Studies of the Martian atmospheric boundary layer and global circulation from combined use of spacecraft data and numerical circulation models

Alexandru Mihai Valeanu

Linacre College

A Thesis submitted for the degree of

Doctor of Philosophy

Atmospheric, Oceanic and Planetary Physics

Department of Physics

University of Oxford

December 2018

Acknowledgements

I would like express my deep gratitude to Professor Peter Read, my supervisor, for being the perpetual tutor that I met in my first week of undergraduate at Trinity College, and for his patient guidance in my development as a researcher. I always cherished his confidence in me and hope to never shatter it.

I would also like to offer my special thanks to Russell Jones for laying the foundations for my programming knowledge, for his immeasurable support in overcoming some of the worst technical challenges of my project and for his immersive discussions. Pierre Augier, Professor Stephen Lewis, Luca Montabone, Liam Steele, Fachreddin Tabataba-Vakili and Roland Young, for their swift and invaluable support, and interesting discussions.

I would also like to thank my closest friends, especially Alexandru Grozavu and David Hansford. You have my gratitude for enduring the difficult times I went through. And of course, deepest thanks to Benish Khan for her unending confidence in me (I finished!).

Finally, I would like to thank my parents, Gabriela and Emil for making everything in my life possible.

Abstract

In this thesis, the near-surface environment on Mars is studied in connection to the general atmospheric circulation. Well-established methods for modelling Earth's atmosphere were invoked, such as embedding a regional mesoscale model into a global atmospheric circulation model, and driving the atmospheric states from the global model, by assimilation of orbital spacecraft data. Embedding a regional model inside a global one for Mars is innovative. But the novelty of our work comes with the assimilation of spacecraft data into such an assembly of numerical models, to produce a regional (GCM+mesoscale model) reanalysis. This achievement is unique in the studies of the Martian atmosphere. However, performing this for the atmosphere of another Solar System planet, came with a collection of challenges. We first proved the viability of our terrestrial model configuration for Mars' atmosphere, by quantifying the direction of energy transfer (potential and kinetic) between scales from both free-running simulations and assimilated reanalyses. We thus produced the first systematic study of the spectral energy budget of the Martian atmosphere, and concluded that energy flows from large scales to small scales. This can be interpreted as suggesting that information flow also goes from large to small scales, to justify using unidirectional information flow from global to embedded model via boundary conditions. Moreover, we discovered that Mars is missing an enstrophy-dominated energy inertial range (i.e. the k^{-3} range in kinetic energy) that is present in Earth's spectral energetics.

The embedded simulations were carried out for 4 typical Martian seasons around Gale Crater, at a resolution of $80 \times 80 \times 60$ grid-boxes of 5km side-lengths, covering a 400km $\times$ 400km region surrounding the crater. Ulteriorly, the assimilated analysis was interpolated to the location of the Curiosity rover (situated in the crater floor) for comparison with in situ measurements. Different methods of decomposing the measured variability into various time-dependent components were employed. Given that atmospheric tides are directly forced in models and

reality, the correlations from the diurnal and semidiurnal atmospheric tides show almost perfect matching between our analysis and observations. The differences come from uncorrelative effects added by the crater topography and incorrect atmospheric dust and aerosol loading in the models. We finalized our pursuit by explaining both datasets (model and observations) from the superposition of correlations. We concluded that atmospheric dust and aerosol loading is responsible for the underlying shape in the data, with perturbations from superimposing the local meteorology. The tidal components that are not well-matched are important because they can lead to improved models and/or analyses.

Contents

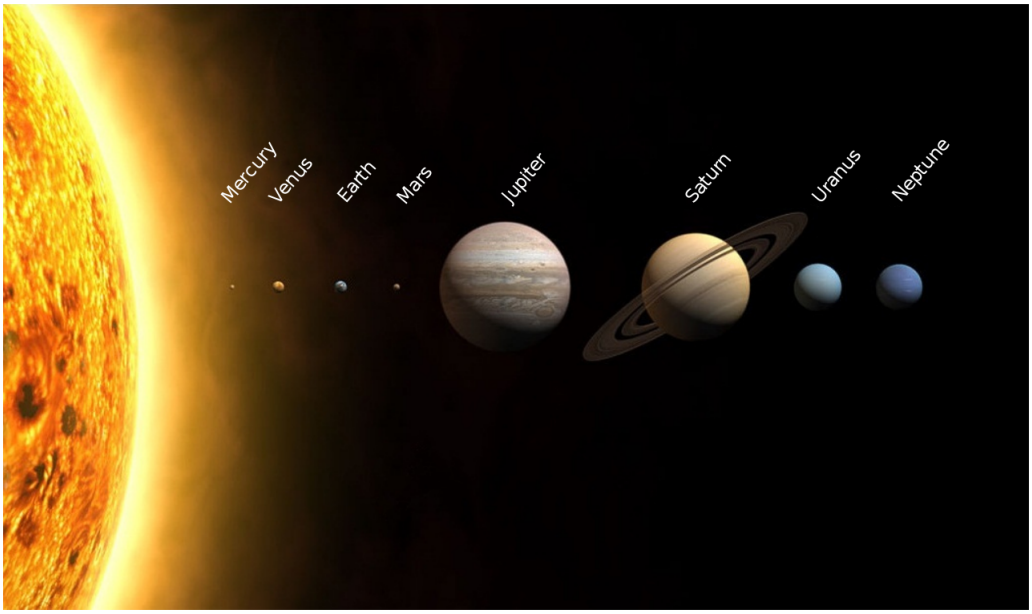


Figure 1.1: The Solar System. All the planets including the Sun are at the correct scale, but orbital-distances are not. ("Planets2013", n.d.)

1 Introduction

The Solar System exhibits quite a remarkable selection of planetary bodies (Figure 1.1) which manifest different structures and environments as dictated largely by the distance to the Sun, the object's spin and its mass. Studying planetary compositions, their geology and atmospheres gives us insight into the Solar System's distant past from the accretion period onwards, as well as valuable knowledge of its present state. For painting a better picture of our Solar System, many spacecraft missions with a wide spectrum of sensors, are currently exploring the neighbouring extraterrestrial objects, with substantial focus on a more recent topic, potential habitability within our Planetary System. The best candidate for habitability, coupled with the current state of our engineering capacity and equipment, is Mars. As a consequence, at the time of writing, there are 8 active missions on Mars with a total of 42 launches to the planet and, as stressed in figure 1.2, many of the recent missions also follow the trend of carrying a lander or rover to its surface. The former USSR, the US, Russia, Japan and ESA (European Space Agency) all

contributed towards Mars exploration (Barlow, 2008 chapter 1.2), though not all focused on acquiring atmospheric data. The plethora of remote sensors cover a wide range of areas of interest including Atmospheric Physics, Biology, Chemistry, Geology, gathering valuable information for adapting future missions and having improved resolution and precision in general. Space missions are getting better at exploring Mars and a new window for planetary science opens with it. One of the recent examples is the Mars Science Laboratory (MSL) Curiosity mission. This successfully landed the Curiosity rover at Mars' Gale Crater on the 6th of August, 2012. Following its main mission of investigating the planet's past habitability, the rover carries a variety of instruments, one of which is dedicated to analysing the atmosphere; this is the Rover Environmental Monitoring Station (REMS). A burgeoning series of research papers followed from this instrument: Pla-Garcia et al (2016), Rafkin et al (2016), Guzewich et al (2016), Martinez et al (2017) just to name a few. REMS is similar to a terrestrial weather station on Mars (more about the instrument in subsection 2.6.2), offering ample opportunity for studying the boundary layer and near-surface of the Martian atmosphere in the vicinity of Gale Crater. Our interests lie on the nomenclature of sensors from the REMS instrument and the UV opacity sensor, as it offers this observer-on-Mars type view of the environment. We intend to analyse the data and deduce the principal physical processes active in Gale Crater using state-of-the-art numerical models. This way Curiosity is no longer blind to its surroundings, but has understanding of the physical processes that take place in the crater. But what does it mean to accurately portray the atmosphere at Gale Crater?

Numerical weather prediction on Mars

The traditional answer comes with the brief history of numerical weather prediction (NWP) (Shuman, 1989, Kalnay et al, 1998) which started with Bjerknes

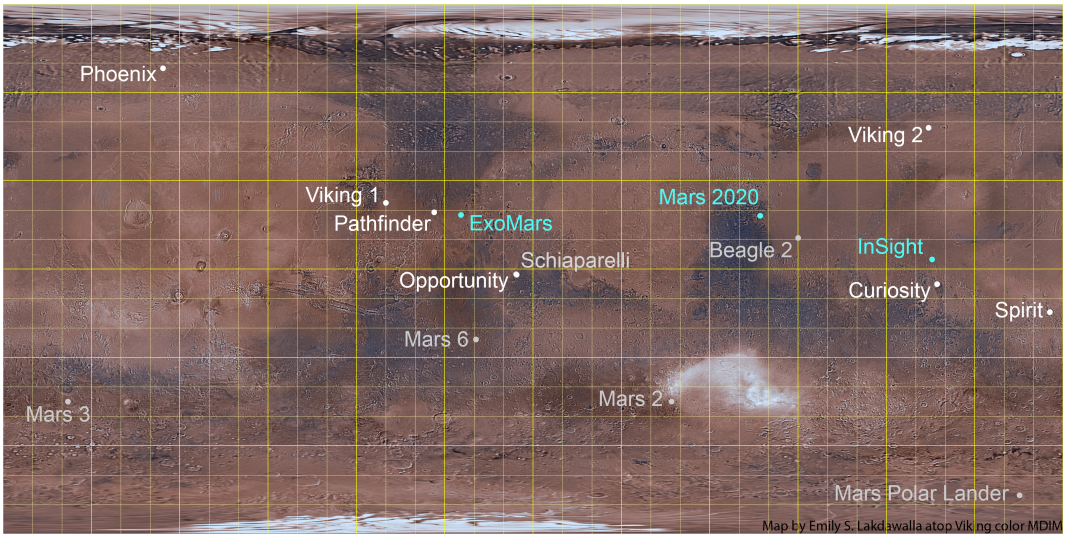


Figure 1.2: A selection of Mars landers placed on a map of Mars. The map contains future missions to Curiosity in blue and the previous missions in white. The missions in faint white are failed missions. ("Map of all Mars landing sites, failed and successful", n.d.)

(1904) (for the importance of his work see Thompson, 1990). He was the first to recognize the same central problem that even state-of-the-art NWP procedures face: to completely forecast the state of the atmosphere one has to start from a comprehensively defined initial state, obtained from observations, and integrate the Navier-Stokes equation, the Boyle-Charles-Dalton equation of state, the continuity equation for mass and the thermodynamic energy equation, subject to properly known boundary conditions. The absence of data processing machines did not stop Bjerknes from putting his theory into application, as he actually received support from the Norwegians to assemble the first NWP program. His design included observation, graphical analysis of the atmosphere and graphically solving the equations. The next and probably the most important milestone in NWP was the paper by Charney et al (1950) and its application to weather forecasting, by numerically solving the primitive equations (known alongside Richardson, 1922, who effectively invented the method of solving PDEs by digital computation). In only six months, the first operational real time NWP program appeared in the

USA. Nowadays, NWP is achieved in a similar fashion but is significantly more accurate due to the improvements in data collecting instrumentation, the advancements in numerical methods and the continuous growth of computational power. Even so, scientists knew that solving the full equations is an impractical approach and that the models need to incorporate only the most important of atmospheric influences; this became a fact especially after the debut of Lorenz's chaos theory. Nevertheless, nowadays the modelling component of NWP is usually done by numerically solving the governing equations through general circulation models (GCMs) and the initial conditions are supplied by data analysis and assimilation techniques.

The study of Mars' atmosphere has thrived with the exploration of the planet's (aforementioned) potential habitability and continues to have crucial importance for understanding its near-surface environment, both now and in its geological past, and for future spacecraft missions. This partly led to the development in Europe of two Mars General Circulation Model (Mars GCM or MGCM, see section 2.1.1) versions: the LMD version (Laboratoire de Météorologie Dynamique) and the UK version (Open University and University of Oxford). The LMD MGCM is currently used for the compilation of the ESA Mars Climate Database (MCD) (Lewis et al, 1999). The UK MGCM is the primary numerical GCM used in the present studies. This GCM code has a recent unique addition from the UK team, which enables it to use data assimilation techniques (Forget et al, 1999, more details are given in section 2.1.1). The approach combines the use of a GCM with spacecraft retrievals of remote sensing data for Mars' atmosphere to produce four-dimensional, multivariate reconstructions of the atmosphere. This technique originates from the operational community in terrestrial meteorology, where they have obtained long-term databases of assimilated atmospheric observations for Earth's climate over periods of decades (this is known as reanalysis in the literature, see section 2.1.1). In the same way, the Oxford and Open University teams

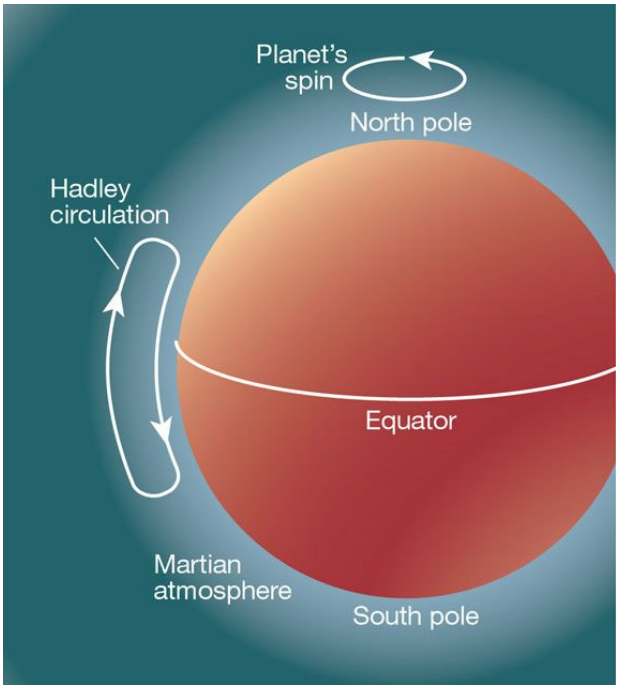


Figure 1.3: A schematic representation of the cross-equator Hadley cell during northern winter. This event takes place twice per year, with the equivalent, but symmetrical, northern summer Hadley cell. The differences between the two cross-equator events are subject to north-south hemispherical dichotomies, coming from orbit eccentricity and variation from the aeroid of the global topography. The general circulation transports atmospheric mass between the two hemispheres, an important feature specific to Mars and not manifested on Earth. (Schematic from Gierasch, 2002)

have begun to achieve a similar goal for Mars and hence we now have the assimilation of the Thermal Emission Spectrometer data (TES) which operates on-board the Mars Global Surveyor spacecraft (MGS) (Albee et al, 2001) and from the Mars Climate Sounder (MCS) instrument on-board the Mars Reconnaissance Orbiter spacecraft (MRO) (Zurek & Smrekar, 2007). Our results are based on the use of the latter instrument. The current "reanalysis" record for Mars extends to ~ 6 Martian years in total ("Mars Analysis Correction Data Assimilation", n.d. or MACDA).

1.1 The general circulation of Mars' atmosphere

Mars has much lower atmospheric density compared to Earth and it lacks heat buffering from a major reservoir of heat, i.e. an ocean system. This results in a low thermal inertia for the Martian atmosphere. Coupled with the low heat capacity of the Martian rocky surface, it makes for a significantly more responsive atmosphere to the thermal solar input, compared to Earth.

Globally, there are two major atmospheric circulations in superposition. First is the thermal tide, which is caused by the selective heating of a spherical object like Mars from a light source such as the Sun (formally known as differential heating) and the planet's spin. This produces a migrating temperature gradient in the whole atmosphere, with the regions of maximum air temperature following the locations on the planet where the Sun is at the zenith (solar radiation is normal to the planetary surface). This crest of the temperature tide advances due to planet spin. Furthermore, the atmospheric wind vector field at the surface tends to follow the maximal temperature region (characterized by low surface pressure) from the surrounding colder regions (high pressure). This gives rise to the so called thermal tide. Another contributor to the general circulation comes from the tilt of the planetary angular momentum vector. This gives rise to an hemispheric dichotomy in the global circulation, with air flowing from the colder to the hotter hemisphere. This particular contribution is Mars specific, occurs during the solstices and is known as the cross-equator Hadley circulation. A depiction of this phenomenon is given in figures 1.3 and 4.4 - left column of plots. During the equinoxes, Mars' atmosphere manifests the double Hadley cell circulation on each side of the equator, a common feature in Earth's atmosphere. This unique seasonal variability transports atmospheric mass, tracers and dust twice per Martian year, from one hemisphere to the other.

Further alterations to the above circulations come from the planet's spin, the

freezing of its atmospheric chemical components and from the occasional dust storm events. The effect of the planet's spin on the global circulation does not do work. It manifests through Coriolis forces acting on the equator-to-pole atmospheric motions, breaking the poleward (meridional) circulation into two Ferrel cells and two polar vortices. The atmospheres of Mars and Earth have different levels of discrepancy. For example, the same Coriolis breaking of meridional motions is found in Earth's atmosphere, but the Ferrel cells are weaker on Mars and the polar vortices manifest a unique seasonal feature, the annular effect of the winter polar vortex (this is actually unique in the Solar System). This effect exhibits a poorly understood annulus type potential vorticity[1] profile with a local minimum near the pole. In comparison, Earth has a monotonically increasing potential vorticity with latitude. Toigo et al (2017) relate this to the second feature which influences the Martian global atmospheric circulation (mentioned above), the condensation of atmospheric CO_2 in the winter polar region. This is yet another special feature of Mars' atmosphere, as CO_2 constitutes 97.32% of the atmosphere's composition, and 3 to 4 trillion tons of it freezes at the polar caps (12% to 16% of the total atmospheric CO_2). Some of the other components in decreasing abundance, are N_2 - 2.7%, Ar - 1.6% and O_2 - 0.13%. The last contributor to the general circulation are dust storm events. This is another unique Martian feature which even made its way in the nickname given to the northern autumn, as "the dust storm season". Multiple dust storm events can occur and even engulf the whole planet in a global dust storm. Some of these contributions

[1]Potential Vorticity or PV is formally defined as $\Pi \equiv \frac{\omega \cdot \nabla \Psi}{\rho}$, where ω and ρ are the usual vorticity and fluid density, while $\Psi = \Psi(\mathbf{x}, t)$ is any arbitrary scalar field. The common choice of Ψ is potential temperature θ, which results in Ertel's PV definition $P \equiv \frac{\omega \cdot \nabla \theta}{\rho}$ which has units of $\frac{K \cdot m^2}{s \cdot kg}$. More specifically, for the annular polar vortex studies mentioned later in this paragraph (Toigo et al, 2017), the scaled PV (Lait, 1994) was used: $PV = \frac{1}{\rho} \times \zeta_a \cdot \nabla \theta \times \left(\frac{\theta}{\theta_0} \right)^{-(1+c_p/R)}$, where ζ_a is absolute vorticity, c_p and R are the usual heat capacity at constant pressure and the universal gas constant, and $\theta_0 = 200K$ is an arbitrary potential temperature.

to the global Martian atmospheric circulation will appear in the MGCM results presented in this chapter, but their formal study based on the MGCM+MMM reanalysis is not the subject of this thesis.

Lastly, the global circulation described here is put over the most extreme topography in the Solar System. This actually has a major impact on all the scales of the atmospheric general circulation, with a global influence on the meridional circulation. The dominant wavenumber 2 global topography forces the mean circulation, intensifying the wavenumber 2 stationary wave (another notable Martian feature). Hence, meridional atmospheric motions may appear opposite to their expected mean (in the MGCM results).

1.2 The motivation for studying Gale Crater

As the crater region is close to the equator, its location motivated Guzewich et al (2016) to analyse atmospheric tides using a mesoscale model designed for Mars, namely the MarsWRF, in comparison to the atmospheric data from REMS. The landing site of Curiosity, unlike the Viking landing sites, is ideal for studying atmospheric tides. This comes from the low latitudes of Curiosity which inhibits the REMS instrument from observing baroclinic waves (in contrast to the high latitude Viking landers). Instead, the meteorology of the region is actively locked to tides. Another aspect is the rich and tight topography of Curiosity's landing site. Gale Crater is not large, with around 154km in diameter, but the topography of its crater features is quite prominent. The central peak, Mount Sharp, is $\approx$ 4.5km above the southern crater floor and $\approx$ 5.5km above the northern crater floor, where Curiosity is approximately located. The crater rim rises to $\approx$ 2km above the north side of the crater floor and the central peak Mount sharp is slightly higher than the south crater rim. All these features are inside a 150km region in diameter, making REMS valuable for studying the effects of topography on the regional

atmospheric circulation, and for studying the smaller scale meteorological effects, such as orographic gravity waves, small scale convection, effects on the planetary boundary layer (PBL) and many others (Pla-Garcia et al, 2016, Rafkin et al, 2016, Kahanpää et al, 2016, and more). Other studies focused on more direct effects of the crater, such as comparing the atmospheric tides in and outside the crater, and concluding that the topography simply amplifies the influence of the circulation outside the crater (Tyler & Barnes, 2013). And many papers focused on the climatology that REMS observes, to try and simulate the environment that the rover encounters (Gomez-Elvira et al, 2014, Rafkin et al, 2016). Other effects include the impact on the local meteorology from anabatic/katabatic wind activities from the crater topography, and anabatic/katabatic winds from the surrounding major landscapes influencing the crater circulation. All these studies are important for understanding the Martian PBL, which in comparison to Earth's PBL, it does not have many meteorological ground stations. In general the Martian atmospheric studies and models are forced to fit quantitatively lower measurements compared to Earth. Nonetheless, the PBL is important for understanding the exchange of heat, momentum, dust and aerosols between the Martian surface and the general circulation. This is particularly important for Mars because direct solar heating induces a "boiling" effect in the atmosphere, lifting the PBL to around 10km on average during the day, compared to the 1km daytime PBL of Earth. This makes the nocturnal low-level jet a general characteristic of the Martian night-time circulation. This ongoing mass of research motivated us to bring a higher resolution reanalysis (through the GCM+mesoscale model technique) to the table, and decisively conduct an analysis of the correlations with the REMS instrument. And since the embedding of a mesoscale model into a global one was extensively used by the scientific community and in conducting the work mentioned above, we took another task into consideration, namely proving that the set-up of two models works for Mars as well (since it is a well-established technique from terrestrial

atmospheric studies). This motivated a full analysis of the spectral energetics of the Martian atmosphere, which will be presented first in this thesis. The aim of this particular project was to understand the exchange of energy between the two main energy reservoirs, the kinetic and available potential reservoirs, and to prove that Mars' atmosphere supports the use of the embedding procedure. The thesis aims to advance the topics presented here (and more), nonetheless, a brief discussion of the technical challenges is relevant (as mentioned), and will follow next.

1.3 Technical challenges for simulating the Martian weather

Studying Gale Crater has proven to be a challenge for the Mars atmospheric modelling community. The reason behind this is the mere scale of the crater, which is smaller compared with typical GCM grids. Our initial work started with boosting and calibrating the MGCM to run at T170, with the unsurprising conclusion that it is impossible to make a pertinent study, even at such a large resolution (typically the MGCM was used at a standard T31 resolution; more about this will will follow at a later point). However, this has not stopped the scientific community from studying Gale Crater, because of its important location on Mars (discussed previously).

In the interim, there is still the question of how we use numerical models to simulate the environment at Gale Crater. The conventional approach is to embed a numerical mesoscale model (MM) into a set of initial boundary conditions. This type of NWP model operates over planetary patches which makes it favourable for local simulations as such a study does not specifically require a global perspective, and all the processing power is thus put directly into simulating the local circulation at the desired site at high resolution. To date, all numerical modelling research on the REMS data has involved the use of a MM. Essentially, the MM

simulates one atmospheric box, where the bottom side is the topography, hence it requires embedding the model into a global simulation by providing the data from a GCM as boundary conditions for the MM (Tyler & Barnes, 2013, Pla-Garcia et al, 2016, Rafkin et al, 2016). Detailed arguments for our GCM+MM set-up can be found in the following chapter (2.1.1), which focuses on the technical description of both models. As a summary, the best resolution that the Oxford global MGCM can produce is not enough to capture the influence from small scale features like Gale Crater itself (e.g. orographic gravity waves). Hence, we used the LMD Mars Mesoscale Model (MMM) with the Oxford MGCM to adopt this same GCM+MM system, but dedicated for Mars. This particular choice of models is sensible as both share same physical parametrizations. The resolution reached with the MGCM+MMM is of 3 km in grid-box lengths, which provide a significant improvement from the $\sim$ 60 km horizontal squares (hydrostatic model) produced by the MGCM alone (at its highest operational resolution; more details about this in 2.1.1). This goes a long way from the previous dataset produced with the MGCM, which was at T31 or almost 300 km in length-scales. The significant increase in resolution is a consequence of the project itself; the need for such a high resolution will be explained in chapter 2. An even better understanding materialises with the use of further weather prediction tools, such as employing data assimilation on the NASA MCS observations. The idea behind this technique is to mathematically extrapolate from the MCS data to Gale Crater using the best association of dynamical core and parametrizations available, and reconstruct the actual weather observed by Curiosity. The implications of such a study are substantial as it would boost our understanding of the connection between the scales; between the microscale observations from REMS (local observer), the mesoscale mechanisms from the MMM (Gale Crater) and the global circulation derived from MGCM reanalysis (from MCS). Additionally, it puts our methods applied for Earth's weather prediction to the severe test, by exposing how well

they perform under Martian atmospheric conditions. The comparison with REMS leads to new insights into the model's representation of Martian meteorology, especially in the boundary layer, which is still a relatively unexplored part of Mars' atmosphere (Rogberg et al, 2010), yet it plays an important role in the surface-atmosphere exchange of mass, heat and momentum. This is in itself an abundant source of novel work, for the reason alone of having the opportunity of doing early investigations of Mars' atmosphere from assimilated data. The MGCM and MMM were initially built to work separately, but also made to work together for this study. Such a tool will have a major impact in estimating the intrinsic limits of the atmosphere's statistically-deterministic property, which will then reveal the feasibility of using such models for more complex spacecraft missions in the future.

1.4 Connecting the scales

Obviously, the main validation of our model-system is the comparison itself with Curiosity. As a validation technique, it serves its purpose (see section 2), however, there is a hidden problem in this set-up. A less superficial way that mathematically explains our set-up, and offers more understanding of its feasibility is required. This is related to the transfer of information (i.e. potential and kinetic energy, vorticity etc.) between the scales, in space and time. Essentially, if there is a causal type of information flow (i.e. energy flux) from microscales, particularly in the boundary layer (where Curiosity is), to the scales of the MCS observations, then our set-up of transferring information from the MGCM to the MMM may be incomplete and unreliable. One may argue that the purpose of data assimilation is to force the atmospheric states of the MGCM to couple with the trend of the MCS observations, independent of the direction of information flow (through the scales). The mesoscale model stage downscales the resolution of the atmospheric states at Gale Crater, hence the MGCM+MMM set-up is just effectively trying to match

MCS' observations, thus rendering the information flow-argument irrelevant. It is only when taking into account the strict one-way influence between the two models that information flow becomes important. If the Gale Crater region somehow injects energy into the atmosphere, which is subsequently propagated to global scales, the set-up will lose information, as the GCM is parametrizing sub-grid scale Physics while the MM would be simulating it in its dynamical core. In the case of Earth's atmosphere, energy is being cascaded to small scales (Lindborg, 2006), making the GCM+MM set-up a widely accepted method in the scientific community. But in the case of Mars, this is unproven as yet. The MGCM was upgraded from the former version's (spectral) resolution of T31L25 to T170L25, in order to couple with the MMM (more details about the need for higher resolutions in sub-section 2.6.2). The demonstration can be provided by computing the first spectrally resolved energetics of the Martian atmosphere.

There are two parts to our study of the scales in spectral space. First, the information gained from such a study is crucial for Mars NWP tools and essential for our novel set-up of simulating from data analysis any volumetric-patch down to 3 km in linear resolution (as already mentioned). Second, there is important knowledge to be extracted from such a study. The computation of the dynamics and pattern of energy distribution across scales for the Martian atmosphere provides a key component in understanding the general atmospheric circulation, in validation of atmospheric models and offers a nice extension to the work on turbulence. Unsurprisingly, the MGCM+MMM set-up passed the information flow test, but for the purpose of its scientific impact, we made a full study of the scales alongside the main course of applying the set-up to Gale Crater, to analyse in situ observations from Curiosity. Attention is thus momentarily moved towards introducing the novel study of the transfer of energy between the scales. Recently, Tabataba-Vakili et al (2015) produced the first Lorenz energy budget for the Martian atmosphere derived from Mars Global Surveyor data for Martian

years 24 to 27. This study was the first to portray the mechanisms governing the general circulation through the framework of energetics and energy reservoirs. Broadly speaking, this is the global scale energy conservation between the globally averaged values for the available potential energy (APE) and kinetic energy (KE) as reservoirs coupled with the averaged injection, dissipation, transfers and conversion of energy. Our objective in the present study is to complete the energy-picture through the multiscale nature of the atmosphere, by rewriting the conservation equations over spherical harmonic coefficients and vertical levels, and also by accounting for the topography in the the formulation of the theory, which is an important factor to consider, especially for Mars. The Red Planet is known to have about half the radius but the same land surface area as Earth (dry land). Mount Olympus, the highest mountain in the Solar System (for example), reaches 22 km above the Martian datum (the datum is equivalent to Earth's sea level, which for Mars is arbitrarily chosen with regard to surface pressure). We will see (section 3.1) that new terms will be created in the conservation of the mentioned fields as a consequence of this method. The extra terms are vertical fluxes of APE and KE, which are relevant for understanding where the injection or dissipation of energy take place; more details in section 3.

This methodology is a continuation of the study done by Tabataba-Vakili et al (2015) using a modified version of Lindborg & Augier (2013). The method is adapted to the Martian atmosphere as it includes the desired mathematics that is previously discussed; more details about their algorithm is presented in the last subsection.

If we are referring to the connectivity between the scales, this study is conveying the first measured energy spectra of Mars' atmosphere and the first Lorenz energy cycle over scales. The spectral energetics analysis is a well known method for understanding both (Frisch, 1995) "microturbulence" (in the boundary layer) and "macroturbulence" at large scales. This is also an important method for

the general circulation of the atmosphere, because of the insight it offers into the way in which the atmosphere transfers energy between scales (some important precedent studies in the scientific community can be found in Frisch (1995), Lindborg (1999), Koshyk (1999), also see 3.4 and appendix D for more details). Nowadays we have a multitude of experiments carried out for Earth's atmosphere, which help us to understand the different patterns we observe in the spectra and fluxes, the energy and enstrophy cascades between atmospheric features of different scales and to have a broader view on atmospheric forcing and dissipation of energy. These mechanisms for the terrestrial atmosphere have a rich theoretical background which started with the work done by Kolmogorov (1941) and Fjørtoft (1953). Therefore, we now possess several heuristic, and very idealized theories which predict: the $k^{-5/3}$, k^{-3} and other energy dependencies in the kinetic energy spectrum (KES), where k is the spectral wavenumber; the "upward" direction of the energy flux to more energetic lower wavenumbers and "downward" enstrophy flux to less energetic high wavenumbers, and as mentioned before, the injection and dissipation of energy with their specific scales. The drive for studying the energy in spectral space for Mars, comes from another convolution of scientific questions.

First of all, we would like to investigate whether there is any evidence for energy or enstrophy cascading inertial ranges in the large-scale and in the mesoscale ranges, and compare to the observations we have for Earth. Additionally, we can formulate quantitatively the spectral energy fluxes and evaluate the validity of the MGCM+MMM system. The spectral energetics exhibit several aspects in common with Earth, like the overall downscale of energy flux and some of the main atmospheric signatures, like Hadley and Ferrel circulations. Remarkably, the atmosphere does manifest the $k^{-5/3}$ law, but only in the regions of the atmosphere which are not pierced by the topography. The calculation of the vertical energy fluxes indeed shows clear energizing in the boundary layer on Mars, which sub-

sequently reaches the upper levels. Here it enters an inertial range described by downscale energy cascade and conversion to kinetic energy. Similar to Earth, these results have a seasonal pattern and also reveal the global dust storm's energetics. This poses more questions, such as: where do the similarities and differences with Earth's atmosphere come from? Why is Mars missing the k^{-3} range? This would greatly boost our understanding of the mechanisms we observe in Mars' general circulation and turbulence of the atmosphere. Furthermore, we would like to progress to analysing the mesoscale range and similarly observe if there is a change in the type of inertial range as observed on Earth. If so, is it associated with the transfer of energy upscale or downscale? Is it dominated by divergent or rotational components of the flow? Before we even search for answers, we know that Mars is in a different regime to Earth, so even after we compute the highest wavenumbers, we might expect interactions between scales to differ from those we see on Earth. Considering the fact that there is no well developed theory for the spectral energy budget in the case of Mars, these questions alone are a requisite for similar background-models as it has been achieved for Earth.

So *why do we want to study the atmosphere in spectral space*? The answer is clear once we look at the spherical harmonics in Figure (1.4). First of all, the energy spectrum we study, is represented by the sums of all energies with the same l index. And as we see from the figure, these sums represent all those energies of features with the same wavenumber pattern. For example, $m = 0$ in the spectrum selects atmospheric features of length-scales of the order of the whole planet in the longitudinal direction (i.e. zonally-symmetric structures that vary only in latitude). This is of course not the same in the latitudes, where as l increases, latitudinal features get richer. But in the end, atmospheric processes, like diurnal variations, zonal winds and even microscale features are longitudinally varying processes. Secondly, we have a rich theoretical background of how the spectral-energies should depend on the l index, as we will see in the following sections.

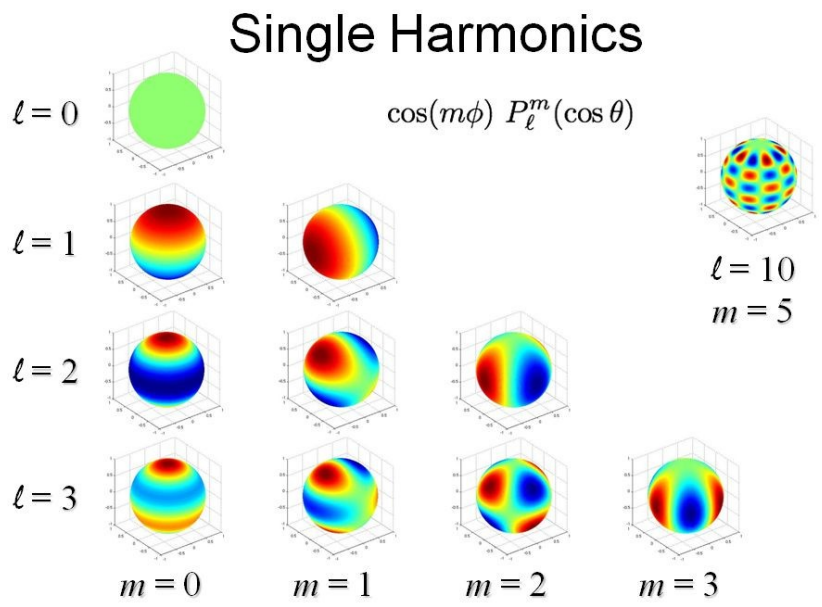


Figure 1.4: Spherical density plots of the spherical harmonics for degrees l raging from 0 to 3 with all the positive orders m. ("Earth's Magnetic Field", n.d.)

The shape of the energies with l depends on many mechanisms that develop in the atmosphere, and by analyzing this l-index to energy mapping, we get insight into how and why the processes occur (we will see this in the next sections).

Depending on the scales we are analysing, certain processes exhibit their mark on the kinetic energy spectrum. Their effects can be classified into injection/dissipation of energy and "non-interaction" processes or cascading of energy. These are fundamentally different mechanisms, in that they can be described by different dynamics. The inertial ranges are regions of *well-behaved turbulence* where energy flows through the scales and is best described by turbulence theory. The injection-dissipation of energy into the kinetic energy spectrum, in contrast, is inferred mathematically through something as traditional as the conservation of energy. Hence a good definition of the energy sources (i.e. the Sun) and the other energy reservoirs (i.e. Potential Energy) along with the energy conversions, transfers and of course the energy fluxes (turbulence cascades) is sufficient to lay out the mathematical basis for the spectrally resolved energetics of the atmosphere.

The energy conservation for the globally averaged quantities of the atmosphere was first theorized by Lorenz (1955).

Why study the energy spectrum of Mars' atmosphere? For many decades, even the energy spectrum for Earth was impossible to analyze. The problem came from the difficulty of synchronously gathering data from the whole atmosphere. Only with the introduction of the Global Atmospheric Sampling Program (GASP), did scientists begin to get a full picture of the energy spectrum. Nowadays, with the new advancements in numerical-modeling (i.e. GCMs), we can assimilate chunks of data from satellites and reconstruct the whole atmosphere (we have seen this in section 2). Thus, with the increasing number of satellites in orbit around Mars, it has recently become feasible to reconstruct an instantaneous synoptic structure of the martian atmosphere and ultimately its energy spectrum. The formal answer to this question is tied to the first question. The energy spectra are a rich source of information and more importantly, the atmosphere on Mars is a new "black box" to be opened, and "testing-site" for more exotic atmospheric processes to be studied.

1.5 A brief overview of the thesis

The results of the thesis are structured in 3 main parts detailing the spectral energetics of the Martian atmosphere first (chapter 3), followed by the MGCM+MMM configuration and reanalysis (chapter 4), and finalising with the Fourier/SSA time variation analysis in connection to the reanalysis (chapter 5). Each of these chapters is self contained, in that it starts with a theoretical introduction on the topic of the chapter and continues with results and interpretations. The appendix offers more background on the theoretical side of some of the chapters and the bibliography can be found in the "References" section 7, preceding the appendix. The motivation behind the structuring of the results chapters comes with the compo-

sition of the thesis. It essentially envelops 3 broad, but interconnected projects which are spacious enough, that introducing the theories in the current chapter proved to be inadequate for the structure of the thesis. This is the reason for the addition of an appendix to the thesis. Therefore, each results chapter will introduce the topics in connection to the literature, prepare the results side, and present our findings and interpretations. Before the self-contained sections, there is an additional chapter introducing the atmospheric circulation models (the MGCM and MMM) and the data sources (MRO-MCS and Curiosity-REMS). This is chapter 2, following next.

2 Tools and methodologies

The present chapter describes the main numerical methods in connection with the projects that were conceptualized in the introduction. The synopsis includes a light introduction of the mathematical basis for our MGCM, data assimilation and MMM, adjusting adamantly to the particularities and necessary adaptations for our methodology. This compilation is structured in two main parts. The first part (2.1) is a description of the dynamical cores of the Mars GCM (MGCM) (2.1.1) and the Laboratoire de Météorologie Dynamique (LMD) Mars Mesoscale Model (MMM; Spiga & Forget, 2009) (2.1.2) coupled with their similar physical cores (2.1.3). In short, the development of the two models share their physical parametrizations because the same MGCM is researched by the LMD and the UK groups in parallel. The difference between the two global models stands in their dynamical cores, the UK MGCM being a pseudo-spectral model in its horizontal discretization, using a spherical harmonic decomposition of the key variables at each model level while the LMD MGCM is a non-hydrostatic real-grid model; the common ground being in their physical parametrizations. Thus, the LMD MMM also has common ground with both the UK and LMD MGCM through its physical parametrizations; more details are provided in the current chapter, but essentially the UK MGCM (or LMD MGCM for this matter) is a best suited candidate for providing boundary conditions to the LMD MMM.

The second part summarises the components of the Data Assimilation (DA) scheme which drives the models through the window of the Mars Climate Sounder (MCS) retrievals coupled with a section on the combined models and outputs. This characteristic of adjusting to spacecraft observations is solely attributed to the MGCM and propagates through their interface to the MMM. To end this chapter, there is a return to the fundamentals of using the output of a global model as boundary conditions to a mesoscale model. Such set-up is well known for

modelling Earth's atmosphere and will be proven viable for Mars in the following chapter. The focus now is on the particularities of our models in the framework of numerical methods.

2.1 The Numerical Models

In order to accurately render Mars' general circulation and exotic planetary boundary layer (PBL) a realistic model was needed, with the best candidate being the UK version of the LMD MGCM (which will be abbreviated MGCM) for its ability to assimilate satellite data. To deliver as much detail as computationally possible for comparison with REMS, we opted to embed the MMM into the output from the MGCM centred around Gale Crater, to spatially downscale the assimilated reanalysis. For such a feat, the MGCM was updated to assimilate satellite data at the spectral resolution T170 (keeping the same number of vertical levels L25) so that its output is feasible for driving the MMM by the provision of high resolution boundary conditions. The interface between the MGCM and MMM automates the use of both models, reducing the time spent on producing a full MGCM+MMM reanalysis to only computational (thus eliminating the time spent on cumbersome manual connections between the models). All the relevant features of the models and their similar physical packages are delivered subsequently.

2.1.1 The Mars Global Model dynamical core

The present section provides a brief introduction to the version of MGCM we exploit in our study. This model is a GCM for the Martian atmosphere produced, upgraded and extensively used between four groups: the Laboratoire de Météorologie Dynamique (LMD) from Paris, the Open University and the University of Oxford from the UK, and the Instituto de Astrofísica de Andalucía from Spain.

The dynamical core of the model is based on an adiabatic, multi-level, primitive-equation atmospheric idealization which uses spherical geometry with representation of the fields at each level as truncated series of spherical harmonics. The dynamics have endured minor changes since their implementation, to account for the evergrowing computational power and the progression of programming features in the computer science community.

Many studies within the groups have investigated the impact of certain physical processes such as radiation, cloud microphysics or dust, on the general circulation; these aspects were coded alongside the dynamical core and were delivered to the MGCM in the form of "physical packages". The packages constitute the physical core of the MGCM and are used in common by both the UK and LMD models, with one important difference. The UK version has its own dynamical core founded on the ideas from Hoskins & Simmons (1975). The linear dynamical operations are all implemented in spectral space as opposed to the finite difference real space used by the LMD; however, the nonlinear operations and physical parametrizations are computed in real space, hence the MGCM is referred to as a pseudo-spectral model. Its upgrades consist in an energy and angular-momentum conserving vertical finite-difference scheme, a semi-Lagrangian tracer advection scheme and a data assimilation scheme. We will deal with formal aspects of the spectral MGCM method in sections 3, 4 and 5. A full exposition of these aspects is not pertinent to the subject of this work, although an introduction is almost imperative. Accordingly, the basic dynamical core is described in the following subsection (2.1.1) and a simple description of the physical packages is provided afterwards (2.1.3). The aim of this section is to clarify the approaches behind our projects, which are systematically rooted to the structure of the MGCM.

Dynamical Core

The crux of the dynamical core is a code written in accordance with the method published by Hoskins & Simmons (1975). A list of GCMs and an accurate description of the method with results are presented in their paper. The model predictions are generated based on the non-dimensional equations of motion for an inviscid, adiabatic, hydrostatic, ideal gas covering a rotating, spherical planet. These are

$$\frac{\partial \zeta}{\partial t} = \frac{1}{1-\mu^2}\frac{\partial}{\partial \lambda}\mathcal{F}_V - \frac{\partial}{\partial \mu}\mathcal{F}_U, \tag{2.1}$$

$$\frac{\partial D}{\partial t} = \frac{1}{1-\mu^2}\frac{\partial}{\partial \lambda}\mathcal{F}_U + \frac{\partial}{\partial \mu}\mathcal{F}_V - \nabla_\sigma^2\left(\frac{U^2+V^2}{2\left(1-\mu^2\right)} + \phi + \bar{T}\ln p_s\right), \tag{2.2}$$

$$\frac{\partial T'}{\partial t} = -\frac{1}{1-\mu^2}\frac{\partial}{\partial \lambda}(UT') - \frac{\partial}{\partial \mu}(VT') + DT' - \dot{\sigma}\frac{\partial T}{\partial \sigma} + \kappa\frac{T\omega}{p}, \tag{2.3}$$

$$\frac{\partial \ln p_s}{\partial t} = -\boldsymbol{V}\cdot\boldsymbol{\nabla}_\sigma \ln p_s - D - \frac{\partial \dot{\sigma}}{\partial \sigma}, \tag{2.4}$$

$$\frac{\partial \phi}{\partial \ln \sigma} = -T. \tag{2.5}$$

where $\mathcal{F}_U = V\zeta - \dot{\sigma}\dfrac{\partial U}{\partial \sigma} - T'\dfrac{\partial \ln p_s}{\partial \lambda}$ and $\mathcal{F}_V = -U\zeta - \dot{\sigma}\dfrac{\partial V}{\partial \sigma} - T'(1-\mu^2)\dfrac{\partial \ln p_s}{\partial \mu}$, ζ is the absolute vorticity, D the divergence, $T = \bar{T}(\sigma) + T'$ the temperature, p_s the surface pressure, ϕ the geopotential (sum of the potential energy for the attractive force and the centrifugal potential energy), σ the vertical coordinate $\sigma = p/p_s$, λ the longitude and $\mu = \sin\theta$, where θ is the latitude. U and V are the longitudinal and latitudinal velocities respectively, multiplied by $\cos\theta$. $\boldsymbol{\nabla}_\sigma$ is the del operator in the horizontal (i.e. along sigma surfaces). Thus, the horizontal advection term is $\boldsymbol{V}\cdot\boldsymbol{\nabla}_\sigma = \dfrac{U}{1-\mu^2}\dfrac{\partial}{\partial \lambda} + V\dfrac{\partial}{\partial \mu}$. The conversion to the SI form of these equations can be obtained by using the radius of the planet a for length scales, the inverse of the angular velocity Ω^{-1} for time scales, the factor $a^2\Omega^2/R$ where R is the gas constant for temperature scales and $p_s = 600Pa$ for pressure scales (which corresponds to the mean surface pressure on Mars; Hourdin et al,

1993). Furthermore, σ values are used for separating vertical layers. Equations (2.1), (2.2) and (2.5) are derived from the momentum equations; (2.3), (2.4) and (2.5) follow from the thermodynamic and continuity equations. These are referred to as the equations of motion or EOMs.

An equivalent way of decomposing the flow into vorticity and divergence is to use a streamfunction ψ and velocity potential χ - a device originating with Helmholtz in the 19th century (Koenigsberger, 1906). The velocity components are also calculated from ψ and χ. These connections may be written as

$$U = -(1-\mu^2)\frac{\partial \psi}{\partial \mu} + \frac{\partial \chi}{\partial \lambda}, \tag{2.6}$$

$$V = \frac{\partial \psi}{\partial \lambda} + (1-\mu^2)\frac{\partial \chi}{\partial \mu}, \tag{2.7}$$

$$\zeta = 2\mu + \frac{1}{1-\mu^2}\frac{\partial V}{\partial \lambda} - \frac{\partial U}{\partial \mu} = 2\mu + \nabla_\sigma^2 \psi, \tag{2.8}$$

$$D = \frac{1}{1-\mu^2}\frac{\partial U}{\partial \lambda} + \frac{\partial V}{\partial \mu} = \nabla_\sigma^2 \chi. \tag{2.9}$$

Even so, the problem of integrating equations (2.1)-(2.5) is highly non-trivial. The right hand side manifests coupling of variables through non-linear differential operators. The procedure for treating these non-linear terms comes with the requirements of working in spectral space. With a reduced number of exceptions, a smooth function on a sphere f at each moment in time can be expanded as a superposition of spherical harmonics Y_n^m

$$f(\lambda, \mu, t) = \mathcal{N}_f \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_n^m(t) Y_n^m(\lambda, \mu), \tag{2.10}$$

where $f_n^m(t)$ are the expansion coefficients at time t and $Y_n^m(\lambda, \mu) = \tilde{P}_n^m(\mu) e^{im\lambda}$ are the spherical harmonics which are related to the associated Legendre functions and $\mathcal{N}_f$ is the normalization constant.

A pure spectral method integrates only the coefficients in spectral space as

used in fluid dynamics. Atmospheric GCMs , instead, use both the values at the grid points and the coefficients in spectral space for calculating nonlinear terms and transforms these values from different spaces at each time step. Since the introduction of the transform method (Orszag, 1970; Eliassen et al, 1970), there has been an increased interest in the use of truncated series of spherical harmonics for horizontal variations of functions on a sphere. On these lines, ζ, D, T', ϕ at each level and $\ln p_s$ are expanded, with the normalizing constants derived from the scaling factors defined above.

The resolution of the model arises from the truncation brought to these expansions. There are several approaches to truncating this sum, depending on what advantages we want to exploit with our model. In our studies we have used the *triangular* truncation, which has the useful property of invariance under rotation. This truncation is denoted by symbol T in the literature and it works by limiting the sum over n in equation (2.10) or in other words, by actually limiting the wavenumbers included. Naming the maximum wavenumber as n_{max}, the resolution of a triangular truncated GCM based on spherical harmonics is denoted by Tn_{max}. In our investigations, we have used $T31$ and $T85$. For example, at $T31$, the harmonic Y_{31}^{31} contains 31 wavelengths in the region $\lambda \in [0, 2\pi)$ (around the Equator), thence, the smallest atmospheric simulated feature can be deduced from the highest wavenumber in the truncation; in the case of Mars, this length scale is 686 km at the equator, or 343 km formal resolution. The technical details of the spectral core are relevant for our first result of the kinetic energy spectrum described at the beginning of the next chapter (3), but are not relevant for the story of this thesis; thence more details on the passing of the variables in the restart system of the MGCM and the resolution are provided in appendix C.

2.1.2 The Mars Mesoscale Model dynamical core

As specified previously, the connectedness between the two models, the UK MGCM and LMD MMM, is supplied by the physical cores. The reason for this feature is the nature of how the MMM was developed, by combination of two particular cores. The dynamical core is adapted from the fully compressible nonhydrostatic Advanced Research Weather Research and Forecasting (ARW-WRF) model developed by the joint efforts of the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) (Skamarock, 2005) for the terrestrial atmosphere. Anticipating from previous parts of this chapter, the physical core is a result of the coupling of the LMD MGCM version and the ARW-WRF, making all three models (the UK MGCM and the two LMD models) share physical parametrisations. More about this in the next subsection.

Attention is presently shifted towards the MMM dynamical core. As its characteristics are imported from the ARW-WRF dynamical core, it is a non-hydrostatic real-grid idealization which constructs the physical fields at each time-step by spatio-temporal integration of the conservative flux-form formulation of the fundamental equations, within the region of interest, from the boundary conditions provided (as is traditional for mesoscale models). The assumption of hydrostatic balance employed by the MGCM is obviously not viable for the MMM dynamical core. At scales where the gravitational acceleration is comparable to the vertical wind accelerations, e.g. when mesoscale convection and convective vortices are resolved, the atmosphere can not be approximated to a thin spherical shell and nonhydrostatic equations are necessary. The conveniently conserved fields from the choice of EOMs-form are tracer, mass, momentum and entropy. The coordinate system of the model is a (x, y, η, t) basis, where x, y and t are the normal horizontal spatial and temporal dimensions and

$$\eta = \frac{p_h - p_{ht}}{p_{hs} - p_{ht}} \tag{2.11}$$

is the terrain following mass-based vertical coordinate which returns to flat surfaces in the upper atmosphere, where p_h is the pressure at level h, and p_{ht} and p_{hs} are the pressures of the top and bottom layers respectively. This coordinate system was proposed by Laprise (1992) and holds many useful features. Firstly, a quick simplification, like nullifying the top pressure $p_{ht} = 0$, leads to the classic $\sigma = p_h/p_{hs}$ levels that the MGCM uses. Secondly, it forces the use of the flux form physical fields

$$\mathbf{V} = \mu\mathbf{u} = (U, V, W)\,, \quad \Omega = \mu\dot{\eta}\,, \quad \Theta = \mu\theta\,, \tag{2.12}$$

where $\mu(x, y) =$ is the mass per unit area within the vertical model column at (x, y), $u = (u, v, w)$ is the covariant velocity vector, $\omega = \dot{\eta}$ is the contravariant vertical scalar velocity and θ is the traditional notation for the potential temperature. By infusing the non-conserved variables $\phi = g$ for the geopotential, p for the pressure and $\alpha = 1/\rho$ for the inverse of the volume (density-inverse) (Laprise, 1992), we can write the flux-form equations

$$\partial_t U + (\nabla \cdot \mathbf{V}u) - \partial_x(p\phi_\eta) + \partial_\eta(p\phi_x) = F_U \tag{2.13}$$

$$\partial_t V + (\nabla \cdot \mathbf{V}v) - \partial_y(p\phi_\eta) + \partial_\eta(p\phi_y) = F_V \tag{2.14}$$

$$\partial_t W + (\nabla \cdot \mathbf{V}w) - g(\partial_\eta p - \mu) = F_W \tag{2.15}$$

$$\partial_t \Theta + (\nabla \cdot \mathbf{V}\theta) = F_\Theta \tag{2.16}$$

$$\partial_t \mu + (\nabla \cdot \mathbf{V}) = 0 \tag{2.17}$$

$$\partial_t + \mu^{-1}[(\mathbf{V} \cdot \nabla\phi) - gW] = 0 \tag{2.18}$$

accompanied by differential equation in η levels

$$\partial_\eta \phi = -\alpha\mu, \tag{2.19}$$

which is also the hydrostatic balance diagnostic equation, and the equation of state

$$p = p_0(R_d\theta/p_0\alpha)^\gamma. \tag{2.20}$$

Here, p_0 is the reference mean surface pressure (i.e. 700 Pa), $\gamma = c_p/c_v = 1.4$ is the generic adiabatic constant and R_d is the gas constant for dry air. The differentiation is defined as

$$(\nabla \cdot \mathbf{V}f) = \partial_x(Uf) + \partial_y(Vf) + \partial_\eta(\Omega f), \tag{2.21}$$

and

$$(\mathbf{V} \cdot \nabla f) = U\partial_x f + V\partial_y f + \Omega\partial_\eta f\,, \quad \forall f \,\text{a scalar field}. \tag{2.22}$$

Additionally, the terms F_U, F_V, F_W and F_Θ from the RHS of 2.13 to 2.16 have the role of forcing the model and incorporate influences from the turbulent mixing, model physics and Mars' rotation, as well as projections on the sphere. The latter is somewhat ambiguous, as the forcing is expected to be physical (and not an abstract projection). In altitude, the projection is on η levels, while horizontally, the ARW-WRF, and by extension the MMM, comes with 3 types of planetary projections which are latitude-dependant. As the dynamical core is performing integration directly in real space, a GCM with this feature would have the traditional poles problem for constant angular grid separation of its grid-boxes (in latitudes-longitudes) (Randall, 1992). Usually real-grid dynamical cores deal with this problem by separating the sphere in unconventional ways, such as using

icosahedral grids (Randall, 1992) or conformal-cubic (McGregor and Dix, 2008). Thus, the elimination of the latitude-singularity at the poles is done by additional computations to convert from the atypical grids to latitudes and longitudes. Fortunately, the MGCM dynamical core works in spectral space, hence it integrates directly to a Gaussian grid which automatically avoids this problem. In the case of the MMM, a full global separation into grids is unnecessary, hence it uses different mappings depending on the latitude. This custom grid-box separation is translated through scaling factors to guarantee a regular grid definition into the prognostic equations (2.13) - (2.18). The three available projections are: Mercator, for the regions around the equator, Lambert Conformal for mid-latitudes and Polar Stereographic for Polar regions. Regarding the main project of this thesis, as Gale Crater was the only priority, the mapping of all our outputs from the MMM are in Mercator, hence straight lines from outputs are also straight on the surface of the sphere ("Mercator projection", n.a.). To finalise the description of the MMM's spatial discretisation, the computations supply grid boxes in the Arakawa C-grid staggering style, meaning that the velocity vectors are half a grid-box away from the thermodynamic fields, both in the horizontal and vertical.

Attention is now moved towards the time dimension, where the integration is accomplished using an adaptive third-order Runge-Kutta method (Sukumar, 2014). Essentially, there are two types of time integration in the dynamical core. The main Runge-Kutta method is used for the lower frequency modes (or slow processes) that are important for the circulation. For the high frequency modes an acoustic integration is performed, with a much lower time step to account for the meteorologically unimportant processes that only compromise numerical stability. The Courant Friedrichs Lewy (CFL) condition is thus satisfied for fast processes, while the computation of slow processes does not impair processing power. A detailed description of the acoustic integration method is unnecessary, hence for more information, it is advisable to visit the NCAR technical notes from 2008

on the ARW-WRF version 3. The two integration instances are accomplished using the time-split integration technique. Each time step from the Runge-Kutta method embodies a full iterative procedure for the acoustic method. The inclusion of the acoustic integration lifts the time step limitations of the main Runge-Kutta method and, to guarantee stability, additional filtering for acoustic modes damp residual instabilities which are suspected to arise from the acoustic integration itself.

2.1.3 Physical Parametrizations

Both the MGCM and MMM are intended to provide realistic simulations of the Martian atmospheric circulation for comparison with atmospheric measurements, so it is important to capture many of the causal sources that make the atmosphere of Mars the way we observe it. Realism is achieved either through generalizing the equations of motion (EOMs) and increasing the spatio-temporal resolution of the dynamical cores, or by the addition of a library of physical parametrizations that represent a range of physical processes in the model atmosphere. The latter solution is more viable, because the additional computational power required for better integration (by the dynamical cores) is replaced by better mathematics (through a physical parametrisations); this is termed as the "gray zone" in the atmospheric modelling community. Hence, the MGCM physical parametrisations modify the integrated fields to include the additional Physics or methodology that is not accounted for in the integration. Depending on the nature of the added schemes, the physical parametrizations can blend in sub-grid scale phenomena such as boundary layer mixing and turbulence, cloud microphysics and dust lifting processes, with resolved (but approximated) physical processes such as radiative transfer and associated heating and cooling. This subsection goes into some of the more relevant physical parametrizations. For a full account of the physical parametrizations, visit (Fita, 2013).

Tracer advection

Preferentially, the models can also simulate dust, water vapour and ice, which are made available through the tracer advection updates. These are implemented using the semi-Lagrangian advection scheme from Newman et al (2002), coupled with the Priestley (1993) scheme for mass conservation. The scheme infers a log-normal distribution for particle sizes, the number density of particles being

$$n(r) = \frac{N}{\sqrt{2\pi}\sigma_0 r} \exp\left(-\frac{\ln^2(r/r_0)}{2\sigma_0^2}\right) \tag{2.23}$$

where r_0 is the geometric mean radius, σ_0 is the standard deviation of the distribution (σ_0 is kept constant), N is the number mixing ratio of particles (i.e. total number of particles per unit mass of atmosphere) and $n(r)dr$ is the number of particles per mass with sizes in the range $[r, r + dr]$.

Radiative transfer

Similar to the previous package, absorption and scattering of radiation by dust and water ice, as well as emission and absoption of radiation by CO_2 can also be simulated using the radiative transfer scheme. These are implemented based on Forget et al (1999) and Madeleine et al (2011). In our T31, T85 and T170 studies we have used a prescribed dust opacity given by the MCS opacity retrievals.

Convection and diffusion

As already mentioned, the dynamical core of the MGCM is fully hydrostatic (no vertical velocities). In order for the model to keep the tracer temperature profile at the adiabatic lapse rate, a convective adjustment scheme has been implemented (Hourdin et al., 1993) which rapidly mixes the unstable layers, redistributing any tracers contained within them in the vertical. Diffusion of momentum, potential temperature and tracers between vertical layers are modelled using the classical

diffusion equation

$$\frac{\partial A}{\partial t} = \frac{1}{\rho_a}\frac{\partial}{\partial z}\left(K\rho_a\frac{\partial A}{\partial z}\right), \tag{2.24}$$

where A represents any tracer or variable, ρ_a is the atmospheric density, K is the diffusion coefficient (computed according to Mellor & Yamada (1982) in a 2.5-level turbulence closure scheme - Forget et al, 1999), z is the altitude and t is time. This scheme has been upgraded to not just adjust the dry adiabat. So lapse rates larger than the dry adiabatic lapse rate are possible now.

Other packages

The water cycle and cloud microphysics (Montmessim et al, 2004), gravity wave drag scheme (Miller et al, 1989), resolved topography (Mars Orbital Laser Altimeter or MOLA data: Zuber et al, 1992, Smith et al, 1998), condensation and sublimation of CO_2 scheme from Forget et al, 1998, etc. The orographic Gravity Wave Drag (GWD) scheme is an important aspect of the high resolution model upgrade, because at T170 it was suspected that the dynamical core will start resolving them. Thence, a brief discussion is given later in the chapter, and an in-depth background in appendix B.

2.2 Data Assimilation

As the data assimilation (DA) scheme is a crucial design for the reanalysis, this section briefly details its characteristics. Moreover, the scheme is arguably the most important addition to the MGCM, as it represents a unique kit originating with the UK team. The DA scheme acts in the same way as the parametrisations of the physical packages, by adjusting the fields from the dynamical core. In this particular scheme, the model is also pushed to mimic observations from regions where satellite data exists and force the rest of the atmosphere from the observed

patches onwards. The outcome gives a complete spatial and temporal coverage of the observational data (commonly known as "reanalysis"), similar in a sense to interpolation but more accurately consistent with both observations and the physics, and the dynamics represented in the model; but more importantly, it fills the gaps outside of the observed interval (where there is missing data). Briefly, assimilating data is usually achieved by doing analysis cycles. At each cycle, observations of the system at the current or past model-timestep (depending on the data-gathering frequency), are combined with the predictions of the model itself. These produce a 'best estimate' of the current state of the system and it is recognized in the literature as an *analysis step*. The modified model state is then advanced to the next time step, where its results become the initial conditions for the next cycle.

The implementation of the DA scheme in the MGCM is based on the analysis correction (AC) method, developed and used by the Met Office (MO) (Lorenc et al, 1991) between 1989-2000 (and was still being used for ocean forecast until 2012). The MO developed this scheme from the need to replace the older Optimum Interpolation (OI) scheme. They required a faster and much simpler DA algorithm which essentially did the same repeated insertion of observational data (Bell and Dickinson, 1987). The repeated insertion approach of the AC DA was kept similar to the OI scheme, but the analysis correction component had a completely new rework. Because of the similarity of the OI's successive correction method (Bergthorsson and Döös, 1955; Cressman, 1959) the new method was named Analysis Correction. The underlying philosophy of the DA method for numerical weather prediction (NWP) consists of finding the initial conditions for each iteration that *best* fit the observations. This optimal forecast has an additional requirement of consistency with the prior knowledge of the atmosphere. The pursuit for the optimized model state has a variety of numerical adaptations in the literature. For a more complete list of the methods and a proof of their

equivalence see Lorenc (1986). This discussion continues in the appendix with the mathematical derivation of the idealized equations; which are subsequently approximated and streamlined to match the AC DA scheme used in our MGCM. This chapter does not go into further detail on the theoretical approach of DA AC method as it is not the subject of this thesis, however, because it is the crux of our reanalysis, the mathematical derivations building up towards the scheme and features that we used to assimilate MCS data, are provided in appendix A (namely the analysis equations and assimilation procedure).

2.3 The Embedding Procedure

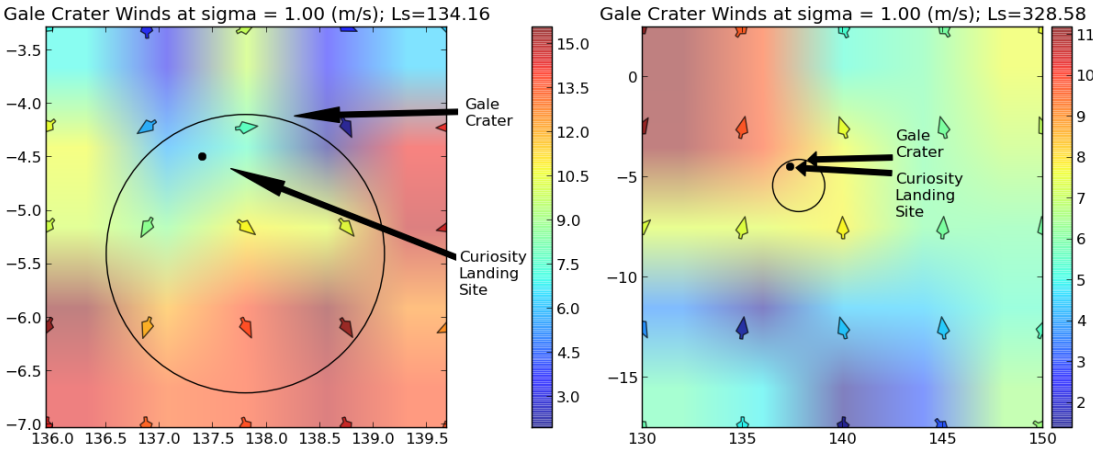


Figure 2.1: The grid of velocities around Gale Crater, from a horizontal layer. The first plot on the left depicts the unit velocity array from a T170L25 run and the plot on the right corresponds to the standard T31L25 MGCM resolution. The corresponding real grid resolutions are: T170→ $384 \times 192 \times 25$ and T31→ $72 \times 36 \times 25$. Their different densities can be interpreted from these arrays. The horizontal coordinates are in latitudes×longitudes and the colour bar is in units of m/s.

The downscaling system is accomplished by initiating the MMM states with the time-dependent boundary conditions from the MGCM, that are by definition lower in resolution compared to the MMM's. The supply of boundary conditions is repeated unidirectionally, which partly motivated the enhancing of the MGCM

resolution. Therefore, the resolution was firstly increased to

- T170L25 spectral resolution, corresponding to a real-grid resolution of $384 \times 192 \times 25$ pixels or 1 grid every ≈ 0.937 of a degree, from its natural
- T31L25 spectral resolution, corresponding to $72 \times 36 \times 25$ or 1 grid-point every 5 degrees.

Their corresponding densities can be deduced from figure 2.1. The reason for only doing case studies comes from the high resolution. Increasing the number of spherical harmonics in the expansion of the physical fields causes a linear increase in the gaussian resolution approximately obeying the relationship $\equiv (3/4) \times ((3T+1)/2+1)$ (see (C.11) from the appendix). So even if the direct relationship between the spatial resolution and computational time is non-intuitive, the CFL condition makes this at least quadratic, as the size of the time-step also has to decrease by the same order as the gaussian resolution to preserve the same causal propagations. For example, 480 dynamical steps per sol is enough for a T31 reanalysis, which has a real-grid resolution of 72×36. This means that for a T170 assimilation, with 384×192 real-grid resolution, requires at least 2560 steps or a model computation each $\approx$ 30 Mars seconds (or $\approx$ 34.68 Earth seconds). In effect, the growth in resource usage permits only case studies at T170.

To obtain the initial conditions, we first produced a T31 reanalysis of the full MCS v4.2 temperature data, with a spin-up period of 400 sols beforehand. The aimed solar longitudes for our MGCM+MMM reanalysis cases were Ls=$\{0, 90, 180, 270\}$, which correspond to the beginning of each season. We limited our analysis to Curiosity's first operational Martian year of REMS measurements, due to the computational power required and the availability of scientific papers covering the first operational year of REMS. Particularly, Rafkin et al (2016) and Pla-Garcia et al, (2016) went through the same endeavour of doing high-resolution regional simulations of Gale Crater, so they only opted for seasonal 6-7 sol case

Operational sol	Ls	Northern season	MY
53	180	autumn	31
196	270	winter	31
350	0	spring	32
543	90	summer	32

Table 1: A table of corresponding Curiosity operational sols, solar longitudes, seasons and Mars year (for chosing MCS temperature measurements). The start of each season was chosen to match the 4 seasonal case studies from Rafkin et al (2016), although we focused on the MCS package and not solar longitude in designing the cases.

studies. To put this into context, they used the NASA Ames GCM to provide boundary conditions to the MarsWRF, and performed 7 nested grids, to push the resolution down to 1 grid-point every 330m, much higher than the 5 km one from our reanalysis. This required supercomputing power, so they only simulated the first 4 seasons. We analysed the same periods, but our approach was different. We opted to enhance the MGCM resolution and performing a 1-nest embedding instead.

The output data from the MGCM was sampled 12 times per sol, to reduce calibration time. Additionally, the output frequency of the MMM was also 12/sol. The T31-interpolated initial conditions for the T170 MGCM reanalysis cases, were taken from around 60 sols prior to the actual boundary condition outputs. This required a 60 sol high-resolution spin-up period, in order for the details in the T31-interpolated initial conditions to stabilise before reaching the desired (5-sol) outputs used for embedding. These final outputs are essentially the boundary conditions for the MMM. The net increase in resolution from the embedding is essentially from $\approx$ 1pixel/55.46km^{-1} in the MGCM, to 1pixel/5km^{-1} in the MMM. Regarding the time of the year, the beginning of the seasons in Curiosity operational sols are provided in table 1.

The parameter space for T170 consisted of 120 configurations with and without data assimilation, spanning the 4 seasonal cases. The parameter space included

the testing of the dynamical and physical timesteps, and the mass budget coefficient among others. One scheme that was given priority, was the orographic gravity wave drag (GWD) scheme. The T170 MGCM resolution can cause the dynamical core to begin resolving orographic gravity waves, thus *double-counting* with the GWD scheme. More details are provided in the next subsection and appendix B. The 120 T170 runs were parallelised and set up on the Advanced Research Computing (named "ARC" by the community) cluster (more details about the cluster can be found in "About Advanced Research Computing", n.d.)

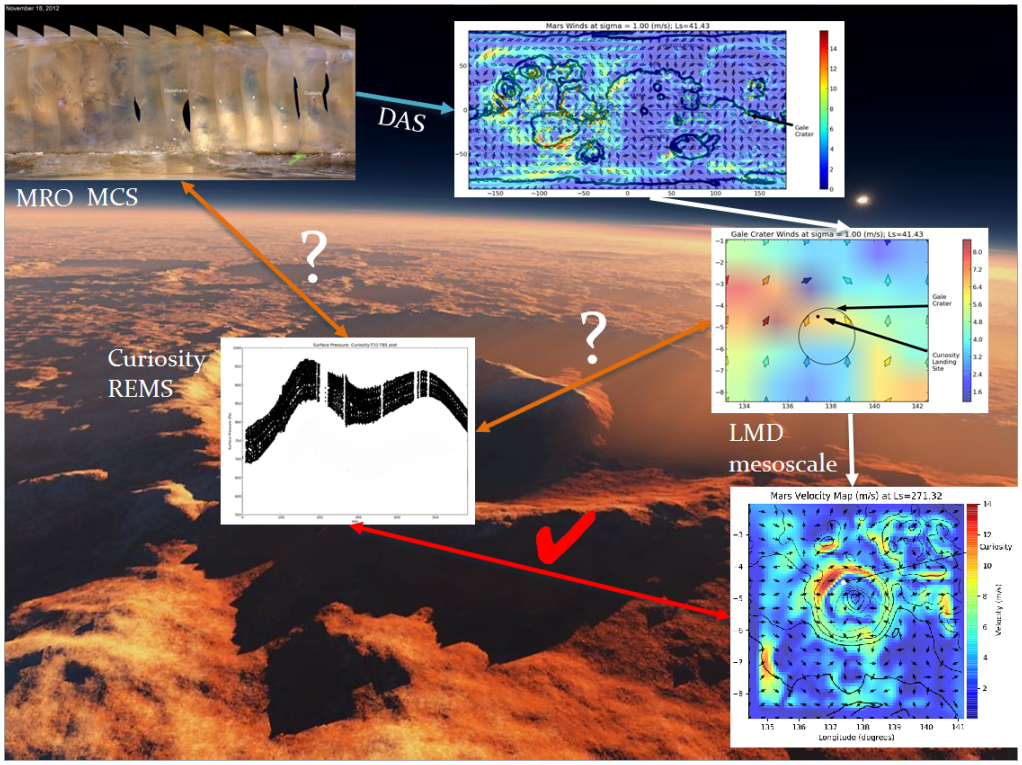


Figure 2.2: The plan for the MGCM+MMM set-up. This figure is a graphical representation of the question "when can we do the MCS-REMS comparison". The orange arrows resemble this question. As noted, the answer to this question is not pertinent for a direct comparison between MCS and REMS, and not enough for a comparison between the MCS→(assimilation)→MGCM and REMS, but starts being viable for MCS→(assimilation)→MGCM→MMM and REMS.

A schematic summary of the MGCM+MMM set-up is provided in figure 2.2. In short, a direct comparison between the MCS and REMS would not be feasible

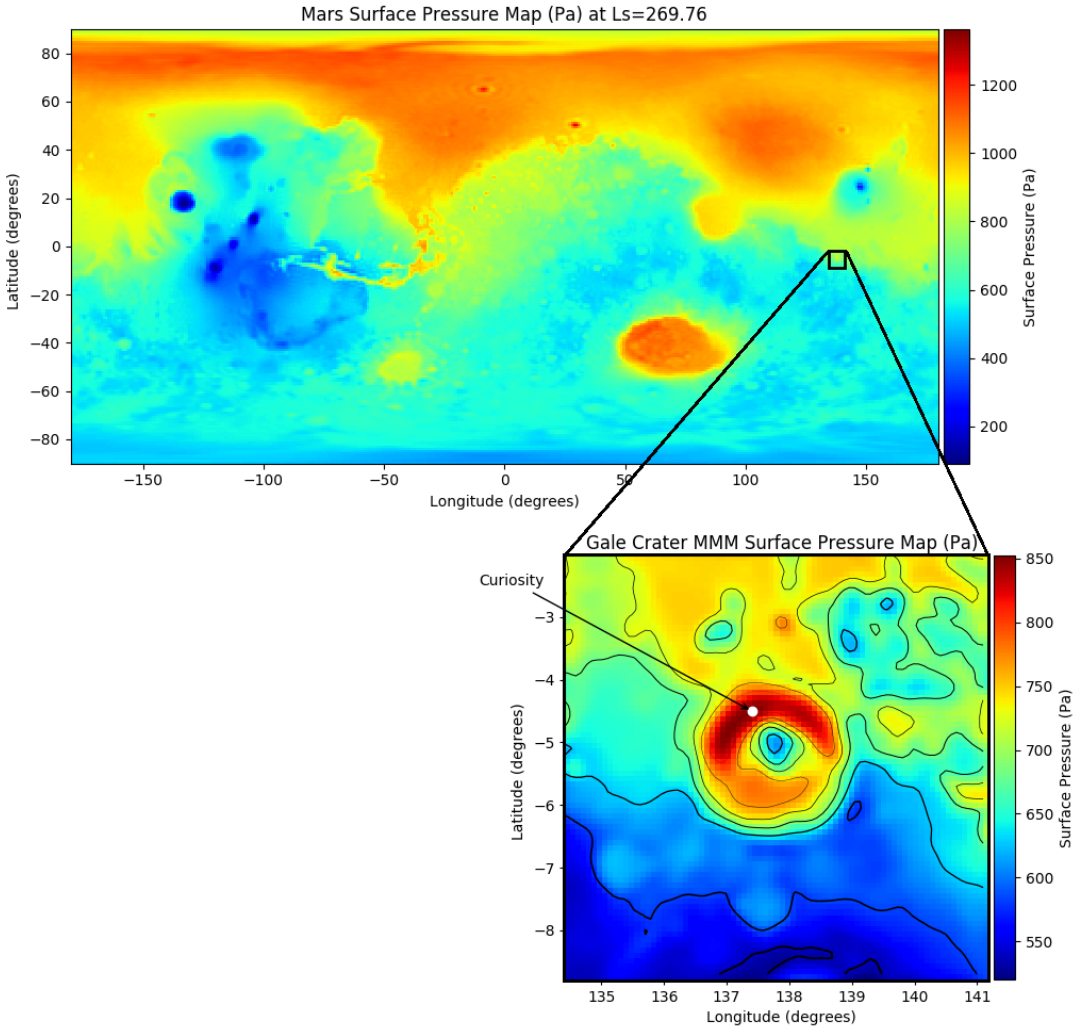


Figure 2.3: The Embedding procedure. The boundary conditions from the MGCM (8x8x25 data grid) continuously drives the MMM, which is at 5 km resolution on a 80x80x60 data grid (the MMM box in the horizontal, is ≈ 6.76 degrees in latitude or longitude). The chosen atmospheric field (from the figure) is the Martian surface pressure. As a result, the two figures resemble Martian topographies at their working resolution. In fact, the crater outline in the MMM figure (bottom) is "drawn" from the mean surface pressure outlines, and used throughout chapters 4 and 5 as a background.

due to the lack of spacio-temporal coverage, while the direct comparison between the MGCM reanalysis (of MCS) and REMS would be too low in resolution to reconstruct the large-scale context of the Curiosity measurements. Figure 2.3 presents the results of the downscaling system. The surface pressure fields from

the T170 MGCM reanalysis correspond to the top plot, and the MGCM+MMM (regional) reanalysis (the embedding) corresponds to the bottom plot. The embedding window is depicted in the global fields (black square), and as mentioned in the introduction, the MGCM+MMM configuration was automated. Hence, any chosen region from the MGCM field is downscaled automatically by the MMM from their interface code.

2.3.1 Embedding methodology

The embedding domain was decided to cover 400×400 km^2 in the horizontal, which essentially contains the $8 \times 8 \times 25$ grid boxes from the MGCM. These drive the $80 \times 80 \times 60$ grid boxes of the MMM. The reason behind this configuration is linked to the Martian general atmospheric circulation. In a similar fashion, the frequency of data injection as MGCM - boundary conditions for the MMM, is also linked to the Martian atmospheric diurnal variability. In this subsection, the schematics of figure 2.2 will be clarified, and the choice of settings for the embedding procedure, will be detailed further.

The MGCM output is the reanalysis. Temperature observations from the MCS are injected in the MGCM, which force the model fields to match observations through the analysis correction technique (presented in the previous section). The first supply of boundary condition to the MMM are also initialising it. Hence, the whole $8 \times 8 \times 25$ - box of data points present in the embedding domain are used in the mesoscale model. The next sets of boundary conditions are supplied each 2 hours to capture the high diurnal variability in atmospheric tides, but only at the rim of the domain box. So the boundary condition points inside the $8 \times 8 \times 25$ - box do not drive the MMM; only the points at the domain edge get injected into the MMM. Essentially, the MGCM boundary conditions initialize the MMM and then dictate the atmospheric fields at the edge of the domain box. In the dynamical core however, the time-step is much lower than the 2 hourly frequency

of the MGCM data injection, so the MMM uses interpolation in time to determine the fields at the boundary for all other iterations. The reason for driving the fields only at the edge of the mesoscale domain comes from the different scales that the two models (MGCM and MMM) compute. The higher resolution processes that are resolved by the MMM dynamical core at its initialization, would drift from the next (lower resolution) data package from the MGCM (after two Martian hours of computational time). Even so, the jump in data density at the domain edge is still a technical problem to overcome. The MMM solves this by having "jump" layers from the boundary box inside, enveloping the MMM domain. These are seen in all MMM outputs of chapters 4 and 5. This is the

The settings of the embedding procedure

As mentioned previously, the horizontal resolution of the MMM is $80 \times 80 \times 60$ grid-boxes covering a 400×400 km^2 (5km lengthwise resolution) around Gale Crater. The absolute limit to the resolution is around 3km-long grid-boxes. Smaller length scales will force the MMM dynamical core to resolve into the gray zone (i.e. the parametrisations), but even the 3km case can compromise model stability, so the 5km resolution was ultimately chosen; which is enough to resolve orographic gravity waves from the crater topography (details in chapter 4). The aim is to include as many surroundings as possible in the MMM domain, so more boundary conditions (from the MGCM-reanalysis) drive the MMM, but not too many. Having a larger forcing region from the reanalysis improves accuracy, but if the MMM domain is too large, it may compromise results. An eddy appearing in the MGCM which is within the domain box, will be "invisible" to the MMM, because the MGCM forces only at the bounding sides of the MMM domain (excluding the first boundary conditions from the MGCM, which are also used to initialise the MMM inside the domain). This is undesirable because the eddy is generated from the reanalysis, and is not driving the MMM. This constrains the domain size to be

smaller than Rossby waves. Another limitation in resolution comes from the more intuitive computation time. The linear size of the MMM domain is approximately twice as big as the crater diameter, and chosen in such a way to avoid cutting major topographical features at the domain edge. The aim is to avoid unwanted interactions. Essentially, the physical fields at the domain edge, can manifest an artificial shock if the higher resolution topography cuts-off to the smoother global topography from the MGCM boundary conditions. The actual embedding was done from packages of $8 \times 8 \times 25$ MGCM points to $80 \times 80 \times 60$ MMM points in a domain size of 400×400 km^2. This takes into account the reasoning above, however, Spiga & Forget (2009) argue that an MMM domain which includes 10×10 boundary points from the MGCM is advisable. This limit is found from Earth models, but we will show that for Mars this is not the case. Chapter 5 will prove that our choice of 8×8 boundary points in the horizontal, is enough for using the embedding procedure in the Martian atmosphere.

The vertical resolution was chosen such that orographic gravity waves are resolved, more details are provided in chapter 4. Compared to the settings in the horizontal resolutions of the MGCM, MMM and domain box, in the vertical, the situation is less flexible. The MGCM vertical resolution is standardized to have 25 vertical levels and model patching was not relevant for our studies. The frequency of boundary conditions is 1 package from the MGCM per 2 Martian hours, and was chosen to account for the large variability in the atmospheric tides.

2.4 The Gravity Wave scheme

The orographic GWD scheme was included in the parameter space for calibrating the T170 runs. We do anticipate its importance from papers like Collins et al (1997). Their study focused on the manifestation of the GWD scheme in the global circulation using the same MGCM (within Prof P.L. Read's group). They

discovered that the scheme (and hence gravity waves) exhibits adiabatic warming of the upper levels in the polar regions, via a dynamically induced circulation. This essentially improved the climatology of the polar regions, which had otherwise, unphysical cold levels.

The calibration of the T170 MGCM concluded that a reduction of 25% in the strength was necessary for the GWD scheme. However, to find the section of the (GWD) scheme that contained a tuning parameter was cumbersome. A full understanding of the theoretical background was required for comprehending parts of the MGCM/GWD code, in order to isolate the tuning parameter. Because of its indirect influence in calibrating the MGCM+MMM, an introduction on how to parametrise orographic gravity waves, ending with the tuning parameter, is provided in appendix B.

2.5 The dust scenarios

The dust cycle is an important aspect of the Martian global circulation, because the atmosphere can actually lift dust particles through a process called *saltation*, and transport it. Usually, dust lifting is a balanced process, because the loaded dust in the atmosphere reduces convection at the surface, which in turn limits the loaded dust. In contrast, a clear atmosphere with little loaded dust, will favour saltation from stronger convection. Under more energetic conditions during the Northern hemisphere autumn (close to the perihelion), dust storm formation is prevalent, sometimes even enveloping the whole planet in a (global) dust storm. These dust storms have a positive feedback effect with the lifting, because they are based on saltation through high wind speeds, which in turn depends on the amount of lifted dust. More energy from the sun gets absorbed by the loaded dust in the atmosphere, increasing wind speeds. Moreover, studies show that the loaded atmospheric dust on Mars is strongly correlated to the semidiurnal

tide amplitude. More details about atmospheric dust and its importance to the Martian general circulation, are found in chapters 4 and 5.

As dust loading in the atmosphere can drastically change the global circulation, many models use so called "dust scenarios", which are essentially climatologies of Martian atmospheric dust (Montabone et al, 2015). In making the MGCM+MMM reanalysis, both the MGCM and the MMM were set to the same dust scenario, the calm (devoid of dust storms) Mars Global Surveyor scenario. Even if these scenarios describe very well the certain atmospheric state (i.e. with or without a storm), the loaded dust is still correlated to the amplitude of the semidiurnal tide. Thus, the MGCM+MMM reanalysis may drift from REMS observations if their atmospheric dust contents do not match. More details can be found in chapter 5. On the other hand, the data for chapter 3 is from MCS temperature and dust assimilation. Instead of a dust scenario, a full dust assimilation and transport scheme was used. The scheme was developed by Ruan (2015) for the MGCM, and will be enabled for future studies with the MGCM+MMM.

2.6 Data sources

2.6.1 The Mars Climate Sounder

The MCS instrument is essentially a radiometer. It collects measurements from low-altitude circular orbit of the electromagnetic spectrum in 9 channels, 5 of which are in mid-infra-red range 11.5 - 25.0 μm (for more information, visit "First Data from Mars Climate Sounder Aboard NASA Mars Reconnaissance Orbiter", n.d.). These are surface - limb measurements with a vertical resolution of 5 km. The different filter channels pass measurements to a retrieval software in different combinations to deduce temperature, dust and water ice data for an altitude ranging up to 80 kilometers, with a random temperature error less than 1K below 50 kilometers and an estimated systematic error of 1-3K. The particular

version of the MCS data that was used in producing the regional (MGCM+MMM) reanalysis, is MCS v4.2.

2.6.2 The Curiosity Rover

This section is structured around understanding and presenting the feature of the Curiosity rover, a good comprehension of the instruments involved and their usefulness, as they span over many scientific areas. Hence the instruments and the mission will be briefly acknowledged, leading to an extension dedicated to the REMS instrument.

REMS is a dedicated weather station on NASA's Mars Science Laboratory (MSL) on Mars, known as the Curiosity rover ("Curiosity Mission Overview", n.d.). The rover was designed to explore the planet's past habitability, to characterize its geology and its past and present climate, and to prepare the "terrain" for future missions (i.e. obtain a more accurate description of Mars' atmosphere and radiation level). Thus, the rover's instruments can be classified into four groups: (a) for analyzing "building blocks of life" chemicals, (b) for investigating the chemical, isotopic, and mineralogical composition of the martian surface plus geological materials, (c) for gathering atmospheric data and (d) for investigating the surface radiation. As our interests focus on recovering and reconstructing atmospheric fields in the vicinity of Curiosity, the data from the full package of sensors that build up REMS is most relevant.

2.6.3 The REMS instrument

The REMS instrument ("Rover Environmental Monitoring Station", n.d. and " Documentation on the REMS", n.d.) was built and supplied for installation on the rover, by the Centro de Astrobiologia and the Spanish government. It consists of six meteorological sensors for measuring pressure, humidity, air temperature, ground temperature, wind speed and direction, and ultraviolet radiation. These

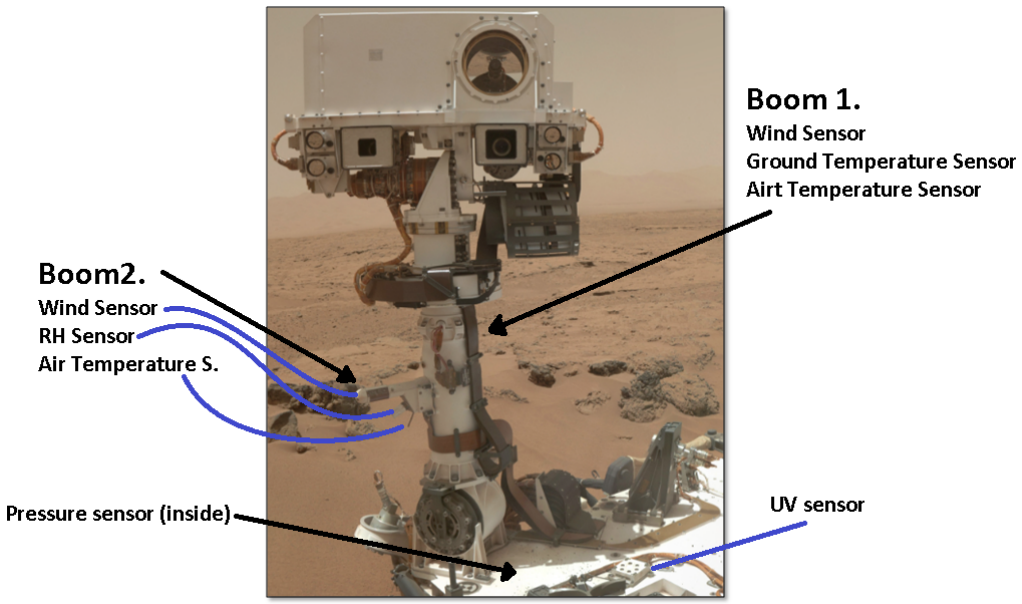
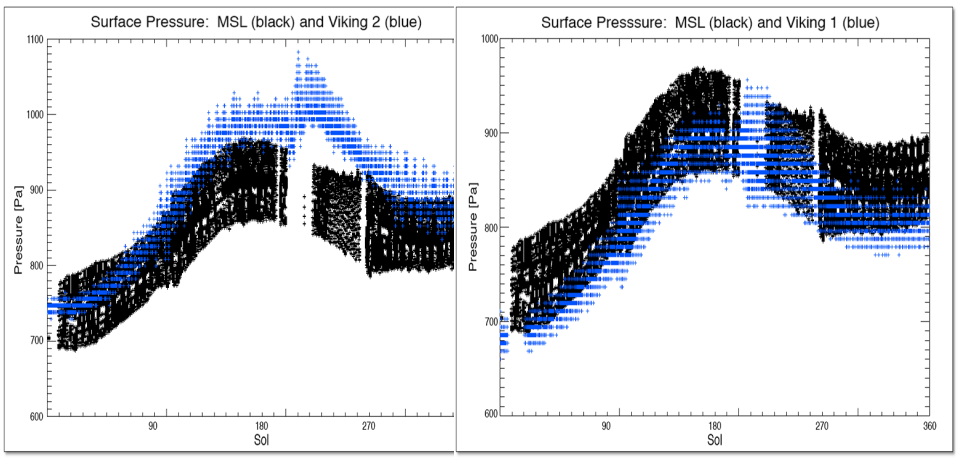


Figure 2.4: The Mast of the Curiosity rover. The REMS instrument includes the two booms on the sides of this. The UV sensor is also visible in this image, and the pressure sensor is inside the body of the rover. (Curiosity, n.d.)

are distributed on the rover over a set of two booms which are on the Remote Sensing Mast (RSM), the UV sensor (UVS) which is on the body and the Instrument Control Unit (ICU) which is located inside the body (Figure (2.4)).



(a) (b)

Figure 2.5: The pressure measurements of REMS and Viking 2 (left) and REMS with Viking 1 (right) over the first 360 sols. (Mischna, 2014)

The two booms are similar in construction and both host wind and air temperature sensors. They are positioned 120 deg apart in order to offer the best correction to the perturbations from the RSM in the wind measurements, and at different heights to avoid mutual interference. Boom 1 also houses the ground temperature sensor and Boom 2 the humidity sensor.

From the rover deck, the UVS measures 6 UV ranges through its six photodiodes (315-350 nm, 280-320 nm, 220-280 nm, 200-370 nm, 230-290 nm and 300-350 nm). These are positioned in the zenith direction and have a field of view of 60 deg. Finally, the ICU hosts the pressure sensor, which is built such that it avoids dust deposition and contamination, through its design and filter.

The design of the instruments is vital for the data we receive, especially for wind measurements (tracking of wind velocity). Unfortunately, due to the **loss of one wind sensor**, during Curiosity's descent, wind data was unavailable in the PDS (The Planetary Atmospheres Data Node, "Welcome to the PDS Atmospheres Node", n.a.) for an extended time. The MSL team calibrated the remaining wind sensor for sole functionality, and recently made the wind data accessible. Nonetheless, it did not stop the scientific community from intensively exploiting the remaining sensors; section 5 is dedicated to the analysis of REMS observations.

The chosen landmark for Curiosity is Gale Crater, which is an impact crater 154 km in diameter, and which has its floor level at 4.5 km below the datum (the conventional zero elevation in Mars' topography). The rover's landing site was close to the northwestern edge of the crater (6 degrees south and 137.4 degrees east). Curiosity's prime mission of one Mars year found its end, and it began its extended mission, heading towards the foothills of Aelis Mons (the center landscape of the crater). The region offers the opportunity of exploring a different environment compared to the previous landers (Golombek et al, 2012). A comparison between the pressure-data received by the rover and the two Viking landers is given in Figures (2.5a) and (2.5b). Judging by the position of the Viking

landers (Figure 1.2) and the position of Curiosity, the pressures differences are as expected: Viking 1's elevation is above Curiosity while Viking 2's is below. The gap in the Curiosity data is due to the occultation of Mars by the Sun. This gap will persist through the rest of REMS' measurements.

This concludes the introduction of the models, data sources and the motivations behind them.

3 The Spectrally resolved energetics of the Martian atmosphere

The spectrally resolved energetics of the atmosphere is an advanced diagnostic method arising from two traditional theories: the global energy budget introduced by Lorenz (1955) and turbulence theory discovered by Kolmogorov (1941). The fields are undergoing constant development since, Kraichnan (1971), Lindborg (1999), Koshyk et al (1999), Boer & Shepherd (1983), Lindborg & Augier (2013) etc. If one is interested in having quantitative knowledge of all energy content and flow that manifests at each scale, behind all circulation mechanisms in the atmosphere, this is possibly the best analysis to make. In our case, the direction of energy flux cascading between adjacent scales would validate (or make it obsolete) the MGCM+MMM for Mars. In addition, this procedure is applied to a reanalysis, making it similar to Gage & Nastrom (1984) but for Mars. The study opens up a new and fruitful theory of comparative planetary atmospherics, with Earth - Mars relation as an opener. Alternatively, there is also a practical use of the spectral energy budget, as a tool for cross-comparison between different Mars reanalyses and model simulations. As Lindborg & Augier (2013) did for the ECMWF and AFES models in their paper, the Mars data assimilation community could benefit from such a tool in improving their Mars GCMs. Obviously this would first require knowledge of what would the *perfect* spectrum be; this is discussed further in the later parts of this chapter.

This chapter follows the trend from the previous one, of having two parts on methodology; although the latter part is significantly more elaborate as it was exhaustively developed. The reason behind this structuring comes from the two distinct methods that can be used with our MGCM for computing the spectral energetics.

Essentially, the MGCM has an important property that can be exploited,

which is its spectral nature. To compute the spectral energetics, one has to bring the physical fields into spectral space, a transformation which is accomplished in the model, though inversely. The mathematics is already is spectral space, only to be summed over in the model to compute the real fields for the output files. Hence, by tapping into the model and generating (and understanding) secondary output files with the spectral coefficients, we managed to lift this method off the ground and compute the energy spectrum directly from the secondary files. The other method is to transform the model output (in real space coordinates) and produce the spectra from there. This might be a more generic method, as it works with any model (as long as the output has the same form as the spectral energetics algorithm).

Returning to the separation of this chapter, a short mathematical introduction into the realm of spectral energetics will prepare the path towards describing the methodology and understanding the results. As already discussed, this section (on methodology) starts with a presentation of the energy spectrum computed directly from the spectral coefficients and continues with the output-dedicated version, with results being displayed independently in each part. The milestones and decisions from the methodologies are relevant for interpreting the results and for setting up the context of the project, so a chronological approach to this chapter is preferred. Finally, the focus will converge on the significance of the study for Mars atmospheric modelling and the novelty of understanding the mechanisms that drive the Martian atmosphere in the framework of energetics.

3.1 Journey through the scales

In appendix D, the flow of mass, momentum and energy equations are derived from first principles, i.e. the flux equation rewritten for mass, momentum and energy; these are the generic equations of motion (or EOMs) (before basis transformations,

redefinitions and approximations). The functionality of the dynamical cores is founded on algorithms from the simplification of the EOMs, the physical cores are parametrizations of additional Physics to the dynamical core and the data assimilation scheme is the application of inverse theory to satellite data to be ingested by the models. The computation of information flow requires a return to the continuum form of the EOMs and a derivation of a diagnostic formulation which also portrays scale interactions. Tabataba-Vakili et al (2015) performed a diagnostic for Mars' atmosphere using the Boer (1989) approach to derive the global Lorenz energy cycle from the MACDA dataset. A generalization from this type of analysis would be the recipe for our diagnostics. This chapter starts from the derivations of appendix D, which invokes an even more theoretical approach that governs both the previous numerical mathematics and the theories which follow.

3.1.1 The energy cycle

The energy balance equations (D.13), (D.15) and (D.17), with the total energy equation (D.11) (appendix D), can be rearranged in the material derivative-form of the flux equation (D.3) as follows (Satoh, 2014)

$$\rho\, d_t\left(\frac{|\mathbf{v}|^2}{2}\right) = -\partial_j(pv_j - \Sigma'_{ij}v_i) + p\nabla\cdot\mathbf{v} - \rho\mathbf{v}\cdot\nabla\Phi - \epsilon, \quad (3.1)$$

$$\rho\, d_t\Phi = \rho\mathbf{v}\cdot\nabla\Phi, \quad (3.2)$$

$$\rho\, d_t u = -\nabla\cdot\mathbf{F}^{\text{heat}} - p\nabla\cdot\mathbf{v} + \epsilon, \quad (3.3)$$

$$\rho\, d_t\left(\frac{|\mathbf{v}|^2}{2} + u + \Phi\right) = -\partial_j(pv_j - v_i\Sigma'_{ij} + F_j^{\text{heat}}), \quad (3.4)$$

where the flux term in the last equation has been moved to the other side as there is no source term for the total energy. Integrating all 4 equations over the atmospheric domain D, the full globally averaged budget equations for energy can

be obtained:

$$\underbrace{\int_D dV\,\rho\, d_t\left(\frac{v^2}{2}\right)}_{d_t K} = \underbrace{\int_D dV\,\{-\partial_j(pv_j-\Sigma'_{ij}v_i)\}}_{=\int_S dS\left(p\mathbf{v}\cdot\hat{\mathbf{n}}+\Sigma'_{ij}v_i n_j\right)\to B} + \underbrace{\int_D dV\,p\nabla\cdot\mathbf{v}}_{W} - \underbrace{\int_D dV\,\rho\mathbf{v}\cdot\nabla\Phi}_{C_G} - \underbrace{\int_D dV\,\epsilon}_{D}, \tag{3.5}$$

$$\underbrace{\int_D dV\,\rho d_t\Phi}_{d_t G} = \underbrace{\int_D dV\,\rho\mathbf{v}\cdot\nabla\Phi}_{C_G}, \tag{3.6}$$

$$\underbrace{\int_D dV\,\rho d_t u}_{d_t I} = \underbrace{\int_D dV\,\{-\nabla\cdot\mathbf{F}^{\text{heat}}\}}_{=-\int_S dS\,\mathbf{F}^{\text{heat}}\cdot\hat{\mathbf{n}}\to F} - \underbrace{\int_D dV\,p\nabla\cdot\mathbf{v}}_{W} + \underbrace{\int_D dV\,\epsilon}_{D}, \tag{3.7}$$

$$\underbrace{\int_D dV\,d_t\left(\frac{v^2}{2}+u+\Phi\right)}_{d_t E_{\text{tot}}} = \underbrace{\int_D dV\,\{-\partial_i\left(pv_j-v_i\Sigma'_{ij}\right)\}}_{B} + \underbrace{\int_D dV\,\left\{-\partial_j F_j^{\text{heat}}\right\}}_{F}. \tag{3.8}$$

Notice that two of the terms are surface integrals emerging from the divergence theorem. B is over the planetary topography accounting for the impact of the interface on the energy budget (i.e. ocean-atmosphere mechanical interactions), while the F term represents the integrated heat flux which has three contributions defined in (D.18). K, G, I and E^{tot} are the global reservoirs of kinetic, potential, internal and total energies, respectively. Their material derivative (left-hand side) is regulated by the following fluxes and conversions (right-hand side): W, D and C_G represent the powers from buoyancy, dispersion and gravity, respectively. Separating reservoirs from fluxes, one can represent them as shown in figure 3.1.

Notice that the reservoirs of internal energy I and gravitational energy G can be combined in a single potential energy reservoir as there is no transfer between them, so $I+G\equiv P$. The typical values for Earth's atmosphere show a high discrepancy between the kinetic energy (KE) and potential energy (PE) reservoirs which make the use of the latter quantity, impractical. Fortunately,

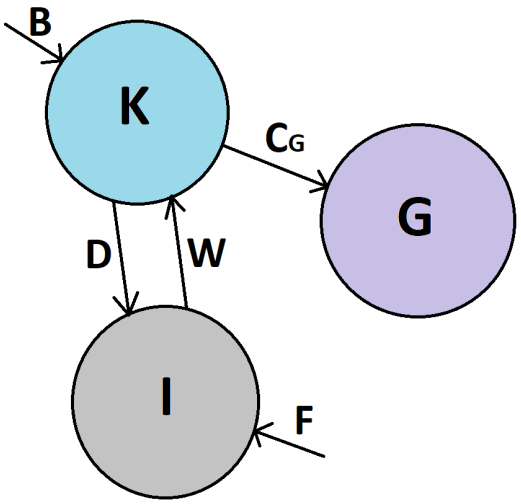


Figure 3.1: The graphical representation of the integrals from equations (3.5)-(3.8).

as it only involves its time derivative in the balance equations, an alternative definition can be used to bring the value of PE down and thus give it a more meaningful definition. The reference state that reduces the PE must be a constant and is certainly not unique. Lorenz was the first to introduce the concept of available potential energy (APE) in 1955 as $A \equiv P - P_{\text{rescale}}$ and offered one such construction for the rescaling PE; here A is the APE. For more details on APE constructions, visit Shepherd (1993). Essentially, the equations evolve into

$$d_t K = B + W - C_G - D, \tag{3.9}$$

$$d_t(P - P_{\text{rescale}}) \equiv d_t A = C_G + F - W + D, \tag{3.10}$$

$$d_t(K + A) \equiv d_t E^{\text{tot}} = B + F. \tag{3.11}$$

Furthermore, notice that the change in E^{tot} is given by the exterior exchanges with the atmosphere, like heat - F and interface - B. This is the energy budget formulation for the globally averaged quantities, which was tackled in Tabataba-

Vakili et al (2015). Up to now, attention has been drawn to the theory of planetary energy reservoirs and the transfers between them. The next section will set up a base for the multiscale nature of the atmosphere.

3.1.2 The energy cycle over scales

There are 3 changes to be made, which bring the budget equations for KE (3.9) and APE (3.10) to the state of proving our MGCM+MMM set-up:

- Break the equations over vertical levels (optional),
- Convert to spherical harmonic spectral space,
- Introduce topography in the equations (optional, but relevant for Mars).

The methodology we are aiming for is achieved in Lindborg & Augier (2013), so we will adopt their approach in a more comprehensive way. The equations can easily be broken into pressure levels

$$\partial_t K(p) = C(p) + L(p) + \partial_p F_{K\uparrow}(p) - D_K(p) + B_K(p), \tag{3.12}$$

$$\partial_t A(p) = G(p) - C(p) + \partial_p F_{A\uparrow}(p) - D_A(p), \tag{3.13}$$

which give rise to two vertical flux terms $F_{K\uparrow}$ for KE and $F_{A\uparrow}$ for APE. These flow through pressure levels, so the energy fluctuations brought to a generic level p are given by their partial derivatives with respect to pressure $\partial_p F_{K\uparrow}(p)$ and $\partial_p F_{A\uparrow}(p)$, respectively. $L(p)$ is the Coriolis term, which will be derived directly in spectral space and with topography. To keep specific, the derivation of the spectral fluxes and reservoirs must start from the altered equations with topography. For a clean-cut progression, we will also set up the spectral space used by Lindborg & Augier (2013).

The real space. First of all, the analysis is accomplished in pressure coordinates, hence four dimensional fields over (time$-t$, spherical$-\phi$, λ, r) from figure D.2, are redefined to the orthonormal basis (time$-t$, latitude$-\phi$, longitude$-\lambda$, pressure$-p$); $\mathbf{x}_h$ are the horizontal coordinates. In this coordinate system, the velocity vector is $\mathbf{v} = (\mathbf{u}, \omega)$ where $\mathbf{u}$ is the horizontal velocity $\mathbf{u} \equiv (u, v)$ and ω is the pressure velocity $\omega \equiv d_t p$, where d_t is the usual material derivative. The horizontal nabla operator is $\nabla_h \equiv (\partial_x, \partial_y)$ and the full 3D version is $\nabla \equiv (\nabla_h, \partial_p)$. For a hydrostatic atmosphere, the connection between the horizontal velocity w and ω is

$$\omega \equiv \underbrace{\partial_t p}_{=0} + u \underbrace{\partial_x p}_{=0} + v \underbrace{\partial_y p}_{=0} + w \underbrace{\partial_z p}_{=-\rho g} = -\rho g w. \tag{3.14}$$

In a similar fashion, the continuity equation (D.5) can be written

$$\begin{aligned} d_t \rho + \rho \nabla \cdot \mathbf{u} = 0 \quad &\Longrightarrow \underbrace{\partial_t \partial_z p}_{0} + (\mathbf{u} \cdot \underbrace{\nabla_h) \partial_z p}_{0} + w \underbrace{\partial_z^2 p}_{0} + \partial_z p (\nabla_h \cdot \mathbf{u} + \underbrace{\partial_z w}_{\partial_p \omega}) = 0 \\ &\Longrightarrow \nabla_h \cdot \mathbf{u} + \partial_p \omega = 0. \end{aligned} \tag{3.15}$$

The fully compressible continuity equation in pressure coordinates conveniently resembles the incompressible one in height coordinates; this is merely an advantage for using pressure coordinates. The hydrostatic equation $\partial_z p = -\rho g$ or simply $p = p(z)$ is also necessary and has been used several times to reach the intermediate and final form of this derivation. Thus, a drawback of changing to this coordinate system is the forced hydrostatic assumption, which might not be accurate at mesoscales. The material derivative in pressure coordinates also has an altered definition, which will be used for the rest of this subchapter:

$$d_t \equiv \partial_t + \mathbf{u} \cdot \nabla_h + \omega \partial_p = \partial_t + \mathbf{v} \cdot \nabla \,. \tag{3.16}$$

Turning to the APE, the geopotential has the following forms

$$\Phi(h) \equiv \int_0^h g\, dz\,, \quad \partial_p \Phi = -\alpha \equiv -\frac{RT}{p}, \tag{3.17}$$

where h is a generic height to which the integral is performed, R is the gas constant and T the temperature. Finally, the addition of topography is done by multiplying the fields by a Heaviside function

$$\tilde{f}(\mathbf{x}_h, p) = \beta(\mathbf{x}_h, p)\, f(\mathbf{x}_h, p)\,, \quad \text{where} \quad \beta(\mathbf{x}_h, p) = H[p - p_S(\mathbf{x}_h)] \tag{3.18}$$

where f is a generic field, H is the classic Heaviside function and the tilde signifies the altered field. H takes the value 0 when the pressure level p pierces the topography when $p > p_S$, and 1 for levels above the topography (i.e. $p < p_S$). The traditional relations involving the Heaviside function that were used in the derivations, are

$$\begin{gathered} \partial_t \beta = \delta_S \partial_t p_S, \quad \nabla_h \beta = \delta_S \nabla_h p_S, \quad \partial_p \beta = -\delta_S, \\ d_t \beta = \delta_S (\partial_t p_S + \mathbf{u} \cdot \nabla_h p_S - \omega_S), \quad \delta_S \equiv \delta(p - p_S), \end{gathered} \tag{3.19}$$

where $p_S \equiv p_S(\mathbf{x}_h)$ is the surface pressure, $\delta(p - p_S)$ is the Dirac delta function and the subscript S denotes functions over the surface (topography). Finally, the EOMs now become

$$\partial_t \tilde{\mathbf{u}} = -\mathbf{v} \cdot \nabla \tilde{\mathbf{u}} - f(\phi)\, \hat{\mathbf{z}} \wedge \tilde{\mathbf{u}} - \beta \nabla_h \Phi + \beta D_{\mathbf{u}}(\mathbf{u}), \tag{3.20}$$

$$\partial_t \tilde{\theta}' = -\mathbf{v} \cdot \nabla \tilde{\theta}' - \tilde{\omega} \partial_p \langle \theta \rangle_r + \tilde{Q}_\theta - \beta \partial_t \langle \theta \rangle_r + \beta D_\theta(\theta), \tag{3.21}$$

where $f(\phi)$ is the Coriolis factor $2\Omega \sin\phi$, $D_{\mathbf{u}}$ and D_θ are diffusion terms, $Q_\theta \equiv$

$\Lambda(p)\dot{Q}/c_p$ where $\dot{Q}$ is the fluctuation of internal energy produced by heating, $\Lambda(p) = (p_0/p)^\chi$ where p_0 is the averaged ground pressure, $\chi \equiv R/c_p \simeq 0.223$ for Mars ($R = 192\,\mathrm{J\,kg^{-1}\,K^{-1}}$ and $c_p = 860\,\mathrm{J\,kg^{-1}\,K^{-1}}$) and $\theta \equiv \Lambda(p)T$ is the potential temperature. These equation, as noted, include topography (β) and account for each of the terms in the balance equations (3.1)–(3.3). The reason behind this formulation of the EOMs is to keep to the notation from Lindborg & Augier (2013). The aim here is to provide a brief but exhaustive derivation of the spectral energetics nomenclature; a concise proof of the methodology from Lindborg & Augier. For completion in defining all the symbols of (3.20) and (3.21), the angle brackets $\langle ... \rangle_r$ denote the representative mean $\langle\theta\rangle_r = \langle\beta\theta\rangle/\langle\beta\rangle$ and $\langle ... \rangle$ is a pressure level mean. The *prime* denotes fluctuations of a field, i.e. $\tilde{\theta}' = \tilde{\theta} - \beta\langle\theta\rangle_r$.

The spectral space. The fields are decomposed into spherical harmonics as follows

$$\theta'(\mathbf{x}_h, p) \propto \sum_{l\geq 0} \sum_{-l\leq m\leq l} \theta'_{lm}(p) Y_{lm}(\mathbf{x}_h), \tag{3.22}$$

where l and m indices are the degree and order of the terms in the series expansion and Y_{lm} is the spherical harmonic function. The decomposition is essentially the same as performed by the MGCM (2.10), but the purpose is different; here we look directly at the spectral coefficients for the information they offer regarding scales, while the MGCM uses them as a mathematical tool to solve the EOMs through the representation of the solutions as spectral series. Triangular truncation is applied here to the sums, but our study of the spectral energetics of the atmosphere follows Lindborg & Augier's paper by carrying out the sum over the orders m and computing the energy cycle over the degrees. The information encoded in such a study is vast, as we will see in the the rest of this chapter or for more information, in appendix D. The mathematical relations used in the derivations regarding

the spherical harmonics functions are: $\nabla_h^2 Y_{lm} \equiv \Delta_h Y_{lm} = -l(l+1)Y_{lm}/a^2$ and $\langle Y^*_{l'm'} Y_{lm} \rangle = \delta_{ll'}\delta_{mm'}$, where $\langle ... \rangle$ is the area-weighted mean defined above, Δ_h is the horizontal Laplace operator. The derivation is done for the kinetic energy only, as the APE balance equation is obtained in a similar fashion and the Heaviside function $\beta(\mathbf{x}_h, p)$ is avoided from here on for clearness.

The full procedure of deriving the KE balance equation including topography is provided in the appendix D. The mean over a generic pressure level can also be decomposed into spherical harmonic coefficients

$$(f,g)_{lm} \equiv \Re\{f^*_{lm} g_{lm}\}\,, \tag{3.23}$$

$$(\mathbf{a},\mathbf{b})_{lm} \equiv \frac{a^2}{l(l+1)} \Re\{(\nabla_h \wedge \mathbf{a})^*_{lm}(\nabla_h \wedge \mathbf{b})_{lm} + (\nabla_h \cdot \mathbf{a})^*_{lm}(\nabla_h \cdot \mathbf{b})_{lm}\}\,, \tag{3.24}$$

where f and g are two generic scalar fields and $\mathbf{a}$ and $\mathbf{b}$ are two generic vector fields. Note that, in a similar fashion to the spectral core of the MGCM, the actual mean over pressure levels can be retrieved from the triangular truncated summation of their specific coefficients in spherical harmonics space (3.23) and (3.24) (above)

$$\langle fg \rangle = \sum_{l \geq 0} \sum_{-l \leq m \leq l} (f,g)_{lm} \quad \text{and} \quad \langle \mathbf{a} \cdot \mathbf{b} \rangle = \sum_{l \geq 0} \sum_{-l \leq m \leq l} (\mathbf{a},\mathbf{b})_{lm}\,. \tag{3.25}$$

The representation of fields in spectral space (or in the case of 2D spherical coordinates, the spherical harmonics space) for solving partial differential equations with no analytical solutions is a well known method in numerical analysis. Nonetheless, to appreciate the potential of this method, one should simply analyse them directly. Using the tools defined here, one can compute the time varying KE spectrum (or KES) from (3.20) by differentiating

$$E_K^{lm}(p) = \frac{(\mathbf{u}\,,\mathbf{u})_{lm}}{2} = \frac{a^2(|\zeta_{lm}|^2 + |d_{lm}|^2)}{2l(l+1)}\,, \tag{3.26}$$

where ζ_{lm} and d_{lm} are the spherical harmonics coefficients for vorticity and divergence. Thus

$$\begin{aligned}
\partial_t E_K^{lm}(p) &= (\mathbf{u}, \partial_t \mathbf{u})_{lm} \left\{ \equiv \partial_t \frac{(\mathbf{u},\mathbf{u})_{lm}}{2} \right\} \\
&= (\mathbf{u}, \overbrace{-\omega\partial_p \mathbf{u} - (\mathbf{u}\cdot\nabla_h)\mathbf{u}}^{-(\mathbf{v}\cdot\nabla)\mathbf{u}} - f(\phi)\hat{\mathbf{z}}\wedge\mathbf{u} - \nabla_h\Phi + D_{\mathbf{u}}(\mathbf{u}))_{lm}
\end{aligned} \tag{3.27}$$

which, after distributing the mean

$$\begin{aligned}
\partial_t E_K^{lm}(p) &= \underbrace{(\mathbf{u}, \nabla_h\Phi)_{lm}}_{C^{lm}(p) - \partial_p(\omega,\Phi)_{lm}} - \underbrace{(\mathbf{u}, \mathbf{v}\cdot\nabla_h\mathbf{u})_{lm} + \partial_p(\mathbf{u},\omega\mathbf{u})_{lm}/2}_{T_K^{lm}(p)} \\
&\quad -\partial_p(\mathbf{u},\omega\mathbf{u})_{lm}/2 \overbrace{-(\mathbf{u}, f(\phi)\hat{\mathbf{z}}\wedge\mathbf{u})_{lm}}^{L^{lm}(p)} + \overbrace{(\mathbf{u}, D_{\mathbf{u}})_{lm}}^{-D_K^{lm}(p)} \\
&= C^{lm}(p) + T_K^{lm}(p) + L^{lm}(p) \overbrace{-\partial_p(\omega,\Phi)_{lm} - \frac{\partial_p(\mathbf{u},\omega\mathbf{u})_{lm}}{2}}^{\partial_p F_{K\uparrow}^{lm}(p)} \\
&\quad - D_K^{lm}(p)\,,
\end{aligned} \tag{3.28}$$

and in a similar fashion to its real-space equivalent (3.12)

$$\partial_t E_K^{lm}(p) = C^{lm}(p) + T_K^{lm}(p) + L^{lm}(p) - \partial_p F_{K\uparrow}^{lm}(p) - D_K^{lm}(p). \tag{3.29}$$

The derivation culminates in the kinetic energy balance equation. The equivalent version for the potential energy has the form

$$\partial_t E_A^{lm}(p) = G^{lm}(p) - C^{lm}(p) + T_A^{lm}(p) + \partial_p F_{A\uparrow}^{lm}(p) - D_A^{lm}(p). \tag{3.30}$$

The lengthy operations are complete, and the resulting terms can be read out of (3.28). To include topography, one has to go back to the EOMs (3.20) and (3.21) to proceed with the same strategy, but making use of the additional relations from Heaviside-function calculus (3.19). The new energy balance equations will generate two more functions in both the KE and APE equations, which are the residuals from manipulating Heaviside and Dirac delta functions (which arise from the differentiation of Heaviside functions). These can be interpreted as the interface terms which account for the influence of the surface to the atmosphere. To our advantage, as we are more interested in energy cascades, or better said in turbulent inertial ranges, we neglected the pressure levels highly impacted by the topography. In these levels (usually the first two levels from the bottom; see subsection 3.3 for more examples), the interface terms are relevant, but offer only information about surface energy injection, rather than information flow. Hence, we ignored the interface terms, as they are also computationally expensive for little usefulness. Nevertheless, the full form of the terms from the balance equations (3.29) and (3.30) including topography, are (PTO):

The energy content

- Kinetic Energy: $E_K^{lm}(p) = \dfrac{(\tilde{\mathbf{u}}, \tilde{\mathbf{u}})_{lm}}{2}$; (*1)

  – The kinetic energy spectrum is a spectral function depicting the content of energy used by dynamical processes at each scale. The local maxima and minima inside the spectrum are markers of scale-dependent injection or dissipation of energy. All adjacent scales in the regions between these sinks and sources, manifest cascades of kinetic energy and remarkably follow power laws. Naturally, the main interaction when energy cascades is turbulence (i.e. nonlinear advective interactions), which governs energy and enstrophy exchanges between large and small scale fluid structures within the KE reservoir.

- Available Potential Energy: $E_P^{lm}(p) = \gamma(p)\dfrac{(\tilde{\theta}', \tilde{\theta}')_{lm}}{2} = \gamma(p)\dfrac{\left|\tilde{\theta}'_{lm}(p)\right|^2}{2}$; (*2)

  – The APE spectrum accounts for all the energy stored in the atmosphere as internal energy, gravitational and centrifugal energy (from the geopotential). The reservoir of potential energy is orders of magnitude larger than the KE reservoir 3.1.1, so the APE reservoir is considered instead. This is a concept introduced by Lorenz which redefines the potential energy to be more useful in this context. This works because the potential energy reservoir itself is not unique and can be renormalised. The shape of the spectrum complements the KE spectrum.

- APE to KE conversion: $C^{lm}(p) = -(\tilde{\omega}, \tilde{\alpha})_{lm}$; (*3)

  – The conversion term is the main scale-dependent transfer between the APE and KE reservoirs. Processes like work done by gravity, buoyancy and diffusion are involved in transferring energy through this term. In spectral form, the conversion is in a cumulative form, so its gradient

from small to large scales is the actual marker for direction of energy transfer.

- Forcing by heating: $G^{lm}(p) = \gamma(\tilde{\theta}', \tilde{Q}'_\theta)_{lm}$; (*4)

  – The external flux of energy from the solar irradiance and cooling to space which only impacts the APE reservoir. The exchange is through shortwave radiative heating from the Sun and longwave radiative cooling into space.

- Nonlinear KE interactions: $T_K^{lm}(p) = -(\tilde{\mathbf{u}}, \mathbf{u} \cdot \nabla_h \tilde{\mathbf{u}} + d \times \tilde{\mathbf{u}}/2)_{lm}$
$+[(\partial_p \tilde{\mathbf{u}}, \omega \tilde{\mathbf{u}})_{lm} - (\tilde{\mathbf{u}}, \omega \partial_p \tilde{\mathbf{u}})_{lm}]/2;$ (*5)

  – The energy flow coming from more complicated correlations, which mathematically involve ensemble means of velocities and velocity variations. This energy flux is the internal exchange of kinetic energy between all adjacent scales of atmospheric structures. As a diagnostic, the flux gets very important because it involves only turbulent interactions. Well-behaved turbulence is the mechanism that may directly feed structures from completely different scales within the KE spectrum, where no sinks or sources have access to the KE reservoir. Working in the blind only with the KE spectrum as in subsection 3.1.3, will demonstrate that even the sign of this term (i.e. the direction of energy cascading) can completely change the mechanisms of atmospheric circulation.

- Nonlinear APE interactions: $T_A^{lm}(p) = -\gamma(\tilde{\theta}', \mathbf{u} \cdot \nabla_h \tilde{\theta}' + d \times \tilde{\theta}'/2)_{lm}$
$+\gamma[(\partial_p \tilde{\theta}', \omega \tilde{\theta}')_{lm} - (\tilde{\theta}', \omega \partial_p \tilde{\theta}')_{lm}]/2;$ (*6)

  – In a similar fashion to the KE cascade above (*5), the APE version flushes energy between scales in the APE reservoir. Again, we notice

correlations between potential temperature variations and their differentials. This accounts for all the isochoric thermodynamic processes (without work) that transport potential energy across scales.

- Coriolis term: $L^{lm}(p) = f_0\{(\breve{\psi}, \breve{d}\sin\varphi + (\partial_\varphi \breve{\chi}\cos\varphi)/a^2)_{lm}$
$$+(\breve{\chi}, \breve{\zeta}\sin\varphi - (\partial_\varphi \breve{\psi}\cos\varphi)/a^2)_{lm}\}; \qquad (*7)$$

  - This term is straightforward as it encapsulates the energy transfer within the KE reservoir due to Earth's rotation. Depends on latitude φ.

- KE vertical flux: $F^{lm}_{K\uparrow}(p) = -(\tilde{\omega}, \tilde{\Phi})_{lm} - (\tilde{\mathbf{u}}, \omega\tilde{\mathbf{u}})_{lm}/2;$ (*8)

  - This accounts for the spectral vertical flux of kinetic energy over the pressure level p from neighbouring levels. This can be compared with other levels to discover if energy is trapped or diverged from certain vertical layers in the atmosphere.

- APE vertical flux: $F^{lm}_{A\uparrow}(p) = -\gamma(\tilde{\theta'}, \omega\tilde{\theta'})_{lm}/2;$ (*9)

  - The spectral vertical flux of potential energy over pressure level p from surrounding levels. Same properties apply from above (*8).

- Diffusion terms: $D^{lm}_K(p) = -(\tilde{\mathbf{u}}, \beta D_{\mathbf{u}}[\mathbf{u}])_{lm}$, $D^{lm}_A(p) = -\gamma(\tilde{\theta'}, \beta D_\theta[\theta])_{lm}$. (*10)

  - The diffusion terms that are not part of KE to APE conversion. These represent the energy sinks across the scales.

- Interface terms: $S^{lm}(p)$ and $J^{lm}(p)$ which are omitted.

– The interface terms are spectral fluxes of energy directly connected to the reservoirs (KE or APE) from the interactions with the planetary surface.

Additionally, $\tilde{A} \equiv \beta A$ is a generic topography-altered field. For the velocity-$\mathbf{u}$ field: $\tilde{\mathbf{u}} \equiv \nabla_h \wedge (\breve{\psi}\hat{\mathbf{z}}) + \nabla_h \breve{\chi}$ is the Helmholtz decomposition into purely rotational and divergent parts of $\tilde{\mathbf{u}}$; $\breve{\zeta} \equiv (\nabla_h \wedge \tilde{\mathbf{u}}) \cdot \hat{\mathbf{z}} = \nabla_h^2 \psi$ and $\breve{d} \equiv \nabla_h \cdot \tilde{\mathbf{u}} = \nabla_h^2 \chi$ are the topography-altered vorticy and divergence respectively; a is the radius of Mars. The factor appearing in the potential energy fluxes and fields is a measure of static stability (equation (3.33)) and is defined as

$$\gamma(p) = R/[-p\partial_p \langle\theta\rangle_r / \sigma^\chi]. \tag{3.31}$$

The potential temperature is defined as $\theta(p) = T(p)/\sigma^\chi(p)$, with $\sigma(p) = p/p_0$ the sigma level and p_0 the surface pressure. Also $\chi = R/c_p \approx 0.2273$. $R = 188.92\text{J/kg/K}$ and $c_p = 831.2\text{J/kg/K}$ are the gas constant and specific heat at constant pressure respectively, for the martian atmosphere.

Note: Not all terms were computed for the energetics analysis, because not all model-outputs include information of the heating or cooling systems. $\tilde{Q}'_\theta$, D_K^{lm} and D_A^{lm} are unknown. Even if we are missing the diffusion terms, effects of diffusion or any other process that is uncovered by our chosen spectra is actually present inside the fields that we have covered. As an example, look at the effects of artificial diffusion which prevents cluttering of energy at lowest scales. All fluxes transport energy inside the atmosphere and connect it to the exterior, however, in models, there is also a need to forcefully extract energy at the end of the resolution. This artificial construct manifests in the fluxes by making them converge to zero at the smallest scales. Figure 3.2 depicts this part of the spectral fluxes.

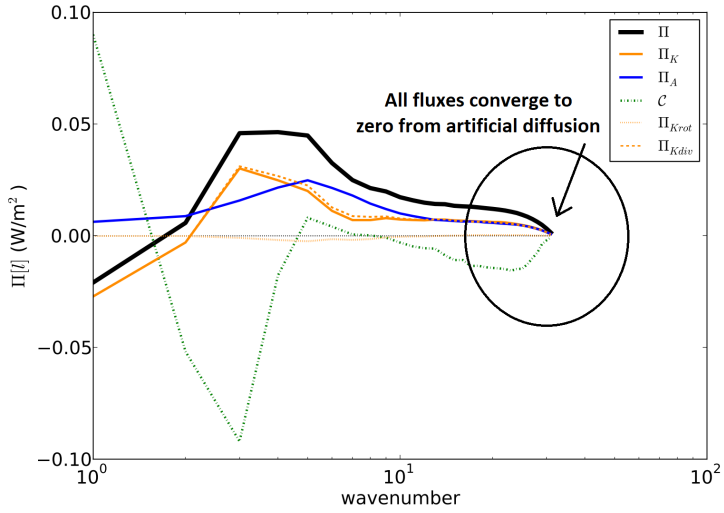


Figure 3.2: All energy fluxes converge to zero at the end of the spectrum due to model diffusion.

3.1.3 The Methodology from spectral fields

Due to the peculiarity of our MGCM, the KES and APES were initially computed in a more efficient and more precise manner compared to the other methodologies used with GCMs throughout the literature. The special procedure of working in spectral space and passing the variables presented in the appendix equation (C.1) between time steps, enables us to directly compute the energy spectra without the use of any indirect methods of computation.

To understand the reasoning behind our methodology, the MGCM is a pseudo-spectral model which has to do a final transformation from spectral to real space to display the output; i.e. to make the output the model does: {spectral $\xrightarrow{\text{output}}$ real}. This means that, to compute the KES and APES, one has to accomplish {spectral $\xrightarrow{\text{output}}$ real $\xrightarrow[\text{APES}]{\text{KES}}$ spectral}. For a spectral space-GCM, this means doing a back-and-forth transformation from spectral space to real space which could compromise the final precision. In the case of a real space-GCM, only one transformation is needed. From real space to spectral space, or {real $\xrightarrow[\text{APES}]{\text{KES}}$ spectral}). There is a workaround in the case of our MGCM by using the second

form of equation (3.26). This form is based on Koshyk et al (1999) but also used by Lindborg & Augier (2013). By summation over the zonal wavenumber m:

$$E_n(p,t) = \frac{1}{4}\frac{a^2}{n(n+1)} \sum_{m=-n}^{n} (|\zeta_n^m|^2 + |D_n^m|^2). \tag{3.32}$$

This methodology frees the KES calculation from the potential loss of precision associated with doing two spectral transforms to get the end result. We are essentially replacing it with {spectral $\xrightarrow{\text{simple sums}}$ KES/APES}, which needs knowledge of vorticity and divergence spectral coefficients. These are already passed between dynamical time-steps, per sigma level, through secondary-binary intermediate files. Equation (3.32) is particularly convenient for further separating the *rotational* and *divergent* parts, and also the *zonal* and *eddy* parts of the end spectrum.

The APES has been implemented according to Augier and Lindborg (2013). The relation presented there follows from summing (*2):

$$A_n(\sigma,t) = \sum_{m=-n}^{n} A_n^m(\sigma,t), \tag{3.33}$$

The terms including the $\gamma(p)$ factor are defined previously (3.31). The 1D APE spectrum is computed with respect to total wavenumber index n, with coefficients A_n and the spectral coefficients are A_n^m. The equivalent relation for the APES used by our code can be derived from (3.33) in the following way

$$A_n^m(\sigma,t) = -\frac{R}{\sigma^{\chi+1}\partial_\sigma \langle T/\sigma^\chi \rangle}\frac{|T_n'^m|^2}{2}. \tag{3.34}$$

Here the variables have the same meaning defined above, and $T_n'^m(\sigma,t)$ are the spectral coefficients of the temperature fluctuations $T'(\sigma,t) = T(\sigma,t) - \langle T(\sigma,t)\rangle$ at sigma level σ. The representative mean profile $\langle\theta\rangle_r = \langle\beta\theta\rangle / \langle\beta\rangle$ has been replaced by the normal σ-level global mean, because the topography is not taken

into account implicitly by the terrain-following sigma coordinate (though note that this is no longer exactly equivalent to a mean along a constant pressure surface); also note that we have changed variables from p to σ coordinates. Hence, the APES is given by summing again like in (3.33). The way this formula works is trickier than the one for KES. As we have seen from chapter one, the spectral form of the temperature fluctuation $T_n'^m$ is transferred between time-steps through a binary file. In the preliminary results computed so far we have extracted the $T_n'^m$ coefficients from the binary files and compute their modulus-squared. We then use the outputted real-grid temperatures to calculate the mean $\langle T/\sigma^\chi \rangle$ for each σ-level and compute the σ-gradient, to calculate the $\gamma(\sigma)$ coefficient. Then we multiply the two results for the same σ levels and time-step t and obtain $A_n^m(\sigma, t)$ which can then be used to calculate the energy spectrum A_n.

The Eddy and Zonal Mean kinetic energy spectra have been extracted from (3.32) in the following way: The zonal mean KES is calculated by separating the $m = 0$ coefficients from the sum, or physically, by separating the zonal mean wavenumbers (which have no features longitudinally) from the total KES. The Eddy KES is the sum of the rest of the coefficients, or physically, the sum over all the turbulent features (or "eddies") from the atmosphere (by filtering the zonal-mean description from the sum). Mathematically, these are given by

$$E_n^{ZM}(\sigma, t) = \frac{1}{4}\frac{a^2}{n(n+1)} \times (\left|\zeta_n^0\right|^2 + \left|D_n^0\right|^2), \tag{3.35}$$

$$E_n^{Ed}(\sigma, t) = \frac{1}{4}\frac{a^2}{n(n+1)} \times \sum_{\substack{m=-n \\ m\neq 0}}^{n} (|\zeta_n^m|^2 + |D_n^m|^2), \tag{3.36}$$

where $E_n^{ZM}(\sigma, t)$ is the zonal mean KES and $E_n^{Ed}(\sigma, t)$ is the eddy KES; their sum gives the total KES from equation (3.32). The energy spectra are also time averaged and where needed, vertically averaged. All methods have been applied

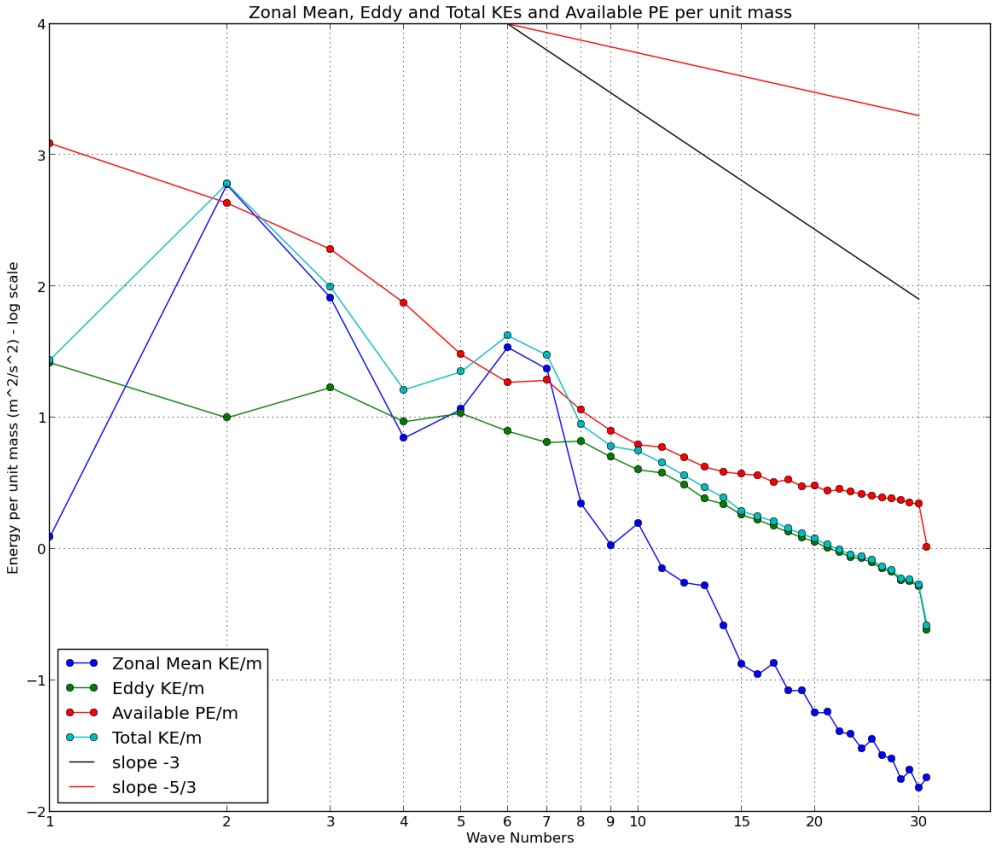


Figure 3.3: The energy spectra computed over the 3 martian years of MCD reanalysis we have at the moment. The algorithms behind each ES plot are given in 3.3

to a T31 assimilation run and T85 free run.

Our applications of this method from the computation of the previously discussed equations for KES (3.35), (3.36) and APES (3.34), (3.33) directly from MGCM binary (spectral) files, are presented next.

3.2 The Energy Spectrum

The calculations are done on the aforementioned spectral coefficients generated by the UK MGCM (section 2.1.1) with assimilated data from the MCS (Mars

Climate Sounder) over 3 martian years (spanning from the middle of martian year 28 to the beginning of 31), at resolution T31 with 25 sigma vertical levels (terrain following). The resolution itself restricts us from extending the spectrum to mesoscale-wavenumbers, therefore mesoscale spectral features are missing from these preliminary studies. Moreover, equations (3.32), (3.34), (3.36) and (3.35) are insufficient for making any assertions about the direction of cascading of energy or enstrophy, and also, the mechanisms behind the forcing of the atmosphere. These are only calculators of overall spectral energies of the kinetic and potential energy reservoirs. Plans for obtaining a more in-depth analysis of the energy spectra are included in the latter part of the chapter. This means implementing energy flux algorithms that quantify the direction and rate of energy transfer between different wavenumber components, and extending the MGCM-resolution and topography to T170, or more (see Conclusions section, 6). Nevertheless, the current stage in our investigations offers us the chance to examine the first Martian energy spectra derived from observations, and to get a glimpse of the turbulent and macroturbulent state of the Martian atmosphere (which will follow in the next section).

The results from the sigma-averaged energy spectra, weighted so that intervals where sigma levels are packed more closely are given lower weights in the spectrum, are presented in Figure 3.3. We will use mass weighting in future studies for computing the spectrum, to automatically weight layers between sigma levels in an appropriate way. Surprisingly, the slope of the total KES departs faster from the -3 law, at roughly wavenumber 10. This suggests that the full martian atmosphere does not really show strong evidence of behaving like an ideal two-dimensional turbulent fluid. One factor which supplies this may follow from the fact that the MGCM simulates an altitude range of $\sim$ 100km in the vertical. With this vertical-thickness, features at wavenumber 10 have an aspect ratio of ~ 10, resulting in a minor approximation breakdown for the assumed two-dimensional

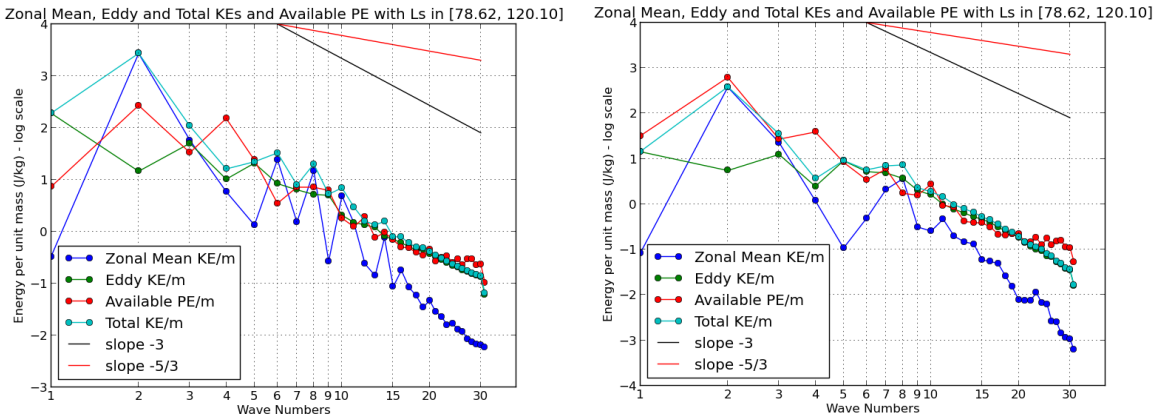


Figure 3.4: the ES computed for the northern summer over the 3 Martian years of assimilated data from MCD. The left plots are the ES at 10 km above the surface and the right plots over 30 km.

atmosphere. Also notice that the flattening of all spectra ends with a steep fall. This is actually a technical issue, which appears from the MGCM (see (C.12) and (C.13)) because of the effects of gridscale dissipation, which is unavoidable due to numerical stability requirements. The energies at this wavenumber (31 in the case of T31) are unphysical and this must be taken into account in our interpretations. Thus, the total number of energies in energy spectra calculations is always reduced by one for the UK MGCM. Another thing to note is the information from the decomposition into eddy and zonal mean KES. Both influence the KES at low wavenumbers, with the zonal mean dominating the spectrum up to wavenumber 7, or wavelength of 3000km. This should be the influence of the large scale circulation: Hadley and Ferrel cells. At larger wavenumbers the eddy KE dominates the spectrum up to the gridbox dissipation range. This region inside the spectrum coincides with the total KES inertial range.

The computation of the energy spectra at altitudes of 10 km (Figure 3.4 right), 30 km (Figure 3.4 left) and 60 km (Figure 3.5) altitude, time-averaged over the northern summer are also presented. These are determined from the spectral coefficients at the sigma levels corresponding to the heights mentioned (10, 30

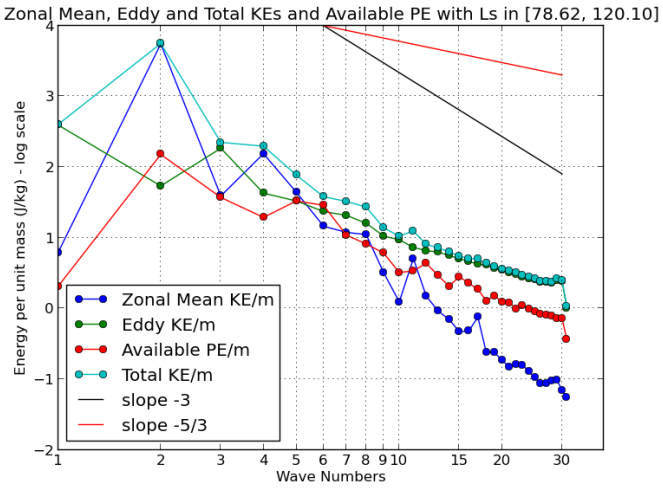


Figure 3.5: The ES at 60 km above the surface for the same conditions as presented in Figure (11).

or 60 kilometers). The case of the 30 km spectrum has a more pronounced -3 law. However, the departure from this law is evident for the other two heights, suggesting that the isotropy approximation is perturbed by forcing at different scales compared to Earth. We will explore this in the second half of the chapter through computing the spectral energy fluxes and the forcing of the atmosphere. Even so, comparing to Figure (3.6) for Earth's atmosphere, the features in the spectrum are clearly restricted to certain vertical levels. This is expected, considering the different energy transfers between scales in the case of the two planets. The features for Mars are, however, left as a question because of the lack of a well developed theory. We intend to explore this in the latter part, because it may well be that there is no classical "inertial range" for either energy or enstrophy with the current results, perhaps because either forcing or dissipation scales do not allow a spectral gap within which the flow can exchange energy conservatively between scales. This is where the spectral fluxes and generation terms can provide more insights.

The computation of the energy spectra at 10 km is for the northern winter and

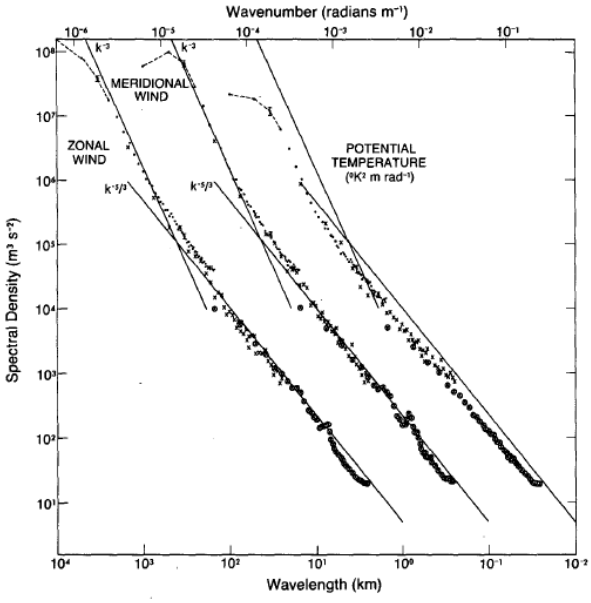


Figure 3.6: The zonal wind, meridional wind and potential temperature spectra computed from the GASP aircraft data. The symbols (•), (○) and (×) represent the spectral averages for long, short and intermediate scale data respectively. (Gage and Nastrom, 1984)

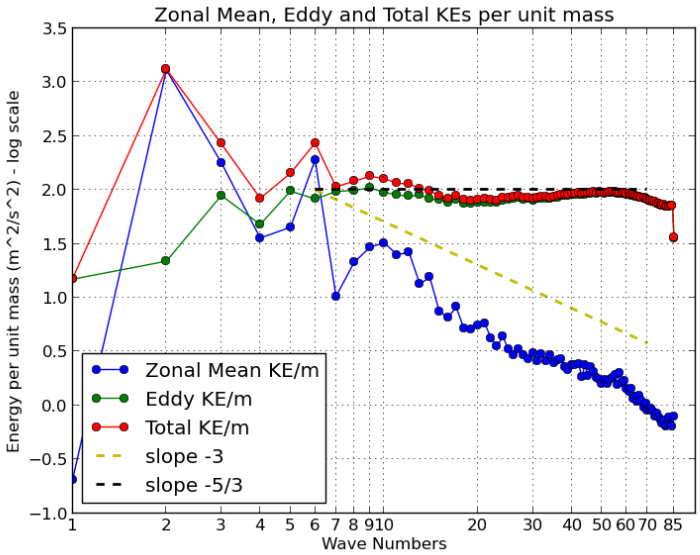


Figure 3.7: The compensated energy spectrum at T85L25 from a free run. Being a compensated spectrum, the horizontal is chosen to represent the -5/3 slope. This emphasizes the persistent energy inertial range from around wavenumber 6-10 to the end of the spectrum (84). Energy is assumed to be cascading in this region according to turbulence theory.

equinox periods. They only show a difference in the APE. This difference is highest for the northern winter and lowest for the northern summer, suggesting that the distance from the Sun and its impact on the local interactions and topography, is the only factor which makes a difference between the seasons. The topography effects are present in the assimilation and not in the actual energy equations in spectral space; in the spectral energy budget includes this feature.

These preliminary results for the Martian atmosphere only "scratch the surface", nevertheless, they do show one major unexpected and unfamiliar pattern of altitude restrictions and early departures from the conventional theories we have for Earth. Additionally, the early $k^{-5/3}$ energy inertial range in the first 31 wavenumbers is unquestionable. However, even if we know that energy should be cascading in this range, this analysis can not conclude on the direction of cascade. So the MGCM+MMM set-up is not formally validated by these calculations, but we can at least extend to 85 wavenumbers to verify how far this inertial range reaches along the spectrum (at larger wavenumbers); see figure 3.7. The spectrum is compensated by 5/3 to emphasise any $k^{-5/3}$ dependence. This is a mathematical manipulation that skews a plot by multiplying all the spectral functions by the compensation factor, which in our case is $\times k^{5/3}$. The T85 energy spectrum was computed for this matter, from a free run over a typical Martian year. Remarkably, the $k^{-5/3}$ range keeps to the end of the spectrum again, going from wavenumber 6-10, or 2200km-3500km to wavenumber 85, or 250km wavelength where it meets the cut-off again. The current results are promising for the MGCM+MMM validation.

3.3 The spectral energy budget

The first scale dependent Lorenz energy budget was computed from a version of the MACDA dataset spanning 3 Martian Years (MY 24 to MY 27). Results only

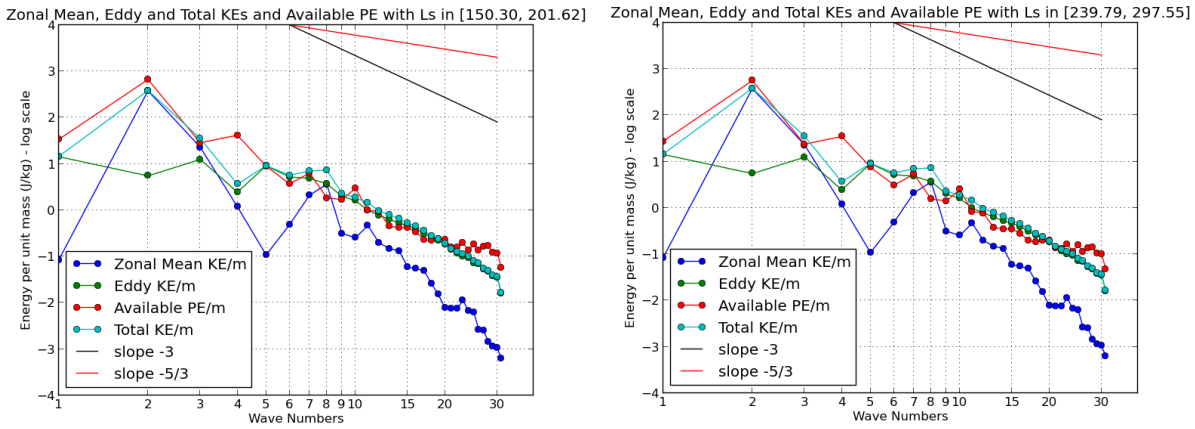


Figure 3.8: The energy spectra at 10 km altitude for the equinoxes and northern winters contained in the 3 Martian years of reanalysis of the MCD data.

include the large scale (small wavenumbers) analysis because the dataset itself is at T31L17 spectral resolution. The theory by Lindborg & Augier (2013) is developed in pressure levels, hence the 17 vertical levels from the MGCM were converted from sigma coordinates, the usual vertical coordinates in the MGCM output. This is also the reason for choosing such a low spectral resolution to do this analysis. To make the transformation from sigma to pressure levels and to do the fast computation of a precise vertical velocity field from a hydrostatic MGCM, are requirements for the Lindborg & Augier analysis. Dr Tao Ruan provided us with a version of the MACDA dataset which already has these transformations; this is essentially the data we used for this study. Nonetheless, we plan on making this link generic for any MGCM output, before releasing the paper on the spectral energetics of the Martian atmosphere (the study of this chapter).

Even at T31L17, the behaviour of the large scale Hadley and Ferrel signatures are manifested in the spectral energetics, with interesting features in the remaining parts of the spectrum. The kinetic energy spectrum shows a $k^{-5/3}$ dependence in the upper levels of the atmosphere for scales smaller than 3500 km, and the energy flux spectrum picks up a downscale energy cascade in this inertial

range. The conversion between the APE and KE mark baroclinic and barotropic energy conversions coupled with large scale circulations (Hadley and Ferrel) and the decomposition into rotational and divergent parts suggest that the enstrophy cascade at low wavenumbers is non existent, although without computing specifically the enstrophy spectrum, this claim inconclusive. The seasonal spectrum over each Martian year picks up the Global Dust Storm (GDS) from autumn Martian Year 25. Outside of this period, the overall energy fluxes are dominated by the divergent part and remarkably in the GDS period, the rotational component becomes comparable to the divergent, resulting in a very energetic and compelling period. The full energy flux is clearly downscale throughout, suggesting that information travels from large to small scales. Finally, different vertical layers have different turbulent behaviours. Inertial ranges can appear in some parts of the spectrum for some vertical levels, while at different elevation, the spectrum can be non-inertial. This means that the vertically - integrated spectrum over the atmospheric domain must be read with caution, because it encompasses inertial regions with injection regions at the same scales, making the spectrum power-laws invalid.

The Lindborg & Augier methodology

The theory that builds up this method was described in section 3.1, which culminated with the mathematical forms for the energy fluxes and reservoirs, in spherical harmonics. The list of spectral terms, along with their equations and physical interpretations are the (*1) to (*10) paragraphs from "the energy content" - table given in section 3.1.2.

The different formalism enabled us to investigate a more general KES. Traditional methods of computing the atmospheric KES usually involved taking samples from the free atmosphere where an inertial range is likely to exist (e.g. the standard Gage and Nastrom, 1984). The modern, Lindborg & Augier (2013)

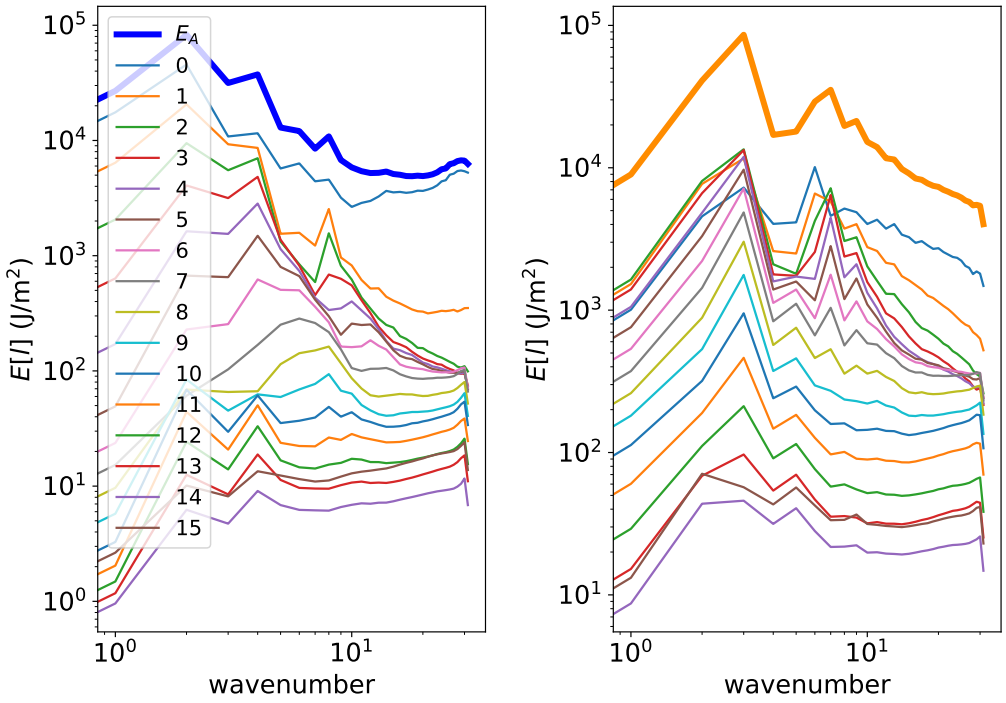


Figure 3.9: The compensated energy spectrum at T31L17 over 3 Martian Years (MY 24 - MY 27). The horizontal is chosen to represent the $k^{-5/3}$ slope and the individual components (thin lines) which construct the full KES or APES (solid lines) are also plotted. The construction of the total KES and APES is done by integrating over vertical pressure levels, i.e. pressure weighted sum over the individual levels. The left plot is the APES and the right plot is the KES.

formalism however, is constructed over pressure levels. As a consequence, topography essentially cuts the first 3 pressure levels, from almost half occupied by topography in the first level, to only Olympus Mons cutting the 3rd pressure level (followed by a clean global wrapping in the 4th pressure level). For this reason, we ignored the first level, as the high surface circulation would be meaningless to the KES, but kept the upper two pressure levels to include the impact of Mars' rich topography in the spectrum.

3.3.1 The Energy Spectrum over 3 Martian years

Figure 3.9 contains the temporally-averaged KES and APES between MY 24 and MY 27 and their decomposition into individual pressure level spectra. Remarkably, the full energy spectrum is featureless from a turbulence perspective. But this should not discourage, as investigating the individual pressure levels, the $k^{-5/3}$ trend is recovered. This means that, inertial ranges can occur locally in the atmosphere, and computing the full atmospheric spectrum can actually be incomplete. As noted in figure 3.9, the influence from the first two pressure levels especially, are corrupting the inertial ranges from the upper levels in the total KES/APES, because the total energy spectra are computed as integrals with respect to pressure (pressure-weighted sums over the vertical levels which are in pressure coordinates) in the vertical (thick blue or dark orange lines). Above the 5th pressure level (≈ 30 Pa), well behaved turbulence starts manifesting right up to the last 5 levels, where it tends to deviate from the $k^{-5/3}$ slope (horizontally, because it's compensated by $k^{-5/3}$). The computation of the two bottom non-inertial layers is a powerful addition of the Lindborg & Augier (2013) method as it elucidates the influence to the spectrum of the topographically-driven circulation, which is an interesting perspective, especially because of Mars' exotic landscape. A noticeable dip (convex function shape) in the KES at wavenumber 4 is accompanied by a peak (concave function shape) in the APES at the same wavenumber, suggesting that a strong correlation is taking place at this wavenumber involving the KE and APE reservoirs. Conversion from KE to APE is the most likely candidate at this length scale. This feature is occurring at all vertical levels, so it most likely has to do with the global circulation. The scale (n=4) and the pronounced effect in the vertical levels close to the surface suggest Hadley and Ferrel signatures in the spectrum.

A decomposition into both eddy and zonal components of the APES and KES

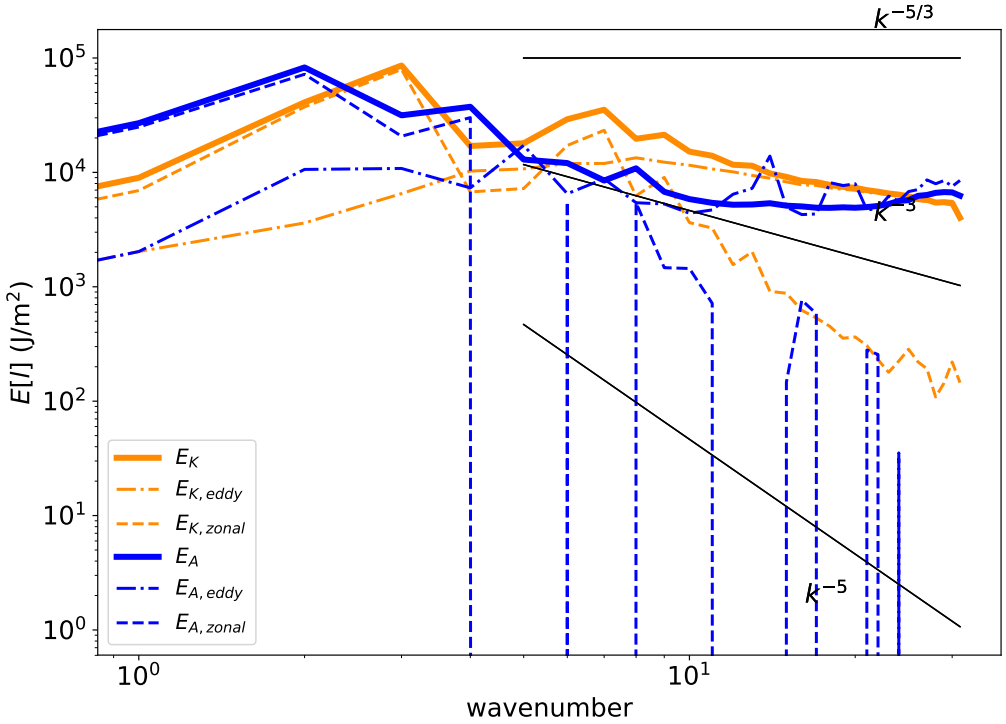


Figure 3.10: The compensated energy spectrum at T31L17 over 3 Martian Years (MY 24 - MY 27). The inertial ranges from the upper atmosphere are masked in the total KES/APES, but the decomposition into eddy and zonal mean energy spectra offers more insight.

is given in figure 3.10. Indeed there is a dip in the zonal KES at n=4, with a small increase in eddy KES and a clear increase in the zonal APES. This reinforces the barotropic energy conversion argument. The total KES and APES are dominated by zonal features at large scales (small wavenumbers) and eddy components dominate the small scales. Earth's atmosphere is similar in this regard. The eddy spectrum is essentially constructed from all the coefficients which are not zonal, i.e. ignoring the m=0 order from figure 1.4, as in equations (3.36) for the eddy KES and (3.35) for the zonal mean KES. So generally, the spectrum is expected to be dominated by the eddy component at large wavenumbers. The most noticeable difference from Earth's KES is the weak presence of the k^{-3} enstrophy inertial range on Mars, or it may even be absent. A spectrum of the enstrophy field would offer more insight and is a feature we would like to add in the future studies of the spectral energetics of the Martian atmosphere.

3.3.2 The Energy Fluxes over 3 Martian years

Causal information is poorly encoded in the energy spectrum, but sets the stage for energy fluxes. The spectral fluxes are thus computed under the same conditions; from the output of a T31L17 reanalysis averaged over 3 Martian years. The collection of analysed spectral fluxes are contained in *the energy content* table from section 3.1.1. Pressure integrated functions from (*3), (*5), (*6) and the rotational and divergent parts of the nonlinear KE interactions (*5), are represented in figure 3.11. The flux notations are defined in the description. The term C is the cumulative conversion from APE to KE and not the total conversion, so the actual converted energy is given by $\Delta_{n1}^{n2}C \equiv C(n_1) - C(n_2)$ between wavenumbers $n_1 < n_2$. Notice that the difference operator Δ is opposite to its standard definition $\Delta x \equiv x_2 - x_1\,,\ \forall x_{1,2}$. Essentially, negative slopes of C denote conversions from APE to KE while positive slopes are KE to APE conversions (over the particular choice of interval $[n_1, n_2]$). The spectrum ends with the forced dissipation from

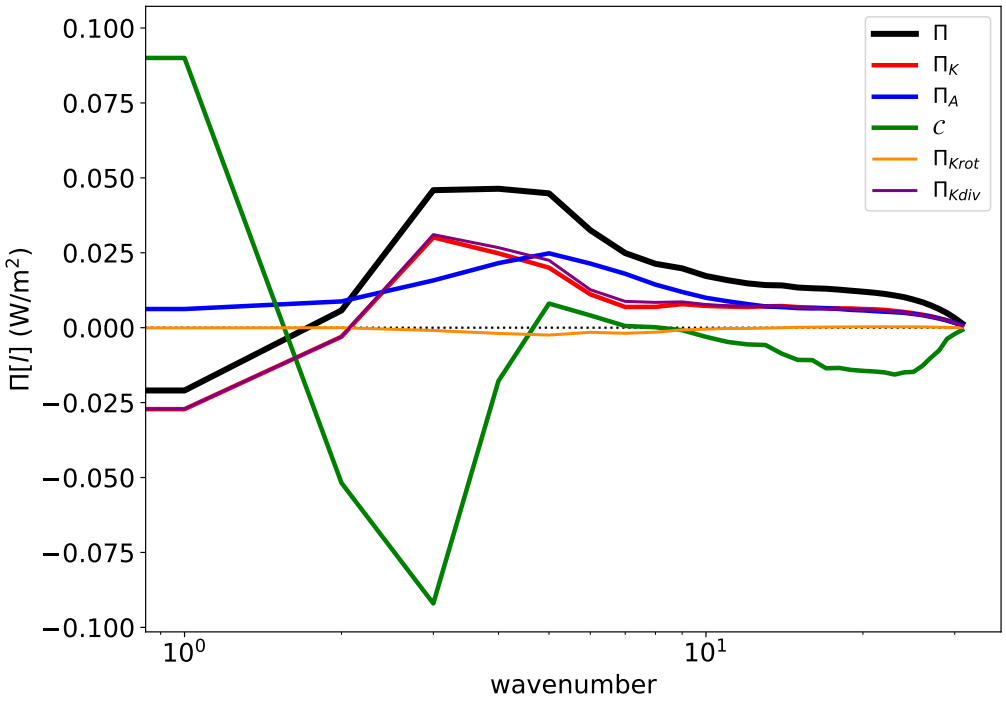


Figure 3.11: The pressure integrated spectral energy fluxes and conversions at T31L17 over 3 Martian Years; MY 24 - MY 27. The symbols are: Π_K and Π_A are the spectral KE and APE fluxes respectively, Π_{Krot} and Π_{Kdiv} are the rotational and divergent parts of Π_K (i.e. $\Pi_K = \Pi_{Krot} + \Pi_{Kdiv}$). Π is the total flux (i.e. $\Pi = \Pi_K + \Pi_A$) and C is the cumulative conversion term between the KE and APE reservoirs.

model resolution, as expected. However, the fluxes should start from zero (not the conversion), but they get deviated from counting the topography-pierced pressure levels in the vertical integral (2 levels are pierced) and ignoring the heavily pierced levels closer to the surface; we wanted to include topographical effects in the spectrum, without significantly modifying the integrated values. Another factor to the fluxes not starting from zero is the small resolution used. The integration done over orders m of spherical harmonic coefficients Y_l^m has a sparse sample of only 61 orders (T31 resolution), which makes the vertical integration (over only 17 levels) miss zero. Additionally, positive values of any of the fluxes Π_x denote downscale cascading, while negative fluxes are upscale.

The averaging of fluxes and cummulative conversion over 3 Martian years in figure 3.11 makes them a reasonable template for Mars. The total flux Π is positive throughout the whole spectrum, suggesting that total energy (APES+KES) cascades down the scales. This trend is followed by both potential and kinetic energy fluxes Π_A and Π_K. Moving through the wavenumbers, the first feature encountered, which is also at the largest scales, is the heavy conversion from APE to KE. This is situated in the interval [1,3] where $\Delta_1^3 C = 0.18\,\mathrm{W/m^2}$ from the APE is converted, making it the largest in the spectrum. Being a spatio-temporal mean, most of the APE is essentially powering the zonal mean KE of the Hadley circulation mixed with baroclinic processes that convert zonal APE to eddy APE to eddy KE. The energy injection into the KE reservoir makes the [1,3] interval non-inertial, and accordingly, the KES and APES from figure 3.9 don't follow any particular theoretical trend (i.e. $k^{-5/3}$, k^{-3}). The next relevant interval is [3,5]. Total flux Π now reaches a plateau as Π_A keeps rising and Π_K decreasing, thus compensating to a constant total flux. The second largest *mass*-conversion in the spectrum follows $\Delta_3^5 C = -0.1\,\mathrm{W/m^2}$, but this time in reverse KE→APE. It is the conversion which creates the fall in KES and the peak in APES in figure 3.9. Now KE is converted back to APE, marking the Ferrel signature, but could also include barotropic processes. Barotropic eddies in the atmosphere decay and boost the zonal wind, hence eddy KE is converted to zonal KE, but not entirely. A portion of the zonal KE also feeds the zonal APE, thus giving an overall KE→APE effect (Simmons & Hoskins, 1978). However, observing the life cycle behaviour of baroclinic eddies (conversion rates in Tabataba-Vakili at al, 2015), it seems that that they tend to equilibrate to almost steady amplitudes for 10s of sols at a time. So another explanation may be needed to account for these fluxes. The spectrum up to wavenumber 5 is an altered version of Earth's energy fluxes as they have the same Hadley signature followed by a Ferrel signature and the KE, APE and total fluxes have the same growth; i.e. all fluxes and conversions are

at different scales and shapes in their absolute value but seem to resemble the signs and first derivative signs from Earth spectra (Lindborg & Augier, 2013). Both planets share these trends in the cumulative conversion term at large scales (for Earth's atmosphere, see Lindborg & Augier, 2013), with one subtle difference which is shown in figure 3.12: barotropic instability dominates the zonal to eddy KE conversion on Mars, while for Earth, this conversion is reversed. This was also seen in simple GCM experiments as we reduce the rotation rate and planetary Thermal Rossby Number (Read et al, 2018). Following is the inertial range where energy is cascaded downscale through turbulence in the upper levels, with a small overall conversion from APE to KE in the lower levels (which corrupt the total KES and APES in figure 3.9 when adding these non inertial levels). From wavenumber 22 to the end of the spectrum, energy dissipation brings all fluxes and conversions to zero.

As C represents the cumulative conversion, its value at $n = 1$ is the total globally averaged conversion taking place in the atmosphere; hence $C^{tot} = 0.09\,\mathrm{W/m^2}$. By adding the ignored pressure level from the surface, the total conversion reaches $C^{tot} = 0.15\,\mathrm{W/m^2}$. This result is in accordance with the sum of all global conversions from Tabataba-Vakili et al (2015) between APE and KE. Their result is presented in figure 3.12 and the total conversion is $C' = C_{Z1} + C_E + C_{Z2} = 0.1596\,\mathrm{W/m^2}$. Finally, the decomposition of Π_K into Π_{Kdiv} and Π_{Krot} proves that the atmosphere is mostly divergent as the rotational part is quite poorly manifested in the spectrum.

The decomposition into eddy and zonal components of all the terms from figure 3.11 are broken in two groups and presented in figure 3.13. The dominant eddy component to the total flux comes from the APE flux, but the zonal components to all fluxes are dominant at large scales. Even the conversion term is mostly dominated by its zonal component, with an exception at wavenumber 5. The two major phenomenon governing the APE→KE conversion from [1,3] are

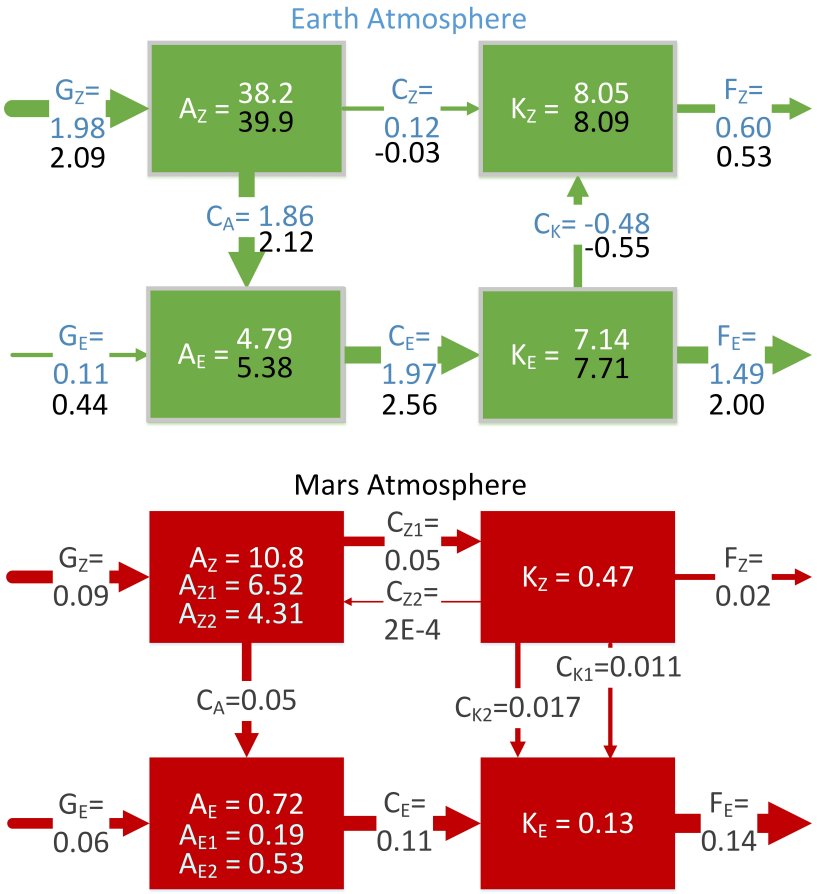


Figure 3.12: The mean energy reservoirs and conversion terms per unit area for Earth's atmosphere (top, Boer & Lambert, 2008) from ERA (black) and NCEP (white/blue) reanalysis (over 17 years, 1979-1995) and for Mars' atmosphere (bottom, Tabataba-Vakili et al, 2015) from the same data used in this chapter. The energies are in 10^5 J/m^2 and conversion terms in W/m^2.

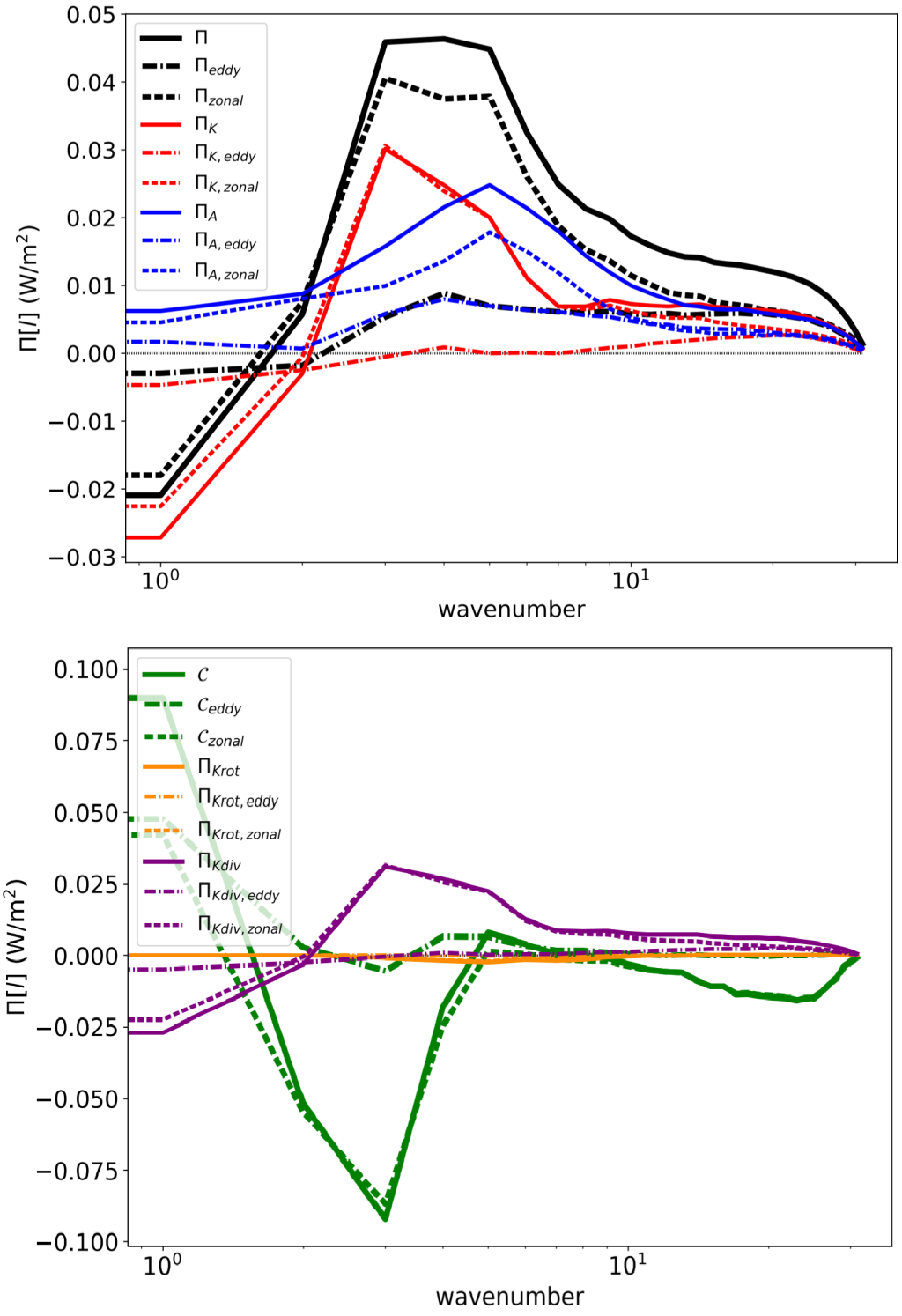


Figure 3.13: The eddy and zonal components of all the terms from figure 3.11.

now quantified. $\Delta_1^3 C_{\text{zonal}} = 0.13\text{W/m}^2$ is the driving energy to the Hadley circulation and $\Delta_1^3 C_{\text{eddy}} = 0.05\text{W/m}^2$ is the amount of baroclinic energy conversion $\text{APE}_{\text{zonal}} \rightarrow \text{APE}_{\text{eddy}} \rightarrow \text{KE}_{\text{eddy}}$.

3.3.3 A seasonal spectral analysis

The focus up to now was on the generic, temporally averaged (over 3 Martian years) picture of Mars' atmosphere. But information from the highly variable spectral features is lost to the average itself. The spectral energy budget is in continuous evolution, with repetition intervals set up by the Sun's synoptic cycles, but even then it is impossible to perfectly repeat two spectral atmospheric states. Thus, the current subsection shifts perspective towards seasonal variability in the spectra, with a glimpse at the global dust storm event (GDSE) from autumn MY 25. But more information can be extracted from the fully time varying wavenumbers.

Together with the global analysis from Tabataba-Vakili et al (2015) of the energy reservoirs and four components of the APE to KE conversion (C_Z for zonal APE to zonal KE, C_E for eddy APE to eddy KE, C_A for zonal APE to eddy APE and C_K for zonal KE to eddy KE; see figure 3.12), our study produces a very dense volume of quantitative flows which have very active dynamical variations that indicate strong turbulent interactions within the spectrum of the Martian atmosphere. The novelty of this comprehensive spectral energy budget study comes from its application to Martian conditions, whether in a model or a reanalysis. Allowing all spectral fields to vary with time, coupled with the time-varying conversions between APE and KE expands the resolution of the changes taking place in the circulation. The full volume of energy spectra, spectral energy fluxes and cumulative conversions from each time frame is ambitious to describe graphically, thence, we opted for a seasonal analysis here.

The first three plots from figure 3.14 are the yearly spectral fluxes and conver-

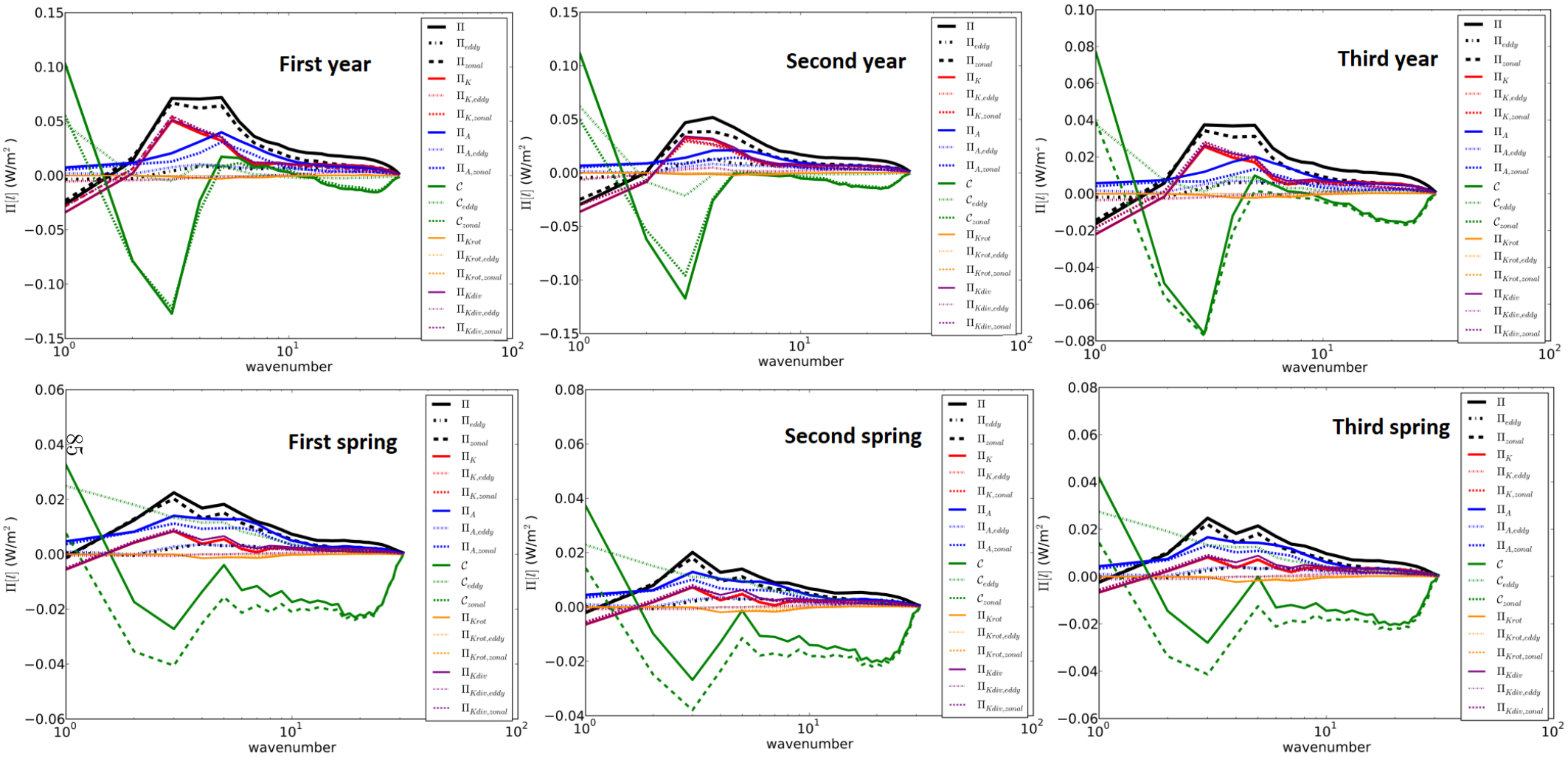


Figure 3.14: The seasonal energy fluxes, cumulative conversion term, decomposition of the KE flux in rotational and divergent components, and the decomposition into zonal mean and eddy components of everything plotted. The yearly and spring values are plotted here.

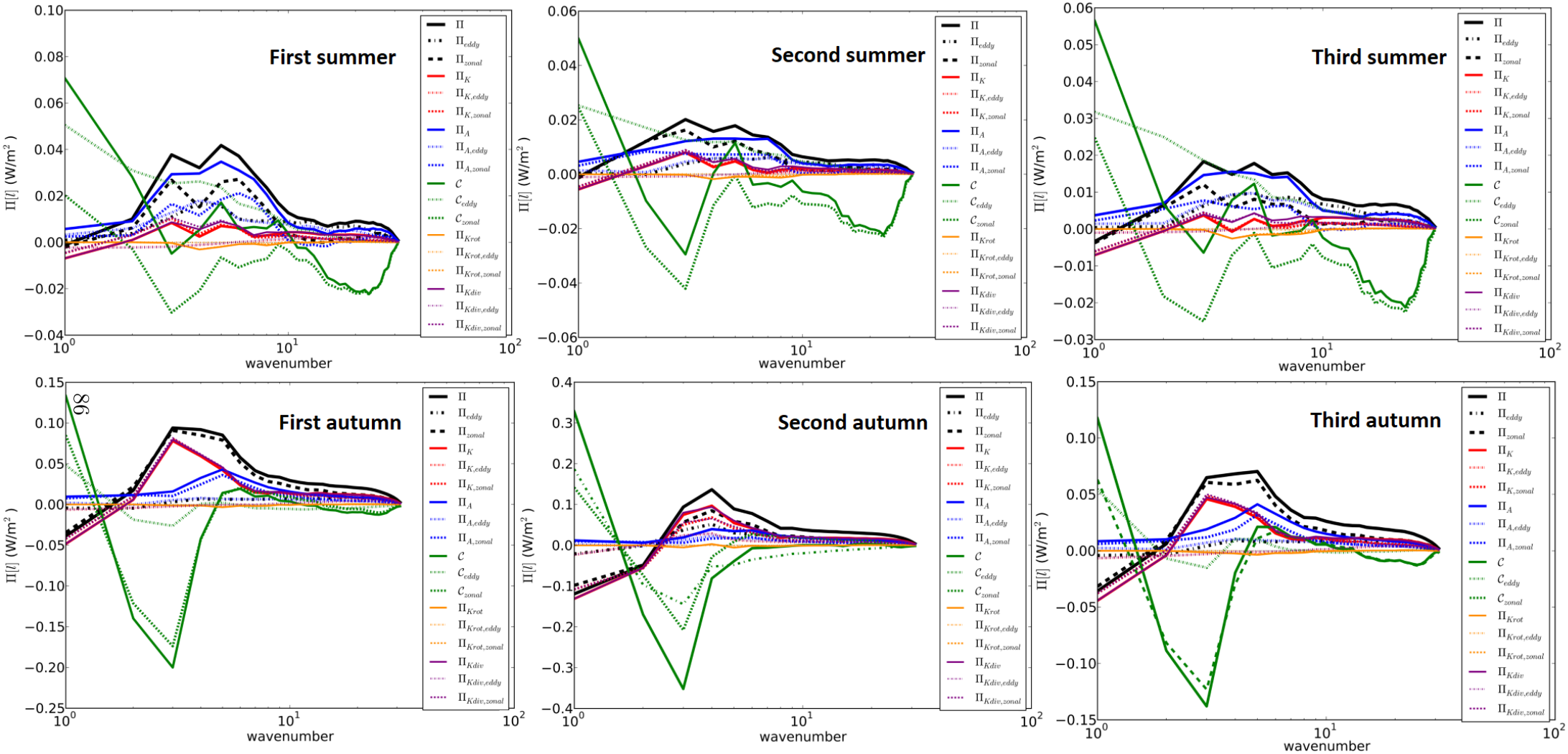


Figure 3.15: Same fields as in figure 3.14. The summer and autumn values are plotted here.

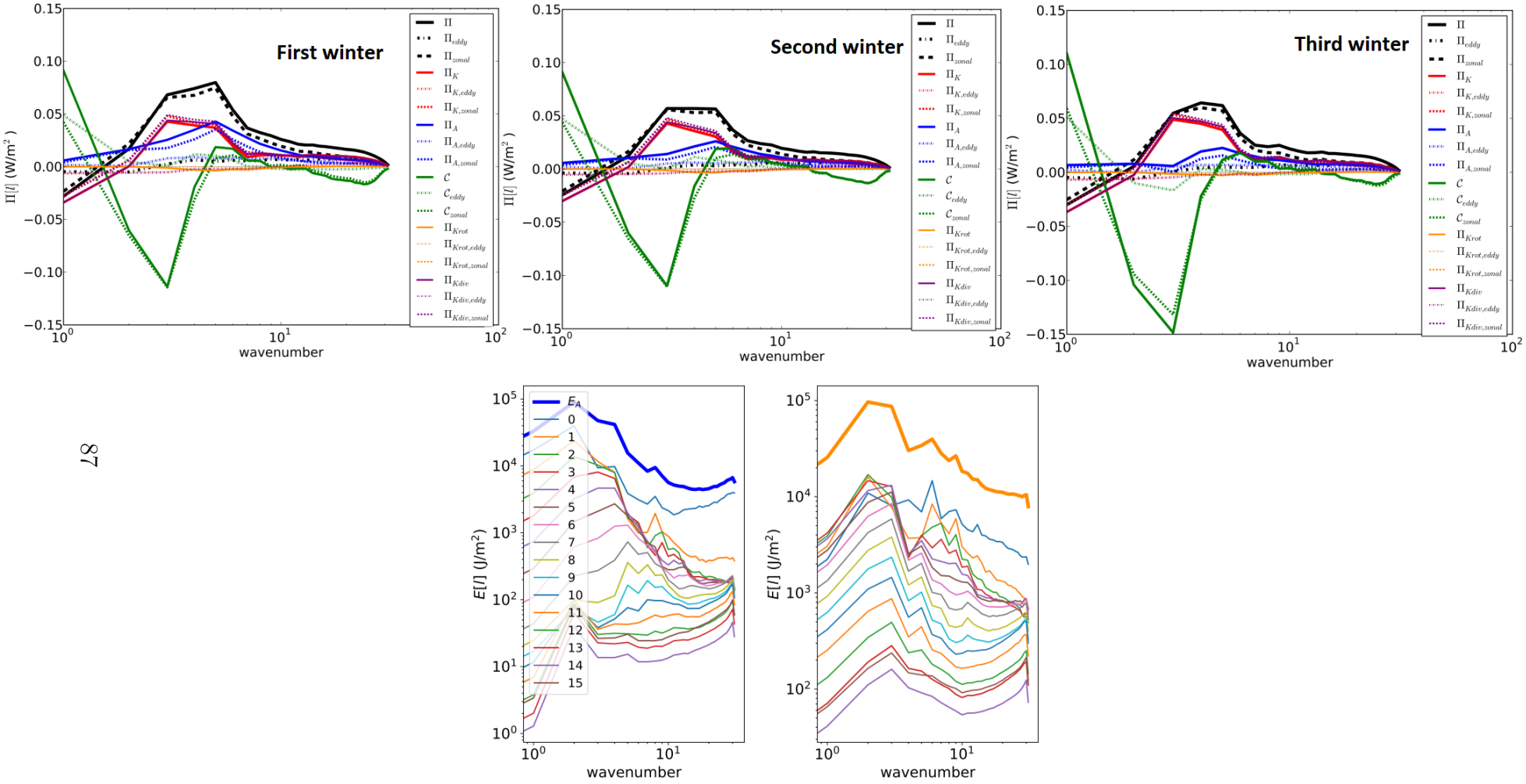


Figure 3.16: Same fields as in figure 3.14 and the APES and KES for second year autumn where the GDSE is picked up. The winter flux values are plotted here.

Period	n=[1,3] (Hadley)		n=[3,5] (Ferrel)		$n > 5$
	X=zonal	X=eddy	X=zonal	X=eddy	X=total
$\Delta C_X^{\mathrm{MY1}}(\mathrm{W/m^2})$	0.18	0.05	-0.13	-0.01	0.04
$\frac{\Delta C_X^{\mathrm{spring}}}{\Delta C_X^{\mathrm{MY1}}}$	0.27	0.25	0.17	$\sim$0.	0.50
$\frac{\Delta C_X^{\mathrm{summer}}}{\Delta C_X^{\mathrm{MY1}}}$	0.27	0.5	0.18	0.16	1.00
$\frac{\Delta C_X^{\mathrm{autumn}}}{\Delta C_X^{\mathrm{MY1}}}$	1.38	1.60	1.37	1.66	0.87
$\frac{\Delta C_X^{\mathrm{winter}}}{\Delta C_X^{\mathrm{MY1}}}$	1.05	1.00	0.88	0.66	1.00
$\Delta C_X^{\mathrm{MY2}}(\mathrm{W/m^2})$	0.15	0.08	-0.10	-0.02	0.02
$\frac{\Delta C_X^{\mathrm{spring}}}{\Delta C_X^{\mathrm{MY2}}}$	0.36	0.15	0.30	$\sim$0.	1.12
$\frac{\Delta C_X^{\mathrm{summer}}}{\Delta C_X^{\mathrm{MY2}}}$	0.36	0.25	0.29	$\sim$0.	0.87
$\frac{\Delta C_X^{\mathrm{autumn}}}{\Delta C_X^{\mathrm{MY2}}}$	2.33	4.00	2.00	0.9	$\sim$0.
$\frac{\Delta C_X^{\mathrm{winter}}}{\Delta C_X^{\mathrm{MY2}}}$	1.00	0.62	0.10	$\sim$0	2.00
$\Delta C_X^{\mathrm{MY3}}(\mathrm{W/m^2})$	0.12	0.04	-0.08	-0.01	0.03
$\frac{\Delta C_X^{\mathrm{spring}}}{\Delta C_X^{\mathrm{MY3}}}$	0.45	0.37	0.30	$\sim$0.	0.66
$\frac{\Delta C_X^{\mathrm{summer}}}{\Delta C_X^{\mathrm{MY3}}}$	0.41	0.30	0.31	0.75	1.06
$\frac{\Delta C_X^{\mathrm{autumn}}}{\Delta C_X^{\mathrm{MY3}}}$	1.54	1.75	1.66	3.00	1.50
$\frac{\Delta C_X^{\mathrm{winter}}}{\Delta C_X^{\mathrm{MY3}}}$	1.83	1.75	2.12	2.00	1.33

Table 2: The conversion values for each MY and season as displayed in Figures 3.14 3.15 and 3.16. $\Delta C_X^{\mathrm{MY1,2,3}}$ are the yearly conversions, where X can be either zonal, eddy or total components for n ={[1,3],[3,5],[5,31]} wavenumber intervals, depending on the column. The seasonal conversions are denoted by $\Delta C_X^{<\mathrm{seasons}>}$, where "seasons" is one of {spring, summer, autumn, winter}. The yearly values are in $\mathrm{W/m^2}$ while the seasonal values in each cell are given as ratios with the yearly value from the specific column (i.e. percentages from the yearly mean). The columns are separated into wavenumbers [1,3] zonal and eddy, [3,5] zonal and eddy and the total conversion for wavenumbers greater than 5.

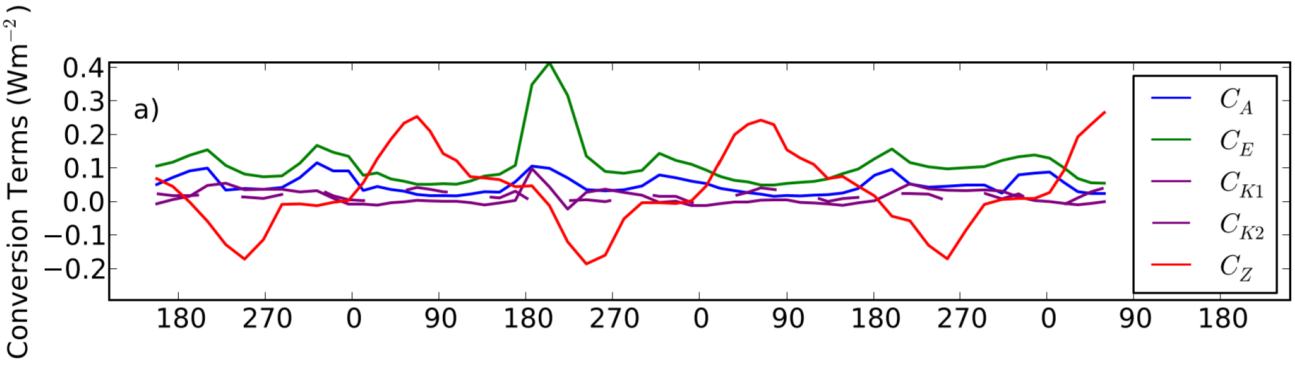


Figure 3.17: The global conversion terms with time over the 3 MY of data; Tabataba-Vakili et al (2015).

sions. General features from each MY keep close to their overall mean from figure 3.11. All the turning points and extrema stay at the same wavenumbers, with variations only in their magnitude. The same Hadley signature converts APE into KE in the interval [1,3], with $\Delta_1^3 C_{\mathrm{MY1}} = 0.23\mathrm{W/m}^2$, $\Delta_1^3 C_{\mathrm{MY2}} = 0.23\mathrm{W/m}^2$, $\Delta_1^3 C_{\mathrm{MY3}} = 0.16\mathrm{W/m}^2$ and $\Delta_1^3 C_{\mathrm{mean}} = 0.18\mathrm{W/m}^2$, followed by the Ferrel signature in the same [3,5] wavenumber interval with $\Delta_3^5 C_{\mathrm{MY1}} = -0.15\mathrm{W/m}^2$, $\Delta_3^5 C_{\mathrm{MY2}} = -0.12\mathrm{W/m}^2$, $\Delta_3^5 C_{\mathrm{MY3}} = -0.09\mathrm{W/m}^2$ and $\Delta_3^5 C_{\mathrm{mean}} = -0.1\mathrm{W/m}^2$ (table 2). The yearly pattern is remarkably consistent throughout the data, even if MY2 contains a GDSE. The effect of the GDSE augments all the eddy components of the flux and conversions, suggesting more eddy-eddy interaction in the atmosphere; typical for such an event. This is correlated mainly to the thermal tide - especially n=2 semi-diurnal (Lewis & Barker, 2005). The baroclinic and barotropic conversions become prominent in MY2. Regarding the seasons, the plots are remarkably specific. Each particular season tends to have the same spectrum with the same season (i.e. spring with spring) from another Martian year (if there is no GDSE) and tends to be different from other seasons (i.e. spring with summer). The standard choice for each season is: 0-90 northern spring, 90-180 northern summer, 180-270 northern autumn and 270-360 northern winter. Northern spring has a weaker overall activity compared to the yearly mean, with a weaker Hadley

and Ferrel signature. From table 2, the Hadley circulation conversion consists of 27%, 36% and 45% of the yearly conversions, for each consecutive MY. Northern summers have weak conversions as well overall, with a peculiar peak in the zonal conversion at wavenumber 9. Mars has the highest eccentricity in the Solar System and reaches aphelion at $L_s = 71$ (where $L_s = 0$ is the conventional start of a MY), which is during northern spring. This makes northern Martian springs longer and gentler than autumns. Solar irradiance varies from $490\mathrm{W/m}^2$ at the aphelion, to $713\mathrm{W/m}^2$ at the perihelion. This drives the seasonal variation to even more extreme differences. Northern springs and summers are significantly milder compared to the yearly mean, as seen from table 2 with much weaker Hadley and Ferrel signatures. Their spectral fluxes and conversions tend to have smaller values across these scales ($n = [1, 5]$). In contrast, autumns and winters are significantly irregular throughout the spectrum, with pronounced large scale conversions, typically above the yearly values. Circulation seems to be more active, especially in the second autumn, which is struck by the GDSE. The zonal component of the Hadley signature is enforced, but the eddy component is so pronounced, that it reaches fourfold its mean value over that year (this is $\frac{\Delta C_{\mathrm{eddy}}^{\mathrm{autumn}}}{\Delta C_{\mathrm{eddy}}^{\mathrm{MY2}}} \approx 4.00$ from table 2); again reflecting an enhanced thermal tide. Energizing the eddy KE produces a growth in eddy activity. This is remarkably followed by a much weaker winter eddy activity, which brings the total fluxes and conversions of the GDS year to their typical yearly values. A GDS year seems to manifest the same total fluxes and conversions as a non-GDS year, but with more energy put into eddy activity. This effect is also seen in the time-varying global conversions from figure 3.17. The non-eddy conversions are particularly weak in MY2 compared to other years, with the exception of C_E during the GDSE itself. Additionally, the KES and APES during the GDSE, which shown in the last plot from figure 3.16, have no well behaved turbulence throughout. This is to be expected, as "waterfall" events should be common during a GDSE (i.e. the turbulent cascade is said to be

non-local).

Final remarks

The spectral energy cycle of the Martian atmosphere is as novel as it is interesting. Features like, large scale conversions and downscale energy cascading allow similarities to be drawn with Earth's atmosphere. However, there are differences and there are unknowns. Mars' low air density and surface pressure brings all energetics to orders of magnitude below what is observed for Earth. Additionally, Mars seems to have a reversed transfer between zonal and eddy KE reservoirs in comparison with Earth, suggesting that barotropic instability dominates this particular conversion (Tabataba-Vakili et al, 2015). Another interesting difference is the absence of an enstrophy inertial range on Mars at large scales, a feature present on Earth. A major influence on Earth's spectral energy fluxes is also the latent heat injections/sinks due to the water-cycle. These can be observed in the spectral energy fluxes of Earth's atmosphere, in the region after the Hadley+Ferrel signatures (Lindborg & Augier, 2013). In the case of Mars, the $n > 5$ features are continuum with $k^{-5/3}$ to the end of the 31 wavenumbers-spectrum. In addition to the missing enstrophy inertial range, it is clear that Mars is in a different regime from Earth, and a complete understanding may require extensions of atmospheric turbulence theory for extra-terrestrial atmospheres.

3.4 Relevance to the MGCM+MMM

Energy spectra and transfers are interesting in their own right. They also happen to provide some justification for 1-way embedding of MMMs into an MGCM. This has intrinsic value to the MGCM+MMM set-up. Thence, the spectral energy budget was employed as a diagnostic tool to compute the scale-wise energy fluxes. These fluxes can be considered as (causal) information flow through the scales of

the atmosphere, and the scales of the MGCM+MMM set-up as well. After running the diagnostics, it turned out that Mars also "accepts" this set-up, however, only for the first 31 wavenumbers. The resolution of the MGCM is T170, meaning that a complete solution to the information flow problem should come with a 170 wavenumber spectrum; this obviously might not be the case. Fortunately, the T85 spectrum in figure 3.7 from subsection 3.1.3 seems to suggest that this downscale inertial range extends further into the mesoscales. For completion of both the information argument and also the full extent of the turbulent study, we are planning to connect the Martian version of the Lindborg & Augier (2013) algorithm to the output of the MGCM directly. This would also be a great addition to the MGCM kit. Extending this idea further, the spectral Lorenz energy cycle analysis can also be a great reanalysis cross-comparison tool.

Lindborg & Augier (2013) used the method for a comparison between two Earth reanalyses, from the ECMWF and AFES, to point out a problem in one of the models, which was subsequently repaired. This requires a template for the Martian spectral Lorenz energy budget (a "perfect" spectrum). Nonetheless, once an agreement is reached, the benefits can be substantial, as the method is essentially in direct link with the dynamics of any Mars GCM that is applied to. Furthermore, the specific scales at which a particular GCM fails to reproduce the energy spectrum and spectral fluxes, can offer insight into exactly which scheme of the GCM needs changing.

4 High resolution reanalysis from the MGCM+MMM system

The REMS era brought a dense volume of research. Its emerging areas of interest include: a. Understanding the diurnal cycle - Gómez-Elvira et al (2014); b. Decomposing and analysing the atmospheric tides - Guzewich et al (2015); c. Convective vortices and dust devils - Kahanpää et al (2016); d. Understanding REMS observations from high resolution simulations - Pla-Garcia et al (2016), Rafkin et al (2016), Tyler & Barnes (2013); and many others. In the current chapter we focus on a broad spectrum of atmospheric motions from the MGCM+MMM reanalysis, specifically:

- Analysing the diurnal tide in connection to the region surrounding the crater. The near-equatorial MMM-domain is dominated by the diurnal tide, which is significantly more prominent on Mars compared to Earth, due to its high diurnal variability. Atmospheric tides have a major impact on the Martian global circulation and Gale Crater is an excellent region for studying the tides. The reanalysis showed similar results to Pla-Garcia et al (2016), Rafkin et al (2016) and Tyler & Barnes (2013).

- Analysing the slope winds at Gale Crater because they constitute the second major component in the local crater circulation. These are coupled to the periodicity of the diurnal tide, but are not part of the tide itself; rather an amplifying effect. Pla-Garcia et al (2016) and Rafkin et al (2016) also analysed the slope winds but from a free running model, and Tyler & Barnes (2013) looked at the amplifying effects to the diurnal tide from the inside the crater.

- Looking at the diurnal variability of the planetary boundary layer (PBL). The PBL is the lowest part of the atmosphere where the planetary surface

exchanges heat, particles (aerosol and dust) and momentum with the general circulation. On Mars, the daytime PBL is highly convective, reaching a height of 10km, much deeper compared to the 1km daytime PBL on Earth. During the night, the Martian PBL drops to 1km, making it tricky to capture in models, but more importantly, it represents a key component in the global circulation of the Martian.

- Capturing instances of the nocturnal low-level jet above Gale Crater. The region between the daytime PBL (10km) and the nocturnal PBL (1km) gets released from friction during the night and develops into a jet-like flow.

- Finding occurrences of daytime convection and convective vortices. Alongside dust devils, these are regular features of the Martian PBL.

- Analysing orographic gravity waves generated from the crater rim at similar instances with Rafkin et al (2016). Gale Crater is a topographically rich landing site for a rover, making it favourable to analyse flow blocking or splitting accompanied by gravity waves in the Martian atmosphere.

Some of the topics are similar to previous studies, but all were driven by the novelty of analysing them from reanalysis.

Overall, the previous chapter tackles information flow between the scales by *categorizing* all turbulent features (e.g. eddies) into their corresponding scales. The method is a leading candidate for extracting information from a reanalysis as a whole, however, it also comes with drawback. Much information is still contained in interpreting the reanalysis directly in real-space, and not just joining it together to be *decomposed* into spectral components. Direct localization of atmospheric processes from a combination of MCS reanalysis, REMS observational data and powerful correlation-extraction methodologies are the main focus of the current and following chapter. State-of-the art methodologies for distilling out information

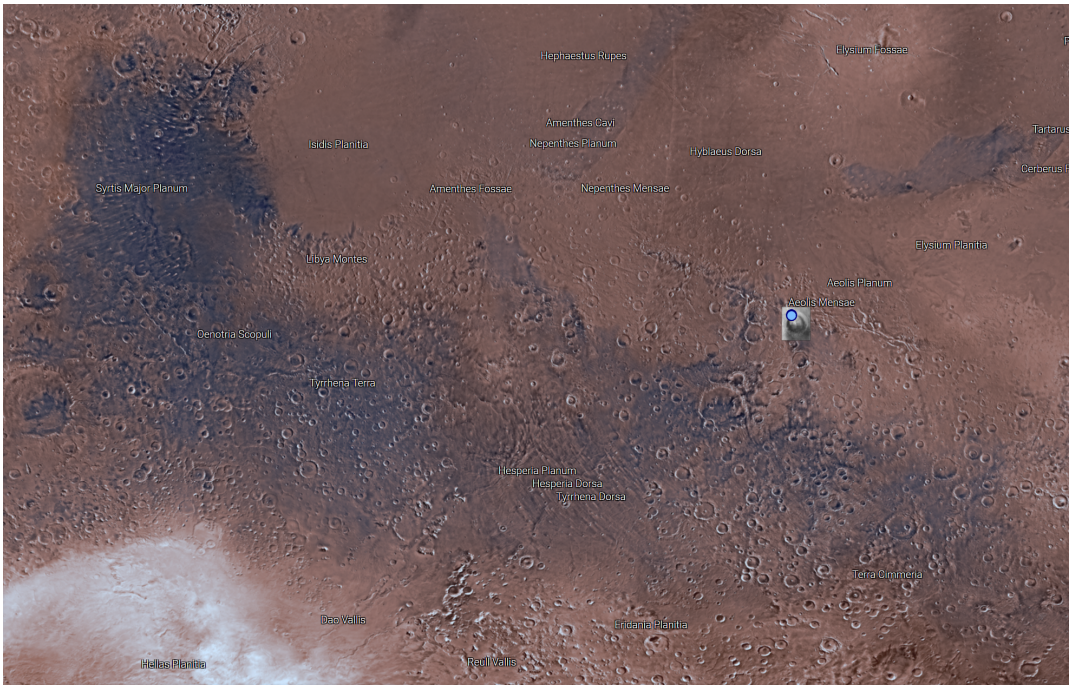


Figure 4.1: Major landmarks around Gale Crater which may have an impact on the crater circulation. ("Explore NASA's Mars Map", n.d.)

is a leitmotif of this thesis.

In this chapter, the results from the MCS$\xrightarrow[\text{Assimilation}]{\text{Data}}$MGCM+MMM configuration will be discussed. A high resolution MGCM+MMM reanalysis (from the mentioned model configuration) is equivalent to opening a window into the planetary environment, so an introduction of the expected Physics (subsection 4.2) is a much needed appetizer. Thus, the present chapter will have two main parts, subsection 4.1 on the Gale Crater circulation and 4.2 for discussing the results from the MGCM+MMM reanalysis.

4.1 Introduction to a crater environment

Gale Crater meteorology is a combination of the global circulation of Mars' atmosphere and the Martian PBL, both from a crater's perspective. Curiosity in particular, is located in the north-western part of the crater floor, which is a plain named Aeolis Palus. The region of the atmosphere where REMS is obtaining

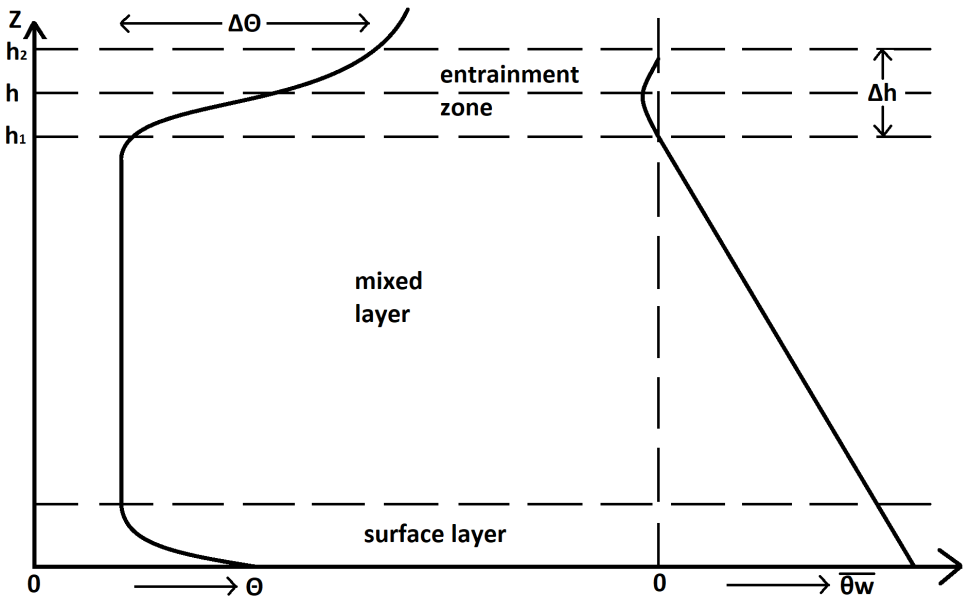


Figure 4.2: The schematic vertical structure of the generic PBL during daytime. (inspired from Read et al, 2017)

observational data is formally named the surface layer, and it is the bottommost component of the Martian PBL. An account of the expectations in this environment is provided in the present section. Influences on the Gale Crater meteorology come from the landscape and the local crater topography. The regional effects from major landscapes such as Elysium Mons and Hellas Impact Basin, are expected to be mostly from anabatic and katabatic winds (a regional map is provided in 4.1); these are discussed in 4.1.2. The local topographical influence is a combination of wind blocking mechanisms (e.g. production of orographic gravity waves) and anabatic/katabatic winds from the crater rim and Mount Sharp, the central crater peak; more details are provided in 4.1.2.

4.1.1 The Martian PBL

The low atmospheric density serves for another distinct and important feature on Mars. First of all, the most abundant atmospheric component is CO_2, which is a potent greenhouse gas. However, the low mass density only raises the mean

planetary temperature by 5K from the effective temperature of 212K (so actual observed mean is 218K). On Earth the effective temperature is 255K, but the greenhouse effect is an additional 33K. Thus on Earth, the actual observed mean temperature is 288K. This means that the surface of Mars is fluctuating greatly from day to night, with a mean variation of 110K. As the planet's surface heats up fast from the Sun (from dawn to noon), the thin atmosphere enters boiling state, with strong superadiabatic convective activity reaching altitudes of 10km during the day. This is the daytime Martian planetary boundary layer or PBL in the literature. The winds from the global circulation will clash with the high convection in the PBL, activating shear and convective instabilities. As a consequence shear and convective turbulence in the PBL accumulates aerosols, dust and water in the atmosphere, which get transported. For this reason, the PBL is an important component for the Martian general circulation. On Earth, this is not the case. The daytime PBL rises up to 1km and is not as active.

Apart from being important in the general circulation, the Martian PBL also plays a key role in sparking dust storm events and in exchange of heat and momentum with the atmosphere above. It is also what Curiosity observes on a circadian basis. Thus, it is paramount to present the daytime PBL effects that are manifested at Gale Crater. Figure 4.2 presents a sketch of the general features of the PBL. The lowest 1km, also called the surface layer, is the transfer region between the surface and the PBL. Here the vertical fluxes of heat and momentum are approximately constant, with a variation of less than 10% from their mean values; more details in Stull (1988) and Monin & Yaglom (1975). Above this region is the mixed layer which is described by strong turbulent mixing governed by convective and sheer instability. At almost 10km in altitude mixing is interrupted and the vertical transfers to the free atmosphere above in an entrainment zone. This is boundary where the PBL passes into the free atmosphere above. In the PBL, the atmosphere exchanges heat, dust and aerosols and momentum with the

planetary surface. Convective vortices are highly active, especially at Gale Crater (Kahanpää et al, 2016), but they seem to not be lifting dust, as only one instance of dust devil was confirmed visually by Curiosity. Lifting of dust requires wind speeds to surpass a threshold value for sand particles get propelled by drag forces. If the process is statistically favourable, enough dragged dust particles will bounce along the surface and force smaller (and harder to lift) dust aerosols into the air; this mechanism is known as saltation (Bagnold, 1941). Kahanpää et al (2016) concluded that the threshold for lifting dust at Gale Crater through convective vortices is rarely met.

At night the Martian PBL shift into the other extreme. Convection ceases and the PBL reduces to a thin stably stratified layer of ≈1km in altitude. Turbulence also descends to the bottom of the stable layer. This extreme reduction in height of all the processes above has another distinct feature, this time on the balance equations of PBL atmospheric motions. The Martian PBL during the night is known to exhibit a nocturnal low-level jet (hundreds of meters above the surface), which can also be found on Earth, but only in relatively flat, desert-like regions. The theoretical description for this effect was suggested by Blackadar (1957). Essentially, the linearised Martian PBL during daytime is govern by the boundary layer equations (Read & Lewis, 2004)

$$\partial_t u - fv + ru = -\frac{1}{\rho}\partial_x p = -fv_g\,, \tag{4.1}$$

$$\partial_t v + fu + rv = -\frac{1}{\rho}\partial_y p = fu_g\,, \tag{4.2}$$

where u and v are the zonal and meridional winds, $f = 2\Omega \sin\phi$ is the Coriolis factor with ϕ denoting the latitude, p and ρ are the usual air pressure and density and r is Rayleigh friction coefficient. As night settles, the PBL rapidly collapses from its usual daytime height of ≈10km to the nocturnal ≈1km, leaving the

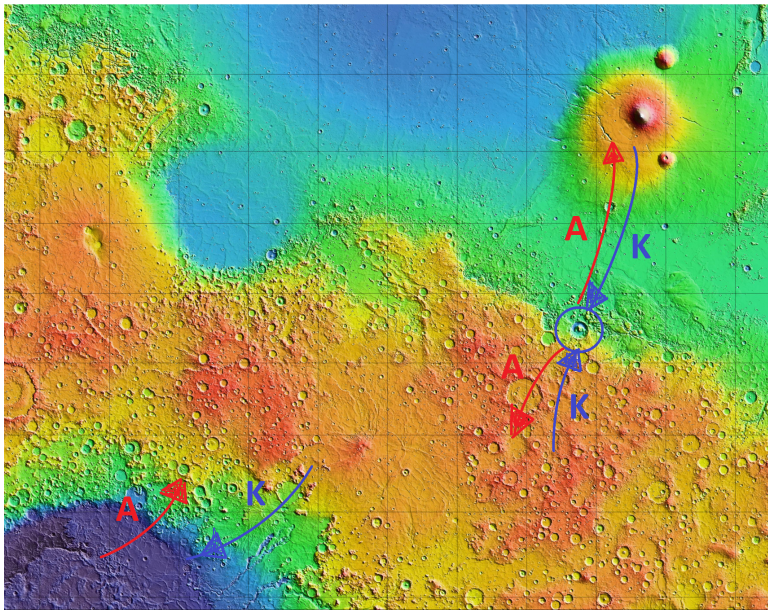


Figure 4.3: The main couplings of anabatic/katabatic winds depicted on a background from the MOLA data. Anabatic wind influences are depicted in red, Katabatic winds in blue. ("The Mars Orbital Laser Altimeter", n.a.)

atmosphere above (i.e. above the descending nocturnal PBL) in a frictionless state. This gradual release from the PBL turbulent hold, gives rise to a jet-like flow, with properties such as jet-nose speed and height. As the daytime PBL is significantly deeper on Mars (compared to Earth), the nocturnal low-level jet has a global presence, with different wind directions dictated by the Coriolis term (which depends on latitude). This phenomenon is restricted to flat regions on Earth.

4.1.2 The regional adjustment of the general circulation at Gale Crater. The Hellas Impact Basin and Elysium Mons

Anabatic and Katabatic winds are thermally driven by orography. These winds are prominent local effects on a mountain's side or over a steep slope (e.g. valleys), and appear from the temperature phase difference between the surface and atmosphere, at these particular landscapes. The ground has low specific heat capacity and absorbs more energy from short-wave radiation compared to the direct

radiative heating of the atmosphere. This is important for generating anabatic winds. The effect accelerates the rise in temperature of the near-surface air which is directly above the steep slope, through conduction and black body radiation. The final ingredient is putting this steep slope with its thin coat of air in an atmosphere. If the slope is high enough, the lapse rate from the atmosphere becomes important. The warm layer of air coating the steep slope is thus in convective instability with the surrounding atmosphere (which is subject to lapse rate, hence getting colder with height) and rises through buoyancy, generating these up-slope anabatic winds. This is also the reason why anabatic winds are strongest at the beginning of daytime insolation, when the surface and near-surface atmospheric temperatures decouple the most. In contrast, when insolation ceases, the lower specific heat capacity of the slope makes it cool faster relative to the atmosphere, through Newtonian cooling. This produces a temperature gradient between the faster-cooling air coating the steep slope and the hotter surrounding atmosphere, creating a negative buoyancy effect. Gravity will drive the coating air, but this time downwards, producing katabatic winds. These are strongest at the beginning of night.

The two major landmarks close to Gale Crater that can manifest anabatic and katabatic winds at a significant magnitude to influence the local crater circulation, are Elysium Mons and the Hellas Impact Basin, as seen in figure 4.3. Additionally, Syrtis Major Planum (north of Elysium Mons) is another significant topographical feature that is even known to manifest low-level jets that cross the equator. Furthermore, the jets occur during periods of zero cross-equator mean flow (i.e. during the equinoxes). This is another orographic forcing that counters the mean circulation, producing counter-intuitive atmospheric motions. This flow is directed to the northern hemisphere and is actually not alone on Mars. A corresponding southward flow appears as well during the equinoxes in the Tharsis plateau region, and exhibits the same cross-equatorial component to the atmo-

spheric circulation in that region (more details are found in Joshi et al, 1997). This further validates the importance of topography as a major component of the general circulation, especially for Mars.

The local topographical effects that alter the weather pattern at Gale. Connections to the PBL

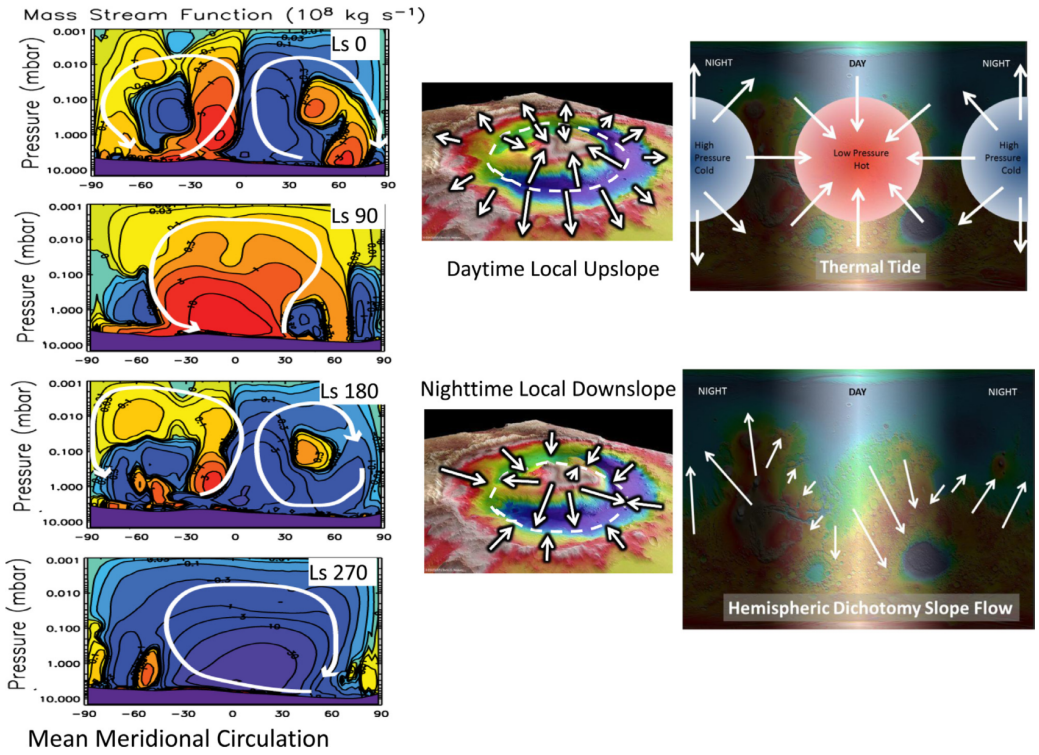


Figure 4.4: The schematic most influential components that drive the circulation at Gale Crater. The zonal section through the general circulation is depicted on the left, the local crater circulation is in the middle and finally, a global schematic of the diurnal tide. (Rafkin et al, 2016)

The same anabatic and katabatic winds make their appearance in the local crater circulation as well. A schematic representation of the winds for Gale Crater are presented in the two middle plots from figure 4.4. The higher crater rim and the central peak (Mount Sharp) represent the steep slopes from the explanation given previously, and are expected to manifest accordingly.

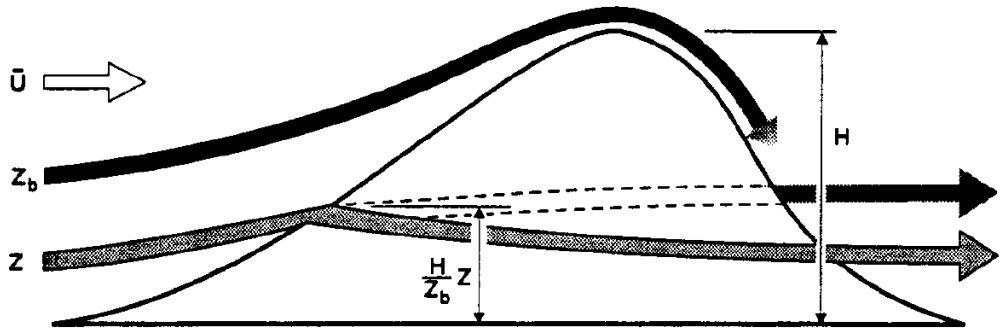


Figure 4.5: Diagram depicting the two situations parametrized by the GWD scheme, for the case of a linear flow over a hypothetical obstacle. H is the height of the obstacle, $\bar{U}$ the velocity of the incoming flow, Z is the layer *depth* defined also in the figure, and Z_b is the boudnary between the two physical situations. Flow breaking and forced flow "over the top", occur for depths $< Z_b$ and $> Z_b$ respectively. (Lott and Miller, 1997)

Other important influences to the local circulation are orographic flow breaking and gravity wave shedding. Both are produced when relatively horizontal winds close to the surface encounter topographic obstacles. Figure 4.5 from Lott & Miller (1997) contains a sketch of this scenario. Simply put, an obstacle can either deflect the incoming flow upwards and produce lee waves on the other side, or split it. The separation between these two distinct outcomes depends on the height of the initial flow, depicted in figure 4.5 by Z_b. It must be noted that these effects can happen at the same time, and the gravity waves produced from flows above Z_b have wavelengths comparable to the obstacle horizontal size and initial amplitudes comparable to its height, but grow (in amplitude) as they dissipate. The boundary height between the two physical situations can be determined analytically and sub-grid scale orographic gravity waves are usually parametrized in GCMs (the MGCM has such a scheme, which is based on Lott & Miller, 1997). A deeper understanding and a theoretical derivation was formalised by Phillips (1984).

4.2 Gale Crater environment from a downscaled reanalysis

4.2.1 The diurnal tide

In the current section, all simulations from the MMM are at 80x80x60 resolution and all simulations from MGCM are at 394x192x25. A post-processing kit for the MMM was used, which was provided by Dr Aymeric Spiga. This was used to perform 2 types of vertical interpolation of the MMM fields to either levels above the local topography (for horizontal maps) or above the aeroid (for vertical maps). All MMM results were processed using this kit.

The near-surface atmospheric temperature of one typical Martian day is presented in figures 4.6 and 4.7. The chosen time of the year is the beginning of Martian autumn, sol 374 or Ls 181.06 to 181.6. There are 12 frames tiled in the figures, sweeping one sol at regular 2 (Martian) hour intervals. Their starting point is from midnight at Curiosity. The frames are depicted by cells containing 3 synchronous plots (hence, each figure contains 6 cells) and evolves from a left cell to a right cell. These cells are thus time frames from 3 different perspectives: a. A global perspective - the top plot is the global near-surface atmospheric temperature from the MGCM (first sigma surface), b. A regional perspective - the lower-left plot is a regional section from the MGCM including Elysium Mons and Hellas Impact Basin, and finally, c. The local perspective - lower-right plot of each cell is the MMM simulation of Gale Crater at 5 m above the local topography. The MMM single embedding box is depicted in the upper (global perspective) MGCM plot by a black rectangular frame, the small circle in the regional plot denotes Gale Crater itself and the white dot in the MMM plot is Curiosity's location in its first Martian Year. The idea behind this graphical set-up is to have a coupled global-regional-local evolution and the reason for choosing the northern autumn reanalysis to present here, stems from the high solar irradiance during this season

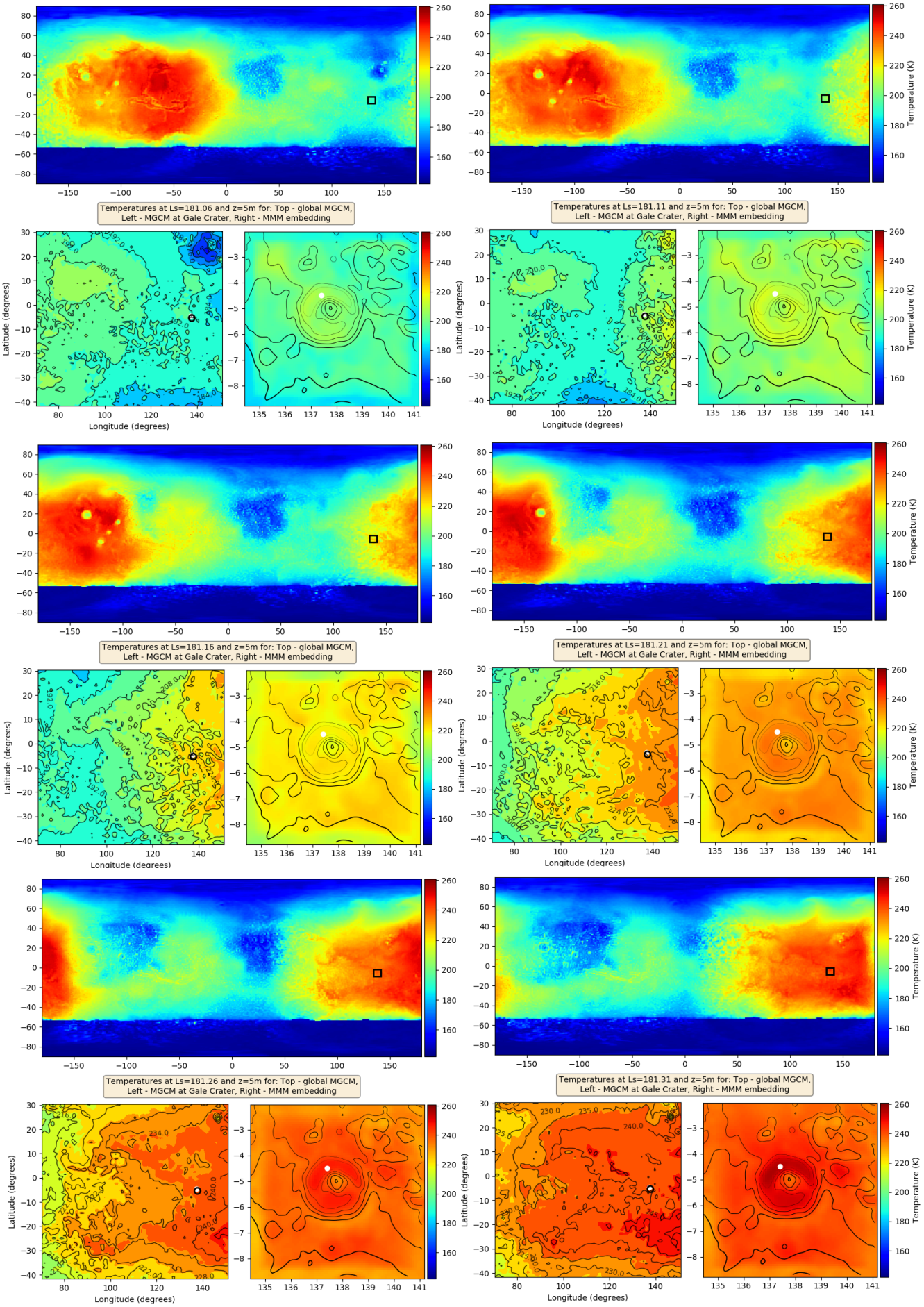


Figure 4.6: The near-surface temperature of the first half of a typical Martian day. The starting time is midnight local Curiosity time. Details are provided in the subchapter.

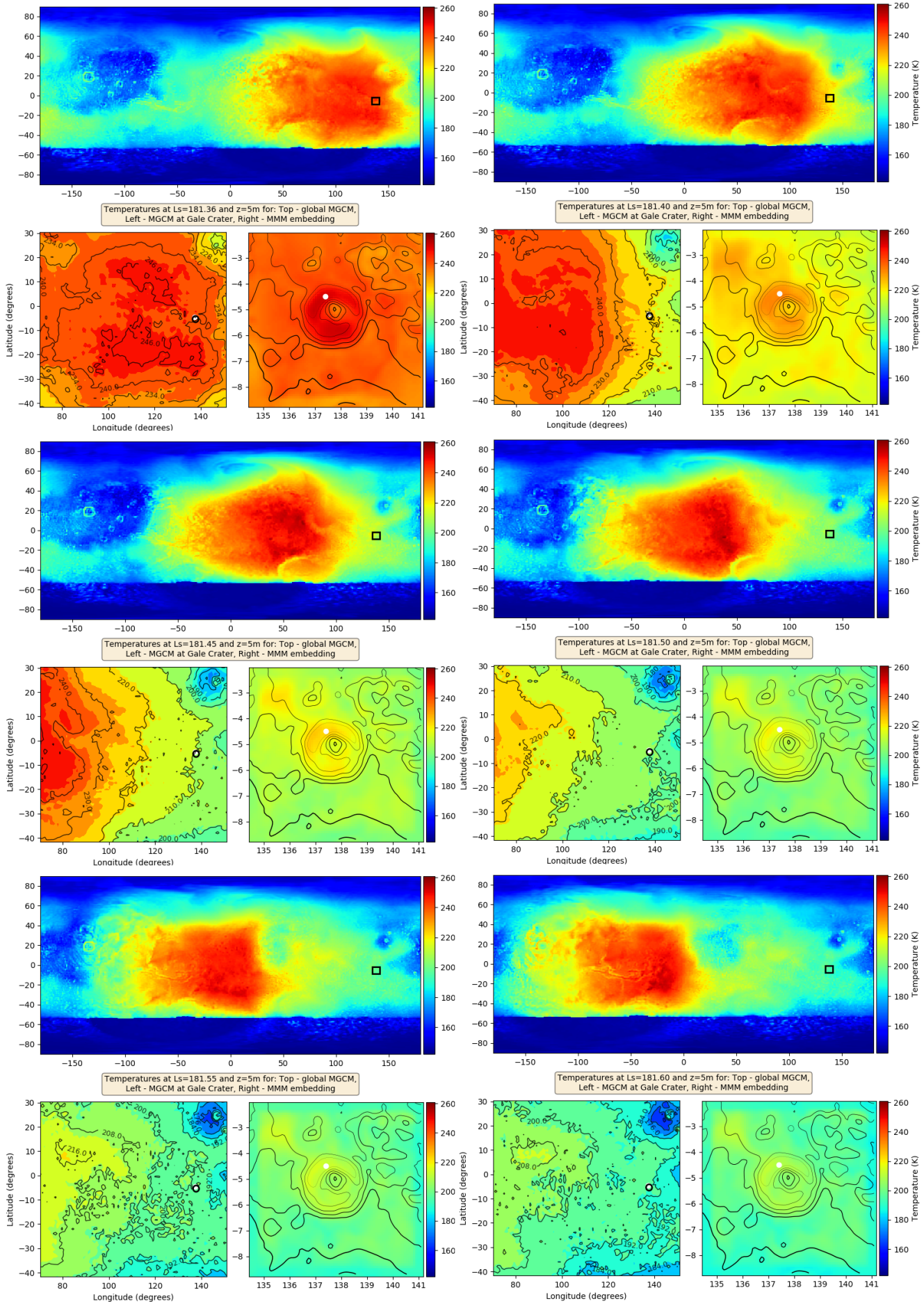


Figure 4.7: The near-surface temperature for the second half of Martian day, continuing from figure 4.6.

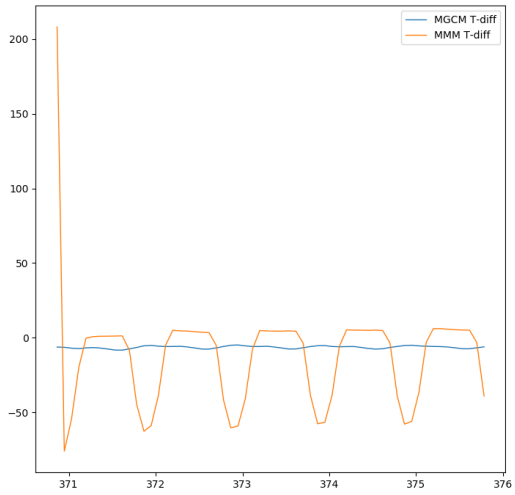


Figure 4.8: The MGCM and MMM averaged atmospheric near-surface temperature minus ground temperature over the same simulated period covered by figures 4.6 and 4.7. This plot is computed for northern autumn.

(Mars's orbit reaches its perihelion at sol 251). Additionally, as the first cell is at 00:00 Curiosity local time, the 12 cells cover 24 Martian hours reaching 00:00 again in the next day.

Near-surface atmospheric temperature couples to the solar tide to form the wave pattern observed in figures 4.6 and 4.7. The T170 spectral resolution is enough for the major landscapes on Mars to perturb the temperatures of the near-surface sigma level (Olympus Mons, Tharsis Volcanic region and different major chasms are visible). A notable characteristic of the higher peaks on Mars, is the appearance of a hot ring of air encircling them during the night. This is most prominent for Olympus Mons, but likewise certainly noticeable for the smaller volcanoes. This phenomenon is interpreted as a down-welling from the top of the volcano, which gets adiabatically compressed as it flows down the volcano (more details of this effect are discussed in Spiga et al, 2011).

The migrating Martian diurnal tide is an exaggerated version of Earth's. A lack of an ocean system does not buffer the heat wave of the tide and the low atmo-

spheric density makes the near-surface atmospheric temperature highly responsive to the surface temperature. As a consequence, both the atmospheric near-surface at and surface temperature diurnal evolutions are almost identical in profile but in a lagged manner. The MMM and MGCM average temperature difference between the atmospheric near-surface and ground temperature for this period are plotted in figure 4.8. As mentioned, the MGCM near-surface atmospheric temperature is generally lagged to the surface temperature ($T_{ambient} - T_{surface} < 0$ throughout), with the usual difference of about -5K. The MMM near-surface to ground temperature difference, however, has a definite diurnal pattern. During the morning heating, the surface decouples from the environment, getting to ≈50K hotter compared to the environment. At night the surface-to-ambient coupling recovers, but this time, the surface temperature falls slightly behind. This is to be expected as the MMM is a regional model, thus it "loses" the information of the solar tide temperature profile from around the planet. As the MGCM is essentially unaffected by this (the surface and ambience diurnal tides are always present on the field-map at any time), it maintains a constant lag throughout. There are two more features to note in figure 4.8, such as the MMM spin-up during the first sol and the incorrect time-stamps from the MGCM (a sol should start where the function from the MMM temperatures is constant, i.e. night-time). The spin-up period for the MMM is approximately 1 sol long, hence the meteorology of the first sol are avoided. The time-stamp problem is discussed thoroughly in chapter 5 and is due to the rounded (but incorrect) Martian year length of 668 sols used in some parts of the MGCM, instead of the more accurate ≈668.522 (with even more undisplayed decimals as a float number) which is used throughout the projects presented in this thesis.

The regional plots from each cell from figures 4.6 and 4.7 show a diurnal tide shaped by the topography, especially Elysium Mons (top-right of the bottom-left regional plot, or 25.02°N and 147.21°E) and Hellas Basin (bottom-left of the same

plot, or 42.4ºS and 70.5ºE). The temperature distribution is stretched towards Hellas Basin as the solar tide arrives at this location (figure 4.6 - bottom left cell), due to the fast decrease in altitude. Similarly, Elysium Mons presents a significant temperature contrast to its surrounding plains during the night. This should generate mild forcing from katabatic winds. Anabatic/katabatic winds from both these landmarks are an expected influence on the Gale Crater region. The MMM - local perspective of each cell is "painting" fine details in the temperature field from the MGCM. The crater floor is consistently warmer compared to its surroundings, implying major topographical influences to the environment inside the crater. Even if the topography of Gale Crater is not as extreme as other regions on Mars, it has a central peak (Mount Sharp) rising 4.5 km above the southern crater floor and 5.5 km above the north side, where Curiosity is located. The crater rim also has a sharp elevation. Its south side is even higher than Mount Sharp and its northern part is ≈2km higher than the crater floor behind it. These are significant features that are not captured in the MGCM T170 topography. Tyler & Barnes (2013) noted that the forcing due to the local topography couples to the diurnal tide. However, they maintain that these are not in phase with the circulation from the topographic dichotomy. The crater boundaries exhibit a stronger vertical motion which is described by Rafkin et al (2016) and also confirmed later in the current chapter.

4.2.2 The velocity field

Figure 4.9 presents the MGCM and MMM horizontal wind field velocities, which are extracted from the first sigma level in the case of the MGCM (top graph), and from an interpolation to a layer 5m above the local topography in the case of the MMM (bottom left graph). The bottom right figure displays the MMM temperature field from the same layer above the local topography. The arrow field depicting velocity vector directions are unit vectors produced from binning

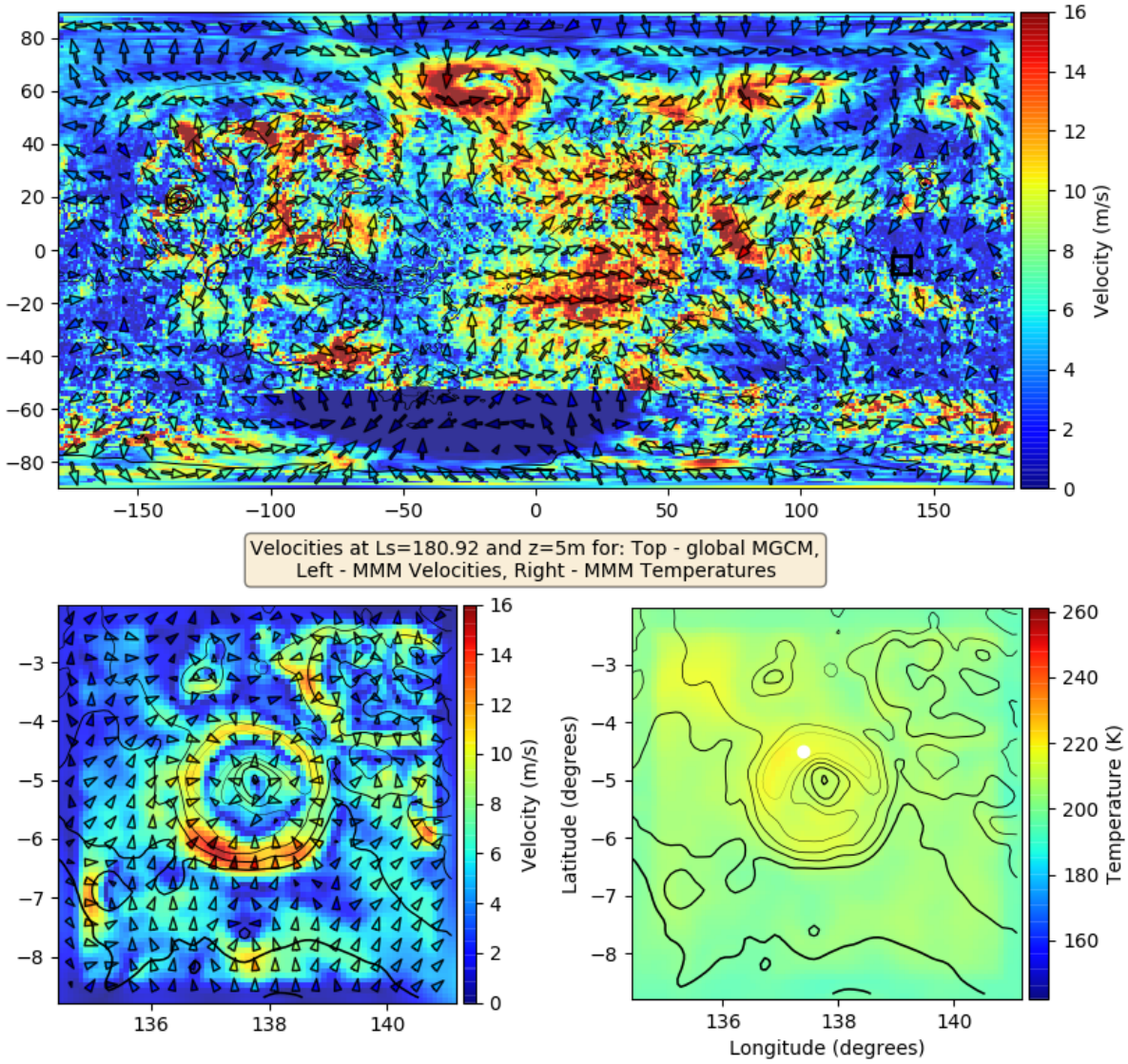


Figure 4.9: Katabatic winds from Mount Sharp and crater rim towards the crater floor. The top plot is a velocity map from the MGCM, while the two bottom plots are MMM velocity (bottom left) and temperature (bottom right) maps. The arrows are binned unit vectors of 10x10 data points from the MGCM and 4x4 points from the MMM. The speed is encoded in the colour map. Local time at Curiosity for this particular figure is 22:00.

(averaging) groups of 4x4 velocity data points for the MMM plot and 10x10 groups for the MGCM plot. The arrow colourings denote wind strengths of the binned groups and the background colour map is the full velocity magnitude field for

both plots. The time of the year from figure 4.9 is again northern autumn and the time of day is the start of night (as noted from the MGCM plot above).

The MGCM top plot in figure 4.9 simulates two notable near-surface quasi-stationary atmospheric gyres that persist for several sols in the northern subpolar region (-50^{o} to 0^{o} longitude and 60^{o} to 110^{o} longitude) and slowly migrate west, opposite to the diurnal tide. These are driven by baroclinic instability from the circulation above; more details can be found in Held & Hoskins (1985). A more important feature is the simulation of the Elysium Mons and south highlands nocturnal katabatic winds that were proposed previously as a regional influence to the Gale Crater circulation. The crater is marked by a black rectangle (which denotes the MMM embedding region) and Elysium Mons can be located $\approx 30^{o}$ in latitude to the north of this. Gale Crater is likewise in the nocturnal circulation regime. Turning to the two MMM plots below in figure 4.9, the same katabatic winds can be spotted from Mount Sharp and the crater rim downwards to the crater floor. This enhances the cooling of the crater floor environment, by adding to the existing temperature decline due to heat loss from the surface layer to the ground (therefore coupling to the ground temperature; figure 4.8) and the long-wave radiation loss (partially) to space. Even so, the crater floor environment is indeed hotter than the surrounding region due to its significantly lower relative altitude. This suggest that the lapse-rate of the column of atmosphere above the crater floor is more important than the cooling from the local circulation.

Anabatic/Katabatic winds

Nocturnal katabatic winds from each season are presented in figure 4.10. In this plot, the arrow field and colour map are the same as in figure 4.9 - MMM wind plot. These have similar strengths, but are superimposed on different regional circulations. In each season, the night starts with north-to-south winds from the circulation outside the crater and adds constructively with the down-slope

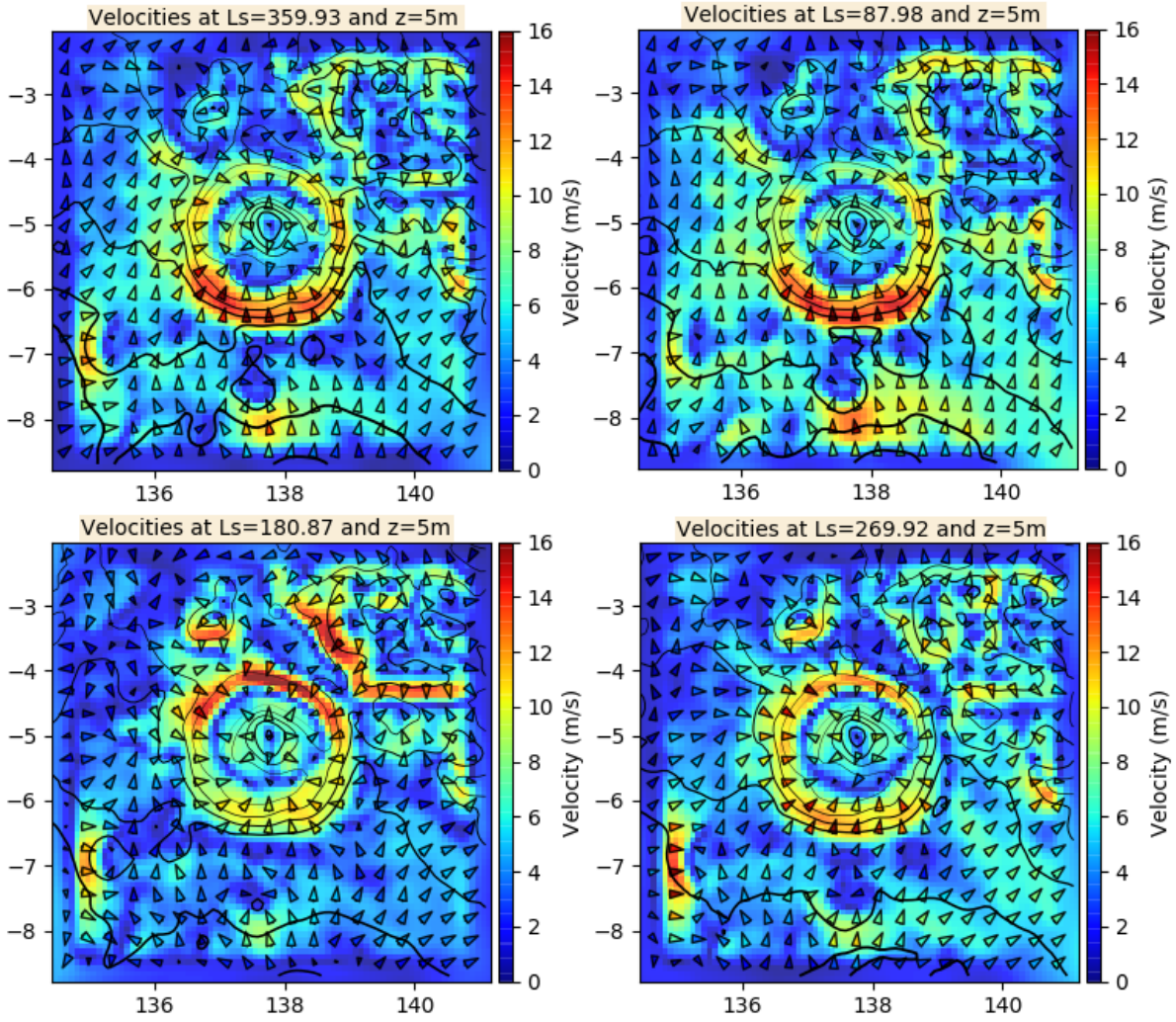


Figure 4.10: Katabatic winds from Mount Sharp and crater rim towards the crater floor. The arrows are binned 4x4 groups of MMM velocities and the local times for each plot from left to right are: top-left 00:00, top-right 02:00, bottom-left 20:00, bottom-right 22:00. The local times are chosen to capture some distinct the change in strength of the katabatic winds around the crater rim.

katabatic winds from the north side of the crater rim (and destructively with the south katabatic winds, weakening them). Instances of nocturnal circulation from the beginning of the night are given in the northern autumn (Ls $\approx 180^o$) and winter plots (Ls $\approx 270^o$). This flow reverses as the night progresses, favouring (constructively) the south katabatic winds, as seen in the northern spring and summer plots (which are taken from later in the night). The reason for choosing

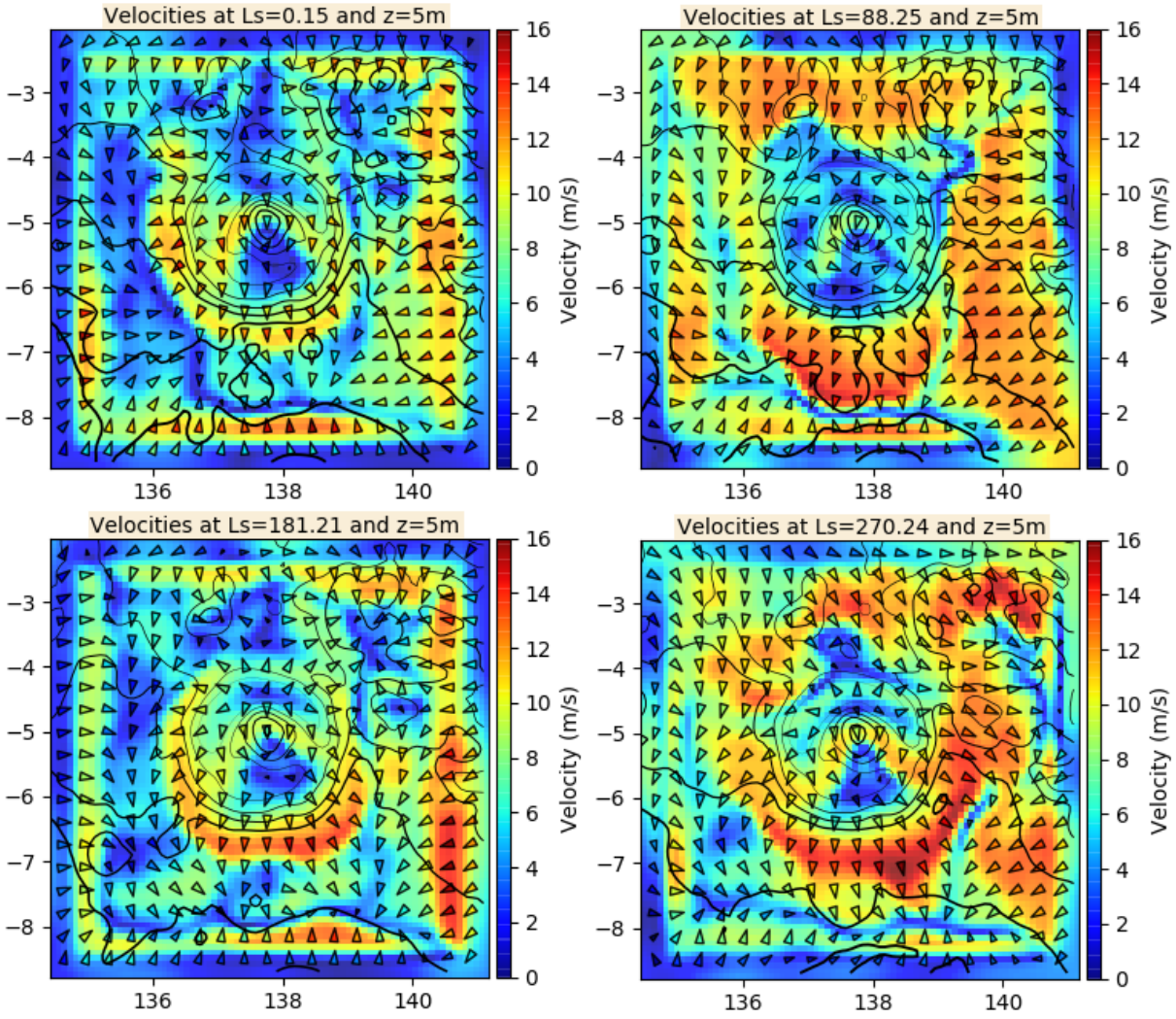


Figure 4.11: Anabatic winds from the crater floor towards the rim and partially towards the top of Mount Sharp. Local time is 10 am in all 4 plots.

different local times was to show the change in wind strength from the north rim side at the start of the night to the south side as time progresses, as this is a recurrent feature, independent of season.

In contrast, morning (strongest) anabatic winds seem to not be the sole contributor to the local circulation during Martian mornings. Figure 4.11 presents the four corresponding typical mornings at each start of season. Intuitively, anabatic winds up the slopes of the crater rim and central mountain should have been the strongest component to the crater circulation (in a similar fashion to the

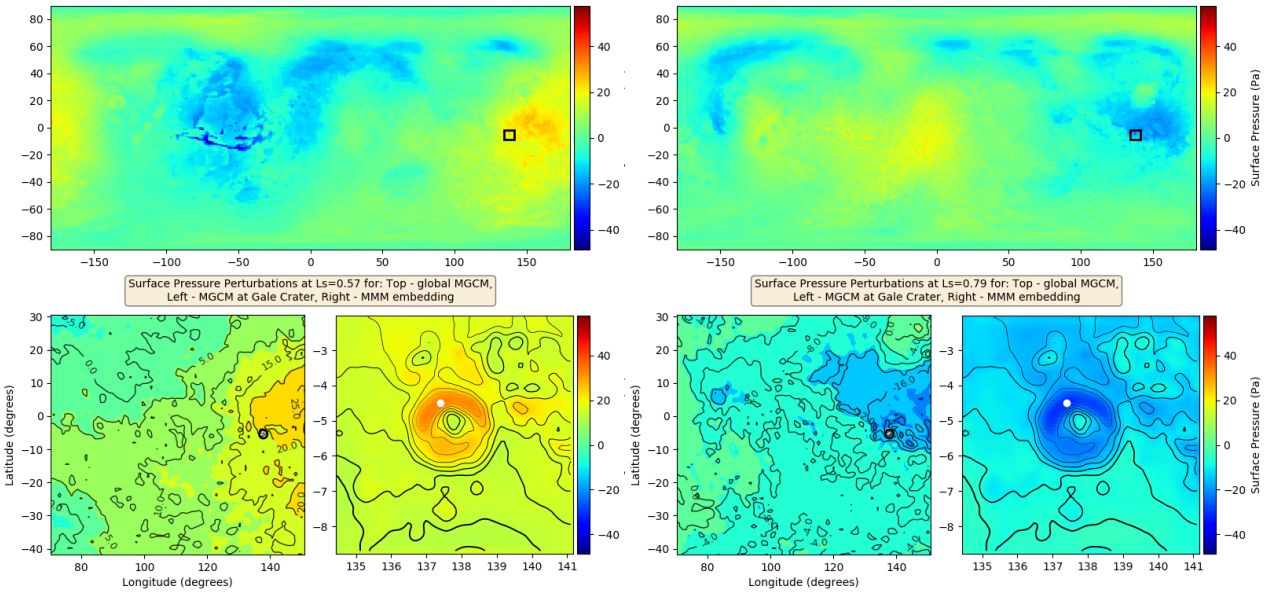


Figure 4.12: Day-night surface pressure perturbation color maps from northern spring. Depicted are the high (left maps) and low (right maps) pressure nodes closest to Gale Crater. These correspond to the two spring plots given at the top of figure 4.13.

four katabatic wind cases presented previously). However, they appear to display a particular scenario of southwards wind splitting, which is superimposed over the morning anabatic winds, in each season. Essentially, winds from the thermal tide which are directed south, split up at Mount Sharp forming the inverted "V" in the colour map. As a result, the anabatic northwards wind immediately south of Mount Sharp reverses direction when combined with the thermal tide wind. The resulting weak flow at the south part of the crater floor can be noticed in each of the four cases.

The start of night persistent southward flow outside the crater is intriguing because it takes place at all seasons. Similarly, the southward flow-splitting from each morning is a clear influence that is external to the crater. For this reason, the day-night variability, including the MGCM-global view for broader context will be investigated next.

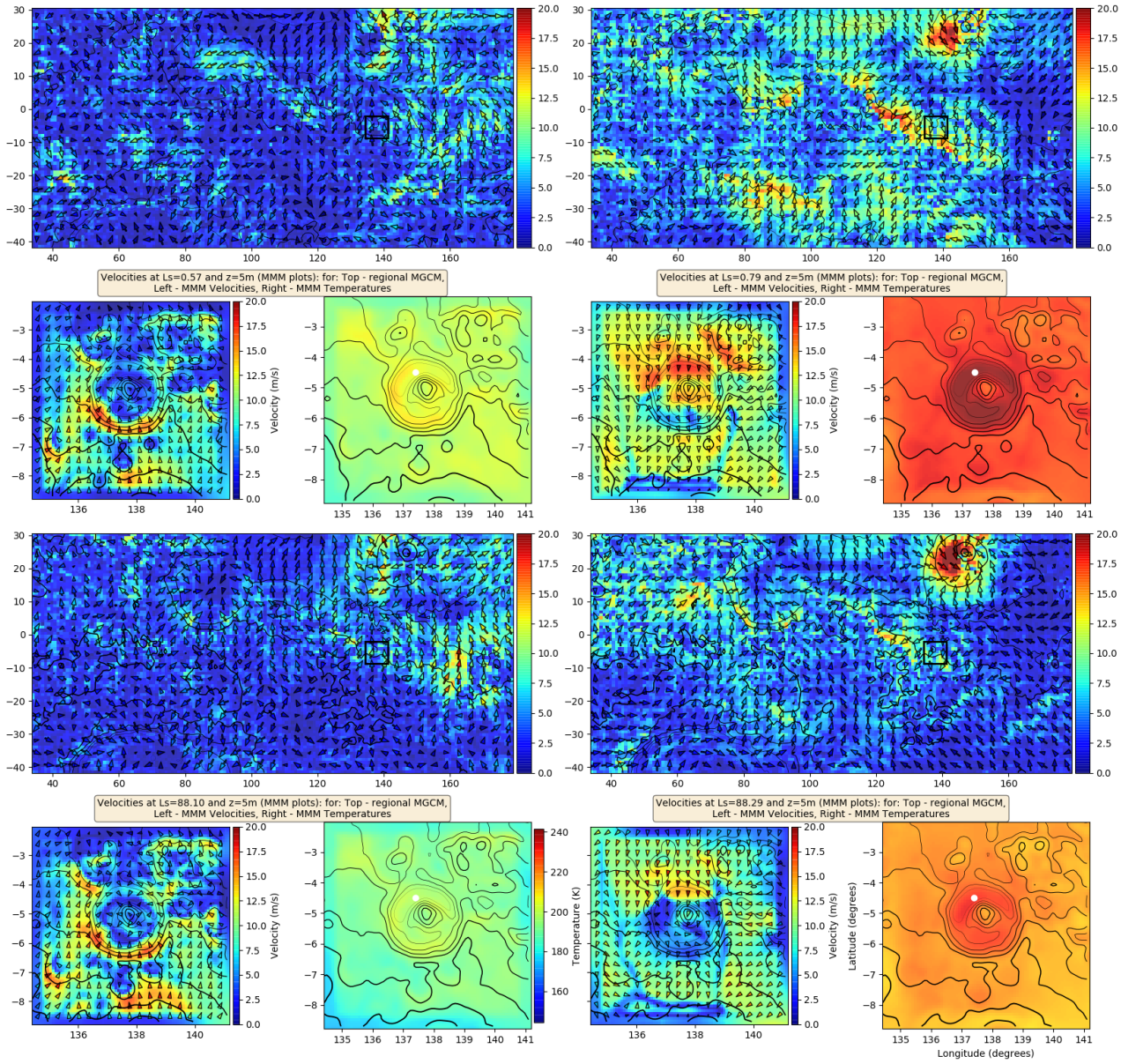


Figure 4.13: Day-night velocity-temperature plots for northern spring and summer. The plots from the left column are taken from the time of day when the high pressure node is closest to Gale Crater. Similarly, the the right column plots are taken from the time of day when the low pressure node is closest to the crater. Additionally, the arrow density is produced from binning groups of 10x10 data points from the MGCM and 4x4 from the MMM.

4.2.3 Day-Night circulation

A day-night analysis of the general and local circulations is provided in figures 4.13 (for northern spring and summer) and 4.14 (for northern autumn and winter).

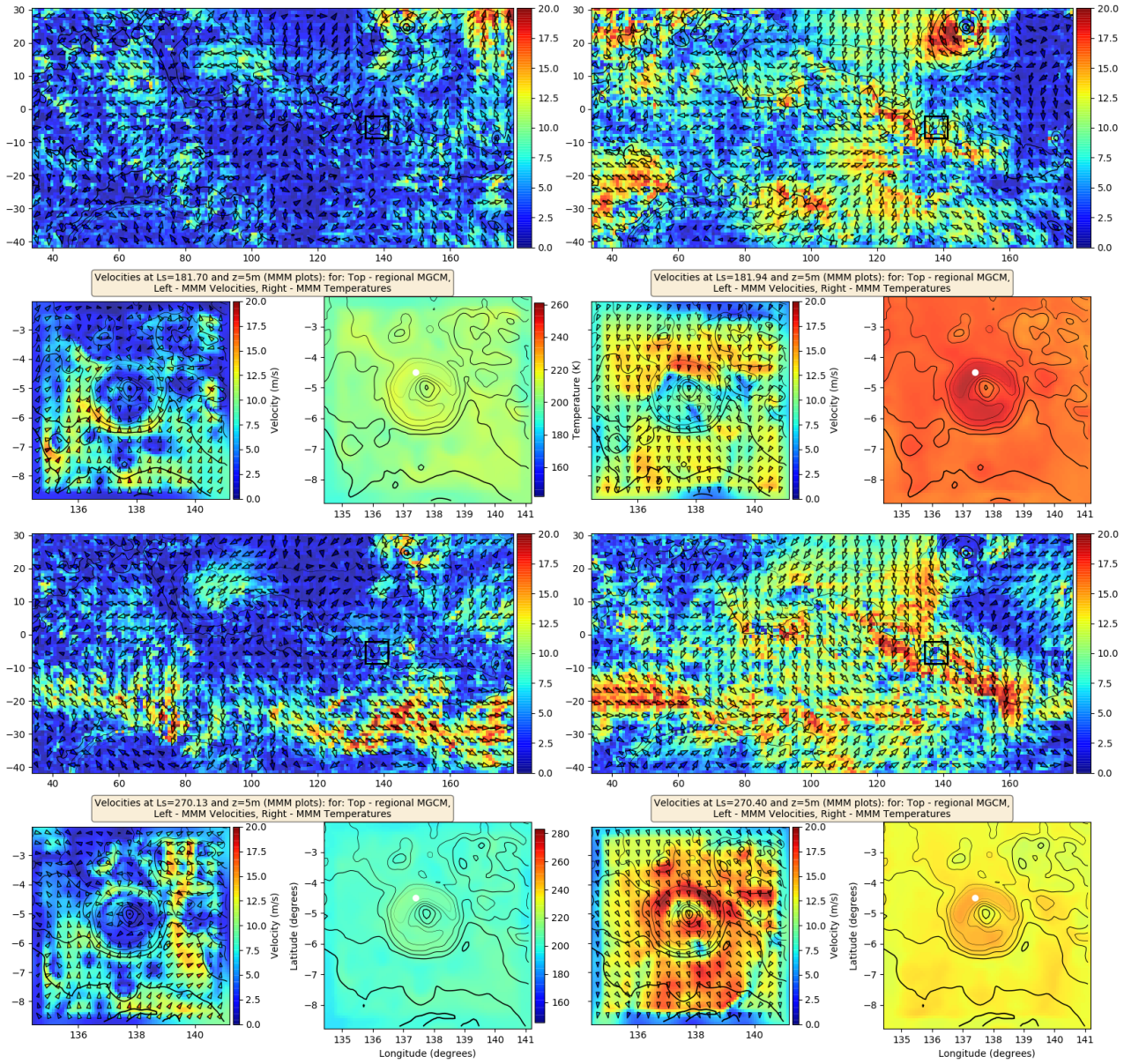


Figure 4.14: Day-night velocity-temperature plots for Northern autumn and winter. These are similar to figure 4.13 but for the other two Martian seasons.

These are series of cells containing three plots as in figures 4.6 and 4.7 for the thermal tide. However, the left column of cells represents the time-frames where the high-pressure node from the tide is the closest to the crater. Similarly, in the right column of cells, Gale Crater is under the low-pressure node. The position of the pressure nodes were located using diurnal pressure-perturbations plots of the same periods. Out of conciseness, all eight pressure-node plots will not be

displayed here. An example of pressure-perturbation plots for northern spring can be seen in figure 4.12, which has plots corresponding to the time of day when the high (left plots) and low (right plots) pressure nodes are closest to the crater. As stated, these were used in selecting velocity maps from figures 4.6 and 4.7. To return, each cell (from the spring-summer figure 4.13 and autumn-winter figure 4.14) has the same structure of plots as in figure 4.9 with the MGCM regional winds plotted above and the MMM winds for clarity and temperatures in the bottom left and right plots, respectively. Essentially, figures 4.13 and 4.14 are sequences of plots (in cells) showing the day-night circulation from each season, both globally and locally.

There is a common set of processes simulated in figures 4.13 and 4.14, which are occurring in the Gale Crater region. At night-time, the katabatic winds from Elysium Mons clash with the katabatic winds from the highlands to the south, creating a calm region of upwelling in the middle. As mentioned before, the high pressure node from the cold part of the diurnal tide is present in that region, making it a point of high pressure. Thus, both down-slope winds are countered by the tidal winds diverging out of the cold node. At Gale Crater, the highland winds are intensifying the south-part of the rim's katabatic flow more intensely than the global tidal effects. During daytime, the anabatic winds from both Elysium Mons and southern highlands are combining destructively in the Gale Crater region, with the highlands winning the trade. The low pressure node is between these two landmarks, thus countering both. At Gale Crater, the winds are predominantly south, and due to the stronger anabatic winds from the highlands. The north part of the crater floor is covered in strong winds which seem to split and rise forced by the presence Mount Sharp. The northern summer night at Gale Crater (from Ls 88.10) is in an almost identical situation to the spring case above it. The regional meteorology in summer is less similar, with the diverging winds from the high pressure node and the highlands both forcing the winds

at Elysium Mons, to the north of the crater. Daytime autumn and winter are changing the regional situation. The low-pressure node winds and highlands upslope winds combine to divert the Elysium Mons anabatic winds predominantly southwards. This makes the Gale Crater region particularly windy, with a strong south component (especially in winter). These winds meet the crater topography to produce local drag, with emphasis on the central Mount Sharp, which bifurcates the wind, and the north side of the crater rim, which deflects flow towards the crater floor. At the global scale, autumn and winter are significantly more active, with large atmospheric gyres in the northern hemisphere around 60N to 80N latitude and flows which overwhelm the topographical circulations throughout the planet, especially during daytime.

These are only cases when the tidal pressure-nodes are closest to the crater. The nodes fall between Elysium Mons and the southern highlands, but as the tide progresses, these effects will differ.

Nocturnal low-level jet

As mentioned earlier in the beginning of the chapter, low-level nocturnal jets are common on Mars due to the significant variability in depth of the daytime PBL, between 1km at night to 10km during the day. At the same time they pose a challenge for Martian atmospheric circulation models, as the variation of the PBL in Earth's case is between 0.5km and 1km. To have an appreciation of the low-level nocturnal jet, one must first look at the boundary layer itself. The expected features about the Martian PBL were discussed in subsection 4.1.1, previously. Essentially, the low-level nocturnal jet is an effect derived from the highly dynamical PBL on Mars. For this reason, height-longitude sections from the latitude of the peak of Mount Sharp are presented in figure 4.15. These 6 plots are divided in two columns, the left one taken from Ls 181.31 (northern autumn), when the temperature at Gale Crater is at its maximum (during the day) and the

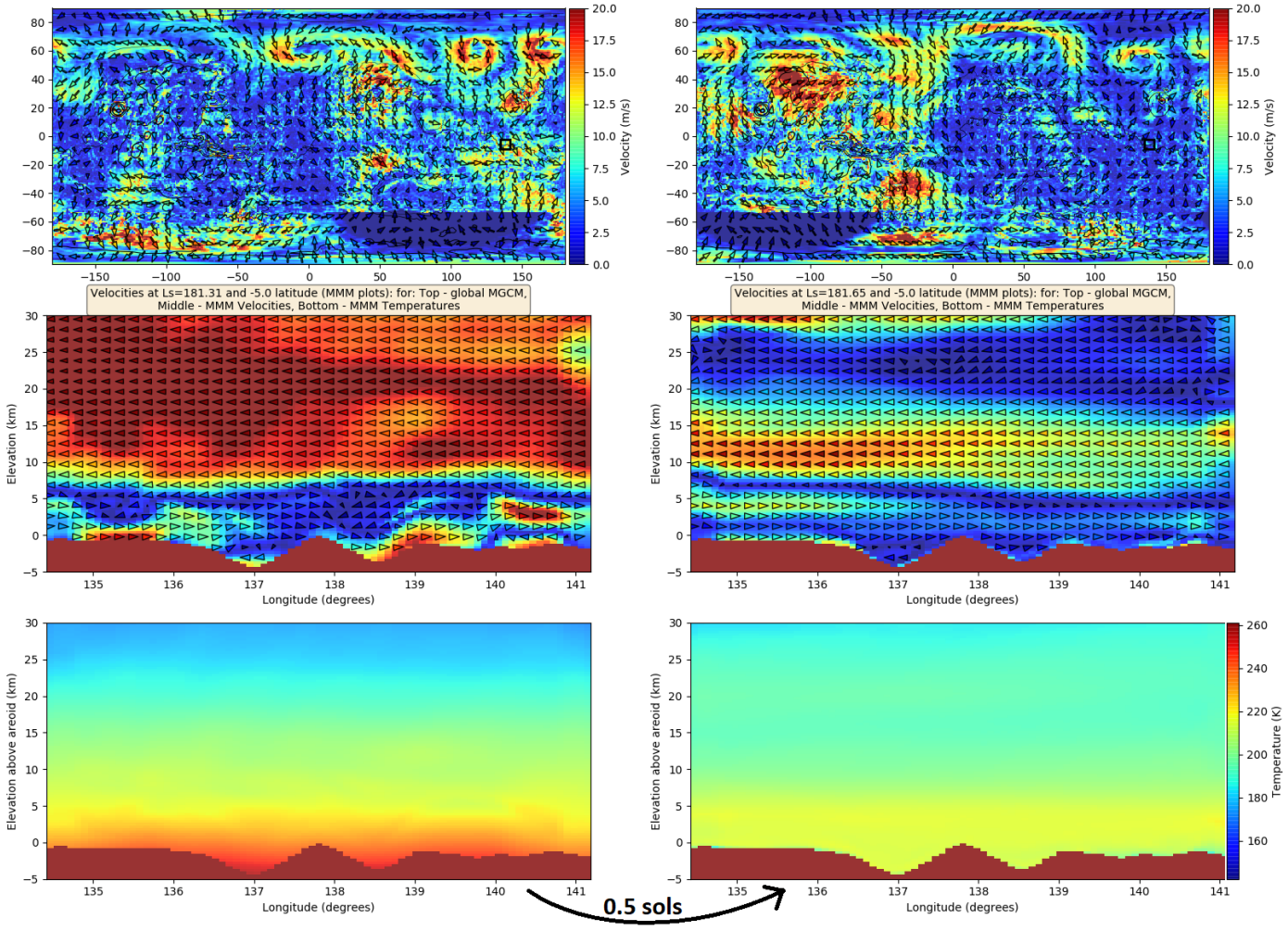


Figure 4.15: The evolution of the planetary boundary layer. On the left, the local time at Curiosity is 14:00, and a developed planetary boundary layer reaches $\approx$ 10km in depth. On the right the local time is 02:00, presenting a collapsed boundary layer with a nocturnal low-level jet developed in it the absence of the forcing from the PBL.

right column corresponds to the lowest temperatures (during the night). In each column, global surface wind activity from the MGCM is shown first (top plot) and has the same details as the previous MGCM plots. The following two are MMM results, interpolated on to vertical levels relative to the aeroid, ranging from -5km to +30km in altitude. The atmospheric fields are zonal cross-sections, extracted from the latitude of Mount Sharp's peak (5°S). The middle plot is the velocity vector field and the bottom plot is the temperature field. The arrow density for

the sideways MMM velocity plots is produced from binning (averaging) groups of 4x2 (in longitude×height) MMM velocities, to capture the finer vertical structure of atmospheric motions.

The temperature field during the day is somewhat continuous from the surface upwards. However, in the mixing layer, instances of turbulent mixing alternating with instances of static stability (Stull, 1991) make it far from continuous. Essentially, the hot surface and colder upper atmosphere, even if they are in convective instability, will not mix in a turbulent fashion. Rather, hot buoyant parcels of air will rise up, from the surface to the entrainment zone, which seems to start from altitudes of 7km and end at 10km. At the same time, negatively-buoyant cold parcels will descend from the entrainment zone to the surface. This is known as static stability; more details are provided in Stull (1991). These are small in scale and are not picked up in the vertical plots. However, larger scale convective plumes are resolved and shown later. The other mixing mechanism is turbulence, where under highly unstable conditions, further mixing occurs; this was discussed in the first half of the chapter. At night, the PBL collapse generated a nocturnal jet at ≈5km and appears to be influenced by the topography. This common occurrence is discussed next.

The Martian nocturnal low-level jet above Gale Crater is captured in a series of vertical plots of the four MMM seasonal cases, and distributed in figure 4.16. These show remarkable variability between the four seasonal cases. However, a more systematic study is required to actually prove that the nocturnal low-level jet is solely seasonal dependent. Nonetheless, the jet is most prominent in the spring (Ls = 0.57). Strong winds reaching up to 20m/s, are contained in a broad 5-6km wide jet around 5km altitude. The relatively calm free atmosphere coupled with the significant height of the jet, leave it unperturbed. The same can not be said about the summer case (Ls 87.98), which displays a weak and perturbed jet at 2km altitude. Close to the second equinox (autumn) at Ls 181.55, the jet seems

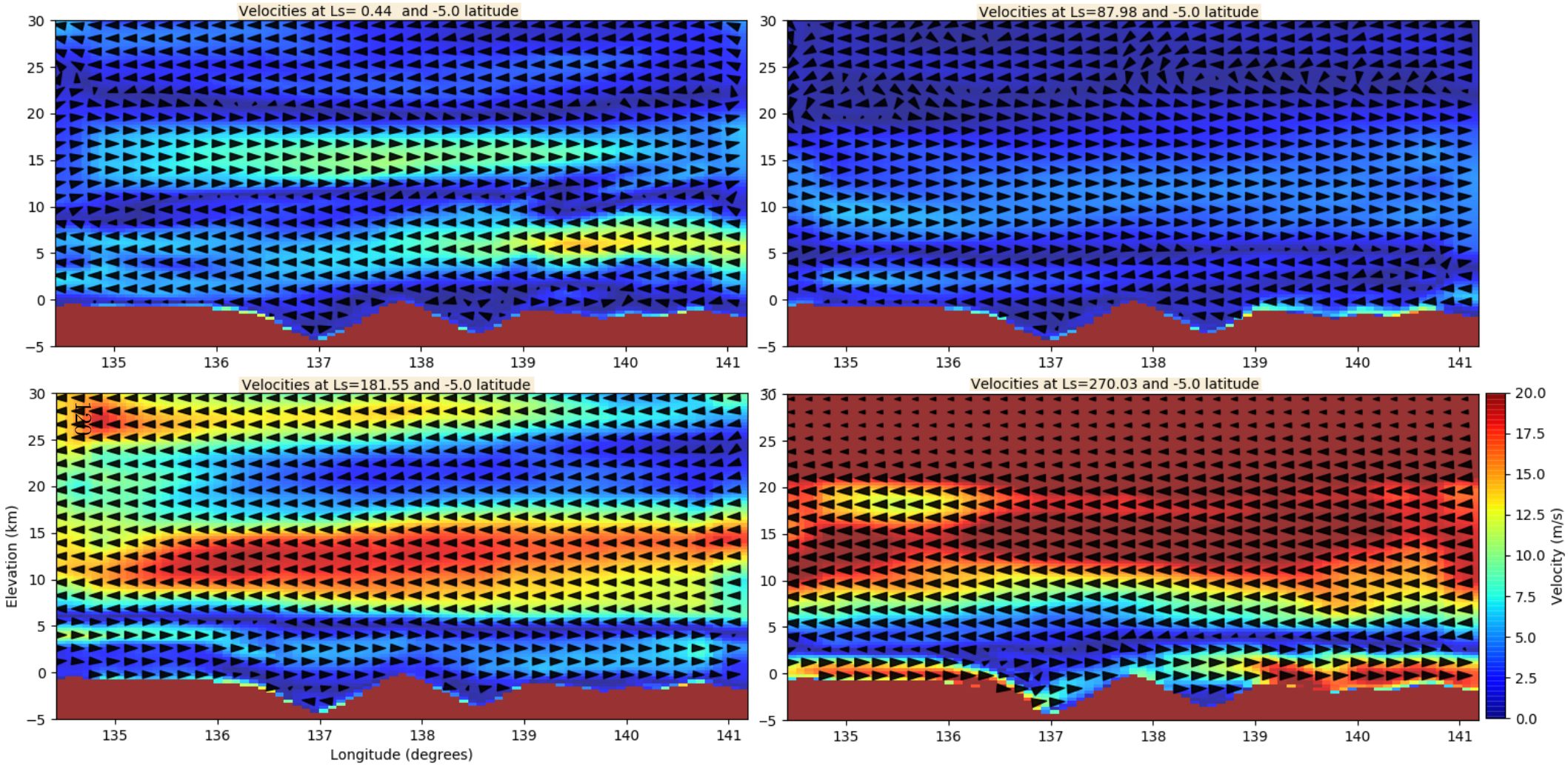


Figure 4.16: Instances of nocturnal low-level jets from each seasonal case. Local Curiosity times are all 00:00 and the plots represent the following seasons: top-left spring, top-right summer, bottom-left autumn, bottom-right winter.

to return higher up, but still below 5 km and is perturbed by the presence of Gale Crater. But the most remarkable is northern winter. The nocturnal low-level jet is so low that it even gets deflected towards the west side of the crater floor. And indeed this is the case for all the nights simulated for the winter case. The free atmosphere winds also descend during the night, and returns to ≈10km in daytime. Another general feature for all four cases is that the jet direction seems to be reversed compared to the free atmosphere which clearly distinguishes it from the free circulation.

Daytime convection

Simple free convection was actually hard to find. As mentioned previously, the horizontal scale of most convective vortices is outside of the resolving capabilities of the current MMM simulations (5km×5km boxes). To put things into perspective, the average dust devil on Earth has a diameter of approximately 10m with a height of 500m. Fortunately on Mars, the scales are bigger, with a diameter of 100m to 1km and altitudes above 5km, but they are still in the grey-zone of the MMM. The best scenario with something resembling convection was found away from the crater, at 139.1° longitude and −8° latitude, in figure 4.17. For this reason, convective vortices and dust devils are normally parametrised.

Gravity waves

Gravity waves are easier to find, as they preserve information about the topography in their wavelengths. More precisely, they preserve the horizontal scale of the topographical obstacle that generated them. Instances of gravity waves are extracted in figure 4.18. This figure contains four plots taken from the four seasonal cases. In the vertical, the extrapolation is again over layers above the aeroid, but this time the vertical range is from -5km to 2km. The extreme aspect ratio is helpful for spotting vertical motions. The orographic gravity waves in all four

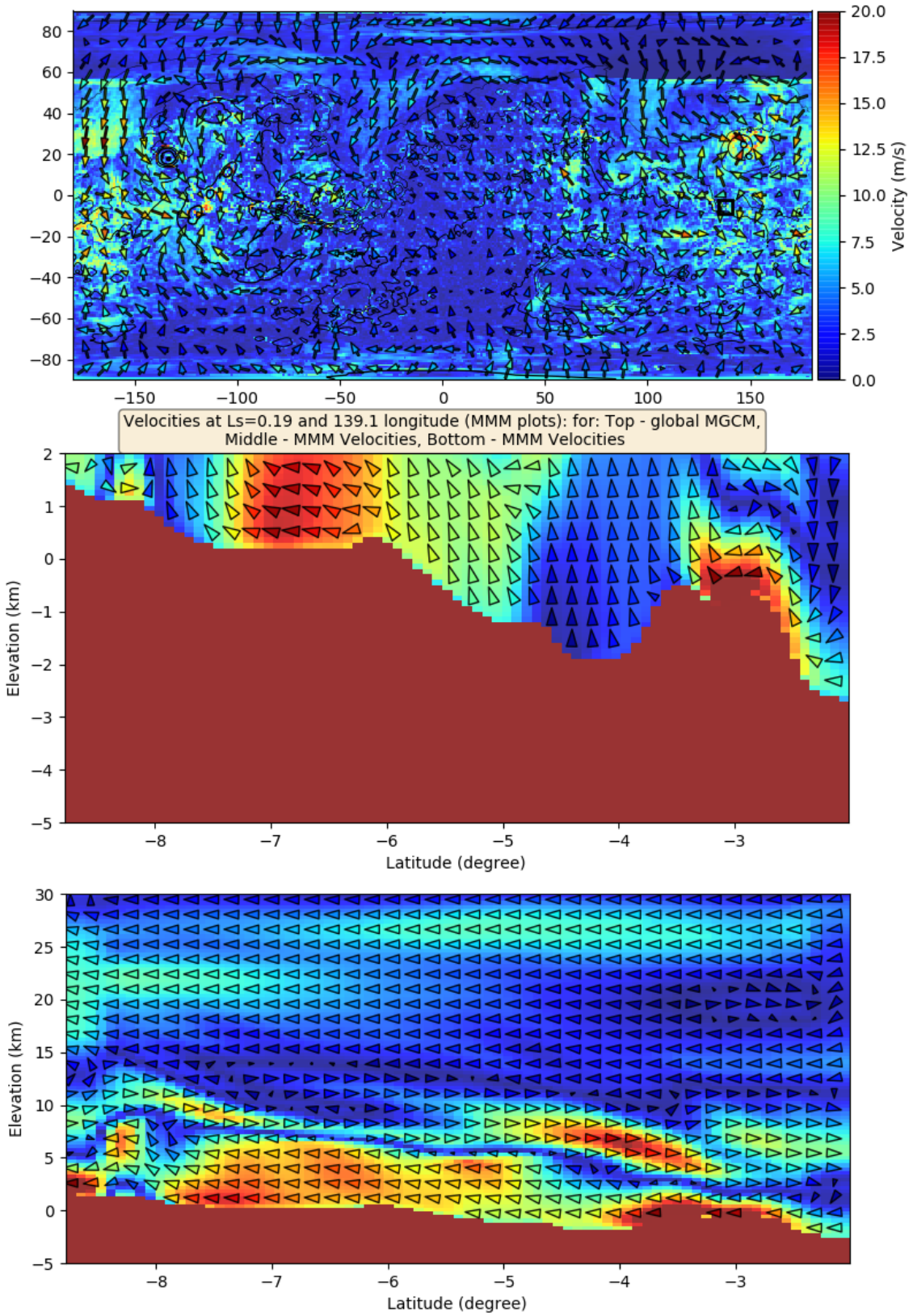


Figure 4.17: An instance indicative of convection at 139.1° in longitude and −8° latitude.

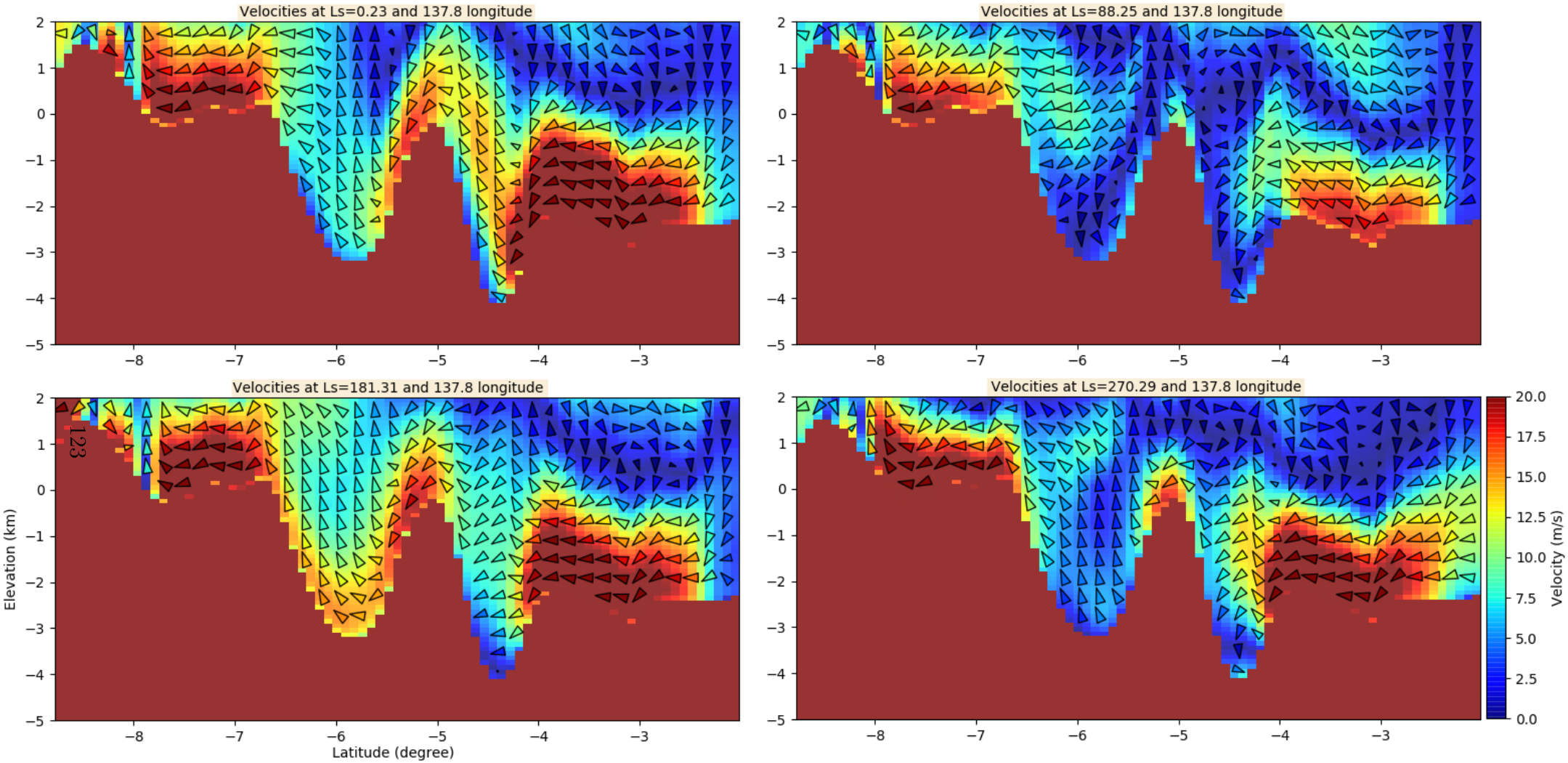


Figure 4.18: Instances of orographic gravity wave activity at the north crater rim. The local Curiosity times and seasons are: top-left spring 14:00, top-right summer 14:00, bottom-left autumn 14:00, bottom-right winter 12:00. Capturing gravity waves was the priority so the times of day are chosen arbitrarily.

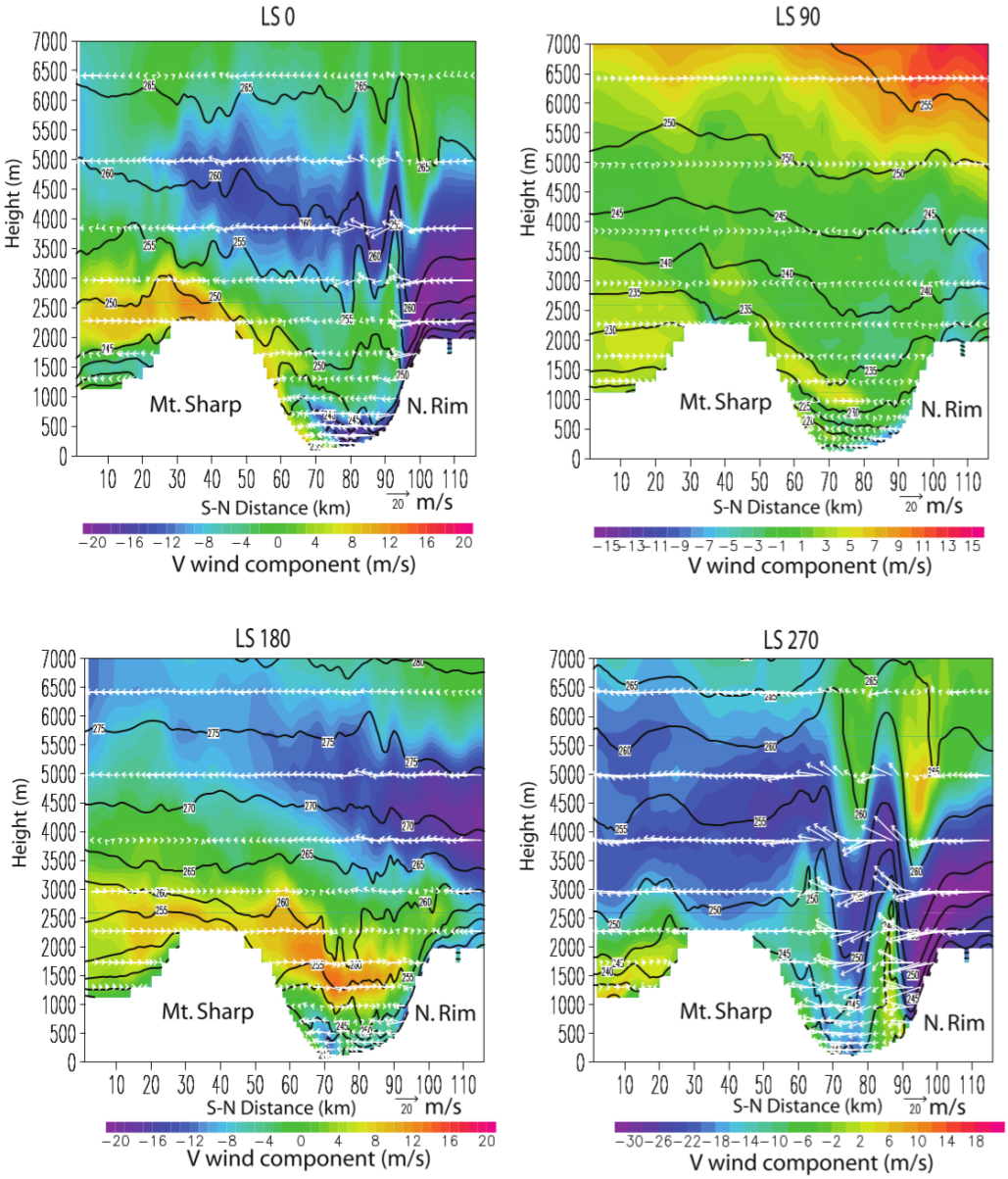


Figure 4.19: Gravity waves simulated using MRAMS (Mars Regional Atmospheric Modelling System), from boundary conditions provided by the NASA Ames GCM. (Rafkin et al, 2016).

cases are not generated by an obstacle, but by a forced Bernoulli-type of downslope flow. Essentially, air has to fill the gap represented by the crater floor, in a near-incompressible fashion, thus deviating downwards. After this point they are dependent on gravity to keep propagating. Moreover, the MMM outputs suggest that vertical motions favour the crater rim and not Mount Sharp. Rafkin et al

(2016) reached the same conclusion, however, we can not generalize as both their study and ours are performed over 5-7 sol intervals.

The plots from Rafkin et al (2016) are given in figure 4.19. These capture orographic gravity waves through a similar set-up with the MGCM+MMM. The NASA Ames GCM provided boundary conditions to the MRAMS mesoscale model at a rate of 1/16 per sol. Additionally, MRAMS was configured for grid nesting, with a total of 7 grids used in the simulation of the environment at Gale Crater. The resolution reached by NASA Ames GCM→MRAMS is 330m, much higher than the 5km resolution that was used with the MGCM+MMM configuration. However, embedding from a reanalysis is the novel addition of our configuration. This addition will even prove to improve prediction, as forcing fields from data assimilation brings the MGCM+MMM closer to REMS' observations compared to the higher resolution NASA Ames GCM→MRAMS configuration. The symbol "+" was used instead of "→" to denote the interfaced configuration of the MGCM+MMM; details were provided in chapter 2.

Gravity waves produced in both studies (from the MGCM+MMM and Rafkin et al, 2016) seem to ascend, thus reducing the surface wind speeds in the crater. Additionally, the waves increase in amplitude as they dissipate closer to Mount Sharp (in both studies) and the surface winds in the crater floor, thought weak, may be reversed; this is the case in the summer and winter plots of figure 4.18 and also appears in Rafkin et al (2016). This effect is known as rotor circulation and is seen produced on Earth as well (Klemp & Lilly, 1975).

5 The meteorology of Gale Crater from Rover Environmental Monitoring Station

The current chapter focuses entirely on making the comparison between the MGCM +MMM reanalysis and REMS observations. To achieve this goal, the data is brought through a series of preparation methods such as post-processing, rebinning and resyncing. The next step is a mathematical decomposition by projection onto families of orthonormal functions. The employed procedures in this chapter are (a) fast Fourier transforms (FFT) and (b) a data-adaptive singular system analysis (SSA) to separate different components of the time variations from the physical fields, as originating from different physical processes (e.g. diurnal tides, topographic circulations or other weather systems). The chapter will end with a return to the results from previous chapter, to connect the in situ analyses to the MMM-embedded reanalyses.

The methodologies will be introduced in the first part of the chapter (section 5.1) along with a brief description of the expected signals in a Martian tropical environment, where Gale Crater is located. Much of the observed variability from the previous chapter was on diurnal periods and shorter. A fitting result for a tropical troposphere, devoid of superdiurnal meteorology. Their systematic application and the comparison problem is tackled in the later parts of the chapter (section 5.2).

5.1 In situ data analysis. A systematic approach

Predicting observations from REMS can prove difficult. A weather forecast approach (from the MGCM+MMM reanalysis) would provide the hands-on style that many papers have used in the past for predicting or comparing with observations; Pla-Garcia et al (2016), Rafkin et al (2016), Tyler & Barnes (2013),

etc. This hands-on comparison could have been applied in the previous chapter, although its omission was intentional. Instead, the previous chapter merely introduced the typical meteorology at Gale Crater, and a more systematic approach to compare directly with in situ measurements is adopted in the present chapter. The comparison is focused on separating the following:

- Effects that are correctly simulated by the reanalysis and are essentially picked up by REMS,
- REMS observations which are not simulated in the reanalysis,
- Incorrect model predictions that do not appear in REMS observations.

The matching meteorology between model and observations would offer a more precise explanation of the observations themselves. The mismatching predictions are not discarded, as they hold valuable information about the Gale Crater environment and the accuracy of the reanalysis (and how the model could be improved).

To tackle the means of comparing model/reanalysis with local measurements, two methods were applied:

- **Fourier decomposition**. Decomposition into Fourier frequencies is essentially equivalent to best fitting a basis of sinusoidal functions to the data. This approach is thus good at setting up correct diurnal and semi-diurnal cycles and to extract the most influential frequency components and their relative phases. A reconstruction algorithm was also implemented to visualise the combination of certain frequency components (through inverse Fourier transform of particular leading frequencies). The primary aim of this approach here was to isolate and compare diurnal and subdiurnal tides between REMS and the reanalysis. The matching cases are essentially well resolved by the MMM. Furthermore, the correctly resolved cycles are an

indication of how well certain atmospheric tides are simulated in the reanalysis. The final aim is to understand atmospheric tides in the framework of the general circulation.

- **Temporal Singular Spectrum Analysis (or SSA) decomposition**. This method creates an orthonormal basis similar to any time series construction (e.g. Fourier series), with the advantage of not imposing a particular set of orthogonal functions. Instead, the SSA creates its own functions numerically, by finding the corresponding basis that represents the given dataset in terms of particular contributions of recurrent patterns to the total variance of the observed or modelled signal. As an example, a pre-imposed Fourier basis may be forcefully rotated to best fit sinusoidal functions to the data. This is not the case for the SSA as it follows an algorithm similar to the Gram-Schmidt process from linear algebra. An interpretation of the method is: the SSA algorithm searches for temporal correlations in the data and creates numerical orthogonal functions based on this. The downside of this method is the lack of an analytical formula to the orthogonal functions. One can proceed to best fit mathematical functions to the SSA numerical functions, but this is not the scope of the algorithm. The power of the SSA method stems from its ability to create its own functions, regardless of the nature of the data. The method is commonly used in finance for its predictive nature, in noise filtering and tackling the comparison between model and observations in atmospheric Physics. As an example, the method was employed by Collins et al (1996) to study baroclinic waves activity from Viking data. For a review of the entire SSA approach, an inspection of Ghil et al (2002) is recommended.

To perform these studies, there is another problem to address. Raw data from the REMS instrument is sampled in a very disordered fashion. The typical sol

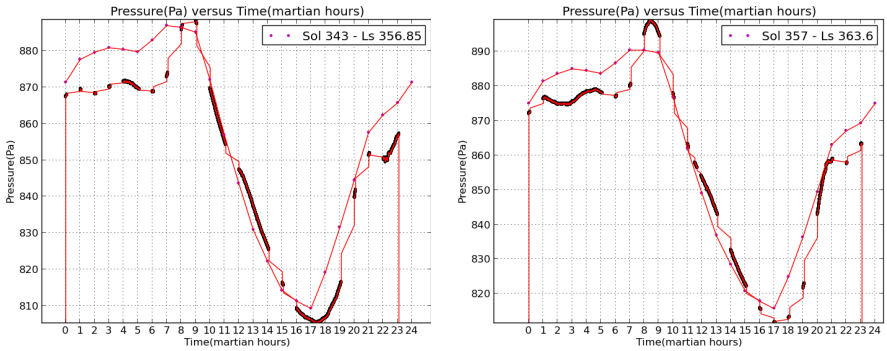


Figure 5.1: The pressure measurements from the REMS instrument (black dots) over the Martian day (sol) 363 (left) and 357 (right) their corresponding moving average. These are compared with their respective MCD surface pressure predictions at Curiosity's landing site.

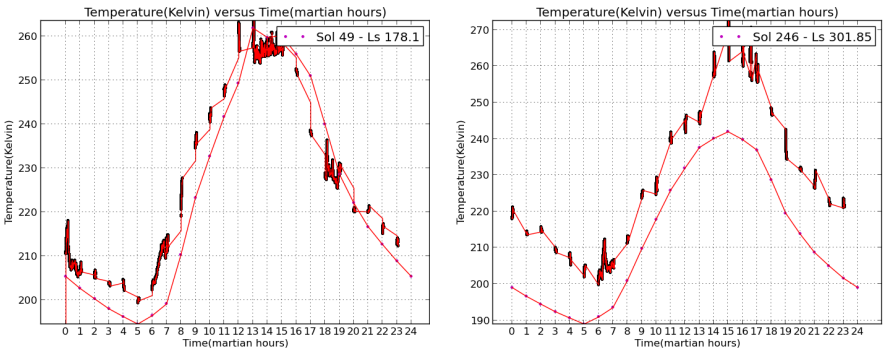


Figure 5.2: The near-surface atmospheric temperature measurements from the REMS instrument (black dots) over the Martian day (sol) 49 (left) and 246 (right) their corresponding moving average. These are compared with their respective MCD temperature predictions at the first vertical level over Curiosity's landing site.

has hourly observational periods. Each hour, a package of observations is taken at a frequency of 1 measurement per second on average, as seen in figure 5.2 - right. These are accompanied by irregular extended intervals of constant high frequency measurements (1 measurement/second) which can last for hours (Martian hours are ≈ 00:01:32 longer than Earth hours and Martian seconds are ≈25 microseconds longer). This high frequency data input can be visible in figure 5.2 - left between 13:00 and 15:00 Curiosity local time, but also in figure 5.1 for a

comparison with surface pressure. This irregularity in the data is a result of mission parameters. In case of relevant events, REMS will switch on the appropriate sensors to capture it. Similarly, the reanalyses (or free runs) have a problem of time-keeping inherited (especially) in large spin-up runs. Essentially, the Martian year after a full 360 degrees in solar longitude gets to ≈668.52 sols, and not the rounded version - 668 as it is seen in some of the MGCM and MCS-postprocessing routines. To perform a systematic study, all data sets had to be converted through back-and-forth rebinning and resyncing algorithms, and are briefly discussed next in 5.1.1. The subsection proceeds with a description of the SSA method in 5.1.2. Moreover, because the main signals from both the observational and model data are dominated by atmospheric tides, a brief introduction to this topic will follow after (in 5.2.1), ending the current subsection.

5.1.1 Preprocessing Data. Rebinning and Synchronization

The over-sampled REMS observations and the regularly sampled MMM output were first brought to a mathematically useful form. First, the irregular (high frequency) Curiosity observations were binned to a regular pattern, at a frequency resembling 10 times the chosen MGCM/MMM output frequency (a tenth of 2 Martian hours = 12 minutes). Second, the model output (which is at the lowest "sampling" - 1 frame every 2 hours) sweeps through the highly resampled Curiosity observations (1 every 12 minutes) to minimise their RMS difference, i.e. to match the MMM's diurnal phase with Curiosity's. This procedure was done using a 1-sol sweeping readjustment routine in both temporal directions. The problem appears only at temporal scales of hours, and fades off over a full year. Surface pressures from REMS accompanied by 3 different model predictions taken from the closest grid point to Curiosity's location, are plotted in figure 5.3. The models are severely underestimating the time mean surface pressures at Aeolis Palus, due to the poor representation of the crater floor elevation in the MGCM topography (at

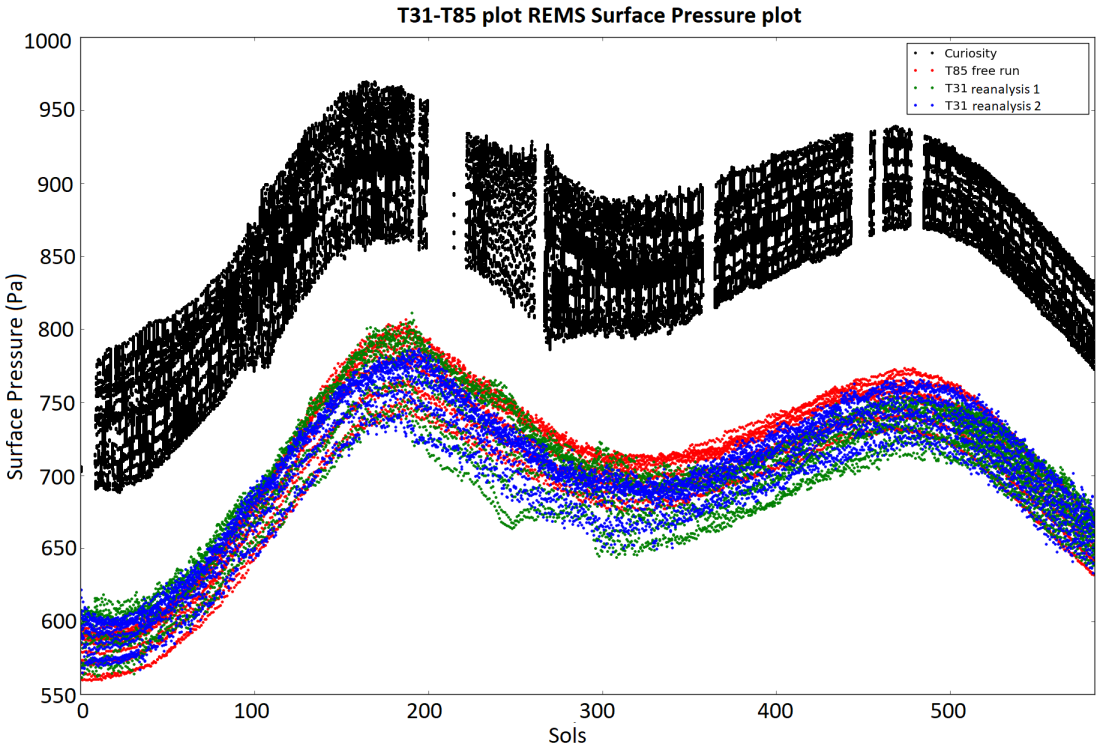


Figure 5.3: Surface pressure values versus time (in sols) from REMS's retrievals (black), our T85 free run (red) over the same Ls, the previous reanalysis performed by Dr L. Montabone over the same Ls (green), our T31 reanalysis at the actual time with Curiosity (blue). The reanalyses values are taken from the closest grid point to the rover's position. Sol 1 on the x-axis is the first sol when Curiosity started gathering data. No preprocessing was applied to the MGCM runs, proving that the intrinsic deviations from MGCM time-keeping is insignificant when viewing a full Martian year. This is not the case for a diurnal analysis, as seen in figure 5.4

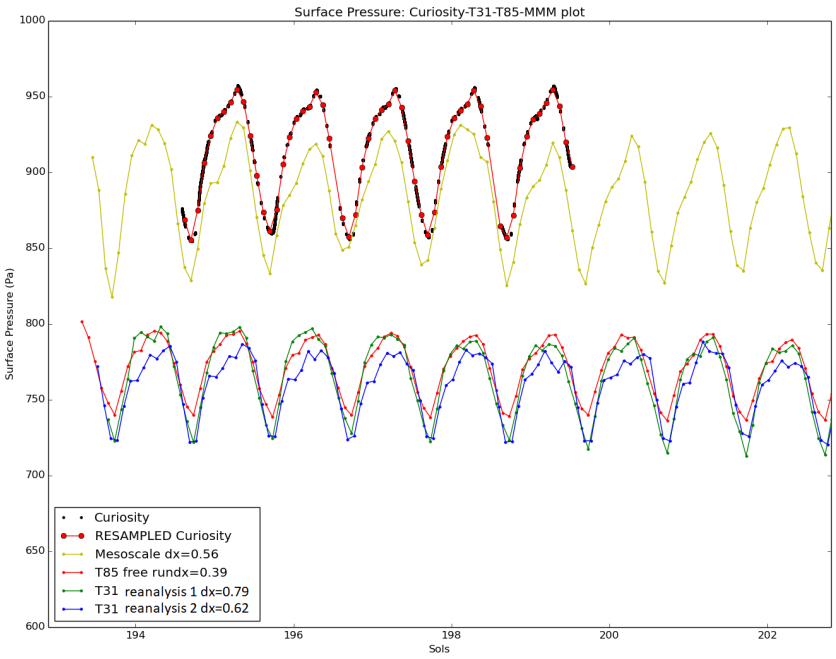


Figure 5.4: The surface pressure data after preprocessing. All model values in this figure are extracted from the closest pixel to Curiosity's location. The black dots are the actual unaltered REMS observations. The large red scatter plot on top of the black dots is the result of rebinning the REMS pressures. Three MGCM simulations are added in red, green and blue scatter plots and the MMM reanalysis is plotted in yellow. The T85 free run (red) and T31 reanalysis 2 (blue) are preliminary test runs made by myself and the T31 reanalysis 1 is done by Dr Luca Montabone. All the model results are subjected to the resynchronisation algorithm and their temporal shifts (in the horizontal) are provided in the legend, in units of sols. In the vertical, the functions are left unaltered. Hence as an example, in the case of the MMM shift $dx = 0.56$, the plot is essentially shifted to the right by 0.56 sols. This figure demonstrates the impact of the improved topography of Gale Crater in the MMM code, but also suggests that the local time at Curiosity is automatically met by the models.

its regular - global resolution; 384x192 for the T170 MGCM runs). Nonetheless, the models accurately capture the yearly trend.

The discrepancy at the diurnal scale is hence a mere problem of incorrect time-stamping in the MGCM and was resolved using the resynchronisation procedure, which was previously described. The final step is another rebinning of the Curiosity data to the 2-hourly rate, thus matching all points computationally (i.e. all

model/observational data points are distributed over the same time-stamps). The results of this procedure are shown in figure 5.4. Notice how the MMM output is also underestimating the depth of the crater floor, but is remarkably more accurate compared to the GCM runs. Thus, even at the 5 km resolution that was used in the embedding, the crater floor is still being placed higher in the model than its real altitude. The resynchronisation values are provided in the legend. After the two resamplings of the Curiosity pressure data and the resynchronisations of the models, the data is ready for the next step, its mathematical decomposition. This preprocessing method is applied to all 4 case studies in this chapter. As convoluted as this method is, it does come with one advantage. Adjustment of the MGCM/MMM time-stamps to the time zone at Gale Crater is not required. Thus, all simulation results are automatically synchronised with Curiosity, and figure 5.4 illustrates this fact.

5.1.2 The Singular Spectrum Analysis of Atmospheric fields. Methodology

The temporal SSA main algorithm was adapted from Golyandina & Zhigljavsky (2013) - Chapter 2. Before proceeding, a number of definitions have to be made. The initial data is put into vector form $X = (x_1, ..., x_N)$, with $N \in \mathbb{N}$ being its length and x_i is the ith data entry. Decomposition is made within a time window K, which in our case is $K = 12$, because there are 12 data points per sol in the reanalysis. In other words, the SSA method constructs the orthonormal basis discussed previously, from functions with periodicity less than or equal to the diurnal period. The choice of K does not exceed the time sequence of one sol because we only modelled 5 to 10 sols per start of season, thus, longer period signals will be too weak to analyse. This does not imply that higher periods than $K = 12$ are ignored from the data. These will be collected in the zeroth order SSA component, which actually defines the general trend of the data X. The last

definition is the *window length* denoted by $L\,(1 < L < N)$, hence $K \equiv N-L+1$. In formal mathematical terms, there are two complementary phases to the algorithm:

- **The Decomposition stage**. Two steps make up this stage:
  - *The Embedding step.* A sequence of lagged vectors of size L are constructed from the data vector X by taking the first L values and sliding them backwards K times, filling the end points with the $N-L+1$ remaining data points from the end of X. This results in $K \equiv N-L+1 = 12$ lagged vectors of length L, defined as

$$X_i = (x_i, ..., x_{i+L-1})^T, \quad (1 \leq i \leq K). \tag{5.1}$$

    These are then put together to form a *trajectory matrix*

$$\mathbf{X} = (X_1, ..., X_K), \tag{5.2}$$

    with the lagged vectors on each column. The resulting trajectory matrix size is $L \times K$.
  - *The Singular Value Decomposition of the trajectory matrix.* The singular value decomposition is performed on the trajectory matrix X. This procedure starts by constructing a covariance matrix as $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ for a typical eigenvalue problem. The eigenvalues $\lambda_1, ..., \lambda_L$ and eigenvectors $U_1, ..., U_L$ are computed for $\mathbf{S}$.

    Normally the rank of the covariance matrix is L, however we used a different definition for the covariance matrix, namely $\mathbf{S} = \mathbf{X}^T\mathbf{X}$. The reason behind this was to produce a basis of $K = 12$ eigenvectors and not $L = N - 11$. The eigenvectors defined in this way will form an orthonormal basis similar to the FFT decomposition. We are allowed to do this, because K and L are interchangeable by definition ($K =$

$N - L + 1$ or $L = N - K + 1$). Other definitions were altered to accommodate our form of covariance matrix, but will not presented here (as swapping transposes and redefining sizes for matrices/vectors is trivial). Proceeding with the original method from Golyandina & Zhigljavsky (2013), the so called *factor vectors* are constructed from the eigenvectors as $V_i \equiv \mathbf{X}^T U_i / \sqrt{\lambda_i}$. Other constructions use the two sets of vectors U_i and V_i. These are the *elementary matrices* defined as $\mathbf{X}_i \equiv \sqrt{\lambda_i} U_i V_i^T$ and the *principal component vectors* defined as $Z_i \equiv \sqrt{\lambda_i} V_i = \mathbf{X}^T U_i$. The matrices $\mathbf{X}_i$ are subsequently used in the reconstruction phase, while Z_i are associated with the principal component analysis.

- **The Reconstruction stage**. Chosen groups of the elementary matrices $\mathbf{X}_i$ are summed together into a resulting matrix $\mathbf{X}_I = \mathbf{X}_{i_1} + ... + \mathbf{X}_{i_p}$, where $I = \{i_1, ..., i_p\}$ is a disjoint subset of the eigenvectors. This step is part of a mathematical procedure called *eigentriple grouping*. Summing all elementary matrices will recover the trajectory matrix X.

  The final step is transforming each two-dimensional computed matrix $\mathbf{X}_I$ into a new one-dimensional time sequence of length N, through a procedure called *diagonal averaging*.

In the second part of the chapter, the SSA method will be applied to the reanalysis and REMS observations. More specifically, the following quantities from the method described above will be extracted:

- *The singular values*. These are computed from the eigenvectors of the covariance matrix, as $a_i \equiv \sqrt{\lambda_i}$. Singular values encode the fraction of the variance associated with their corresponding eigenvectors. As the crux of the SSA is optimisation of the orthonormal basis, each eigenvector is constructed in such a way to minimize remaining variance. Thus, the resulting time series

from reconstruction of the first n eigenvectors will have the largest variance compared to any other orthonormal basis (e.g. Fourier basis).

- *The eigenvector basis.* This is the optimised orthonormal basis of (discrete) functions.

- *The reconstructions.* This stage is operationally similar to the inverse FFT for the Fourier components. It reconstructs time series from any choice of eigenvectors, essentially filtering out the rest of the eigenvectors from the initial sequence (hence its use in signal filtering).

- *Principal components.* The intuition behind the principal components comes from their construction. Define a generic element from the principal component Z_i as $\{z_i\}_j$. Then $\{z_i\}_j$ is formed from the scalar product between its corresponding unit vector U_i and a vector of $K = 12$ consecutive data points from X, that follow from the jth element (i.e. the first element from the $K = 12$ consecutive points is the jth element in X). This sweeps over the window length L within X from, as $1 \leq j \leq L$. In other words, Z_i is formed from the projections of all diurnal windows onto the basis U_i. Thus, if the correlations are in the diurnal window of length K, the principal components are the statistically-uncorrelated time-series with each element $\{z_i\}_j$ representing the sum of all U_i-type correlations around it.

The Fourier transform theory is essentially a lower dimensional case of the spherical harmonics decomposition, which was described in chapter 2. A presentation of it is thus omitted.

5.2 The meteorology of Gale Crater. A systematic approach

As mentioned at the beginning of the chapter, time-variations of atmospheric fields at Gale Crater are in resonance with the diurnal period, unless any major dust storm event occurs. This is to be expected considering its near equatorial location on Mars. Hence, a formal, but brief explanation along with an introduction of atmospheric tides, is provided first. Thereafter, results from the comparison of REMS temperatures and pressures, with their corresponding reanalysis values will be presented in two parts. First is the comparison between REMS observations and local reanalysis time series through the analysis of the FFT and SSA components, following with a study of the atmospheric tides reconstructed only from the latter (SSA) method. The last part is done in a similar fashion to chapter 4 and SSA decomposition is used because of its superiority to the FFT. There are no constraints over the chosen orthonormal basis (it basically chooses the best basis to fit the data), making it a better candidate for analysing the reconstructions.

5.2.1 A tropical Martian environment. Atmospheric tides

A non-tidally locked planet like Mars subjected to solar heating produces a generic oscillation over the whole atmosphere. This is the main diurnal tide and it migrates around the planet with a period locked to the rotational period of the planet (sun-synchronous); however, it is not the only atmospheric oscillation. Non-migrating (or non-sun-synchronous) and subdiurnal (migrating) tides are also mixed in. They are produced by the coupling of the sun-synchronous thermal tide with spatial patterns in the Martian surface properties, such as topography, thermal inertia and surface albedo, but are also influenced by the dust and aerosol loading in the atmosphere. These are especially important for analysing REMS observations because Curiosity is located under a region of weak horizontal tem-

perature gradient conditions of the Martian atmosphere, namely in the tropical region. In contrast, the Viking landers are further north in the high-latitude regions. This made observations from the landers particularly good for studying baroclinic waves in the upper atmosphere (Collins et al, 1996). However, where Curiosity is located (i.e. Gale Crater), diurnal variations are dominant over longer timescale synoptic variability and are strongly influenced by atmospheric tides (Wilson & Hamilton, 1996; Guzewich et al, 2015). Therefore, the comparison between REMS observations and the MGCM+MMM reanalysis starts with atmospheric tides.

Atmospheric tides are at the core of this chapter because only temperature and pressure observations from REMS will be analysed, which will be dominated by the tidal signals. The loss of one of the two velocity sensors has put a delay on the release of the velocity data (from the remaining velocity sensor). The velocity data was indeed made available by the time of this thesis, however, we decided to not use it due to the high uncertainty in the data. Not only does the rover's body, component functionality and movements influence the actual wind direction, but the redesign from two velocity sensors to only one adds more bias to the data, making it undesirable for this systematic analysis.

A final remark about atmospheric tides is their separation into migrating and non-migrating. These can be distinguished by placing at least two meteorological stations at different longitudes. Unfortunately, with the addition of the most recent lander on Mars, NASA's InSight on the 26th of November 2018, the migrating from non-migrating properties of tides are still not distinguished. This is because the landing site is at 4.5^oN and 135.0^oE which is further north than Curiosity, but at similar longitude (Curiosity's landing coordinates: 4.5^oS 137.5^oE). Nonetheless, their separation is simulated by the MGCM+MMM and can be analysed in connection to REMS observations.

5.2.2 Analysis of FFT and SSA components

All properties from the FFT and SSA analyses will be presented for the northern autumn case first, and then move on with the seasonal particularities. Additionally, all the one-dimensional functions from the reanalysis are plotted from the closest pixel to Curiosity's location in the case of the surface pressure field, and the closest pixel from a terrain following layer 5m above the local topography in the case of temperatures; these features are valid for all the results of this chapter, and thus considered specified unless a statement to the contrary is made.

The FFT components

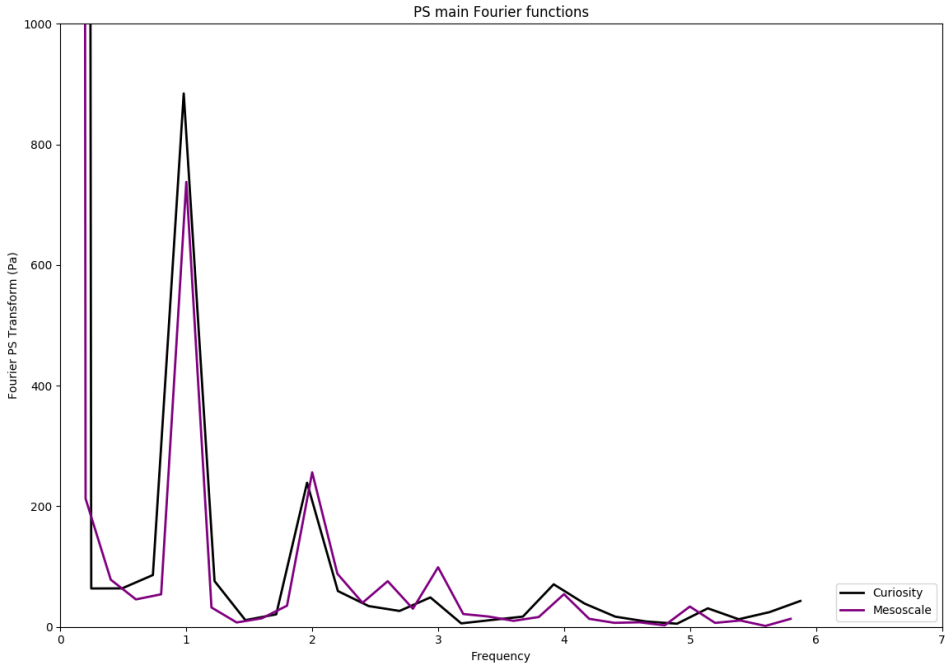


Figure 5.5: The surface pressure Fourier transform from REMS observations (black line) and MMM predictions (purple line). The interval is 5 sols taken from the beginning of northern autumn. The MMM values are taken from the closest pixel to Curiosity. The diurnal and semidiurnal tide signals are prominent but slightly offset due to the small number of sols (5 sols) and the low sampling rate (12 values per sol).

In figure 5.5, the Fourier transforms of surface pressure is computed for REMS

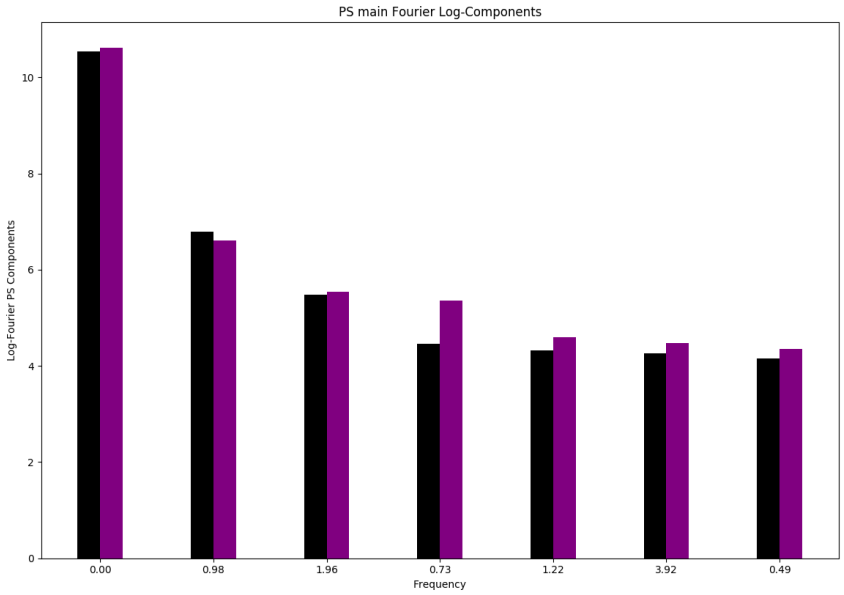


Figure 5.6: The main frequencies of the Fourier transform presented in figure 5.5 in descending order. The scale is logarithmic.

observations and the reanalysis. The leftmost peak is the non-harmonic signal, which corresponds to their mean values. The following prominent peaks are in the vicinity of diurnal frequencies 1 and 2 (i.e. 1 corresponds to a period of 1 sol, and 2 corresponds to a period of 0.5 sol etc). These are associated to with diurnal and semi-diurnal tides. The discrepancies start with the higher frequencies. The terdiurnal signal for the REMS surface pressures appears to be weaker than in the reanalysis. Even the quadiurnal signal is stronger in the reanalysis than in the REMS observations. The graph also shows a frequency 5 signal in both REMS measurements and the reanalysis. Both the terdiurnal and quadiurnal tides are much weaker compared to the rest of the spectrum, as Haberle et al (2014) noted. Figure 5.6 shows a descending selection of the strongest signals in logarithmic scales and figure 5.7 is a plot of the first 5 of these. The frequencies on the x-axis of figure 5.6 are REMS-only. The plots from figure 5.7 are produced by applying an inverse FFT; these are the FFT *reconstructions* (this terminology is borrowed from the SSA reconstructions). The first 3 plots contain the mean and the diurnal

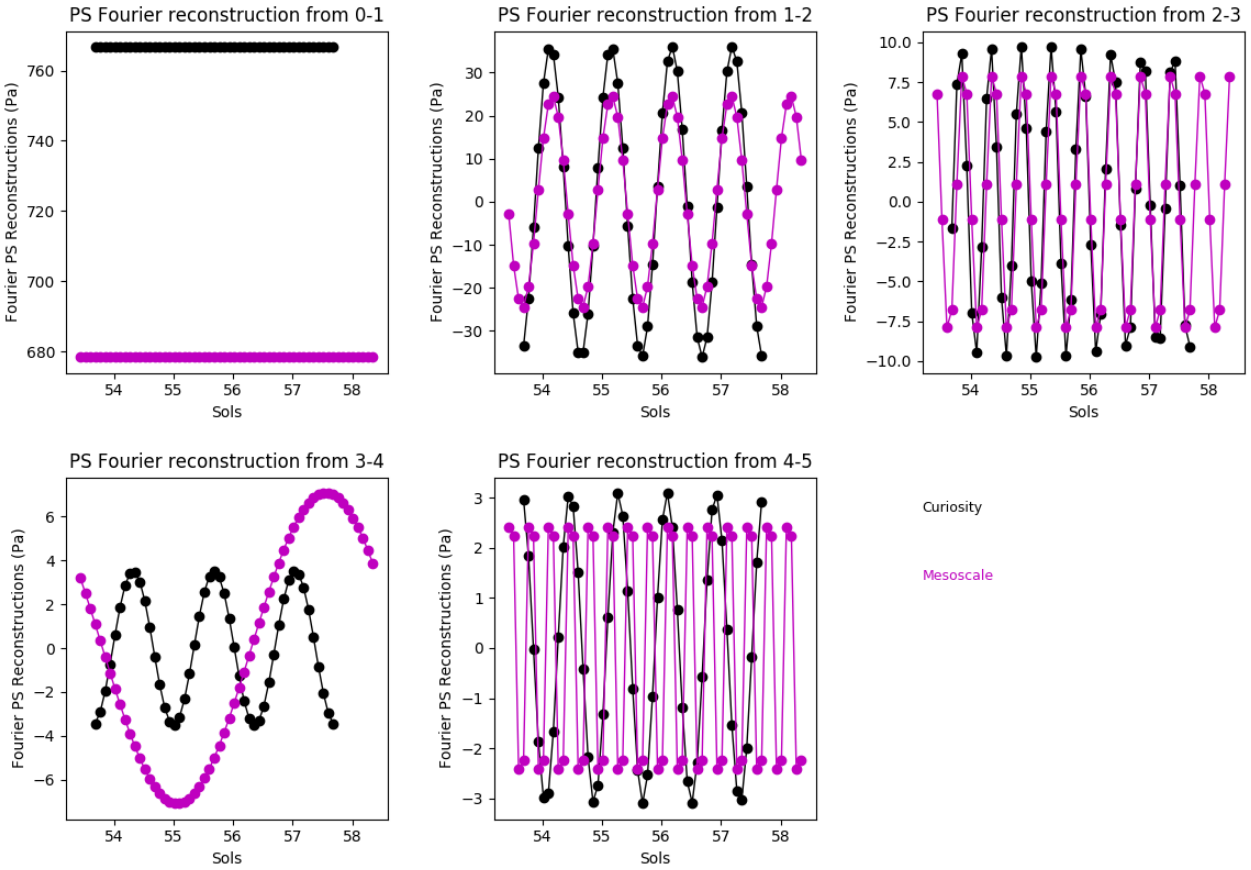


Figure 5.7: The inverse Fourier transform of the main surface pressure components from figure 5.6 in descending order from right to left.

and semidiurnal tides. The lower magnitude for the reanalysis in the first plot is due to the finite resolution of the topography; it represents the crater floor higher than it is in reality at Curiosity's location. The diurnal tide amplitude is $\approx 28\%$ weaker for the reanalysis, compared to REMS measurements. The semidiurnal tide is slightly better simulated (in the reanalysis), but does not vary in amplitude as seen in reality (REMS measurements). On an annual scale, Lewis & Barker (2005) and Guzewich et al (2015) showed that there is a strong correlation between the dust loading in the Martian atmosphere and the amplitudes of the diurnal and semidiurnal tides. Taking this into account for the semidiurnal tide, suggests that the variability in atmospheric dust loading is poorly represented in the reanalysis,

but the mean dust loading is relatively accurate. Although, from both diurnal and semidiurnal tides, it seems that the reanalysis does indeed underestimate atmospheric dust loading. The 4th plot shows the 0.73 REMS-frequency from the Fourier frequencies plot 5.6 and remarkably a 6.5 sol period oscillation in the reanalysis. Their magnitudes can indeed be seen in the Fourier transform from figure 5.5 at frequencies 0.73 - REMS and 0.15 - reanalysis. In a similar fashion, the last plot of figure 5.7 shows oscillations corresponding to the frequency 1.22 for REMS and the terdiurnal tide for the reanalysis, with amplitudes smaller than 4Pa. The contributions from plots 4 and 5 prove that indeed only the diurnal and semidiurnal tides are the most prominent in the Martian atmosphere, while the rest of the spectrum (ignoring the zero frequency component) is weak and varying in strength from non-tidal meteorology.

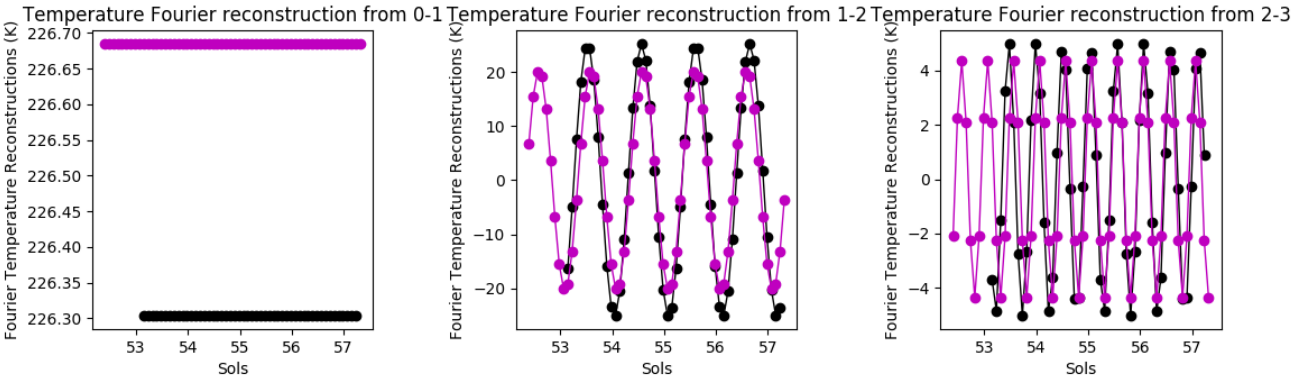


Figure 5.8: The inverse Fourier transform of the main temperature components in descending order from right to left.

Figure 5.8 contains the Fourier transform of the temperature signals, similar to 5.7. And in a similar fashion to the surface pressure, the diurnal and semidiurnal signatures dominate the spectrum. The agreement between model and observations is much better compared to surface pressures. The mean temperatures are within 0.5K agreement, and the two main thermal tides still seem to indicate

reanalysis underestimation, however, within 16% for the diurnal 5% and for the semidiurnal tides. This also suggest that the amplitude of the pressure tide is amplified by the lower altitude of Curiosity in comparison to the underestimating reanalysis topography. Furthermore, the lower altitude of the crater floor does amplify the diurnal pressure tide. This is accordance with Tyler & Barnes (2013) findings, that the crater topography simply amplifies the diurnal tide, but a more technical part must be taken into account. Data analysis in the MGCM is performed from MCS temperature observations, which forces the temperature field to be more realistic.

The SSA components

Atmospheric tides computed from the SSA (i.e. the eigenvector basis) are generally less ambiguous compared to their FFT counterparts. For instance, the surface pressure eigenvectors from figure 5.9 have almost matching diurnal and semidiurnal orthogonal functions, with increasing discrepancies for shorter tidal periods (the extent of the x-axis is 1 sol). It must be noted that the first component is the general trend (i.e. the first eigenvalue, denoted as eigenvalue 0 everywhere in our study to emphasize its characteristic as a *trend* component), and the remaining nine are coupled even and odd oscillatory functions (but not sinusoidal). In pairs, they represent the main atmospheric oscillations. Thence, in descending order from left to right, the following eigenvalue groups are coupled opposite-parity functions of the same frequency: (1,2), (3,4), (5,6), (7,8). The reanalysis and REMS observations may appear to be in antiphase ($\approx 180^o$ phase difference), even if their frequencies and parity essentially match (as is the case for the diurnal and semidiurnal tides). This is a feature of the SSA method, as it does not assume any a priori information about starting value of the second derivatives of the orthonormal functions. And because both eigenvectors of a parity pair can be constructed as if they are both multiplied by a -1, they cancel out

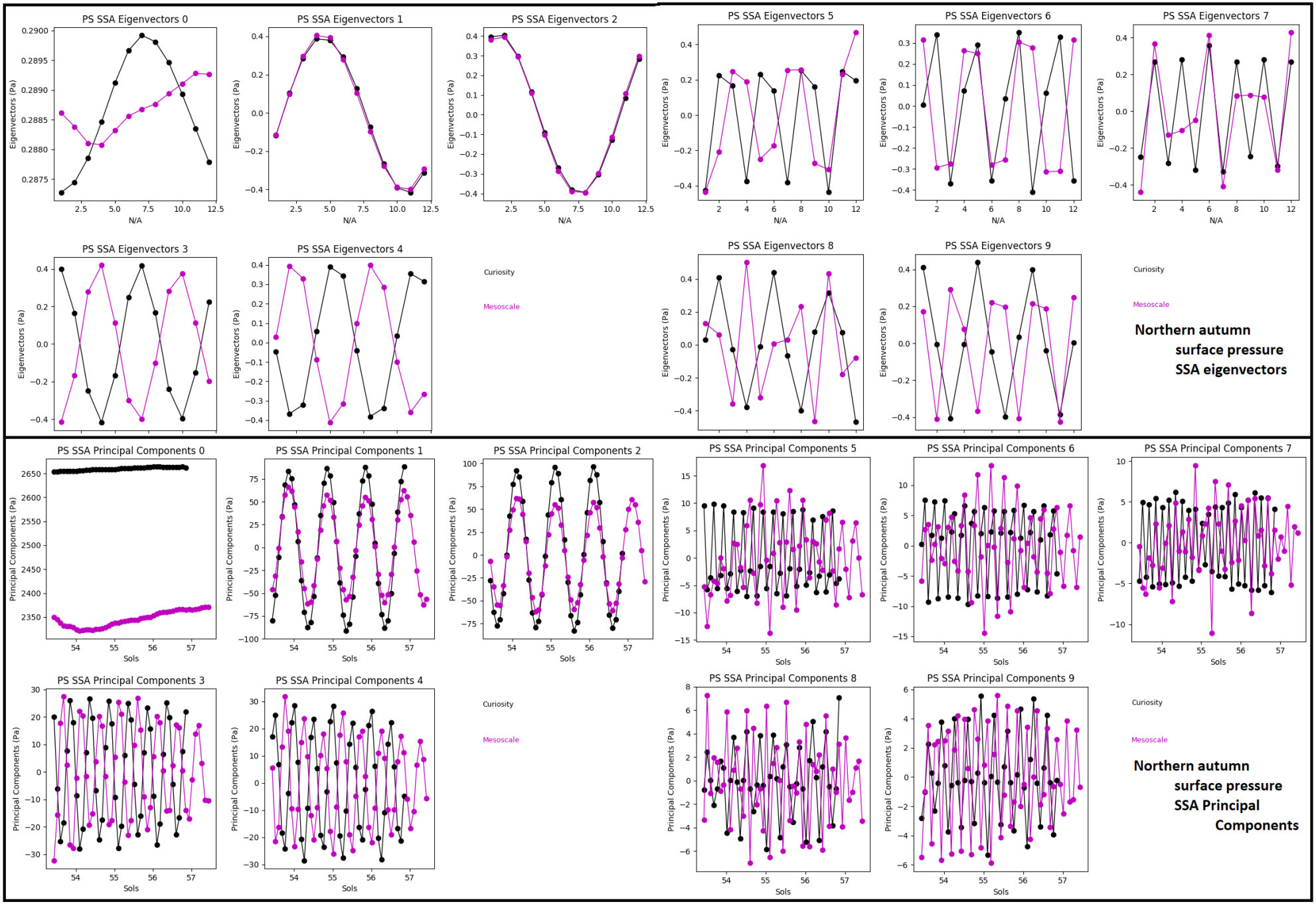


Figure 5.9: Tiles of northern autumn surface pressure SSA eigenvalues and principal components.

when involved in the reconstructing process of their particular atmospheric tide, i.e. their tidal properties are preserved correctly. Additionally, to investigate the phase difference between the surface pressure and thermal tides, the resynchronisation algorithm was applied only to the reanalysis surface pressures, and the same temporal displacements were used for temperatures (i.e. temperature SSA vectors from figure 5.10 are set to the local time from the pressure tide). Focusing on individual components, eigenvector 0 appears to pick up an oscillatory trend for the case of REMS observations, with a period surpassing 2 sols, while the SSA always starts with a descending part for surface pressures, and a corresponding ascending part over the same interval, for temperatures. This essentially captures the signal of the short spin-up period of the MMM. It can even be noted that this build-up period is slightly longer than one sol (what was suggested in the LMD MMM user manual), by about a quarter of a sol. The trend (zeroth) principal component tells a different story. Encompassing all superdiurnal influences, the pressure seems to monotonously increase over the 5 sol period to reach its maximum in the northern winter season; a very stable and predictable weather pattern up to this point. The temperature trend is where non-tidal meteorology is more active and it would be interesting to compare with the results from previous chapter (in the next subsection). Moving forward to the first parity pair (1,2) or in other words the diurnal tide, the matching of the periods and amplitudes of the REMS observations and reanalysis eigenvectors is uncanny. This agreement in diurnal correlations is not detected in the FFT diurnal component. It seems that the discrepancy in the FFT diurnal tide comes from the (uncorrelative) principal components in the SSA analysis. The diurnal signal (i.e. the (1,2) eigenvector pair) is carried out over the sample window at larger amplitudes for REMS (i.e. the (1,2) principal component pair). This suggests that the reanalysis indeed captures diurnal correlations correctly, via matching eigenvectors, but carries them over a smaller amplitude in the 5 sol period, via unmatching principal compo-

nents. So the amplification of the diurnal tide from the topography that Tyler & Barnes (2013) noted, simply boosts the diurnal principal components for REMS measurements, but the correlations are the same (i.e. eigenvectors match with the reanalysis). We know this, because the principal components are uncorrelated functions, as stated in the theory section. Tyler & Barnes's (2013) discovery is indeed validated. A final remark about the (1,2) eigenvector pair, is that there is indeed an offset of the phase between the reanalysis surface pressure tide and its temperature tide, which is not found in REMS observations. This offset is of the order of 2 Martian hours, which corresponds to the chosen MGCM/MMM output period. This was noticed in chapter 4 as well when tracking the pressure nodes around Mars, in the analysis of the day-night circulation in the crater (subsection 4.2.3). The offset seems to be intrinsic to the MGCM and propagating to the MMM.

The next eigenvector parity-pair (3,4), which is the semidiurnal signal, is again in a remarkable good agreement between the reanalysis and REMS, both in period and amplitude. A small difference comes in the form of the function itself. The reanalysis surface pressure eigenvectors seem to fall-off slightly faster compared to REMS observations, after they both reach their peaks. Hence, an overestimation in the reanalysis semidiurnal oscillations of the surface pressure drop-off, after a local maximum (or minimum), might be due to an overestimation of the semidiurnal variation of atmospheric dust and aerosol loading in the model, over the same interval (after a local maximum or minimum). In the temperature eigenvectors, this discrepancy is manifested as a smaller gradient succeeding a local maximum (or minimum) in the semidiurnal tide. These differences are small though. In contrast, the reanalysis pressure-to-temperature phase difference (of 2 hours for the diurnal tide), is even smaller in the semidiurnal tide phase difference, although still present. This is a compelling result for the MGCM+MMM configuration. Principal components on the other hand, manifest similar results with the FFT

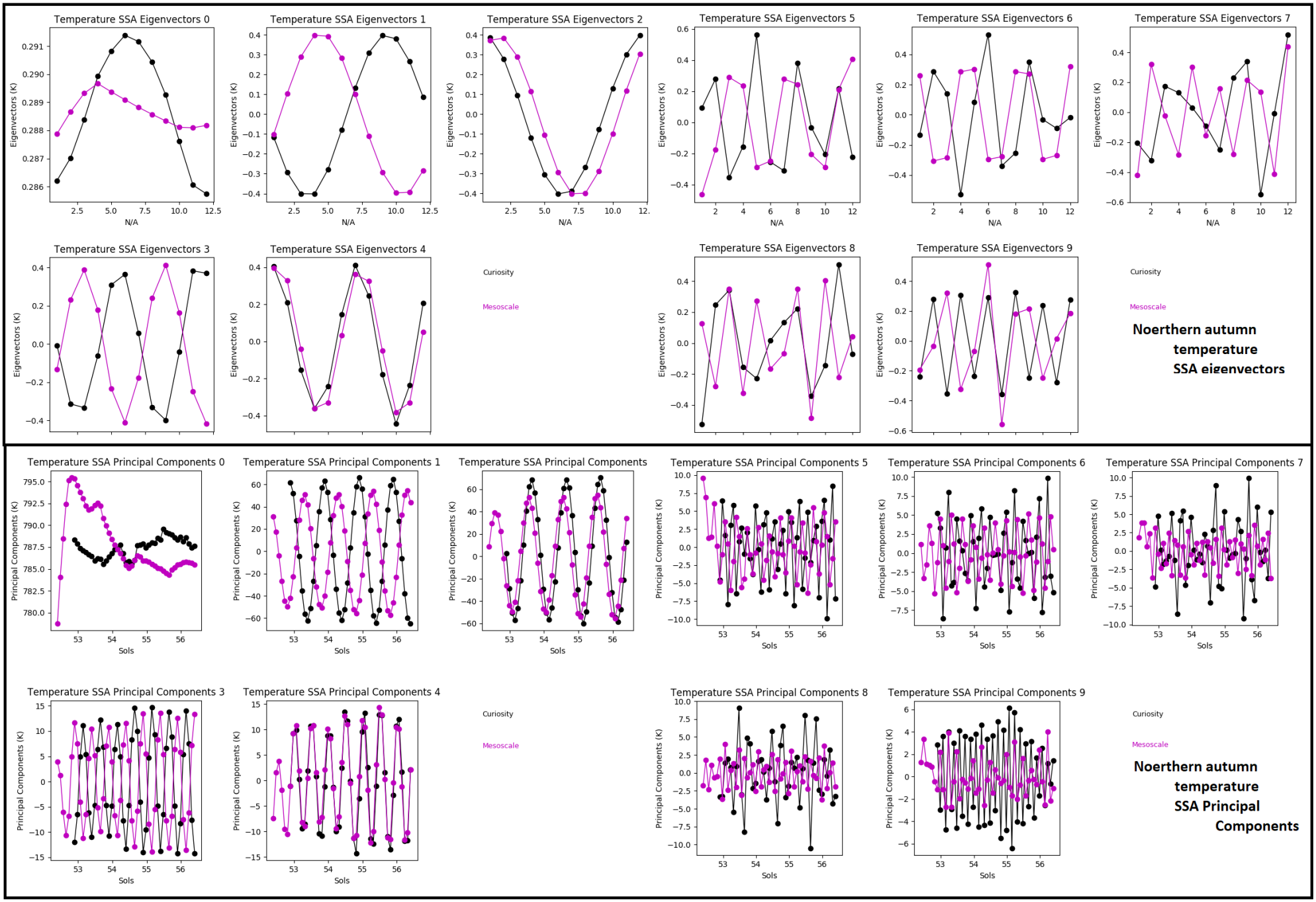


Figure 5.10: Tiles of northern autumn temperatures SSA eigenvalues and principal components.

semidiurnal component. Surface pressure amplitudes in the case of the reanalysis, are slightly underestimated compared to REMS, however, the agreement is actually improved compared to the FFT. The decrease in amplitude is actually present in the reanalysis SSA-decomposition, a feature not appearing in the FFT semidiurnal component (from figure 5.7). The surface pressure amplitude difference varies from 0 to 4 Pa and both the reanalysis and REMS overall amplitudes in the semidiurnal principal components have an oscillatory behaviour, with a maximum amplitude at the beginning of the day, to a smaller one in the next semidiurnal maximum of the same day. This can be related again to the semidiurnal dust loading, which falls off when reaching night-time. It is also important to note that the Gale Crater topography only amplifies the diurnal tide, while the semidiurnal tide amplification is more likely due to dust loading; although this fact needs more investigation. So the local topography and atmospheric dust and aerosol loading do not change tidal properties, but merely amplify them.

Next parity-pairs of eigenvectors are significantly noisier. The terdiurnal tide (parity pair (5,6)) is indeed the next strongest correlation for both the reanalysis and REMS observations, unlike the case of the FFT analysis, where other oscillations were generally stronger. This is the case because the SSA method is significantly more organised in determining (and filtering) correlations, especially at short, 5 sol long datasets. Features like the spin-up or superdirunal meteorology are not influencing the main subdiurnal signals, which is not the case for the FFT analysis. The spin-up of the MMM for example, is neatly added as an eigenvector 0 signal in the SSA, leaving the diurnal tide pure (filtered), and strongly matching between the two datasets (REMS and the reanalysis), something impossible for the FFT. Nevertheless, the terdiurnal tide is present, but noisy and weak compared to the diurnal and semidiurnal tides; a known characteristic of the Martian atmosphere. The next eigenvector pairs are even noisier, but nonetheless have matching, but varying periods and the same order of magnitude, between

the reanalysis and REMS observations.

Seasonal variability

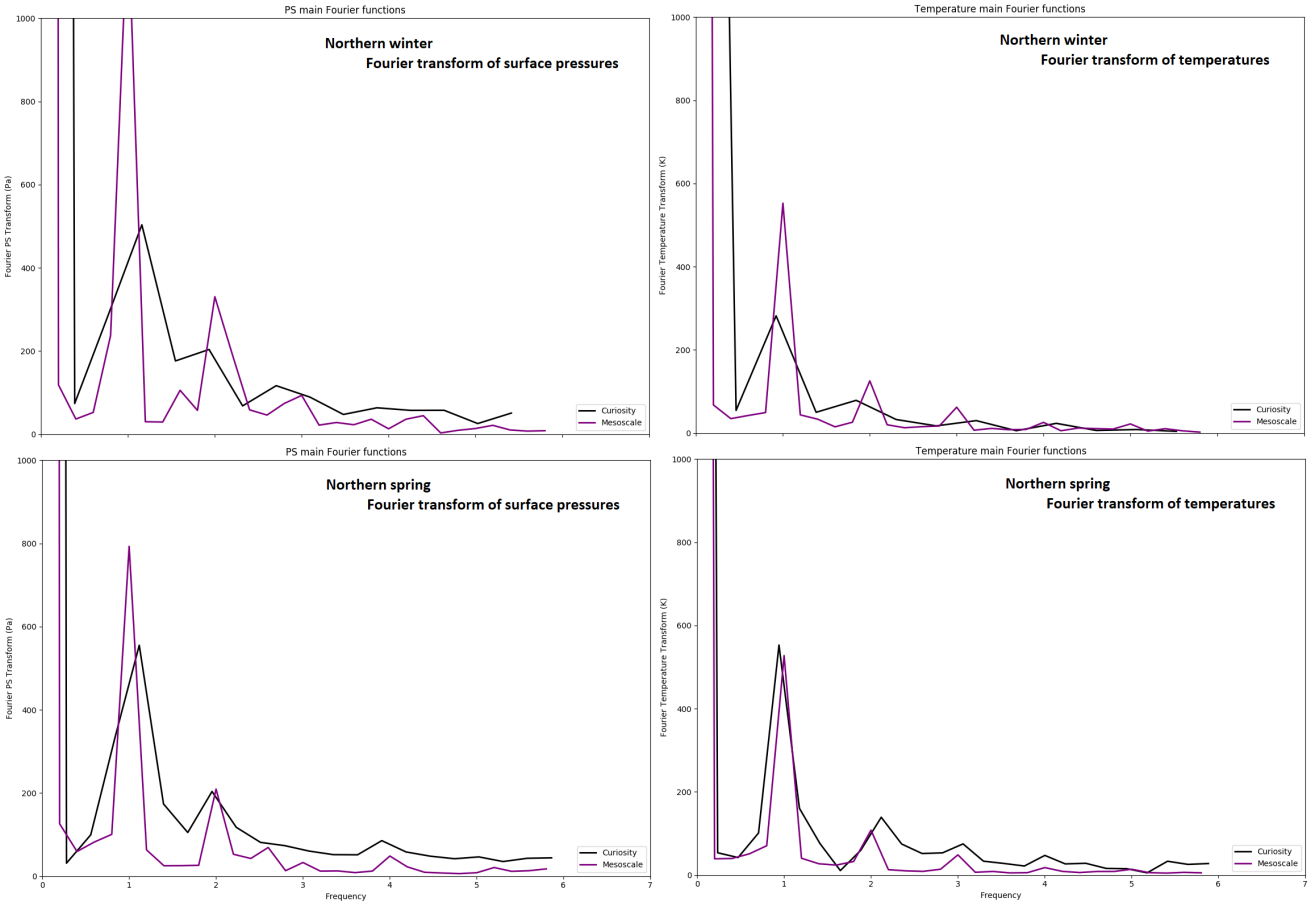


Figure 5.11: The Fourier transform for northern winter (top two plots) and spring (bottom two plots). The left plots are surface pressures and right plots are temperatures. Each individual plot is constructed in the same way as the northern autumn Fourier transform from figure 5.5.

Next is the analysis of northern spring and winter. Summer was omitted from the FFT+SSA decomposition due to the short overlap period between the reanalysis and REMS temperatures, which corresponds to less than 1 sol and covers the spin-up period of the MMM, thus making the summer decomposition unreliable. We are planning to make optimisations to the MGCM+MMM configuration, by allowing multiple embeddings and a smaller, but more efficient MGCM resolution.

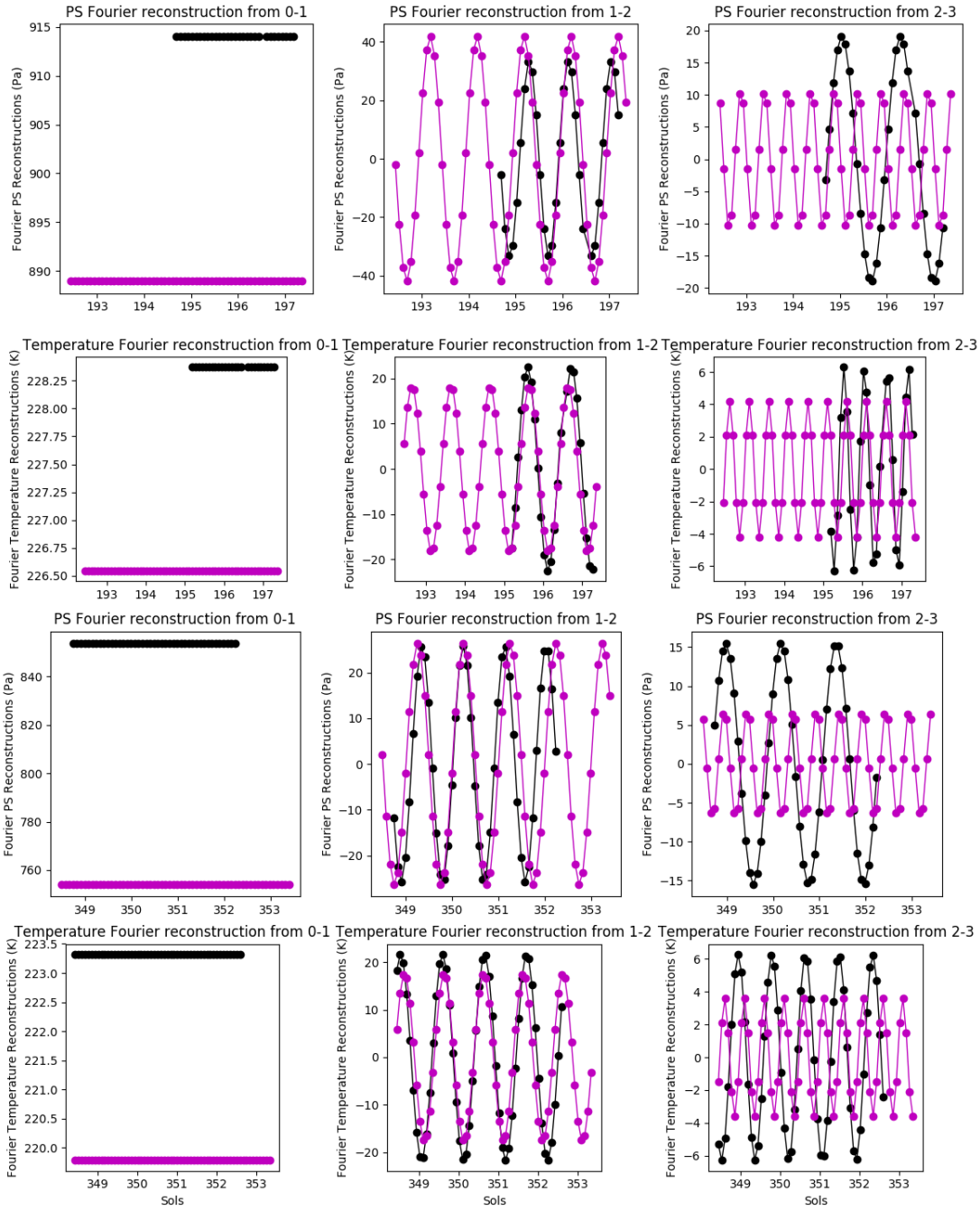


Figure 5.12: The inverse Fourier transform of the main components from the 4 Fourier transform plots presented in figure 5.6. From top to bottom, each row represents the reconstructions of northern winter surface pressures, northern winter temperatures, northern spring surface pressures and northern spring temperatures. Each column from left to right, contains the reconstructed components of the mean and diurnal and semidiurnal tides.

This will greatly extend the (currently 5 sol) reanalysis intervals and eliminate overlapping problems, which are recurrent in our short seasonal datasets (e.g. northern winter REMS temperatures from figure 5.12). More details will be discussed in the Conclusions chapter.

Figure 5.11 contains the Fourier transforms of northern winter surface pressures (top-left) and temperatures (top-right), and northern spring surface pressures (bottom-left) and temperatures (bottom-right). Each individual plot is designed similarly to the Fourier transform from 5.5. The winter surface pressure and temperature frequency spectra show an overestimation of the diurnal and semidiurnal signatures for the reanalysis in comparison to REMS, with the remaining parts of the spectrum favouring the other direction (the reanalysis underestimates REMS temperatures and surface pressure). Northern spring on the other hand, presents a less prominent reanalysis diurnal and semidiurnal tide. Plotting the main 3 signatures from the Fourier transforms in the same way as figures 5.7 and 5.8, one obtains the tiles from figure 5.12. In this figure, the mean (zeroth component) reanalysis values are organized in the left column. These reasonably follow the autumn case, with the reanalysis underestimating REMS surface pressures by 25 Pa for northern winter and up to 100 Pa in the case of spring. The MMM results are quite closer to the MGCM values and further from REMS observations for the spring surface pressure mean. Temperatures otherwise, show a better agreement, with a difference between the reanalysis and REMS of less than 2K for northern winter and less than 4K for spring. The diurnal tide periods seem to not perfectly match the period of one sol. This is due to the small number of diurnal oscillations present in the signal, making the FTT method problematic at decomposing the correct periods. Nonetheless, in comparison to the autumn case, the surface pressure diurnal tide from the reanalysis overestimate the corresponding REMS observations by almost 10 Pa in the case of northern winter and correctly approximates REMS observations in the northern spring case. As

was the case with northern autumn, the temperature diurnal tide shows a better agreement compared to the pressure tide between the reanalysis and REMS for both winter and spring, with the same $\approx$ 5K underestimation as before. Reanalysis and REMS temperatures are thus generally in a better agreement regardless of season. The left column, which should have contained the semidiurnal signature, shows other components instead, making the FFT quite unreliable for short period analysis.

Fortunately, with the SSA method, this is not the case. Figure 5.14 contains SSA components constructed analogous to 5.10, for northern spring temperatures. The same time shifts extracted from the resynchronisation algorithm applied to northern autumn surface pressures is kept for spring as well, to detect any long term offset in the reanalysis. Indeed a 2 hour phase difference between the reanalysis and REMS values, appears in both the eigenvectors and principal components. This can be recognised from the northern spring temperatures, where the diurnal tide shows matching REMS and reanalysis functions. But as there was a constant phase difference of 2 Martian hours between the temperature and surface pressure diurnal tides, and the spring temperatures match, this means that the spring surface pressures are offset by two hours between the reanalysis and REMS, and this can be seen in figure 5.13. So indeed there is a small phase difference propagating in the MGCM from time-keeping, as mentioned at the beginning of the chapter. The semidiurnal tide generally preserves the characteristics from northern autumn throughout spring. Agreement in period and amplitude between the reanalysis and REMS surface pressures and temperatures is evident in the northern spring, same as it was for the autumn season. Additionally, the principal components preserve the phase characteristics of the eigenvalues, but also the uncorrelative amplifications that were present in northern autumn as well (in principal components). The surface pressure diurnal tide contains the amplifying signature of the topography (Tyler & Barnes, 2013) and the semidiurnal tide contains the dust and

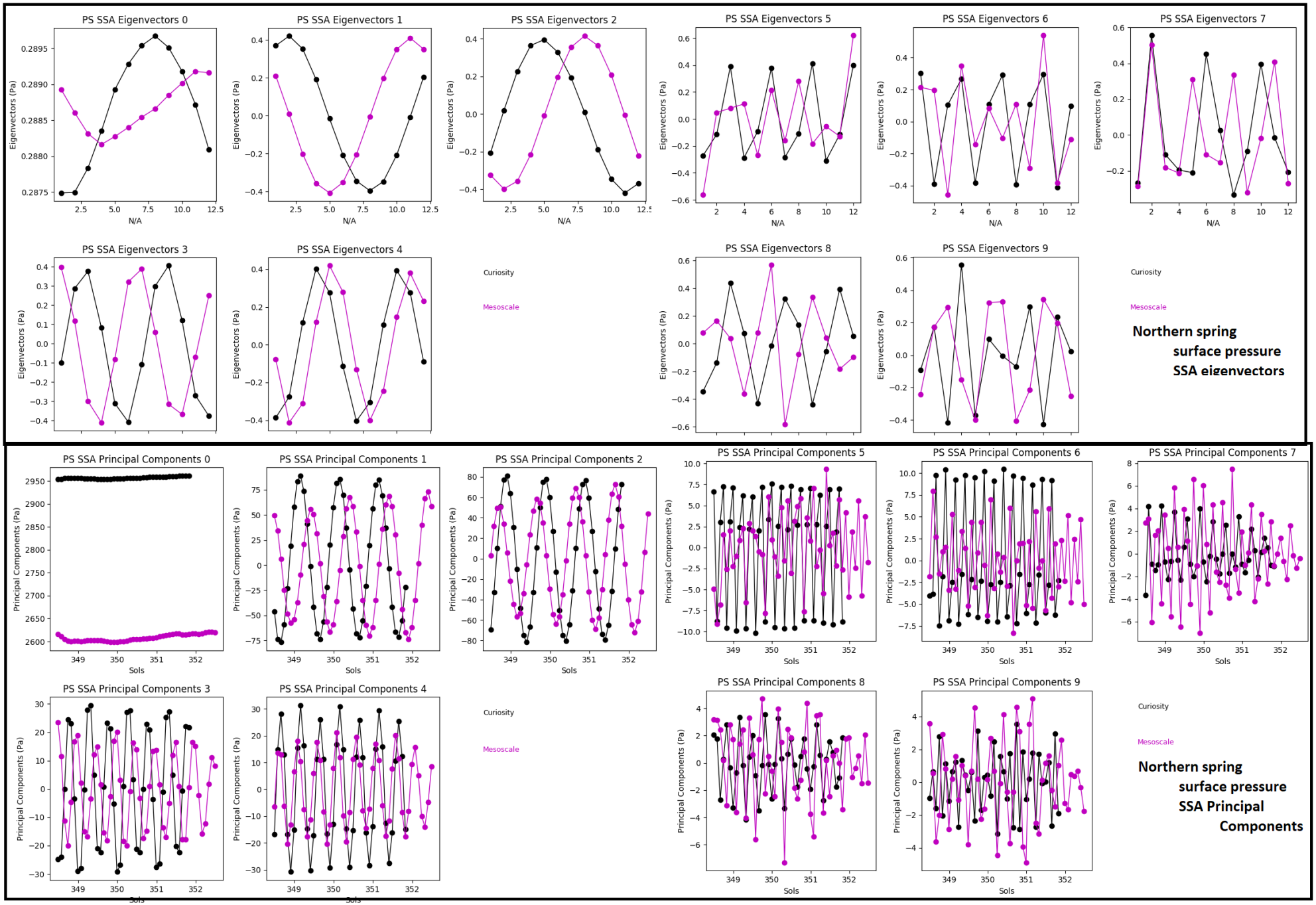


Figure 5.13: Tiles of northern spring surface pressure SSA eigenvalues and principal components.

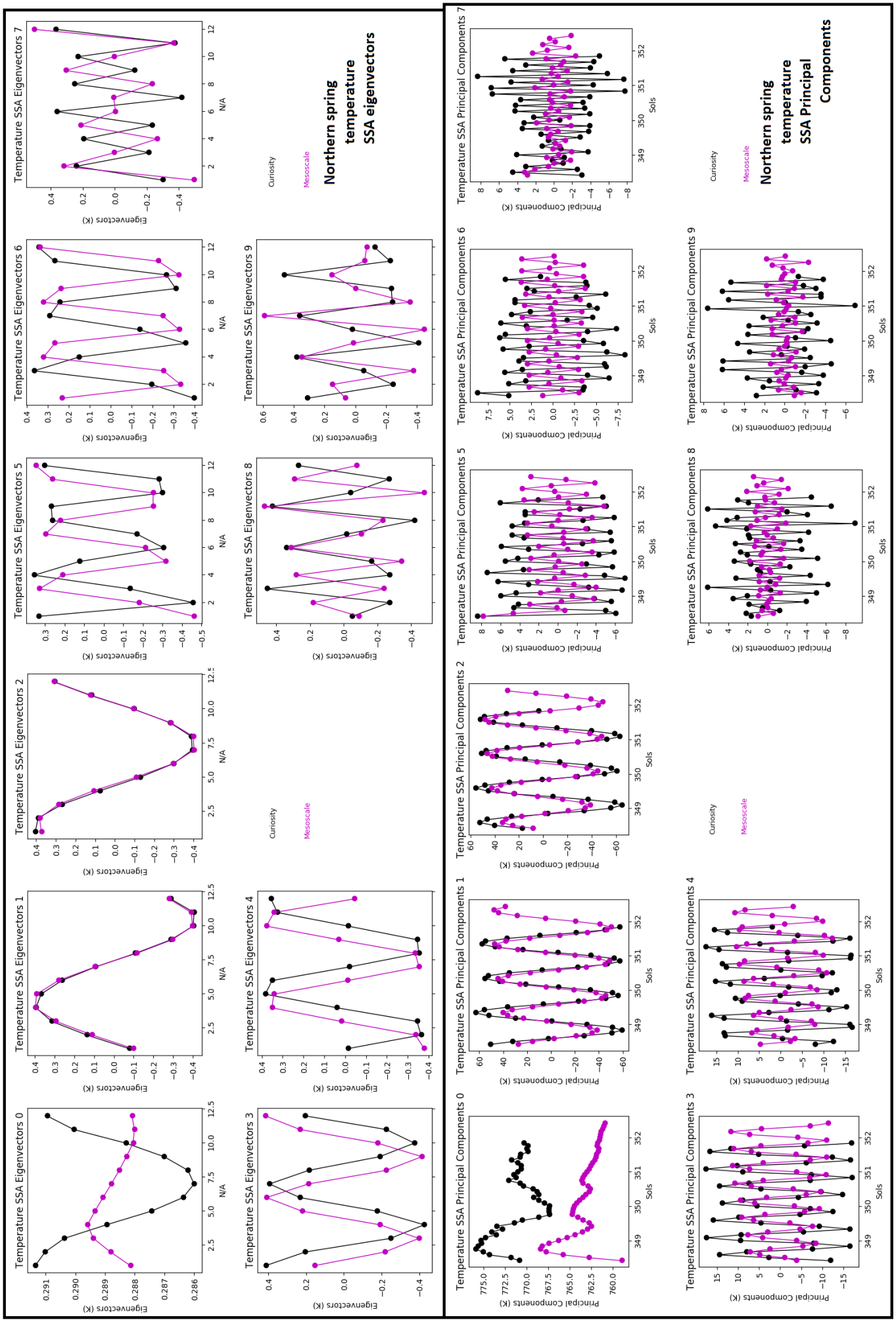


Figure 5.14: Tiles of northern spring temperature SSA eigenvalues and principal components.

aerosol signature (Lewis & Barker, 2005 and Guzewich et al, 2015), as it was the case in the autumn 5-sol dataset. Both these signatures are present in the principal components, and not in the typically matching eigenvectors. Generally, the atmospheric tidal characteristics in SSA eigenvectors and principal components are preserved from the ones manifesting in northern autumn. Thus a discussion of both northern spring and winter will be similar with the one from autumn, so the winter values are left for the next part. A full seasonal discussion using solely the SSA method is given in the next subsection, because it couples the results from chapter 4 as well, making the seasonal analysis more decisive. And as noticed, the FFT components are generally unreliable, so only the SSA reconstructions will be used. Additionally, the trend reconstructions will be elucidated by comparing with the meteorology from chapter 4.

5.2.3 SSA Reconstructions of the general trend coupled with the diurnal, semidiurnal and terdiurnal tides

Previously, the eigenvectors showed remarkable agreement between the reanalysis and REMS datasets for the diurnal and semidiurnal tides, implying that the disagreements in amplitude come from uncorrelative principal components. Here the SSA eigenvectors and principal components from the previous subsection were passed through the last phase of the SSA, namely the reconstruction of the trend and diurnal, semidiurnal and terdiurnal parity pairs. Additionally, the previous subsection also explained the phase difference in the reanalysis between the temperature and surface pressure diurnal tides, and the offset between seasons coming from time-keeping. Both are caused by the MGCM and propagated to the MMM. In this section, the focus is on the meteorology, so the resynchronisation method is applied to all results hereafter (as a phase difference will only make the analysis cumbersome). This is also the reason why mean REMS values are shifted down to match the mean reanalysis values, so that the values in the colour bars from

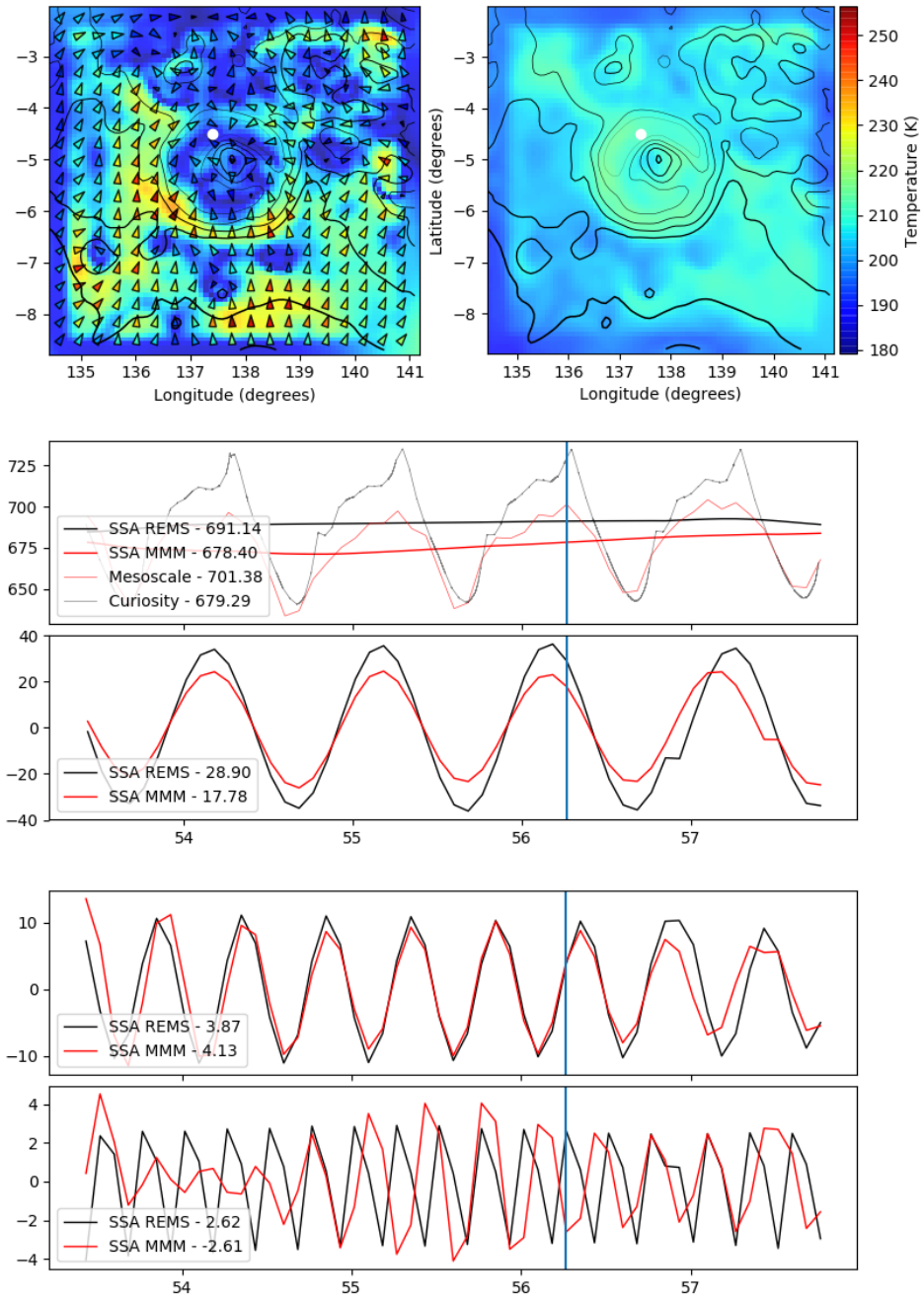


Figure 5.15: The SSA decomposition of surface pressures from the MMM and Curiosity-REMS from the start of northern autumn. The first two colour maps are MMM velocities (top left plot) and temperatures (top right) taken from a terrain-following layer, 5m above the local surface. The following 4 plots contain SSA reconstructions of REMS observations (black lines) and MMM values (red lines) taken from the closest pixel to Curiosity's location (-4.5 latitude and 137.4 longitude) during its first operational year (corresponding to Martian year 31). The time of day selected for the velocity and temperature fields at the top is indicated by the vertical blue line in the SSA reconstructions. Non-binned values are denoted by faint lines and only appear in the first SSA plot. Additionally, the values from the reanalysis in this plot are shifted downwards for clarity, and the full unaltered reconstructions can be found in figure 5.16. The SSA plots in descending order are reconstructions from the following groups of eigenvectors: 0 - general trend, (1,2) - diurnal tide, (3,4) - semidiurnal tide, (5,6) - terdiurnal tide.

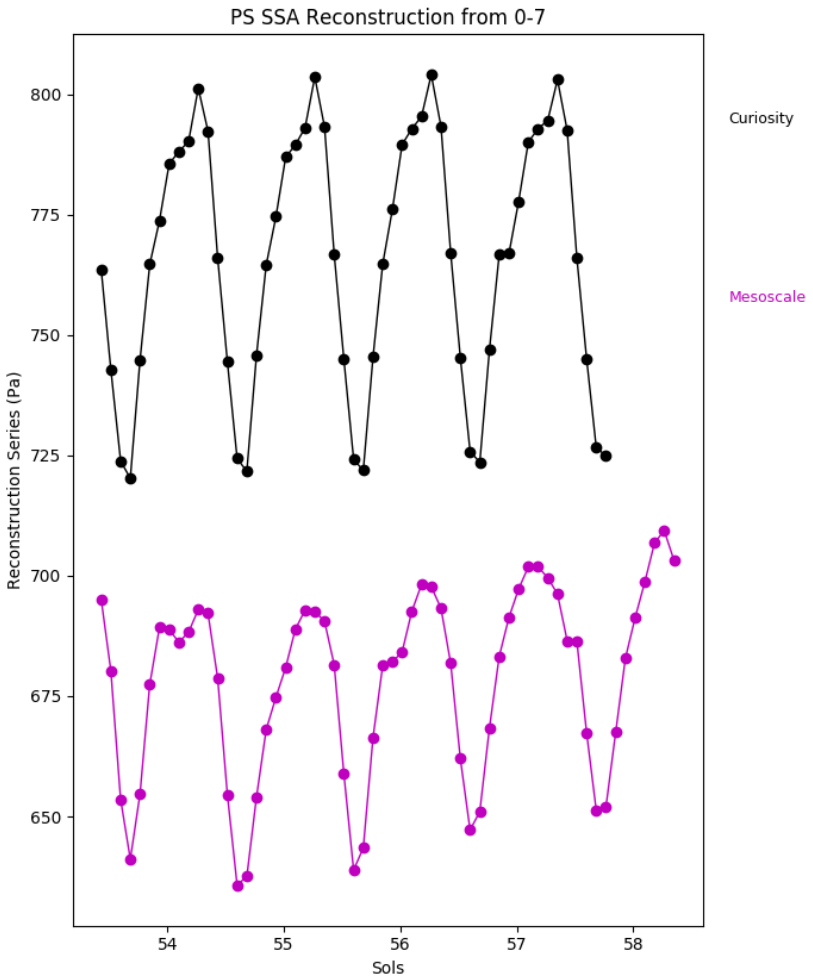


Figure 5.16: The SSA reconstructions from combining all decompositions in figure 5.15. The trend values (from first principal component) in this figure are not shifted as in 5.15 in order to provide the realistic reconstructions.

the reanalysis match the ones from the SSA reconstructions.

Nonetheless, the autumn SSA reconstructions are showed in figure 5.15 for surface pressures. The chosen time captures the peak in the pressure tide, which is prominent especially for REMS, in other words the high pressure node is closest to Curiosity. The peak in the REMS observations seems to be produced by the largest rise in the semidiurnal tide (inflection point), while the terdiurnal tide is at its maximum and the diurnal tide is close to its maximum. On the other hand, the corresponding peak in the reanalysis seems to be smaller due to the general antiphase in the terdiurnal tide between REMS and the reanalysis. Recall

from previous chapter, that the winds were in a competing phase when the high pressure node is close. The katabatic winds from Elysium competed with the katabatic winds from the southern highlands, and the high surface pressure node is right in the middle of them, thus countering both. This made for a relatively calm environment (with weak winds) at Gale Crater, and is manifested mostly in the terdiurnal component (and in the smaller variance components that are not shown). In the pressure tide, it seems that the maximal pressure node does not follow the peak of the diurnal tide. This is due to the semidiurnal tide, which has its through two hours before the peak of the diurnal tide and its inflection point 2 hours after the maximal diurnal tide. Hence, the semidiurnal pressure tide has an inflection right when the diurnal tide is at its maximum. In other words, the left side of the peak of the diurnal tide presents a decrease in the semidiurnal tide to its minimal value, while to the right of the diurnal tide, there is a strong increase in the semidiurnal tide (passing through its inflection point) which manifests that periodic peak in both the REMS and reanalysis data. The terdiurnal tide amplifies the REMS peak and counters the reanalysis peak. Additionally, there is a very weak rise in the surface pressure trends of both the REMS observations and the reanalysis, althought these are relatively uneventful, suggesting again that the atmospheric circulation at Gale Crater is indeed strongly influenced by tides. The reconstructions of the 4 tidal components are given in figure 5.16, without the vertical shift in the trend component (zeroth eigenvector). Previous studies by Lewis & Barker (2005) and Guzewich et al (2015) suggest that these features in the diurnal and semidiurnal tides are correlated to the dust loading. Thus, coupled to previous studies, our work shows that the delicate balance in the phases of the diurnal and semidiurnal components of dust and aerosol loading can explain the shape in both the reanalysis and REMS observations. The discrepancy between REMS and the reanalysis comes from terdiurnal tide differences and the amplitude difference of $\approx$ 10Pa, which is displayed in the diurnal tide.

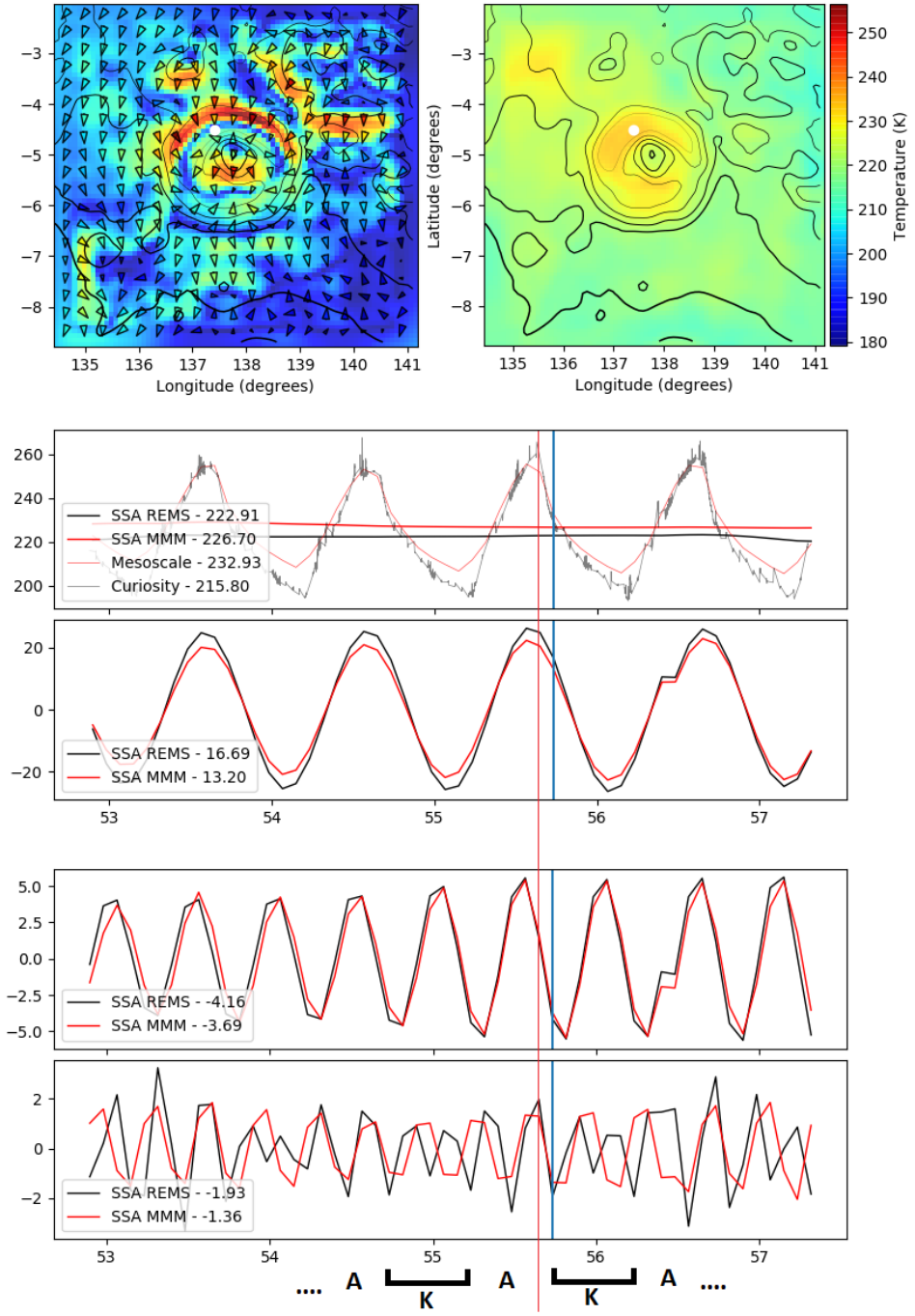


Figure 5.17: The SSA decomposition of temperatures from the MMM and Curiosity-REMS from the start of northern autumn. This figure is similar to 5.15 with one difference. The MMM temperatures shown in the bottom 4 plots (the SSA reconstructions) are extracted from Curiosity's location (same as in figure 5.15) from the layer 5m layer above the local topography. Hence the temperature simulated by the MMM is not exactly at the height of the REMS temperature sensor on Curiosity, nevertheless, it is the closest. The additional vertical red line is at the maximum of the REMS thermal tide and the "K"-marked and "A"-marked intervals katabatic and anabatic wind activity.

Reconstructions of temperatures are given in figure 5.17. The matching of the form between REMS and reanalysis temperature time series is in better agreement compared to surface pressures. The peaks in the thermal tide seem to have the diurnal and semidiurnal tide maxima almost in phase. The fast drop after reaching its peak temperature, in both the reanalysis and REMS observations, are caused again by the inflection point of the semidiurnal tide, which stands in the interval between the red and blue vertical lines in figure 5.17. Because the peak of the semidiurnal tide has a faster drop to the right than to the left side, this produces a faster drop to the right of the overall thermal tide. This is also coupled with the minimum point in the terdiurnal tide which seems to be in sync with the inflection point of the semidiurnal tide, thus favouring the fast temperature drop. This is similar to the surface pressure case, in that the inflection point of the semidiurnal tide is in phase with the extreme point in the terdiurnal tide. Moving forward to the snapshot depicted in figure 5.17, the katabatic wind activity from the crater topography is captured. The katabatic winds keep for the whole duration of the night-time semidiurnal tide peak, denoted by the "K" interval in the figure, following to the right of the blue line. This meteorological activity depicted by katabatic winds seems to only influence the terdiurnal tide, as the two "K"-intervals in the terdiurnal tide structure are very similar to each other, both in the reanalysis and REMS values. The anabatic wind activity is brief and highly correlated to the peaks of the diurnal and semidiurnal tides at and before the red vertical line. These come from the outside circulation which is driven by the presence of the minimal pressure node in the vicinity of the Gale Crater region, thus bringing strong winds to the crater which win over the local circulation. Their manifestation is essentially at the constructive interference between the thermal diurnal and semidiurnal tides, in figure 5.17. The results from the reanalysis from this period (beginning of northern autumn) are nonetheless matching the REMS observations (both in pressures and temperatures), because

the SSA eigenvalues which depict diurnal correlations, match between the two datasets.

In the other seasons the SSA reconstructions do seem to tell a different story. Figures 5.18 and 5.19 are surface pressure and temperature SSA analyses from northern spring which are equivalent to figures 5.15 and 5.17 shown previously for northern autumn. The sample size for the SSA method is significantly shorter, hence the method is less reliable, but certainly much better than the FFT method (presented in the previous subsection, 5.2.2). For this reason, structures like the one to the left side of the vertical blue line in figure 5.18, for the surface pressure diurnal tide, are not physical. The SSA method is simply trying to cope with the abrupt ending of the tide at the end of the depicted period. Nonetheless, the snapshot in the 2 MMM plots at the top (which is represented by the blue line in the SSA reconstructions below) is from the end of a high wind activity during day time. The diurnal, semidiurnal and terdiurnal tides are all at their lowest, hence the arrival of the pressure node brings this high wind activity period, which ends when the pressure is at its maximum. In the next time step (after 2 Martian hours) the winds from the circulation outside the crater cease in strength, and anabatic winds from the local topography return as the dominant circulation. In the temperature field from figure 5.19, the high wind activity corresponds to the rising of the diurnal, semidiurnal and terdiurnal temperature tides, and this period repeats in other sols as well, at the same spots. The thermal terdiurnal reconstructions show phase variations between REMS and the reanalysis, from being in resonance to being in antiphase, suggesting that local meteorology is variable between the two datasets. Nonetheless, there is a periodicity in the REMS temperature terdiurnal tide, with two higher amplitude peaks repeating at the start and at the end of the two and a half sol interval. This return to higher terdiurnal amplitude is not picked up by the reanalysis signature. The semidiurnal tide reconstructions for both surface pressures and temperatures shows a poor agree-

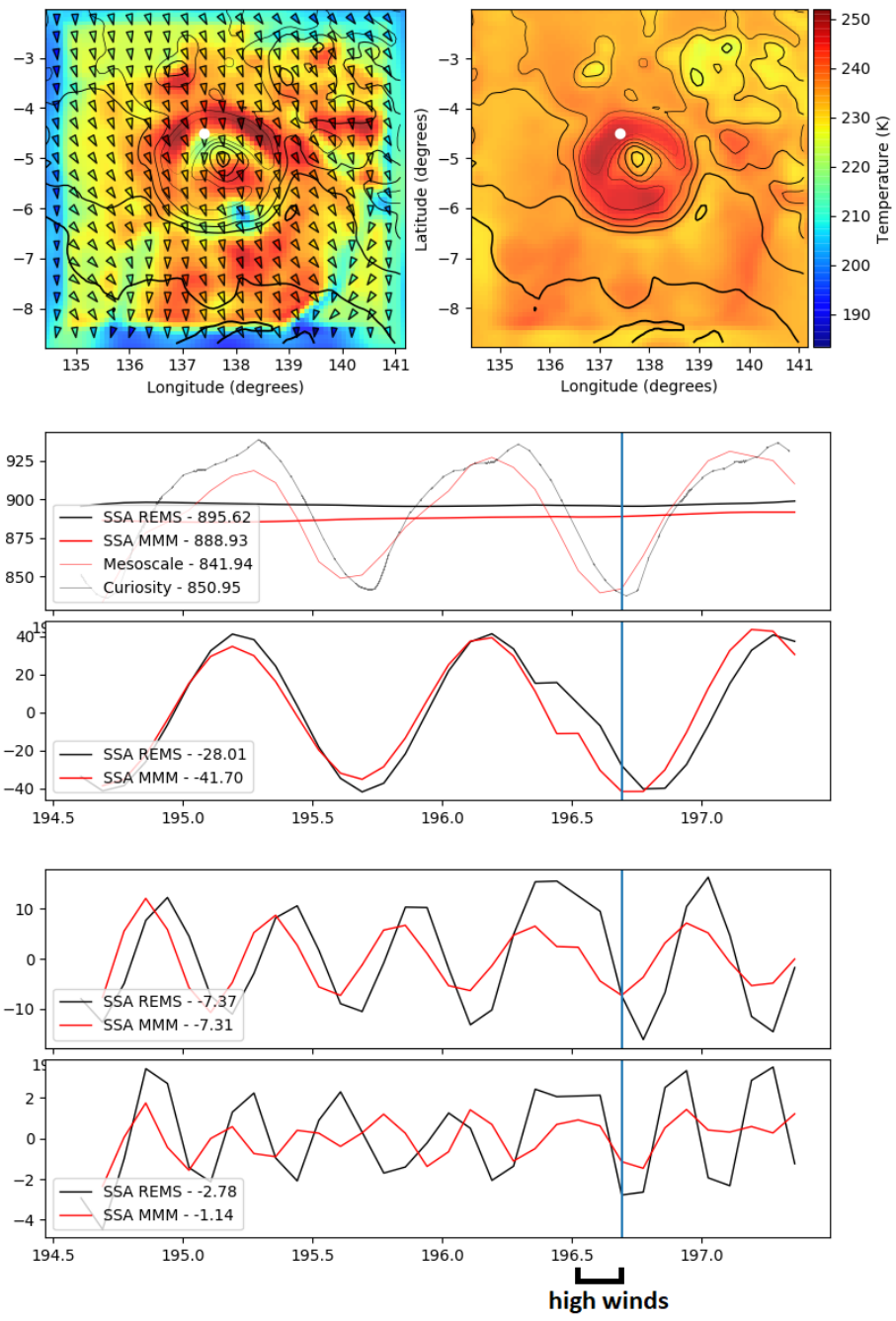


Figure 5.18: The SSA decomposition of surface pressures from the MMM and Curiosity-REMS from the start of northern winter. This figure is similar to 5.15. The chosen snapshot in the top two MMM plots is from the maximal wind activity at Gale Crater, from this short period. The extent of the period that is shown in this figure is shorter compared to northern autumn or spring, because the overlap with REMS is over these particular sols from the reanalysis. The high wind activity frames are depicted at the bottom of the figure

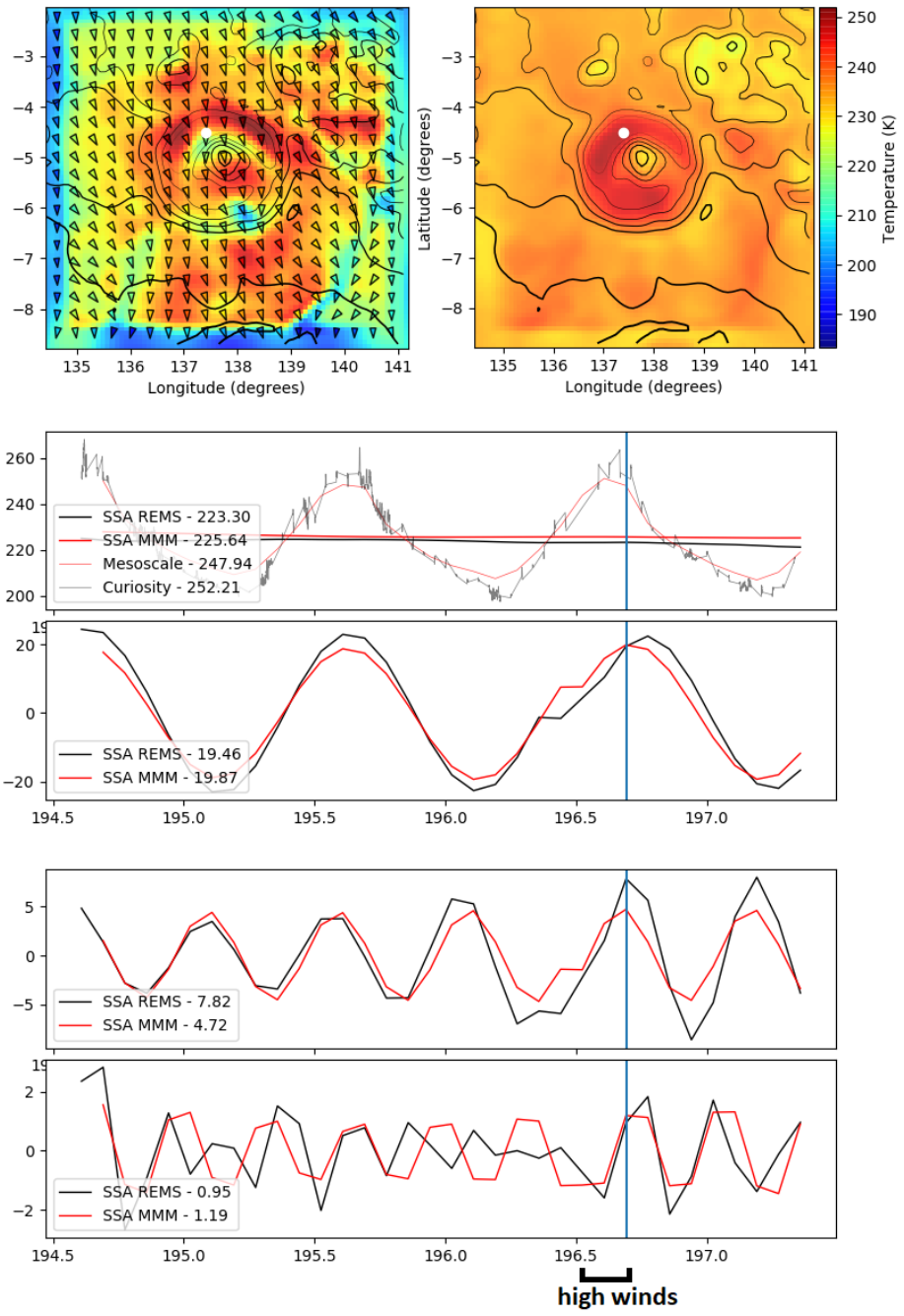


Figure 5.19: The SSA decomposition of temperatures from the MMM and Curiosity-REMS from the start of northern winter. This figure is similar to 5.18, but for temperatures.

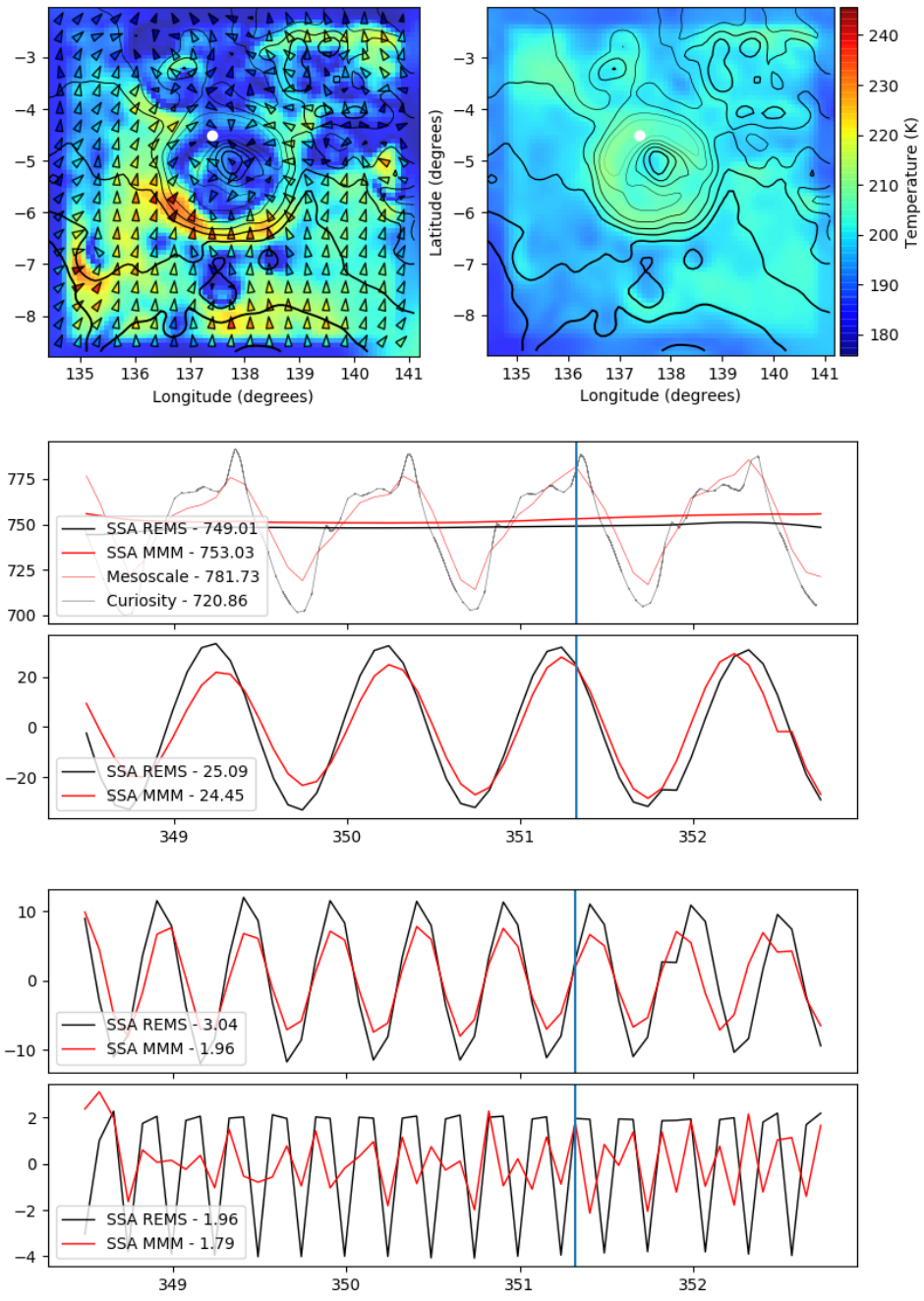


Figure 5.20: The SSA decomposition of surface pressures from the MMM and Curiosity-REMS from the start of northern spring. This figure is similar to 5.15.

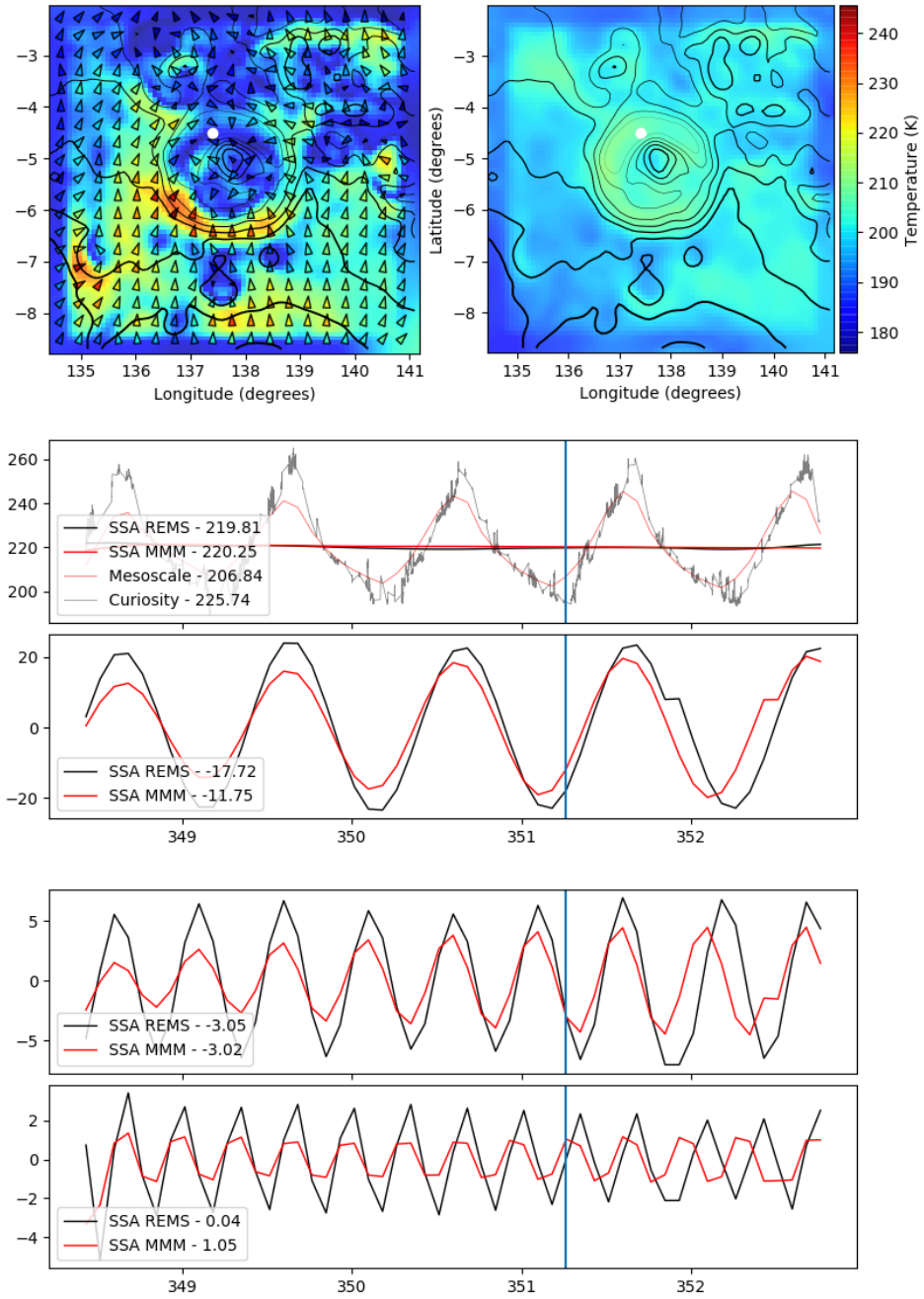


Figure 5.21: The SSA decomposition of temperatures from the MMM and Curiosity-REMS from the beginning of northern spring. This figure is similar to 5.20, but for temperatures instead.

ment between REMS and the reanalysis values, in comparison to the northern autumn period, suggesting that reanalysis dust and aerosol loading in this period is out of phase and lower in its semidiurnal component.

The northern spring season shows a much better agreement, even in the terdiurnal tide. The beginning of spring is depicted in figures 5.20 for the surface pressure, and 5.21 for temperatures. Compared to the previous seasons, the REMS pressures and temperatures show a much consistent oscillatory behaviour in the terdiurnal tide. Similarly, the reanalysis terdiurnal tide is also in phase for the beginning of the 5 sol period. The anabatic and katabatic wind activity and the daytime wind periods, are in the same temporal order in the diurnal patterns as the two previous seasons. The major differences between REMS observations and the reanalysis in the 3 cases presented in this chapter, usually comes from the terdiurnal tide, as their eigenvectors are much more variable. Agreement between REMS and the reanalysis can be reached, as seen in the northern spring terdiurnal tide reconstructions. Their terdiurnal eigenvector pair from figure 5.14 also shows a better agreement compared to the other seasons. But these correlations have a higher variability between the seasons. Nonetheless, the SSA method is indeed a powerful tool in making comparisons between model data and in situ observational data; and the MGCM+MMM reanalysis, even if it is at a lower resolution compared to the NASA Ames GCM $\rightarrow$ MRAMS simulation, it does shows remarkable agreement in its SSA eigenvector decomposition, which are the template for extraction of correlations. The difference comes only from the uncorrelative amplitudes depicted by SSA principal components.

5.3 Final remarks. An overview of the methodology

As mentioned before in subsection 5.2.2, the MGCM+MMM was not optimized for longer runs, hence the reanalysis case studies involve mostly 5 sol packages

(due to the computationally expensive tasks). Nonetheless, the MGCM+MMM shows strong potential in constructing the environment at Gale Crater, with robust agreement with REMS observations in both the diurnal and semidiurnal tides. This agreement between reanalysis and REMS values, is seen in the SSA eigenvector pairs for the diurnal and semidiurnal signals. As the eigenvectors are constructed from a diurnal window of correlations, this means that the model internal correlations are designed correctly. The differences between model and observations come from uncorrelated effects for the first two tidal components, which manifest amplification effects to the signals of the eigenvectors. The terdiurnal tide is the trickiest to capture in the model, but it does show signs of agreement with REMS in one of the seasons (northern spring). Additionally, the short windows of overlap with REMS measurements and the spin-up period for the MMM, proved to be challenging for the FFT method, however, it did reconstruct the diurnal and semidiurnal tides to a certain extent. The SSA method on the other hand, was efficient at extracting the correct signals, isolating the spin-up and the amplifying effects of the crater, and dust and aerosol atmospheric loading. Some small artefacts in the SSA reconstructions did appear in the northern winter analysis, as the overlap period between the reanalysis and REMS was slightly longer than 2.5 sols.

Nonetheless, the superposition of the diurnal, semidiurnal and terdiurnal tides offer an explanation to the evolution of the surface pressures and temperatures detected by REMS, and the local meteorology simulated by the reanalysis does seem to be periodic and correlated to the tides as well. In the time window of maximum surface pressure tide, there is usually a destructive interference between the diurnal and semidiurnal tidal components, with a sharp peak in pressure occurring two hours after the maximal diurnal tide pressure. This was established to be from the influence of the semidiurnal tide, where the pressure reaches the inflection point. Similarly, we discovered a constructive interference in the tem-

perature diurnal and semidiurnal tides, which gives rise to the peak shape of the thermal tide. Relating to the previous studies by Lewis & Barker (2005) and Guzewich et al (2015), namely that the diurnal and semidiurnal tides are strongly correlated to the atmospheric dust and aerosol loading, we conclude that these are the main influences to the REMS surface pressure measurements. Furthermore, anabatic and katabatic winds are the strongest signal from the local crater circulation, for most of a typical sol, and in a short interval of 6 hours before the constructive interference from the thermal diurnal and semidiurnal tides is set, the regional anabatic winds from the southern highlands dominate the circulation at Gale Crater. An interesting exercise would be to accomplish the same analysis between the reanalysis and REMS, but for velocities. However, due to the unreliability of velocity data from REMS, this would not make for a pertinent analysis. Nevertheless, the backbone of the REMS observations come from diurnal and semidiurnal signals and are correctly simulated by the reanalysis, and an analysis of the local meteorology from the reanalysis would require valid local velocity measurements from REMS.

5.3.1 Sources of uncertainties in the reanalysis

The topic of what can compromise reanalysis precision was indirectly discussed throughout this current chapter. To summarise, the major uncertainties in the reanalysis come from dust loading and incorrect representation of the topography. The amplitude of the semidiurnal tide in the reanalysis is lower compared to REMS. This is an indication that loaded dust is underestimated in the model, as both atmospheric dust and semidiurnal tide amplitude are strongly correlated (Lewis & Barker, 2005). The other possible reason is related to the slope winds, although without a study of the REMS winds, this can not be tested. The Conclusions chapter presents an alternative method of clarifying this discrepancy and testing whether slope winds are the cause or dust. The other aspect is the gen-

eral underestimation of surface pressures in both the MGCM and the MMM. The MGCM representation of the topography at Gale Crater is highly smoothed due to the low working resolution of a global model. The MMM significantly improves the representation of the crater, however, it still has pixels of 5×5 km^2 to construct a crater of 150km in diameter; the surface pressure is still underestimated. Additional sources of uncertainties appear in the model parametrizations, either as internal numerical errors or incorrect assessment of parameter space. These are minimised by the model calibrations (discussed in subchapter 2.3) and the data assimilation scheme (which repairs the model state by convergence to observations).

6 Conclusions and Future work

This thesis is a continuation of the ongoing pursuit of understanding the Martian near-surface atmosphere in connection to the general circulation, from the comparison with in situ observations; specifically the REMS instrument on-board NASA's Curiosity rover (Gómez-Elvira et al, 2014, Guzewich et al, 2015, Kahanpää et al, 2016, Pla-Garcia et al, 2016, Rafkin et al, 2016, Tyler & Barnes, 2013 and many others). Our contribution is twofold, by first engaging in proving that Mars' atmosphere supports the use of an embedded mesoscale model into a GCM (more details provided in chapter 3) and second, using the Oxford MGCM's data assimilation capabilities to produce a MGCM+MMM reanalysis of the environment at Gale Crater (described in chapter 4). The rest of our work involved designing state-of-the-art tools for performing the first spectral energy budget of the Martian atmosphere (also in chapter 3), and achieving the comparison between the reanalysis and REMS observations, from SSA and FFT decomposition (details provided in chapter 5).

6.1 The spectral energetics of the Martian atmosphere

This work involved analysing the atmosphere on Mars from the framework of spectral energy transfers between the two main energy reservoirs (kinetic and available potential energy), with the inclusion of Martian topography and the separation over pressure levels as well. On Earth, this diagnostic shows two inertial ranges in the energy spectrum, namely a large-scale enstrophy dominated cascade characterised by a k^{-3} dependence in the spectrum, followed by an energy cascade at smaller scales with a $k^{-5/3}$ dependence. For more details, Lindborg & Augier (2013) is a good start as it is also the paper on which our methodology is based on. Additionally, for a better understanding, appendix D and chapter 3 provide a more in-depth presentation. Nonetheless, the Martian atmosphere appears to

manifest only a $k^{-5/3}$ in the first 31 wavenumbers. The total energy flux was proven to be positive over this range, as is the case on Earth, but with differences in both spectral ranges and energy amplitudes. Moreover, even if the spectrum does not reach 170 wavenumbers, which is the resolution of the MGCM in our studies, it is reassuring that Mars seems to cascade energy from large to small scales, enabling the use of our MGCM+MMM configuration. An important feature that is worth pursuing, is performing the spectral analysis to mesoscales. Already Mars' atmosphere showed an important property, such as the absence of an enstrophy inertial range. The field of atmospheric turbulence would benefit from learning from extraterestrial atmospheres, which are in different circulation regimes compared to Earth.

We did not carry our analysis to 170 wavenumbers because the output from the MGCM is in sigma levels and being a hydrostatic model, and there was not enough time to compute vertical velocities (which can be computed diagnostically from the MGCM fields). As these are a requirement for the spectral diagnostics, we opted to apply the procedure to a version of the MACDA dataset spanning from the middle of Martian year 28 to the beginning of 31. We will extend the procedure to apply to any hydrostatic GCMs to extend the study to 170 wavenumbers, when publishing the work of chapter 3.

6.2 The MGCM+MMM configuration

Due to the intensive processing power needed for T170 MGCM runs, the reanalysis intervals were 5 sols long, situated in time at the beginning of each Martian season. In doing the reanalysis, the UK/LMD MGCM and LMD MMM were interfaced in order for any horizontal patch from the MGCM to be downscaled by the MMM. This interface is still under development as the restart system of the LMD MMM was not automated yet. Nonetheless, the reanalysis reproduced a

nomenclature of atmospheric processes at Gale Crater, including orographic gravity waves, the depth of the planetary boundary layer and the low-level nocturnal jet, the local anabatic and katabatic winds from the crater rim and central peak as well as the influence of the crater on external winds (more details are provided in 4). The resolution of the reanalysis consisted of $80 \times 80 \times 60$ grid-boxes spanning a 400km $\times$ 400km region around Gale Crater. This resolution is much lower compared to the previous studies by Rafkin et al (2016) (330m grid-boxes), nonetheless, our reanalysis performed at least comparable to their simulations, proving the advantage of assimilating spacecraft data in the MGCM+MMM configuration.

The MGCM+MMM will be optimized for longer runs in our future studies, by reducing the computationally intensive T170 resolution in the MGCM and performing multiple embeddings of the MMM, which is also the parallelized component of the MGCM+MMM configuration (it can run on a cluster). In addition, our studies were based on the reanalysis of MCS temperatures only. However, different fields were provided for the data assimilation kit in the past, such as the assimilation of water-ice (Steele et al, 2014), column dust optical depths (used for the production of the MACDA dataset; "Mars Analysis Correction Data Assimilation", n.d.) and dust retrievals (Ruan, 2015). Enabling these features alongside the mentioned optimisations would be invaluable for future reanalyses using the MGCM+MMM configuration.

6.3 The FFT/SSA decomposition of REMS measurements

When comparing with REMS observations, the reanalysis temperatures were interpolated to a terrain-following level, 5m above the local topography. Hence, the time-series of surface pressures and temperatures were produced from the closest grid-point to Curiosity's location. The time-dependent rebinned REMS observa-

tions and the sequences from the reanalysis, were decomposed into Fourier and SSA components. The eigenvectors from the SSA method showed almost perfect correlations between the reanalysis and REMS observations in the diurnal and semidiurnal tidal components. The principal components were similar to the FFT components, proving that the differences in the first two tidal components between the reanalysis and REMS is almost completely uncorrelative. As the correlations are virtually identical between model and observations, we proved a previous result by Tyler & Barnes (2013), that the differences come from amplification effects from the crater topography. Taking previous studies into account (Guzewich et al, 2015 and Lewis and Barker, 2005), we connected the differences in the principal components to atmospheric dust and aerosol loading. An SSA analysis of REMS dust observations would be a practical addition to our analysis, however, the response of the first two tidal components to atmospheric dust and aerosol loading is so efficient that the amplitude of the semidiurnal tide was even used in estimating the global mean aerosol loading (Lewis and Barker, 2005, Zurek, 1980). The other tidal components were influenced by local meteorological effects and their matching between model and observations was situational (in the spring season), but certainly not perfect. Comprehending the terdiurnal tide without wind data is difficult, however, weather patterns do seem to leave periodic marks in the terdiurnal tide (more details are provided in subchapter 5.2.3).

The SSA method will also be used to quantify the improvements to the atmospheric tides, brought by the MMM to the reanalysis. Essentially, the MGCM reanalysis will be decomposed into SSA eigenvectors and principal components, to be compared to their values from the MGCM+MMM reanalysis (and REMS). Additionally the SSA can be used to formally validate the results from Haberle et al (2018). They claim to find very weak signals of northern hemisphere baroclinic waves in the REMS data, from comparison between Viking observations. The SSA will bring a complete answer to the claim through it's powerful filtering property

(more details in chapter 5). We also plan on extending the capabilities of our SSA algorithm to include space as well. To obtain spatio-temporal correlations from a fully temporal SSA only needs extension of the covariance matrix; the rest of the methodology remains unaltered. This would open up more opportunities, such as formally investigating the mechanisms for the amplification of the diurnal tide due to topography (at Gale Crater).

6.4 Final remarks

The exercise of setting up these studies provided one additional insight, different from the scientific ones discussed above, but nonetheless relevant for the community of Mars atmospheric modelling. These techniques can actually be used more extensively. The spectral energetics package can be used as a cross-comparison kit between global reanalyses, because it extracts information relevant to the model dynamics. Once this diagnostic tool is made widely available, it would be interesting to converge on a "correct" set of energy spectrum and spectral energy fluxes for the Martian atmosphere to be used as a template for model/reanalysis improvements. Additionally, such a "correct" set can can broaden our understanding of atmospheric turbulence. Finally, the matching of SSA eigenvectors between models and in situ observations also provides a workshop for model/reanalysis improvements, or at least to filter out unwanted influences from reanalyses (e.g. detecting the extent of a spin-up period; more details in subchapter 5.2.3).

7 References

- Albee *et al.* Overview of the Mars Global Surveyor mission. *J. Geophys. Res.*, **106**, pp. 23165-23945 (2001).
- "About Advanced Research Computing". http://www.arc.ox.ac.uk/content/about-advanced-research-computing (n.d.).
- Bagnold, R.A. The Physics of Blown Sand and Desert Dunes, Methuen, New York (1941).
- Barlow, N. Mars: An introduction to its Interior, Surface and Atmosphere. Cambridge University Press, print (2008).
- Barnes, S.L. Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memorandum ERL NSSL-62 (1973).
- Bell, R.S. & Dickinson, A. The Meteorological Office operational numerical weather prediction system. *Sci. Paper*, **41**. Meteorological Office (1987).
- Bergthorsson, P. & Döös, B.R. Numerical weather map analysis. *Tellus*, **7**, pp. 329-340 (1955).
- Bjerknes, V. Das Problem der Wettervorhersage, betrachtet vom Stanpunkt der Mechanik und der Physik. *Meteor. Zeits.*, **21**, pp 1-7 (1904).
- Blackadar, A. Boundary-layer wind maxima and their significance for the growth of nocturnal inversion. *Bull. Am. Meteorol. Soc.*, **38**, 283-290 (1957).
- Boer, G.J. On exact and approximate energy equations in pressure coordinates. *Tellus A*, **41(2)**, pp. 97-108 (1989).
- Boer & Shepherd. Large-scale two-dimensional turbulence in the atmosphere. *J. Atmos. Sci.*, **40**, pp. 164-184 (1983).

- Charney, J.G., Fjørtoft, von Neuman J. Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, pp. 237-254 (1950).
- Charney, J.G. Geostrophic turbulence. *J. Atmos. Sci.*, **28**, pp. 1087-1095 (1971).
- Clark, T.L. & Gall, R. Three-dimensional numerical model simulations of airflow over mountainous terrain. A comparison with observations. *Mon. Wea. Rev.*, **110**, pp. 766-791 (1982).
- Collins, M., Lewis, S.R., Read, P.L. Baroclinic Wave Transitions in the Martian Atmosphere. *Icarus*, **120**, pp. 344-357 (1996).
- Collins, M., Lewis, S.R., Read, P.L. Gravity wave drag in a global circulation model of the Martian atmosphere: parameterisation and validation. *Adv. Space Res.*, **19(8)**, pp. 1245-1254 (1997).
- Crapper, G.D. A three-dimensional solution for waves in lee of mountains. *J. Fluid Mech.*, **6**, pp. 51-76 (1959).
- Cressman, G.P. An operational objective analysis scheme. *Mon. Weather Rev.*, **87**, pp. 367-374 (1959).
- Curiosity(rover). http://en.wikipedia.org/wiki/Curiosity_(rover) (n.d.).
- Curiosity mission overview. http://mars.jpl.nasa.gov/msl/mission/overview (n.d.).
- Documentation on the REMS. http://atmos.nmsu.edu/PDS/data/mslrem_1001/DOCUMENT/ (n.d.).
- Earth's Magnetic Field. http://principles.ou.edu/mag/earth.html (n.d.).

- Eliassen *et al.* On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. *Inst. of Theor. Met, Univ. of Copenhagen*, Report No. 2 (1970).
- "Explore NASA's Mars Map": https://mars.nasa.gov/maps/explore-mars-map/fullscreen/ (n.a.).
- "First Data from Mars Climate Sounder Aboard NASA Mars Reconnaissance Orbiter": http://www.spaceref.com/news/viewsr.html?pid=20143 (n.a.).
- Fita, L. LMDZ physical schemes. A brief description for the LMDZ-B configuration. At: http://lmdz.lmd.jussieu.fr/developpeurs/notes-techniques/ressources/LMDZphys.pdf (2013).
- Fjørtoft, R. One the changes in the spectral distribution of kinetic energy for two-dimensional, nondivergent flow. *Tellus*, **5**, pp. 225-230 (1953).
- Forget *et al.* CO_2 Snowfall on Mars: Simulation with a General Circulation Model. *Icarus*, **131**, pp. 302-316 (1998).
- Forget *et al.* Improved general circulation models of the Martian atmosphere from the surface to above 90 km. *J. Geophys. Res.*, **104**, pp. 24155-24176 (1999).
- Frisch, U. Turbulence. Cambridge University Press, print (1995).
- Gage & Nastrom. A Climatology of Atmospheric Wavenumber Spectra of Wind and Temperature Observed by Commercial Aircraft. *J. Atmosph. Sci.*, **42**. pp. 950-960 (1984).
- Ghil, M., Allen, R.M., Dettinger, M.D., Ide, K., Kondrashov, D. Advanced spectral methods for climatic series. *Rev. Geophys.*, **40(2)**, pp. 1-41 (2002).

- Gierasch, P. The north-south martian divide. *Nature*, **416**, pp. 269-270 (2002).

- Golombek, M., et al. Selection of the Mars Science Laboratory landing site. *Space Science Reviews*, **170**, pp. 641-737 (2012).

- Gomez-Elvira *et al.* Curiosity's rover environmental monitoring station: Overview of the first 100 sols. *J. Geophys. Res. Planets*, **119**, pp. 1680-1688 (2014).

- Golyandina, N., Zhigljavsky, A. Singular Spectrum Analysis for Time Series. Springer (2013).

- Gossard, E.E., & Hooke, W.H Waves in the Atmosphere. Elsevier, print (1975).

- Gradshteyn, I.S. & Ryzhik, I.M. Table of Integrals, Series, and Products. Academic press, print (1965).

- Guzewich *et al.* Atmospheric Tides in Gale Crater, Mars. *Icarus*, **268**, pp. 37-49 (2016).

- Haberle, R.M., et al. Preliminary Interpretation of the REMS pressure data from the first 100 sols of the MSL mission. *J. Geophys. Res.*, **119**, pp. 440-453 (2014).

- Held, I.M., Hoskins, B.J. Large-Scale Eddies and the General Circulation of the Troposphere. *Adv. Geophys.*, *28A*, pp. 3-31 (1985).

- Hoskins & Simmons. A multi-layer spectral model and the semi-implicit method. *J. R. Met. Soc*, **101**, pp. 637-655 (1975).

- Hourdin *et al.* Meteorological variability and the annual surface pressure cycle on Mars. *J. Atm. Sc.*, **50**, pp. 3625-3640 (1993).

- Joshi *et al.* Seasonal Variations in low level flow in the NASA-Ames Mars GCM. *Adv. Space Res.*, **19(8)**, pp. 1261-1265 (1997).

- Kahanpää *et al.* Convective vortices and dust devils at the MSL landing site: Annual variability. *J. Geophys. Res. Planets*, **121**, pp. 1514-1549 (2016).

- Kalnay, E., Lord, S., McPherson, R. Maturity of operational numerical weather prediction: the medium range. *Bull. Amer. Meteor. Soc.*, **79**, pp. 2754-2769 (1998).

- Klemp, J.B. & Lilly, D.K. Mountain waves and momentum flux. WMO GARP Publ. Ser., **23**, pp. 115-141 (1980).

- Klemp, J.B. & Lilly, D.K. The dynamics of wave induced downslope windstorms. *J. Atmos. Sci*, **32**, pp. 320-339 (1975).

- Koenigsberger, L. Hermann von Helmholtz. Clarendon Press Oxford, pp. 357 (1906).

- Kolmogorov, A.N. Dissipation of Energy in the Locally Isotropic Turbulence *a,b and c. Math. and Phys. Sc.*, **434**, No. 1890, pp. 15-17 (1941).

- Koshyk *et al.* Kinetic energy spectrum of horizontal motions in the middle-atmosphere models. *J. Geophys. Res.*, **104**, pp 27,177-27,190 (1999).

- Kraichnan, R.H. Inertial-range transfer in two- and three-dimensional turbulence. *J. Fluid. Mech.*, **47**, pp. 525-535 (1970).

- Lait, L.R. An alternative form for potential vorticity. *J. Atmos. Sci.*, **51**, pp. 1754-1759 (1994).

- Laprise, R. The Euler Equations of Motion with Hydrostatic Pressure as an Independent Variable. *Mon. Wea. Rev.*, **120**, pp. 197-207 (1992).

- Law, K.J.H., Stuart, A.M., Zygalakis, K.C. Data Assimilation: A Mathematical Introduction. Springer (2015).

- Lewis, S.R., Barker, P.R. Atmospheric tides in a Mars general circulation model with data assimilation. *Adv. Space R.*, **36(11)**, 2162-2168 (2005).

- Lewis, S.R., Collins, M., Read, P.L., Forget, F., Hourdin, F., Fournier, R., Hourdin, C., Talagrand, O. & Huot, J,-R. A climate database for Mars. *J. Geophys. Res.*, **104**, pp. 24177-24194 (1999).

- Lewis *et al.* A Climate Database for Mars. *J. Geophys. Res.*, **104** (1999).

- Lindborg, E., Augier, P. A New Formulation of the Spectral Energy Budget of the Atmosphere, with Application to Two High-Resolution General Circulation Models. *J. Atmos. Sci.*, **70**, pp. 2293-2308 (2013).

- Lindborg, E. Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *J. Fluid Mech.*, **338**, pp. 259-288 (1999).

- Lindborg, E. The energy cascade in a strongly stratified fluid. *J. Fluid Mech.*, **550**, 207-242 (2006).

- Lorenc, A.C. Analysis methods for numerical weather prediction. *Quart. J. R. Met. Soc.*, **112**, pp. 1177-1194 (1986).

- Lorenc, A.C., Bell, R.S., Macpherson, B. The Meteorological Office analysis correction data assimilation scheme. *Q. J. R. Meteorol. Soc*, **117**, pp. 59-89 (1991).

- Lorenz, E.N. Available potential energy and the maintenance of the general circulation. *Tellus*, **7**, pp. 157 (1955).

- Lott, F. & Miller, M.J. A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Q. J. R. Meteorol. Soc.*, **123**, pp. 101-127

(1997).

- Lyra, G. Theorie der stationären Leewellenströmung in freier Atmosphäre. *Z. Ang. Math. Mech.*, **23**, pp. 1-28 (1943).

- Madeleine *et al.* Revisiting the radiative impact of dust on Mars using the LMD Gloval Climate Model. *J. Geophys. Res. (Planets)*, **116**, pp. 11010 (2011).

- "Map of all Mars landing sites, failed and successful". http://planetary.s3.amazonaws.com/assets/images/4-mars/2018/20181109_mars_lander_map.jpg (n.d.).

- "Mars Analysis Correction Data Assimilation" (MACDA): http://www2.physics.ox.ac.uk/research/geophysical-fluid-dynamics/macda (n.d.).

- Martinez *et al.* The Modern Near-Surface Martian Climate: A Review of In-situ Meteorological Data from Viking to Curiosity. *Space Sci Rev*, **212**, pp. 297-338 (2017).

- McGregor & Dix. An Updated Description of the Conformal-Cubic Atmospheric Model. In: Hamilton, K., Ohfuchi, W. (eds) High Resolution Numerical Modelling of the Atmosphere and Ocean. Springer, New York, NY (2008).

- Mellor, G.L. & Yamada, T. Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys and Space Phys.*, **20**, pp. 851-875 (1982).

- "Mercator projection": https://en.wikipedia.org/wiki/Mercator_projection (n.a.).

- Miller *et al.* Parametrization and influence of subgrid-scale orography in general circulation and numerical weather prediction models. *Meteorol. Atmos. Phys.*, **40**, pp. 84-109 (1989).

- Monin, A., Yaglom, A. Statistical Fluid Mechanics. MIT Press, Cambridge, Mass (1975).

- Montabone *et al.* Eight-year climatology of dust optical depth on Mars. *Icar*, **251**, pp 65-95 (2015).

- Montmessim *et al.* Origin and role of water ice clouds in the Martian water cycle as inferred from a general circulation model. *J. Geophys. Res. (Planets)*, **109**, E10004 (2004).

- Newman, C.E., Lewis, S.R, Read, P.L.. Modelling the Martian dust cycle, 1. Representations of dust transport processes. *J. Geophys. Res (Planets)*, **107**, pp. 5123 (2002).

- Orszag, S.A.. Transform method for calculation of vector coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.*, pp 890-895 (1970).

- Palmer, T.N., Shutts, G.J., Swinbank, R. Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. *Quart. J. R. Met. Soc.*, **112**, pp. 1001-1039 (1986).

- Philips, N.A. A coordinate system having some special advantages for numerical forecasting. *J. Meteor*, **14**, pp. 194-184 (1957).

- Phillips, D.S. Analytical Surface Pressure and Drag for Linear Hydrostatic Flow over Three-Dimensional Elliptical Mountains. *J. A. Sci.*, **41**, pp. 1073-1084 (1984).

- Pla-Garcia *et al.* The Meteorology of Gale Crater as Determined from Rover Environmental Monitoring Station Observations and Numerical Modeling. Part I: Comparison of Model Simulations with Observations. *Icarus*, **280**, pp. 114-138 (2016).
- "Planets2013". https://commons.wikimedia.org/wiki/File:Planets2013.svg (n.d.).
- Priestley, A. A Quasi-Conservative Version of the Semi-Lagrangian Advection Scheme. *Mon. Weather Rev*, **121**, pp. 621 (1993).
- Queney, P. Theory of perturbations in stratified currents with applications to air flow over mountain barriers. Dept. of Meteor., University of Chicago, Misc. Rep. No. 23, print (1947).
- Queney, P. The problem of air flow over mountains: A summary of theoretical studies. *Bull. Amer. Meteor. Soc.*, **29**, pp. 16-26 (1948).
- Queney, P., Corby, G.A., Gerbier, N., Koschmieder, H., Zierep, J. The airflow over mountains. WMO Tech, Note No. 34, print (1960).
- Rafkin *et al.* The meteorology of Gale Crater as determined from Rover Environmental Monitoring Station observations and numerical modeling. Part II: Interpretation. *Icarus*, **280**, pp. 114-138 (2016).
- Randall, D.A. Development of an advanced finite-difference atmospheric general circulation model. United States: N. p. (1992).
- Read *et al.* The Martian Planetary Boundary Layer. In: Haberle, M., Clancy, R.T., Forget, F., Smith, M.D., Zurek, R.W. (eds) The atmosphere and Climate of Mars. Cambridge: Cambridge University Press, print (2017).
- Read, P.L. & Lewis, S.R. The Martian Climate Revisited. Chichester, UK, Springer, print (2004).

- Read, P.L., Wang, Y., Tabataba-Vakili, F., Augier, P., Lindborg, E., Valeanu, A., Young, R.M.B. Comparative terrestrial atmospheric circulation regimes in simplified global circulation models: II. energy budgets and spectral transfers. *Q. J. R. Meteorol. Soc.*, **1**, pp. 1-19 (2018).

- Richardson, L.F. Weather Prediction by Numerical Process. Cambridge University Press: Cambrdige, MA. Reprinted 2006 by Cambridge University Press with a new introduction by Peter Lynch. (1922).

- Rogberg, L.P.H.T., Read, P.L., Lewis, S.R. & Montabone, L. Assessing atmospheric predictability on Mars using numerical weather predictions and data assimilation. *Quart. J R. Meteorol. Soc.*, **136**, pp. 1614-1635 (2010).

- Rover Environmental Monitoring Station. http://mars.jpl.nasa.gov/msl/mission/instruments/ environsensors/rems/ (n.d.).

- Ruan, T. The climate of Mars from assimilations of spacecraft data. PhD thesis, University of Oxford (2015).

- Satoh, M. Atmospheric Circulation Dynamics and General Circulation Models. Springer, print (2014).

- Shepherd, T.G. A unified theory of available potential energy. *Atmosphere-Ocean*, **31**, pp. 1-26 (1993).

- Shuman, F.G. History of numerical weather prediction at the National Meteorological Center. *Wea. Forecasting*, **4**, pp. 286-296 (1989).

- Simmons, A.J., Hoskins, B.J. The life Cycles of Some Nonlinear Baroclinic Waves. *J. Atmos. Sci.*, **35**, pp. 414-432 (1978).

- Skamarock, *et al.* A description of the Advanced Research WRF, version 2. (2005).

- Smith *et al.* Topography of Northern Hemisphere of Mars from the Mars Orbiter Laser Altimeter. *Science*, **279**, pp. 1686 (1998).
- Smith, R.B. The influence of mountains on the atmosphere. *Advances in Geophysics.*, **21**, pp. 87-230 (1979).
- Smith, R.B. Linear theory of stratified hydrostatic flow past an isolated mountain. *Tellus*, **32**, pp. 348-364 (1980).
- Spiga, A & Forget, F. A new model to simulate the Martian mesoscale and microscale atmospheric circulation: Validation and first results. *J. Geophys. Res. (Planets)*, **35**, pp L15201 (2008).
- Spiga, A., Forget, F., Madeleine, J.-B., Montabone, L., Lewis, S.R., Millour, E. The impact of martian mesoscale winds on surface temperature and on the determination of thermal inertia. *Icar*, **212**, pp 504-519 (2011).
- Steele, L.J., Lewis, S.R., Patel, M.R. The radiative impact of water ice clouds from a reanalysis of Mars Climate Sounder data. *Geoph. R. Lett.*, **41(13)**, pp. 4471-4478 (2014).
- Stull, R.B. An Introduction to Boundary Layer Meteorology, Springer, Dordrecht, Netherlands (1988).
- Stull, R.B. Static Stability - An Update. *Bull. Amer. Meteor. Soc.*, **72**, pp. 1521-1530.
- Sukumar, S.S. A Study on Third Order Runge-Kutta Techniques for Solving Practical Problems. *Walailak Jour. Sci. & Tech.*, **11**, pp. 679-686 (2014).
- Tabataba-Vakili, F., Read, P.L., Lewis, S.R., Montabone, L., Ruan, T., Wang, Y., Valeanu, A., Young, R.M.B. A Lorenz/Boer energy budget for the atmosphere of Mars from a "reanalysis" of spacecraft observations. *Geophys. Res. Lett.*, **42**, 8320-8327 (2015).

- "The Mars Orbital Laser Altimeter". http://mola.gsfc.nasa.gov/images.html (n.d.).

- Thompson, P..D. Charney and the revival of NWP. In *The Atmosphere - a challenge*, Lindzen, E., Lorenz, E., Platzman, G., editors. *Am. Meteor. Soc., Boston*, pp. 93-119 (1990).

- Toigo, A.D., Waugh, D.W., Guzewich, S.D. What causes Mars' annular polar vortices? *Geophys. Res. Lett.*, **44**, pp. 71-78 (2017).

- Tyler, D. & Barnes, J.R. Mesoscale Modeling of the Circulation in the Gale Crater Region: An Investigation into the Complex Forcing of Convective Boundary Layer Depths. *Mars J.*, **8**, pp. 58-77 (2013).

- "Welcome to the PDS Atmospheres Node". https://pds-atmospheres.nmsu.edu (n.a.).

- Wilson, R.J., Hamilton, K. Comprehensive model simulation of thermal tides in the martian atmosphere. *J. Atmos. Sci.*, **53**, pp. 1290-1326 (1996).

- Wurtele, M.G. The three-dimensional lee wave. *Contrib. Atmos. Phys.*, **29**, pp. 242-252 (1957).

- Zuber *et al.* The Mars Observer layer altimeter investigation. *J. Geophys. Res*, **97**, pp. 7781-7797 (1992).

- Zurek, R.W. Surface pressure response to elevated tidal heating sources: Comparison of Earth and Mars. *J. Atmos. Sci*, **37**, pp. 1132-1136 (1980).

- Zurek & Smrekar. An overview of the Mars Reconnaissance Orbiter (MRO) science mission. *J. Geophys. Res*, **112** (2007).

Appendices

A Analysis equations

Consider a certain forecast model state as a rank N_x vector $\mathbf{x}_{t'}$, containing all the quantities we wish to estimate at the particular time t'. The basis for this vector will be denoted $\mathbf{x}$ and the individual quantities are N_x orthonormal vectors. In a similar fashion, the observed data is denoted by the vector $\mathbf{y}$ of dimension N_y, and the operator which relates $\mathbf{y}$ to the model basis $\mathbf{x}$ is denoted $\mathbf{K}$, such that

$$\mathbf{y} = \mathbf{K}(\mathbf{x}). \tag{A.1}$$

Because the observations are known, our quest is now translated into an inverse problem of finding the solution in $\mathbf{x}$-space that satisfies equation (A.1). The $\mathbf{K}$ operator is known in inverse theory terminology as the "forward model" or "forward process". The numerical approach to solving (A.1) clearly requires having a strategy for picking the best model state out of a range of possible states. Our pursuit is now reduced to a form of numerical variational principle. The most general approach for obtaining an equation to be minimized, is to turn to Bayesian analysis. Hence, we start with Bayes' theorem: For a given probability for event A to occur, knowing that event B has already occurred ($P(A|B)$), the following relation has to be satisfied:

$$P(A|B) \propto P(B|A)P(A), \tag{A.2}$$

where $P(B|A)$ is the probability of B occurring given that A has occurred and $P(A)$ is the probability of even A to occur independently. By applying Bayes' theorem multiple times for the set model states and observations, we can use it for our multidimensional case. Hence, we can attribute A to the event $\mathbf{x} = \mathbf{x}_t$

and B to $\mathbf{y} = \mathbf{y}_o$, where the subscripts t and o indicate the true state $\mathbf{x}_t$ and the observed state $\mathbf{y}_o$ respectively. Then, we can generalize equation (A.2) to a discrete probability distribution function (PDF) equation, by writing

$$P(\mathbf{x} = \mathbf{x}_t|\mathbf{y} = \mathbf{y}_o) \propto P(\mathbf{y} = \mathbf{y}_o|\mathbf{x} = \mathbf{x}_t)P(\mathbf{x} = \mathbf{x}_t). \tag{A.3}$$

The same dimensions from equation (A.1) are retained in (A.3). The probability $P(\mathbf{x} = \mathbf{x}_t)$ must also include the prior knowledge of the atmosphere since $\mathbf{x}_t$ can be thought of as the model state before the actual measurements are taken (independent of $\mathbf{y}$). Following this strategy, In the nomenclature of inverse methods, the best estimate for the most likely model state in the absence of observations, is the *background* or $\mathbf{x}_b$; this is considered to be known. The PDF of the deviation from the background, prior to the moment of observation input, should satisfy the following relation

$$P(\mathbf{x} = \mathbf{x}_t) = P_b(\mathbf{x} - \mathbf{x}_b). \tag{A.4}$$

Furthermore, considering the observational errors, we can write the PDF for the deviation of the measured values $\mathbf{y} = \mathbf{y}_o$ from the actual (*true*) values $\mathbf{y}_1 = \mathbf{y}_t$ in terms of their conditional PDF as follows

$$P(\mathbf{y} = \mathbf{y}_o|\mathbf{y}_1 = \mathbf{y}_t) = P_o(\mathbf{y}_o - \mathbf{y}_t). \tag{A.5}$$

For completeness, the errors in the forward model $\mathbf{K}$ are also adopted in the scheme. This will also introduce a deviation in the model state from transforming to observation space $\mathbf{y}_f = \mathbf{K}(\mathbf{x}_t)$. Thus, the error in $\mathbf{K}$ should introduce this additional relation

$$P(\mathbf{y}_1 = \mathbf{y}_t|\mathbf{x} = \mathbf{x}_t) = P_f(\mathbf{y}_1 - \mathbf{y}_f). \tag{A.6}$$

For exact $\mathbf{K}$, we have $\mathbf{y}_f = \mathbf{y}_t$ and hence this equation becomes

$$P(\mathbf{y} = \mathbf{y}_o|\mathbf{x} = \mathbf{x}_t) = P_o(\mathbf{y}_o - \mathbf{y}_t) = P_{of}(\mathbf{y}_o - \mathbf{y}_t). \tag{A.7}$$

To get a relation between ranges of probabilities, we can use the differential form $dP(A) = P(A|B)P(B)db$, which represents the probablity between the corresponding parameters b and $b + db$ for the event B. Hence, the integration of this relation will be

$$P(\mathbf{y} = \mathbf{y}_o|\mathbf{x} = \mathbf{x}_t) = \int P(\mathbf{y} = \mathbf{y}_o|\mathbf{x} = \mathbf{x}_t, \mathbf{y}_1 = \mathbf{y}_t)P(\mathbf{y}_1 = \mathbf{y}_t|\mathbf{x} = \mathbf{x}_t)d\mathbf{y}_1. \tag{A.8}$$

or using the above formalism along with $\mathbf{y}_f = \mathbf{K}(\mathbf{x}_t) \neq \mathbf{y}_t$

$$P_{of}(\mathbf{y}_o - \mathbf{y}_f) = \int P_o(\mathbf{y}_1 - \mathbf{y}_o)P_f(\mathbf{y}_1 - \mathbf{y}_f)d\mathbf{y}_1. \tag{A.9}$$

Substituting equations (A.4) and (A.6) into Bayes' theorem (A.3) produces:

$$P_a(\mathbf{x}) \propto P_{of}(\mathbf{y}_o - \mathbf{K}(\mathbf{x}))P_b(\mathbf{x} - \mathbf{x}_b), \tag{A.10}$$

$$P_a(\mathbf{x}) \propto \left\{ \int P_o(\mathbf{y}_1 - \mathbf{y}_o)P_f(\mathbf{y}_1 - \mathbf{K}(\mathbf{x}))d\mathbf{y}_1 \right\} P_b(\mathbf{x} - \mathbf{x}_b), \tag{A.11}$$

where we will define P_a as the actual PDF for the analysis. In other words, the best estimate will either be the expectation value or the most probable value. These are defined as

$$\mathbf{x}_a = \int \mathbf{x}P_a(\mathbf{x})d\mathbf{x} \tag{A.12}$$

or

$$\mathbf{x}_a \text{ for } dP_a(\mathbf{x}_a) = 0. \tag{A.13}$$

Note that $P_b(\mathbf{x} - \mathbf{x}_b)$ and $P_{of}(\mathbf{y}_o - \mathbf{K}(\mathbf{x}))$ represent the background and observational errors respectively. The two events should be uncorrelated, hence, in deriving (A.11), it is assumed that $P_b(\mathbf{x} - \mathbf{x}_b)$ and $P_{of}(\mathbf{y}_o - \mathbf{K}(\mathbf{x}))$ are independent functions. A more thorough derivation which considers their joint PDF is possible, but omitted here.

Our DA scheme assumes Gaussian PDFs, hence the unknown PDFs in equation (A.11) are

$$P_b(\mathbf{x} - \mathbf{x}_b) \propto \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)\right], \tag{A.14}$$

$$P_o(\mathbf{y} - \mathbf{y}_o) \propto \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{y}_o)^T \mathbf{O}^{-1}(\mathbf{y} - \mathbf{y}_o)\right], \tag{A.15}$$

$$P_f(\mathbf{y} - \mathbf{y}_f) \propto \exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{y}_f)^T \mathbf{F}^{-1}(\mathbf{y} - \mathbf{y}_f)\right], \tag{A.16}$$

where $\mathbf{B}$, $\mathbf{O}$ and $\mathbf{F}$ are error covariance matrices for the background, instrument and $\mathbf{x}$ to $\mathbf{y}$ interpolation method, and can be shown to satisfy the following relations

$$\mathbf{B} = \left\langle (\mathbf{x}_b - \mathbf{x}_t)(\mathbf{x}_b - \mathbf{x}_t)^T \right\rangle, \tag{A.17}$$

$$\mathbf{O} = \left\langle (\mathbf{y}_o - \mathbf{y}_t)(\mathbf{y}_o - \mathbf{y}_t)^T \right\rangle, \tag{A.18}$$

$$\mathbf{F} = \left\langle (\mathbf{y}_t - \mathbf{K}(\mathbf{x}_t))(\mathbf{y}_t - \mathbf{K}(\mathbf{x}_t))^T \right\rangle, \tag{A.19}$$

with P_{of} from equation (A.9) is defined as

$$P_{of}(\mathbf{y}_o - \mathbf{y}_f) \propto \exp\left[-\frac{1}{2}(\mathbf{y}_o - \mathbf{y}_f)^T (\mathbf{O} + \mathbf{F})^{-1}(\mathbf{y}_o - \mathbf{y}_f)\right]. \tag{A.20}$$

Now substituting (A.17), (A.18) and (A.19) in (A.14), (A.15) and (A.16), the later group into (A.11) and carrying out either of (A.12) or (A.13), we get that

$$P_a(\mathbf{x}) \propto \exp\left[-\frac{1}{2}\{\mathbf{y}_o - \mathbf{K}(\mathbf{x})\}^T(\mathbf{O}+\mathbf{F})^{-1}\{\mathbf{y}_o - \mathbf{K}(\mathbf{x})\} - \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right] = \text{const.} \tag{A.21}$$

or the penalty functional $\mathbf{J}$ to be maximized

$$\mathbf{J} = \{\mathbf{y}_o - \mathbf{K}(\mathbf{x})\}^T(\mathbf{O}+\mathbf{F})^{-1}\{\mathbf{y}_o - \mathbf{K}(\mathbf{x})\} + (\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b). \tag{A.22}$$

The best fit to the observations can now be derived from this equation. This result includes observational (A.5), model (A.4) and numerical ($\mathbf{K}$) (A.6) errors. Hence

$$\mathbf{x}_a = \mathbf{x}_b + \left\{\mathbf{B}(\partial_{\mathbf{x}}\mathbf{K})^T(\mathbf{O}+\mathbf{F})^{-1}(\partial_{\mathbf{x}}\mathbf{K}) + \mathbf{I}\right\}^{-1}\mathbf{B}(\partial_{\mathbf{x}}\mathbf{K})^T(\mathbf{O}+\mathbf{F})^{-1}\{\mathbf{y}_o - \mathbf{K}(\mathbf{x}_b)\} \tag{A.23}$$

is one form of the best fit and its deviation from the true model state. The error covariance matrix of this result is given by

$$\left\langle(\mathbf{x}_a - \mathbf{x}_t)(\mathbf{x}_a - \mathbf{x}_t)^T\right\rangle = \mathbf{B} - \mathbf{B}(\partial_{\mathbf{x}}\mathbf{K})^T\left((\partial_{\mathbf{x}}\mathbf{K})\mathbf{B}(\partial_{\mathbf{x}}\mathbf{K})^T + \mathbf{O} + \mathbf{F}\right)^{-1}(\partial_{\mathbf{x}}\mathbf{K})\mathbf{B}. \tag{A.24}$$

The crux of the DA scheme is the penalty functional from equation (A.22). It is computationally favourable to use iterative methods to find the solution to $\mathbf{x}$ from $\mathbf{J}$ as we will see in the next subsection. Additionally, note that the model state $\mathbf{x}$ can vary from the simple traditional form of a single grid-point field to

vast multivariate 3D-grid-point and spectral fields.

Data Assimilation

As mentioned at the end of the last subsection, the AC DA scheme which built in the MGCM, is an iterative approach derived, approximated and made efficient from the minimization of the penalty functional. The successive correction design comes from Bergthorsson & Döös, (1955), Cressman (1959), Barnes (1973). The $i+1$ iteration is given by

$$\mathbf{x}[i+1] = \mathbf{x}[i] + \mathbf{QW}(\mathbf{y}[i] - \mathbf{K}(\mathbf{x}[i])), \tag{A.25}$$

$$\mathbf{y}[i+1] = \mathbf{y}[i] + (\mathbf{O}+\mathbf{F})^{-1}\mathbf{Q}\{\mathbf{y}[i] - \mathbf{K}(\mathbf{x}[i])\}, \tag{A.26}$$

where $\mathbf{W}$ is a $N_x \times N_y$ dimensional matrix of weights depending on observational errors and the observations to model grid point distance. $\mathbf{Q}$ is a $N_x \times N_x$ dimensional matrix containing factors of normalization. This is derived from (A.23) which can also be written in iterative form as

$$\mathbf{x}[i+1] = \mathbf{x}[i] + \mathbf{Q}(\mathbf{W}\{\mathbf{y}_o - \mathbf{K}(\mathbf{x}[i]) + \mathbf{x}_b - \mathbf{x}[i]\}), \tag{A.27}$$

with $\mathbf{Q}$ and $\mathbf{W}$ replace the following

$$\mathbf{W} = \mathbf{B}(\partial_{\mathbf{x}}\mathbf{K})^T(\mathbf{O}+\mathbf{F})^{-1}, \tag{A.28}$$

$$\mathbf{Q} = (\mathbf{W}(\partial_{\mathbf{x}}\mathbf{K}) + \mathbf{I})^{-1}. \tag{A.29}$$

The difference between the the modified model state solution used by the DA scheme (A.25) and the actual iterative solution (A.27) is the omission of the $\mathbf{x}_b - \mathbf{x}[i]$ term, which forces the analysis to the background, and the implementation

of error correction (A.26), making it a two-stage process (it also finds the best estimate of the true value $\mathbf{y}_o$). This method (which the AC MGCM also uses) is designated the one dimensional variational analysis or 1DVAR in the literature (see Law et al, 2015 for more details). Next, we will present a brief but holistic view on the functionality of DA in NWP.

Assimilation technique

The derivation from this appendix already exposes some of the means of doing DA. These are coupled with other important parts in the iterations. A summary of the main steps and our progression through them in order to accomplish formal DA includes:

- **The Prior**. This is a necessary step which precedes the assimilation process. The DA scheme gets computationally expensive fast, because it takes into account the full range of atmospheric motion speeds. Including the *prior* information that the atmosphere is usually slowly varying, ensures maximal iteration speed with marginally decreased precision. Our DA scheme already considers this.

- **Quality Control**. Problems evaluating the expected accuracy arise in an iterative analysis system. Observational data checking and approximately Gaussian error distribution, helps in the simplification of the analysis algorithm. The quality control for the MCS data was conducted by Dr L. Montabone.

- **Model+DA**. This is the main DA procedure. As already anticipated from the previous section, we can break each iteration into a sequence of three steps

  - **Analysis**. The observational data is organized and summarized before

constructing the best model state; these are then used in the construction through an analysis process. This technique is commonly known in NWP systems as a forward DA.

- **GCM Initialization**. The GCM runs using the best model state from the previous step as initial conditions.
- **GCM Forecast**. After a timestep preset by the assimilation scheme, the model feeds the forecast to the analysis correction and *waits* for the next set of initial conditions (hence, repeat from first step).

These steps are not necessarily independent. As suggested in figure A.1, most of the DA schemes indeed use the clearly separated sequence depicted in figure A.1c. The AC MGCM actually adopts a continuous repeated insertion of observations as in A.1a. The OI scheme is an infrequent repeated insertion scheme portrayed by figure A.1b. The final DA method A.1d, acknowledges the prior knowledge about the atmosphere differently. In this case, there are two models running simultaneously, one which computes the classic best estimate of the model states, while the other continuously adjusts the prior knowledge from our physical understanding of atmospheric behaviour; in other words, the motion from the second model or the adjoint model, is described by a specific set of equations.

B Gravity Wave Drag Scheme

The MGCM orographic gravity wave drag (GWD) scheme parametrizes the influence of the mesoscale topogrophy (on a lower spatial scale compared to the grid point dimensions at which the MGCM does computations) on the surface layer atmospheric fields. For completeness, the GWD scheme includes both wave blocking at the base of obstacles as well as the gravity wave production over the

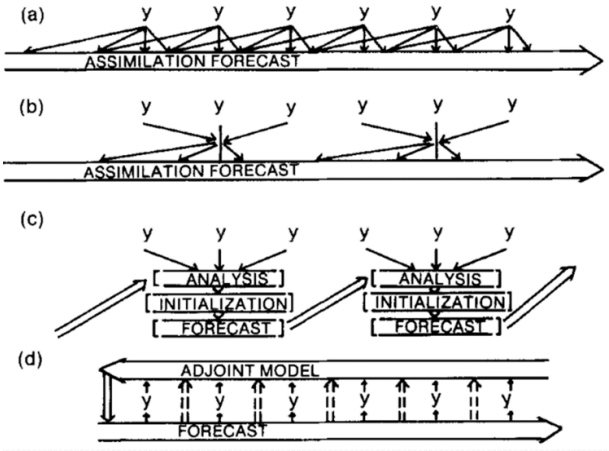


Figure A.1: Different data assimilation strategies. In each diagram, time proceeds from left to right, observations are denoted by y and the arrows have the following meaning: "$\downarrow$" information insertion as values, "$\Downarrow$" information field flow. Each DA scheme has the following characteristics: (a) forward model by continuous repeated insertion (same as our AC DA scheme used in the MGCM), (b) forward model by periodic repeated insertion (same as the older OI scheme from the MO), (c) forward model by periodic independent analysis-initialization-forecast sequences (used in most DA schemes), (d) DA by construction of the adjoint of forecast model; the dashed arrows denote the information passed from the model to solve nonlinearities. (Lorenc et al, 1991)

top of these features. Thus, the parametrization is based on two papers which tackle these two distinct cases; these are Palmer et al (1986) and Lott asnd Miller (1997). The two physical situations can be understood from figure B.1: if the incoming flow is at a low enough altitude ($< Z_b$), it will get divided by the hypothetical obstacle into two streams which "hug" it; in contrast, the higher altitude part ($> Z_b$) of the flow will be *forced* over the obstacle and produces gravity waves from the vertical motion of the fluid.

The parameter which distinguishes between the two situations is the non-dimensional mountain height (Lott and Miller, 1997)

$$H_n = \frac{\bar{N} H}{\bar{U}}, \tag{B.1}$$

where the maximum obstacle height is H, the Brunt-Väisälä frequency is $\bar{N}$ and the wind speed is $\bar{U}$. The border between the two situations is depicted by the closest streamline which goes over the mountain. If $\bar{U}$ and $\bar{N}$ are height independent, it can be shown that the layer depth, Z_b is given by

$$Z_b = H \max\left(0, \frac{H_n - H_{nc}}{H_n}\right), \tag{B.2}$$

where H_{nc} is of order unity and represents the critical non-dimensional mountain height. The derivation of the wave splitting is considerably simpler compared to the other case because it is based on the Rayleigh drag formula in the presence of an obstacle. In contrast, the GWD derivation is considerably more convoluted and requires a more sophisticated approach. The next two sections cover the phenomenon in more detail.

B.0.1 Gravity wave production theory

The theory for the gravity wave production by "over the top" flow (figure B.1) of the mountain flow drag part of the scheme is illustrated in this section; its

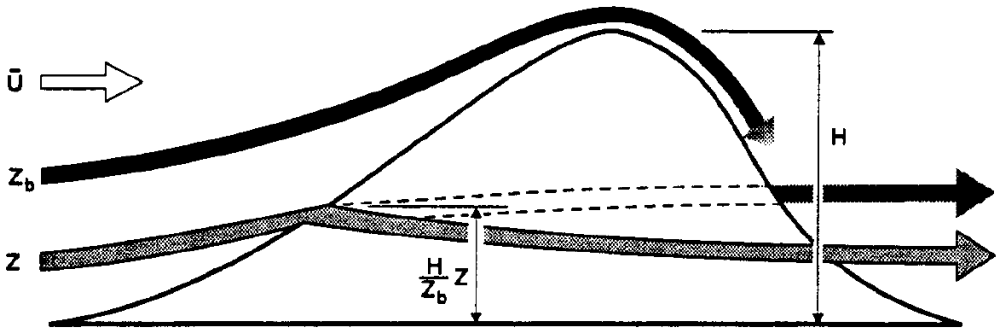


Figure B.1: Diagram depicting the two situations parametrized by the GWD scheme, for the case of a linear flow over a hypothetical obstacle. H is the height of the obstacle, $\bar{U}$ the velocity of the incoming flow, Z is the layer *depth* defined also in the figure, and Z_b is the boudnary between the two physical situations. Flow breaking and forced flow "over the top", occur for depths $< Z_b$ and $> Z_b$ respectively. (Lott and Miller, 1997)

derivation is almost entirely from Phillips (1984). The physical ensemble includes an inviscid, dry and steady idealized flow over a mountain top. The equations considered are the horizontal momentum, hydrostatic, continuity and adiabatic equations which are linearized about the general flow trend (from the global circulation of the atmosphere), which in turn is assumed to be an isothermal upstream of constant speed $\bar{U}$ in the $+x$ direction (see Phillips, 1984 for the story behind the choice of physical assumptions). These equations are

$$\bar{\rho}\bar{U}u'_x + p'_x = 0, \tag{B.3}$$

$$\bar{\rho}\bar{U}v'_x + p'_y = 0, \tag{B.4}$$

$$g\rho' + p'_z = 0, \tag{B.5}$$

$$\bar{\rho}(u'_x + v'_y + w'_z) + \delta_1\bar{U}\rho'_x + \delta_2\bar{\rho}_z w' = 0, \tag{B.6}$$

$$\delta_3\bar{U}p'_x + \bar{p}_z w' - \bar{c}^2(\bar{U}\rho'_x + \bar{\rho}_z w') = 0, \tag{B.7}$$

where u', v' and w' are the velocity perturbations to the incoming U-flow due

the mountain top region; these are in the x, y and z directions respectively. ρ' and p' are the density and pressure perturbations from the upstream $\bar{\rho}$ and $\bar{p}$. $\bar{c}$ is the linear sound speed of the general flow. δ_i where $i = 1, 2, 3, 4$ will be shortly assigned either 0 or 1 depending on further approximations. Finally, the subscripts x, y and z everywhere in this subchapter denote derivatives with respect to these dimensions. The equations for the background U-flow considered by Phillips are

$$\bar{\rho} = -\frac{\bar{p}_z}{g} = \bar{\rho}_0 e^{-z\delta_4 \bar{H}_0^{-1}}, \tag{B.8}$$

$$\bar{c}^2 = \gamma_0 g \bar{H}_0 = \gamma_0 R \bar{T}, \tag{B.9}$$

where $\bar{T}$ is the temperature of the upstream flow, $\bar{\rho}_0$ is the surface density, $\bar{H}_0$ density scale height, γ_0 the specific heat ratio (gamma factor) and R is the dry air gas constant; these are all constant. The approximations to the fluid equations for small displacements at the mountainous region are justifiable for small amplitude waves (equations (B.3)-(B.7)). The point of the GWD scheme is to parametrize orographic gravity waves at smaller scales than the MGCM can compute. Hence, we are only interested in deviations from the large *global* winds, making the linearization a good starting strategy for our situation. Furthermore, the solutions to these equations are subject to the constraint that the flow does not break or overturn and thus follows the mountain profile. Continuing in the same fashion, consider η to be the vertical displacement of the general flow far upstream and h the terrain displacement from the elevation of the surrounding area. Hence, the linearized boundary condition is

$$\eta(z = 0) = h. \tag{B.10}$$

As noted, η is calculated at $z = 0$ instead of $z = h$. Hence considering

$$w' = \frac{D\eta}{Dt} = D_t\eta \approx \bar{U}\eta_x, \tag{B.11}$$

the adiabatic equation (B.7) can be expressed in terms of η

$$\delta_3 \frac{g}{\bar{c}^2} p' + p'_z = -\bar{\rho}\bar{N}^2\eta, \tag{B.12}$$

where $\bar{N}$ is the Brunt-Väisälä (buoyancy) frequency

$$\bar{N}^2 = -g\left(\frac{g}{\bar{c}^2} + \frac{\bar{\rho}_z}{\bar{\rho}}\right) = \left(1 - \frac{1}{\gamma_0}\right)\frac{g}{\bar{H}_0} \tag{B.13}$$

calculated for $\delta_4 = 1$. After many lines of algebra, one can combine the coupled perturbation equations to obtain a single differential equation in η

$$\eta_{xxxx} + \left(\frac{\delta_1 - \delta_4(1+\delta_2)}{\bar{H}_0} + (\delta_3 - \delta_1)\frac{g}{\bar{c}^2}\right)\eta_{xxz} + \frac{\delta_4}{\bar{H}_0}\left(\frac{\delta_2\delta_4 - \delta_1}{\bar{H}_0} + (\delta_1 - \delta_2\delta_3)\frac{g}{\bar{c}^2}\right)\eta_{xx} + \frac{\bar{N}^2}{\bar{U}^2}(\eta_{xx} + \eta_{yy}) = 0. \tag{B.14}$$

This is a good place to start choosing the δ_i coefficients. This equation (B.14) is essentially the same for $\delta_i = 1, \forall i$ in its domain with the simultaneous conditions $\delta_1 = \delta_2 = 0$ (continuity equation neglect), $\delta_3 = 0$ (pressure fluctuation neglect) (Gossard and Hooke, 1975) and $\delta_1 = 1$. It is also shown in Phillips (1984) that the maximum pressure variation with these conditions compared to the assumption of $\delta_i = 0, \forall i \in (1, 2, 3, 4)$, changes by at most 5-10%, apart from when calculating the Brunt-Väisälä frequency; hence making it a good approximation. By Fourier transforming (B.14), Queney (1948) has solved the two dimensional case while Wurtele (1957) and Smith (1980) have solved the three dimensional situation. Employing the same notation from Philips (1984), the Fourier transform of the height $h(x, y)$ is

$$\hat{h}(k,l) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)e^{-i(kx+ly)}dxdy, \tag{B.15}$$

while the displacement η is defined as

$$\eta(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{h}(k,l)e^{i(kx+ly+mz)}dkdl. \tag{B.16}$$

Furthermore, the pressure is obtain from substituting (B.16) into (B.12) and carrying out the integral over z to obtain

$$p' = i\bar{\rho}_0\bar{N}^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\hat{h}(k,l)}{m(k,l)}e^{i(kx+ly)}dkdl. \tag{B.17}$$

These are good solutions for (B.14) with the boundary condition (B.10) provided that we have the relation

$$m = \frac{\bar{N}}{\bar{U}}\frac{\sqrt{k^2+l^2}}{k} \tag{B.18}$$

where k, l and m are wavenumbers in the horizontal and vertical. The final form of the pressure and height perturbations are

$$p'(x,y) = i\bar{\rho}_0\bar{N}^2\int_{-\infty}^{\infty}\frac{\tilde{h}(k,y)}{m(k)}e^{ikx}dk \tag{B.19}$$

and

$$\tilde{h} = \frac{1}{2\pi}\int_{-\infty}^{\infty} h(x,y)e^{-ikx}dx \tag{B.20}$$

The pressure perturbation (B.19) is the solution to equations (B.3)-(B.7) subject to the constraints presented before. It describes the pressure change at horizontal orographic slices (x-y) of any three dimensional shape. The most physically accurate choice for the shape that the scientific community has found solutions to, is the height dependent elliptical cross section shape given in polar coordinates

by

$$h(x, y) = h_0 H(R), \tag{B.21}$$

where

$$R = \sqrt{X^2 + Y^2}, \quad X = x/a = R\cos\theta, \quad Y = y/b = R\sin\theta \tag{B.22}$$

and H is given by

$$H(R) = (1 + R^2)^{-\mu}. \tag{B.23}$$

Here x and y satisfy the equation for an ellipse represented by the first formula of (B.22); R, X, Y and H are dimensionless and θ is the polar angle. x and y are the real distances, fixing a and b to be scaling lengths in the horizontal. The aspect ratio $\gamma = a/b$ controls the eccentricity ϵ in the following way

$$\epsilon = \begin{cases} (1 - \gamma^2)^{1/2} & \text{if} \quad \gamma \leq 1, \\ \left(1 - \dfrac{1}{\gamma^2}\right)^{1/2} & \text{if} \quad \gamma \geq 1 \end{cases} \tag{B.24}$$

μ is a parameter specifying the reduction of the mountain circumference with height as well as the mountain geometry. A systematic description of the different shapes produced by the μ and γ combination can be found in Phillips (1984). Various values for μ were used by the scientific community: Crapper (1959) did studies for $\mu = 1/2$ and circular cross section $\gamma = 1$; Clark & Gall (1982) utilized $\mu = 1/2$; Smith (1980) used $\mu = 3/2$.

Further transformation to the height from equation (B.15) are imperative to reach the final drag. These are coordinate transformations to real space polar (R, θ) and spectral transformations to wavenumber space polar as well (ρ, ϕ), described by

$$\rho = \sqrt{\kappa^2 + \lambda^2}, \quad \kappa = ka = \rho\cos\phi, \quad \lambda = lb = \rho\sin\phi. \tag{B.25}$$

This gives the normalized Fourier transform for the height $\hat{H}(\rho)$

$$\hat{H}(\rho) = \frac{\hat{h}(k,l)}{h_0 ab} = \frac{1}{2\pi}\int_0^\infty RH(R)J_0(\rho R)dR, \tag{B.26}$$

where the Bessel function appears after solving the integral over θ (Gradshteyn & Ryzhik, 1965).

The flow drag

The expected drag force is written in a two-component basis. The basis is rotated such that the drag component D' is employed in the U-flow direction and the application of the transversal component T' is done perpendicular to the U-flow. The drag is only due to momentum transfer with the mountain and does not involve any other type of forcing (like friction). Hence, invoking the surface slope $\nabla h = (\partial_x h, \partial_y h)$, and the pressure equation

$$\mathbf{F}' = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p'(x,y)\nabla h dxdy \tag{B.27}$$

one can use (B.17) and (B.18) and obtain

$$\mathbf{F}' = 4\pi^2\bar{\rho}_0\bar{U}\bar{N}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{k\mathbf{k}}{\sqrt{k^2+l^2}}|\hat{h}(k,l)|^2 dkdl, \tag{B.28}$$

where $\mathbf{k} = (k,l)$. The final step consists of applying this equation to a certain shape (B.21) and transforming to polar coordinates (B.25)

$$\mathbf{F} = (D,T) = \frac{\mathbf{F}'}{\bar{\rho}_0\bar{U}\bar{N}h_0^2 b}, \tag{B.29}$$

where

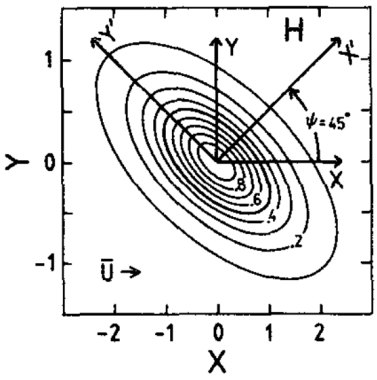


Figure B.2: Contour plot of the obstacle elliptical shape with height, with the definition of the two bases involved in the GWD derivation: the flow basis (X, Y) and the obstacle basis (X', Y'). The obstacle basis is rotated by the angle ψ to the flow direction. Note: ψ is not constrained to 45°. (Philips, 1984)

$$\begin{cases} D = G(B\cos^2\psi + C\sin^2\psi) \\ T = G(B - C)\sin\psi\cos\psi \end{cases}, \tag{B.30}$$

with the G coefficient given by

$$G = 16\pi^2 \int_0^\infty \rho^2[\hat{H}(\rho)]^2 d\rho. \tag{B.31}$$

The factors B and C are dependent on the cross sectional shape (as is G) and can be shown (Phillips, 1984) to have the form

$$B(\gamma) = \begin{cases} \dfrac{E(\epsilon) - \gamma^2 K(\epsilon)}{\epsilon^2}, & \gamma < 1 \\ \dfrac{K(\epsilon) - E(\epsilon)}{\gamma\epsilon^2}, & \gamma > 1 \end{cases} \quad \text{and} \quad C(\gamma) = \gamma B(\frac{1}{\gamma}), \tag{B.32}$$

where K and E are complete elliptic integrals which depend on the eccentricity (B.24). The angle ψ with the other geometrical parameters can be understood from figure B.2. Essentially, the large scale wind $\bar{U}$ defines the X-axis direction, while the ellipse basis (X', Y') is rotated by the angle ψ to the flow.

There have been a multitude of studies which tackle the parametrization of

the orographic drag force (B.29). The ones which are related to Phillips' work include Lyra (1943) and Queney (1947) and (1948) which also considered a larger scale incoming flow but over a two-dimensional, infinitely long mountainous ridge. Smith (1980) used the same conditions (treated by Phillips as well) and worked out the solutions to the three dimensional case but for circular cross sections. As expected, his solutions can be recovered by imposing the aspect ratio of $\gamma = 1$ and shape parameter $\mu = 1/2$ in the drag formula (B.29). For a comprehensive survey of the extensive literature, see Queney et al (1960), Smith (1979), Klemp and Lilly (1980).

B.0.2 Flow blocking theory

The theory for the orographic flow blocking part of the scheme is illustrated in this section; it is based almost entirely on Lott and Miller (1997). The derivation is more transparent because it involves manipulating the Rayleigh obstacle drag formula

$$\mathbf{F}_b(z) \approx -\bar{\rho} C_d l(z) \frac{\bar{U}\bar{\mathbf{U}}}{2}, \tag{B.33}$$

where the drag vector $\mathbf{F}$ is parallel to the velocity vector $\mathbf{U}$, $l(z)$ is the height-dependent obstacle width as *seen* by the incoming flow and C_d is the drag coefficient; this varies depending upon the object's geometry, friction and drag. Equation (B.33) is a modified form of Rayleigh's drag formula for each z-layer. The drag formula involves the projected area and not the width.

The hypothetical ridge shape is again elliptical, with a and b being the parameters in the cross-obstacle direction and along the obstacle direction, the aspect ratio is defined as

$$\gamma = a/b \leq 1 \tag{B.34}$$

and the height is equated to

$$h(x,y) = \frac{H}{1 + x^2/a^2 + y^2/b^2}, \tag{B.35}$$

where H is the maximum height (figure B.1). The approximate eccentricity at each layer is employed from Lott and Miller (1997) as

$$\epsilon'(a',b') \approx \epsilon(a,b)\left(\frac{Z_b - z}{z + \mu'}\right)^{1/2} \tag{B.36}$$

where μ' is the elevation above the mountain valleys at level $z = 0$ (the mean orography of the model, see Lott and Miller, 1997). The width projection can now be approximated to

$$l(z) \approx 2\max(b\cos\psi, a\sin\psi)\left(\frac{Z_b - z}{z + \mu'}\right)^{1/2}, \tag{B.37}$$

where ψ retains a similar meaning from last section (i.e. the angle between the normal ridge direction and the U-flow direction). From here we generalize the ridge to encompass a uniform distribution of subgrid scale obstacles. First consider the case of a normal flow ($\psi = 0$) over a square ridge of length scale L. In the normal direction, the fluid faces $L/2a$ obstacles, while along the ridge it goes over $L/2b$ obstacles. The combined influence changes $l(z)$ to

$$l(z) = L\left(\frac{Z_b - z}{z + \mu'}\right)^{1/2}. \tag{B.38}$$

Furthermore, if we rotate the ridge by the obstacle angle ψ as a whole, the flow will encounter roughly $L/2b$ or $L/2a$ obstacles, depending if it progresses parallel or normal to the ridge. Hence, the new ridge width becomes

$$l(z) = \frac{L^2}{2}\left(\frac{Z_b - z}{z + \mu'}\right)^{1/2}\max\left(\frac{\cos\psi}{a}, \frac{\sin\psi}{b}\right). \tag{B.39}$$

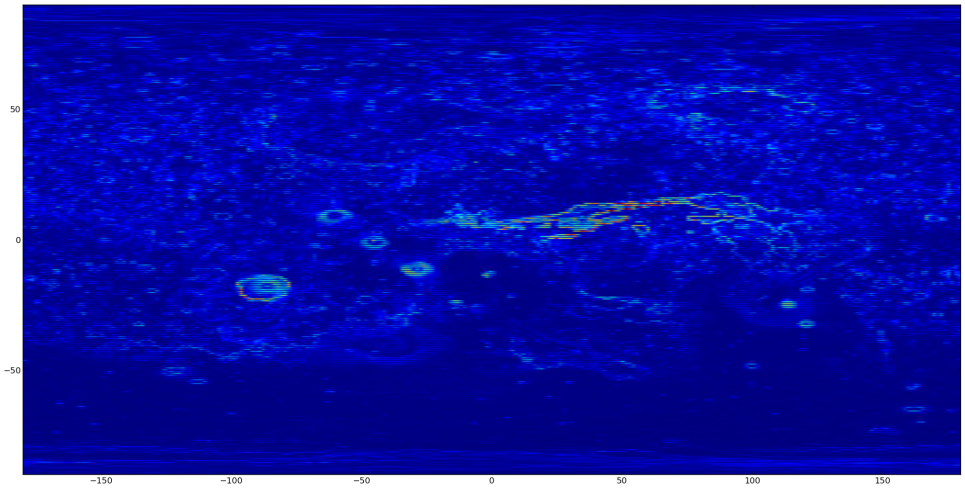


Figure B.3: Sub-grid scale standard deviation taken from the GWD scheme at T170 resolution. This parameter is one of the 4 main variables used to describe the low grid-scale topography.

Now we change the ellipse parameters into the subgrid scale parameters $a \approx \mu/\sigma$ and $a/b = \gamma$, and modulate the drag coefficient C_d in equation (B.33) to account for the different angles ψ encountered by the flow, by using the parameter

$$r = \frac{\cos^2\psi + \gamma \sin^2\psi}{\gamma \cos^2\psi + \sin^2\psi}. \tag{B.40}$$

The drag formula (B.33) becomes

$$\mathbf{F}_b(z) = -C_d \max\left(2 - \frac{1}{r}, 0\right) \bar{\rho} \frac{\sigma}{2\mu'} \left(\frac{Z_b - z}{z + \mu'}\right)^{1/2} \max(\cos\psi, \gamma \sin\psi) \frac{\bar{U}\bar{\mathbf{U}}}{2}, \tag{B.41}$$

where σ accounts for the obstacle slope. The factor $\max(\cos\psi, \gamma\sin\psi)$ can be shown to be similar to the D coefficient in equation (B.30). This modification greatly helps the computational side of the scheme because we can use the D factor in calculating both the GWD and flow-breaking parts of the scheme. Hence, using

the notation from Phillips (1984) the drag has the final form

$$\mathbf{F}_b(z) = -C_d \max\left(2 - \frac{1}{r}, 0\right) \rho \frac{\sigma}{2\mu'} \left(\frac{Z_b - z}{z + \mu'}\right)^{1/2} (B\cos^2\psi + C\sin^2\psi)\frac{\bar{U}\bar{\mathbf{U}}}{2}. \tag{B.42}$$

where B and C are related to elliptical integrals as in (B.32).

B.0.3 The full scheme

The wave-blocking drag from equation (B.42) is already in the form the scheme uses. The GWD equation from (B.29) is also modified to include successive obstacles within a ridge. Adding the contributions together as we did in the last subsection, we for obtain for the "over the top" flow drag, the following formula

$$\mathbf{F}_{gwd} = \bar{\rho}\bar{U}\bar{N}\mu'\sigma G\{B\cos^2\psi + C\sin^2\psi, (B - C)\sin\psi\cos\psi\} \tag{B.43}$$

where H and b in Phillips' drag formula (B.29) have been replaced with $2\mu'$ and μ'/σ. The B and C coefficients involve elliptical integrals (B.32) and hence, for computational reasons, are approximated to $B = 1 - 0.18\gamma - 0.04\gamma^2$ and $C = 0.48\gamma + 0.3\gamma^2$ in the scheme. The parameters μ', γ, σ and ψ (the standard deviation, anisotropy, slope and orientation respectively) are calculated from the full Mars Orbiter Laser Altimeter (MOLA) data ($1 \times 2\text{km}^2$ grid-point size or $1/64^\circ \times 1/32^\circ$). An account for the standard deviation μ' passed in the model by the T170 assimilating MGCM, is given in figure B.3. The MGCM also uses the MOLA data in modeling the atmosphere, but interpolated and smoothed to lower resolutions. This confirms the assumptions in equations (B.3)-(B.7) and the requirement for working in first order.

C Further details on the MGCM

This section starts with giving more details on the theoretical side of the MGCM, and finalises with a description of how the variables are stored in the restart system. This is relevant for the kinetic energy spectrum calculation directly from spectral coefficients from section 3.1.3, but is also an informative addition to 2.1.1 from the "Tools and methodologies" chapter.

The essence of spectral space, comes with the liberty of working with spectral coefficients instead of differential operators. Hence, after expanding ζ, D, T', ϕ and $\ln p_s$

$$\{\zeta, D, T', \phi, \ln p_s\}(\lambda, \mu, t) = \mathcal{N}_x \sum_{n=0}^{n_{max}} \sum_{m=-n}^{n} \left\{\zeta_n^m, D_n^m, T'^m_n, \phi_n^m, (\ln p_s)_n^m\right\}(t) Y_n^m(\lambda, \mu), \tag{C.1}$$

and in a similar way, the streamfunction and velocity potential

$$\{\psi, \chi\}(\lambda, \mu, t) = \mathcal{N}_x \sum_{n=0}^{n_{max}} \sum_{m=-n}^{n} \{\psi_n^m, \chi_n^m\}(t) Y_n^m(\lambda, \mu), \tag{C.2}$$

one can use the connection relations (2.6) - (2.9) to also relate the expansion coefficients in the same fashion. Thus, from (2.6)-(2.10) the following relations can be obtained:

$$U_n^m = (n-1)\epsilon_n^m \psi_{n-1}^m - (n+2)\epsilon_{n+1}^m \psi_{n+1}^m + im\chi_n^m, \tag{C.3}$$

$$V_n^m = -(n-1)\epsilon_n^m \chi_{n-1}^m + (n+2)\epsilon_{n+1}^m \chi_{n+1}^m + im\psi_n^m, \tag{C.4}$$

$$\zeta_n^m = -n(n+1)\psi_n^m, \tag{C.5}$$

$$D_n^m = -n(n+1)\chi_n^m, \tag{C.6}$$

where $\epsilon_n^m = \sqrt{\frac{n^2-m^2}{4n^2-1}}$. The sums can and are rewritten as: $\sum_{n=0}^{n_{max}} \sum_{m=-n}^{n} \equiv \sum_{m=-n_{max}}^{n_{max}} \sum_{n=|m|}^{n_{max}}$. Following these steps, the coefficients from (C.1) can also be re-

lated, with the exception of the non-linear terms in the EOMs.

We can treat the non-linear terms on the right hand side of the equations of motion in the same manner:

$$\mathcal{F}_U = V\zeta - \dot{\sigma}\frac{\partial U}{\partial \sigma} - T'\frac{\partial \ln p_s}{\partial \lambda} = \sum_{m=-n_{max}}^{n_{max}} A_m e^{im\lambda} \tag{C.7}$$

$$\mathcal{F}_V = -U\zeta - \dot{\sigma}\frac{\partial V}{\partial \sigma} - T'(1-\mu^2)\frac{\partial \ln p_s}{\partial \mu} = \sum_{m=-n_{max}}^{n_{max}} B_m e^{im\lambda} \tag{C.8}$$

where A_m and B_m are the truncated Fourier series coefficients. To obtain the Fourier coefficients for any of the series (C.1), (C.2), (C.7), (C.8) at the latitude μ we simply project onto the zonal Fourier harmonics:

$$X^m(\mu) = \frac{1}{2\pi}\int_0^{2\pi} \mathcal{X}(\mu,\lambda)e^{-im\lambda}d\lambda. \tag{C.9}$$

To obtain the actual spectral coefficients for equations (C.1) and (C.2) an additional projection is required, this time onto the space of associated Legendre polynomials:

$$X_n^m = \frac{1}{2}\int_{-1}^{1} d\mu X^m(\mu)\tilde{P}_n^m(\mu). \tag{C.10}$$

Up to this point we can use the standard mathematics presented above to mutually convert between our physical parameters and EOMs in real space and their equivalent in Fourier and Spectral space. Additionally, we can use the spectral form of the connection formulas (2.6)-(2.9) to obtain relations between the corresponding coefficients (C.3)-(C.6).

The procedure of the transform method for the inviscid, adiabatic, hydrostatic, ideal gas equations can be summarized as follows:

1. Start by assuming that $\zeta_m^n(t_i)$, $D_m^n(t_i)$, $T'^n_m(t_i)$, $\phi_m^n(t_i)$ and $(\ln p_s)_m^n(t_i)$ are

given (from a previous timestep t_i).

2. Calculate $\psi_m^n(t_i)$ and $\chi_m^n(t_i)$ and hence all the linear terms in the spectral version of (2.1)-(2.5) (see A1) using the connections (C.3)-(C.6).

3. Calculate $\zeta_m(\mu_j)$, $D_m(\mu_j)$, ${T'}_m(\mu_j)$, $\phi_m(\mu_j)$ and $(\ln p_s)_m(\mu_j)$ using the Legendre transform of the latitude grid points (i.e. expansions (C.7) and (C.8)).

4. Calculate $\zeta(\lambda_l, \mu_j)$, $D(\lambda_l, \mu_j)$, $T'(\lambda_l, \mu_j)$, $\phi(\lambda_l, \mu_j)$ and $(\ln p_s)(\lambda_l, \mu_j)$ using the Fourier transform of the longitude grid points (reach the form of expansions (C.1) in the end).

5. Calculate the coefficients of the non-linear terms $A_m(\mu_j)$ and $B_m(\mu_j)$ using (C.7) and (C.8) with the outcome from the previous points. We have to use the real grid functions because of the nature of (C.7) and (C.8).

6. Use the Gauss-Legendre quadrature method to integrate over the Gaussian latitudes μ_n, and obtain ζ_m^n, D_m^n, ${T'}_m^n$, ϕ_m^n and $(\ln p_s)_m^n$ for time t_{i+1}. Repeat.

The real grid resolution is dictated by the anti-aliasing condition: we need waves with a larger scale than the resolvable scale in order to represent them accurately and unambiguously. If we fail to do this, such waves which are generated by nonlinear interactions will be falsely represented as waves with a larger scale than the resolution. The spectral model is free from aliasing if a number of sufficient grid points are used. The sufficient condition for aliasing-free grid points is to take

$$N_{lon} \geq 3n_{max} + 1 \quad \text{and} \quad N_{lat} \geq (3n_{max} + 1)/2 \tag{C.11}$$

where n_{max} is the largest Fourier wavenumber as we have seen before, N_{lon} is the number of longitudes and N_{lat} is the number of latitudes. As such, the lon×lat grid for nonlinear terms in a T31 simulation has 96×48 grid points ($48 = \frac{3\times31+1}{2} + 1$).

The linear terms can have a coarser sampling of 72×36 grid points – which is arguably a more correct representation of the equivalent resolution in grid space.

In the vertical, the coordinate system of Phillips (1957) is employed, in which σ-coordinates are used. The model's layers are unevenly spaced, the numbers getting richer the closer you get to the surface in order to better represent the interactions and boundary layer processes between the surface and the atmosphere. The last layer in the vertical extends to $\sim$ 100km above the surface, and the last three layers work as 'sponge layers' in order to reduce unphysical reflections of vertical waves from the upper boundary. The total number of vertical layers used in our preliminary studies is 25.

Ultimately, the model computations, over any chosen time interval, end up outputting netCDF files, containing real grid point atmospheric parameters. Essentially, the model uses the spectral space variables or a combination of them (depending on the outputted parameters), and transforms them in real space. The linking information between the run-time steps are passed in the form of machine-specific binary files (which are memory efficient, but are nontrivial to decode, as the files themselves provide no clue as to the format) containing an the variables described in equations (C.1). Normally, the total number of coefficients for triangular truncation at a certain σ-level and time-step is expected to be $(n_{max} + 1) \times (n_{max} + 2)/2$, which for a T31 resolution is 528. However, the MGCM stores only $n_{max}^2/2$ or 512 for T31. This technical feature will prove important for the energy spectra calculations (section 3), and appears from the triangular truncation approach of the MGCM. The coefficients are manipulated and stored in vectors constructed from packages of coefficients of the same n-index and symmetry (global or hemispherical), in the following way

$$A_0^0, A_2^0, A_4^0, ..., A_1^1, A_3^1, ..., ..., A_{n_{max}-1}^{n_{max}-1} \quad , \qquad \text{- hemispherical package} \tag{C.12}$$

$$A_1^0, A_3^0, A_5^0..., A_2^1, A_4^1, ..., ..., A_{n_{max}}^{n_{max}-1} \quad , \qquad \text{- global package} \tag{C.13}$$

where A is an arbitrary spectral coefficient from the MGCM. Notice that the odd coefficients (global) stop at $n_{max}-1$ such that the packages have the same number of coefficients. This improves efficiency and order in the computations, but makes the use of these variables more difficult, because the upper limit in the coefficient-"triangle" (from the triangular truncation) is actually jagged instead of straight (for normal T-truncation). These were decoded and used in subsection 3.1.3.

D The theory of flows and fluctuations

This section is an informative addition to the derivation from the spectral energetics chapter, namely the one from subsection 3.1. Mathematically, this section should precede subsection 3.1.

The derivations presented in this appendix are an homage to the book Satoh (2014), for the fundamental nature of deriving the backbone equations from atmospheric physics: the continuity equation, Navier-Stokes and the energy balance equations (as well as entropy, angular momentum, vorticity equations). These can all be derived from just flow and fluctuation arguments over different physical parameters (as presented in Satoh, 2014). The next step is deriving the spectral energetics equations from the energy flow equations (in connection to Satoh, 2014). Additionally, the Navier-Stokes equations were also written in the most general form for rotating frames.

Consider a generic physical quantity A^* dispersed through space with a distribution given by the quantity per unit volume $A \equiv dA^*/dV$ (A can even be

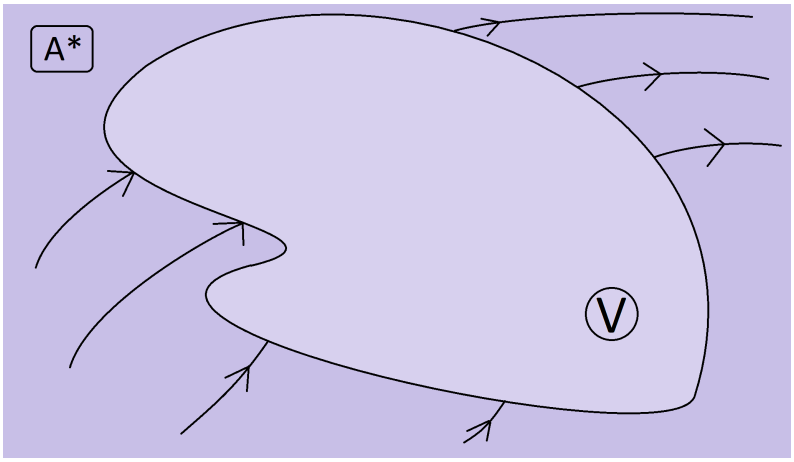


Figure D.1: The flow of A^* through the selected closed surface separating the volume V. The change of A^* within V must balance the flow of A^* passing through V and the sinks or sources of A^* inside V

spatially discontinuous). For an arbitrary choice of closed surface within our domain, we can assert the flow and fluctuations of quantity A^* through that volume. The amount of A^* must be conserved, hence we expect the change in A^* within the closed surface to balance the net flux of A^* through the surface and any source-sink process inside. The formal relationship for A^* is

$$\frac{\partial}{\partial t}\int_V A\,dV = -\oint_{\partial V}\mathbf{F}\cdot d\mathbf{S} + \int_V \sigma[A]\,dV, \tag{D.1}$$

where the change in A^* over time is on the LHS (integrated density over volume), the crossing flux through the surface and the production or depletion of A^* are on the RHS. Using the divergence theorem, equation (D.1) can be written only in terms of volume integrals

$$\int_V dV\,\partial_t A = -\int_V dV\,\nabla\cdot\mathbf{F} + \int_V dV\,\sigma[A] \tag{D.2}$$

which has to hold for any V. Thus, in differential form, the conservation equation gets the following format

$$\partial_t A + \nabla \cdot \mathbf{F} = \sigma[A]. \tag{D.3}$$

A more useful procedure is to define the quantity A^* per unit mass as $a \equiv dA^*/dm$ and use the mass density ρ to connect to $A = \rho a$. By taking advection out of $\mathbf{F}$ and redefining to $\mathbf{F} \to A\mathbf{v} + \mathbf{F}$ where $\mathbf{F}$ is now the rest of the flux (i.e. diffusion), the traditional conservation equation is reached

$$\partial_t(\rho a) + \nabla \cdot (\rho a \mathbf{v}) + \nabla \cdot \mathbf{F} = \sigma[\rho a]. \tag{D.4}$$

We can transport many quantities and generate EOMs from here, as long as we have good definitions for $\mathbf{F}$, $\sigma[...]$ and a.

Mass transport. Let $a = 1$ hence $A = \rho$, $\mathbf{F} = 0$ (only mass) and $\sigma = 0$ (no sources/sinks), then

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \quad \text{or} \quad d_t \rho + \rho \nabla \cdot \mathbf{v} = 0 \quad \text{where} \quad d_t \equiv \partial_t + \mathbf{v} \cdot \nabla\,, \tag{D.5}$$

where d_t is the material derivative. This equation is obviously the *continuity equation*.

Momentum transport. Let $A_i = \rho v_i$ and $a_i = v_i$, thus substituting in $F_{ij} = \Pi_{ij} = (\rho v_i)v_j - \Sigma_{ij}$, where Π_{ij} is the momentum flux density tensor and Σ_{ij} is the stress tensor which alters at each point in space (i) the momentum field vector (j). The pressure tensor can be extracted from the stress tensor as follows $\Sigma_{ij} = -p\delta_{ij} + \Sigma'_{ij}$, leaving behind the viscous stress tensor Σ'_{ij} which has the formula

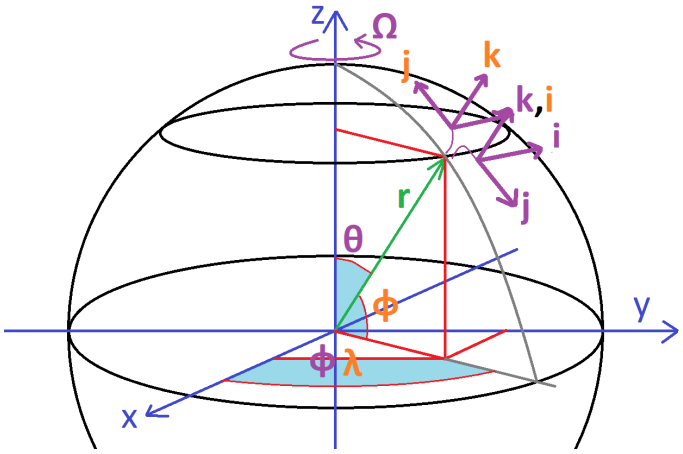


Figure D.2: Comparison between the spherical coordinates (r, ω, ϕ) and (λ, ϕ, r). The purple (i, j, k) basis is the mirror of the orange basis against the zonal axis. This discrepancy comes from the complementary latitude angles purple-θ and orange-ϕ.

$$\Sigma'_{ij} = \eta\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\nabla\cdot\mathbf{v}\right) + \zeta\delta_{ij}\nabla\cdot\mathbf{v} \quad \text{where} \quad \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}. \tag{D.6}$$

The δ_{ij} is the unit tensor or Kronecker's delta while η and ζ are coefficients of viscosity (pressure and temperature dependant). The source term of momentum is gravity, hence $\sigma[\rho v_i] = -\rho g_i$. Putting everything together in the conservation equation (D.3)

$$\partial_t(\rho v_i) + \partial_i \Pi_{ij} = \sigma[\rho v_i], \tag{D.7}$$

which after the use of the continuity equation (D.5), it winds up to

$$\partial_t v_i + (v_j\partial_j)v_i = -\frac{1}{\rho}\,\partial_i p + \frac{1}{\rho}\,\partial_j \Sigma'_{ij} - \partial_i \Phi, \tag{D.8}$$

$$\text{where} \quad \frac{1}{\rho}\,\partial_j\Sigma'_{ij} = \frac{\eta}{\rho}\nabla^2 v_i + \frac{1}{\rho}\left(\frac{\eta}{3} + \zeta\right)\partial_i(\nabla\cdot\mathbf{v}) \quad \text{and} \quad \Phi = gz. \tag{D.9}$$

The Navier-Stokes equation is conservation of momentum; Φ is the gravita-

tional potential energy. The formal atmospheric momentum equations emerge once we set them up in a rotating frame of angular velocity Ω and change to spherical coordinates (λ, ϕ, z). After more meticulous calculus

$$(\partial_t + \frac{u}{a\cos\phi}\partial_\lambda + \frac{v}{a}\partial_\phi + w\partial_z)\begin{bmatrix} u \\ v \\ w \end{bmatrix} + \begin{bmatrix} -\frac{uv\tan\phi}{r} + \frac{uw}{r} \\ \frac{u^2\tan\phi}{r} + \frac{vw}{r} \\ -\frac{u^2+v^2}{r} \end{bmatrix} + 2\mathbf{\Omega}\wedge\mathbf{v} = -\frac{1}{\rho}\nabla p - \nabla(\Phi - \frac{1}{2}\Omega^2 X^2) + \frac{1}{\rho}\nabla\cdot\mathbf{\Sigma}' \quad \text{(D.10)}$$

where (u, v, w) is the velocity vector, $\mathbf{X} = \mathbf{r} - (\mathbf{r}\cdot\mathbf{\Omega})\mathbf{\Omega}/\|\mathbf{\Omega}\|^2$ and $-\Omega^2X^2/2$ is the centrifugal potential energy. These are the basis for our simplified hydrostatic EOMs for the MGCM, (2.1)-(2.5) and will play a role in deriving the transport of energy.

Energy transport. Define the total energy per unit volume as

$$\rho e^{tot} = \frac{1}{2}\rho v^2 + \rho\Phi + \rho u, \quad \text{(D.11)}$$

where the energy components are: $\rho v^2/2$ the kinetic energy, $\rho\Phi$ the potential energy and ρu the internal energy; u is the internal energy per unit mass. Then let $A = \rho e^{\text{tot}}$, $\mathbf{F} = \mathbf{F}^{\text{tot}}$ and $\sigma[\rho e^{\text{tot}}] = 0$, because total energy must be conserved. The flux equation (D.3) yields

$$\partial_t(\rho e^{tot}) + \nabla\cdot\mathbf{F}^{\text{tot}} = 0. \quad \text{(D.12)}$$

where $\mathbf{F}^{tot}$ is the flux density vector field of total energy. The full expression of the flux requires the balance of the individual components. The transport of kinetic energy can be derived from the momentum equations (D.8) as follows

$$\partial_t\left(\frac{\rho v^2}{2}\right)+\partial_j\left\{\left(\frac{\rho v^2}{2}+p\right)v_j-\Sigma'_{ij}v_i\right\}=p\partial_i v_i-\epsilon-\rho v_i\partial_i\Phi, \tag{D.13}$$

$$\text{where}\quad \epsilon\equiv\Sigma'_{ij}\partial_i v_i=\frac{1}{2}\eta\left(\partial_j v_i+\partial_i v_j-\frac{2}{3}\delta_{ij}(\partial_k v_k)\right)^2+\zeta(\partial_k v_k)^2 \tag{D.14}$$

is the dissipation rate of kinetic energy, which is expectedly formulated from the viscous stress tensor (D.6). The potential energy transport is less cumbersome to derive and comes from imposing static gravity $\partial_t\Phi=0$, hence

$$\partial_t(\rho\Phi)+\nabla\cdot(\rho\Phi\mathbf{v})=\rho\mathbf{v}\cdot\nabla\Phi\equiv\sigma\left[\rho\Phi\right], \tag{D.15}$$

where the source term $\sigma[\rho\Phi]$ is evidently the work by gravity. The internal energy transport is the final component of the total energy conservation equation (D.11) and thus the source term-$\sigma[\rho u]$ can be extracted from imposing no source of total energy (which stems from conservation) as follows

$$\sigma[\rho e^{tot}]\equiv\sigma\left[\frac{\rho v^2}{2}\right]+\sigma[\rho\Phi]+\sigma[\rho u]=0,\quad\text{thus}\quad\sigma[\rho u]=-p\nabla\cdot\mathbf{v}+\epsilon, \tag{D.16}$$

$$\therefore\partial_t(\rho u)+\nabla\cdot(\rho u\mathbf{v}+\mathbf{F}^{\text{heat}})=-p\nabla\cdot\mathbf{v}+\epsilon, \tag{D.17}$$

where $\mathbf{F}^{\text{heat}}$ is the heat flux, which stands for the flux density vector of internal energy leaving out the advection of internal energy $\rho u\mathbf{v}$. For dry air, the flux is equivalent to the sum of the radiative and heat conduction fluxes $\rightarrow$ $\mathbf{F}^{\text{heat}}=\mathbf{F}^{\text{rad}}+\mathbf{F}^{\text{therm}}$. If the design of the planetary circulation includes a hydrological cycle, then the transport of internal energy must include a contribution from water vapour transport. On these assumptions, the heat flux is further divided into the latent heat flux from water vapour and the sensible heat flux due to thermal diffusion $\rightarrow$ $\mathbf{F}^{\text{therm}}=\mathbf{F}^{\text{lh}}+\mathbf{F}^{\text{sh}}$. Putting all together

$$\mathbf{F}^{\mathrm{heat}} = \mathbf{F}^{\mathrm{rad}} + \mathbf{F}^{\mathrm{lh}} + \mathbf{F}^{\mathrm{sh}}. \tag{D.18}$$

Summa summarum, transporting certain fields through the conservation equation (D.3) and rewriting the resulting differential equations on a spherical shell, the full EOMs for all the chapters of this thesis can be recovered. To obtain the particular forms from the MGCM, MMM or the spectral energetics equations, the use of prior knowledge to approximate will suffice. This discussion essentially continues in 3.1.