Low-order dynamical behavior in the martian atmosphere: Diagnosis of general circulation model results

Article

Accepted Version

Low-order dynamical behaviour in the Martian atmosphere: Diagnosis of general circulation model results

Oscar Martínez-Alvarado
E: O.MartinezAlvarado@reading.ac.uk, T: +44 (0)118 378 7995, F: +44 (0)118 378 8905
Mathematical Institute, University of Oxford
24-29 St Giles, Oxford, OX1 3LB, UK
Now at: Department of Meteorology, University of Reading
Earley Gate, PO Box 243, Reading, RG6 6BB, UK

Irene M. Moroz
E: moroz@maths.ox.ac.uk, T: +44 (0)1865 270514, F: +44 (0)1865 273583
Mathematical Institute, University of Oxford
24-29 St Giles, Oxford, OX1 3LB, UK

Peter L. Read
E: p.read1@physics.ox.ac.uk, T: +44 (0)1865 272082, F: +44 (0)1865 272923
Clarendon Laboratory, University of Oxford
Parks Road, Oxford, OX1 3PU, UK

Stephen R. Lewis
E: S.R.Lewis@open.ac.uk,, T: +44 (0)1908 655803, F: +44 (0)1908 654192

Luca Montabone
E: montabone@atm.ox.ac.uk, T: +44 (0)1865 272902, F: +44 (0)1865 272923
Department of Physics and Astronomy, The Open University
Walton Hall, Milton Keynes, MK7 6AA, UK

March 10, 2009

71 pages, 22 figures, 3 tables

Proposed running head:

Low-order dynamics in the Martian atmosphere

Corresponding author:

Oscar Martinez-Alvarado

Current address:

Department of Meteorology

University of Reading

Earley Gate, PO Box 243

Reading, RG6 6BB

United Kingdom

E-mail: O.MartinezAlvarado@reading.ac.uk

Telephone: +44 (0)118 378 7995

Fax: +44 (0)118 378 8905

Abstract

The hypothesis of a low dimensional Martian climate attractor is investigated by the application of the proper orthogonal decomposition (POD) to a simulation of Martian atmospheric circulation using the UK Mars general circulation model (UK-MGCM). In this article we focus on a time series of the interval between autumn and winter in the northern hemisphere, when baroclinic activity is intense. The POD is a statistical technique that allows the attribution of total energy (TE) to particular structures embedded in the UK-MGCM time-evolving circulation. These structures are called empirical orthogonal functions (EOFs). Ordering the EOFs according to their associated energy content, we were able to determine the necessary number to account for a chosen amount of atmospheric TE. We shown that for Mars a large fraction of TE is explained by just a few EOFs (with 90% TE in 20 EOFs), which apparently support the initial hypothesis. We also show that the resulting EOFs represent classical types of atmospheric motion, such as thermal tides and transient waves. Thus, POD is shown to be an efficient method for the identification of different classes of atmospheric modes. It also provides insight into the nonlinear interaction of these modes.

Keywords: Mars; atmospheres, dynamics; data reduction techniques.

1 Introduction

A number of observational studies and numerical simulations have suggested that the dynamical behaviour of the Martian atmosphere may be more regular and coherent than its terrestrial counterpart. For example, the analysis of meteorological information sent back to the Earth by Viking Lander 2 during two consecutive Martian years showed that transient baroclinic wave activity in the northern hemisphere was characterised by rather similar parameters, suggesting similar dynamics, during spring and autumn in both years (Barnes, 1980, Barnes, 1981). This regularity in the behaviour of transient waves was also found in simulations carried out using the NASA Ames General Circulation Model (GCM) (Barnes *et al.*, 1993).

An early series of numerical experiments with the UK Mars GCM (UKMGCM) suggested that, without the influence of the diurnal cycle, the Martian atmosphere would display a behaviour characterised by coherent baroclinic waves during late autumn, winter and early spring with periods of approximately 2.2 or 5.5 sols (1 sol = 1 Martian solar day). When the diurnal cycle was included, the baroclinic waves no longer displayed a regular behaviour. Instead, they evolved in a more irregular way which was, nevertheless, influenced to a large extent by episodic oscillations with periods of 2.6 and 5.6 sols (Collins *et al.*, 1996). Both numerical experiments produced the same zonal structure for the baroclinic waves. Thus, the high-frequency baroclinic waves displayed a horizontal structure with zonal wavenumber two in both simulations while the low-frequency waves had a zonal wavenumber-1 wave structure (Collins *et al.*, 1996). In this manner thermal tides and

baroclinic waves appear to be spatially resonant, which would suggest that complex non-linear interactions between these different kind of motions are to be expected (Collins *et al.*, 1996). In fact these results can be interpreted as a complex dynamical pattern controlled by two coexisting equilibria. The system would tend to stay near one of the equilibria until it was perturbed by the diurnal cycle which would then cause the system to switch erratically from one equilibrium state to another (Read and Lewis, 2004).

These findings constitute rather intriguing features of the Martian atmosphere considering the high complexity that would be expected in such a system. The enhanced regularity observed in the dynamical behaviour of the atmosphere of Mars in comparison with its terrestrial counterpart suggests the hypothesis of a relatively low-dimensional underlying attractor of the Martian climate. In turn, this hypothesis points towards the possibility of analysing the Martian atmosphere in the context of low-order diagnostic and, in principle, prognostic models.

Low-order models based on proper orthogonal decomposition (POD) have been under development for the terrestrial case for some time (e.g. Schubert, 1985). However, work on low-dimensional models in the context of planetary atmospheres is still sparse despite the potential significance that this kind of research could have for the study of extraterrestrial atmospheres and comparative planetology.

One example of these studies is given by the work of Whitehouse *et al.* (2005a, 2005b). They analysed the possibility of representing an atmospheric dataset by a small number of modes (Whitehouse *et al.*, 2005a). Their atmospheric dataset was produced by a simplified GCM (SGCM) for a planet

with no topography (smooth surface with no variation in thermal properties) and simple physical parametrizations such as Newtonian cooling and Rayleigh friction. The parameters in the model were tuned as for resembling the behaviour of the atmosphere of Mars as described by Collins and James (1995). Whitehouse *et al.* (2005a) carried out a two-stage decomposition on the quasi-geostrophic (QG) horizontal streamfunction. First, they decomposed this field in a series of vertical modes based on the QG vertical structure equation. After this first decomposition, they compiled a partially reduced dataset by choosing those vertical modes that explained most of the energy content in the original dataset. The reduced dataset was subject to further reduction by means of POD, a statistical technique for extracting coherent structures from experimental or simulated data according to a suitably defined generalised energy (see Holmes *et al.*, 1996, and Section 3.1 below for a more complete description of POD). The complete procedure allowed the representation of $\mathcal{O}(10^3)$ ensemble members by a set of $\mathcal{O}(50)$ EOFs. Dynamically, this effectively amounts to the reduction from $\mathcal{O}(10^4)$ degrees of freedom in the original model down to $\mathcal{O}(50)$ degrees of freedom in the fully reduced dataset (Whitehouse *et al.* (2005a).

In a second article, Whitehouse *et al.* (2005b) introduced the construction of dynamical reduced-dimension models for the SGCM based on the decomposition described above. They managed to construct successful models with 80 degrees of freedom in a combination of vertical and horizontal modes derived during the diagnostic analysis previously described.

Closely following the programme established by Whitehouse *et al.* (2005a, 2005b), this and a forthcoming article explore the hypothesis of a relatively

low-dimensional attractor underlying the Martian climate by searching for low-order models capable of representing the essential dynamics of the Martian atmosphere. In particular, this article presents a diagnostic analysis of simulated data by the UK-MGCM. In contrast with the SGCM used by Whitehouse *et al.* (2005a, 2005b), the UK-MGCM is a state-of-the-art model comprising complex parametrizations for a number of relevant physical processes such as radiative heat transfer, surface processes, sub-grid dynamics and polar carbon dioxide condensation and sublimation. It also includes a realistic representation of the Martian topography as measured by the Mars Orbiter Laser Altimeter (MOLA) aboard the Mars Global Surveyor (MGS) spacecraft (Smith *et al.*, 1999).

The approach of the present study is focused on the use of the primitive equations of dynamic meteorology as underlying dynamics rather than the QG theory approach of previous studies (Whitehouse *et al.*, 2005a, 2005b). This also positions our study closer to the latest work in the terrestrial context (Achatz and Opsteegh, 2003). The UK-MGCM and the climatology of the UK-MGCM simulation to be analysed will be described in Section 2. As a consequence of this approach, the analysis was carried out by means of a fully three dimensional version of POD, which involved the decomposition of three fields (two horizontal velocity components and a thermodynamic field) dependent on three spatial dimensions and time. The full description of this method is given in Section 3.1. In Section 4, the energy distribution as extracted by POD is presented. Section 5 investigates the relation between POD-modes, or empirical orthogonal functions (EOFs), and components of atmospheric motions such as thermal tides and baroclinic waves, that have

been the subject of extensive studies. By establishing this relationship we seek to shed light on the appropriate physical interpretation of the EOFs for this problem. Furthermore, we intend to show that thermal tides and transient waves are intimately related not only by sharing similar energy contents but also by being part of common coherent structures, which confirms and extend the findings by Collins *et al.* (1996). Finally, a summary and conclusions are given in Section 6.

2 The UK Mars GCM

The UK-MGCM is a pseudo-spectral model jointly developed at the University of Oxford and the Open University in collaboration with the Laboratoire de Météorologie Dynamique (LMD) in Paris. It is based on the terrestrial model of Hoskins and Simmons (1975). The resolution used in the simulation was jagged T31 (triangular with maximum total wavenumber 31) with 25 unevenly-spaced σ-levels in the vertical. The uneven spacing in the vertical allows for an enhanced resolution near the surface (Forget *et al.*, 1999). The uppermost full level is located at $\sigma = 5.6 \times 10^{-5}$, approximately corresponding to 100 km assuming a constant scale height $H = 10$ km. This vertical extension allows the model to capture the fully developed Martian Hadley cell, which can reach heights of around 80 km (Wilson, 1997).

The large-scale topographic features on Mars play an essential role in the evolution of the general circulation. For example, coupled with the diurnal tide, it is well-known to excite waves that resemble non-migrating tides (Zurek, 1976). In particular, an eastward-propagating diurnal Kelvin wave

and an westward-propagating wavenumber-3 diurnal wave arise from this interaction. The diurnal Kelvin wave seems to have an influence on processes such as the development of dust storms (Montabone *et al.*, 2007). Therefore, it is important to have an accurate representation of the Martian orography. The UK-MGCM uses the highly accurate Mars topography dataset as measured by the Mars Orbiter Laser Altimeter (MOLA) aboard NASA's Mars Global Surveyor spacecraft (MGS) (Smith *et al.*, 1999).

The physical processes in the model include radiative heat transfer, surface processes, sub-grid dynamics and carbon dioxide mass exchange between the polar caps and the atmosphere (Forget *et al.*, 1999; Read and Lewis, 2004, and references therein). The later is largely responsible for the significant surface pressure variations along the Martian year. This study, however, is focused on just one season for which the global mean surface pressure varies only 0.03%. This amount can be considered small if compared with the 30% peak-to-peak variation encountered for the full year. Under these conditions we have neglected surface pressure variations in this study. Nevertheless, we recognize that including these variations should be a priority for future studies.

2.1 UK-MGCM dataset climatology

In order to reduce the amount of information to be handled by POD, thereby reducing the computational expense, only 10 unevenly spaced sigma-levels were selected from the sigma-levels in the UK-MGCM. Table 1 shows the sigma-levels in the UK-MGCM. The levels that were retained are enclosed

in a box. The lower levels were chosen so that the first scale height ($H_0 = 10$ km, on Mars) was well represented. Thus, the first four levels lay within the first 10 km and the fifth is just above. The uppermost levels in the UK-MGCM were discarded to avoid including the sponge layer. The horizontal resolution was set to a regular longitude-latitude grid with $I = 64$ grid points in the zonal direction and $J = 32$ grid points in the meridional direction for vorticity, potential vorticity and streamfunction.

The UK-MGCM time series over the chosen sigma-levels was pre-processed using the procedure outlined in the Appendix. The pre-processing produced time series of horizontal velocities, square root of potential temperature and an ensemble-averaged surface pressure to be used as time-independent surface pressure during the analysed period.

The UK-MGCM dataset the we used in this study covers a 90-sol interval from $L_s = 267.7\,^\circ$ (sol 511) to $L_s = 323.2\,^\circ$ (sol 601). Here L_s is the areocentric longitude and gives the position of the planet in its orbit around the sun. $L_s = 0, 90, 180, 270\,^\circ$ correspond to the northern spring equinox, summer solstice, autumn equinox and winter solstice, respectively. Thus, this dataset corresponds to the transition between autumn and winter in the northern hemisphere. The simulation was carried out with a dust scenario which prescribes the total dust optical depth at each latitude and time of the year. This scenario (MGS dust scenario) has been derived from MGS observations and produces temperature profiles in the model which are close to those observed by the Thermal Emission Spectrometer aboard MGS (Smith *et al.*, 2001). The time series was sampled every two hours to make an ensemble of 1080 elements.

2.1.1 Zonal mean

Figure 1 shows the temporal and zonal average of temperature and zonal wind in the UK-MGCM for the period under consideration. Figure 1 also shows the meridional mass transport streamfunction χ defined, in sigma-coordinates, as (e.g., Gill, 1982) **Figure 1**

$$\overline{p_s v} = -\frac{1}{\cos\phi}\frac{\partial \chi}{\partial \sigma}, \quad \overline{p_s \dot{\sigma}} = \frac{1}{\cos\phi}\frac{\partial \chi}{\partial \phi}, \tag{1}$$

where $\overline{(\cdot)}$ indicates zonal average so that positive streamfunction indicates counter-clockwise circulation.

Figure 1 presents a zonal mean atmospheric structure with characteristics typical of winter solstice. A steep horizontal temperature gradient near the surface in the winter hemisphere gives rise to a very strong winter westerly jet in the middle atmosphere above the northern hemisphere, reaching peak speeds of 140 m s^{-1} on average. This jet is accompanied by a weaker easterly jet above the equator. The temperature cross-section (Fig. 1) exhibits the so-called winter polar warming that takes place in the middle atmosphere between 50 and 80 km (Read and Lewis, 2004). This phenomenon is due to adiabatic compression in the downwards side of the Hadley cell, which on Mars extends up to 80 km above the surface (Wilson, 1997, Forget *et al.*, 1999). Notice also the strong hemispheric asymmetry of the Martian solstice general circulation. This is characterized by a strong quasi-global Hadley cell that extends from almost 60 ° in the summer hemisphere to the pole in the opposite hemisphere (Forget *et al.*, 1999).

2.1.2 Atmospheric tides

There are two important features of the UK-MGCM which have an impact on the global circulation: the presence of diurnal and seasonal cycles and the inclusion of topography. For this study, only the diurnal cycle and topography are relevant since the seasonal cycle is not being considered due to the assumption of time-independent surface pressure and the short duration of the time series being analysed.

The presence of the diurnal cycle gives rise to atmospheric thermal tides. These are atmospheric motions whose characteristic periods are harmonics of 1 sol. They can be classified as sun-synchronous, or migrating, and longitude-dependent, or non-migrating, tides. The most significant migrating tides are the diurnal and semidiurnal tides (see, e.g. Wilson and Hamilton, 1996).

The diurnal tide is a wave travelling westwards, characterised by a zonal wavenumber one and a period of 1 sol. The semidiurnal tide is also a westward travelling wave, though it is characterised by a zonal wavenumber two and a period of half a sol. The diurnal tide is mainly the response of the atmosphere to surface heating because of its short vertical wavelength and surface-trapped structure. The semidiurnal tide, on the other hand, has a much longer vertical wavelength and reflects the vertically integrated radiative response of the atmosphere. Hence, the activity of the semidiurnal tide has a central influence on the atmospheric response to radiatively active phenomena such as dust storms, when the vertical radiative properties are changed due to the enhanced presence of dust loadings (Lewis and Barker, 2005).

On Mars an eastward-propagating diurnal Kelvin wave characterised by a zonal wavenumber one and a resonant period close to 1 sol (e.g., Zurek, 1976, Forbes, 2004, Lewis and Barker, 2005) plays the role of a non-migrating tide. It arises from the interaction between the diurnal tide and topography dominated by a strong zonal wavenumber-2 pattern (Zurek, 1976). Hence, it can only be simulated in a model which includes appropriate topographic information (Forbes, 2004, Read and Lewis, 2004). The interaction between the diurnal tide and topography also gives rise to a westward propagating wavenumber-3 wave with a period of 1 sol.

Figure 2 is a graphical representation on a period-wavenumber space of atmospheric planetary waves present on average in the UK-MGCM dataset as seen in two different fields (a negative period represents a wave travelling westwards). Figure 2(a) corresponds to surface pressure (cf. Fig. 5.9 in Read and Lewis, 2004); Fig. 2(b) presents the same information as seen in the temperature at an approximate height of 600 m above the planetary surface. These figures were obtained by a double Fourier transform (in time and longitude) of a surface pressure (or temperature) time series around the equator (Read and Lewis, 2004, Lewis and Barker, 2005). For the Fourier transform in time we used a rectangular window of 20 sols in order to get a sufficient resolution in frequency. The window was slided along the time axis in steps of 2 Martian hours. The resulting time series of spectral coefficients was then averaged in time. **Figure 2**

The strongest signal in both fields corresponds to the westward-propagating diurnal tide (Period = -1 sol, $m = 1$). The semidiurnal tide (Period = -0.5 sol, $m = 2$) and the diurnal Kelvin wave (Period = +1 sol, $m = 1$) are present

in both fields as well as the westward propagating diurnal wavenumber-3 wave (Period = -1 sol, $m = 3$). However, these waves have different impact on each field. In the surface pressure signal, the diurnal Kelvin wave is the second strongest, followed by the semidiurnal tide. In contrast, in the temperature field at the level of observation, the semidiurnal mode is the second most significant followed by the westward propagating diurnal wavenumber-3 wave.

A similar analysis at mid-latitudes in the northern hemisphere reveals a completely different atmospheric structure at mid-latitudes. Figure 3(a) shows a frequency-wavenumber diagram for atmospheric planetary waves in surface pressure on a longitudinal circle at 64.3 °N; Fig. 3(b) presents similar information for temperature at an approximate height of 600 m above the planetary surface on the same longitudinal circle. At this level the most important signal comes from slow eastward-moving waves with periods ranging from approximately 1.8 to 5.5 sols and zonal wavenumbers between 1 and 3. These waves correspond to baroclinic activity taking place at mid-latitudes in the northern hemisphere similar as found in previous studies using this same MGCM (Collins *et al.*, 1996) and perhaps related to the baroclinic activity first detected in the analysis of Viking 1 and 2 surface pressure data (Barnes, 1980; Barnes, 1981) although the wavenumbers predicted by the UK-MGCM are somewhat smaller than those envisaged by Barnes (1980). As on the equator, an indication of vertical structure can be deducted from the dissimilarities in the waves at the surface and at slightly higher levels in the atmosphere. For example, the maximum power contribution to the surface pressure comes from zonal wavenumber-1 waves, whereas zonal wavenumber-

Figure 3

2 waves contribute the most to the power observed in temperature at 600 m above the surface. Furthermore, the diurnal tide is still perceptible at the surface whereas its signal has vanished when observed at an altitude of only 600 m above the surface.

An important difference (not captured by the figures) between the dynamics at the equator and at mid-latitudes is that there was a constant exchange of power between waves at different frequencies and zonal wavenumbers at mid-latitudes for the time interval under analysis. In contrast, the power distribution at the equator did not vary considerably and the picture remained pretty much the same as that shown in Fig. 2.

3 Analysis methodology

The analysis methodology that was used in this work consisted of a combination of proper orthogonal decomposition (POD) and short-time Fourier analysis. This combination allows us to identify the distribution of the energy (via POD) and relate this to the tidal modes and natural modes that occur in the atmosphere (via Fourier analysis). In this section we review these two methods and explain the way in which we have combined them in order to achieve our objective.

3.1 Proper orthogonal decomposition

Proper orthogonal decomposition is a statistical technique that originally was intended as a compression tool to represent most of the variance in a given dataset using the smallest number of modes (Karhunen, 1946, Loève,

1945). Then, it was found to be useful in the study of turbulence in fluids (Lumley, 1970) and more recently as a tool, in combination with Galerkin projection techniques, to construct low-order dynamical models of very high-dimensional systems whose attractors are believed to be restricted to a lower dimensional space (Holmes *et al.*, 1996). The systems to which the technique has been applied range from fluid turbulence (e.g., Smith *et al.*, 2005) and atmospheric physics (e.g., Achatz and Opsteegh, 2003) to control theory (e.g., Rowley *et al.*, 2001). The resulting POD-Galerkin models have been proved to be successful in at least some cases at capturing the most important features of specific systems using a limited number of POD modes, sometimes called empirical orthogonal functions (EOFs). This article is focused on the analysis of an atmospheric dataset derived from simulations with the UK-MGCM. In a forthcoming paper we will combine the results presented here with the Galerkin projection of the primitive equations to construct low-order dynamical models.

POD can be set to optimize either variance or other forms of generalized energy depending on the definition of a norm to measure the space of states. The correlation or kinetic-energy (KE) norm (used in statistical analysis of data and incompressible homogeneous fluid studies) is one example. Another one, incorporating greater generality, is the total-energy (TE) norm. The description of a stratified and/or compressible fluid requires not only the velocity components but also one or more thermodynamic variables. This is an example where the TE norm is useful since it allows the incorporation of variables of different kinds in its definition. Atmospheric models constitute a second example where the TE norm has been found to be an appropriate

choice (Achatz and Opsteegh, 2003, Whitehouse *et al.*, 2005a). In agreement with this findings, we use atmospheric TE as a norm in this study.

Assuming hydrostatic balance, atmospheric TE for an air column is given by

$$e = \int_{z_0}^{\infty} \rho \left(\frac{1}{2} \mathbf{v} \cdot \mathbf{v} + c_v T + gz \right) dz, \tag{2}$$

where ρ is density, $\mathbf{v}$ is horizontal velocity, T is temperature, z is height and z_0 is topographic height. Also, $c_v = 629$ J kg^{-1} K^{-1} is the heat capacity at constant volume and $g = 3.72$ m s^{-2} is the acceleration due to gravity. The first term in the integrand of Eq. 2 represents kinetic energy, while the second term represents internal energy and the third one represents gravitational potential energy. It can be shown that the last two terms can be written as atmospheric total potential energy (TPE) plus a surface term (e.g. Houghton, 2002)

$$\int_{z_0}^{\infty} \rho(c_v T + gz)\, dz = \int_0^{p_s} \frac{c_p}{g} T\, dp + z_0 p_s, \tag{3}$$

where $c_p = 820$ J kg^{-1} K^{-1} is the heat capacity at constant pressure, p is local pressure and p_s is surface pressure. Restricting our study to just one season and assuming that the variation of p_s can be neglected in a first order approximation, the surface term becomes effectively a time-independent quantity and does not need to be included for our purposes since its only effect would be to offset the energy level.

Since we want to express TE as a norm it is necessary to write it as a quadratic form of the state variables. This can be achieved by introducing the new thermodynamic variable $\tau = \sqrt{R\theta}$ (Achatz and Opsteegh, 2003) where $R = 192$ J kg^{-1} K^{-1} is the gas constant per unit mass and $\theta = T(p_0/p)^{\kappa}$ is

the potential temperature ($p_0 = 500$ Pa is a constant). Moreover, in order to obtain a non-dimensional version of our equations, the planetary radius a and the reciprocal rotation rate Ω^{-1} ($a = 3.3960 \times 10^6$ m, $\Omega = 7.088 \times 10^{-5}$ s^{-1}) were used as length and time scales, respectively. A complete list of non-dimensionalizing factors is given in Table 2. Thus, total atmospheric energy (assuming hydrostatic stability) is given, in dimensionless form, by

$$E = \int_A p_s \left(\frac{\mathbf{v} \cdot \mathbf{v}}{2} + \frac{p^\kappa}{\kappa} \tau^2 \right) dx, \tag{4}$$

where p and p_s are now pressure and surface pressure normalised by p_0 and $\kappa = R/c_p$. Here, x represents the set of areographical coordinates, namely λ and ϕ representing longitude and latitude, respectively, and $\sigma = p/p_s$ acting as a terrain-following vertical coordinate. Hence, $dx = \cos\phi \, d\phi \, d\lambda \, d\sigma$ is a volume element in areographical coordinates. The integration is carried over the whole atmosphere A. The energy scale is $E_0 = (a^2\Omega)^2 p_0/g = 89.82$ EJ (1 EJ $= 10^{18}$ J). The form of the new dimensionless thermodynamic variable is $\tau = \theta^{1/2}$.

Defining the state vector as $\psi = (\mathbf{v}, \tau)^T$, Eq. 4 can be identified with a squared norm

$$E = ||\psi||^2 = (\psi, \psi). \tag{5}$$

This leads directly to the definition of the inner product as

$$(\psi^{(1)}, \psi^{(2)}) = \int_A p_s \left(\frac{1}{2} \mathbf{v}^{(1)} \cdot \mathbf{v}^{(2)} + \frac{p^\kappa}{\kappa} \tau^{(1)} \tau^{(2)} \right) dx, \tag{6}$$

where the superindices (1) and (2) just indicate two different states and the

surface pressure p_s is assumed to be time-independent.

Within this framework the EOFs constitute the solutions to the eigenvalue equation (e.g. Berkooz *et al.*, 1993)

$$\int_A \mathbf{K}(x,x')\mathbf{E}(x')\,\vartheta^{(k)}(x')\,dx' = \Lambda^{(k)}\vartheta^{(k)}(x), \tag{7}$$

where $\Lambda^{(k)}$ and $\vartheta^{(k)}$ play the role of the k-th eigenvalue and eigenfunction, respectively. The energy matrix $\mathbf{E}$ is the weighting function in the definition of the inner product and can be expressed as (cf. Eq. 6)

$$\mathbf{E} = p_s \begin{pmatrix} 1/2 & 0 \\ 0 & p^\kappa/\kappa \end{pmatrix}. \tag{8}$$

The kernel $\mathbf{K}$ is the correlation function defined in terms of the elements of an ensemble of realizations of the system $\mathcal{S} = \{\psi^{(k)}\}_{k=1}^{M}$ containing M elements. Thus, $\mathbf{K}$ is given by

$$\mathbf{K}(x,x') = \langle \psi(x)\,\psi^T(x') \rangle, \tag{9}$$

where $\langle\cdot\rangle$ denotes the average over $\mathcal{S}$. In this study, $\mathcal{S}$ is given by the UK-MGCM dataset described in Section 2.1 and, consequently, the average over the ensemble of realizations is equivalent to a time average over the period under analysis.

The state vector can be approximated by means of the finite sum

$$\psi_N = \sum_{j=1}^{N} a_j \phi^{(j)}. \tag{10}$$

The energy contained by the approximate state vector ψ_N is, on average, (Holmes *et al.*, 1996)

$$\langle E \rangle_N = \sum_{j=1}^{N} \Lambda^{(j)}. \tag{11}$$

The upper index N in the last two expressions indicates the order of the approximation. It is possible to show that a decomposition in terms of EOFs converges more rapidly than any other linear expansion (Berkooz *et al.*, 1993; Holmes *et al.*, 1996)). This energy content is, therefore, the largest possible. The only part that remains to be determined are the coefficients a_j. By taking the inner product of $\vartheta^{(k)}$ and Eq. 10 we obtain

$$a_k = (\vartheta^{(k)}, \psi). \tag{12}$$

These time-dependent coefficients are called the principal components (PCs) of ψ and contain all the information about the evolution of the system.

There are two points that deserve further comment. The first is concerned with the definition of TE in atmospheric dynamics. It can be argued that this definition should be based on available potential energy (APE) rather that TPE since APE is the only fraction of potential energy available to be transformed into kinetic energy (Lorenz, 1955, Lorenz, 1960). However, APE is only well-defined over isentropic or barotropic surfaces (Lorenz, 1955) which do not constitute terrain-following coordinate systems. The use of this kind of coordinate systems is desirable due to the major role that the Martian topography plays in the dynamics of its atmosphere. The choice of TPE, on the other hand, allows the use of the terrain-following sigma-coordinates

without major modifications. Nevertheless, we will see in the following sections that the thermal field splits into two parts. The first one is given by the first and most energetic EOF. This part corresponds to the background potential energy, unavailable to be transformed into kinetic energy. The second part is given by higher order EOFs. This part is comparable with APE in the sense that interactions between these EOFs correspond with exchange of APE and KE.

The second point that deserves further comment is the assumption of the time independence of surface pressure. This assumption is required to give a proper definition to Eq. 6 since a time-dependent inner product would prevent the EOFs from being representative of a global time-independent system. However, this assumption could constitute a very strong constraint in the case of the Martian atmosphere, where the surface pressure can exhibit variations up to 30% peak-to-peak throughout a Martian year (Read and Lewis, 2004). The main reason for this variation is the condensation and sublimation of CO_2 in the polar caps (Hourdin *et al.*, 1993). This process might be taken into account by including the contribution of the ice caps to the internal energy in the expression for TE. However, for the time being this is left as a next step for developing long-term atmospheric models. Instead, in order to avoid a possible inconsistency coming from the assumption of time-independent surface pressure, this study is restricted to the analysis of one particular season, during which p_s remains roughly constant. The considered period is the transition from autumn to winter in the northern hemisphere, since it is during this interval when the transient baroclinic activity is intense. The enhanced intensity in baroclinic activity seems to be closely related to a

reduction in atmospheric predictability in comparison with the period from northern mid-spring to mid-autumn, when the atmosphere appears to be remarkably predictable (Newman *et al.*, 2004).

3.2 Short-time Fourier analysis and tidal extraction

POD and the TE norm are useful for identifying the atmospheric modes that contain most of the total energy in the atmosphere. However, these modes lack a clear physical interpretation. In an attempt to provide this interpretation, we intend to establish a relationship between the EOFs and tidal atmospheric motions. We must recall at this point that POD is not explicitly designed to extract atmospheric tides. Instead, the tidal modes should be expected to appear distributed over more than one EOF.

In order to determine the distribution of tidal modes over EOFs, it is useful to extract the tides (and other wave motions) from the raw UK-MGCM dataset. This is done by means of a new decomposition based on short-time Fourier analysis. The method is an extended version of the method outlined by Lewis and Barker (2005), which consists of a double Fourier transform in longitude and time. Unlike Lewis and Barker (2005), who only analyzed surface pressure along the equator, we performed this analysis along every discrete latitudinal circle and height. Retaining only the frequency n and wavenumber s of interest, an inverse Fourier transform was applied to recover the tides over physical grid points.

This procedure was applied to horizontal velocity $\mathbf{v}$ and temperature T. In order to have the three scalar components in the state vector, the tidal

component of τ was then computed as

$$\tau' = \frac{T'}{2P\Theta^{1/2}}, \tag{13}$$

where T' is the tidal component of temperature and $\Theta^{1/2}$ is the τ-component of the background state, defined as the average projection of the UK-MGCM dataset over the first and most energetic EOF, $\Psi = \langle a_1 \rangle \vartheta^{(1)}$.

Once tides and transient waves have been extracted, their distribution over N EOFs was computed as

$$E_k^{s,n} = \frac{1}{Z} \left\langle (a_k^{s,n})^2 \right\rangle, \tag{14}$$

where $a_k^{s,n}(t) = (\vartheta^{(k)}, \psi_{s,n}(t))$ is the projection of the wave component with frequency n and wavenumber s, denoted as $\psi_{s,n}(t)$, over the k-th EOF at time t and

$$Z = \sum_{k=1}^{N} \left\langle (a_k^{s,n})^2 \right\rangle$$

is a normalization factor where N as the number of retained EOFs.

4 Eigenspectrum

To begin the POD analysis of the UK-MGCM dataset, we look at the empirical eigenspectrum based on total energy (TE). Figure 4 shows the empirical eigenvalues for the autumn-winter transition in the northern hemisphere. As expected, the first mode contains by far the majority of the atmospheric TE on average, the energy content decreasing as the EOF index increases.

The UK-MGCM eigenspectrum splits into several well-defined groups of modes of comparable energy content. Using the diagram in Fig. 4 and Fourier analysis, it is possible to characterise and classify these groups according to their energy and the dominant periods present in their corresponding principal component (PC) time series. There are at least sixteen groups as shown in Table 3. These can include either one, two or three modes. We can refer to the components of these groups as singlets, doublets and triplets, respectively. A continuous tail starts with EOF28.

Figure 4

Table 3

Singlets can be interpreted as non-periodic trends or standing waves. The first and second singlets carry information about the seasonal trend. Since the models that will be constructed do not consider the seasonal cycle, these features are not expected to be resolved by the low-order models. When the seasonal trend is removed, the first and second singlets reveal the influence of the diurnal cycle. The ratios $\mathrm{PC1}_{rms}/\lambda_1^{1/2} = 1.35 \times 10^{-4}$ and $\mathrm{PC2}_{rms}/\lambda_2^{1/2} = 1.23 \times 10^{-1}$, where X_{rms} is the quadratic mean of X and λ_k is the k-th empirical eigenvalue, show the relative importance of the diurnal cycle to each of these modes. This influence is more noticeable in PC2 than in PC1.

Doublets represent travelling waves. There are two special cases which are characterised by a single period. These are the pairs EOF3–4 and EOF9–10, characterised by a diurnal and semidiurnal period, respectively, which seem to indicate a relation between these groups and the diurnal and semidiurnal tides. Other pairs can be classified by their secondary period, which can either be the same for the two components of the pair or different for each member.

Figure 5 shows the normalised cumulative eddy TE for the UK-MGCM

during the period analysed. It was normalised with respect to the average eddy TE in 1080 POD modes. The first apparent characteristic is the rapid growth in the first modes, reaching 80% with only 7 EOFs and 90% in 20 EOFs. This small number of required modes is remarkable if we compare it to the required modes to represent the terrestrial general circulation. **Figure 5**

The number of terrestrial modes vary depending on the dataset that is taken as a basis for the POD analysis. For example, the analysis of observed geopotential height at 700 hPa for winter in the northern hemisphere in a band between 20 °N and 60 °N shows that 17 modes are required for explaining 86% variance and 36 modes explain 96% variance (Kimoto and Ghil, 1993). These numbers are certainly very comparable to the results we are presenting here for Mars. However, our study is closer in nature and method to the work by Achatz and Branstator (1999) and Achatz and Opsteegh (2003). Achatz and Branstator (1999) carried out the simultaneous analysis of global streamfunction at 200 and 700 hPa from the Community Climate Model of the National Center for Atmospheric Research (NCAR) for perpetual January using a TE norm. They found a fairly flat spectrum and showed that as many as 500 EOFs were needed to explain 90% of TE (Achatz and Branstator, 1999). This number is certainly very different even in order of magnitude to that found in the present work. Achatz and Opsteegh (2003) extended the study by Achatz and Branstator (1999) by including the seasonal cyclone and a more complete atmospheric state defined not only by streamfunction but, as in our case, by horizontal velocity and the square root of potential temperature. Interestingly enough the number of EOFs to explain 90% of TE was again about 500 (Achatz and Opsteegh, 2003). This

result was interpreted as an indication of the response of the terrestrial atmosphere being dependent on the same set of EOFs throughout the year. The significance of each of these relatively low-order EOFs would vary depending on the season (Achatz and Opsteegh, 2003).

Despite the different order of magnitude of the required number of EOFs to represent 90% of TE on the Earth and Mars, the analysis of the Martian atmosphere by POD shows similar results as those just described for the Earth in the sense that essentially the same number of modes explain similar amount of total energy for different times of the Martian year. These modes might not be the same as those presented here. Consider for example the transient baroclinic-barotropic activity which is active during winter in each hemisphere. As we shall show in Section 5 some of the EOFs in the present analysis are closely related to this transient wave activity. The analysis of a similar period for winter in the southern hemisphere would yield similar patterns in the opposite hemisphere. Assuming that we actually need 20 EOFs to explain 90% of TE in each season leads us to conjecture that approximately 80 modes would be required throughout the Martian year. This number is still small when compared to the 500 EOFs needed in the terrestrial case.

This result suggests that the region of phase space where the UK-MGCM evolves can be represented by a few modes. This feature is due to the presence of the diurnal cycle which tends to accumulate the energy in modes largely influenced by solar heating, in particular, EOF3–4 as we shall see in the next section.

5 EOFs and atmospheric motions

In this section, we consider some of the leading EOF groups, and seek to identify them with some of the most important components of atmospheric motion. The components to be considered include atmospheric tides and transient waves.

5.1 Diurnal tide

EOF3 and EOF4 constitute the third group, sharing a similar structure and period. The corresponding coefficients also share a similar behaviour, clearly showing the influence of the diurnal cycle. For example, we focus our attention first on EOF3 (Fig. 6). This is essentially a zonal wavenumber-one wave, stronger in the southern hemisphere as expected, since the original dataset corresponds to the transition from spring to summer in that hemisphere (Wilson and Hamilton, 1996). The vertical structure of EOF3 is as expected for a diurnal tide as well (cf. Wilson and Hamilton, 1996, Read and Lewis, 2004). The maximum amplitude is located near the surface with a vertically propagating wavenumber-one wave above the equator. Away from the tropics, the wavenumber-one wave is vertically trapped in the lower atmosphere.

Figure 6

The decomposition of the diurnal tide ($n = -1$ sol^{-1} and $s = 1$) into a projection over the first 40 EOFs using the procedure outlined in Section 3.2 is shown in Fig. 7. As expected given the structural resemblance between EOF3–4 and the diurnal tide, the pair EOF3–4 contains 97.0% of the tidal TE whereas the next most energetic group, EOF18–19, contains only 1.11%.

Figure 7

In conclusion, the pair EOF3–4 is a good approximation to the diurnal tide in the reduced space. Alternatively, we can say that the diurnal tide is the most energetic atmospheric mode after the background state, in the dataset under analysis.

5.2 Semidiurnal tide

The pair EOF9–10 consists of zonal wavenumber-two waves (Fig. 8(a)). Like the semidiurnal tide, which has a vertical wavelength $>$ 100 km (Zurek, 1981, Read and Lewis, 2004), the pair EOF9–10 also has a long vertical wavelength (Fig. 8(b)). However, the semidiurnal tide ($n = -2$ sol^{-1} and $s = 2$) is not quite as well represented by only one EOF pair. Performing the same decomposition analysis as for the diurnal tide case, we find that the pair EOF9–10 indeed explains 92.8% of TE in the semidiurnal tide (Fig. 9). The second most energetic pair is the pair EOF5–6, which contains 2.7% of TE. This is still significant, especially since the pair EOF5–6 are more energetic than EOFs 9–10 when the full atmospheric motion is considered. In other words, the importance of EOF5–6 for representing the semidiurnal tide is hidden by the strong transient component in the same EOF pair.

Figure 8

Figure 9

The amplitude of the pair EOFj–k can be defined as

$$A_{j-k} = \sqrt{a_j^2 + a_k^2}, \tag{15}$$

where a_j is the j-th principal component. Plotting this diagnostic as a function of time can help to visualize the contribution of the waves to a certain motion. Thus, Figure 10 shows that indeed the contribution of EOF5-6 to

the semidiurnal tide is still significant when compared to the contribution from EOF9-10. However, this signature is weak compared with the already reported maxima for EOF5–6 in their PC time series at periods of 3.6 and 3.2 sols (cf. Table 3). This case reveals the nature of the EOFs as combinations of waves that can be further separated by means of appropriate techniques. This topic will be revisited in Section 5.4 where transient baroclinic waves are analysed by projection onto the EOF basis. EOF5–6 also contributes to the diurnal tide's energy. However, in this case this pair is relatively unimportant to represent the diurnal tide as we have already seen in the previous section.

Figure 10

As an additional comment to the decomposition of the diurnal and semidiurnal tide, we must mention, for example, the case of pair EOF9-10 whose components appear to explain different amounts of TE in the diurnal tide decomposition (Fig. 7). EOF9 and EOF10 are conjugated in the sense that they together can represent travelling waves of zonal wavenumber two. However, the difference in TE in the components of this pair indicates that such a wave is not a component of the diurnal tide. Therefore, these EOFs behave independently like high-order perturbations which, in this sense, are unessential for the description of the semidiurnal tide. Similarly, a zonal wavenumber-1 wave with the characteristics given by the pair EOF3–4 is not a component of the semidiurnal tide. From this point of view we can considered that POD did a remarkable job in the separation of the diurnal and semidiurnal components of the Martian thermal tides. This result is robust in the sense that this has been the case in the analysis of other times of the Martian year by these methods.

5.3 Diurnal resonant Kelvin and wavenumber-3 waves

As discussed in Section 2.1.2, the interaction between the diurnal tide and the topography, dominated by a zonal wavenumber-2 pattern, gives rise to different waves among which the diurnal Kelvin and the zonal wavenumber-3 waves are the most efficiently excited. The diurnal resonant Kelvin wave is a zonal wavenumber-1 wave which propagates eastwards with period of 1 sol. The second wave that arises from the interaction of the diurnal tide with topography is a zonal wavenumber-3 mode which propagates westwards also with period of 1 sol.

The decomposition of the diurnal Kelvin wave ($n = 1$ sol^{-1} and $s = 1$) over 40 EOFs is shown in Fig. 11(a). The modes that contain the largest amount of TE is the pair EOF3–4. However, this group only explains 36.0% of TE in this wave. The rest is distributed over various modes without clear localised energy peaks. On the other hand, the diurnal wavenumber-3 wave ($n = -1$ sol^{-1} and $s = 3$) shows clear peaks over the pairs EOF3–4 (representing 34.3% of TE) and EOF18–19 (representing 38.7% of TE), already identified with westward-travelling waves with a dominant period of 1 sol.

Figure 11(a)

5.4 Non-tidal transient waves

Generally, transient waves arise as a consequence of mixed barotropic and baroclinic instabilities and, as for the Earth, they are responsible for much of the energy transport polewards from the equator. These waves are not typically characterised by a definite wavenumber or period. Hence, the separation of

these waves by means of a double Fourier transform in space and time is not as clear as in the case of tidal components. Figure 12 shows an attempt to extract projections of transient waves from the UK-MGCM dataset onto the EOFs by subtracting the tides from the complete fields. From this figure we can see that the most important contribution comes from the pairs EOF5-6 and EOF7–8, representing 57.3% and 15.8% of TE, respectively. EOF11–12, 13–15 form a set of secondary contributory EOFs to the non-tidal transient-wave representation.

Figure 12

EOF5–6, with a leading period of 3.6 sol and a second common period of 3.2 sol, is an example of an EOF pair where both components have the same secondary characteristic period. This group comprises disturbances that resemble baroclinic waves. These waves were expected since the season under analysis corresponds to an active period of baroclinic activity in the northern hemisphere (typically from late autumn to early spring in each hemisphere; Read and Lewis, 2004). Figure 13 shows contours of sigma-velocity and superimposed vectors of horizontal velocity. Recalling that $w \propto -\dot{\sigma}$, it can be seen that air masses moving northwards go upwards whereas air masses moving towards the equator go downwards. This structure is maintained from the surface to an altitude of about 40 km. However, the thermal wave extends to higher altitudes (Fig. 14). As we have seen in Section 5.2, the structure that rises into higher altitudes is in fact part of the semidiurnal wave that appears as a secondary component of EOF5–6 with a period of 0.5 sols.

Figure 13

Figure 14

Figure 15 shows a longitude-latitude map and a longitude-vertical section of EOF7 as representative of the pair EOF7–8. This pair consists of a

Figure 15

wavenumber-1 travelling wave with characteristics of a baroclinic wave located around 70°N. In fact, the structure of this wave is similar to that of the pair EOF3–4. The sloping convective structure is maintained up to an altitude of 30 km whereas the thermal wave reaches heights up to 60 km at 70°N and almost 70 km at around 80 °N.

The amplitude evolution of each of the waves described by the EOF5–6 and EOF7–8, computed using Eq. 15, is shown in Figure 16. Also shown is the contribution of the pair EOF3–4 to the transient motion (i.e. once thermotidal motions have been removed). Unlike the diurnal and semidiurnal tides that are fairly constant (not shown here), these two transient waves exhibit large variations along the period under study. In general EOF5–6 dominates over EOF7–8 apart from the short intervals between the lines A and B and C and D where the amplitude of the latter is slightly larger than that of EOF5–6.

EOFs 11–12 are an example of a pair in which each component has its own secondary frequency. Both exhibit sloping convection near the surface (Fig. 17). However, these two modes have very different structures. The surface wave in EOF11 seems to have a dominant zonal wavenumber two whereas the wave in EOF12 has a dominant zonal wavenumber three extending up to around 40 km. Figure 18 shows these two modes at approximately 11 km and 32 km above the surface. From this point of view, EOF11 appears as a wave with dominant zonal wavenumber one restricted to higher latitudes in the northern hemisphere. In turn, EOF12 appears to have the structure of a zonal wavenumber-3 wave near the surface, although this is lost at around $z \simeq 30$ km, constraining baroclinic wave-like structures to the

Figure 17

Figure 18

lower atmosphere, where horizontal temperature gradients are stronger.

Each component in EOF13-15 has its own secondary period. Figure 19 shows longitude-latitude maps of τ in these three modes at approximately 1.9 km and 32 km above the surface. Near the surface, all of them exhibit a zonal wavenumber-3 wave structure. These waves can be associated with sloping convection and, therefore, these modes can be identified with baroclinic waves. However, around $z \simeq 30$ km this structure is lost in EOF14 and 15 to be replaced by zonal wavenumber-one waves in both modes. EOF13 reaches around 40 km keeping the wavenumber-3 structure.

Figure 19

As we can see from these examples, EOF groups with more than one secondary period have a very complex structure and, rather than interpreting them as individual travelling waves, they must be thought of as mixtures of waves that modulate the general circulation. In fact, this same statement is strictly valid for any other EOF group but, as we have seen before (e.g. Sections 5.1 and 5.2), when all the components in a group are defined by the same periods (leading and secondary), its characterisation becomes simpler.

6 Discussion and conclusions

Through the identification of the most energetic EOFs and examining the atmospheric motions represented by them, we have achieved a clearer physical interpretation of the coherent structures represented by these EOFs. In particular, we have shown that the leading EOFs can be regarded as actual representations of those atmospheric motions known to be primary components of the general circulation on Mars, namely thermal tides and

baroclinic-barotropic transient waves (e.g., Read and Lewis, 2004).

Figure 20 shows together the mean decompositions of the different kinds of motions analysed in Section 5. The solid line represents the eigenvalues obtained from POD. Recall that these represent the average TE represented by each EOF. The relative contribution due to the different types of wave motion can be inferred from this figure as well. The average TE explained by the first two EOFs is, as mentioned before, due to the background state, while the average TE comprised by the rest of the modes is due to either tidal motion or transient waves. The atmospheric tides that have a significant contribution towards TE are the diurnal tide (EOF3–4 and EOF18–19) and the semidiurnal tide (EOF9–10). Nevertheless, transient waves are responsible for a non-negligible amount of TE for every mode, even for those where the diurnal and semidiurnal tides play the leading role. This is a clear indication of a strong relationship between thermal tides and transient baroclinic waves, a conclusion that is supported by previous studies, where the diurnal cycle has been shown to play an important role for the stability of baroclinic waves (Collins *et al.*, 1996).

Figure 20

On the other hand, the diurnal resonant Kelvin wave and the wavenumber-3 wave account for an amount of TE that is approximately two orders of magnitude less than the amount explained by transient waves. This result could suggest that these waves are less significant in the global context from the energy point of view at this time of the year. However, there are studies which suggest that these waves can play an important role for the development of both local and global phenomena close to the autumn equinox (Lewis and Barker, 2005, Montabone *et al.*, in preparation).

Figure 20 provides an alternative tool for deciding what modes should be retained when constructing a low-order model. For instance, it shows that the diurnal pair EOF3–4 explains about the same amount of TE as the transient pair EOF5–6. A dynamical interpretation of these observations can be obtained by looking at the phase portraits described by these modes. Figure 21 shows the phase portrait of the transient component in EOF5–6 and the diurnal component in EOF3–4. This subsystem evolves over a perturbed torus and, thus, shows features of a quasiperiodic system with at least two intrinsic frequencies given by the diurnal cycle and the characteristic frequency of the transient waves. Furthermore, constructing a phase portrait with the amplitudes of the transient component in EOF5–6 and EOF7–8 and the diurnal component of PC-3 we find the characteristic two-centred phase portrait described in earlier works (Collins *et al.*, 1996; Read *et al.*, 1998). This might be a further indication of a relationship between thermal tides and transient baroclinic activity. However, this would need to be confirmed, for example, by constructing a dynamical system including these modes and evaluating their actual interaction.

Figure 21

Figure 22

This work is comparable, though not directly, to the work by Martínez-Alvarado *et al.* (2005). Following the analysis by Whitehouse *et al.* (2005a) on the SGCM, Martínez-Alvarado *et al.* (2005) carried out a decomposition of the same UK-MGCM dataset in vertical and horizontal modes, using the quasi-geostrophic vertical structure equation, respectively. The norm employed in POD was a TE-norm, although it was defined as the sum of KE and APE, rather than total PE, as in the present study. Despite these differences in the method, the results are similar, especially when considering

the structure of the first ten modes comprising thermal tides and baroclinic waves. In Martínez-Alvarado *et al.* (2005), EOF1 and EOF2 were characterised by a diurnal period corresponding to EOF3 and EOF4 in the present study. Furthermore, their EOF3 and EOF4 mainly corresponded to transient waves as EOF5 and EOF6 do in this work. The apparent shift in order is simply due to the fact that in Martínez-Alvarado *et al.* (2005) the fields were separated into zonal (axisymmetric) and eddy parts.

So far we have shown that a dataset of atmospheric variables extracted from the UK-MGCM can be represented to a certain level of accuracy by a few leading EOFs. We have also shown that these leading EOFs represent the waves that have previously been identified as main components of the Martian general circulation. A hypothesis yet to be tested is whether the dynamical interaction of these few EOFs alone is able to produce a realistic evolution resembling the actual dynamics of the Martian atmosphere. We shall explore this hypothesis in a forthcoming paper.

It is appropriate here to point out that the modes are not expected to be invariant with respect to the dataset used in their derivation. Instead, a different set of EOFs is likely to be extracted from a different dataset corresponding not only to different seasons but also to the same period over different conditions, for example, under different dust loadings. The extreme case would be the study of planet-encircling dust storms which occur from time to time on Mars and are know to have an important impact over the thermal tides (Lewis and Barker, 2005). Indeed, the authors have started the analysis of this case finding that again the number of EOFs needed to give an accurate representation of the atmospheric state in terms of energy

is similar to the number presented here. The patterns represented by these EOFs are altered by the sudden load of dust. This, however, is beyond the scope of the present paper.

Acknowledgements

OMA thanks the National Council for Science and Technology of Mexico (CONACYT) for the support provided to this project via the doctoral scholarship 179183.

Appendix

As explained in the text, the UK-MGCM data was transformed into time series of zonal and latitudinal velocity and square root of potential temperature on the 10 sigma-levels used for POD. Each level was discretized in a 64×32 regular longitude-latitude grid, corresponding to squares of angular length $5.625°$ in either direction and equivalent to a maximum arc length of $s = 333$ km. The variables were distributed over a C-grid (Arakawa and Lamb, 1977).

The variation in the global average surface pressure is only 0.03% during the time interval under analysis, justifying the assumption of time-independent surface pressure. This assumption imposes conditions on the velocity components. Under these circumstances, sigma-velocity is

$$\dot{\sigma} = -\frac{1}{p_s} \int_0^{\sigma} \nabla \cdot (p_s \mathbf{v})\, d\sigma_1. \tag{16}$$

Integrating Eq. 16 from the surface to infinity with the boundary conditions $\dot{\sigma} = 0$ at $\sigma = 0$ and $\sigma = 1$, we conclude that

$$\nabla \cdot (p_s [\![\mathbf{v}]\!]) = 0, \tag{17}$$

where $[\![\cdot]\!] = \int_0^1 \cdot \, d\sigma$ so that $[\![\mathbf{v}]\!]$ gives the barotropic velocity.

Since surface pressure is time-dependent in the UK-MGCM, Eq. 17 is not satisfied in general. To make the data consistent with a time-independent surface pressure, a pre-processing of the UK-MGCM time series was required. The procedure is now described.

The streamfunction ψ and velocity potential χ are computed by solving the Poisson equations $\nabla^2 \psi = \zeta - f$ and $\nabla^2 \chi = D$ in spectral space to take advantage of the properties of the spherical harmonics, in particular, of their being eigenfunctions of the Laplacian. In terms of ψ and χ, velocity is given by

$$\mathbf{v} = \nabla \chi + \mathbf{k} \times \nabla \psi. \tag{18}$$

Since we are interested in the product $p_s [\![\mathbf{v}]\!]$, we integrate Eq. 18 from top ($\sigma = 0$) to bottom ($\sigma = 1$) and multiply it by p_s to obtain, after some algebra,

$$p_s [\![\mathbf{v}]\!] = p_s \nabla [\![\chi]\!] - \mathbf{k} \times [\![\psi]\!] \nabla p_s + \mathbf{k} \times \nabla (p_s [\![\psi]\!]). \tag{19}$$

The filtered barotropic velocity can be defined as

$$[\![\mathbf{v}]\!]_f = [\![\mathbf{v}]\!]_o - [\![\mathbf{v}]\!]', \tag{20}$$

where $[\![\mathbf{v}]\!]_o$ is the original barotropic component in the UK-MGCM dataset and $[\![\mathbf{v}]\!]'$ is a residual component. Multiplying Eq. 20 by p_s and taking the divergence we obtain

$$\nabla \cdot (p_s \nabla [\![\chi]\!]_o - \mathbf{k} \times [\![\psi]\!]_o \nabla p_s - p_s [\![\mathbf{v}]\!]') = 0, \tag{21}$$

where we have made use of Eq. 19 applied to the original barotropic velocity and the fact that the filtered barotropic velocity must satisfy Eq 17.

By choosing $[\![\mathbf{v}]\!]' = \nabla [\![\chi]\!]_o - \mathbf{k} \times [\![\psi]\!]_o \nabla(\ln p_s)$, the velocity can be computed using the formula

$$\mathbf{v} = \frac{1}{p_s} \mathbf{k} \times \nabla(p_s [\![\psi]\!]) + \nabla(\chi - [\![\chi]\!]_o) + \mathbf{k} \times \nabla(\psi - [\![\psi]\!]). \tag{22}$$

The surface pressure in Eq. 22 is an ensemble average of the corresponding time series in the original dataset. This average surface pressure is also used in the definition of the matrix $\mathbf{E}$. The first term on the right hand side of Eq. 22 is the non-divergent barotropic velocity. The last two terms together represent the baroclinic velocity.

References

[1] Achatz, U., Branstator, G., 1999. A two-layer model with empirical linear corrections and reduced order for studies of internal climate variability J. Atmos. Sci. 56, 3140–3160.

[2] Achatz, U., Opsteegh, J.D., 2003. Primitive-equation-based low-order

models with seasonal cycle. Part I: Model construction. J. Atmos. Sci. 60, 465–477.

[3] Arakawa, A., Lamb, V.R., 1977. Computational design of the basic dynamical processes of the UCLA General Circulation Model. Methods Comput. Phys. 17, 173–265.

[4] Barnes, J.R., 1980. Time spectral analysis of midlatitude disturbances in the Martian atmosphere. J. Atmos Sci. 37, 2002–2015.

[5] Barnes, J.R., 1981. Midlatitude disturbances in the Martian atmosphere: A second Mars year. J. Atmos Sci., 38, 225–234.

[6] Barnes, J.R., Pollack, J.B., Haberle, R.M., Leovy, C.B., 1993. Mars atmospheric dynamics as simulated by the NASA Ames general circulation model. J. Geophys. Res. 98 (E2), 3125–3148.

[7] Berkooz, G., Holmes, P., Lumley, J.L., 1993. The proper orthogonal decomposition in the analysis of turbulent flows. Annu. Rev. Fluid Mech. 25, 539–575.

[8] Collins, M., James, I.N., 1995. Regular baroclinic transient waves in a simplified global circulation model of the Martian atmosphere. J. Geophys. Res. Planet 100 (7), 14421–14432.

[9] Collins, M., Lewis, S.R., Read, P.L., Hourdin, F., 1996. Baroclinic wave transitions in the Martian atmosphere. Icarus 120, 344–357.

[10] Forbes J.M., 2004. Tides in the middle and upper atmospheres of Mars and Venus. Advances in Space Research 33, 125–131.

[11] Forget, F., Hourdin, F., Fournier, R., Hourdin, C., Talagrand, O., Collins, M., Lewis, S.R., Read, P.L., 1999. Improved general circulation models of the Martian atmosphere from the surface to above 80 km. J. Geophys. Res. Planet 104, 24155–24176.

[12] Gill, A.E. 1982. Atmosphere-ocean dynamics. Academic Press.

[13] Holmes, P., Lumley, J.L., Berkooz, G. 1996. Turbulence, coherent structures, dynamical systems, and symmetry. Cambridge Univ. Press, Cambridge, UK.

[14] Hoskins, B.J., Simmons, A.J., 1975. A multilayer spectral model and the semi-implicit method. Quart. J. R. Met. Soc. 101, 637–655.

[15] Houghton, J.T., 2002. The Physics of Atmospheres. Cambridge Univ. Press, Cambridge, UK.

[16] Hourdin, F., Le Van, P., Forget, F., Talagrand, O., 1993. Meteorological variability and the annual surface pressure cycle on Mars. J. Atmos. Sci. 50, 3625–3640.

[17] Karhunen, K., 1946. Zur spektral Theorie stochastischer Prozesse. Ann. Acad. Sci. Fenicae A1, 34.

[18] Lewis, S.R., Barker, P.R., 2005. Atmospheric tides in a mars general circulation model with data assimilation. Advances in Space Research 36, 2162–2168.

[19] Loève, M. 1945. Functions aleatoire de second ordre. C. R. Acad. Sci. Paris.

[20] Lorenz, E.N., 1955. Available potential energy and the maintenance of the general circulation. Tellus 7, 157–167.

[21] Lorenz, E.N., 1960. Energy and numerical weather prediction. Tellus 12, 364–373.

[22] Lumley, J.L. 1970. *Stochastic tools in turbulence*. Academic Press.

[23] Martínez-Alvarado, O., Moroz, I.M., Read, P.L., Lewis, S.R. 2005. Reduced order models of the Martian atmospheric dynamics. In: Proceedings of the 5th EUROMECH Nonlinear Dynamics Conference. EUROMECH.

[24] Kimoto M. and M. Ghil. 1993. Multiple flow regimes in the northern hemisphere winter. Part I. Methodology and hemispheric regimes. J. Atmos. Sci. 50 (16), 2625–2643.

[25] Montabone, L., Martinez-Alvarado, O., Lewis, S.R., Read, P.L., Smith, M.D. 2007. Meteorology of the 2001 global dust storm on Mars in an assimilation of Thermal Emission Spectrometer data from Mars Global Surveyor. In: Seventh International Conference on Mars, Pasadena, California.

[26] Newman, C.E., Read, P.L., Lewis S.R. 2004. Investigating atmospheric predictability on Mars using breeding vectors in a general-circulation model. Q. J. R. Meteorol. Soc. 130, 2971–2989.

[27] Read, P.L., Collins, M., Früh, W.G., Lewis, S.R., Lovegrove, A.F., 1998. Wave interactions and baroclinic chaos: A paradigm for long timescale

variability in planetary atmospheres. Chaos, Solitons and Fractals 9 (1/2), 231–249.

[28] Read, P.L., Lewis, S.R. 2004. The Martian climate revisited. Springer-Praxis.

[29] Rowley, C.W., Colonius, T., Murray, R.M., 2001. Dynamical models for control of cavity oscillations. AIAA paper 2001–2126.

[30] Schubert, S.D., 1985. A statistical-dynamical study of empirically determined modes of atmospheric variability. J. Atmos. Sci. 42 (1), 3–17.

[31] Smith, D.E., Zuber, M.T., Solomon, S.C., Phillips, R.J., Head, J.W., Garvin, J.B., Banerdt, W.B., Muhleman, D.O., Pettengill, G.H., Neumann, G.A., Lemoine, F.G., Abshire, J.B., Aharonson, O., Brown, C.D., Hauck, S.A., Ivanov, A.B., McGovern, P.J., Zwally, H.J., Duxbury, T.C., 1999. The global topography of Mars and implications for surface evolution. Science 284, 1495–1503.

[32] Smith, M.D., Pearl, J.C., Conrath, B.J. and Christensen, P.R., 2001. Thermal Emission Spectrometer results: Mars atmospheric thermal structure and aerosol distribution. Journal of Geophysical Research-Planets, 106, 23929-23945.

[33] Smith, T.R., Moehlis, J., Holmes, P., 2005. Low-dimensional modelling of turbulence using the proper orthogonal decomposition: a tutorial. Nonlinear Dyn. 41, 275–307.

[34] Whitehouse, S.G., Lewis, S.R., Moroz, I.M., Read, P.L., 2005a. A simplified model of the Martian atmosphere. Part 1: a diagnostic analysis. Nonlinear Proc. Geophys. 12, 603–623.

[35] Whitehouse, S.G., Lewis, S.R., Moroz, I.M., Read, P.L., 2005b. A simplified model of the Martian atmosphere

[36] Wilson, R.J., 1997. A general circulation model simulation of the Martian polar warming. Geophys. Res. Lett. 24, 123–126.

[37] Wilson, R.J., Hamilton, K., 1996. Comprehensive model simulation of thermal tides in the Martian atmosphere. J. Atmos. Sci. 53, 1290–1326.

[38] Zurek, R.W., 1976. Diurnal tide in the martian atmosphere. J. Atmos. Sci. 33, 321–337.

[39] Zurek, R.W., 1981. Inference of dust opacities for the 1977 Martian great dust storms from Viking Lander 1 pressure data. Icarus 48, 202–215.

Table 1: Sigma-levels and equivalent altitudes in the original dataset. The levels that were kept to perform the POD are boxed.

Level	σ	z (km)	Level	σ	z (km)
1	0.000056	97.902	14	0.455263	7.869
2	0.000250	82.940	15	0.600911	5.093
3	0.000678	72.964	16	0.731006	3.133
4	0.001542	64.747	17	0.831173	1.849
5	0.003199	57.449	18	0.899629	1.058
6	0.006290	50.688	19	0.942591	0.591
7	0.011955	44.266	20	0.968097	0.324
8	0.022181	38.085	21	0.982737	0.174
9	0.040323	32.108	22	0.990977	0.091
10	0.071666	26.357	23	0.995562	0.045
11	0.123568	20.910	24	0.998098	0.019
12	0.203945	15.899	25	0.999500	0.005
13	0.316690	11.498			

Table 2: Non-dimensionalizing factors

Dimension	Factor	Value
Length	a	3.3960×10^6 m
Time	Ω^{-1}	1.4108×10^4 s
Velocity	$a\Omega$	240.7 m/s
Temperature	$(a\Omega)^2/R$	303.1 K
Geopotential	$(a\Omega)^2$	5.794×10^4 m^2/s^2
Pressure	p_0	500 Pa
Frictional force	$a\Omega^2$	1.7062×10^{-2} m/s^2
Diabatic heat	$a^2\Omega^3/\kappa$	17.614 J(kg $\cdot$ s)$^{-1}$

Table 3: MGCM leading EOF groups and characteristic periods. When doublets and triplets in a pair or trio have different second periods, these appear in parentheses. The corresponding frequencies were computed with a resolution of 0.0111 sol^{-1}.

Group	Type	EOFs	Leading period (sol)	Secondary period (sol)
1	single	1	seasonal	(very weak) diurnal
2	single	2	seasonal	(weak) diurnal
3	pair	3, 4	diurnal	
4	pair	5, 6	3.6 ± 0.14	3.2 ± 0.12
5	pair	7, 8	4.5 ± 0.23	4.1 ± 0.19
6	pair	9, 10	semidiurnal	
7	pair	11, 12	1.73 ± 0.034	$(10.0 \pm 1.24,$ $1.83 \pm 0.38)$
8	trio	13 to 15	1.63 ± 0.030	$(1.47 \pm 0.024,$ $1.73 \pm 0.034,$ $1.66 \pm 0.031)$
9	pair	16, 17	10.0 ± 1.24	$(1.63 \pm 0.030,$ $1.01 \pm 0.011)$
10	pair	18, 19	diurnal	
11	single	20	12.8 ± 2.1	8.2 ± 0.81
12	pair	21, 22	1.80 ± 0.036	1.70 ± 0.031
13	pair	23, 24	4.7 ± 0.26	$(5.6 \pm 0.37,$ $2.9 \pm 0.096)$
14	single	25	7.5 ± 0.67	10.0 ± 1.24
15	pair	26, 27	1.12 ± 0.014	1.38 ± 0.021
16	continuous	≥ 28	various	various

Figure captions

Figure 1: Temporal and zonal average of temperature (gray scale), zonal wind (white) and meridional mass transport streamfunction (black) in the UK-MGCM dataset. Solid, dashed and dotted lines represent positive, negative and null contours, respectively. The separation between contours is $\Delta u = 27.0$ m s^{-1} for zonal velocity, and $\Delta\chi = 129.2$ kg s^{-3} for mass streamfunction.

Figure 2: Planetary waves in the UK-MGCM dataset as seen in (a) surface pressure, and (b) temperature at $\sigma = 0.943$ (approximate height $z = 600$m above the planetary surface). Both figures correspond to an equatorial longitude circle. Negative periods indicates westward travelling waves.

Figure 3: Planetary waves in the UK-MGCM dataset as seen in (a) surface pressure, and (b) temperature at $\sigma = 0.943$ (approximate height $z = 600$m above the planetary surface). Both figures correspond to a longitude circle at 64.3 °N. Negative periods indicates westward travelling waves.

Figure 4: Eigenvalues obtained from POD of the UK Mars GCM.

Figure 5: Cumulative energy in the UK Mars GCM normalised with respect to the total energy in 500 EOFs (without considering energy in the first EOF).

Figure 6: Square root of potential temperature in UK-MGCM EOF3: (a) Longitude-latitude map at $\sigma = 0.6$ ($z \simeq 5.1$ km above the planetary

surface) (white contours represent topography); (b) latitude-altitude section at 36.6°W, and longitude-altitude sections (c) at the equator and (d) 34.9°S.

Figure 7: Mean diurnal tide decomposition over 40 EOFs.

Figure 8: Square root of potential temperature, in arbitrary units, in the UK-MGCM EOF9: (a) Longitude-latitude map at the lower-most sigma-level and (b) longitude-altitude section at 17.8°N.

Figure 9: Mean semidiurnal tide decomposition over 40 EOFs.

Figure 10: Amplitude evolution (defined in text) of the pairs EOF9-10 and EOF5–6 in the semidiurnal tide decomposition.

Figure 11: Mean distribution over 40 EOFs of (a) the diurnal Kelvin wave and (b) westward propagating diurnal wavenumber-3 wave.

Figure 12: Mean distribution of transient waves over 40 EOFs.

Figure 13: MGCM EOF5 sigma-velocity contours on a longitude-latitude map at $z \simeq 32$ km. The arrows represent horizontal velocity.

Figure 14: Longitude-altitude sections at 64.3 °N showing (a) sigma-velocity and (b) square root of potential temperature in UK-MGCM EOF5.

Figure 15: (a) Longitude-latitude map at $z \simeq 32$ km, and (b) longitude-altitude section at 70.2 °N of the square root of potential temperature in UK-MGCM EOF7.

Figure 16: Amplitude evolution of the pairs EOF5–6 and EOF7–8 in the transient wave decomposition. Also included is amplitude evolution of the transient component in the pair EOF3–4.

Figure 17: Sigma-velocity longitude-altitude sections at 70.2 °N in UK-MGCM (a) EOF11 and (b) EOF12.

Figure 18: Longitude-latitude maps of square root of potential temperature, in arbitrary units, at $\sigma = 0.3167$ ($z \simeq 11$ km) (left column) and $\sigma = 0.0403$ ($z \simeq 32$ km) (right column) for UK-MGCM (a) EOF11 and (b) EOF12.

Figure 19: Contours of square root of potential temperature in UK-MGCM (a) EOF13, (b) EOF14 and (c) EOF15 at $\sigma = 0.8312$ ($z \simeq 1.9$ km) (left column) and $\sigma = 0.0403$ ($z \simeq 32$ km) (right column).

Figure 20: Mean distribution of tidal and non-tidal atmospheric motions over 40 EOFs.

Figure 21: Phase portrait of the transient component in EOF5–6 and the diurnal component in EOF3.

Figure 22: Phase portrait of the transient component amplitude in the pairs EOF5–6 and EOF7–8 and the diurnal component in PC-3. The two centres of evolution are labelled with the letters A and B.

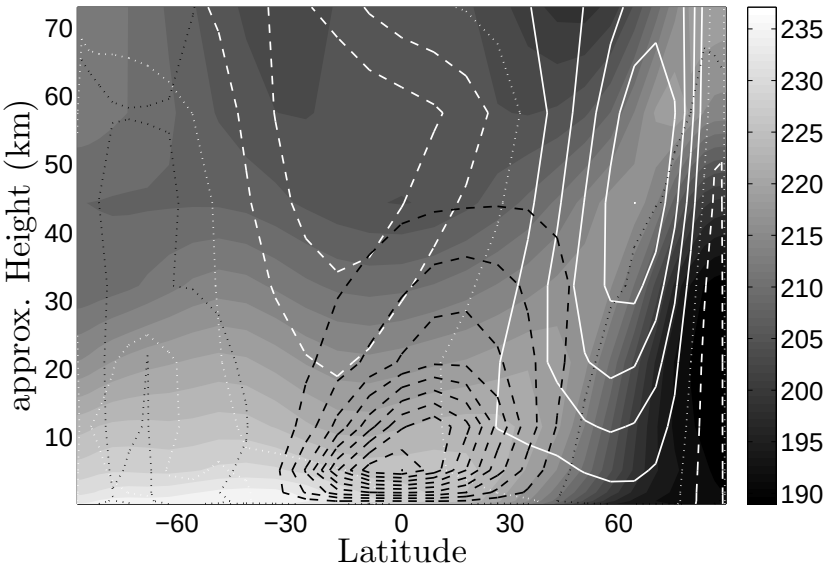
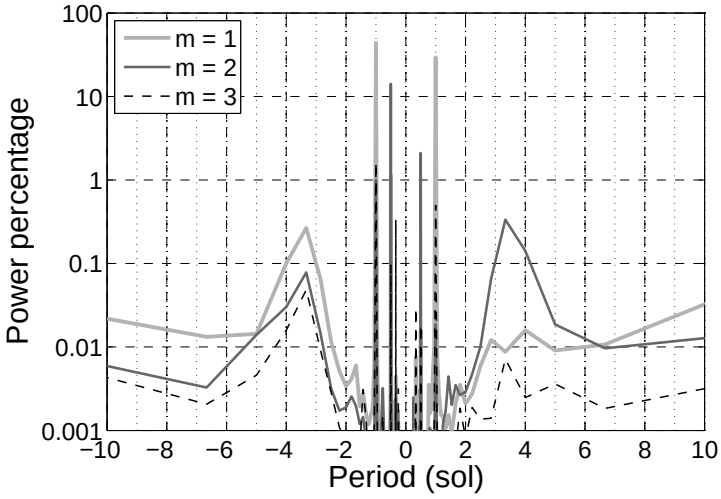
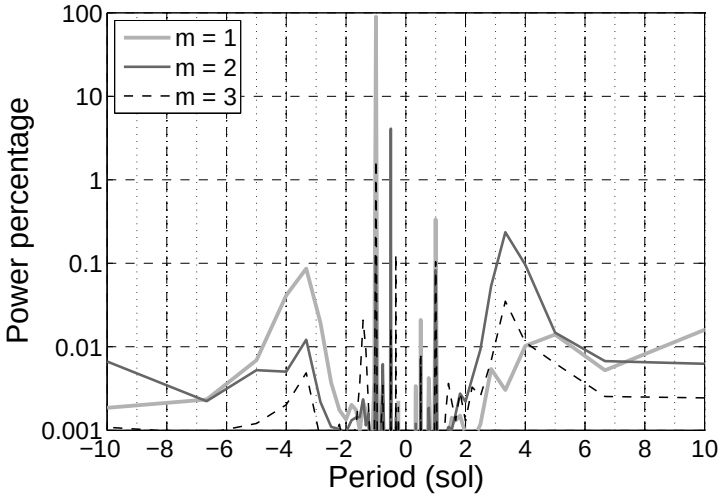


Figure 1: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

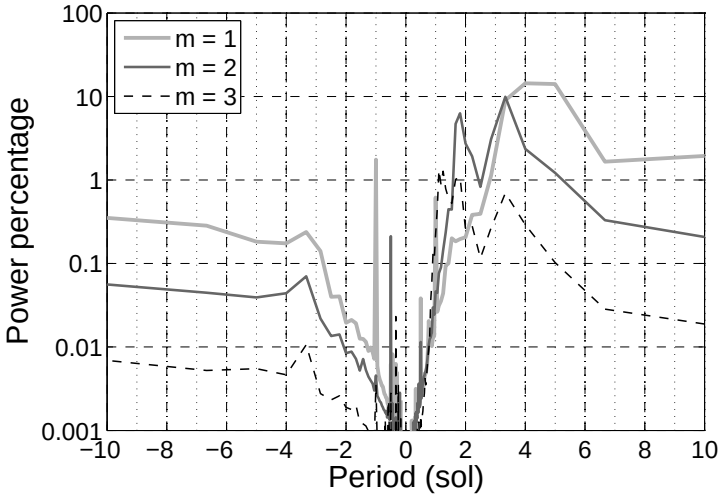


(a)

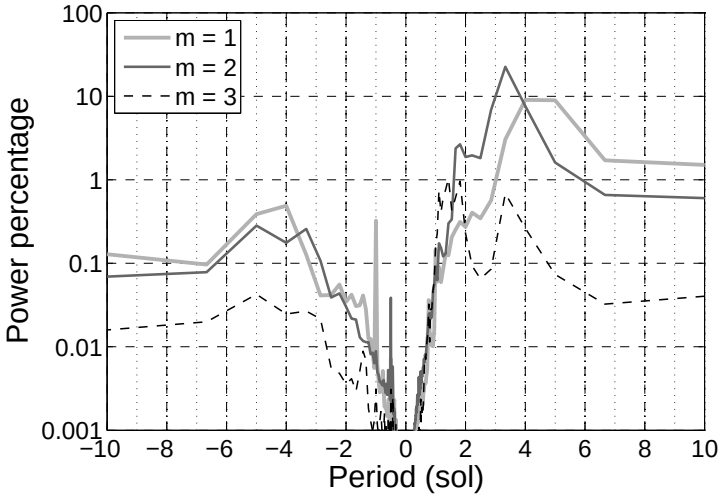


(b)

Figure 2: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a)



(b)

Figure 3: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

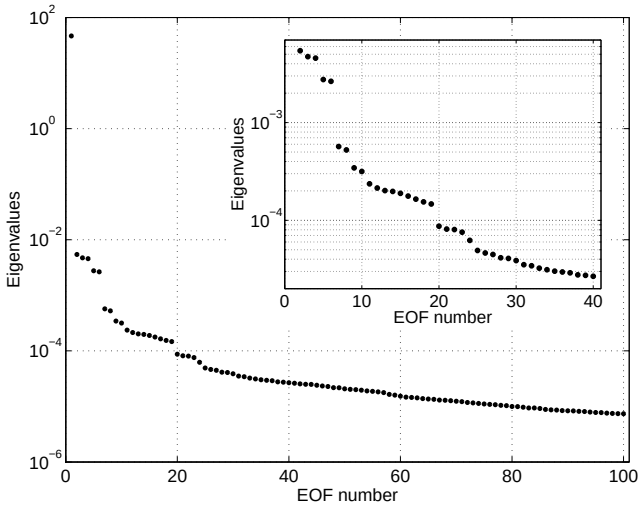


Figure 4: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

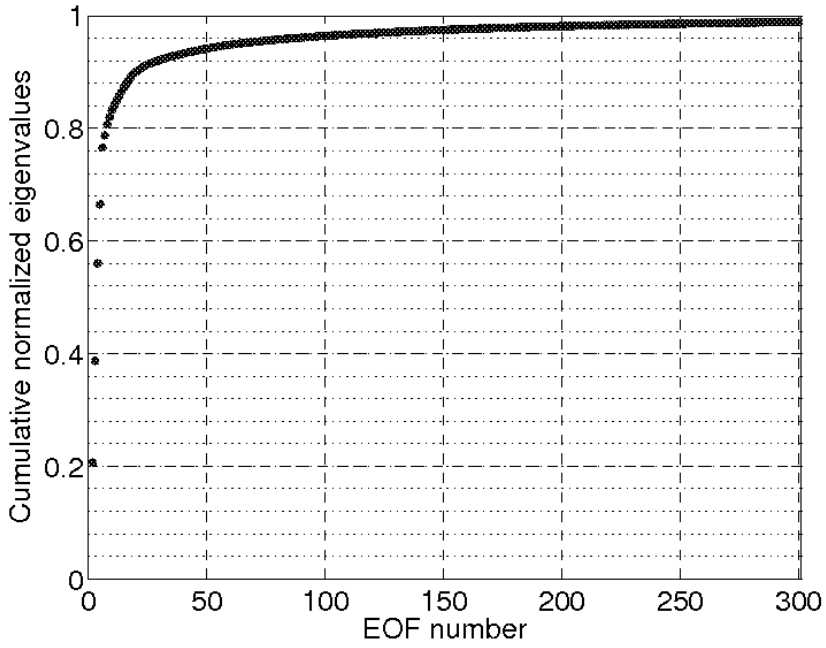
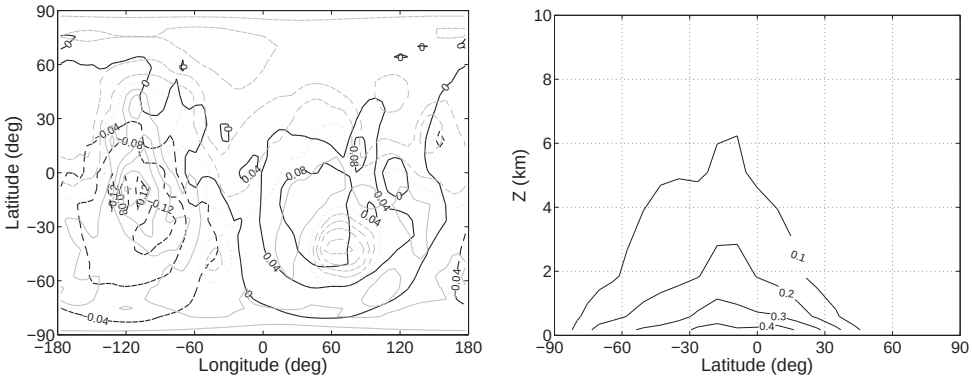


Figure 5: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a) $z \simeq 5.1$ km (b) 36.6 °W

(c) Equator (d) 34.9 °S

Figure 6: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

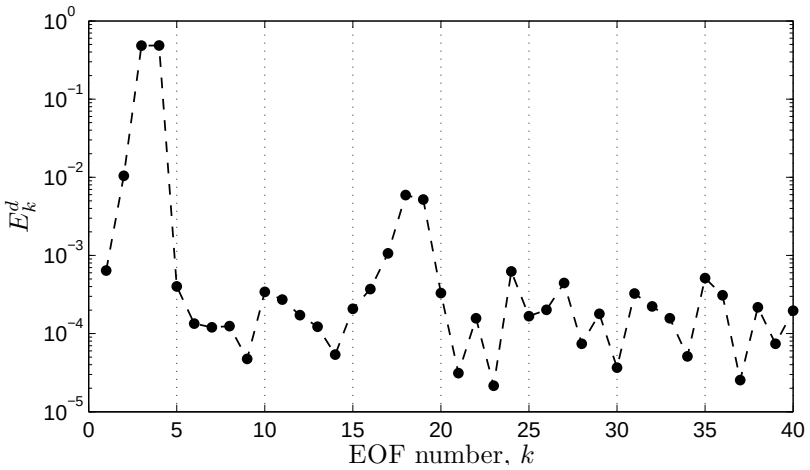
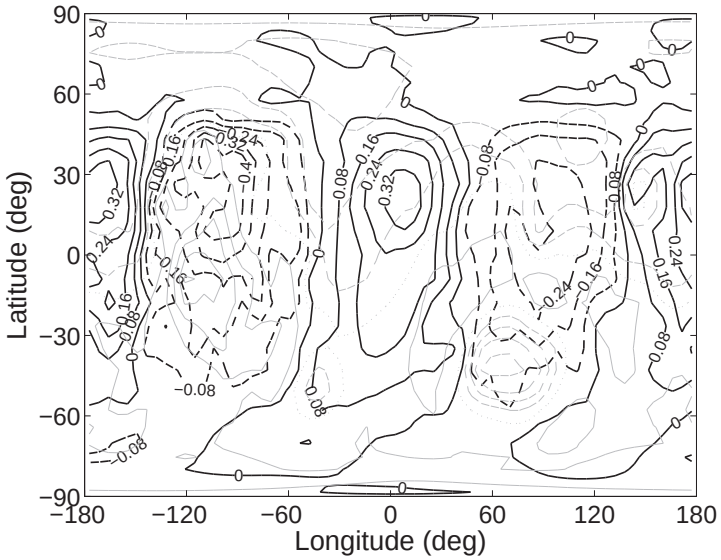
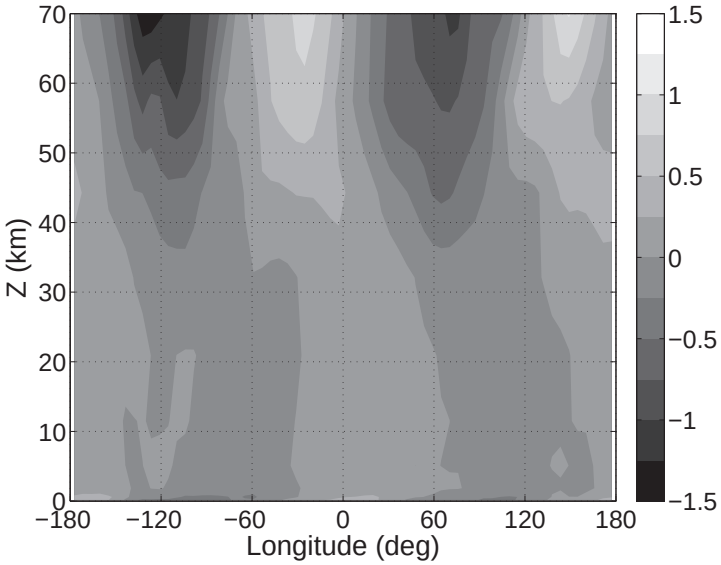


Figure 7: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a) $z \simeq 174$ m



(b) 17.8°N

Figure 8: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

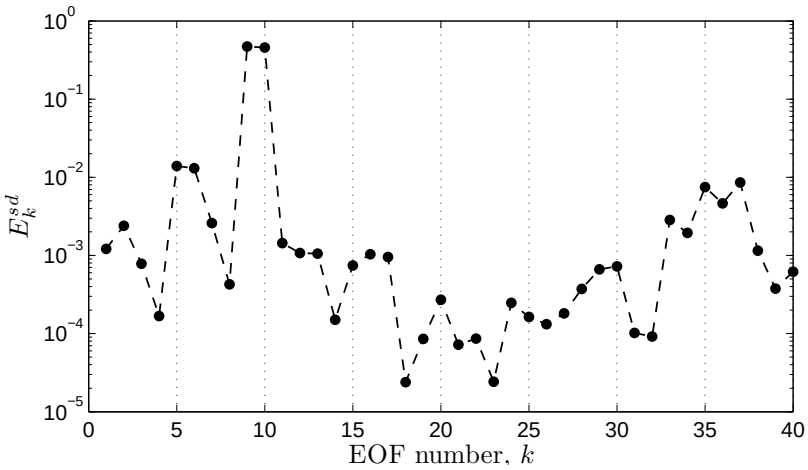


Figure 9: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

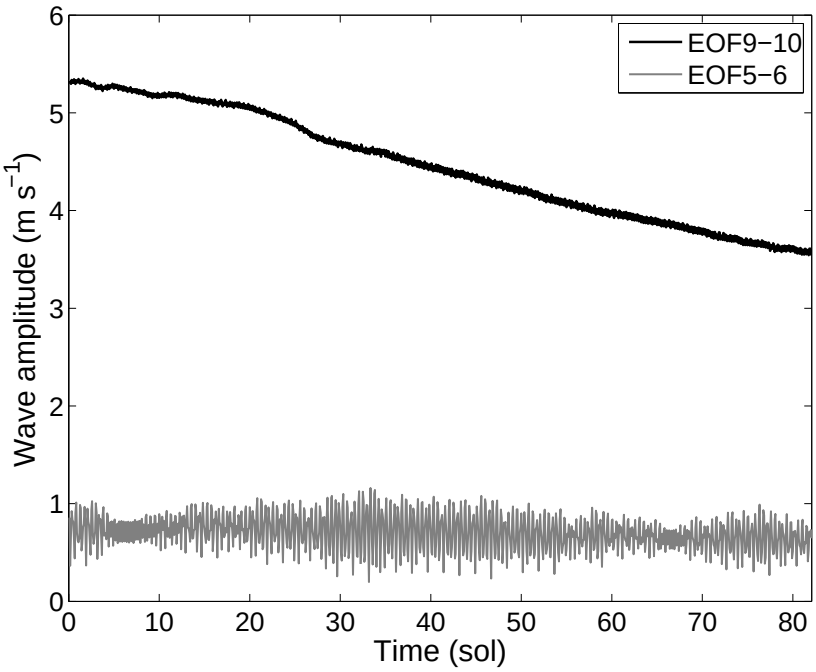
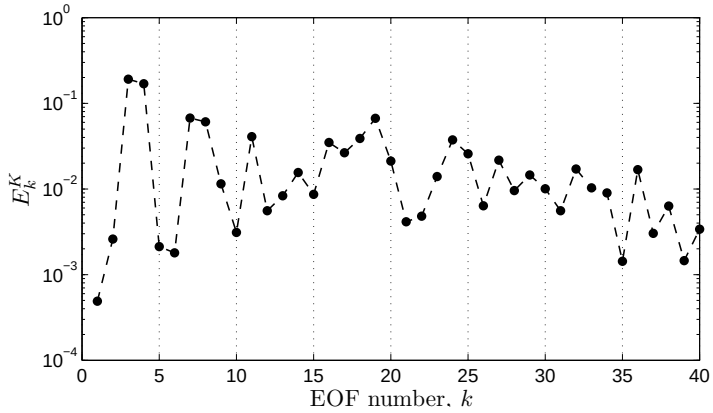
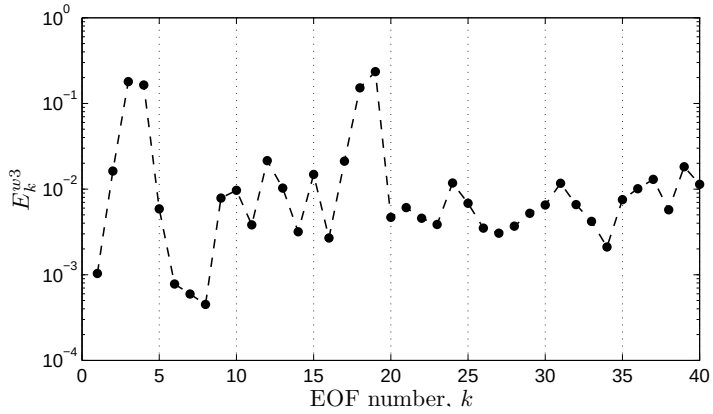


Figure 10: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a) Diurnal Kelvin wave



(b) Diurnal wavenumber-3 wave

Figure 11: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

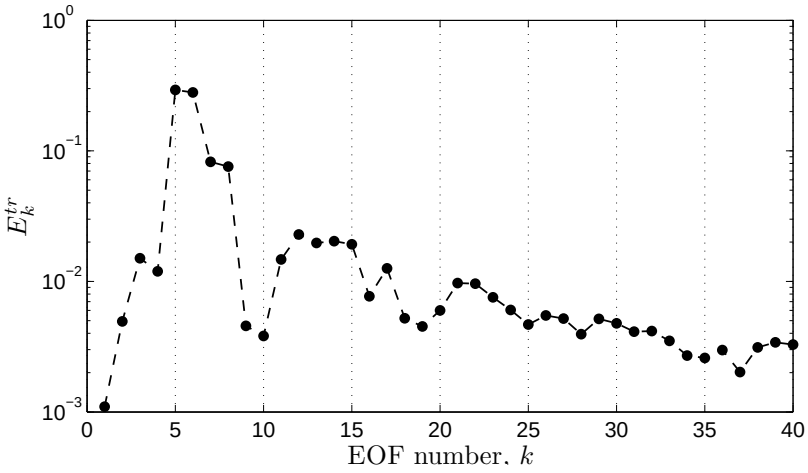


Figure 12: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

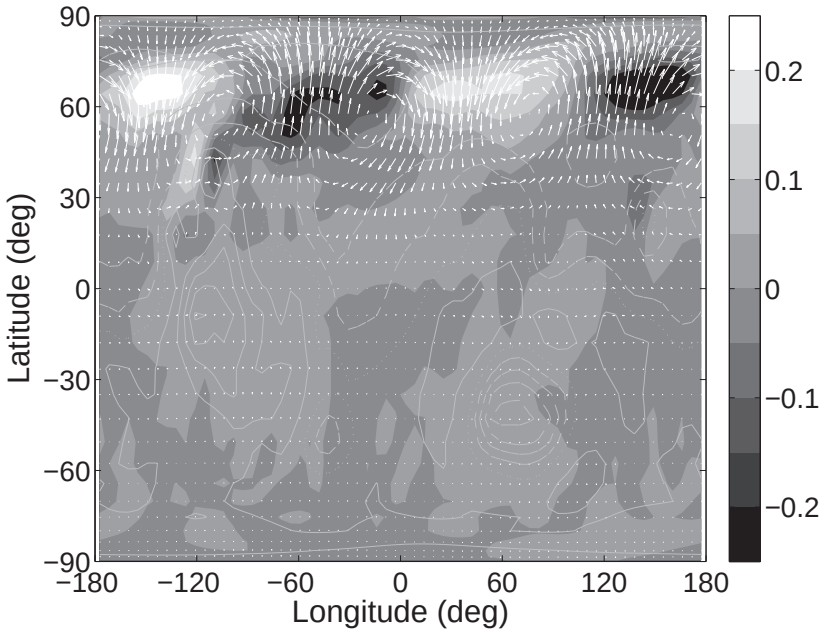
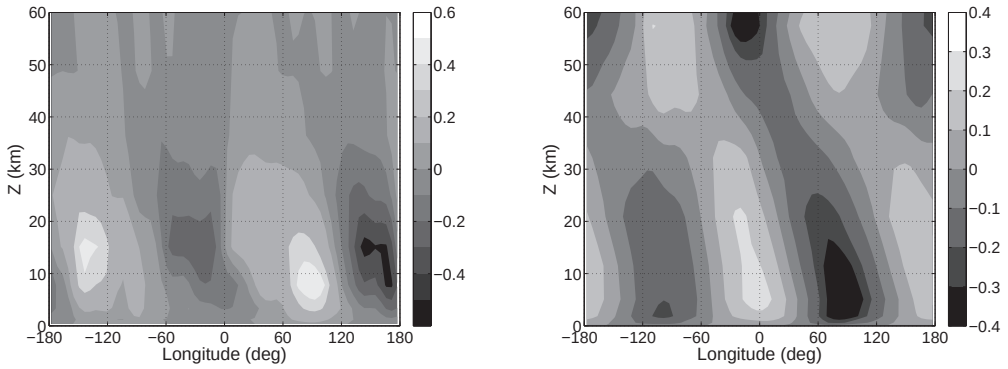


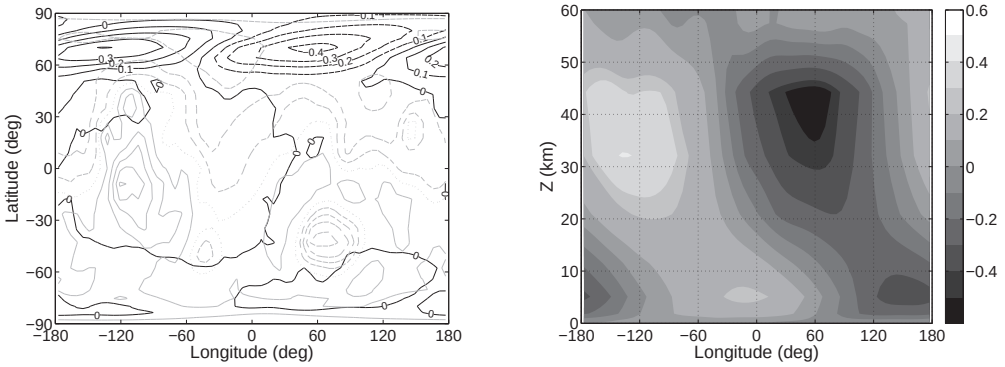
Figure 13: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a) Sigma-velocity

(b) Square root of potential temperature

Figure 14: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a) $z \simeq 32$ km (b) 70.2 °N

Figure 15: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

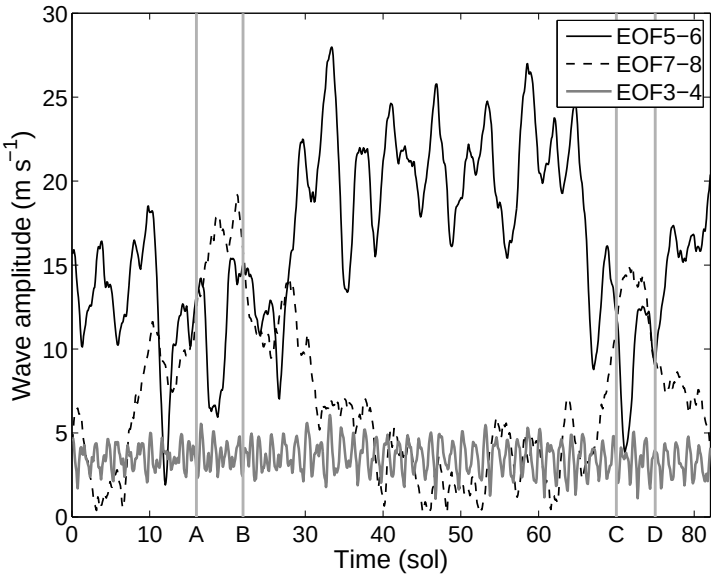
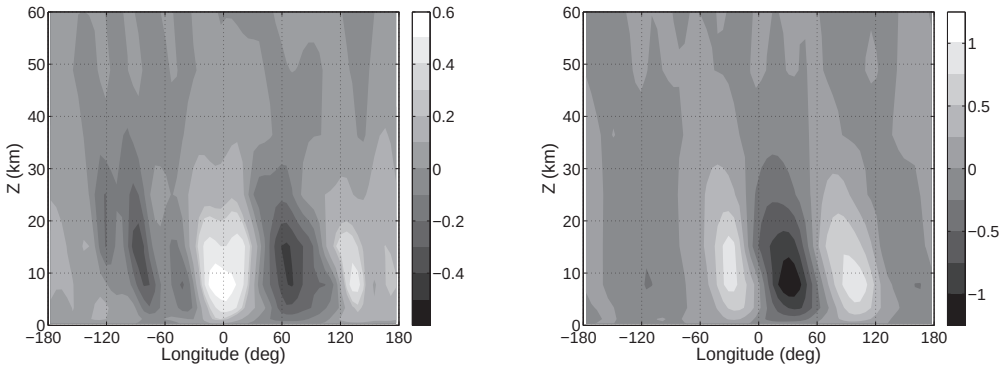
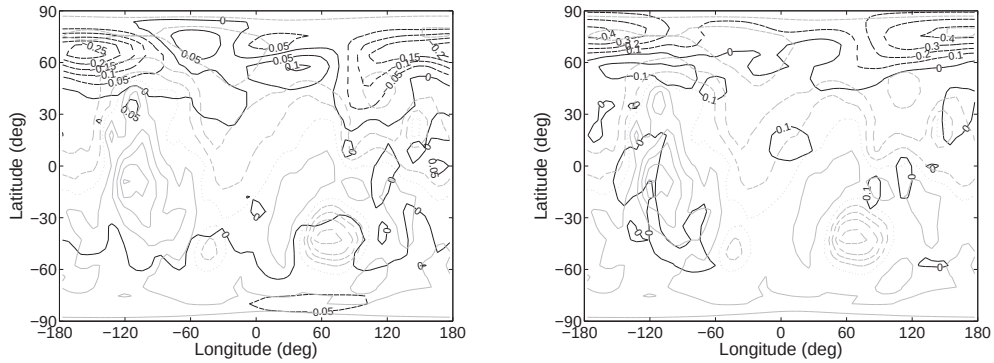


Figure 16: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

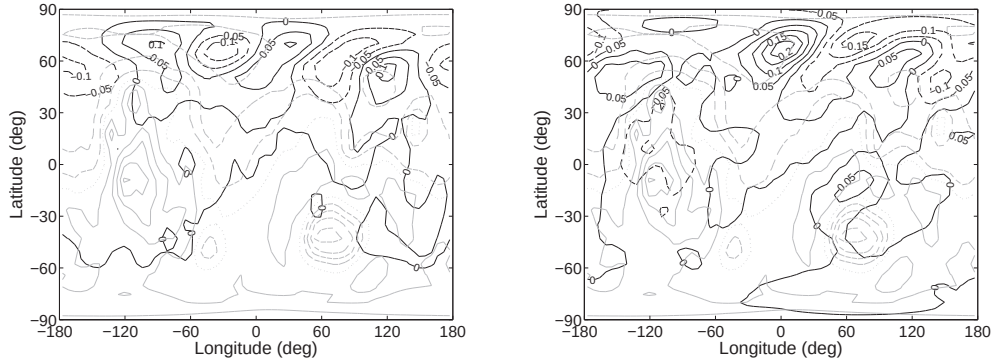


(a) EOF11 (b) EOF12

Figure 17: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

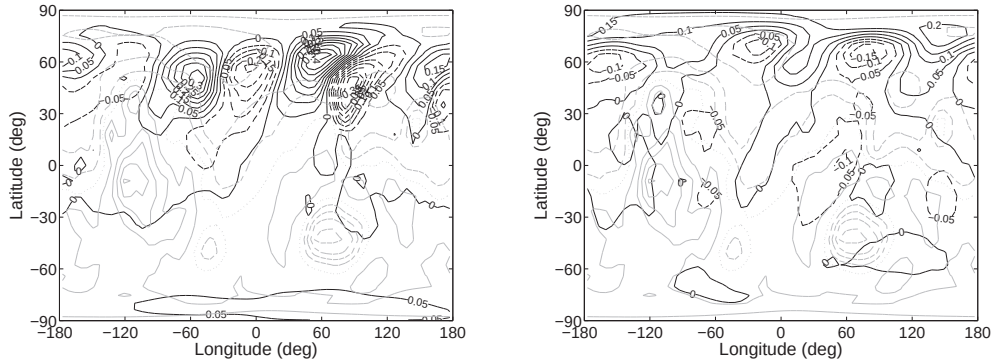


(a) EOF11

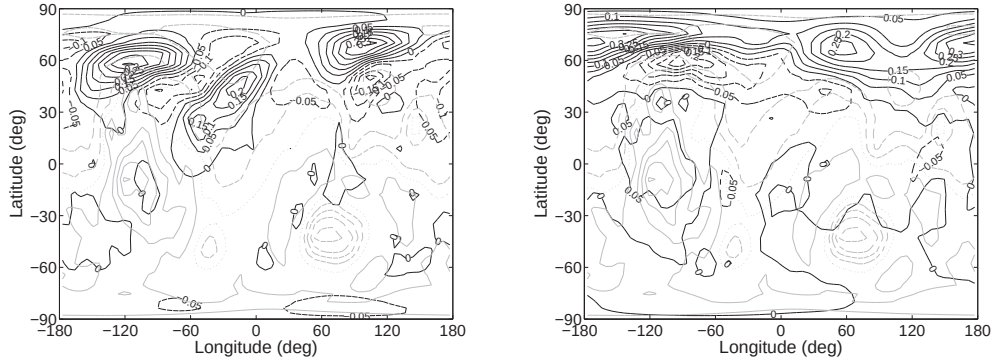


(b) EOF12

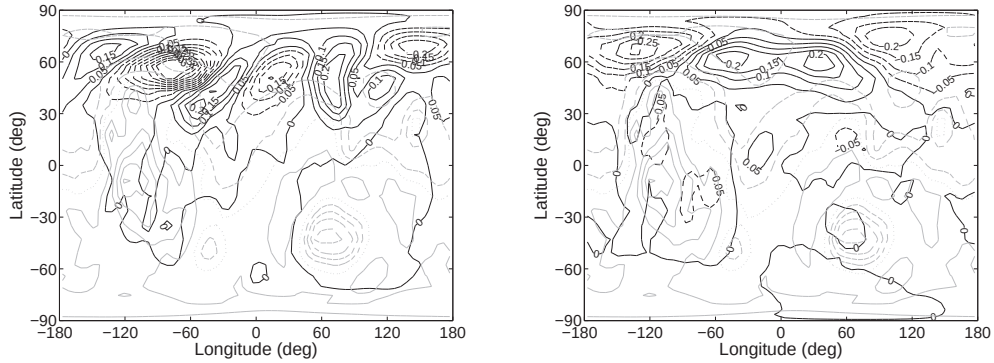
Figure 18: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.



(a) EOF13



(b) EOF14



(c) EOF15

Figure 19: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

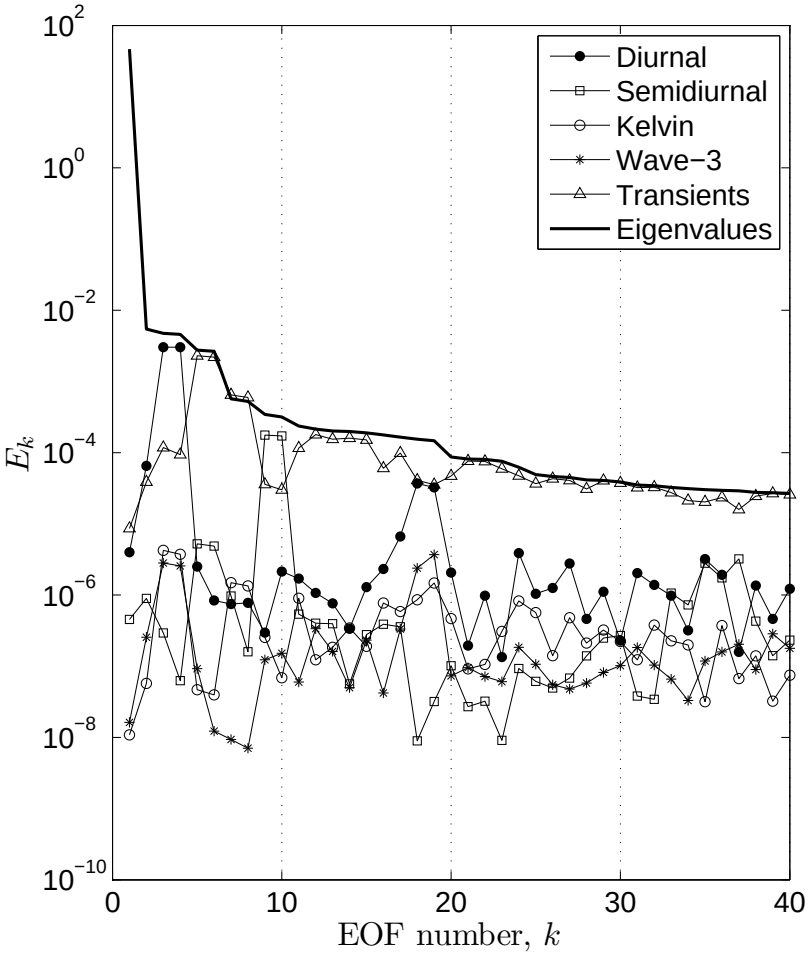


Figure 20: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

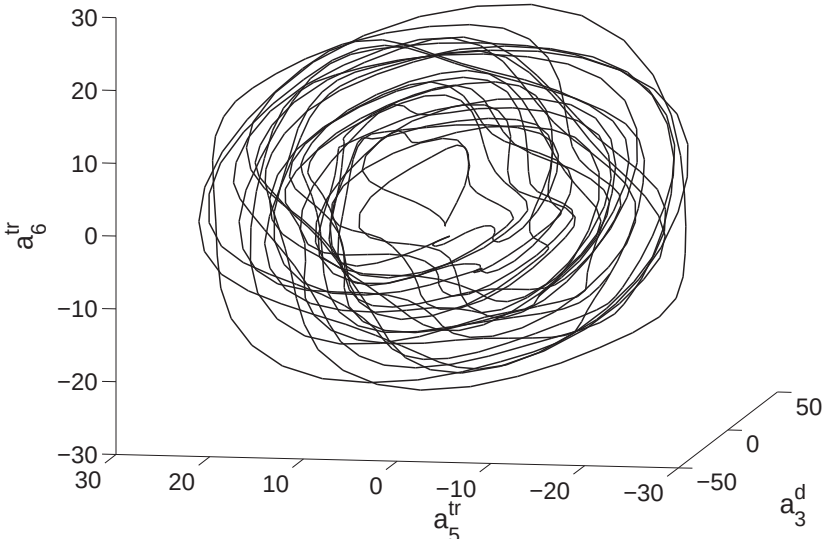


Figure 21: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.

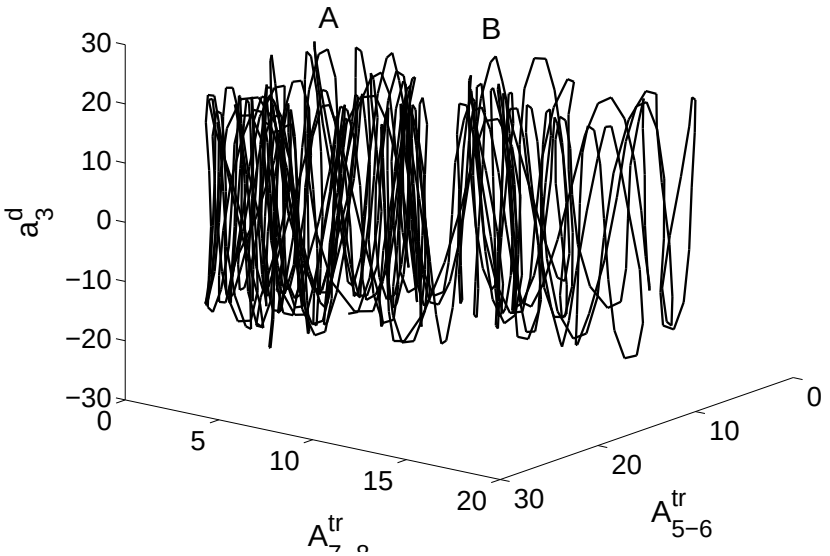


Figure 22: O. Martínez-Alvarado, I. M. Moroz, P. L. Read, S. R. Lewis and L. Montabone.