hickey, and Schubert, 2013). This appears to be roughly consistent with what is seen in Figure 4.2b. Since temperature is not measured, we note that the increase in kinetic energy and hence thermal velocity associated with these perturbations is the likely source for the increase in the O/CO_2 ratio at much lower column densities than is the case for the average atmospheric profile shown. This is seen to be the case well below the observed perturbation peak, where there is indication that perturbations are still reaching large amplitudes, seen as noise in the average profile. Even before O begins to dominate over CO_2, the ratio for the orbits of interest is closer to 0.5 (solid curve) at the reference level as compared to the $\sim$ 0.25 (dot-dash curve) average of the non-perturbed passes. In the absence of a direct measure of temperature data, we interpret this as evidence of enhanced energy transport into this region by the perturbations, resulting in an enhancement of diffusive separation and hence the O/CO_2 ratio for orbits with a significant exospheric perturbation. Such an effect is also seen in our molecular kinetic simulations for upwardly propagating perturbations in the exobase region of an $O + CO_2$ Mars-like atmosphere, where the kinetic temperature can be directly measured in the simulated atmosphere (Leclercq et al., 2018).

As mentioned previously, the averaged amplitudes in Figure 4.2a begin to decrease at a column density of $\sim 2 \times 10^{18}$ m^{-2}. This occurs at an altitude above the column density for the peak in the average production of escaping O due to photodissociation of O_2^+ (10^{-8}g/cm^2 $\sim 4 \times 10^{18}$atoms/m^2) (Lillis et al., 2017), which is of the order of our reference exobase. Escape can act to cool the atmosphere, limiting the growth and inhibiting the upward propagation of the perturbation energy. However, we also note that the O/CO_2 ratio in Figure 4.2b eventually begins to decrease with decreasing column density (increasing altitude) for both the perturbed and unperturbed profiles, although it is both more prominent and at higher column densities for the perturbed orbits. This is primarily due to ballistic transport of O (Deighan, J et al., 2015; Bougher et al., 2014), which, of course, includes any escaping O. In order to cause the observed local damping of the ratio, this transport must occur over length scales larger than the estimated horizontal wavelengths. The $\sim 2 \times 10^{18}$ m^{-2} column density suggests a mean free path between collisions $\sim$ 2.5 times that at the exobase, roughly consistent with the ratio of suggested wavelengths to the exobase scale height. As these large wave-like perturbations dissipate, the heat flux is directed downwards and horizontally to the unperturbed regions of the exosphere, resulting in cooling above peak amplitudes (Hickey, Walterscheid, and Schubert, 2011). This process is likely enhanced by the photochemical production of hot O at the atomic column densities of interest. Therefore, the composition and structure of this rarefied region of Mars' atmosphere is significantly modified by perturbation-induced heating from below and cooling by transport above the peak amplitude.

4.5 Conclusions

In conclusion, large density perturbations in the exobase region of Mars atmosphere are shown to be much better organized by atomic column density than altitude of occurrence. This roughly removes the dependence on scale height so that the amplitudes of these perturbations peak at roughly the same column density above a common reference level such as the nominal atomic exobase. In this way it is readily seen that the local heating by perturbations cause the O/CO_2 ratio to be enhanced in this region of the atmosphere. While atmospheric perturbations with large amplitudes can penetrate the exosphere, the amplitude growth with increasing altitude (decreasing atomic column density) is eventually limited by ballistic transport. This transport is enhanced due to heating produced by the upwardly propagating perturbation and acts to cool the atmosphere locally. The perturbations examined here are seen to die out above $\sim 10^{18}$atoms/m^2, although they affect composition up to much higher altitudes (smaller atomic column densities). These exospheric perturbations have scale height dependent amplitudes and amplitude growth with altitude consistent with upwardly propagating gravity waves. Although they can affect the escape rate, in the absence of direct measurement of temperature, molecular kinetic simulations are required. What is clear from the results presented is that the nature of the pulse and the composition indicate that the collision rate slowly decreases above the nominal exobase until energy flow is dominated by horizontal molecular transport and, possibly, escape. More details on the wave properties such as frequency and phase speed and their effect on Mars atmosphere will require further simulations and data analysis, currently in progress.

Chapter 5

Molecular Kinetic Simulations of Atmospheric Perturbations

Paper submitted to Icarus in March 2019 as *Molecular Kinetic Simulations of Transient Perturbations in a Planet's Upper Atmosphere*, authors Ludivine Leclercq (UVA), Hayley N. Williamson (UVA), Robert E Johnson (UVA), Orenthal J. Tucker (NASA GSFC), Lucia Tian (UVA), and Darci Snowden (Central University Washington). My contribution to this paper was in the form of providing both the theoretical basis and suggestion of how best to modify the model so that it represent a realistic wave, including modeling the perturbation as a multi-period pulse with a frequency approximating the atmospheric BV frequency. Additionally, I contributed the NGIMS data used as a comparison to the model result and provided information about the amplitudes of perturbations seen both above and below the exobase.

5.1 Introduction

The physics and chemistry of the exobase region of the upper atmospheres of exoplanets, solar system planets, planetary satellites, and Kuiper Belt Objects (KBOs) can determine their long-term evolution. The behavior of this region has been shown to be sensitive to the molecular composition and the transition from a collision dominated fluid-like regime to a nearly collisionless corona from which escape occurs. Based on MAVEN (Mars Atmosphere and Volatile Evolution) data at Mars and Cassini data at Titan, significant variations in the density structure with altitude have been observed in their upper atmospheres. Such perturbations are generally interpreted as gravity waves, which are certainly generated in these atmospheres (e.g., Snowden et al., 2013; Yiğit et al., 2015; England et al., 2017; Terada et al., 2017). The density data is typically analyzed using continuum fluid descriptions of the atmosphere which have been shown to fail well below a planet's exobase (e.g., Volkov et al., 2011; Volkov and Johnson, 2013; Tucker and Johnson, 2009; Tucker et al., 2012; Tucker et al., 2016; Johnson, Volkov, and Erwin, 2013a; Johnson, Volkov, and Erwin, 2013b). On the other hand, molecular kinetic simulations, which are numerical solutions to the Boltzmann equation, can be used to describe the transition from the collision dominated to the nearly collisionless regime giving the thermal structure of the upper atmosphere and the escape rate. Such simulations are especially important as the atmospheric temperature is typically not measured but is extracted from density vs. altitude data assuming local thermodynamic equilibrium. Since continuum models can fail even when the mean free between collisions is a very small fraction of the scale height (e.g., (Tucker et al., 2013; Tucker et al., 2016), the processes used for temperature extraction in the transition region need to be examined.

Molecular kinetic simulations have been used extensively to determine the *steady state* behavior of an atmosphere in which the relaxation time scales are short compared to day/night and seasonal time scales. In this paper, the Direct Simulation Monte Carlo (DSMC) method (Bird, 2013) is used to study *transient* events that propagate through the transition region and into the exosphere with emphasis on mass separation and on extraction of the local temperature which is directly calculated in such simulations. The region of interest is a few scale heights below the nominal exobase to a few scale heights above where collisions can be ignored. Perturbations can be produced by transient solar events affecting the absorption of short wavelength radiation, by a heat pulse due to a transient flux of the ambient plasma and pick-up ions, or by a gravity wave formed at depth propagating into this region. We do not try to describe how the observed density perturbations are produced, rather, our goal is to better understand the implications of the density perturbations observed in the transition region which can affect our interpretation of the heating and atmospheric evolution. We first describe the simulations. Then, in section 5.3, we simulate perturbations in two atmospheres, O only and O+CO_2. We use Mars-like properties, although the results are meant to be general. Finally, in section 5.4, we show that the simulated temperature profile can differ significantly from temperature profile extracted from density variations with altitude. This indicates to us that molecular kinetic simulations might be required to understand spacecraft observations that exhibit significant density perturbations in the transition region of an atmosphere.

5.2 Model

5.2.1 Description of the DSMC

In the DSMC method, the motion of atmospheric molecules is followed, subject to gravity and mutual collisions using a large number of numerical particles. Each numerical particle represents a very large number of physical particles, called the statistical weight w_s (Bird, 2013). Our simulation domain is composed of 55 cells whose sizes range from 6 to 7 km depending on the altitude, with the bottom and top boundaries at 100 and 450 km. These values are subsequently varied to be sure that their choice does not affect the outcome with the cell-sizes of order or smaller than local mean free path. The density n_s and the temperature T_s for species s in cell i are computed as:

$$n_s(i) = \frac{N_s(i) w_s}{V_i} \tag{5.1}$$

$$T_s(i) = \frac{m_s}{3k_B} \left(\frac{\sum_{p=1}^{N_s(i)} v_p^2 w_s}{n_s(i) V_i} - \langle v_{p_s}(i) \rangle^2 \right) \tag{5.2}$$

where $N_s(i)$ is the number of test particles of type s with mass m_s in cell i, k_B is the Boltzmann constant, v_p is the velocity of the particle p and V_i is the volume of cell i. $\langle v_{p_s}(i) \rangle$ is the average velocity in the cell i for the species s. Particles of species s are assigned a weight $w_s = N_s/N_p$, where N_s is the total column density and N_p is the total number of test-particles created at the initialization. These particles are initially distributed to obtain a barometric density profile with velocities chosen from a Maxwell-Boltzmann (MB) distribution. At each time step, dt ∼0.5 s, particles are ejected from the lower boundary using an upward flux, $\Phi_{0_s} = n_{0_s} \langle v_s \rangle /4$, with n_{0_s} the density and $\langle v_s \rangle = \sqrt{\frac{8k_B T_s}{m_s \pi}}$ the average velocity. Reducing the time step to dt ∼0.1 s did not affect our results. The velocity of the particles entering from the lower boundary is chosen from a Maxwell-Boltzmann Flux distribution (Smith et al., 1978). Particles with energy smaller than the escape energy that cross the upper boundary are assumed to be ballistic. Their trajectories are still computed at each time step until they return to the simulation domain, where they collide with other particles. That is, we track the trajectories of all the particles, even beyond the simulation domain in which we compute the density and temperature. Such particles are often simply reflected. This procedure is adequate when simulating a *steady state* atmosphere but fails when simulating *transients*. Consistent with the 1D methods of extracting temperature from density profiles, we present results of 1D simulations, applicable when the horizontal scale of the perturbations is much larger than the local scale height. This of course eventually breaks down in the exosphere, as discussed below, and multi-dimensional simulations are in progress. We used a number of cross section estimates but only show results using cross sections from Lewkow and Kharchenko, 2014 recently applied at Mars Leblanc et al., 2017. Since the results are broadly applicable, they can be applied to other atmospheres by scaling (e.g., Johnson et al., 2015).

5.2.2 Simulations parameters

The effect of perturbations are calculated in either an O or an O+CO_2 atmosphere using gravity and densities like those in Mars upper atmosphere. After the atmosphere reaches steady state, a perturbation is generated by creating a density or a temperature pulse at 150 km of altitude where the atmosphere is collisional. For the simple O atmosphere the density at the lower boundary (100 km) is 10^{10} cm^{-3}, with a temperature of 270 K giving a scale height of ∼40 km, an exobase at ∼230 km and a mean free path at the lower boundary of ~ 1.5 km using an average O+O cross section of 4.5×10^{-16} cm^2. In the multi-component atmosphere, the density at the lower boundary for O and CO_2 respectively are 1.6×10^8 cm^{-3} and 2.9×10^{10} cm^{-3} with a temperature 270 K. Such parameters give a CO_2 scale height of ~ 15 km, a CO_2 exobase at ~ 200 km, and a mean free path for CO_2 at the lower boundary of ~ 0.03 km using an average $CO_2 + CO_2$ cross section $\sim 10^{-14}$ cm^2.

For a Mars-like atmosphere of the type being simulated, MAVEN data indicate that density amplitudes, $(n - n_0)/n_0$, of the order of or greater than $\sim 50\%$ are often observed propagating into the transition region (Terada et al., 2017), where n_0 is the background density. Figure 5.1 shows an example of NGIMS (Neutral Gas and Ion Mass Spectrometer, Mahaffy et al., 2015b) density profile, n, vs. altitude for O (in blue) and CO_2 (in red). From the density data in the left panel, we computed a background density, n_0, shown in black using a polynomial fit. Density amplitudes were then calculated from the measured and background densities and shown in the second panel from the left of Figure 5.1, . These amplitudes are seen to reach 20% and 50% for O and CO_2 respectively. Such relatively large perturbations, often seen at these altitudes, have motivated the simulations described here. Snowden et al. Snowden et al., 2013 developed a method to extract temperature profiles from density measurements below Titan's exobase. This method, described in section 5.4, has also been applied to MAVEN data (Yiğit et al., 2015; England et al., 2017) often over a broad range of altitudes even into the exosphere. Both T_0 and T extracted from n_0 and n by this method are displayed on the third panel of Figure 5.1. The temperature excursions calculated from this data are seen in the right hand panel to have amplitudes of 30-40%. Such large density amplitudes, observed on many MAVEN passes through Mars upper atmosphere (e.g., Chapters 4 and 6), and the corresponding extracted temperature amplitudes in Figure 5.1 are used here as a guide for the simulations that show the effect of such perturbations in the transition region of an atmosphere as discussed below.

To understand the propagation of perturbations in the transition region of an atmosphere, we first generate a very large initial density perturbation by adding particles in the cell at 150 km in each time step maintaining a density 2 times the initial local density for a relatively short time, ∼25 s, while maintaining the initial local temperature. Although such a pulse is clearly artificial, it is seen in Figure 2 to rapidly relax resulting in density perturbations of the order of those observed in Figure 5.1. Reducing this to amplitude of the size observed did not the change the implications as discussed below. We then initiated a heat pulse, produced by increasing the velocity of particles in the cell at 150 km in each time step maintaining a MB distribution with a temperature $T = 300$ K also for ~ 25 s. A difference of 30 K from the background corresponding to a temperature perturbation of $\sim 10\%$, which is smaller than the excursions for the extracted temperature in Figure

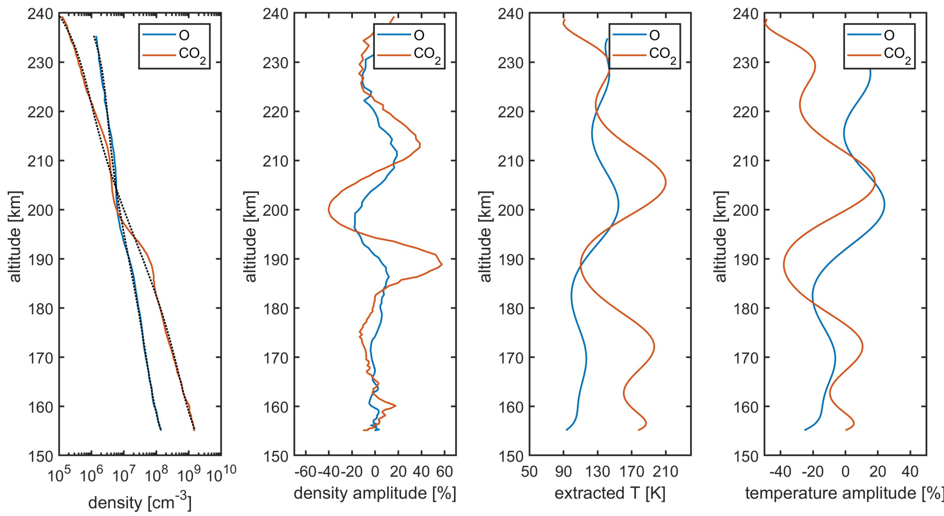


FIGURE 5.1: NGIMS data for O (blue) and CO_2 for orbit number 5854, from October 5, 2017. The local time of the orbit ranged from 14.2-15.0 hours covering latitudes $-11°$ to $29°$. From the left to the right: 1) density n in cm^{-3}, with black dotted lines the fitted densities; 2) the density amplitude, $(n - n_0)/n_0$; 3) the temperature extracted from the measured (colored) and fitted background (black) densitise using the method described in Snowden et al., 2013; 4) extracted temperature amplitudes $(T - T_0)/T_0$.

5.1.

Both extreme and modest perturbations, as well as changes in the pulse length, all of which correspond pressure variations, were found to exhibit similar behavior as they propagate through the transition region. To further test this, we also simulated a wave-like perturbation occurring at the lower boundary of the simulation regime for a pure O atmosphere. This was done by varying the incoming flux with time, t, at the lower boundary as $\Phi_s(t) = \Phi_{0_s}[1 + A\sin(B(t - t_0))]$ where t_0 is the start time. The amplitude, A, was varied from 0.05 to 0.25. Since the simulations all result in the same conclusions, we only present the results obtained for an amplitude $A = 0.25$ for which the density variations are of the order of amplitudes observed in Mars upper atmosphere (e.g., Yiğit et al., 2015; Terada et al., 2017) and are consistent with the data in Figure 5.1. In addition we varied the frequency B over a large range from about 1/4 of the the Brunt-Väisälä (BV) frequency to a few times that frequency, with $\omega_{bv} = \sqrt{-\frac{g}{n}\frac{dn}{dz}}$. g the gravitational acceleration and $(dn/dz)/n$ the inverse of the scale height of the background atmosphere giving a period, $2\pi/B \simeq 670$ s. Simulations were run varying the surface flux for a single period and for a number of periods which is equivalent to varying the local pressure.

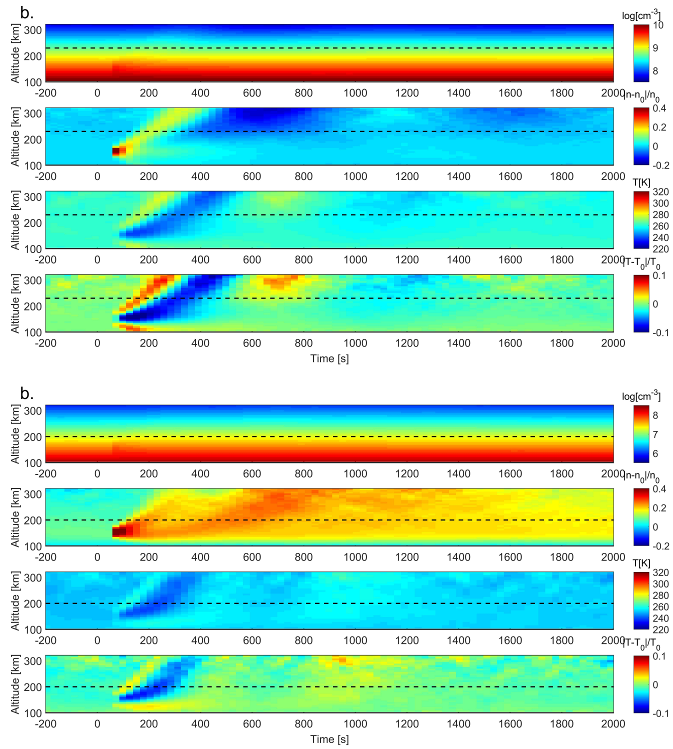


FIGURE 5.2: Evolution of O in an O atmosphere (top panels) and an O+CO_2 atmosphere (bottom panels) vs. time and altitude following a density pulse $2n_0$ for 50 time steps, $\sim$ 25 s. From top to bottom: density in cm^{-3} ; $(n - n_0)/n_0$; temperature in K ; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes described in text.

5.3 Perturbation propagation

The temporal evolution and subsequent relaxation of a relatively large initial density pulse is shown in Figure 5.2 for an O (top panels) and an O+CO_2 (bottom panels) atmosphere. The panels (top to bottom) show the density and its amplitude $(n - n_0)/n_0$, and the temperature and its amplitude $(T - T_0)/T_0$, with n_0 and T_0 the average, steady state values at time 0. The dotted lines indicate the nominal exobase altitudes described above. As the perturbations are produced in the collisional regime, in which the mean time between collision is short compared to the perturbation time, we find that the speed distribution stays close to a MB distribution during the perturbation so the gas in this region is roughly in local equilibrium. Therefore, the upward and downward particle flux from the perturbed region is approximately the MB flux $\Phi(i) \sim n_s(i) < v_i > /4$. Since the faster particles dominate the flow across any boundary, the corresponding energy flux for an MB distribution is $\sim (2k_bT_s)\Phi(i)$ and not $\sim (3k_bT_s/2)\Phi(i)$. Therefore, heat is transiently removed faster than particles following a perturbation, a kinetic effect seen in the simulations when the mean free path between collisions is not negligible. If the mean free path is indeed very small compared to any atmospheric length scale, this difference is equilibrated locally by collisions so that a thermal conductivity can be used. That is not the case in the transition region (e.g., Tucker et al., 2016) and it is seen that, even though the very large perturbation rapidly relaxes, the temperature enhancement precedes the density pulse.

The pulses (enhancements) propagate upward and downward, locally heating the atmosphere while cooling the perturbed region. At a time $t \sim 250$ s, it is seen that the downward propagating pulse appears to be 'reflected'. This feature is only marginally modified by either increasing the height of the perturbation or lowering the boundary. Therefore, the effect is due to the increase in the collision rate in the high density regime below the perturbation and disappears when collisions are suppressed. In a fluid dynamic sense, the perturbation is constrained by the buoyant force in this stable region of the atmosphere. It is also seen that the oxygen density remains larger than the steady state density even after ~ 2000 s in the multi-component atmosphere as seen at Mars in the perturbed upper atmosphere (Williamson et al., submitted). In these simulation we chose an O density so that the atoms experience roughly the same number of collisions in the lower atmosphere in both cases. However, collisions of O with the much heavier CO_2 result in a longer residence time in the lower atmosphere, a slower approach to steady state, and a smaller O temperature amplitude. The pulse amplitude, $(n - n_0)/n_0$, continues to grow above the exobase, a feature seen at Mars but becomes suppressed at high altitudes on Mars (e.g., Yiğit et al., 2015; Terada et al., 2017, Chapter 4).

Figure 5.3 compares the temporal evolution of CO_2 in an O+CO_2 atmosphere perturbed by a relatively large density pulse (top panels) and a modest heat pulse (bottom panels) for comparison. The temperature peak is again seen to precede the density peak for both perturbations. Following the pulse in the mixed atmosphere, the CO_2 component approaches steady state faster than the O component in Figure 5.2. The CO_2 stabilizes faster as they are heavier, have a much larger cross section, and are confined gravitationally to the higher density region in which the collision frequency is highest.

Finally, Figure 5.4a) shows the propagation in the transition region of a wave-like perturbation produced at the lower boundary for 5 BV periods. As the density at the lower boundary increases and decreases, the collision rate varies affecting the local temperature. It is seen that the wave pattern becomes roughly stable and dies out in ~ 2 BV periods, which is ~ 20 minutes in this model atmosphere. Figure 5.4a) shows the temporal evolution of the density and temperature amplitudes with altitude. The wave amplitudes are seen to increase with altitude in the transition region as expected, and, as shown in Figure 5.4b), the temperature pulse *again precedes* the density pulse at all altitudes due to the more rapid transport of the fastest molecules in the transition region of an atmosphere. We also find, not surprisingly, that the time separation between peaks grows slowly with altitude as the mean free path between collisions increases. Since these results are 1D they

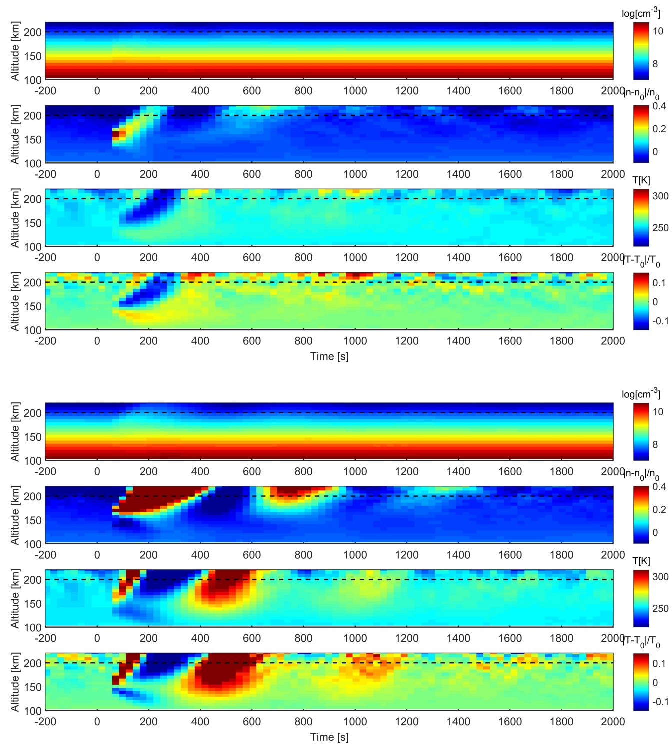


FIGURE 5.3: Evolution of the CO_2 component in an $O+CO_2$ atmosphere vs. time and altitude following a ($\sim$ 25 s) pulse at 150 km: top panels, density pulse ($2n_0$); bottom panels, heat pulse ($\Delta T \sim 30$ K). Individual panels as in Figure 5.2.

eventually break down at a few scale heights above the exobase.

5.4 Temperature extraction

In the extensive analysis of the upper atmosphere of Mars (Yiğit et al., 2015; England et al., 2017; Liu et al., 2017; Walterscheid, Hickey, and Schubert, 2013) the observed variations in the vertical structure of the density vs. altitude, interpreted as gravity waves, were used to extract the temperature structure using the 1D method in Snowden et al., 2013. The hydrostatic law was used to calculate a pressure vs. altitude profile from smoothed and extrapolated NGIMS density data. Based on the ideal gas law, that profile was subsequently used to extract the local temperature vs.altitude (England et al., 2017; Liu et al., 2017). As the perturbations propagated into the region above the nominal exobase ($\sim$200 km) Yiğit et al., 2015 and others cautioned the method could be problematic. Using the results in Figures 5.2 to 5.4 we show that these cautionary remarks are correct.

The integration of pressure vs. altitude from the measured density data requires a value for the pressure, P_u, at the upper limit of the data, n_u. Assuming $P_u = n_u k_B T_u$, the

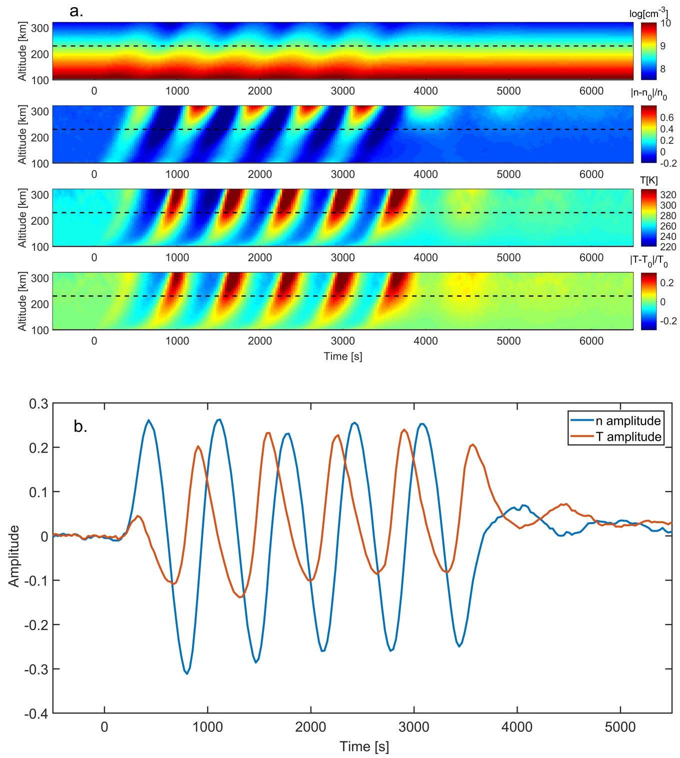


FIGURE 5.4: Response of an O atmosphere in altitude and time: a) injection flux at the lower boundary varies for 5BV periods (period$\sim$ 660s) with panels as in Figure 5.2. b)$(n - n_0)/n_0$ in blue and $(T - T_0)/T_0$ in red, in function of time, at 230km (exobase): time difference between peaks increases slowly with altitude.

temperature at the upper boundary, T_u, is estimated using:

$$\frac{d \log n(r)}{dr} = \frac{mg(r)}{k_B T_u} \left(\frac{\alpha}{C_p} - 1 \right) \tag{5.3}$$

where C_p is the specific heat, $g(r)$ is the gravitational acceleration, n the density and r the distance to the center of the body (Snowden et al., 2013). Although α was varied from 0 to ± 0.5 to take into account uncertainties in the extrapolation, we only show profiles using $\alpha = 0$. Changing alpha changes the temperature values at the highest altitudes but does not improve the agreement with the simulations.

Figure 5.5 shows the steady state density and temperature (solid lines) from our O and O+CO_2 simulations. Figure 5.5b) and e) confirm that the density and temperature amplitudes at steady state are nearly zero and the extracted and simulated temperatures are in agreement to within the uncertainties. In the following, the DSMC simulated kinetic temperature is compared to the temperature extracted from the calculated density profile in Eq. 5.1 which are in rough agreement in steady state (solid and dotted lines in 5.5c) and f)). In the following discussions, only density values below the nominal exobase altitudes

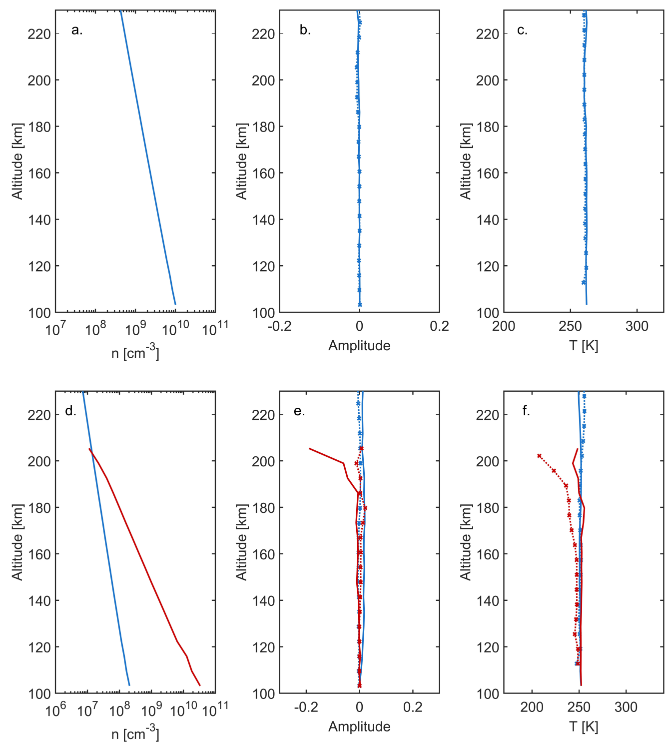


FIGURE 5.5: Steady state: top 3 panels O atmosphere; bottom 3 panels $O + CO_2$, O (blue), CO_2 (red): a),d) Simulated density; b),e) $(n - n_0)/n_0$ (solid), $(T - T_0)/T - 0$ (dotted). c),f) Simulated temperature (solid), extracted temperature (dotted). Weight differences between O and CO_2 account for differences in statistics above 180 km.

are compared due to the 1D nature of the simulations and the extraction method.

Figure 5.6 shows our key results. The vertical density and temperature profiles from Figures 5.2 and 5.3 at 55s after the perturbation are displayed. The top panels are for a density pulse in the O atmosphere (density in blue, temperature in red). The middle panels are from a temperature pulse in a two component atmosphere (O in blue, CO_2 in red). The bottom panels are extracted from the wave perturbation results in Figure 5.4a) using a profile of the first vertically propagating pulse, ~ 720 s after the perturbation. From the left to the right, the panels give the density, the amplitude, and the extracted and simulated temperature profiles. In all cases the thermal wave is seen to precede the density wave causing a transient thermal depression in the perturbed region, with thermal peaks propagating away from the region as discussed. In contrast to this, the extracted temperature simply follows the form of the pressure wave gradient. For the density pulse a difference of $\Delta T \sim -90$ K with respect to the steady state atmosphere is seen in Figure 5.6c), overestimating the local cooling of the atmosphere ($\Delta T \sim -60$ K). When the perturbation is due to even a modest heat pulse (red curves), the kinetic and extracted temperature profiles are almost out of phase. In particular, between 180 km and 220 km the model shows the atmosphere is heated by the perturbation with a temperature increase of $\Delta T \sim 50$ K while the extracted temperature shows it cooled ($\Delta T \sim -70$ K). These temperatures are also in serious disagreement in the O+CO_2 atmosphere (middle panels: O in blue, CO_2 in red). Below 160 km the simulations predict a small thermal perturbation whereas the extracted T reaches about 340 K for each species, which would require heating by $\Delta T \sim 70$ K. Above 160 km, the simulated temperatures peak at ~ 10 K for each species. The extracted temperature on the other hand requires local cooling of $\Delta T \sim -50$ K for the O and ~ -100 K for the CO_2 component. Finally, for a wave like perturbation propagating into this region from the lower atmosphere (bottom panels), the extracted and simulated temperatures, in the bottom right hand panel, not only disagree but are also out of phase. Therefore, the published thermal profiles in the upper atmospheres of Mars and Titan that are extracted from the density profiles should be re-examined using a molecular kinetic model.

The results presented above show suggest there can be a significant disagreement between the simulated and extracted temperatures at a given time for each simulation. However, we find the differences persist and are determined by the local density. Figure 5.7 shows results obtained from the GW like perturbation of a pure O atmosphere for 5 BV periods. From the top to the bottom, the three panels give the temporal evolution of the simulated temperature, the temporal evolution of the temperature extracted from the simulated densities following the method in Snowden et al., 2013 and the difference between the simulated and extracted temperatures. To compute the evolution of the extracted temperature in time, we used the simulated density to compute a temperature. These results clearly show that there can be a significant disagreement between the simulated and extracted temperatures during the perturbations propagation. These temperatures gradually come into agreement as the atmosphere again returns to steady state after the perturbations terminate. As this is a 1D simulation, the results are most relevant a scale height or so below the nominal exobase (here 190km) above which the dissipation/ dispersion of O increases. But it is seen that the extracted temperatures are "out of phase" and, mostly too large, even at low altitudes while the atmosphere is perturbed.

5.5 Summary

Molecular kinetic simulations were carried out to describe the propagation of perturbations through the transition region of a single component and a two component atmosphere. The amplitude of the perturbations simulated were primarily guided by the MAVEN NGIMS data, but not meant to reproduce those observations. Our primary goal was to examine the implications for determining the local temperature as a disturbance propagates through the transition region of a planet's atmosphere. In the absence of a significant perturbations or wave activity, we showed that such simulations reproduce the temperature profile extracted from the density vs. altitude data below the nominal exobase as verified in Figure 5. However, this was found not to be the case when density or temperature perturbations propagate through a planet's transition region. Because the mean free path between collisions is not negligible, the temperature pulse is out of phase with the density pulse, unlike what is found using 1D continuum models to extract temperature from measured density profiles in this region. Therefore, published temperature profiles extracted from density data below the exobase, but in the transition region, could be incorrect, possibly affecting our understanding of the physics and chemistry in this region. We show this is the case for small and relatively large perturbations, and even when the density variations are driven by wave-like perturbations from below at a variety of amplitudes. However, we have also shown that as the perturbation frequency becomes much smaller than the BV frequency (i.e, the periods much longer) then the actual temperature fluctuations become much smaller. However, the difference between the simulated temperatures and the temperatures extracted from the density become even larger and are still out of phase in this region of an atmosphere. Not surprisingly, in the two component atmosphere, the heavy species quenches faster than the light species, and, although the density amplitude grows as the perturbations propagate upward through the transition region, the growth with altitude differs from what is expected from linear theories (e.g., Hines 1960). Well above the nominal exobase, the observed amplitudes at Mars eventually decrease (Chapter 4), requiring multi-dimensional MK simulations to better understand the propagation across the exobase. However, the results presented here, which are generally applicable, are a cautionary note, suggesting that molecular kinetic simulations might be needed in some instances to better interpret measured density perturbations in the transition region of a planet's atmosphere, which is work in progress.

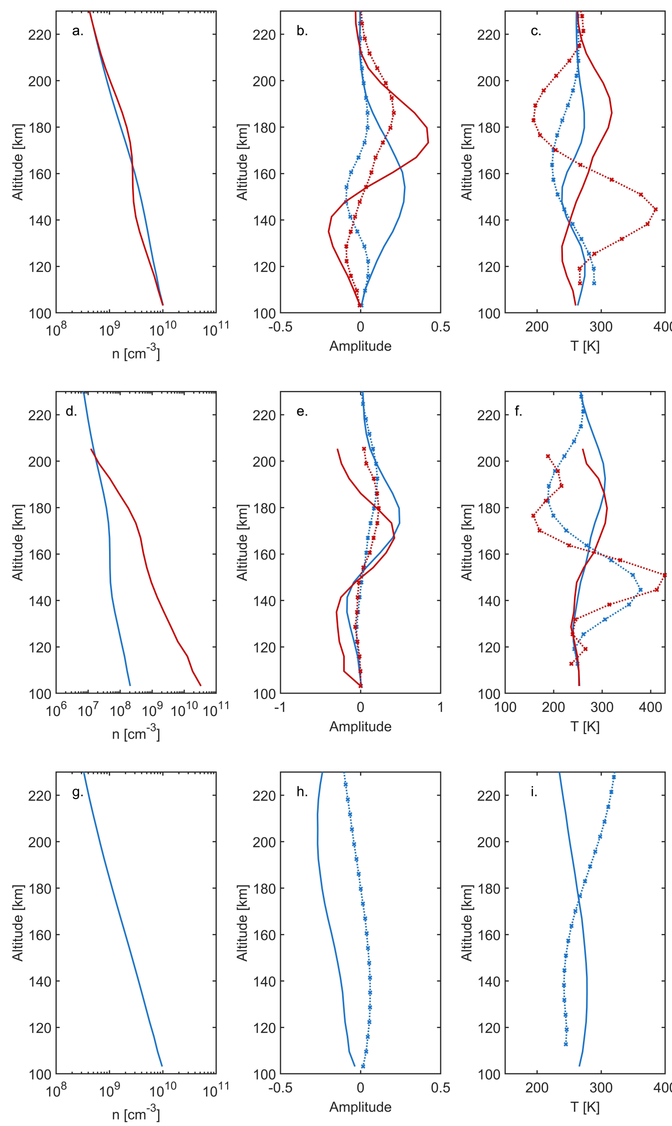


FIGURE 5.6: a)b)c) O atmosphere. Results obtained $\sim$ 55s after a density pulse (blue) and heat pulse (red). d)e)f) results obtained $\sim$ 55s after a heat pulse in a two component atmosphere (O blue, CO_2 red). g)h)i) wave perturbation in an O atmosphere $\sim$ 720 s after initiation (first pulse in Figure 3). a)d)g) Density (cm^{-3}); b)e)h) simulated n (solid) and T (dotted) amplitudes; c)f)i) simulated (solid) and extracted (dotted) temperatures in K for each species.

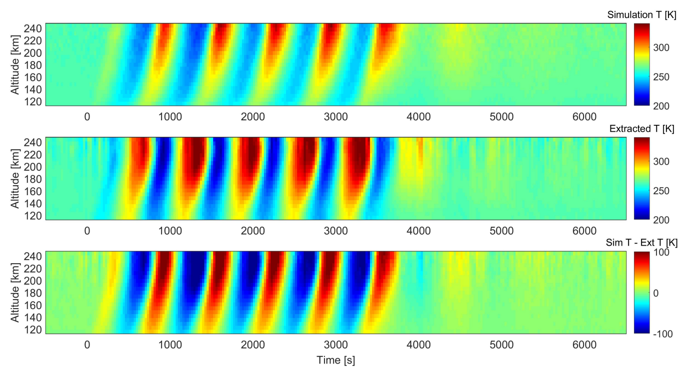


FIGURE 5.7: Evolution of temperature in degrees K calculated from the simulation of a pure O atmosphere in Figure 5.4 in which flux at the lower boundary varies at the $\sim BV$ frequency starting at time equal to zero for 5 periods. Top: simulated temperature; middle: the temperature extracted from the density profile; bottom; difference between the simulated and extracted temperatures.

Chapter 6

Amplitude and Location Trends of Exospheric Perturbations

6.1 Introduction

As discussed in Chapter 4, we have found 252 examples of MAVEN orbits where the NGIMS data show perturbations in the exosphere that satisfy the follow criteria:

1. There must be a perturbation with an Ar or CO_2 peak amplitude greater than 40% of the background density, where amplitude percentage is defined by $\frac{n_{data}-n_0}{n_0} \cdot 100$. O amplitude is not considered here, as it is generally half that of Ar due to its smaller mass.

2. The profile exhibits only one peak with Ar or CO_2 amplitude larger than 40% of the background density. Most profiles also have small amplitude perturbations with amplitudes below this limit, assumed to be of the same wave train.

3. The peak amplitude must occur above 5×10^{18} atoms/m^2, a rough estimate of the onset of the transition into the exosphere as discussed below. This ensures the features examined dissipate in the nearly ballistic regime.

These perturbations, most likely gravity waves, may have a variety of potential causes. These include topographic origin, although waves generated in the lower atmosphere are unlikely to propagate to exospheric altitudes; winds, which have been detected in recent NGIMS wind campaigns at these altitudes; and other sources of temperature changes, as any disturbance in an atmosphere can generate an internal wave, which then propagates under the force of gravity. Here, I show that these perturbations, while large in amplitude and high in altitude, have much in common with gravity waves observed at lower altitudes (e.g. England et al. (2017), Terada et al. (2017), and Yiğit et al. (2015)).

In this chapter, I examine the distribution of orbits containing an exospheric perturbation as defined above in a variety of coordinates, including local time, latitude and longitude in MSO and MSE coordinates (defined in Chapter 3 and included to see if ion precipitation is a possible generation mechanism), and the solar wind magnetic field and proton velocity vectors. Solar wind variables are included to discuss the possibility that solar wind interactions with the neutral atmosphere are a generation mechanism for these exospheric perturbations, as their large amplitudes and high altitudes are suggestive of their generation high in the thermosphere. Waves generated lower in the atmosphere are subject to viscous dissipation (Hines, 1960). Thus, while amplitude of the wave increases with altitude, quasi-linear studies show that waves with these large amplitudes that propagate from the lower to the middle atmosphere saturate in the middle atmosphere, well below the area of interest for this study (Fritts and Dunkerton, 1984). This indicates that the perturbations may begin to propagate much higher in the atmosphere than, for example, an orographic wave, making the addition of energy to the atmosphere by the solar wind a possible source of the initial instability.

In addition to the distribution of perturbation orbits, the amplitudes of the perturbations in the three species of interest, O, Ar, and $CO_2$are compared to the variables described above. With the exception of the comparison to solar wind variables, this is similar to the work done in Terada et al. (2017), which provides a valuable comparison to a known set of thermospheric gravity waves. While criteria 1 sets a lower limit for the amplitude of the exospheric perturbations, there is still a wide range of amplitudes, with the maximum amplitudes reaching approximately 150% of the background profile. All three species are examined because, as previously stated, the differing masses result in a smaller amplitude for O.

It has been well-established that the amplitude of thermospheric gravity waves is inversely proportional to background atmospheric temperature (Terada et al., 2017), even as it also grows with altitude. As shown in Appendix A, using first order perturbation theory for the Taylor-Goldstein plane wave equations produces a wave equation where the amplitude goes as $\sim e^{z/2H}$, where H is the scale height defined as kT/mg. At the altitudes examined in this study, the O scale height is significantly larger than that of Ar or CO_2, giving a correspondingly smaller amplitude. By analyzing the amplitudes of the exospheric perturbations, trends in local time (a proxy for background temperature) can be compared to those seen for thermospheric gravity waves to determine if the perturbations are a subset of the larger population of gravity waves. As with the distribution, the change in amplitude with respect to location and solar wind variables can also serve as a potential clue to the perturbation generation mechanism.

6.2 Data

For this work, the NGIMS level 2, version 08, revision 01 neutral density data was used. In the NGIMS level 2 data product, instrument particle counts have been converted to mass densities. Spacecraft ephemeris data such as latitude, longitude, altitude, local solar time, and solar zenith angle are also included. Thus species-specific density profiles are produced; here, as before, the focus is on CO_2, Ar, and O. Additionally, I use the MAVEN Key Parameters (KP) *in situ* data set, version 12, revision 1 to retrieve solar wind variables, as the KP data files collate data from multiple instruments, including MAG and SWIA. Of particular interest for this work were the solar wind magnetic field and the solar wind velocity vector, as both as indicative of solar activity and could potentially have an effect on the wave-like structures seen in the data.

I first sought to ascertain if the assumed exospheric gravity waves are more likely to occur under certain conditions, i.e. if they are more common at particular local times or under certain solar wind conditions. This could, in theory, offer insight into the cause of these large amplitude, high altitude waves that are able to propagate into the exosphere without dissipating. Rather than simply see how many perturbations were observed in a particular region, the data is normalized by binning the number of observations, then dividing the number of perturbation observations in a particular bin by the total number of NGIMS observations in the same bin. This gives the ratio $\frac{n_{pert}}{n_{tot}}$. The ratio is taken rather than the relative difference $\frac{n_{pert}-n_{tot}}{n_{tot}}$ so that the range of values increases from 0 to 1 instead of -1 to 0 for ease of plotting.

However, while looking at when perturbations are likely to occur can give information about the generation mechanism, it is most likely that these perturbations are generated by a number of mechanisms, as gravity waves can occur when there is any type of instability in the atmosphere, whether due to orographic flow or temperature gradients higher in the atmosphere. Therefore the peak amplitudes of the perturbations in all three species is examined as a function of various conditions, as it is possible that a particular variable may not contribute to perturbation generaly, but may affect the resulting amplitude.

The amplitudes were obtained by fitting a 5^{th} degree polynomial for the full density vs amplitude profile for all three species. This is similar to the fitting method used in England et al. (2017) and Yiğit et al. (2015), although we chose to use a 5^{th} degree polynomial rather than a 7^{th} degree polynomial as it better captured long-wavelength perturbations such as many of those found in the case study orbits. This fit is used as a smooth background profile and find the amplitude relative to the background, with the amplitude being $\frac{n_{data}-n_0}{n_0} \cdot 100$, where n_0 is the density of the background fit. This gives an amplitude that is a percent difference between the idealized smooth background and the data. These amplitudes are then compared to a variety of variables.

Both frequency of occurrence and amplitude variation are examined versus the following variables:

- Local solar time (LST) and solar zenith angle (SZA)
- Mars-Solar-Orbital (MSO) coordinates
- Mars-Solar-Electric (MSE) coordinates

- Interplanetary magnetic field (IMF) vector in MSO coordinates
- Solar wind velocity vector in MSO coordinates

MSO coordinates, obtained from the KP data and the NASA Navigation and Ancillary Information Facility (NAIF) SPICE solar geometry software, refer to the Mars-fixed coordinate system wherein the X unit vector points towards the Sun, the Y unit vector is anti-parallel to the direction of the orbit, and the Z unit vector completes the right-handed system. In the MSE coordinate system, the X unit vector remains the same as in MSO coordinates, but the direction of the XZ plane is found by calculating the positive direction of the solar wind convective electric field vector. This vector is found using Ampere's Law and the IMF and velocity vectors given in the KP data: $\vec{E} = \vec{v} \times \vec{B}$. The Z unit vector is taken to be that orthogonal to X in the plane of the electric field, and the Y unit vector completes the right-handed system. Chapter 3 discusses in detail how this configuration means ions preferentially precipitate in the MSE southern latitudes, meaning that a higher occurence of perturbations or larger amplitudes in souther MSE latitudes could indicate that the exospheric perturbations are at least partially effected by ion precipitation. For both MSE and MSO coordinates, longitude is roughly equivalent to local time, while in MSO coordinates latitude is roughly equivalent to geographic latitude.

6.3 Results

6.3.1 Frequency of occurrence

First, I examine where the perturbations are more likely to occur compared to the number of MAVEN orbits as described in the text above.

Figure 6.1 shows the ratio of perturbation orbits to total orbits for LST (top panel) and SZA (bottom panel). A cubic fit to the result is also included for both panels to serve as a general trend guide. The perturbations are generally more likely on the nightside, indicated by low/high LST and high SZA. However, the ratio also spikes near the terminators, around 6 and 18 LST and 90° SZA, likely due to the instability induced by the temperature gradient across the terminator. The higher occurrence on the nightside is consistent with lower amplitude thermospheric gravity waves as seen in Terada et al. (2017).

I then examine the likelihood of perturbations in MSO and MSE coordinates in Figure 6.2. In both panels, the ratio of perturbation orbits to total orbits are binned by longitude on the X axis and latitude on the Y axis. Color then indicates the value of the ratio, so brighter colors indicate bins where perturbations are more likely relative to the total number of orbits in that bin. Noon in both panels is located at 0° longitude. The distribution is fairly evenly spread in MSE coordinates, while in MSO coordinates, the northern dayside hemisphere is relatively empty of exospheric perturbations, consistent with the lack of perturbations at low SZA seen in Figure 6.1.

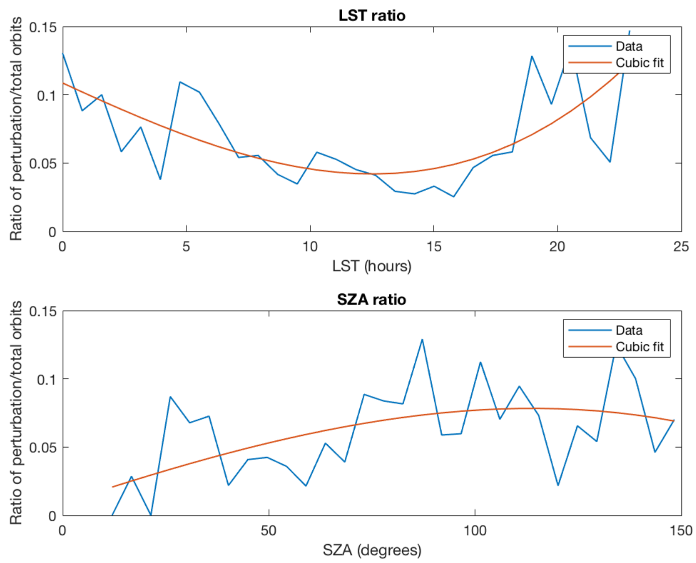


FIGURE 6.1: The ratio of number of perturbation orbits to total orbits, found by separating both into 30 bins then taking the ratio of each bin. The top is binned by LST and the bottom is binned by SZA. The red line indicates a cubic fit to roughly show the trend.

Similarly to Figure 6.1, Figure 6.3 shows the ratio of perturbation orbits to total orbits as a function of the IMF vector as measure by the MAG instrument and retrieved from the KP data files. There is no trendline included. The X, Y, and Z directions are in MSO coordinates as described in Section 6.2. The X axis indicates the full range of the data, as the ratio indicated in the graph is often zero. The figure shows that perturbations are more likely at relatively low IMF magnitudes. They are also more likely for negative B_Z, i.e. when the Z component of the IMF points to the ecliptic south. However, the most obvious trend is that perturbations are unlikely when the IMF magnitude is high. It is possible that this is due to the general lack of observations at solar extremes, as the Sun has been largely quiescent for the duration of the MAVEN mission. Thus while I have attempted to normalize the data, with few total observations for these magnitudes, seeing a perturbation will be unlikely, as they only comprise a small percentage of all the density observations.

Figure 6.4 shows the ratio as a function of the solar wind proton velocity given by the SWIA instrument retrieved from the KP data set. The panels are the same as in Figure 6.3. The top panel shows that perturbations are more likely when the Mars-directed component of the solar wind is low, although this is not a strong trend and perturbations appear even at much higher velocities. For comparison, the average v_x velocity is ≈ 400 km/s. This would seem to indicate that the X component of the solar wind velocity is not a strong factor in the occurrence of exospheric perturbations. Likewise, there are also no

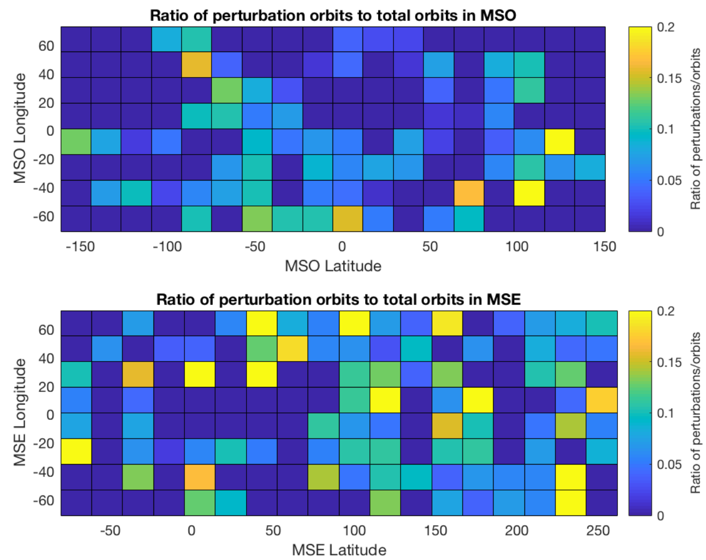


FIGURE 6.2: The ratio of number of perturbation orbits to total orbits, in MSO and MSE latitude and longitude. The data for perturbation orbits and all orbits was binned into $9^\circ \times 9^\circ$ bins, and then the ratio of the two is taken. In both the top and bottom, 0 longitude indicates noon local solar time. The color indicates the ratio value.

strong trends for the Y and Z components, as perturbations are visible for the full range of magnitudes.

6.3.2 Trends in amplitude

Next, I examine trends in the amplitude of the exospheric perturbations, which produces slightly different results than in 6.3.1. As before, the species-specific amplitudes are plotted as a function of LST and SZA. As described in Chapter 4, due to differences in species scale heights in the thermosphere and exosphere, each species will have a different amplitude, here defined as the percent relative difference between the polynomial background fit and the NGIMS data. The O amplitude is consistently approximately half that of Ar and CO_2 due to its lighter mass. Thus, in this section, the amplitudes are plotted for the individual species. In the following figures, blue dots represent the O amplitudes, red open circles the Ar amplitudes, and yellow asterisks the CO_2 amplitude. Again as before, a cubic polynomial fit to the CO_2 amplitudes is included as a general guide of trend.

In Figure 6.5 the overall trend is consistent with that seen in Figure 6.1, in that amplitudes are larger on the nightside. However, in the top panel there are also high amplitudes near noon LST. This will be discussed further later.

To study changes in amplitude relative to location in MSO and MSE coordinates, the perturbation observations are binned by their latitudes and longitudes in both coordinate

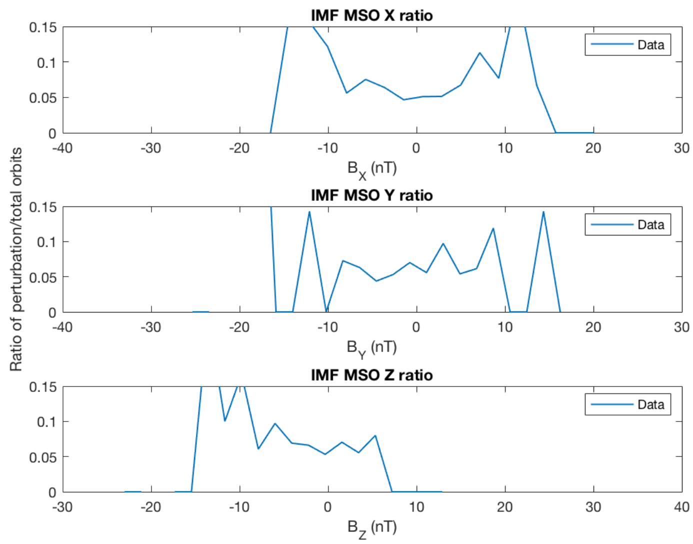


FIGURE 6.3: The ratio of number of perturbation orbits to total orbits binned by the IMF vector in MSO coordinates. X, Y, and Z are as described in the text.

systems. I then find the mean amplitude per species in each bin. This is similar to the process described in Chapter 3 for creating mean density maps, with the exception that these are average amplitude percentage maps, not average density maps. So then in Figures 6.6 and 6.7, the X axis indicates longitude, the Y axis latitude, and color the mean amplitude for that particular bin. The panels are, from top to bottom, the amplitudes for O, Ar, and CO_2. A dark blue bin with a value of zero indicates no perturbations detected in that bin. This can be because there were no MAVEN orbits in that location, so the interpretation of these figures will ignore these bins.

In MSO coordinates, there are slightly larger mean amplitudes in the nightside (high longitudes), which would be expected given the amplitudes as a function of LST shown in Figure 6.5, as high longitudes in MSO correspond to night LST values. There are also higher amplitudes in the MSO polar regions evident in all three species.

For MSE coordinates shown in Figure 6.7, there are again generally higher amplitudes on the nightside (high longitudes). Amplitudes, particularly for CO_2 are also slightly higher in the northern MSE latitudes, i.e. the region where the solar wind convective electric field is directed away from the planet. The northern high longitudes in MSE coordinates roughly correspond to the region were ion escape due to the ion polar plume is most likely (Dong et al., 2015).

Figure 6.8 shows the species amplitudes as a function of IMF magnitude and direction. The panels are, from top to bottom, the MSO X direction, Y direction, and Z direction.

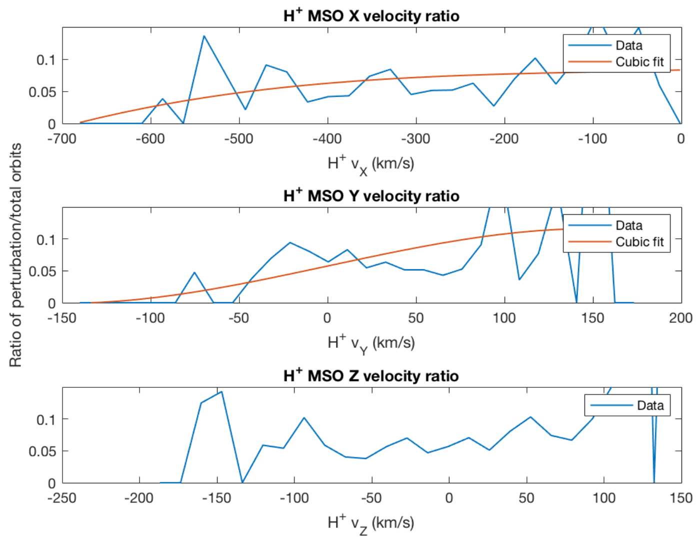


FIGURE 6.4: The ratio of number of perturbation orbits to total orbits binned by the solar wind velocity vector in MSO coordinates. X, Y, and Z are as described in the text. The red line is a cubic fit to the data as a rough indication of trend.

There is no real trend in amplitude dependence on the IMF vector, except possibly that extreme IMF magnitudes are associated with smaller amplitudes.

Likewise, in Figure 6.9 there is also no evident correlation between species amplitude and solar wind proton velocity in either magnitude or direction. Thus solar wind velocity and magnetic field likely have no effect on the amplitudes of these exospheric perturbations, even if they possibly affect the generation mechanism of the perturbations.

6.4 Discussion

Overall, there are not any trends in the perturbation locations and amplitudes that vary significantly from what would be expected for gravity waves, e.g. as in Terada et al. (2017). There is evidence that the perturbations occurrence frequency and amplitudes are inversely proportional to background temperature, so that there are both more and larger perturbations on the nightside of the planet. This is evident in Figures 6.1, 6.5, 6.2, and 6.6 and is indicative that these perturbations are a subset of the gravity waves ubiquitous in the Martian atmosphere, as both other gravity wave data and theory predict that amplitude will be inversely proportional to scale height and, hence, temperature (Terada et al., 2017; England et al., 2017; Hines, 1960; Midgley and Liemohn, 1966).

However, there are two key features that are slightly different from the gravity waves observed in the thermosphere. Firstly there is an increase in perturbation occurrence near

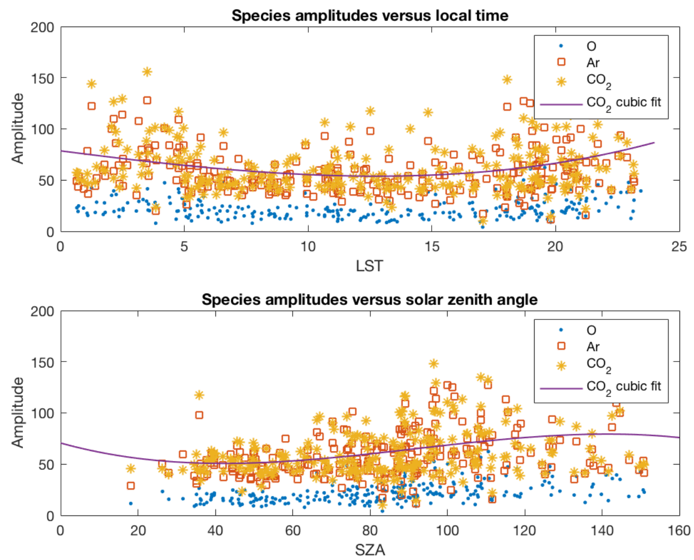


FIGURE 6.5: The species-specific amplitudes as a function of LST (top) and SZA (bottom). The line in both panels is a cubic fit to the CO2 amplitudes as a rough indication of trend.

the solar terminators in the top panel of Figure 6.1. This is likely because there is a large temperature gradient at the terminator, which is one of many generation mechanisms for a gravity wave, as a temperature gradient introduces an atmospheric instability, which may then propagate (Hodges, 1967). Therefore the peak in perturbation observations at the terminators supports the theory that these exospheric perturbations are indeed large-scale gravity waves that have been able to propagate past the exobase. Additionally, Figure 6.5 shows an increase in wave amplitude near noon LST, contrary to the expected inverse dependence on background temperature. Generally, large amplitude waves are less likely in warmer temperatures because the likelihood of wave saturation is increased when the background temperature is higher (Hickey, Walterscheid, and Schubert, 2011). At higher temperatures, thermal conductivity and viscous forces increase, making a wave more likely to be damped. Therefore on the dayside atmosphere, it is likely that the only waves able to escape dissipation at lower altitudes are those with sufficiently large amplitude to avoid thermal and viscous forces, yet not so large that wave breaking or saturation occurs. Therefore the increase in amplitude seen near noon LST is likely due to the narrow range of wave parameters that will permit high altitude propagation at higher background temperature.

Both the distribution and amplitude in MSO coordinates are again consistent with

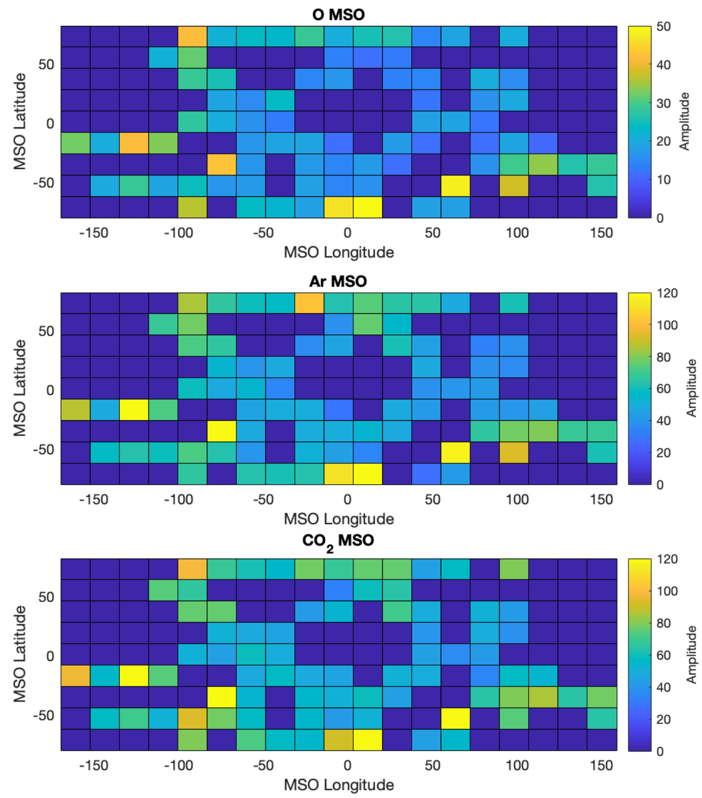


FIGURE 6.6: The species-specific amplitudes as a function of MSO latitude and longitude. Color here indicates the mean amplitude in each 9° × 9° bin. 0° longitude is equivalent to noon LST and ±90° are roughly equivalent to the north and south geographic poles.

prior studies of thermospheric gravity wave populations, that the exospheric perturbations are both more likely and larger amplitudes on the nightside of the planet. Interestingly, the perturbations are also more likely and larger amplitude in the MSO polar regions, with the equator here corresponding roughly to the orbital plane of the planet (see Chapter 3 for a full explanation of MSO coordinates). It is possible that cross-polar flow at high altitudes, seen commonly on Earth as a cause of polar stratospheric clouds (Fritts and Alexander, 2003; Miller et al., 2015), is a generation mechanism for these exospheric perturbations. Indeed, Hunsucker (1982) shows that large scale traveling ionospheric disturbances induced by long wavelength gravity waves frequently originate in the polar regions then propagate equatorward. Additionally, three-dimensional global circulation models of the Martian atmosphere show a strong temperature gradient across the poles, with high latitudes often having high temperatures in the upper thermosphere around

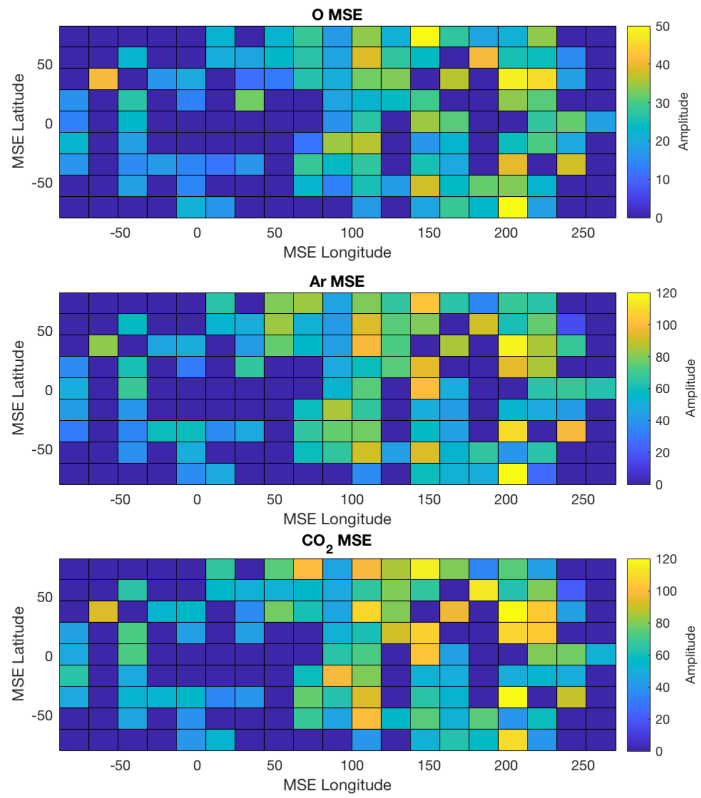


FIGURE 6.7: The species-specific amplitudes as a function of MSE latitude and longitude. Color here indicates the mean amplitude in each 9° × 9° bin. 0° longitude is equivalent to noon LST and latitude is described in the text.

190 km (Valeille et al., 2009). The large temperature gradients and flow velocities at altitudes near those where for exospheric perturbations are therefore a likely explanation of the enhanced distribution and amplitudes for polar latitudes seen in Figures 6.2a and 6.6.

The purpose of examining the distribution and amplitudes of the exospheric perturbations in MSE coordinates was to determine if ion precipitation could be a generation mechanism for these perturbations, as precipitating ions are likely to deposit their energy, potentially causing neutral sputtering, near the exobase (Johnson and Luhmann, 1998). Because one way for a wave to occur at these high altitudes without dissipating is for it to also be generated at high altitudes, I search for potential generation mechanisms that occur in the upper thermosphere or lower exosphere. In MSE coordinates, the XZ plane is determined by the direction of the solar wind convective electric field, given by $\vec{E} = -\vec{v} \times \vec{B}$. The X vector points towards the Sun, and the Z vector is then chosen to be

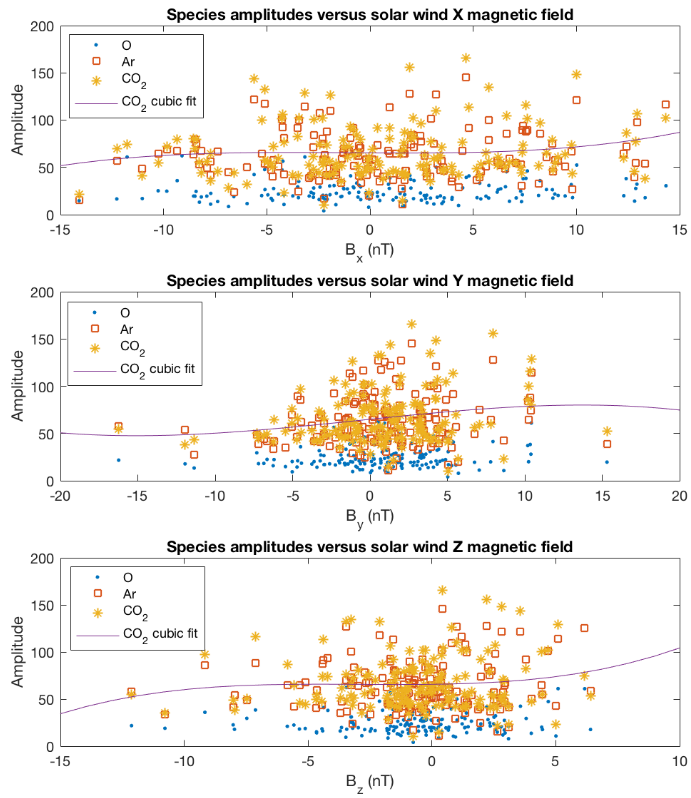


FIGURE 6.8: The species-specific amplitudes as a function of IMF vector in MSO coordinates. The red line indicates a cubic fit to the CO_2 amplitudes as an indicator of trend.

that orthogonal to X in the plane of the electric field. As a result, latitude in MSE coordinates indicates whether the solar wind convective electric field is pointing away from the planet (the +E or positive latitudes) or towards the planet (-E or negative latitudes). Because of this configuration, ion precipitation is more likely to occur in the negative latitudes as the ions travel along the electric field lines. Longitude, as with MSO coordinates, roughly corresponds to local time. In Chapter 3 MSE coordinates were similarly used as a way to look for the effects of ion precipitation on neutral species.

If ion precipitation was a common cause of these exospheric perturbations, we would expect to see either more or larger amplitude perturbations in the negative MSE latitudes. However, Figure 6.2b shows no trends in the distribution of the perturbations in MSE coordinates. This does not necessarily preclude ion precipitation from being a source of the perturbations, but it is clearly not the sole or most likely cause. Figure 6.7 shows broadly that amplitudes grow with increasing longitude, corresponding with similar trends in Figure 6.5. However, Figure 6.7 shows that, particularly for CO_2 (panel c), amplitudes

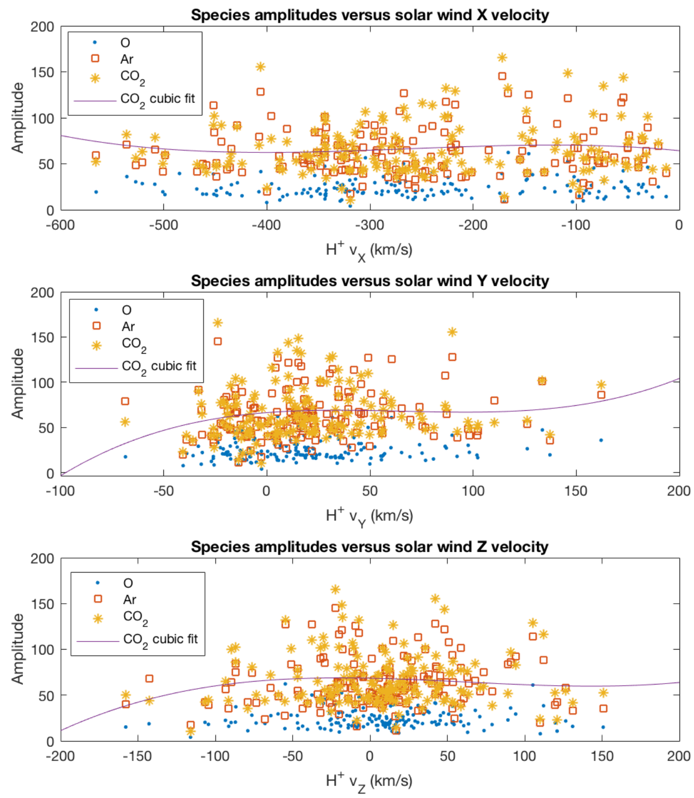


FIGURE 6.9: The species-specific amplitudes as a function of solar wind velocity vector in MSO coordinates. The red line indicates a cubic fit to the CO_2 amplitudes as an indicator of trend.

are on average larger in the MSE positive latitudes, where ion precipitation is less likely than ion escape. Ion precipitation in the negative latitudes could potentially induce turbulence, thus preventing the propagation of the exospheric perturbations and accounting for the smaller amplitudes in the latitudes where precipitation would be expected. In general, however, ion precipitation does not appear to be a strong contributor to exospheric perturbations.

For both solar wind magnetic field and proton velocity there are no obvious trends in either perturbation occurrence or amplitude. Table 6.1 shows for reference the average values of the solar wind variables shown here, taken from (Halekas et al., 2017) for solar wind velocity and Curry et al. (2015) for IMF strength. In Figure 6.3, it is apparent that for all three magnetic field directions, perturbations are more likely to occur when the field strength is low. However, it is possible this is due to sampling bias, as solar activity has been generally low during the MAVEN mission (McComas et al., 2013), so there are few orbits with extreme magnetic field strength, making the likelihood of a perturbation

TABLE 6.1: Average values of solar wind variables from Curry et al. (2015) and Halekas et al. (2017)

IMF component	Mean magnitude (nT)	$\vec{v}_p$ component	Mean magnitude (km/s)
B_x	-2.5	v_x	-408
B_y	2.8	v_y	23.7
B_z	1.0	v_z	-0.9

occurring simultaneously low, as they only occur in a few percent of the total MAVEN orbits. The solar wind magnetic field also does not appear to have a correlation with the perturbation amplitudes as shown in Figure 6.8.

There is a slight indication as seen in Figure 6.4a that perturbations are more likely for quiescent solar wind conditions, below the average v_x shown in Table 6.1. This may be due to higher velocity solar wind, and hence increased solar wind dynamic pressure, compressing the solar wind interaction region and ionosphere, shown to occur even in the absence of extreme solar weather (Halekas et al., 2017). This compression of the ionosphere would potentially change the neutral atmosphere enough that large amplitude waves are unlikely to propagate. Panel b of Figure 6.4 shows a higher likelihood of perturbation occurrence for high positive v_y values, but this is likely due to sampling bias, as not many observations were made with high v_y values. In panel c there is no correlation between the v_z magnitude and perturbation occurrence. This trend continues for the perturbation amplitudes, as neither Figure 6.8 or Figure 6.9 show any correlation between the solar wind and perturbation amplitudes.

6.5 Summary

In conclusion, the location and amplitude of the large amplitude exospheric perturbations studied here indicate they are most likely a special subset of the larger population of upper atmospheric gravity waves. Like those seen in the upper thermosphere, both the amplitude and likelihood of occurrence are inversely proportional to the background temperature. A possible generation mechanism is the upper atmospheric temperature gradient present both across the terminators and polar regions, known to induce large amplitude, long wavelength gravity waves in the ionosphere of Earth. While the possibility of ion precipitation as a possible perturbation source cannot be ruled out, comparing the perturbation data in MSE coordinates does not show this as a dominant cause of the perturbations. Likewise, it is unlikely that the solar wind at present plays a large role in exospheric perturbation generation, although atmospheric compression at high solar wind velocities may impede wave propagation.

Chapter 7

Characterizing Perturbation Parameters

7.1 Introduction

I have previously discussed characterizing the exospheric perturbations seen in NGIMS data by amplitude and location in local time, geodetic coordinates, and solar wind parameters. All of these indicate the perturbations are likely high altitude, large amplitude gravity waves able to propagate into the almost collisionless regime of the Martian atmosphere. Because gravity waves are an important mechanism for the vertical transport of energy, it is also likely that these exospheric perturbations deposit energy in the exosphere, which, as Walterscheid, Hickey, and Schubert (2013) show could affect the Jeans escape rate. However, the amount of energy deposited is highly dependent on wave parameters such as frequency and phase speed, so in order to understand how these perturbations are affecting the exosphere the perturbations should be characterized in terms of wave parameters.

This task is complicated by the elliptical spacecraft trajectory. The spacecraft, with the exception of very close to periapsis, travels predominately vertically, so it is generally assumed that the density profile found by NGIMS presents a vertical slice through the atmosphere. The MAVEN average orbital speed is $\sim$ 4.2 km/s and so the spacecraft takes on the order of a couple of minutes to complete the inbound leg of an NGIMS density profile. Because typical gravity and acoustic wave phase speeds are on the order of a hundred m/s (Hickey, Walterscheid, and Schubert, 2011; Walterscheid, Hickey, and Schubert, 2013), the density profile is most accurately described as a snapshot of a given wave packet in position space. This makes it possible to roughly estimate the dominant wavenumber spectrum for a perturbation, but difficult to estimate the dominant the wave frequency spectrum. Here I endeavor to find the wave frequencies and horizontal phase speeds using spectral analysis of the perturbations to obtain a wave number, then guided by the previously described DSMC model, which does describe the wave as a function of time, to determine the frequency of a wave with a similar wavenumber in an atmosphere with the same background density and temperature. These wave parameters can then be combined with the calculated acoustic cutoff frequency and BV frequency to estimate horizontal phase speed from the dispersion relation, the real part of equation 2.51. Having a better understanding of the general wave characteristics of the exospheric perturbations studied in this work is essential to understanding how they deposit energy in the exosphere, as models such as Hickey, Walterscheid, and Schubert (2011) among others show that phase speed and frequency can vastly alter the wave-induced atmospheric heat flux. Being able to estimate the wave frequency by approximating the wave with the DSMC model can also indicate whether the wave is a gravity wave or acoustic wave as discussed in Chapter 2, which affects energy propagation.

7.2 Data Analysis

7.2.1 Obtaining the wavenumber

The first step in finding the wave parameters is calculating the wavenumber spectrum. As stated above, due to the relative difference between the spacecraft speed and estimated wave phase speed, an exospheric perturbation density profile shows changes in density as a function of altitude rather than a function of time. Generally when doing spectral analysis of a wavefunction, the amplitude data is transformed to frequency space and thus a frequency spectrum is obtained. However, for this method to work, the wave must be sampled in time, unlike our perturbation profiles. As a result, the analysis below will be used to obtain a wavenumber spectrum.

The Lomb-Scargle periodogram is used to obtain this wavenumber spectrum(Scargle, 1982), given by

$$P_X(k) = \frac{1}{2}\left[\frac{\left(\sum_j X_j \cos k(z_j - \tau)\right)^2}{\sum_j \cos^2 k(z_j - \tau)} + \frac{\left(\sum_j X_j \sin k(z_j - \tau)\right)^2}{\sum_j \sin^2 k(z_j - \tau)}\right]$$

where X is the density signal, assumed to be a sum of noise and actual observation, z is the altitude the signal was taken, k the wavenumber, and τ is given by

$$\tan(2k\tau) = \frac{\sum_j \sin 2kz_j}{\sum_j \cos kz_j}$$

Using the Lomb-Scargle periodogram has advantages over either a Fourier transform or normal periodogram because it allows for the signal sampling to be unevenly spaced. With the amplitude vs altitude profiles discussed in Chapter 4 the data can be processed using the Lomb-Scargle periodogram function in the MATLAB Signal Processing Toolbox, which gives an output of power/(1/km) versus 1/km. The Lomb-Scargle method of spectral analysis has also been previously used to determine apparent gravity wave wavelengths from NGIMS data (England et al., 2017) and showed apparent wavelengths up to 300 km, much larger than those seen here. However, they are examining all NGIMS orbits, instead of only selecting those 252 orbits with a perturbation that fits our criteria.

Figure 7.1 shows an example of a Lomb-Scargle periodogram of the Ar amplitude profile from orbit 5962, which occurred on October 25, 2017. The data for orbit 5962 was taken near LST 12 and geographic latitude -6.7°. Like the periodograms for all perturbation orbits, the figure shows a prominent low wavenumber peak with power nearly an order of magnitude larger than higher wavenumber peaks, which is likely noise in the amplitude profile. In fact, this particular case has a relatively small difference between the two highest peaks in the spectrum; the mean difference between the power of the two highest peaks in all 252 spectra is 1.01×10^4, indicating that for all the spectra, there appears to be a predominant peak that is larger than all other peaks by several orders of magnitude. With such strong dominance of a single wavenumber, this likely indicates that the perturbations do not consist of multiple superimposed waves but a single wave train. The mean dominant wavenumber for all perturbation spectra is 0.038 km^{-1}, which gives a mean perturbation wavelength of 29.6 ± 9.5 km. These wavelengths are consistent with those found in Terada et al. (2017) for thermospheric gravity waves. While as mentioned previously the spacecraft speed is roughly an order of magnitude larger than a typical phase speed, the density profile is not a perfect snapshot and so there is some uncertainty to the wavenumber; however, this uncertainty is likely to be small because of the large difference between spacecraft and phase speed.

7.2.2 Inherent atmospheric frequencies

For calculating the dispersion relation and determining the type of wave, the acoustic cutoff frequency and the Brunt-Väisälä frequency must be found. The equations for these along with the speed of sound, given in Chapter 2 as equations 2.53 and 2.52, are:

$$\omega_a = \frac{\gamma g}{2C_s} \tag{7.1}$$

$$\omega_{bv} = \frac{(\gamma - 1)^{1/2}\, g}{C_s} \tag{7.2}$$

$$C_s = \sqrt{\gamma g H} \tag{7.3}$$

Here γ is the ratio of specific heats; g is the gravitational acceleration, which for around 200 km altitude is $\sim$ 3.4m/s^2; H is the scale height; ω_a is the acoustic cutoff frequency, ω_{bv} is the Brunt-Väisälä frequency, and C_s is the sound speed. While these values vary between species due to the differences in scale heights, the sound speed and frequencies

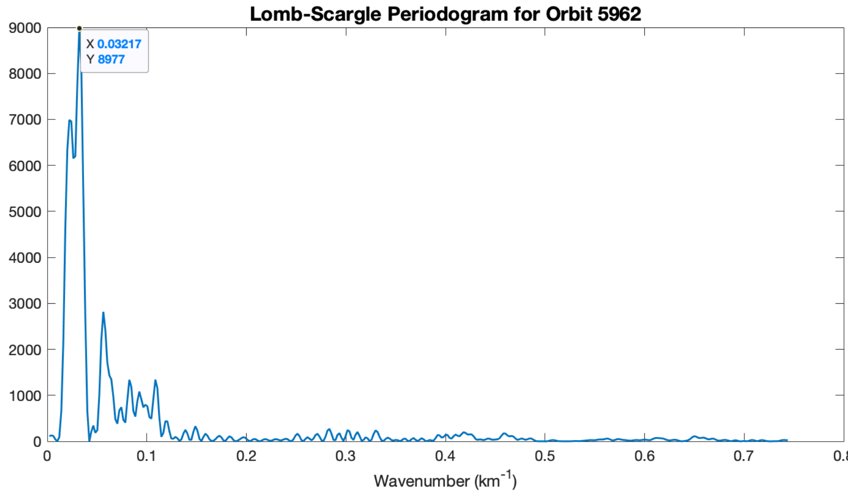


FIGURE 7.1: The Lomb-Scargle periodogram for MAVEN orbit 5962 Ar density vs altitude data on October 25, 2017. The x axis in km^{-1} shows vertical wavenumber and the y axis shows the resulting dimensionless power per wavenumber. The wavenumber with the highest power is highlighted with a text box, showing at x = 0.032 and y = 8977.

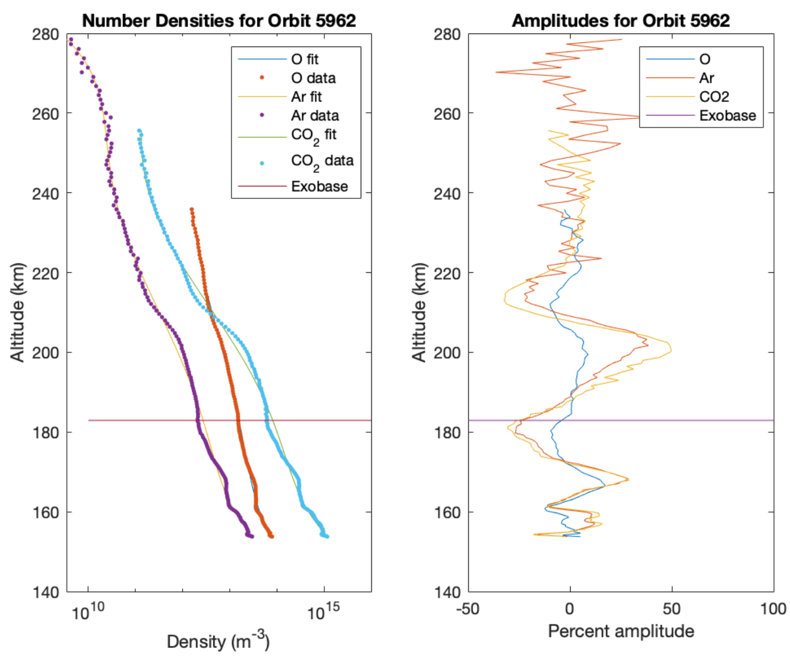


FIGURE 7.2: The density and amplitude vs altitude corresponding to the above Lomb-Scargle periodogram for orbit 5962. The lefthand panel shows the densities of O, Ar, and CO_2 versus altitude, with lines indicating the polynomial background profile. The righthand panel shows the species amplitudes found by taking the relative difference of the data and the background profile. The Ar amplitude profile is used in the above periodogram.

TABLE 7.1: Average perturbation orbit values of ω_a and ω_{bv} as calculated from NGIMS data for all 252 perturbations

	Mean	Median	Standard Deviation
ω_a (mHz)	10.6	10.4	1.8
ω_{bv} (mHz)	9.22	9.01	1.6
C_s (m/s)	221	220	41

for Ar are calculated, as it is non-reactive and so does not fractionate in the NGIMS instrument unlike CO_2. Because its mass is similar to CO_2 it therefore serves as a more easily calibrated proxy for CO_2 and can be used to calculate an average value for the atmosphere.

An atmospheric scale height near the exobase must be found in order to find the above variables. This is accomplished by fitting the Ar altitude-density profile such that

$$n_1 = n_0\, e^{-(z_1 - z_0)/H} \tag{7.4}$$

$$\implies H = \frac{z_1 - z_0}{\log\left(\frac{n_0}{n_1}\right)} \tag{7.5}$$

where 0 indicates the lower altitude and 1 indicates the higher altitude. For H_{Ar}, the Ar background profile densities and altitudes for a 20 km region around the approximate exobase are used to obtain an approximate exobase scale height for each perturbation orbit. At the relevant altitudes for the exospheric perturbations, the Ar density is negligible compared to the O and CO_2 densities, which are roughly equal. Therefore, to find the ratio of specific heats, I assume a half O and half CO_2 gas mixture, which gives $\gamma = 1.34$. With these values, I can then calculate an approximate sound speed near the exobase for each perturbation orbit and subsequently an acoustic cutoff and BV frequency. From equations 2.52 and 2.53, the difference between ω_a and ω_{bv} will be $\sim 0.14\sqrt{\frac{1}{H}}$ for our γ. Thus for small scale heights, the difference between the two is larger and vice versa. Because of the dependence on scale height, ω_a and ω_{bv} vary with local time, having generally larger frequencies on the nightside and lower frequencies on the dayside. The mean, median, and standard deviation of both frequencies and the sound speed are in table 7.1.

7.3 Modeling the Wave

To get an idea of the effects produced by the perturbations, a perturbation is simulated at three different frequencies in a mixed O-CO_2 atmosphere with similar densities and scale heights to those seen in the NGIMS data. The simulation software used here is the same as the multispecies simulation results shown in Chapter 5. For the particular case study below, we input background densities and scale heights corresponding to orbit 5962, shown in Figure 7.1. The calculated acoustic cutoff frequency for this orbit is 9.65×10^{-3} Hz and the BV frequency is 8.40×10^{-3} Hz, so a perturbation is simulated at frequencies of 7 mHz to simulate a gravity wave, 9 mHz to simulate an evanescent wave, and 11 mHz to simulate an acoustic wave. Similarly to Chapter 5, these pulses were simulated using a sinusoidal function for 5 periods and allowed to propagate vertically. The perturbations are introduced after the simulated atmosphere reaches a steady state, with both density and molecular kinetic temperature subsequently sampled throughout the model run.

As in Chapter 5, the evolution of the perturbation in time is seen by plotting the density, density amplitude, temperature, and temperature amplitude as a function of altitude and time. The figures below show the O densities and temperatures followed by the CO_2 densities and temperatures for each frequency. The modeling for these figures was run by Lucia Tian, with atmospheric parameters for the model, such as steady state density/temperature (an approximation derived from the scale height fit as described above) and wave frequencies chosen above from the NGIMS Ar data.

First is shown an example of a gravity wave with frequency lower than the atmospheric BV frequency. The wave is introduced for five full periods then subsequently allowed to subside. The frequency of the wave is 0.007 Hz, equivalent to a period of $T = 2\pi/f \sim 898s \sim 15$ min. Because the wave frequency $\omega < \omega_{bv}$, the dispersion equation 2.51 can be solved for horizontal phase speed, i.e. $c = \omega/k$, where k is the horizontal wavenumber and m is the vertical wavenumber found from our spectral analysis of the NGIMS amplitude profile for orbit 5962:

$$k_z^2 = k_x^2 \frac{\omega_{bv}^2 - \omega^2}{\omega^2} \tag{7.6}$$

$$\implies \frac{\omega^2}{k_x^2} = \frac{\omega_{bv}^2 - \omega^2}{k_z^2} \tag{7.7}$$

$$\implies c^2 = \frac{\omega_{bv}^2 - \omega^2}{k_z^2} \tag{7.8}$$

Using the values for $\omega = 0.007Hz$, $k_z = 0.032$, and $\omega_{bv} = 0.0084Hz$ (i.e. wave frequency, vertical wavenumber, and BV frequency) for the modeled wave gives a phase speed of 144 m/s. This corresponds to a relatively fast gravity wave such as those seen in the Mars atmosphere by Hickey, Walterscheid, and Schubert (2011). They show that waves with similar phase speeds and periods in a Mars-like atmosphere typically induce cooling in the upper atmosphere greater than any heating produced at lower altitudes. This can be roughly seen in Figure 7.3 and Figure 7.4 bottom panels, where the temperature amplitude is largely negative above the exobase. Perhaps because the model is one dimensional, there does not appear to be a wavelength evident in the vertical direction,

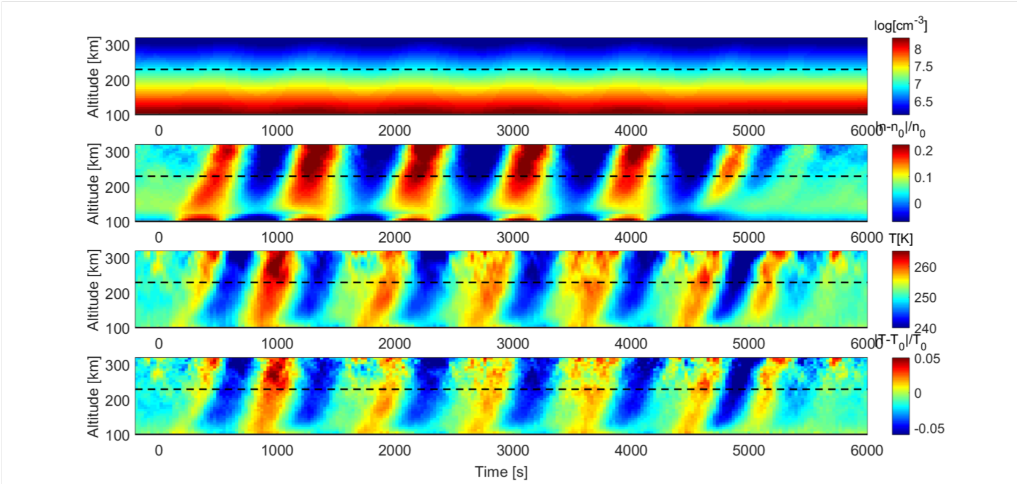


FIGURE 7.3: The O density and temperature evolution in a mixed atmosphere for a perturbation with frequency 0.007 Hz, corresponding to a gravity wave. From top to bottom: density in cm^{-3}; $(n - n_0)/n_0$; temperature in K; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes.

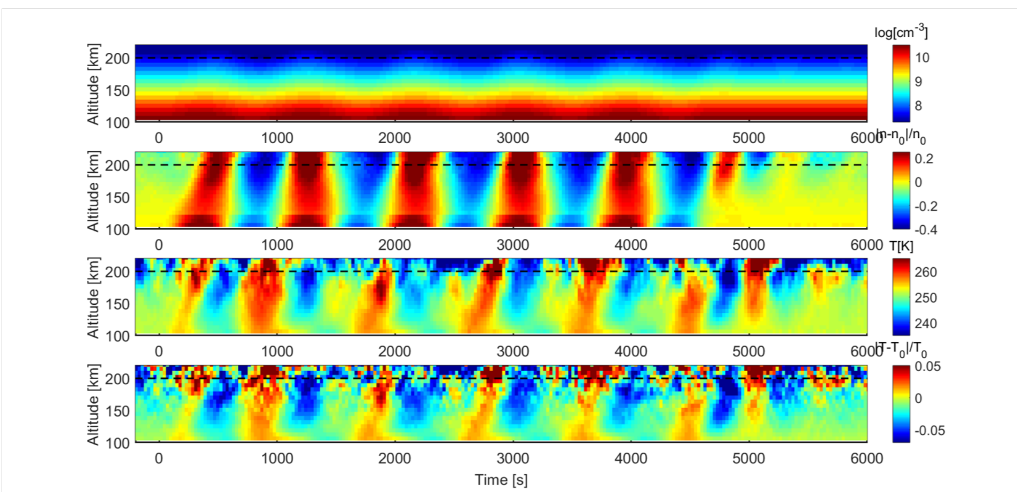


FIGURE 7.4: The CO_2 density and temperature evolution in a mixed atmosphere for a perturbation with frequency 0.007 Hz. From top to bottom: density in cm^{-3}; $(n - n_0)/n_0$; temperature in K; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes.

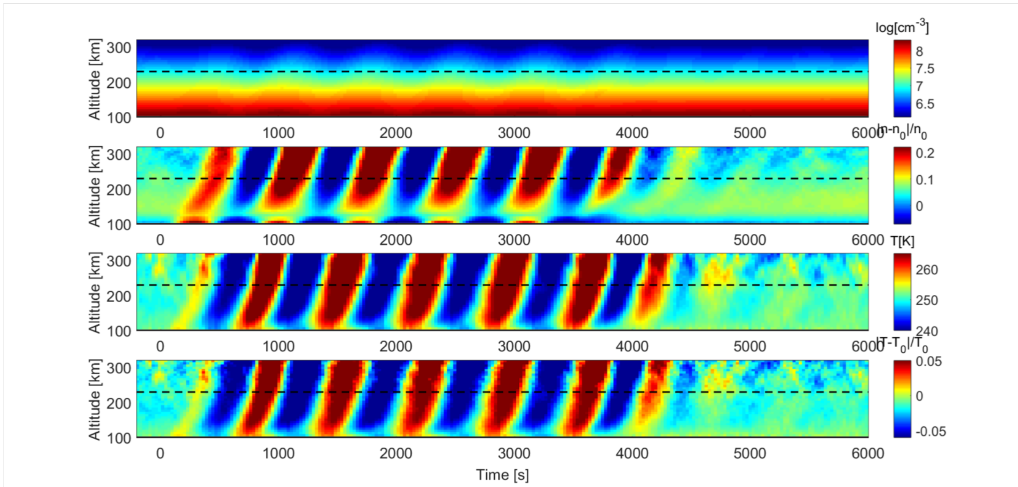


FIGURE 7.5: The O density and temperature evolution in a mixed atmosphere for a perturbation with frequency 0.009 Hz, corresponding to an evanescent wave. From top to bottom: density in cm^{-3}; $(n - n_0)/n_0$; temperature in K; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes.

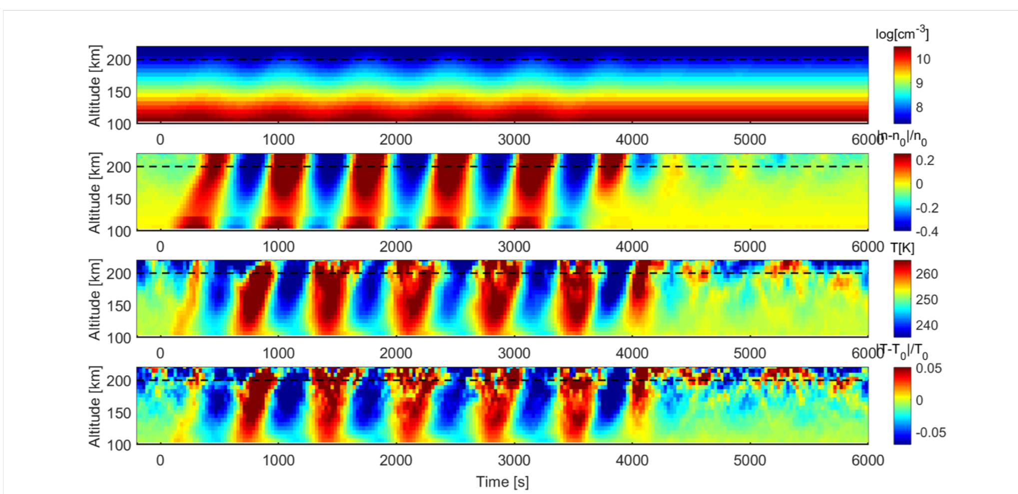


FIGURE 7.6: The CO_2 density and temperature evolution in a mixed atmosphere for a perturbation with frequency 0.009 Hz. From top to bottom: density in cm^{-3}; $(n - n_0)/n_0$; temperature in K; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes.

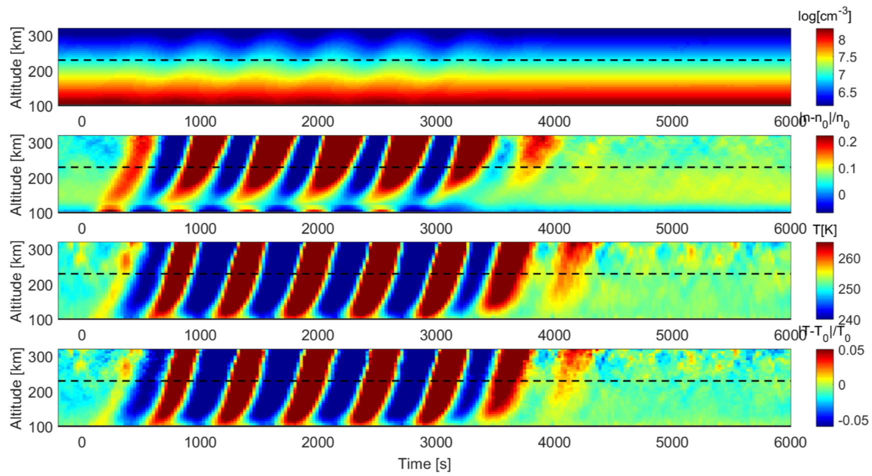


FIGURE 7.7: The O density and temperature evolution in a mixed atmosphere for a perturbation with frequency 0.011 Hz, corresponding to an acoustic wave. From top to bottom: density in cm^{-3}; $(n - n_0)/n_0$; temperature in K; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes.

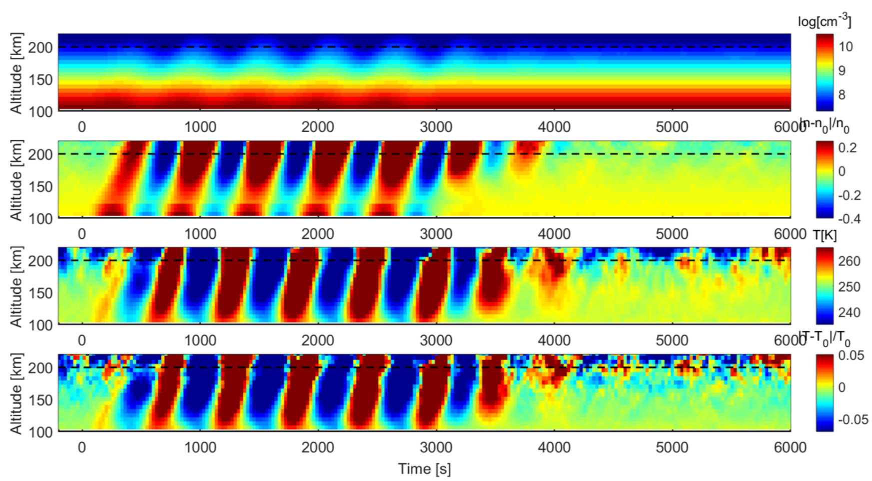


FIGURE 7.8: The CO_2 density and temperature evolution in a mixed atmosphere for a perturbation with frequency 0.011 Hz. From top to bottom: density in cm^{-3}; $(n - n_0)/n_0$; temperature in K; $(T - T_0)/T_0$. Black dotted lines indicate the nominal exobase altitudes.

making it difficult to compare to the NGIMS data.

Next, we model an evanescent wave, i.e. a frequency between the BV frequency and acoustic cutoff frequency. For an evanescent wave, wave theory predicts an imaginary vertical wavenumber m, seen as one possible solution to equation 2.51. As described in Chapter 2, this leads to a wave that does not propagate vertically but has an exponentially decaying amplitude in the vertical direction, although it may propagate horizontally. In the model, we introduce a sinusoidal pulse with a frequency of 0.009 Hz for 5 periods, corresponding to a period of $\sim$ 11.6 min. In Figures 7.5 and 7.6 after five periods evident in the amplitude maxima and minima, there are some lingering negative amplitudes for temperature and density above the exobase indicative of potential cooling. As with the case where $\omega < \omega_{bv}$ in Figures 7.3 and 7.4, there is not a good correspondence between the model results and the data as seen in Figure 7.2 or a discernible wavelength.

Finally, we model a wave with a frequency of 0.011 Hz, corresponding to a low frequency acoustic wave with a period of $\sim$ 9.5 min. This period is slightly longer than the 8 min period acoustic waves in Walterscheid, Hickey, and Schubert (2013), the longest acoustic period chosen in the paper; however, it is still above the acoustic cutoff frequency. Because this is an acoustic wave, the phase speed is the speed of sound in the atmosphere, which was calculated to be $\sim$ 236 m/s. Comparing these parameters to those seen in Walterscheid, Hickey, and Schubert (2013) gives the expectation of a peak in heating at approximately 250 km of roughly 5 K, which integrates to approximately 200 K/day. This is roughly consistent with Figures 7.7 and 7.8, where there is heating persisting above the exobase after the five periods has passed, for example at approximately 4000 s and 200-300 km for O and 150-200 km for CO_2 in the third and fourth panels of Figures 7.7 and 7.8, where the temperature amplitude is approximately 5%.

In examining the second panels of Figures 7.7 and 7.8, there is some suggestion of a vertical wavelength that is smaller than the boundaries of the model. However, this is still significantly longer than the calculated wavelength for orbit 5962 of approximately 31 km, found by taking the inverse of the Lomb-Scargle dominant wavenumber. Despite this, the acoustic wave frequency scenario best approximates the data in terms of wavelength, as the cases shown in Figure 7.3, 7.4, 7.5, and 7.6 have wavelengths that exceed the model boundaries. Additionally, while there is cooling at high altitudes for CO_2 at later times, there is no such cooling for O. This could be the source of the increased O/CO_2 ratio seen in Chapter 4, as the wave appears to be preferentially heating the O, increasing the diffusive separation between the two species. Therefore the increased ratio seen in the NGIMS data is likely driven by these acoustic waves.

To examine the heating and cooling present in the model, I calculate average temperatures at 220 km altitude for CO_2 and 300 km for O beginning after the model reached steady state to immediately before the perturbation is introduced, for one full period while the atmosphere is being perturbed (to include both an amplitude maximum and

TABLE 7.2: Temperature of O and CO_2 averaged as described in the text before, during, and after each perturbation type with the standard deviation for each value

Species	ω (Hz)	Before (K)	Stdev	During (K)	Stdev	After (K)	Stdev
O	0.007	255	0.94	249	8.8	255	4.1
CO_2	0.007	249	9.6	249	14	230	6.1
O	0.009	255	0.94	246	7.5	253	2.3
CO_2	0.009	249	9.6	235	11	243	8.4
O	0.011	255	0.94	252	4.9	262	1.4
CO_2	0.011	249	9.6	244	16	250	5.1

minimum), and for 200 s beginning at 1100 s after the perturbation pulse ends, to see if there is any lingering heating or cooling, shown in Table 7.2. I average the temperatures of the species at different altitudes due to the increased scale height of O over CO_2. At 300 km altitude, where the O temperature is averaged, the CO_2 density is low and so the model does not have good statistics, so the CO_2temperature average must be at a lower altitude closer to the CO_2 exobase. For all three frequencies, there is no heating of CO_2 after the perturbation, with cooling evident for the 0.007 Hz case. The acoustic wave frequency produces the most heating of O, around a few K, with the other frequencies having no ongoing effect on the O temperature. While this ΔT_O is small, only a few degrees, as stated previously this is consistent with other models such as Walterscheid, Hickey, and Schubert (2013). Additionally, the standard deviation for the mean O temperature after the perturbation is 1.4 K, so the 7 K change is statistically significant. There is significant cooling, around 20 K for CO_2, for the wave with the lowest frequency, which is consistent with what is shown in Hickey, Walterscheid, and Schubert (2011). Essentially, slower, longer wavelength gravity waves produces a downward-directed heat flux at high altitudes that dominates the upward heat flux at lower altitudes, leading to net cooling. The smallest change between both the O and CO_2 before and after values are for the evanescent wave, consistent with wave theory predictions that an evanescent wave does not transport energy in the same way as a propagating wave.

7.4 Summary

While the simulation results presented above are only for a single case study perturbation, they are suggestive that the exospheric perturbations in the NGIMS data are high altitude acoustic gravity waves, as the simulated acoustic wave is the only case that shows preferential heating of O, required to match the increased O/CO_2 ratio seen in the NGIMS data. Both Hickey, Walterscheid, and Schubert (2011) and Walterscheid, Hickey, and Schubert

(2013) show that waves with high phase speeds, such as acoustic waves, are able to propagate higher in the atmosphere and escape dissipation at lower altitudes. As such, these fast waves are more capable of depositing energy in the upper atmosphere, which likely correlates with the increased O/CO_2 ratio seen for orbits with an exospheric perturbation. Additionally, the acoustic wave case is the only one with a suggestion of a vertical wavelength, while for the other two cases the vertical wavelength clearly exceeds the boundaries of the model.

However, our model is limited and cannot give a full understanding of these perturbations. Because it is one-dimensional, it is not able to account for the horizontal transport of O above the exobase, which could be responsible for the dissipation of these perturbations in the exosphere. As a result, the perturbations in the model continue to increase in amplitude up to the maximum altitude of the model in concordance with gravity wave physics. As the highest increase in the O/CO_2 ratio is in the region where the perturbations dissipate (Chapter 4), it is clear that wave dissipation plays a crucial role in energy transport, supported by theory and previous models (Charney and Drazin, 1961; Hodges, 1967; Fritts and Dunkerton, 1984; Hickey, Walterscheid, and Schubert, 2011). So while Figures 7.7 and 7.8 show that these perturbations can raise the temperature of the atmosphere by a few percent, the total effect of the how these perturbations change the exosphere cannot be completely quantified without including dissipation.

Additionally, as mentioned above, these model results are based on a single exospheric perturbation; it is possible that other perturbations are gravity waves, not acoustic waves. Because the perturbations appear to be characterized by a predominant peak in the Lomb-Scargle spectral analysis, it seems evident, however, that these perturbations are in fact propagating and not evanescent waves. Additionally, the standard deviation of the perturbation wavelengths is low, indicating that despite the differing amplitudes of the perturbations, these waves have roughly similar wavelengths. This makes it unlikely that these perturbations differ wildly in frequency due to the dispersion relation between wavelength and frequency. For these reasons, despite the limited modeling, it would seem that the exospheric perturbations are consistent with acoustic waves, with frequencies greater than the acoustic cutoff frequency of the background atmosphere. Therefore, based on both our model and results seen in Walterscheid, Hickey, and Schubert (2013), these waves likely increase the temperature of the exosphere on the order of 5-10 K locally. Our two-species model shows no change in the temperature of CO_2 as the wave dissipates in time for an acoustic wave, which could explain why there is an increase in the O/CO_2 ratio in Chapter 4. These exospheric acoustic waves are of interest for further study, as both the NGIMS data and DSMC model show they can alter the composition and temperature of the exosphere.

Chapter 8

Conclusions and Future Work

Our first goal as stated in Chapter 1 is to determine if ion precipitation into the upper atmosphere had a measurable effect on global average neutral densities. I accomplished this goal in Chapter 3 by binning NGIMS observations of O, Ar, and CO_2 at an altitude of 180-220 km by latitude and longitude, then taking the mean density of each bin. By looking at these average densities in two separate coordinate systems, one of which is based on the direction of the solar wind convective electric field, I is able to examine the changes in neutral densities near the exobase as a function of solar insolation (MSO coordinates) and ion precipitation (MSE coordinates). When the mean densities are plotted in MSO coordinates, the average densities in MSO coordinates for all three species match what would be expected from GCMs such as that shown in Bougher et al. (2014) and Valeille et al. (2009). Generally, for Ar and CO_2 there is a significant decrease in density on the nightside of the planet, while the density gradient across the terminator for O is much lower due to ballistic transport.

The average densities are then displayed in MSE coordinates, where ion precipitation is more likely in the negative latitudes or southern hemisphere. Doing so makes it possible to see if ion precipitation is affecting the neutral atmosphere by comparing the average densities of the positive and negative latitudes. There is a small increase in density in the negative latitudes, potentially indicating heating due to the ion precipitation, but this increase is within the standard deviation of the average density per hemisphere and so not statistically significant. However, it is suggestive that ion precipitation might have a lasting continuous effect on the neutral upper atmosphere, even if it is small, which is what would be expected in the present epoch. However, current methods of estimating the solar wind velocity and magnetic field while MAVEN has its apoapsis in the Martian magnetotail are ongoing; having such a solar wind proxy measurement would allow for the calculation of MSE coordinates of densities measured in the subsolar region. When MAVEN has its periapsis on the dayside, it does not directly measure the solar wind and so MSE coordinates for those orbits have not been calculated in this work. Being able to do so would extend the coverage of the MSE average density map and perhaps better show the effect of the ion precipitation, making this a possible avenue for future work.

For the second goal, there are multiple objectives, including setting criteria for defining an "exospheric perturbation" in the NGIMS dataset and examining all of the data to find examples of these perturbations. The criteria chosen are outlined in Chapter 4. The primary criterion, the amplitude, is found by creating a smooth background fit to the data by using a high order polynomial, then finding the relative difference between the fit and the data. The O amplitudes are roughly half that of Ar and CO_2, consistent with gravity wave theory, which predicts that wave amplitude is inversely proportional to scale height. As O has a larger scale height than Ar or CO_2, it subsequently has a smaller amplitude. When categorized by column density rather than altitude these perturbations dissipate at a consistent column density roughly equivalent to that where O ballistic transport becomes a dominant exospheric process, approximately 2×10^{18} m^{-2}. The dissipation at this column density indicates that as these exospheric perturbations propagate upwards they are likely dissipated by ballistic transport, rather than the processes such as eddy turbulence responsible for wave dissipation at higher densities.

When organizing the data by column density to remove day/night temperature differences, orbits with an exospheric perturbation show an increase in the O/CO_2 ratio, which in the absence of temperature data is interpreted as heating due to wave dissipation. This heating would increase the rate at which the species diffusively separate, leading to the observed ratio change. However, this alone did not provide a full understanding of these perturbations, so I also studied the distribution of both perturbation occurrence and amplitude through the upper atmosphere. This goal, shown in Chapter 6, is accomplished by plotting the ratio of perturbation orbits to total orbits and the perturbation species amplitudes as a function of multiple variables, including local time, latitude and longitude in MSO and MSE coordinates, and solar wind velocity and magnetic field. Again, there is not any evidence that the solar wind had a significant effect on the occurrence or the amplitude of the perturbation, but there are distinct distributions in local time and latitude/longitude. Perturbations are both more likely and had larger amplitudes on the nightside, much like the inverse relation of amplitude to background temperature seen for lower altitude gravity waves. Additionally, there are larger amplitudes in positive MSE latitudes, where ion escape rather than ion precipitation would be expected. This could potentially indicate that ion precipitation damps exospheric perturbations, perhaps by adding enough heat that the perturbations dissipate before reaching the exobase.

I finally characterized the perturbation data by using a Lomb-Scargle periodogram to retrieve the wavenumber of each perturbation. Our spectral analysis indicates that these perturbations are most likely a single wave train due to the strong dominance of a single vertical wavenumber. In addition to our data analysis, I also used the DSMC model developed by Ludivine Leclercq and implemented by Lucia Tan to simulate perturbations of similar amplitude. This model, unlike the data, can give a direct temperature measurement, which Chapter 5 shows differs drastically from a temperature profile calculated using the hydrostatic equation for a non-steady state atmosphere. Because it is not possible to calculate the temperature profile for our exospheric perturbations, the model creates a pulse that propagates vertically through the 1D atmosphere and the resulting temperature profile can be analyzed.

The frequency of this pulse can be varied in the model, so this work shows results for frequencies typical of a gravity wave, an evanescent wave, and an acoustic wave, chosen by calculating the BV and acoustic cutoff frequencies for a sample exospheric perturbation. Using the dispersion relation from Chapter 2 and the vertical wavenumber from the data, a phase speed is calculated for the gravity and acoustic wave cases. When the three different frequency waves are compared to the data, the exospheric perturbations are most consistent with acoustic waves, which are able to propagate to high altitudes due to their fast phase speed and high frequency, which are less likely to become saturated due to viscous forces. The model also shows deposition of heat high in the atmosphere, with a preference towards heating O, accounting for the increased O/CO_2 ratio in orbits with a perturbation compared to those without. In the future, we aim to simulate perturbations with a wider range of frequencies, to see if the model can better approximate the short wavelengths seen in the NGIMS data.

This work leads us to conclude that the large exospheric perturbations seen in NGIMS data are consistent in appearance and behavior with large amplitude acoustic gravity waves. While the typical fluid equations are not valid in the nearly collisionless regime of the exosphere, there is still amplitude growth with altitude, as well as scale height dependent amplitude, perturbations occurring in all local times and latitudes, and a single dominant wavenumber, all of which are consistent with acoustic gravity waves. These waves then dissipate above the exobase, a process known to be an important source of energy for the upper atmosphere of Earth. Due to the one dimensional limitations of both the data and the model, exact calculation of the energy transfer due to wave dissipation is not possible at present, but based on model results and comparisons to fluid models such as Walterscheid, Hickey, and Schubert (2013) it is likely that these perturbations can add on the order of 5-10 K to the atmosphere. While this is likely not enough to significantly increase the Jeans escape, the data clearly show that it is enough to alter the atmospheric composition, perhaps indicating that the dissipation of the waves adds an additional, larger amount of energy to the exosphere that cannot yet be accounted for. In order to better quantify the effect of dissipation, two dimensional modeling is being developed and implemented which will lead to a better understanding of all perturbation frequencies beyond one case study.

Appendix A

Gravity Waves: A One-Dimensional Derivation

A.1 Introduction

While atmospheres generally follow predictable trends, they are also inherently non-linear systems, initiated by multiple types of possible perturbations and disturbances. One of the most common types of disturbance is a wave feature known as a gravity wave, which is a wave generated in a fluid with gravity or buoyancy as the restoring force. Gravity waves transfer momentum throughout the atmosphere, changing the local temperature and density, and are seen in the atmospheres of many planets. A common example of a non-atmospheric gravity wave is surface ocean waves, which are generated by the interface between winds and the ocean surface, with water buoyancy as the restoring force.

Gravity waves typically fall into one of three categories, determined by their cause (Green, 1999):

- Buoyancy waves caused by convective instabilities
- Topographic waves caused by air flow over topographic changes, and
- Dynamic instability waves

Their appearance and properties are largely determined by the local boundary conditions as well as general atmospheric dynamics, which we will examine.

A.2 One dimensional atmospheric dynamics

To understand the atmospheric phenomena of gravity waves, we must first understand the underlying physics of an atmosphere. In this section, we will derive the various governing equations for a one dimensional atmosphere. For all of these equations, we begin by examining a parcel of air.

A.2.1 State and hydrostatic equations

We assume that our atmosphere is a perfect or ideal gas. As such, the ideal gas law applies, which states that

$$p = \rho RT = nk_bT \tag{A.1}$$

with p being pressure, ρ mass density, R the ideal gas constant, T the temperature, n the number density, and k_B the Boltzmann constant (Brown, 1991).

We can arrive at the 1D hydrostatic equation by considering the force balance on a stationary parcel of air (a moving parcel of air will be addressed later). For this parcel to be stationary, the upward directed force must be balanced by the downward directed force. The upward directed force will be the upward pressure, while the downward directed force will be the downward pressure plus the weight of the parcel, the mass times the gravitational acceleration. Let us define the weight as

$$mg = (\rho V)g = \rho A\ \delta z\ g \tag{A.2}$$

where A is the surface area of the parcel and δz is the change in height of the parcel, ρ the mass density, V the volume, m the mass of the air, and g the gravitational acceleration. Then our force balance equation setting the upward force equal to the downward force is

$$F_{up} = p_{up}A = F_{down} = \rho A\ \delta z\ g + p_{down} \tag{A.3}$$

where F_{up} indicates the upward directed force and F_{down} indicates the doward directed force found above. Let Δp indicate the change in pressure across δz, $p_{down} - p_{up}$. Then,

$$\Delta pA = -\rho A\delta zg \tag{A.4}$$

As $\delta z \to 0$, we can cancel the A on both sides and divide by δz to get

$$\to \boxed{\frac{dp}{dz} = -\rho g} \tag{A.5}$$

is the hydrostatic equation. We can integrate this to get pressure as a function of altitude, using the ideal gas law as stated above to set

$$\rho = \frac{pm}{k_bT} \tag{A.6}$$

with m as the molecular mass. Substituting this into our differential equation, we get

$$\frac{dp}{dz} = -\frac{pmg}{k_b T} \tag{A.7}$$

$$\frac{dp}{p} = -\frac{mgdz}{k_b T} \tag{A.8}$$

$$\int \frac{1}{p} dp = \int -\frac{mg}{k_b T} dz \tag{A.9}$$

which gives the solution

$$p(z) = p_0 e^{-\frac{mg}{k_b T} z} \tag{A.10}$$

where $p(z)$ is the pressure as a function of altitude and p_0 is the pressure at a chosen altitude (often the surface). One typically defines a new variable, the scale height, to be $H = \frac{k_b T}{mg}$, so that our solution becomes (Brown, 1991):

$$p(z) = p_0 e^{-z/H} \tag{A.11}$$

A.2.2 The mass continuity equation

Imagine, as stated above, a parcel of air. We know that flow into the parcel must equal flow out of the parcel, as our parcel is not changing in mass. So we define outward flow as

$$\vec{\boldsymbol{u}} \cdot \hat{n} \tag{A.12}$$

where $\vec{\boldsymbol{u}}$ is the flow velocity vector and $\hat{n}$ is the vector normal to the surface of the parcel. We will begin our derivation with a general number of dimensions, then show how it applies to one dimension. With this expression for outward flow, our equation of flow in equals flow out becomes

$$\int\int\int \frac{\partial \rho}{\partial t}\, dV = -\int\int \rho\, \vec{\boldsymbol{u}} \cdot \hat{n}\, dA \tag{A.13}$$

where $\rho\, \vec{\boldsymbol{u}} \cdot \hat{n}$ is the flow out of the parcel times the mass density. We can use the divergence theorem on the right hand side:

$$\int\int\int \frac{\partial \rho}{\partial t}\, dV = -\int\int \bigtriangledown \rho \vec{\boldsymbol{u}}\, dV \tag{A.14}$$

$$\implies \int\int\int \left(\frac{\partial \rho}{\partial t} + \bigtriangledown \rho \vec{\boldsymbol{u}} \right) dV = 0 \tag{A.15}$$

$$\implies \frac{\partial \rho}{\partial t} + \bigtriangledown \rho \vec{\boldsymbol{u}} = 0 \tag{A.16}$$

Now we change to one dimension, the z dimension, to look at how the atmosphere changes with altitude. This gives us for our differential equation:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u_z)}{\partial z} = 0 \tag{A.17}$$

Expanding the second derivative,

$$\frac{\partial \rho}{\partial t} + u_z \frac{\partial \rho}{\partial z} + \rho \frac{\partial u_z}{\partial z} = 0 \tag{A.18}$$

Given the definition of total derivative as $Da/Dt = \partial a/\partial t + (\vec{\boldsymbol{u}} \cdot \bigtriangledown)a$ or in one dimension $Da/Dt = \partial a/\partial t + u_z \partial a/\partial z$, we have for a one dimensional vertical atmosphere our mass continuity equation (Brown, 1991),

$$\boxed{\frac{D\rho}{Dt} + \rho \frac{\partial u_z}{\partial z} = 0} \tag{A.19}$$

A.2.3 Conservation of momentum equation

For our parcel of air, we know momentum must be conserved. From Newton's Second Law, we also know that the change in momentum where momentum is equal to mass times flow velocity is equal to the sum of the forces on the parcel, i.e.

$$\frac{D(m\vec{\boldsymbol{u}})}{Dt} = \sum \vec{\boldsymbol{F}} \tag{A.20}$$

For our fluid parcel with uniform density we can write the mass as mass per unit volume, i.e. density, so the left hand side becomes

$$\frac{D(m\vec{\boldsymbol{u}})}{Dt} = \frac{D(\rho\vec{\boldsymbol{u}})}{Dt} = \frac{\partial(\rho\vec{\boldsymbol{u}})}{\partial t} + (\vec{\boldsymbol{u}} \cdot \bigtriangledown)(\rho\vec{\boldsymbol{u}}) \tag{A.21}$$

For one dimension, $\bigtriangledown = \frac{\partial}{\partial z}$ and $\vec{\boldsymbol{u}} = u_z$. So the above becomes

$$\frac{D(\rho\vec{\boldsymbol{u}})}{Dt} = u_z \frac{\partial \rho}{\partial t} + \rho \frac{\partial u_z}{\partial t} + \rho u_z \frac{\partial u_z}{\partial z} + u_z \frac{\partial(\rho u_z)}{\partial z} \tag{A.22}$$

However, from equation A.17 above, we know that $\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u_z)}{\partial z} = 0$, so we can cancel those terms, leaving us with

$$\frac{D(\rho\vec{\boldsymbol{u}})}{Dt} = \rho \frac{\partial u_z}{\partial t} + \rho u_z \frac{\partial u_z}{\partial z} = \rho \left(\frac{\partial u_z}{\partial t} + u_z \frac{\partial u_z}{\partial z} \right) \tag{A.23}$$

Now we integrate over our control volume, in this case in the z dimension, to get (per unit volume)

$$\int \rho \left(\frac{\partial u_z}{\partial t} + u_z \frac{\partial u_z}{\partial z} \right) dz = \sum F_z \tag{A.24}$$

So now we need to find the right hand side to complete the equation. For a parcel of air, there will be both internal body forces and surface forces, i.e.

$$\sum \vec{F} = \rho \vec{F}_B + \vec{F}_S \tag{A.25}$$

Because our parcel is uniform, the body forces act uniformly on each element; hence it can be represented by a force per unit volume acting on the center of the parcel in the direction of the vector $\vec{F}_b$. We include the density because in the vertical dimension, the only body force will be the weight force per unit mass acted upon the parcel by gravity, so $\rho \vec{F}_B = -\rho g$. Therefore it remains to find the surface forces (again in three dimensions first, then generalizing to one). We define a stress tensor σ_{ij} such that

$$\vec{F}_s = \int \int \sigma_{ij} \cdot \hat{n} \, dA \tag{A.26}$$

By the divergence theorem,

$$\implies \vec{F}_s = \int \int \int (\bigtriangledown \cdot \sigma_{ij}) \, dA \tag{A.27}$$

or, in one dimension,

$$\vec{F}_s = \frac{\partial \sigma_{ij}}{\partial z} \tag{A.28}$$

where σ_{ij} is our stress tensor. Now we combine all of our terms to get

$$\int \left(\rho \left(\frac{\partial u_z}{\partial t} + u_z \frac{\partial u_z}{\partial z} \right) + \rho g - \frac{\partial \sigma_{ij}}{\partial z} \right) \, dz = 0 \tag{A.29}$$

Take the limit as $\delta z \to 0$ such that

$$\rho \frac{Du_z}{Dt} = -\rho g + \frac{\partial \sigma_{ij}}{\partial z} \tag{A.30}$$

We assume that the atmosphere is a Newtonian fluid. For a Newtonian fluid, p is pressure, the stress tensor σ_{ij} is defined by

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij} \tag{A.31}$$

where $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$, and τ_{ij} is the tensor force due to viscous transfer of momentum. The derivative of τ_{ij} in the z dimension is defined as

$$\frac{\partial \tau_{ij}}{\partial z} = \mu \frac{\partial^2 u_z}{\partial z^2} \tag{A.32}$$

where μ is the dynamic coefficient of viscosity.

Therefore, for our one-dimensional case, $i = j = 3$ and so

$$\frac{\partial \sigma}{\partial z} = -\frac{\partial p}{\partial z} + \mu \frac{\partial^2 u_z}{\partial z^2} \tag{A.33}$$

Substituting this into our equation, we get (Brown, 1991):

$$\boxed{\frac{\partial p}{\partial z} = -\rho g - \rho \frac{D u_z}{Dt} + \mu \frac{\partial^2 u_z}{\partial z^2}} \tag{A.34}$$

Note here that if u is 0 or constant we arrive at the familiar 1D hydrostatic equation $\frac{\partial p}{\partial z} = -\rho g$. Additionally, since we are focusing on the z dimension, we may ignore Coriolis forces.

A.2.4 Conservation of energy equation

From the First Law of Thermodynamics, we know that for our parcel of air, the change in internal energy is equal to the work done on the parcel plus the total change in heat, i.e.

$$\delta E/\delta t = \delta W/\delta t + \delta Q/\delta t \tag{A.35}$$

where W is the work and Q is the total heat. Let us expand each term. The total energy is equal to the kinetic energy plus the internal potential energy. Let us denote the internal energy as ρe and the kinetic energy as $\frac{1}{2}\rho u^2$. Then

$$\delta E/\delta t = \int\int\int \frac{\partial}{\partial t}\left(\rho(e + 1/2\, u^2)\right)\, dV + \int\int \rho(e + 1/2\, u^2)\vec{\boldsymbol{u}} \cdot \hat{n}\, dA \tag{A.36}$$

where the first term is the internal change in energy and the second is the change in energy due to flow in or out. Now we progress to the work term. We know the rate of work $\delta W/\delta t$ is force times velocity, so we get

$$\delta W/\delta t = -\int\int\int \vec{\boldsymbol{F}}_b \cdot \vec{\boldsymbol{u}}\, dV - \int\int \vec{\boldsymbol{u}} \cdot (\sigma \hat{n})\, dA \tag{A.37}$$

where again, σ is our stress tensor. Finally, we write out the change in heat, using R as a source/sink parameter and $\vec{\boldsymbol{K}}$ as the heat conduction.

$$\delta Q/\delta t = \int\int\int \rho R\, dV - \int\int \vec{\boldsymbol{K}} \cdot \hat{n}\, dA \tag{A.38}$$

Now let us write out the full equation in three dimensions, using the divergence theorem, where again e represents internal kinetic energy and $\frac{1}{2}u^2$ represents internal potential energy.

$$\int\int\int \frac{\partial}{\partial t}\left(\rho(e+1/2\,u^2)\right)+\bigtriangledown\cdot\left(\rho(e+1/2\,u^2)\right)\,dV \\ =\int\int\int (\vec{\boldsymbol{F}}_b\cdot\vec{\boldsymbol{u}})+(\bigtriangledown\cdot\sigma\vec{\boldsymbol{u}})+(\rho R)-(\bigtriangledown\cdot\vec{\boldsymbol{K}})\,dV \tag{A.39}$$

Take the limit as $\delta V\to 0$.

$$\frac{\partial}{\partial t}\left(\rho(e+1/2\,u^2)\right)+\bigtriangledown\cdot\left(\rho(e+1/2\,u^2)\right)=(\vec{\boldsymbol{F}}_b\cdot\vec{\boldsymbol{u}})+(\bigtriangledown\cdot\sigma\vec{\boldsymbol{u}})+(\rho R)-(\bigtriangledown\cdot\vec{\boldsymbol{K}}) \tag{A.40}$$

We will deal with the left hand side first.

$$\rho\frac{\partial}{\partial t}\left(e+1/2\,u^2\right)+\left(e+1/2\,u^2\right)\frac{\partial\rho}{\partial t}+\rho\left(\vec{\boldsymbol{u}}\cdot\bigtriangledown\left(e+1/2\,u^2\right)\right)+\left(e+1/2\,u^2\right)(\bigtriangledown\cdot\rho\vec{\boldsymbol{u}}) \\ =\rho\left(\frac{\partial}{\partial t}\left(e+1/2\,u^2\right)+\vec{\boldsymbol{u}}\cdot\bigtriangledown\left(e+1/2\,u^2\right)\right)+\left(e+1/2\,u^2\right)\left(\frac{\partial\rho}{\partial t}+\bigtriangledown\cdot\rho\vec{\boldsymbol{u}}\right) \tag{A.41}$$

or, in one dimension,

$$\rho\left(\frac{\partial}{\partial t}\left(e+1/2\,u^2\right)+u_z\frac{\partial}{\partial z}\left(e+1/2\,u^2\right)\right)+\left(e+1/2\,u^2\right)\left(\frac{\partial\rho}{\partial t}+\frac{\partial}{\partial z}\rho u_z\right) \tag{A.42}$$

But from A.17 we know the second term is zero. So we can rewrite the left hand side as

$$=\rho\left(\frac{\partial}{\partial t}+u_z\frac{\partial}{\partial z}\right)\left(e+1/2\,u^2\right)=\rho\frac{D}{Dt}\left(e+1/2\,u^2\right) \tag{A.43}$$

So then, changing to the z dimension only and using $F_b=\rho g$, our energy equation is (Brown, 1991):

$$\boxed{\rho\frac{D}{Dt}\left(e+1/2\,u^2\right)=\rho g u_z+\frac{\partial}{\partial z}(\sigma_z u_z)+\rho R-\frac{\partial K}{\partial z}} \tag{A.44}$$

A.2.5 Adiabatic lapse rate

Now let us look at the case of an adiabatic dry atmosphere, so $\delta Q/\delta t=0$ and the viscosity is low enough for us to consider the atmosphere incompressible, therefore all the work is done by pressure (remembering that $\sigma=-p+\tau$). For this stationary parcel, we define the heat capacity to be

$$c_v=\frac{dE}{dT}\to dE=c_v\,dT \tag{A.45}$$

where E is the energy and T is the temperature. Then our initial equation becomes

$$0 = c_v \frac{dT}{dt} + p \frac{dV}{dt} \tag{A.46}$$

Canceling the dt, we can write for a dry, adiabatic atmosphere,

$$c_v \, dT = -p \, dV \tag{A.47}$$

From the Ideal Gas Law,

$$p = \rho R T \tag{A.48}$$

$$\implies pV = \rho V R T = mRT \tag{A.49}$$

where p is the pressure, V is volume, ρ is mass density, R is the molar gas constant, T is temperature, and m is mass, since density times volume equals mass. Taking the derivative of both sides gives the differential equation

$$p \, dV + V \, dp = mR \, dT \tag{A.50}$$

$$\implies -p \, dV = \frac{V \, dp}{m} - R \, dT \tag{A.51}$$

$$\implies -p \, dV = \frac{dp}{\rho} - R \, dT = c_v \, dT \tag{A.52}$$

$$\implies \frac{dp}{\rho} = (c_v + R) \, dT \tag{A.53}$$

Let us define $c_v + R = c_p$. Then we take the z-derivative of both sides.

$$\frac{1}{\rho} \frac{dp}{dz} = c_p \frac{dT}{dz} \tag{A.54}$$

From equation A.5 we know $\frac{dp}{dz} = -\rho g$ so

$$\boxed{\frac{dT}{dz} = -\frac{g}{c_p} = -\Gamma_{ad}} \tag{A.55}$$

which we define to be the dry adiabatic lapse rate. Solving this equation for T we get

$$T = T_0 - \Gamma_{ad} z \tag{A.56}$$

So the temperature in an adiabatic atmosphere falls off linearly with altitude (Green, 1999).

Let us take the equation $\frac{dp}{\rho} = c_p\, dT$ and divide by T. This gives

$$c_p dT/T = dp/(\rho T) = R dp/p \tag{A.57}$$

where p is pressure. If we integrate this from the pressure at the surface p_s to some p and a "potential temperature" Θ to T we get an equation for potential temperature:

$$\Theta = T\,(p_s/p)^{R/c_p} \tag{A.58}$$

i.e.

$$c_p d\Theta/\Theta = c_p dT/T - R dp/p \tag{A.59}$$

and from equation A.55 we can write

$$\frac{dT}{dz} + \Gamma_{ad} = T/\Theta \frac{d\Theta}{dz} \tag{A.60}$$

This means that we can write δQ from the first law of thermodynamics as

$$\delta Q = c_p\; T/\Theta\; d\Theta \tag{A.61}$$

The definition of entropy is $ds = \delta Q/T$ so we get

$$ds = c_p\; d\Theta/\Theta \tag{A.62}$$

which means that lines of constant potential temperature are equivalent to lines of constant entropy (Brown, 1991). This will be important later when we look at the physics of gravity waves.

A.3 Gravity Wave Physics

Now that we have examined the physics of a one-dimensional atmosphere, we can apply this to gravity waves. First, however, we will go over some wave equations and parameters.

A.3.1 Wave Parameters

Let us first begin by defining some characteristics of a wave. Since gravity waves propagate by their nature in more than one direction, the following equations will be in two dimensions, x and z, where z is altitude.

First, we define wavelength λ as the distance between two subsequent peaks or troughs of a wave. We can then define the more useful wavenumber as $k = 2\pi/\lambda$. However, waves can have wavelengths and hence wavenumbers in all three (or for the purpose of this paper, two) dimensions, so we will define the vector wavenumber as

$$\vec{k} = k_x \hat{x} + k_z \hat{z} \tag{A.63}$$

We can also define the wave period τ as the time it takes for a wave to oscillate once, and hence the frequency as $\omega = 2\pi/\tau$.

Now that we have defined the wavenumber and frequency, we can define a sinusoidal wave with the equation $A\ \cos(k_x x - \omega t)$ where A is the amplitude of the wave, i.e. the height from a trough to a peak. Then we define the angle described by $k_x x - \omega t$ as the phase angle ϕ, or, in more than one dimension, $\phi = \vec{k} \cdot \vec{r} - \omega t = k_x x + k_z z - \omega t$. One of the ways we can measure the speed of wave propagation is to find the phase speed, i.e. the speed with which a point of constant phase moves.

$$\frac{\partial \phi}{\partial t}|_\phi = \vec{k} \cdot \frac{d\vec{r}}{dt} - \omega = 0 \tag{A.64}$$

$$\implies \frac{d\vec{r}}{dt} = \frac{\omega}{\vec{k}} \tag{A.65}$$

is the phase speed. We can also think of a group of wave fronts, or a wave packet, propagating in a direction. The speed and direction of the wave packet determines the rate of energy propagation from the packet, so it is important to determine this as well.

For a wave packet, the rate of change of ϕ with time is the frequency ω, but ϕ also changes with x, since this is a packet rather than a single wavefront. So we have

$$\frac{\partial \phi}{\partial t} = \omega, \ \frac{\partial \phi}{\partial x} = -k_x \tag{A.66}$$

$$\implies \partial \phi = \omega \partial t = -k_x \partial x \tag{A.67}$$

$$\implies \frac{\partial \omega}{\partial x} + \frac{\partial k_x}{\partial t} = 0 \tag{A.68}$$

Let us treat ω as a function of k_x. Then

$$\frac{\partial k_x}{\partial t} + \frac{d\omega}{dk_x}\frac{\partial k_x}{\partial x} = 0 \tag{A.69}$$

So we will define the group velocity as $\frac{d\omega}{dk_x}$, giving us

$$\frac{\partial k_x}{\partial t} + u_g \frac{\partial k_x}{\partial x} = 0 \tag{A.70}$$

$$\implies \frac{Dk_x}{dt} = 0 \tag{A.71}$$

So the wavenumber is constant at the group velocity (Nappo, 2013).

A.3.2 The Brunt-Väisälä Frequency

We already know from equation A.55 that $\frac{dT}{dz} = -\Gamma_{ad}$ where T is temperature, z is altitude, and Γ_{ad} is the adiabatic lapse rate and that the potential temperature $\Theta = T\,(p_s/p)^{R/c_p}$. From the definition of logarithmic differentiation, we know that $f'/f = (ln(f))'$, so let us do this to find the derivative of our potential temperature function.

$$ln(\Theta) = ln(T) - \frac{R}{c_p}ln(P) + \frac{R}{c_p}lnP_s \tag{A.72}$$

$$\implies d(ln(\Theta))/dz = \frac{1}{T}\frac{\partial T}{\partial z} - \frac{R}{c_p}\frac{1}{P}\frac{\partial P}{\partial z} \tag{A.73}$$

$$\implies \frac{1}{\Theta}\frac{\partial \Theta}{\partial z} = \frac{1}{T}\frac{\partial T}{\partial z} - \frac{R}{c_p}\frac{1}{P}\frac{\partial P}{\partial z} \tag{A.74}$$

We can then use the hydrostatic equation and the ideal gas law to replace $P = \rho RT$ and $\frac{\partial P}{\partial z} = -\rho g$, giving the result

$$\frac{1}{\Theta}\frac{\partial \Theta}{\partial z} = \frac{1}{T}\left(\frac{\partial T}{\partial z} + \frac{g}{c_p}\right) \tag{A.75}$$

If we let γ indicate the temperature gradient and use the definition of the adiabatic lapse rate, we get

$$\frac{1}{\Theta}\frac{\partial \Theta}{\partial z} = \frac{\Gamma_{ad} - \gamma}{T} \tag{A.76}$$

Now, say that we want to find the acceleration of our air parcel in an adiabatic regime due to gravity or the buoyant force, which is the true restoring force for a gravity wave. The buoyant force is equal to gravity times the mass of the air displaced by the parcel minus the mass of the air parcel, so by Newton's Second Law, for movement of a distance δz, we have

$$m_p\frac{d^2(\delta z)}{dt^2} = -g(m_p - m_a) \tag{A.77}$$

where m_p is the mass of the parcel and m_a is the mass of the displaced air (we have reversed the sign from the description above). If we assume that the volume of the parcel is equal to the volume of the displaced air, we can then change our equation to look at mass per unit volume, or density. We can then use the ideal gas law as above to replace density with temperature:

$$\frac{d^2(\delta z)}{dt^2} = -g\,\frac{\rho_p - \rho_a}{\rho_p} = -g\,\frac{T_a - T_p}{T_a} \tag{A.78}$$

Across the height δz, we can expand the temperatures to first order, giving $T(z+\delta z) = T_0 + \frac{\partial T}{\partial z}\delta z$. Canceling the constant T_0 that will appear in both since $T_0 >> frac\partial T\partial z\delta z$, we get:

$$\frac{d^2(\delta z)}{dt^2} = -\frac{g}{T_a}\left(\frac{\partial T_a}{\partial z} - \frac{\partial T_p}{\partial z}\right) \tag{A.79}$$

However, as stated, the parcel is moving adiabatically, so we know $-\frac{\partial T_p}{\partial z} = \Gamma_{ad}$. As above, we set the $\frac{\partial T_a}{\partial z} = \gamma$, which gives us

$$\frac{d^2(\delta z)}{dt^2} = \frac{-g}{T}(\Gamma_{ad} - \gamma)\delta z \tag{A.80}$$

From our above equation for the derivative of the potential temperature, we can rewrite this as:

$$\frac{d^2(\delta z)}{dt^2} = \frac{-g}{\Theta}\frac{\partial \Theta}{\partial z}\delta z \tag{A.81}$$

As long as our potential temperature gradient is positive, this is an equation for simple harmonic motion of the form

$$\delta z(t) = Ae^{i\omega_{bv}t} + Be^{-i\omega_{bv}t} \tag{A.82}$$

where the frequency ω_{bv}, which is called the Brunt-Väisälä frequency, is

$$\boxed{\omega_{bv} = \sqrt{\frac{g}{\Theta}\frac{\partial \Theta}{\partial z}}} \tag{A.83}$$

(Nappo, 2013) or, since for an isothermal atmosphere (i.e. in the exosphere) the temperature gradient is zero,

$$\omega_{bv} = \sqrt{g\Gamma/T} \tag{A.84}$$

which are all easily known variables.

A.3.3 Dispersion Relation and Energy

In the previous subsection, we derived the Brunt-Väisälä frequency, which is the frequency with with a parcel of air with move up and down adiabatically. Thus, the buoyancy restoring force will be

$$\vec{F} = -\omega_{bv}^2 \delta s \tag{A.85}$$

where δs is along the path of the parcel.

Now imagine we have a wave field in an adiabatic regime. We have already defined phase in a wave packet; we know there will be lines of constant phase in the field that form some angle to the vertical. Let that angle be θ. Then the restoring force along these lines of constant phase will be

$$\vec{F} = -\omega_{bv}^2 \delta z \cos\theta \tag{A.86}$$

$$\delta z = \delta s \cos\theta \tag{A.87}$$

$$\implies \vec{F} = -\omega_{bv}^2 \cos^2\theta \delta s \tag{A.88}$$

This means our equation for the oscillation of the parcel of air takes the form

$$\frac{\partial^2 \delta s}{\partial t^2} = -\omega_{bv}^2 \cos^2\theta \delta s \tag{A.89}$$

which has the solution

$$\omega^2 = \omega_{bv}^2 \cos^2\theta \tag{A.90}$$

$$\implies \cos^2\theta = \frac{\omega^2}{\omega_{bv}^2} \tag{A.91}$$

where ω is the frequency of oscillation. We know the angle θ the lines of constant phase make to vertical must be related to the ratio of the horizontal and vertical wavenumbers k_x and k_z, given the definition of phase. So then

$$\tan\theta = \frac{k_z}{k_x} \tag{A.92}$$

$$\implies \tan^2\theta = \frac{k_z^2}{k_x^2} = \frac{1-\cos^2\theta}{\cos^2\theta} \tag{A.93}$$

$$\implies \frac{k_z^2}{k_x^2} = \frac{1-\frac{\omega^2}{\omega_{bv}^2}}{\frac{\omega^2}{\omega_{bv}^2}} \tag{A.94}$$

Solving for the frequency ω, we get

$$\boxed{\omega = \pm\frac{\omega_{bv} k_x}{(k_x^2 + k_z^2)^{1/2}}} \tag{A.95}$$

which is our dispersion relation in two dimensions (Linzen, 2008). To get the dispersion relation in the vertical dimension, let us assume $k_x << k_z$, i.e. the horizontal wavelength is much longer than the vertical wavelength. Then we can cancel k_x from the above equation, leaving us with

$$\omega = \pm\frac{\omega_{bv}}{k_z} \tag{A.96}$$

For energy, we can treat the air parcel of the wave as undergoing simple harmonic motion, thus having the same energy per unit volume. With the amplitude A analogous to the length of a pendulum and a pendulum velocity equal to length times its oscillation frequency ω, we can write the velocity as $A\omega$. Therefore, using density instead of mass to get kinetic energy per unit volume, we have:

$$KE = \frac{1}{2}\rho A^2\omega^2 \tag{A.97}$$

Alternatively, we can see that this applies by taking the form $y(t) = Ae^{i(k_x x - \omega t)}$ and taking the second derivative with respect to time, then averaging out the exponential, which will give us the same result.

A.4 Waves in the Upper Atmosphere

While it is outside the scope of this appendix, it is possible to use perturbation and instability theory to further examine energy transfer by a gravity wave throughout the atmosphere. This produces, using the assumption that density gradients can be ignored called the Boussinesq approximation, a version of the fluid dynamics equations known as the Taylor-Goldstein equations. In this section, we will qualitatively explain this, while leaving the mathematical derivations to more in-depth studies. While many studies over the last few decades, particularly on Earth, have looked at how exactly gravity waves deposit energy into the atmosphere, the source remains not completely understood; instead, there are several ways this is likely. However, it is known that this does happen, as the effects of gravity waves on the lower, middle, and upper atmosphere are seen in a variety of data, such as both *in situ* detections and remote measurements using radio sounding.

The most likely way gravity waves deposit energy is though breaking into turbulence–hence why instability theory is crucial to understanding energy transport and deposition. The turbulence from the wave breaking (analogous to an ocean wave) then induces a drag force on the local atmospheric movement. As the local atmosphere feels the effect of the drag force, energy is added to the area. This theory is supported by data showing that gravity waves and turbulence almost always coexist in an otherwise stable atmosphere. However, it is somewhat difficult to define what exactly constitutes wave instability that will then lead to energy deposition, as wave instability and turbulence are not exactly the same thing. Additionally, turbulence does not have a clear mathematical definition either, due to it being a regime where chaotic motion is dominant. Instead, turbulence is thought of as something qualitatively identifiable, and indeed remains a large unsolved problem in fluid mechanics. This adds significantly to the difficulty of discovering how gravity waves break and potentially change the energy of the surrounding atmosphere.

One way to define wave instability is to examine a system with two vertically arranged parcels of fluid. These parcels then have their positions exchanged. If the total energy of the system increases after this exchange takes place, the system is stable, as work has been done on it. However, if the total energy of the system decreases, the system has done the work, and therefore it is unstable. This serves as the basis of stability analysis and can determine how instabilities develop through time. One important parameter in stability analysis is known as the Richardson number, given as

$$Ri = \frac{\omega_{bv}^2}{\left(\frac{\partial u}{\partial z}\right)^2} \tag{A.98}$$

representing the ratio between the turbulence kinetic energy produced by buoyancy and that produced by shear, where ω_{bv} is the BV frequency and u is the flow velocity. The important thing to note here is that a flow is stable if its Richardson number is greater than $1/4$. If it is negative, a convective instability occurs, whereas if it is between zero and $1/4$, a dynamic instability occurs. In a convective instability, convective vertical motion will occur. A dynamic instability includes all other well-known fluid instabilities, such as Kelvin-Helmholtz and Rayleigh-Taylor. Using perturbation theory's definition of $\frac{\partial u}{\partial z}$ for a vertical linear wave perturbation, we can achieve the result that the Richardson number actually goes as $1/\omega_{bv}$ (Nappo, 2013), where ω_{bv} is the Brunt-Väisälä frequency. So in regions where ω_{bv} decreases with height, upward propagating gravity waves become more stable. Because ω_{bv} is dependent on the potential temperature gradient, this occurs in regions like the stratosphere or in the upper thermosphere and exosphere, indicating that gravity waves can stably propagate at high altitudes. In fact, this indicates that in the upper thermosphere (roughly >120 km on Earth), upward propagating gravity waves will be more stable than in lower atmospheric regions.

However, despite the analysis that upper atmospheric gravity waves should be quite stable, the perturbation theory solution to the Taylor-Goldstein equations for a plane wave gives an equation for vertical velocity of the form

$$w_1(x, z, t) = \hat{w}(z)e^{z/2H}e^{i(k_x x - \omega t)} \tag{A.99}$$

where w_1 is the perturbed component of the velocity, H is the scale height, and $\hat{w}(z)$ is a sinusoidally varying amplitude component. This equation shows that, due to the first exponential term, the amplitude of a gravity wave increases exponentially with altitude and lower density (Nappo, 2013; Fritts and Alexander, 2003). Because the Richardson number increase allows them to continue to propagate, the amplitudes will grow large enough for the wave to break as it reaches progressively higher altitudes. Because of this, gravity waves can generate turbulence and modulate Richardson numbers in the upper atmosphere, hence transporting energy. This means gravity waves play a critical role in global circulation by depositing energy from the troposphere into the mesosphere and thermosphere; even, as we have seen in Mars data, continuing to propagate and deposit energy above the exobase before dissipating. However, global circulation models for Earth and Mars generally have horizontal resolutions too large to capture the scales of gravity waves and so cannot model them or their energy contributions accurately due to computational constraints. One way to compensate for this is to parameterize the effects of gravity waves; however, as we have seen, the behavior especially in the upper atmosphere is highly non-linear and so difficult to parameterize accurately. Hydrodynamic models where gravity waves are instigated then allowed to propagate can help with determining what parameters a GCM should use to include their effects.

In the upper atmosphere where species are not well-mixed and different species have varying scale heights, the amplitude of gravity waves can be different for species of different masses, depending on the frequency of the wave. For a linear, plane wave perturbation, we can write the change in density over a background density as a function of the flow divergence and the flow's vertical advection:

$$\frac{\rho_1}{\bar{\rho}} = \frac{i}{\omega}\left(\frac{w_1}{H} - \nabla \cdot v_1\right) \tag{A.100}$$

where ρ_1 is the density perturbation, ω the frequency, w_1 is the vertical perturbation velocity, H the scale height, and v_1 the velocity perturbation. For waves with long wavelengths, the first term will dominate and so the amplitude will vary with $1/H$, meaning lighter species will have smaller amplitudes. For small wavelengths, the second term will dominate, and so all species should have the same amplitude, as the velocity perturbation is independent of species. This is seen in gravity wave data from Mars, where waves in the upper thermosphere will have smaller amplitudes in species such as O and N_2 than for heavier species like Ar and CO_2(England et al., 2017).

One of the reasons for this non-linearity is that the atmosphere cannot be assumed to be homogeneous and the decrease of atmospheric density with altitude cannot be ignored, making the Boussinesq approximation that is used in the Taylor-Goldstein equations invalid. This means that, rather than a single gravity wave or wave packet, the upper atmosphere has a spectrum of waves with frequencies ranging from planetary scale to the Brunt-Väisälä frequency. Additionally, as the amplitude grows exponentially as stated above, the wave can break, which will almost always occur high in the atmosphere. Because of this, only the high frequency, long vertical wavelength gravity waves will be able to propagate up to the highest altitudes, as the long vertical wavelength can help compensate for the exponentially increasing amplitude. A 2009 review of gravity wave coupling of the lower and upper atmosphere also found that gravity waves propagate to higher altitudes during times of high solar activity than they do at solar minimum due to higher thermospheric temperatures (Vincent, 2009). This could also indicate that it is possible for gravity waves to propagate to higher altitudes on the dayside of planets, since the thermosphere is both at higher altitudes and warmer there than the nightside.

So, in conclusion, while the momentum transfer between the lower and upper atmosphere by gravity waves is important, it is on too small of a horizontal scale to be captured by global circulation models. However, this does not mean it is not happening; on the contrary, gravity waves will be more easily visibly in the upper atmosphere due to their amplitude increasing with decreasing density. It is also easy for gravity waves to deposit their energy at high altitudes because instability theory shows that the waves are likely to break and induce turbulence in the thermosphere or exosphere. Additionally, we also know that for a wave to reach high altitudes without breaking beforehand, it must have a high frequency and long vertical wavelength, >10 km. This is also aided by a warmer thermosphere, such as one heated by solar activity. Unfortunately, solving the fluid equations analytically is difficult due to the non-linear behavior of turbulence, especially in the upper atmosphere, but we can apply these concepts to data, as we will show in the next section.

A.5 Applications to Data

There are several methods of detecting gravity waves in the atmosphere. One unusual method was to use the Day/Night band on the NASA/NOAA Suomi satellite, for detecting nightglow, to see gravity waves in the upper atmosphere (Miller et al., 2015). More typical methods include ground-based pressure and temperature measurements; weather balloon soundings with pressure sensors; and remote measurements such as radar, Doppler radar, and lidar, in addition to airglow as previously mentioned (Nappo, 2013). At Mars, gravity waves can be seen in *in situ* density measurements of the atmosphere, such as those made by the Neutral Gas and Ion Mass Spectrometer (NGIMS) aboard the MAVEN spacecraft (Yiğit et al., 2015), which is similar to pressure measurements on an Earth station or balloon in type. This is the predominant type of data we will be discussing in this section. However, there are also radio occultation data available from the Radio Science Experiment on Mars Global Surveyor that show vertical wave structures (Creasey, Forbes, and Hinson, 2006).

Since models show that gravity waves play a similarly important role in energy transport in Mars' atmosphere as they do at Earth's, several papers have examined the presence of gravity waves in NGIMS data. Because of the speed with which the satellite traverses the atmosphere, the data has a spatial scale for waves of roughly 20 km. For all of these studies, the data used is usually between 220-160 km, as data resolution for background subtraction is not ideal at high altitudes, and the spacecraft travels predominately horizontally near its periapsis (nominally at 150 km) and so cannot distinguish between vertical and horizontal density perturbations (Yiğit et al., 2015). Gravity waves are seen throughout these altitudes (Yiğit et al., 2015; England et al., 2017; Terada et al., 2017).

The typical process to analyze the gravity waves present is to first fit the background density. There are several methods for doing this; Yigit et al. uses a seventh-order polynomial fit to the log of the density (Yiğit et al., 2015), while other methods include using a moving average or least-squares fit (Terada et al., 2017). This background density profile is removed to retrieve the perturbations, which are then normalized with the background density to get a percentage amplitude of the form $\delta\rho/\rho_0$ where ρ is the density of the data and ρ_0 is a background density (Yiğit et al., 2015). England et al. (2017) then uses a second two-step fitting method to retrieve the wavelength and frequency of the waves that involves using spectral analysis to get the wavelength, then a least-squares fit to the amplitude and wavelength to retrieve frequency and phase. Additionally, using the method outlined in Snowden et al. (2013), a background temperature profile and temperature perturbation is also found (England et al., 2017; Terada et al., 2017). England et al. (2017) found that the majority of the waves in the upper thermosphere had amplitudes of $\leq 10\%$. Terada et al. (2017) found that gravity waves with these small amplitudes tend to have vertical wavelengths of $\lambda_z \simeq 20 - 40km$. However, there are also some waves with amplitudes up to 100% of the background density, which have much larger wavelengths of $\lambda_z \simeq 100 - 200km$.

Interestingly, the data show larger amplitudes near periapsis and higher densities (Yiğit et al., 2015), contrary to what theory would predict. The amplitude is also inversely proportional to perturbations in temperature, which suggests that the reason for the diminishing amplitudes with height is due to developing convective instabilities (Terada et al., 2017). These instabilities limit the amplitude growth and induce wave breaking, as discussed in the previous section. This also explains why those waves with large amplitudes also have large wavelengths, since the theory in the last section shows that long wavelength waves are more resistant to these types of instabilities. However, none of the currently published papers on these upper thermospheric waves calculate the Richardson number, which could confirm the idea that convective instabilities are inhibiting wave propagation at high altitudes. This is counter to gravity waves at Venus and Earth, where radiative damping, among other processes such as molecular diffusion, tend to be the cause of upper atmospheric gravity wave dissipation (Terada et al., 2017).

Finally, we can examine the trends in location of gravity waves at Mars as found in MAVEN data. Yiğit et al. (2015) found that gravity waves were more common in the nightside local times, where mean atmospheric scale heights were lower. This is most likely due to the dayside atmosphere being more stable. Terada et al. (2017) goes into more detail on the various location trends of gravity wave detections, including geographic, Mars-Solar-Orbital (MSO), and Mars-Solar-Electric (MSE) latitude and longitude; local time; solar zenith angle; and various solar variables such as solar wind dynamic pressure, density, and velocity. MSO coordinates refers to a coordinate system where the x vector points towards the Sun and the z vector to the ecliptic north, while in the MSE coordinate system the xz plane is defined by the direction of the solar wind electric field, which makes it a good indicator for particle precipitation and energy deposition.

Terada et al. (2017) finds that the amplitude of thermospheric gravity waves is highly dependent on and inversely proportional to the background thermospheric temperature, as discussed above. They also find that this is correlated with a solar zenith angle and solar EUV flux dependence, since both of these things affect the thermospheric temperature, and that amplitudes on the nightside are roughly twice that of those on the dayside. They do see larger amplitudes in the northern geographic hemisphere, but conclude that this is in fact due to day/night and seasonal coverage and is not an actual correlation. Their results indicate that other than the dependence on atmospheric temperature, there is no real location dependence on the appearance of gravity waves in the thermosphere of Mars, indicating that gravity waves are likely not predominately topographically generated (or, at least not those detected in the thermosphere). They suggest both upward propagation from the lower atmosphere and deposition of energy from precipitating particles are possible sources, but are unable to distinguish between the two.

A.6 Conclusion

In conclusion, we have examined the physics of both a steady state atmosphere and perturbations where gravity or buoyancy is the restoring force. We have then looked at how these physics apply to specifically the upper atmosphere, where the lower densities and larger scale heights affect the way waves propagate. Additionally, this region is where gravity waves largely deposit their energy due to turbulence and wave instabilities, making gravity waves an important method of coupling the lower and upper atmosphere. While much of the theoretical work has been done with Earth in mind, the theory also applies to Mars, where in situ data shows multiple detections of a wide spectrum of gravity wave frequencies and amplitudes. Using this data, one can fit the wave parameters and then apply the theory to understand that convective instabilities drive the damping and dissipation of gravity waves at high altitudes, although there are also detections of large amplitude gravity waves propagating even above the exobase. However, as of yet there has been limited work on proving that convective instabilities are responsible for this, so this remains an area of interest for further study.

Finally, we can discuss how to incorporate this theory and data analysis into a gravity wave simulation. One way to do this is to create a hydrodynamic fluid model, where the simulation creates a fluid governed by the physics outlined above and numerically solves the equations. This is the most efficient way to model a perturbation in a fluid, and works well for a variety of situations. However, this method only works well for situations where the altitude does not vary greatly, or the fluid is not rarefied, as then the fluid does not behave as cohesively. So for the upper atmosphere, some sort of particle simulation is required to accurately model large changes in density with height and low density regimes. In these models, particles obey gravitational laws, as well as having collisions and collisional cross sections. While these models are more accurate, they are also much more computationally expensive, making them unfavorable in the higher density regimes where hydrodynamic models can be used.

Both types of models can be used to study the creation and propagation of gravity waves by inducing a perturbation. In the hydrodynamic model, this can be done using the perturbation equations outlined previously. In the particle model, this perturbation can take the form of directly adding velocity to the particles in the desired location. While solving for the energy transport of a gravity wave analytically is difficult, energy can be directly measured in a model, so it is much easier to find how much energy the wave is carrying. In doing so, the models can form a comparison to the data, which unlike the model is only a snapshot in time. A model, on the other hand, can have the gravity wave evolve over time until it breaks or dissipates, and can also show what initial conditions produce breaking or dissipation, making it a valuable tool in the study of gravity wave, a kind of bridge between theory and data.

Appendix B

Low Beta Regions on the Nightside Ionosphere of Venus

B.1 Introduction

I examine what I have termed low β regions in the nightside Venusian ionosphere, where β is plasma pressure over magnetic pressure. Frequently termed ionospheric holes, they are visible in Pioneer Venus Orbiter (PVO) ion mass spectrometer (OIMS) and magnetometer (OMAG) data. Beginning with Brace et al. (1982), many attempts have been made to study the structure and origin of these regions, as they appear to be unique. The regions consist of an area where the O+ density, as well as other ion densities, are depleted over an order of magnitude and also have a well-ordered magnetic field within the region, which Marubashi et al. (1985) concluded originated in the IMF, and I agree with based on this examination. Figure B.1 shows a typical density depletion indicative of a low β region, with OIMS data taken during orbit 56, with the depletion most evident for O^+. Figure B.2 shows the corresponding magnetic field for the density depression evident in Figure B.1 in Venus-Solar-Orbital (VSO) coordinates, described below.

Brace et al. (1982) suggested the general theory that the magnetic field in these regions is a result of IMF draping around the dayside ionosphere of the planet and the localized plasma depletion is due to the plasma being pulled out along the field lines (Hoegy and Grebowsky, 2010). Marubashi et al. (1985) presented a model in which the field lines are convected by the cross-terminator longitudinal plasma flow, which I have also assumed to be true. These plasma depletions were previously termed ionospheric holes; however, I have chosen the term low β region, as it more accurately represents the physical phenomenon and includes the magnetic field aspect of the structure.

The dataset came from the PVO orbits where periapsis was both on the nightside and near 150 km. This includes orbits 19-130, 243-354, and 467-578 for which the data are of good quality. I began by re-examining the dataset of low β regions used by Hoegy and Grebowsky (2010) and refined it using additional criteria. While previous studies have

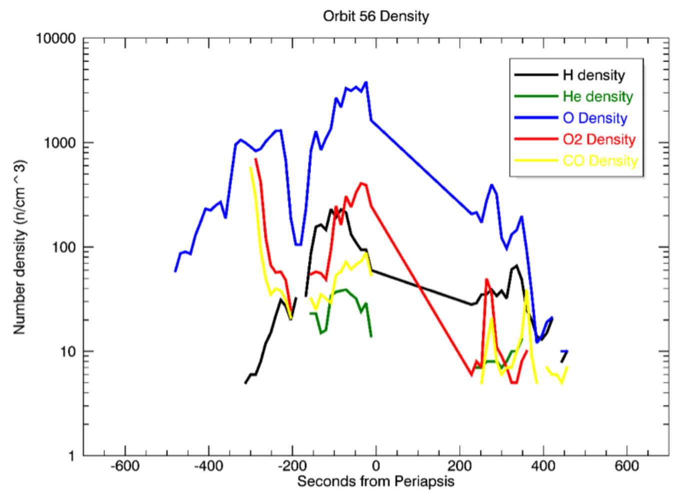


FIGURE B.1: Ion density depression from Orbit 56 taken from OIMS data, with O^+ density in blue, O_2^+ density in red, CO^+ density in yellow, H^+ density in black, and He^+ density in green.

ascertained the presence and location of a region primarily by the ion density, I have chosen to use the magnetic field as the primary criterion. I have done so because plasma deficiencies might not always be present in the lower regions of the ionosphere, where the plasma does not get pulled out (below approximately 200 km) or at high altitude, where the density is too low for any significant change to occur. Additionally, a plasma deficiency might be present and be considered a region, but lack the well-ordered magnetic field necessary to categorize it as such. Therefore, in my treatment of these regions, I have taken the β parameter to be key in categorizing the regions by examining both the plasma density and magnetic density.

I have also excluded those regions in Hoegy and Grebowsky (2010) which had a magnetic field that was not consistent in direction. Some local plasma depletions have magnetic fields, but the direction may change inside the region, which is not consistent with IMF draping in the ionosphere. Therefore, I have restricted this study to 43 regions, approximately one-third of which appear in pairs. This is thought to be due to draping of lines with opposite polarity, as the field direction is generally different for each member of a pair, as seen in Brace et al. (1982). They also concluded that the regions occur in pairs, but the spacecraft did not always pass through both due to the orientation of the orbit.

In this paper, I seek to recreate the IMF direction based coordinate system used in Marubashi et al. (1985), with some corrections. In that paper, the y-coordinate, and hence the local time, was reversed. While I have also created an IMF dependent coordinate system in which I examine these regions, it does not reverse the y coordinate and as such gives a new method to study trends in the locations of these regions. In order to help

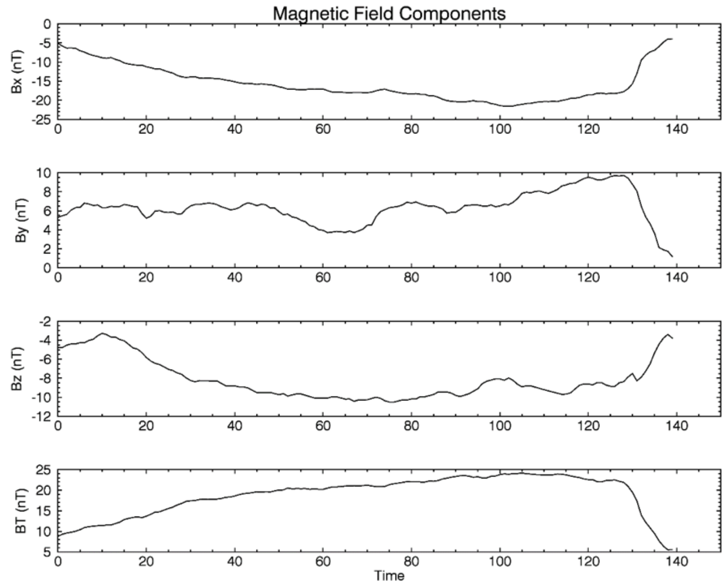


FIGURE B.2: The magnetic field components in VSO coordinates from the ion density depression in Figure B.1. From top to bottom: B_x component, B_y component, and B_z component.

determine the origin of these structures, I also take the curl of the magnetic field vector component, using Ampere's Law to examine a component of the currents present in the ionosphere around the low β regions.

B.2 Data

The data used for this paper comes from the OIMS and OMAG instruments on PVO. Additionally, spacecraft ephemeris data are used to obtain the location of the low β regions. The original ephemeris data are in Venus Solar Orbital (VSO) coordinates, where the x axis points towards the Sun, the z axis is the ecliptic north, and the y axis is perpendicular to both. The average region is between one and two minutes across, which translates to several hundred kilometers. Figure B.3 shows the locations of the regions in VSO coordinates. In the plot, red indicates a paired region and black indicates a single region. Each line represents the part of PVO's orbit that passed through a low β region.

There is a slight bias towards the dawn side in the original data, as discussed in Marubashi et al. (1985), due to a lack of telemetry when the orbiter's periapsis was on the dawn side during a conjuction of Venus and the Sun. Therefore this can be discounted as a lack of sufficient data, not a trend. The regions are fairly evenly split between the northern and southern hemisphere. It may be noted that the regions appear larger in this plot than in previous papers; this is due to choosing the magnetic field as the marker of the region rather than the plasma density. The magnetic field often extends outside the plasma depletion region, particularly for regions that the spacecraft passed through at higher altitudes. Therefore, regions closer to the poles or equator, where the spacecraft passed below the limit of the plasma depletion, appear larger when considering the magnetic field boundaries. Therefore this plot slightly different from those in previous papers.

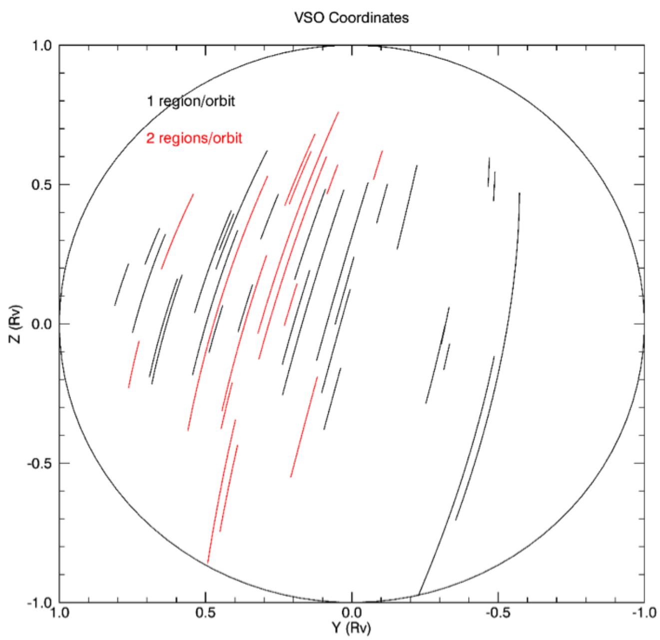


FIGURE B.3: VSO locations of all the low β regions on the nightside of the planet; dawn is on the lefthand side of the plot.

In Luhmann et al. (1982), which also uses the magnetic field as the basis for the regions' locations, the appearance of a magnetic field without a corresponding plasma depletion is discussed as the result of low altitude. That paper concludes that regions on the outbound side of periapsis have a larger plasma depletion than those on the inbound side due to the spacecraft being at higher altitudes. Because periapsis of the spacecraft was at 17 degrees north, inbound regions in the northern hemisphere will be measured at lower altitudes than outbound regions. For this reason they conclude that the magnetic field is a better indicator of size of the low β region than the plasma depletion, since the magnetic field does not disappear at low altitudes. Luhmann and Russell (1992) also notes that while plasma depletions are only visible above 200 km, magnetic fields can be seen much lower in the atmosphere. It is thought that, below around 200 km, the plasma is too dense to be affected as strongly by the magnetic field as at higher altitudes.

Additionally, Luhmann et al. (1982) analyzes the angle of the magnetic field from radial direction. I have also done this and concluded that many of the low β regions do not have a magnetic vector that points predominately in the radial direction. While they do not give an average angle, I surmise that while it would be close to zero, it would still deviate significantly from the radial. While Table B.1 shows that the majority of the regions have either radially directed or sunward pointing magnetic field vectors, there is also significant overlap between regions with both radial and sunward components.

This supports the conclusion in Luhmann and Russell (1992) that the radial field discussed in Brace et al. (1982) and Marubashi et al. (1985) is actually a sunward/antisunward directed field that merely appears radial near the antisolar point. However Luhmann and Russell (1992) also note that in their magnetic vector plots, multiple regions do have a significant horizontal component which is not seen outside of the region in the undisturbed nightside ionosphere. The origin of this horizontal component, which is not consistent with the direction of the IMF draping, will be discussed in the portion of this paper that examines the currents near the edges of the low β regions.

TABLE B.1: Number of regions where B_i contributes $> 70\%$ of B_{tot}

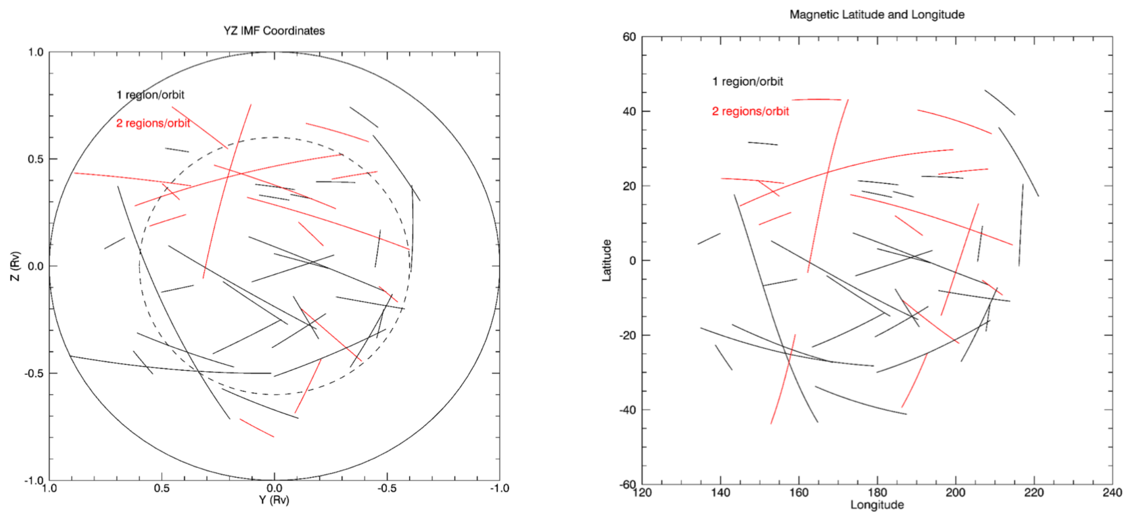
B component	B_x	B_y	B_z	B_{sun}
Number of regions	15	12	21	30

B.3 Results

B.3.1 Coordinate Transformation

In order to examine the effects of IMF angle on the locations of the regions, I created a coordinate system dependent on the IMF clock angle, which is measured from the positive z axis in the YZ plane of VSO coordinates. The clock angle at Venus changes frequently, often within the span of an orbit, and is measured outside the ionopause. For this dataset, I have used the inbound ionopause crossing clock angle for regions on the inbound side of periapsis and the outbound clock angle for outbound regions. There are some cases where either the inbound or outbound clock angle is not available. In this case, the other clock angle measurement was used. If the region was paired, then both regions are then rotated by the same clock angle.

In the IMF coordinates, I have taken the original location of each region in VSO coordinates and rotated it by the respective clock angle, using a rotation matrix. This creates a map of their locations dependent on the direction of the IMF, and is therefore more helpful in determining origin. By setting the IMF direction to be the constant axes in the coordinate system, I have created a system with both a magnetic equator and magnetic poles. In this system, the equator is defined by the point where the plane of the IMF touches the



(A) Rotated coordinates

(B) Magnetic latitude and longitude

FIGURE B.4: Locations of the low β regions in the transformed coordinates. Panel (a) shows the locations in the rotated YZ coordinates as described in the text, and panel (b) shows the locations in the magnetic latitude and longitude. The solid circle in panel (a) indicates the rough boundary of the planet as a reference and the dashed circle indicates the tailward region of the ionosphere. Such a reference is not possible in panel (b).

surface of the planet and the poles are defined by a 90 degree surface arc from the equator. Therefore, I also have a magnetic latitude and longitude. Figure B.4 shows the locations of the regions in the rotated YZ coordinates and in the magnetic latitude and longitude. In this plot, Z is the magnetic north-south axis and Y is the magnetic east-west axis. For the YZ coordinates, dawn is negative Y and dusk is positive Y. For the latitude and longitude plot, longitudes less than 180 degrees are on the dusk side and those greater than 180 degrees are on the dawn side.

After the coordinate transformation, I can observe that the regions are now located roughly within a circle near the antisolar point, which covers over half of the nightside hemisphere. Interestingly, I find that those on the outside of this circular area seem to have slopes in the YZ plane that correspond with the tangent to a circle, while those in the center have slopes closer to horizontal. Figure B.5 shows a histogram of the degree of each regions' slope, measured from the horizontal. As is evident in the histogram, a majority of the regions have slopes under 45 degrees. The larger slopes are from those evident in Figure B.4 near the edge of the circular area. It is also noteworthy that none have slopes greater than 70 degrees from the YZ horizontal, so there are now no vertically oriented regions. However, there is clearly not a Gaussian distribution of slope angles, indicating an element of randomness in the orientation of the regions.

In Brace et al. (1982), it was surmised that there were always two regions, with one

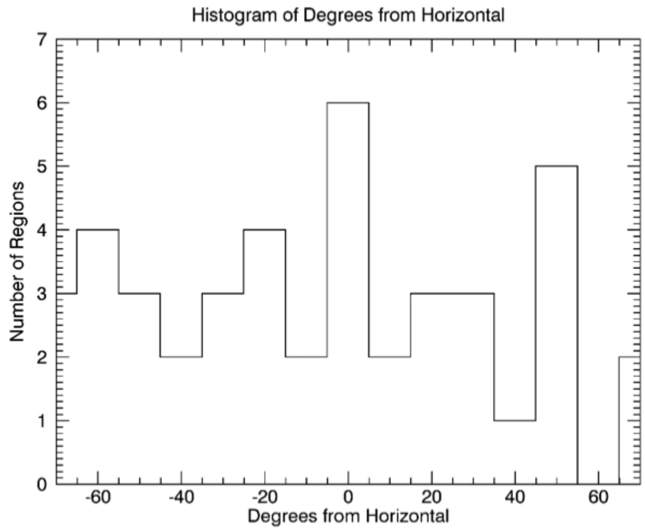


FIGURE B.5: A histogram of the region's slopes in the new coordinate system as defined in the text.

north and one south of the equator, and orbits where only one appeared were due to the spacecraft passing outside of the second region. It is unclear if the new coordinate system sheds any light on this problem. The locations and slopes of the paired regions do not appear to be fundamentally different from those of the single regions. However, regions always being paired seems unlikely due to the difference in numbers of the single regions versus those that are paired. As seen in both Figure B.3 and Figure B.4, there are significantly more single regions. Therefore, it is possible the regions are more transient than believed by Brace et al. (1982).

This is also supported by the inconsistent locations of the regions in the IMF coordinates. If the low β regions were permanent features, as suggested in previous papers, it is unlikely the locations would still be highly variable in the IMF coordinates, as permanent features' locations should only depend on the IMF clock angle. The variability in location could potentially be indicative of other effects contributing to the origin of the low β regions, such as induced currents in the plasma. However, this cannot be proven from the new coordinate system alone.

B.3.2 Magnetic Field Curl Analysis

Using Ampere's Law, I took the curl of the magnetic field vector to examine the currents in the area surrounding the low β regions. To calculate a component of the current using the curl, I interpolated the spacecraft ephemeris data to be the same resolution as the magnetometer data, which was taken of very short intervals on the order of milliseconds. There are several issues with this, as the interpolation is not an accurate representation of

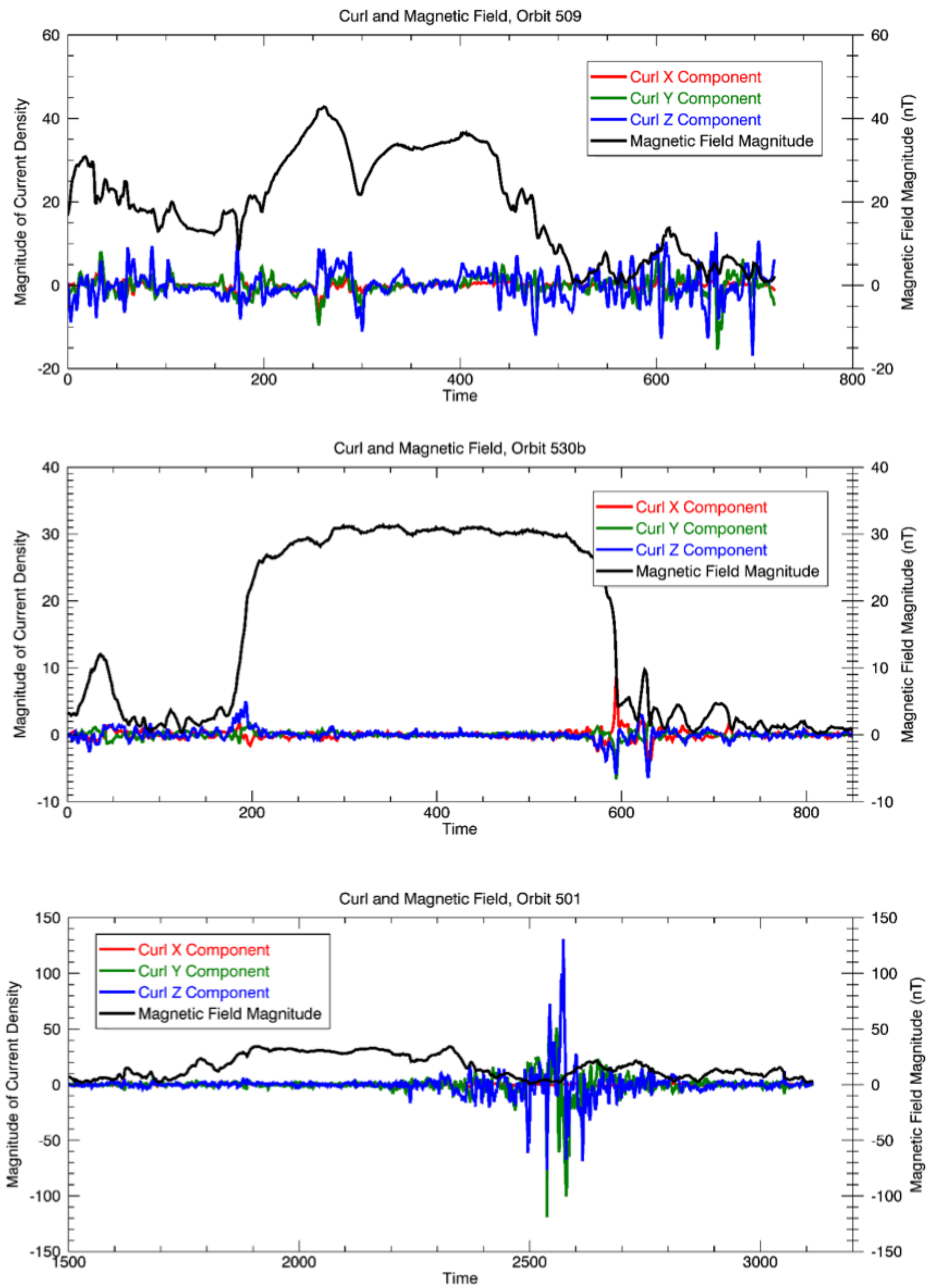


FIGURE B.6: Three examples of the components of the curl vector from a low β region overlaid on the total B magnitude in black.

the spacecraft path and thus introduces rounding errors into the resulting current vector. Potentially due to these errors, the resulting curl data is very noisy. However, some of this noise is introduced by the inherent instability of the magnetic field outside of the low β regions. Figure B.6 shows an example plot of the curl components overlaid on the total magnetic field magnitude to show the boundaries of the low β region.

The noise is evident in Figure B.6, especially outside the boundaries of the region where the magnetic field is not well-ordered. This makes sense, as the curl will naturally be larger where the magnetic field is more variable. It is also worth noting that the curl is small throughout the magnetic field plateau of the low β region. The lack of currents inside the region is due to the ordering of the magnetic field and the low density of the plasma inside the region, which in some orbits is close to a vacuum. When I average the magnitudes of the curl components over each low β region, I can see (shown in Figure B.7) that the averages for the Y and Z components are substantially higher than the X component of the current.

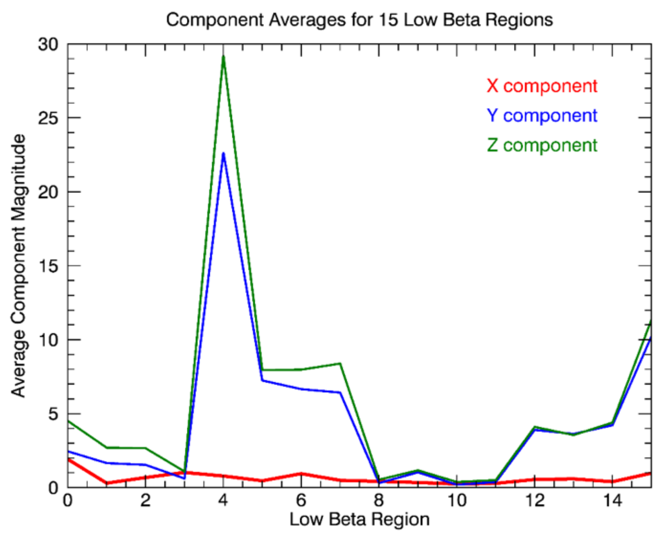


FIGURE B.7: Curl component magnitude averages for 15 low β regions

I only examine the curl components of 15 low β regions, due to the quality of the data. The 15 regions used generally have strong radial fields, and so a horizontal current near the boundaries is expected. To determine if this is a general phenomenon, deriving the curl of regions with, for example, a strong horizontal field is necessary. It is possible that, in those cases, induced currents play a stronger part in shaping the orientation of the magnetic field, such that it deviates from the expected IMF direction.

If the boundary currents are contributing to the magnetic field, it could explain the variability found in both location and slope angle in the IMF coordinates. As mentioned earlier, if the features were permanent results of IMF draping and convecting into the nightside hemisphere, more obvious trends would be expected in an IMF centric coordinate system. Therefore, it is possible that as the plasma flows across the terminator to the chemical sink at the antisolar point (Knudsen et al., 1980), it piles up against the convected field lines and begins to flow around the field lines, likely in a circular manner. As the plasma flows around the field lines, currents could be generated by separation of ions and electrons due to their mass differences. Thus horizontal currents are created around the edges of the low β regions.

However, it is also possible for the created currents to induce a magnetic field that does not point sunward or antisunward. In many cases, while the current is almost completely horizontal, it still can have a radial component. Therefore, this diagonal current vector will then help to induce a magnetic field in a different direction than expected from IMF draping alone. This would account for not only the low β regions with a predominately horizontal magnetic field vector, but also the magnetic field vectors in Luhmann and Russell (1992) with a noticeable horizontal component. The presence of currents also supports the theory of plasma piling up against IMF draped field lines as the origin of the low β regions, while accounting for anomalous behavior that cannot be explained by the draping model alone.

B.4 Conclusion

In summary, I here examine the presence of low β regions in two ways: firstly, by mapping their location in coordinate systems defined by the solar magnetic field, the probable cause of the low β regions, and by calculating the curl of the magnetic field vector in the ionosphere, producing an electric field vector. I find that, unlike in previous publications, these features are likely transient and dependent on quickly-changing solar wind conditions. However, there is no real trend in location in magnetic latitude and longitude, making it unlikely that they are caused by any particular magnetic field configuration, even if they are slightly more common when the magnetic field vector is predominately pointed towards the Sun.

The variability in magnetic coordinate location of these regions may be due to the effects of strong horizontal electric currents at the boundaries of the 15 regions where I had sufficient data quality for finding the curl of the magnetic field vector. These boundary currents may then induce an additional magnetic field that is then detected by the spacecraft, adding an element of randomness to the magnetic field vector that would eliminate any easily found correlation with the IMF. Additionally, the data is limited by the track of the spacecraft, which gives no indication of the true three-dimensional extent of the regions both in ion density and magnetic field. Therefore, it is unlikely I would be able to ascertain the true cause of these low β regions without newer additional data. However, these remain a feature not seen on any other planet and so of great interest for further studies in planetary ionospheric reactions to the solar wind.

Bibliography

Baker, Victor R (2001). "Water and the martian landscape". In: *Nature* 412, pp. 228–236.

Benna, Mehdi and Meredith Elrod (2017). "Neutral Gas and Ion Mass Spectrometer (NGIMS) PDS Software Interface Specification". In: *PDS Atmospheres Node*.

Bird, G. A. (1994). *Molecular gas dynamics and the direct of gas flows simulation*. Clarendon Press. Oxford.

Bird, G.A. (2013). *The DSMC Method*. CreateSpace Independent Publishing Platform, p. 300. ISBN: 1492112909.

Bougher, S et al. (1999). "Mars Global Surveyor aerobraking: Atmospheric trends and model interpretation". In: *Advances in Space Research* 23.11, pp. 1887–1897. ISSN: 0273-1177. DOI: https://doi.org/10.1016/S0273-1177(99)00272-0. URL: http://www.sciencedirect.com/science/article/pii/S0273117799002720.

Bougher, S. et al. (2014). "Mars Global Ionosphere-Thermosphere Model : Solar cycle , seasonal , and diurnal variations of the Mars upper atmosphere". In: *JGR-Planets*, pp. 311–342. DOI: 10.1002/2014JE004715.Received.

Bougher, S et al. (2015). "Early MAVEN Deep Dip campaign reveals thermosphere and ionosphere variability". In: *Science* 350.6261.

Brace, L H et al. (1982). "Holes in the Nightside Ionosphere of Venus Pioneer was have been encountered swept placed into a highly eccentric orbit with an inclination of nightside during the second and third altitude of 11 a periapsis purpose of this paper is to outline our current". In: *Journal of Geophysical Research* 87, pp. 199–211. DOI: 10.1029/JA087iA01p00199.

Brain, D. A. et al. (2015). "The spatial distribution of planetary ion fluxes near Mars observed by MAVEN". In: *Geophysical Research Letters* 42.21, pp. 9142–9148. ISSN: 19448007. DOI: 10.1002/2015GL065293.

Brain, D. A. et al. (2016). "Atmospheric escape from unmagnetized bodies". In: *Journal of Geophysical Research: Planets*, pp. 2364–2385. ISSN: 21699097. DOI: 10.1002/2016JE005162. URL: http://doi.wiley.com/10.1002/2016JE005162.

Bretherton, F. P. (1969). "Waves and Turbulence in Stably Stratified Fluids". In: *Radio Science* 4.12, pp. 1279–1287. ISSN: 1944799X. DOI: 10.1029/RS004i012p01279.

Brown, Robert A. (1991). *Fluid Mechanics of the Atmosphere*. 1st. Academic Press, Inc. ISBN: 0121370402.

Charney, J. G. and P. G. Drazin (1961). "Propagation of planetary-scale disturbances from the lower into the upper atmosphere". In: *Journal of Geophysical Research* 66.1, pp. 83–109. ISSN: 01480227. DOI: 10.1029/JZ066i001p00083. URL: http://doi.wiley.com/10.1029/JZ066i001p00083.

Chassefière, Eric and François Leblanc (2004). "Mars atmospheric escape and evolution; interaction with the solar wind". In: *Planetary and Space Science* 52.11, pp. 1039–1058. ISSN: 00320633. DOI: 10.1016/j.pss.2004.07.002.

Connerney, J E P et al. (2015). "The MAVEN Magnetic Field Investigation". In: *Space Science Reviews* 195.1, pp. 257–291. ISSN: 1572-9672. DOI: 10.1007/s11214-015-0169-4. URL: https://doi.org/10.1007/s11214-015-0169-4.

Creasey, John E, Jeffrey M Forbes, and David P Hinson (2006). "Global and seasonal distribution of gravity wave activity in Mars' lower atmosphere derived from MGS radio occultation data". In: *Geophysical Research Letters* 33.1, n/a–n/a. ISSN: 1944-8007. DOI: 10.1029/2005GL024037. URL: http://dx.doi.org/10.1029/2005GL024037.

Cui, J et al. (2014). "Density waves in Titan's upper atmosphere". In: *Journal of Geophysical Research: Space Physics* 119, pp. 490–518. DOI: 10.1002/2013JA019113.Received.

Curry, Shannon M et al. (2015). "Comparative pick-up ion distributions at Mars and Venus : Consequences for atmospheric deposition and escape". In: *Planetary and Space Science* 115, pp. 35–47. ISSN: 0032-0633. DOI: 10.1016/j.pss.2015.03.026. URL: http://dx.doi.org/10.1016/j.pss.2015.03.026.

Deighan, J et al. (2015). "MAVEN IUVS observation of the hot oxygen corona at Mars". In: *Geophysical Research Letters* 42, pp. 9009–9014. DOI: 10.1002/2015GL065487.Received.

Delisi, D.P. and I. Orlanski (1975). "On the role of density jumps in the reflexion and breaking of internal gravity waves". In: *Journal of Fluid Mechanics* 69, pp. 445–464. DOI: 10.1017/S0022112075001516.

Dong, Y et al. (2015). "Strong plume fluxes at Mars observed by MAVEN: An important planetary ion escape channel". In: *Geophysical Research Letters* 42, 8942–8950. ISSN: 19448007. DOI: 10.1002/2015GL065346.Received.

Dong, Y. et al. (2017). "Seasonal variability of Martian ion escape through the plume and tail from MAVEN observations". In: *Journal of Geophysical Research: Space Physics* 122.4, pp. 4009–4022. ISSN: 21699402. DOI: 10.1002/2016JA023517.

Dundas, Colin M et al. (2018). "Exposed subsurface ice sheets in the Martian mid-latitudes". In: *Science* 359.6372, pp. 199–201. ISSN: 0036-8075. DOI: 10.1126/science.aao1619. URL: http://science.sciencemag.org/content/359/6372/199.

Dunn, Patrick (2015). "Mars Atmosphere and Volatile Evolution (MAVEN) Mission In Situ Instruments Key Parameters PDS Archive Software Interface Specification". In: pp. 1–51.

Elrod Meredith; Navas, Tiffany; Benna Mehdi (2015). *NGIMS L2 Data Collection (Version 8)*. URL: https://atmos.nmsu.edu/PDS/data/PDS4/MAVEN/ngims_bundle/l2/.

England, S. L. et al. (2017). “MAVEN NGIMS observations of atmospheric gravity waves in the Martian thermosphere”. In: *Journal of Geophysical Research: Space Physics* 122.2, pp. 2310–2335. ISSN: 21699402. DOI: 10.1002/2016JA023475.

England, Scott L et al. (2016). “Simultaneous observations of atmospheric tides from combined in situ and remote observations at Mars from the MAVEN spacecraft”. In: *Journal of Geophysical Research: Planets* 121.4, pp. 594–607. DOI: 10.1002/2016JE004997. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JE004997.

Eparvier, F G et al. (2015). “The Solar Extreme Ultraviolet Monitor for MAVEN”. In: *Space Science Reviews* 195.1, pp. 293–301. ISSN: 1572-9672. DOI: 10.1007/s11214-015-0195-2. URL: https://doi.org/10.1007/s11214-015-0195-2.

Fang, Xiaohua et al. (2013). “The importance of pickup oxygen ion precipitation to the Mars upper atmosphere under extreme solar wind conditions”. In: *Geophysical Research Letters* 40.10, pp. 1922–1927. ISSN: 00948276. DOI: 10.1002/grl.50415. URL: http://doi.wiley.com/10.1002/grl.50415.

Fox, J. L. (2004). “CO2+ dissociative recombination: A source of thermal and nonthermal C on Mars”. In: *Journal of Geophysical Research: Space Physics* 109.A8, pp. 1–9. ISSN: 21699402. DOI: 10.1029/2004JA010514.

Fox, Jane L. and Aleksander B. Hać (2009). “Photochemical escape of oxygen from Mars: A comparison of the exobase approximation to a Monte Carlo method”. In: *Icarus* 204.2, pp. 527–544. ISSN: 00191035. DOI: 10.1016/j.icarus.2009.07.005.

Fritts, D C (1984). “Gravity Wave Saturation in the Middle Atmosphere: A Review of Theory and Observations”. In: *Rev. Geophys.* 22.3, pp. 275–308.

Fritts, David C. and Timothy J. Dunkerton (1984). *A Quasi-Linear Study of Gravity-Wave Saturation and Self-Acceleration*. DOI: 10.1175/1520-0469(1984)041<3272:AQLSOG>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1984)041%3C3272:AQLSOG%3E2.0.CO;2.

Fritts, David C, Ling Wang, and Robert H Tolson (2006). “Mean and gravity wave structures and variability in the Mars upper atmosphere inferred from Mars Global Surveyor and Mars Odyssey aerobraking densities”. In: *Journal of Geophysical Research: Space Physics* 111.A12. DOI: 10.1029/2006JA011897. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JA011897.

Fritts, Dc David C. and Mj Alexander (2003). “Gravity wave dynamics and effects in the middle atmosphere”. In: *Reviews of Geophysics* 41.1, p. 1003. ISSN: 8755-1209. DOI: 10.1029/2001RG000106. URL: http://www.agu.org/pubs/crossref/2003/2001RG000106.shtml%5Cnhttp://doi.wiley.com/10.1029/2001RG000106.

Garcia, Raphael F. et al. (2017). “Finite-Difference Modeling of Acoustic and Gravity Wave Propagation in Mars Atmosphere: Application to Infrasounds Emitted by Meteor Impacts”. In: *Space Science Reviews* 211.1-4, pp. 547–570. ISSN: 15729672. DOI: 10.1007/s11214-016-0324-6. URL: http://dx.doi.org/10.1007/s11214-016-0324-6.

Geller, Marvin A., Hiroshi Tanaka, and David C. Fritts (1975). "Production of Turbulence in the Vicinity of Critical Levels for Internal Gravity Waves". In: *Journal of the Atmospheric Sciences* 32.11, pp. 2125–2135. ISSN: 0022-4928. DOI: 10.1175/1520-0469(1975)032<2125:POTITV>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1975)032%3C2125:POTITV%3E2.0.CO;2.

Green, John (1999). *Atmospheric Dynamics*. 1st. Cambridge University Press. ISBN: 0521249759.

Griffiths, David J. (1995). *Introduction to Quantum Mechanics*. Upper Saddle River, NJ: Prentice Hall. ISBN: 0131244051.

Halekas, J S et al. (2015). "The Solar Wind Ion Analyzer for MAVEN". In: *Space Science Reviews* 195.1, pp. 125–151. ISSN: 1572-9672. DOI: 10.1007/s11214-013-0029-z. URL: https://doi.org/10.1007/s11214-013-0029-z.

Halekas, J. S. et al. (2017). "Structure, dynamics, and seasonal variability of the Mars-solar wind interaction: MAVEN Solar Wind Ion Analyzer in-flight performance and science results". In: *Journal of Geophysical Research: Space Physics* 122.1, pp. 547–578. ISSN: 21699402. DOI: 10.1002/2016JA023167.

Hara, Takuya et al. (2017). "MAVEN observations on a hemispheric asymmetry of precipitating ions toward the Martian upper atmosphere according to the upstream solar wind electric field". In: *Journal of Geophysical Research: Space Physics*, pp. 1–19. ISSN: 21699380. DOI: 10.1002/2016JA023348. URL: http://doi.wiley.com/10.1002/2016JA023348.

Hickey, M. P., R. L. Walterscheid, and G. Schubert (2011). "Gravity wave heating and cooling of the thermosphere: Sensible heat flux and viscous flux of kinetic energy". In: *Journal of Geophysical Research: Space Physics* 116.12. ISSN: 21699402. DOI: 10.1029/2011JA016792.

Hines, C O (1960). "Internal Atmospheric Gravity Waves at Ionospheric Heights". In: *The Upper Atmosphere in Motion*. Vol. 18. American Geophysical Union, pp. 248–328. ISBN: 9780875900186. DOI: 10.1029/GM018p0248.

Hodges, R. R. Jr. (1967). "Generation of Turbulence in the Upper Atmosphere by Internal Gravity Waves". In: *Journal of Geophysical Research* 72.13.

— (1969). "Eddy diffusion coefficients due to instabilities in internal gravity waves". In: *Journal of Geophysical Research* 74, pp. 4087–4090. DOI: 10.1029/JA074i016p04087.

Hoegy, Walter R. and Joseph M. Grebowsky (2010). "Venus nightside ionospheric holes". In: *Journal of Geophysical Research: Space Physics* 115.12, pp. 1–10. ISSN: 21699402. DOI: 10.1029/2010JA015675.

Holton, J R (2015). "STRATOSPHERE/TROPOSPHERE EXCHANGE & STRUCTURE | Global Aspects". In: *Encyclopedia of Atmospheric Sciences (Second Edition)*. Ed. by Gerald R North, John Pyle, and Fuqing Zhang. Second Edi. Oxford: Academic Press, pp. 257–261. ISBN: 978-0-12-382225-3. DOI: https://doi.org/10.1016/B978-0-

12-382225-3.00394-7. URL: http://www.sciencedirect.com/science/article/pii/B9780123822253003947.

Hunsucker, Robert D. (1982). "Atmospheric gravity waves generated in the high latitude ionosphere: A review". In: *Reviews of Geophysics* 20.2, pp. 293–315. ISSN: 19449208. DOI: 10.1029/RG020i002p00293.

Izakov, M. N. (1978). "The Martian upper atmosphere structure from the Viking spacecraft experiments". In: *Icarus* 36.2, pp. 189–197. ISSN: 10902643. DOI: 10.1016/0019-1035(78)90103-3.

Jakosky, B. M. et al. (2015). "The Mars Atmosphere and Volatile Evolution (MAVEN) Mission". In: *Space Science Reviews*, pp. 3–48. ISSN: 15729672. DOI: 10.1007/s11214-015-0139-x. URL: http://dx.doi.org/10.1007/s11214-015-0139-x.

Jakosky, B. M. et al. (2017). "Mars' atmospheric history derived from upper-atmosphere measurements of 38Ar/36Ar". In: *Science* 355.6332, pp. 1–14. ISSN: 10959203. DOI: 10.1126/science.aai7721.

Jakosky, Bruce M. et al. (1994). "Mars atmospheric loss and isotopic fractionation by solar-wind-induced sputtering and photochemical escape". In: *Icarus* 111, pp. 271–288. ISSN: 0019-1035. DOI: 10.1006/icar.1994.1145.

Johnson, R. E. (1990a). *Energetic Charged-Particle Interactions with Atmospheres and Surfaces*. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642-48377-6 978-3-642-48375-2. URL: http://link.springer.com/10.1007/978-3-642-48375-2 (visited on 04/28/2015).

— (1994). "Plasma-Induced Sputtering Of An Atmosphere". In: *Space Science Reviews* 69, pp. 215–253. ISSN: 0717-6163.

Johnson, R E and M Liu (1996). "The loss of atmosphere from Mars." In: *Science (New York, N.Y.)* 274.5294, pp. 1932–1933. ISSN: 0036-8075. DOI: 10.1126/science.274.5294.1932b.

Johnson, R. E. and J. G. Luhmann (1998). "Sputter contribution to the atmospheric corona on Mars". In: *Journal of Geophysical Research-Planets* 103, pp. 3649–3653.

Johnson, R. E. et al. (2015). "VOLATILE LOSS AND CLASSIFICATION OF KUIPER BELT OBJECTS". In: *The Astrophysical Journal* 809.1, p. 43. ISSN: 1538-4357. DOI: 10.1088/0004-637X/809/1/43. URL: http://stacks.iop.org/0004-637X/809/i=1/a=43?key=crossref.2df71f738c24327b39a770fb752e1d04 (visited on 04/03/2018).

Johnson, Robert E (1990b). *Energetic Charged-Particle Interactions with Atmospheres and Surfaces*. English. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642-48375-2 3-642-48375-5. URL: http://dx.doi.org/10.1007/978-3-642-48375-2 (visited on 04/28/2015).

Johnson, Robert E., D. Schnellenberger, and M. C. Wong (2000). "The sputtering of an oxygen thermosphere by energetic O+". In: *Journal of Geophysical Research* 105.E1, pp. 1659–

1670. ISSN: 0148-0227. DOI: 10.1029/1999JE001058. URL: http://www.agu.org/pubs/crossref/2000/1999JE001058.shtml.

Johnson, Robert E., Alexey N. Volkov, and Justin T. Erwin (2013a). "ERRATUM: "MOLECULAR-KINETIC SIMULATIONS OF ESCAPE FROM THE EX-PLANET AND EXOPLANETS: CRITERION FOR TRANSONIC FLOW" (2013, ApJL, 768, L4)". In: *The Astrophysical Journal* 779.2, p. L30. ISSN: 2041-8205, 2041-8213. DOI: 10.1088/2041-8205/779/2/L30. URL: http://stacks.iop.org/2041-8205/779/i=2/a=L30?key=crossref.c1804cb179a616ad47eeb25c3dd62724 (visited on 04/03/2018).

— (2013b). "MOLECULAR-KINETIC SIMULATIONS OF ESCAPE FROM THE EX-PLANET AND EXOPLANETS: CRITERION FOR TRANSONIC FLOW". In: *The Astrophysical Journal* 768.1, p. L4. ISSN: 2041-8205, 2041-8213. DOI: 10.1088/2041-8205/768/1/L4. URL: http://stacks.iop.org/2041-8205/768/i=1/a=L4?key=crossref.ebf4657d9cfb48a1b84eced73f233bfb (visited on 04/03/2018).

Kharchenko, V. et al. (2000). "Energy transfer in collisions of oxygen atoms in the terrestrial atmosphere". In: *Journal of Geophysical Research: Space Physics* 105, pp. 899–24.

Kloos, J L et al. (2018). "Interannual and Diurnal Variability in Water Ice Clouds Observed from MSL Over Two Martian Years". In: *Journal of Geophysical Research: Planets* 123.1, pp. 233–245. DOI: 10.1002/2017JE005314. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JE005314.

Knudsen, William C et al. (1980). "Transport of ionospheric O+ ions across the Venus terminator and implications". In: *Journal of Geophysical Research* 85.A13, 7803–7810. ISSN: 0148-0227. DOI: 10.1029/JA085iA13p07803. URL: http://www.agu.org/pubs/crossref/1980/JA085iA13p07803.shtml.

Leblanc, F. and R. E. Johnson (2001). "Sputtering of the Martian atmosphere by solar wind pick-up ions". In: *Planetary and Space Science* 49.6, pp. 645–656. ISSN: 00320633. DOI: 10.1016/S0032-0633(01)00003-4.

Leblanc, F. and R.E. Johnson (2002). "Role of molecular species in pickup ion sputtering of the Martian atmosphere". In: *Journal of Geophysical Research* 107.E2, p. 5010. ISSN: 0148-0227. DOI: 10.1029/2000JE001473. URL: http://doi.wiley.com/10.1029/2000JE001473.

Leblanc, F et al. (2015). "Mars heavy ion precipitating flux as measured by Mars Atmosphere and Volatile EvolutioN". In: *Geophysical Research Letters* 42, pp. 9135–9141. ISSN: 19448007. DOI: 10.1002/2015GL066170.Received.

Leblanc, F. et al. (2017). "On the Origins of Mars' Exospheric Nonthermal Oxygen Component as Observed by MAVEN and Modeled by HELIOSARES". In: *Journal of Geophysical Research: Planets*, pp. 1–28. ISSN: 21699097. DOI: 10.1002/2017JE005336. URL: http://doi.wiley.com/10.1002/2017JE005336.

Leclercq, Ludivine et al. (2018). "Propagation of Transient Perturbations into a Planet's Exosphere: Molecular Kinetic Simulations". In: *The Astrophysical Journal Letters*.

Lee, Yuni et al. (2015). "A comparison of 3-D model predictions of Mars' oxygen corona with early MAVEN IUVS observations". In: *Geophysical Research Letters* 42, pp. 9015–9022. DOI: 10.1002/2015GL065291.Received.

Lewkow, N. R. and V. Kharchenko (2014). "PRECIPITATION OF ENERGETIC NEUTRAL ATOMS AND INDUCED NON-THERMAL ESCAPE FLUXES FROM THE MARTIAN ATMOSPHERE". In: *The Astrophysical Journal* 790.2, p. 98. ISSN: 0004-637X, 1538-4357. DOI: 10.1088/0004-637X/790/2/98. URL: http://stacks.iop.org/0004-637X/790/i=2/a=98?key=crossref.e1d1418269b4a590ea9503c78aa33442 (visited on 04/12/2017).

Lillis, R. J. et al. (2015). "Characterizing Atmospheric Escape from Mars Today and Through Time, with MAVEN". In: *Space Science Reviews*, pp. 357–422. ISSN: 15729672. DOI: 10.1007/s11214-015-0165-8.

Lillis, Robert J. et al. (2017). "Photochemical escape of oxygen from Mars: First results from MAVEN in situ data". In: *Journal of Geophysical Research: Space Physics* 122.3, pp. 3815–3836. ISSN: 21699402. DOI: 10.1002/2016JA023525.

Linzen, Richard (2008). "Internal Gravity Waves : Basics". In: *12.810 Dynamics of the Atmosphere*. Massachusetts Institute of Technology: MIT OpenCourseWare. Chap. 8, pp. 149–172. URL: https://ocw.mit.edu/courses/earth-atmospheric-and-planetary-sciences/12-810-dynamics-of-the-atmosphere-spring-2008/lecture-notes/chapter_8.pdf.

Liu, Guiping et al. (2017). "Longitudinal structures in Mars' upper atmosphere as observed by MAVEN/NGIMS". en. In: *Journal of Geophysical Research: Space Physics* 122.1, pp. 1258–1268. ISSN: 21699380. DOI: 10.1002/2016JA023455. URL: http://doi.wiley.com/10.1002/2016JA023455 (visited on 03/28/2018).

Luhmann, J. G., R. E. Johnson, and M. H. G. Zhang (1992). "Evolutionary impact of sputtering of the Martian atmosphere by O(+) pickup ions". In: *Geophysical Research Letters* 19.21, pp. 2151–2154. ISSN: 00948276. DOI: 10.1029/92GL02485.

Luhmann, J. G. and D. S. Russell (1992). "Magnetic fields in Venus nightside ionospheric holes - Collected Pioneer Venus Orbiter magnetometer observations". In: *Journal of Geophysical Research* 97, pp. 10267–10282. ISSN: 0148-0227. DOI: 10.1029/92JE00790.

Luhmann, J. G. et al. (1982). "Pioneer Venus Observations of Plasma and Field Structure in the Near Wake of Venus". In: *Journal of Geophysical Research* 87, pp. 9205–9210. DOI: 10.1029/JA087iA11p09205.

Mahaffy, Paul R. et al. (2015a). "Structure and composition of the neutral upper atmosphere of Mars from the MAVEN NGIMS investigation". In: *Geophysical Research Letters* 42, pp. 8951–8957. DOI: 10.1002/2015GL065329.

Mahaffy, Paul R. et al. (2015b). "The Neutral Gas and Ion Mass Spectrometer on the Mars Atmosphere and Volatile Evolution Mission". In: *Space Science Reviews* 195.1-4, pp. 49–

73. ISSN: 0038-6308. DOI: 10.1007/s11214-014-0091-1. URL: http://link.springer.com/10.1007/s11214-014-0091-1.

Marubashi, K. et al. (1985). "Magnetic field in the wake of Venus and the formation of ionospheric holes". In: *Journal of Geophysical Research* 90, pp. 1385–1398. ISSN: 0148-0227. DOI: 10.1029/JA090iA02p01385.

MAVEN Insitu Key Parameters Data Bundle (2019).

McComas, D. J. et al. (2013). "Weakest Solar Wind of the Space Age and the Current "Mini" Solar Maximum". In: *The Astrophysical Journal* 779.1, p. 2. ISSN: 0004-637X. DOI: 10.1088/0004-637X/779/1/2. URL: http://stacks.iop.org/0004-637X/779/i=1/a=2?key=crossref.7267c4409c973e925804ba4ba07b3c0e.

Medvedev, Alexander S. et al. (2015). "Cooling of the Martian thermosphere by CO 2 radiation and gravity waves: An intercomparison study with two general circulation models". In: *Journal of Geophysical Research: Planets* 120.5, pp. 913–927. ISSN: 21699097. DOI: 10.1002/2015JE004802. URL: http://doi.wiley.com/10.1002/2015JE004802.

Michael, M. and R. E. Johnson (2005). "Energy deposition of pickup ions and heating of Titan's atmosphere". In: *Planetary and Space Science* 53.14-15, pp. 1510–1514. ISSN: 00320633. DOI: 10.1016/j.pss.2005.08.001.

Midgley, J E and H B Liemohn (1966). "Gravity waves in a realistic atmosphere". In: *Journal of Geophysical Research* 71.15, pp. 3729–3748. ISSN: 2156-2202. DOI: 10.1029/JZ071i015p03729. URL: http://dx.doi.org/10.1029/JZ071i015p03729.

Miller, Steven D. et al. (2015). "Upper atmospheric gravity wave details revealed in nightglow satellite imagery". In: *Proceedings of the National Academy of Sciences* 112.49, E6728–E6735. ISSN: 0027-8424. DOI: 10.1073/pnas.1508084112. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1508084112.

Nappo, Carmen J. (2013). *An Introduction to Atmospheric Gravity Waves*. 2nd. Elsevier. ISBN: 9780123852236.

Nier, A O and M B McElroy (1977). "Composition and structure of Mars' Upper atmosphere: Results from the neutral mass spectrometers on Viking 1 and 2". In: *Journal of Geophysical Research (1896-1977)* 82.28, pp. 4341–4349. DOI: 10.1029/JS082i028p04341. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JS082i028p04341.

Rahmati, A. et al. (2015). "MAVEN insights into oxygen pickup ions at Mars". In: *Geophysical Research Letters* 42.21, pp. 8870–8876. ISSN: 19448007. DOI: 10.1002/2015GL065262.

Scargle, J. D. (1982). "Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data." In: *Ap.J.* 263, pp. 835–853. DOI: 10.1086/160554.

Slipski, M et al. (2018). "Variability of Martian Turbopause Altitudes". In: *Journal of Geophysical Research (Planets)* 123, pp. 2939–2957. DOI: 10.1029/2018JE005704.

Smith, Gerald R. et al. (1978). "Monte Carlo modeling of exospheric bodies: Mercury". en. In: *Journal of Geophysical Research* 83.A8, p. 3783. ISSN: 0148-0227. DOI: 10.1029/

JA083iA08p03783. URL: http://doi.wiley.com/10.1029/JA083iA08p03783 (visited on 04/28/2015).

Snowden, D. et al. (2013). "The thermal structure of titan's upper atmosphere, I: Temperature profiles from Cassini INMS observations". In: *Icarus* 226.1, pp. 552–582. ISSN: 00191035. DOI: 10.1016/j.icarus.2013.06.006. URL: http://dx.doi.org/10.1016/j.icarus.2013.06.006.

Terada, Naoki et al. (2017). "Global distribution and parameter dependences of gravity wave activity in the Martian upper thermosphere derived from MAVEN/NGIMS observations". In: *Journal of Geophysical Research: Space Physics* 122.2, pp. 2374–2397. ISSN: 21699402. DOI: 10.1002/2016JA023476.

Thiemann, E. M B et al. (2015). "Neutral density response to solar flares at Mars". In: *Geophysical Research Letters* 42.21, pp. 8986–8992. ISSN: 19448007. DOI: 10.1002/2015GL066334.

Tolstoy, Ivan (1963). "The Theory of Waves in Stratified Fluids Including the Effects of Gravity and Rotation". In: *Review of Modern Physics* 35.1, pp. 207–230.

Tucker, O.J. et al. (2012). "Thermally driven escape from Pluto's atmosphere: A combined fluid/kinetic model". en. In: *Icarus* 217.1, pp. 408–415. ISSN: 00191035. DOI: 10.1016/j.icarus.2011.11.017. URL: http://linkinghub.elsevier.com/retrieve/pii/S0019103511004441 (visited on 04/03/2018).

Tucker, O.J. et al. (2013). "Diffusion and thermal escape of H2 from Titan's atmosphere: Monte Carlo simulations". en. In: *Icarus* 222.1, pp. 149–158. ISSN: 00191035. DOI: 10.1016/j.icarus.2012.10.016. URL: http://linkinghub.elsevier.com/retrieve/pii/S001910351200423X (visited on 03/16/2018).

Tucker, Orenthal J. and R.E. Johnson (2009). "Thermally driven atmospheric escape: Monte Carlo simulations for Titan's atmosphere". en. In: *Planetary and Space Science* 57.14-15, pp. 1889–1894. ISSN: 00320633. DOI: 10.1016/j.pss.2009.06.003. URL: http://linkinghub.elsevier.com/retrieve/pii/S0032063309001597 (visited on 03/16/2018).

Tucker, Orenthal J. et al. (2016). "Examining the exobase approximation: DSMC models of Titan's upper atmosphere". en. In: *Icarus* 272, pp. 290–300. ISSN: 00191035. DOI: 10.1016/j.icarus.2016.02.044. URL: http://linkinghub.elsevier.com/retrieve/pii/S0019103516001226 (visited on 03/16/2018).

Tully, C. and R.E. Johnson (2001). "Low energy collisions between ground-state oxygen atoms". In: *Planetary and Space Science* 49.6, pp. 533–537. ISSN: 00320633. DOI: 10.1016/S0032-0633(01)00002-2. URL: http://linkinghub.elsevier.com/retrieve/pii/S0032063301000022.

Valeille, Arnaud et al. (2009). "Three-dimensional study of Mars upper thermosphere/ionosphere and hot oxygen corona: 1. General description and results at equinox for solar low conditions". In: *Journal of Geophysical Research* 114.E11, E11005. ISSN: 0148-0227. DOI: 10.1029/2009JE003388. URL: http://doi.wiley.com/10.1029/2009JE003388.

Vignes, D. et al. (2000). "The Solar Wind interaction with Mars: Locations and shapes of the Bow Shock and the Magnetic Pile-up Boundary from the observations of the MAG/ER experiment onboard Mars Global Surveyor". In: *Geophysical Research Letters* 27.1, pp. 49–52. ISSN: 00948276. DOI: 10.1029/1999GL010703.

Vincent, RA (2009). "Gravity wave coupling from below: A review". In: *CAWSES: Selected Papers From the 2007 Kyoto ...* 1960, pp. 279–293. URL: http://www.terrapub.co.jp/onlineproceedings/ste/CAWSES2007/pdf/CAWSES_279.pdf.

Volkov, Alexey N. and Robert E. Johnson (2013). "THERMAL ESCAPE IN THE HYDRODYNAMIC REGIME: RECONSIDERATION OF PARKER's ISENTROPIC THEORY BASED ON RESULTS OF KINETIC SIMULATIONS". In: *The Astrophysical Journal* 765.2, p. 90. ISSN: 0004-637X, 1538-4357. DOI: 10.1088/0004-637X/765/2/90. URL: http://stacks.iop.org/0004-637X/765/i=2/a=90?key=crossref.84575feafdbe3d6bee393747fb89603a (visited on 04/03/2018).

Volkov, Alexey N. et al. (2011). "Thermally-driven atmospheric escape: Transition from hydrodynamic to Jeans escape". In: *The Astrophysical Journal* 729.2, p. L24. ISSN: 2041-8205, 2041-8213. DOI: 10.1088/2041-8205/729/2/L24. URL: http://stacks.iop.org/2041-8205/729/i=2/a=L24?key=crossref.760587099e0e36c7c158577f97db9e9d (visited on 03/28/2018).

Waite, J Hunter et al. (2005). "Ion Neutral Mass Spectrometer Results from the First Flyby of Titan". In: *Science* 308.5724, pp. 982–986. ISSN: 0036-8075. DOI: 10.1126/science.1110652. URL: https://science.sciencemag.org/content/308/5724/982.

Walterscheid, R. L., M. P. Hickey, and G. Schubert (2013). "Wave heating and jeans escape in the martian upper atmosphere". In: *Journal of Geophysical Research E: Planets* 118.11, pp. 2413–2422. ISSN: 01480227. DOI: 10.1002/jgre.20164.

Withers, Paul et al. (2015). "Changes in the thermosphere and ionosphere of Mars from Viking to MAVEN". In: *Geophysical Research Letters* 42.21, pp. 9071–9079. ISSN: 19448007. DOI: 10.1002/2015GL065985.

Yiğit, Erdal et al. (2015). "High-altitude gravity waves in the Martian thermosphere observed by MAVEN/NGIMS and modeled by a gravity wave scheme". In: *Geophysical Research Letters* 42.21, pp. 8993–9000. ISSN: 19448007. DOI: 10.1002/2015GL065307.