Transient teleconnection event at the onset of a planet-encircling dust storm on Mars

Journal Article

How to cite:


For guidance on citations see FAQs

© 2009 The Authors
Version: Version of Record
Link(s) to article on publisher’s website:

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Transient teleconnection event at the onset of a planet-encircling dust storm on Mars

O. Martínez-Alvarado¹, L. Montabone², S. R. Lewis², I. M. Moroz³, and P. L. Read⁴

¹Department of Meteorology, University of Reading, UK
²Department of Physics and Astronomy, The Open University, UK
³Mathematical Institute, University of Oxford, UK
⁴Department of Physics, University of Oxford, UK

Received: 15 April 2009 – Revised: 27 August 2009 – Accepted: 7 September 2009 – Published: 30 September 2009

Abstract. We use proper orthogonal decomposition (POD) to study a transient teleconnection event at the onset of the 2001 planet-encircling dust storm on Mars, in terms of empirical orthogonal functions (EOFs). There are several differences between this and previous studies of atmospheric events using EOFs. First, instead of using a single variable such as surface pressure or geopotential height on a given pressure surface, we use a dataset describing the evolution in time of global and fully three-dimensional atmospheric fields such as horizontal velocity and temperature. These fields are produced by assimilating Thermal Emission Spectrometer observations from NASA’s Mars Global Surveyor spacecraft into a Mars general circulation model. We use total atmospheric energy (TE) as a physically meaningful quantity which weights the state variables. Second, instead of adopting the EOFs to define teleconnection patterns as planetary-scale correlations that explain a large portion of long timescale variability, we use EOFs to understand transient processes due to localised heating perturbations that have implications for the atmospheric circulation over distant regions. The localised perturbation is given by anomalous heating due to the enhanced presence of dust around the northern edge of the Hellas Planitia basin on Mars. We show that the localised disturbance is seemingly restricted to a small number (a few tens) of EOFs. These can be classified as low-order, transitional, or high-order EOFs according to the TE amount they explain throughout the event. Despite the global character of the EOFs, they show the capability of accounting for the localised effects of the perturbation via the presence of specific centres of action. We finally discuss possible applications for the study of terrestrial phenomena with similar characteristics.

Keywords. Atmospheric composition and structure (Aerosols and particles) – Meteorology and atmospheric dynamics (General circulation)

1 Introduction

Dust is an atmospheric component that plays a more important role in the Martian climate than in that of the present day Earth. The dust content in the atmosphere has a considerable influence on the atmospheric forcing since the absorption of solar radiation is enhanced when the atmospheric dust load is heavier. One of the most striking features of the Martian atmosphere is its capacity to develop and sustain dust storms that virtually cover the whole planet, a phenomenon known as “planet-encircling dust storm”. The processes that start, sustain and cause a dust storm to decay have been subject to extensive studies involving observations (Smith et al., 2002; Strausberg et al., 2005; Cantor, 2007), Mars general circulation models (MGCMs) (Newman et al., 2002a,b; Toigo et al., 2002; Basu et al., 2006), and data assimilation (Lewis and Barker, 2005; Montabone et al., 2005, 2007). The study of Martian dust storms is extremely beneficial for terrestrial climate studies, as Mars is an alternative natural laboratory to test our knowledge of the dynamics and effects of terrestrial dust storms. Improving our capability to forecast these phenomena is vital for populations living in desert areas of our planet, as well as for aiding the safe landing of spacecraft on Mars.

Teleconnections represent another feature that is not exclusive to the Earth system. The signature of a transient teleconnection event has recently been identified on Mars during the onset of the 2001 planet-encircling dust storm. Montabone et al. (2008) have identified a long-range radiative-dynamical coupling between the site where a regional dust storm (about
a thousand kilometres across) had its explosive growth (near the largest Martian impact crater, the Hellas Planitia basin) and the area where subsequent storms appeared (in the region of the highest Martian volcanoes, the Tharsis Plateau), on the opposite side of the planet (see the topographic map of Mars in Fig. 1). This coupling has the characteristics of a transient teleconnection event, which can create a causal relationship between phenomena occurring at distant locations, through the propagation of some kind of mediating signal.

In contrast with teleconnection patterns, which can be thought as vast atmospheric regions dynamically coupled over long time scales, a transient teleconnection event couples distant regions over much shorter time scales. Typical time scales for teleconnection patterns on the Earth range from inter-annual to inter-decadal. In studies of the terrestrial atmospheric dynamics, proper orthogonal decomposition (POD), also known as principal component analysis (PCA), has been used to characterise teleconnection patterns (Preisendorfer, 1988). In this case, teleconnection patterns are described by the first empirical orthogonal functions (EOFs), when the analysis is performed using time series that span several years (Barnston and Livezey, 1987).

The objective of this paper is to extend the use of POD to characterise a transient teleconnection event during the initial stage of the 2001 planet-encircling dust storm on Mars (Cantor, 2007; Montabone et al., 2008), which occurred at a time scale in the order of a few sols (1 sol = 1 Martian day ≃ 1 terrestrial day + 40 min).

In this paper, we use a four-dimensional dataset (three spatial coordinates and time) of the Martian atmospheric state at the time of the onset and development of the 2001 planet-encircling dust storm. Nadir thermal profiles and columnar dust optical depth observations from the Thermal Emission Spectrometer (TES) on board NASA’s Mars Global Surveyor (MGS) (Conrath et al., 2000; Smith et al., 2000b; Smith, 2004) were assimilated into the United Kingdom Mars general circulation model (UK-MGCM) (Lewis and Read, 1995; Forget et al., 1999; Lewis et al., 2007). We apply to this dataset a combination of POD and Fourier analysis as described in Martínez-Alvarado et al. (2009), in order to extract the coherent structures that better represent the observed teleconnection. Martínez-Alvarado et al. (2009) used this technique to investigate the hypothesis of a low dimensional Martian climate attractor in the UK-MGCM. In contrast, in this paper we take into account the transient character of the studied phenomena by performing a time-evolving POD.

The rest of the article is organised as follows. In Sect. 2, Martian meteorology is briefly reviewed. Section 3 describes the 2001 planet-encircling dust storm as our case study. A complete description of the data and the methodology is given in Sect. 4. The results of the application of this novel technique to the study of the case-study are described in Sect. 5. In Sect. 6 we draw the conclusions and we discuss the possible applications of this work to the study of the dust storms and transient teleconnection events on the Earth.

2 Review of relevant large-scale Martian meteorology

The lack of oceans and tenuous atmosphere on Mars makes this planet more susceptible to stronger temperature variations and to a more rapid response to solar radiative forcing than the Earth (Read and Lewis, 2004). These conditions are ideal for sustaining planetary-scale waves with periods that are harmonics of the solar day directly forced by solar heating. These waves are known as thermal tides and constitute an important component of the Martian atmospheric dynamics. The most important thermal tides on Mars include the migrating, sun-synchronous, diurnal and semi-diurnal tides (Forbes, 2004). The semi-diurnal tide, especially, exhibits a strong sensitivity to vertically-integrated solar heating. Since dust in the atmosphere enhances its radiative absorption properties, the semi-diurnal tide magnitude also shows a strong correlation with atmospheric dust content (Lewis and Barker, 2005). The interaction of the diurnal solar heating with the large Martian topography gives rise to non-migrating (longitude-dependent) tides. The most important of these has the characteristics of an eastward propagating diurnal mode (Zurek, 1976). It has been shown that also this wave can be affected by the enhanced presence of dust in the atmosphere during major dust storms (Wilson and Hamilton, 1996; Lewis and Barker, 2005; Montabone et al., 2008). This diurnal non-migrating tide has a period close to that of a Kelvin mode in the Martian atmosphere.

Despite several differences between Mars and the Earth, the dynamics of the Martian atmosphere exhibits remarkable analogies to the Earth’s. Like on the Earth, the energy transfer from the equator polewards is carried out by waves generated via barotropic and baroclinic instabilities.
Barotropic instability is associated with the reduction of horizontal shear, and has been related to the circumpolar jet in the middle Martian atmosphere (Michelangeli et al., 1987). Evidence for baroclinic activity on Mars have been found in surface pressure time series from the Viking Lander 2 mission (Barnes, 1980, 1981) and in Mars GCM simulations (Barnes et al., 1993; Collins et al., 1996). These baroclinic waves define the structure of the Martian “storm tracks”, which result in atmospheric fronts that might initiate high latitude dust storms (Wang, 2007). These fronts are remarkably similar to the terrestrial counterpart, with water vapour and clouds playing the role of the Martian dust. Other dynamical features closely resemble terrestrial phenomena: dust storms apart, a “dry” monsoon circulation can be recognized on Mars, as well as western boundary currents (Read and Lewis, 2004), polar vortices (McConnachie et al., 2003), and many mesoscale phenomena related to topographic features and radiative effects typical of desert locations on the Earth (Spiga and Forget, 2009).

3 Case study

Planet-encircling dust storms are a major atmospheric phenomenon on Mars. Since the early ground-based observations of the planet in the XIX century, 6 global dust storms have been fully documented, and at least other 2 might potentially be considered as global storms (Martin and Zurek, 1993; Zurek and Martin, 1993). Many others might have gone undetected before the spacecraft era. However, it is clear that planet-encircling storms on Mars are characterised by high intermittency and year-to-year variability.

Observations show that every Martian year during northern autumn and winter the dust activity increases, raising the probability of observing small (sizes of few tens kilometres across) and large (sizes of few hundreds kilometres across) dust storms. This well observed phenomenon coincides with the greater forcing in the atmosphere during the period around perihelion (which on Mars occurs at arocentric longitude\(^1\) \(L_s=251^\circ\)).

Planet-encircling dust storms, however, are not observed every year, showing an estimated intermittency of two to three Martian years (Zurek and Martin, 1993). The size and exact season for the occurrence of global events also vary significantly within the observed group of planet-encircling storms. A common characteristic of all global events is their high impact on the thermal and dynamical structure of the lower atmosphere for long times (few tens of sols), which is due to the enhanced absorption of visible and infrared radiation by the large loading of mineral aerosol at planetary scale. The Hadley circulation and the baroclinic waves at high northern latitudes are among those dynamical features which are deeply affected by the presence of a planet-encircling dust storm (Montabone et al., 2005; Wang, 2007).

The 2001 planet-encircling dust storm was the combination of a series of multiple, isolated, regional-scale storms (about a thousand kilometres across) occurring between June and August that year. These storms merged and evolved rapidly over a period of a few sols, contributing to create a diffuse and global dust haze lasting for several sols (Cantor, 2007). The Martian Orbiter Camera (MOC) and the TES both on board of the MGS spacecraft provided images, profiles of temperature and values of columnar dust opacity during the event (Smith et al., 2001; Smith, 2004). These observations, together with radio occultation observations performed with the high gain antenna have made this storm the first planet-encircling storm for which a variety of good quality data with planetary coverage exists. The comprehension of such a global event requires the study of several associated meteorological phenomena operating at different scales, one of which (the transient teleconnection event) is the topic of the present study.

The key events which characterised the evolution of the storm have been described by Strausberg et al. (2005) and Cantor (2007). These include the development of a large regional storm around the northern rim of the largest Martian impact crater (Hellas Planitia basin) and the onset of subsequent regional storms in the region of the highest Martian volcanoes (the Tharsis Plateau). Figure 2 shows three snapshots during the development of the 2001 planet-encircling dust storm, using MGS/TES total column dust opacity at the reference surface pressure of 610 Pa. These correspond to the dust outburst over Hesperia Planum and Tyrhena Terra at \(L_s=186^\circ\) (Fig. 2a) and the activation of secondary lifting centres around the Tharsis Plateau only three sols later, around \(L_s=189^\circ\) (Fig. 2b). The third snapshot (Fig. 2c) shows the dust storm once it is fully developed 39 sols after the initial major regional storm.

Hesperia Planum and Tyrhena Terra on one side of the planet and the Tharsis Plateau on the other might have been teleconnected through a large-scale radiative-dynamical coupling (Montabone et al., 2008). Thus, after the first regional dust storm developed near Hellas Planitia, dust particles aloft produced a strong heating in the lower atmosphere through the absorption of incident visible radiation. Such anomalous, localized forcing at equatorial latitudes modified the cyclic solar forcing and the stationary forcing provided by the interaction of solar heating with the large-scale Martian topographic features. In particular, three components of the Martian thermal tides resulted deeply modified: the westward-propagating diurnal and semi-diurnal components, and the eastward-propagating diurnal non-migrating component. This interference was associated with a sudden zonal phase shift of the large-scale surface pressure and wind...
patterns. Strong local slope winds in the favourable locations of the Tharsis Plateau might have developed new dust lifting centres that evolved very quickly into new regional dust storms. This kind of teleconnection would be extremely efficient on Mars because of its thin atmosphere, the crucial role played by the dust in the atmospheric forcing, and the lack of water vapour that could limit the growth of dust clouds.

It is worth mentioning here that Zurek and Leovy (1981) have already recognized the key role played by the westward and eastward-propagating diurnal tidal modes at the onset of planetary-encircling dust storms on Mars by analysing surface pressure data from the Viking Landers during the 1977 dust storms. Their linear analysis of the atmospheric tides revealed an anomaly in the diurnal component at the onset of the 1977 planetary encircling dust storm. Although the analysis of a single point surface pressure time series could not distinguish between westward and eastward-propagating diurnal modes, they argued that the sudden drop in the diurnal tide amplitude could be consistent with the enhancement of the eastward-propagating mode and the resulting destructive interference with the westward-propagating mode. Montabone et al. (2008) and the present paper contribute to extend the idea further, recognizing the role played by these two components of the tides in establishing a transient teleconnection in a global, fully non-linear model with data assimilation.

4 Data and methodology

4.1 Dataset

In this study we are interested in the processes that occurred during the onset of the 2001 planet-encircling dust storm. We focus our analysis on the storm that initiates around Tyrrhena Terra and Hesperia Planum (northeast of Hellas Planitia basin, see Fig. 1), and on the link between this primary dust lifting region and the onset of a second dust storm around Daedalia Planum (Montabone et al., 2008).

The dataset consists of output from the UK-MGCM (Forget et al., 1999; Lewis et al., 1999). This is a state-of-the-art model developed at the University of Oxford and now jointly with the Open University in the UK. It is based on the spectral dynamical solver by Hoskins and Simmons (1975). The UK-MGCM includes a parameterisation package based on that developed for the Laboratoire de Météorologie Dynamique (LMD) MGCM. It comprises a variety of relevant processes for the Martian meteorology such as radiative heat transfer, surface processes, sub-grid dynamics and carbon dioxide mass exchange (Forget et al., 1999). The model also incorporates a high-resolution Mars topography dataset as measured by the Mars Orbiter Laser Altimeter (MOLA) on board the MGS spacecraft (Smith et al., 1999).

The dataset under analysis ranges from $L_s=168.95^\circ$ to $L_s=262.10^\circ$ and represents the period from late winter to the end of spring in the Southern Hemisphere. This corresponds to the transition from the period of remnant baroclinic activity in the Southern Hemisphere to the start of the baroclinic activity season in the Northern Hemisphere. This period is

Fig. 2. Instantaneous plots of (equivalent visible) dust optical depth at the reference pressure of 610 Pa assimilated at model resolution. (a) $L_s=186^\circ$ (sol 383 from the beginning of the Martian year), (b) $L_s=189^\circ$ (sol 386), and (c) $L_s=210^\circ$ (sol 422) during the initial stage of the 2001 planet-encircling dust storm. White contours represent topography with solid, dotted and dashed lines indicating contours above, at and below the reference areoid, respectively. Blank patches indicate regions where data is unavailable.
Terrestrial weather forecasting is carried out using data assimilation to find the state of the atmosphere that is most consistent with all available observations and with an atmospheric model as an initial state for a numerical forecast. We have applied this technique to Mars using the MGS/TES observations (remotely-sensed temperature profiles and dust opacity values) and the UK-MGCM. The data assimilation technique allows us to analyse the complete evolution in three spatial coordinates and time of the Martian atmosphere during the period of this planet-encircling dust storm. The analysis includes variables that could not be directly observed by the MGS spacecraft, such as surface pressure and winds.

The data assimilation is conducted using a modified form of the sequential analysis correction scheme (Lorenc et al., 1991) with parameters tuned for the specific case of Mars. Lewis et al. (2007) summarises the details of the technique, and Montabone et al. (2006) presents a validation of the data assimilation for Mars using independent observations (MGS radio occultations). We only mention here that, at each sample interval, the technique combines information from both present and past observations of MGS/TES thermal profiles and columnar dust optical depth, which are retrieved from nadir soundings below 40 km altitude (see Smith et al., 2000a, for details of the retrieval technique). The GCM at T31 resolution (corresponding to a triangular truncation with maximum zonal wavenumber 31 with a 96×48 grid for nonlinear products and 64×32 grid for physical processes) and 25 vertical levels on a stretched grid is then used to produce a time-evolving analysis of the atmospheric state.

Since only total column dust opacities can be retrieved from nadir soundings, the vertical distribution of the dust has to be prescribed, according to the following equation in terms of volume mixing ratio (cf. Montabone et al., 2006),

$$ q = q_0 \exp \left( \alpha - \left( \frac{p_{\text{ref}}}{p} \right)^{b/z_{\text{max}}} \right), $$

for pressure $p \leq p_0$, where $p_{\text{ref}}$ is a reference pressure (on Mars we choose the average surface pressure of 610 Pa), and with $q=q_0$ for $p > p_{\text{ref}}$. $q_0$ is calculated to give the assimilated total optical depth at the appropriate latitude, longitude and time at the reference pressure. $\alpha$ and $b$ are free parameters with values $\alpha=0.007$, $b=70$ km, and $z_{\text{max}}$ (the top of the dust layer on Mars) varies with the time of the year and latitude according to Eq. (2) in Montabone et al. (2006). Furthermore, we mention that the retrieved dust opacities from MGS/TES are in the infrared (wavelength around 1075 cm$^{-1}$), whereas the GCM radiation scheme computes dust heating rates based on a mean visible opacity. Total dust opacities have therefore been converted from the infrared to a mean visible value by multiplying by a factor of 2.0, according to Clancy et al. (1995, 2003). Data assimilation for the Martian atmosphere has been carried out using the available MGS/TES observations during the spacecraft mapping phase, namely between May 1999 and August 2004.

The output from the data assimilation has been pre-processed in order to obtain three-dimensional time series of horizontal velocity and temperature. These variables were sub-sampled on a regular longitude-latitude grid with 64×31 grid points (approximately corresponding to 5.6° and 5.8°, respectively). In the vertical direction, the data were discretized over 10 unevenly spaced levels using terrain-following sigma-coordinates ($\sigma = p/p_s$, where $p_s$ is surface pressure) as shown in Table 1. The number of levels was chosen in order to maintain the computational cost within practicable limits. Horizontal velocity, temperature and surface pressure are considered here as state variables, containing all the information about the state of the atmosphere at any time. The state of the atmosphere was measured by its total atmospheric energy (TE) as explained in the next sub-section.

### 4.2 Methodology

The objective of our analysis is the identification of a minimal reduced space retaining the most relevant effects of the onset of the 2001 planet-encircling dust storm over the atmosphere. In particular, we are interested in retaining the transient teleconnection event that allowed for the quasi-simultaneous activation of various dust lifting centres. The analysis of the data has been carried out using a combination of a TE-based proper orthogonal decomposition (POD) (Achatz and Opsteegh, 2003) and spatio-temporal short-time Fourier analysis (STFA), in a similar way to that described in Martínez-Alvarado et al. (2009).

POD has been used to extract a series of coherent structures, also called empirical orthogonal functions (EOFs). Unlike typical POD, where empirical orthogonal functions (EOFs) are ordered according to the amount of explained variance, TE has been used to weight the resulting modes. This choice has been shown to be useful in previous atmospheric studies (e.g. Achatz and Opsteegh, 2003; Whitehouse et al., 2005). It gives a physically meaningful measure to the atmosphere’s state, which is given in terms of variables of different physical nature. Hence, a dynamical variable such
as horizontal velocity is combined with a thermal variable such as temperature. Assuming hydrostatic balance, TE for an air column is computed using the expression

\[ e = \int_0^p \frac{1}{g} \left( \frac{1}{2} v \cdot v + c_p T \right) dp + \frac{1}{g} \Phi_S p_s \]  

(2)

where \( v \) is horizontal velocity, \( T \) is temperature, \( p \) is local pressure, \( p_s \) is surface pressure, \( \Phi_S \) is geopotential height at the surface (topography), \( g = 3.72 \) m/s\(^2\) is the acceleration due to gravity, and \( c_p = 820 \) J kg\(^{-1}\) K\(^{-1}\) is the heat capacity at constant pressure. These values are those appropriate for Mars.

In order to make the expression for the energy a quadratic form in terms of the state variables, Eq. (2) can be non-dimensionalised by choosing the planetary radius \( a \) and the reciprocal rotation rate \( \Omega^{-1} \) \( (a = 3.396 \times 10^9 \) m, \( \Omega = 7.088 \times 10^{-5} \) s\(^{-1}\) as length and time scales, respectively.

Furthermore, by introducing the new variables \( \alpha = \sqrt{\frac{g}{a}} \mu = \alpha u, \nu = \alpha v, \) and \( \tau = \alpha \sqrt{T} \), where \( u \) and \( v \) are the zonal and meridional velocity components, respectively, we can use Eq. (2) to define an inner product for the vector space of atmospheric states \( \psi = (\mu, \nu, \tau, \alpha)^T \), where \( \psi^T \) represents the transpose of \( \psi \). The expression for the inner product is

\[ \langle \psi^1, \psi^2 \rangle = \int_A \left( \frac{1}{2} \mu^1 \mu^2 + \frac{1}{2} \nu^1 \nu^2 + \frac{1}{\kappa} \tau^1 \tau^2 + \Phi_S \alpha^1 \alpha^2 \right) dx, \]  

(3)

where the superindices (1) and (2) just indicate two different states and the integration is carried over the whole atmosphere. Thus, the EOFs constitute the solutions to the eigenvalue equation (Berkooz et al., 1993)

\[ \int_A K(x, x') E(x') \varphi^{(k)}(x') dx' = \Lambda^{(k)} \varphi^{(k)}(x), \]  

(4)

where \( \Lambda^{(k)} \) and \( \varphi^{(k)} \) play the role of the \( k \)-th eigenvalue and eigenfunction, respectively. The energy matrix \( E \) is the weighting function in the definition of the inner product and can be expressed as (cf. Eq. 3)

\[ E = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/\kappa & 0 \\ 0 & 0 & 0 & \Phi_S \end{pmatrix}. \]

(5)

The kernel \( K \) is the correlation function given by

\[ K(x, x') = \langle \psi(x) \psi^T(x') \rangle, \]

(6)

where \( \langle \cdot \rangle \) denotes the average over the UK-MGCM dataset or, equivalently, an average over the time interval under analysis. The eigenvalue \( \Lambda^{(k)} \) represents the average TE explained by the \( k \)-th EOF. The EOFs are sorted in decreasing order according to their associated eigenvalues. Further details about the method can be found in Martínez-Alvarado et al. (2009).

### Table 2. Thermal tides retrieved by spatio-temporal short-time Fourier analysis.

<table>
<thead>
<tr>
<th>Thermal tide</th>
<th>Zonal wavenumber</th>
<th>Frequency (sol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diurnal</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Semidiurnal</td>
<td>-2</td>
<td>2</td>
</tr>
<tr>
<td>Diurnal non-migrating</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Diurnal wave-3</td>
<td>-3</td>
<td>1</td>
</tr>
</tbody>
</table>

In turn, spatio-temporal STFA has been used to isolate specific wave components embedded within the original dataset and, consequently, also within the EOFs. Given that the sampling rate is a sample every 2 Martian hours, the maximum resolved frequency is \( f_{\text{max}} = 0 \) sol\(^{-1}\). The frequency resolution is given by the length of the time-window used to perform the STFA. In this study we have used a time-window of 8 sols equivalent to a frequency resolution \( \Delta f = 0.125 \) sol\(^{-1}\). Zonally, 64 grid points around a longitudinal circle allows for a resolution of zonal wavenumbers from 1 to 31. The wave components retrieved were various thermo-tidal components characterised by specific zonal wavenumber and frequency. The waves that were extracted in this form appear in Table 2. The difference between the original dataset (after removing the time-mean state) and the sum of the thermal tides listed in Table 2 was grouped in a fifth group labelled as transient waves.

The TE distribution of the various waves over the EOFs was computed as

\[ E_{k,n}^{\psi} = \frac{1}{Z} \left( a_{k,n}^s \right)^2, \]

(7)

where \( a_{k,n}^s(t) = \langle \varphi^{(k)}, \psi_{s,n}(t) \rangle \) 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

\[ Z = \sum_{k=1}^N \left( a_{k,n}^s \right)^2 \]

(8)

is a normalization factor where \( N \) is the number of retained EOFs, which in our case was set to 100 modes. Further details about the spatio-temporal STFA can also be found in Martínez-Alvarado et al. (2009).

### 4.3 EOFs in a transient regime

An essential step in our study is the projection of the original dataset onto the EOF linear space in order to obtain time series of the corresponding time-dependent coefficients. These coefficients are also called principal components (PCs). The \( k \)-th PC \( a_k \) is given by

\[ a_k = \langle \varphi^{(k)}, \psi \rangle. \]

(9)
The instantaneous energy associated with a particular EOF is given by the square of the instantaneous value of the corresponding PC.

Every state in the original dataset can be reconstructed in terms of a linear combination of EOFs as

$$\psi = \sum_{j=1}^{N} a_j \phi^{(j)}, \quad (10)$$

where $N$ is the number of EOFs that are retained. In a steady-state system, the accuracy of the reconstruction in a least-squares sense would be ensured by including those EOFs associated with the largest energy-like content (variance, TE, etc.). However, if the system is transient, then the relative significance of the EOFs may be subject to large variations throughout the time interval spanned by the complete original dataset. This can effectively alter the order of the modes obtained from the analysis of the complete dataset.

An alternative method to circumvent this problem would be to perform the analysis over partial time intervals in order to find the coherent structures relevant to each partial time interval. However, this procedure would prevent us from extracting those modes that might be important in the past or future intervals with respect to the current partial time interval, if they were not significant enough to be detected as one of the most significant modes at this time.

A better alternative is to allow for dynamical EOF indices describing the instantaneous significance of each EOF. The $k$-th dynamical EOF index can be defined as the ordinal number corresponding to the $k$-th EOF after all the EOFs have been re-sorted according to their average TE content in decreasing order. The average TE content in this context should be computed over a given time window. This should be short in comparison with the total time spanned by the complete dataset, but it should be long enough to include the behaviour of the target waves. In this work we have used a time window of 5 sols which is sufficient to observe waves of up to 10 sols if these were perfectly sinusoidal, given that energy is a positive definite quantity.

5 Results

5.1 Whole-dataset analysis

Figure 3 shows the distribution of the energy in the whole dataset. The black curve represents the POD eigenvalues or the average TE explained by each EOF. The amount of energy decreases rapidly from EOF1, which represents the background state and, therefore, explains most TE including unavailable potential energy. The rest of the modes represent the fraction of TE that is actually available for being transformed from kinetic to potential energy and vice versa. In order to represent 90% TE (not including the background energy explained by EOF1), it is necessary to include the first 21 EOF (i.e. EOFs 2 to 22). This is in the same order of magnitude as the number found by Martínez-Alvarado et al. (2009).

EOF2 and EOF5 are seasonal corrections to the background state. These represent seasonal changes to the zonal mean state and circulations near the retreating southern ice cap, respectively. The rest of the modes are combinations of various atmospheric motions such as thermal tides and mid-latitude transient waves.

The diurnal and semidiurnal tides are explained to a large extent by two pairs of EOFs in quadrature. Thus, the diurnal tide is represented by EOF3–4, which accounts for 93.3% TE in this tide. Similarly, EOF6–7 represents the semidiurnal tide, explaining 92.4% TE in this tide. In contrast, transient waves are almost uniformly distributed throughout the modes, with EOFs 13 to 15 being a slightly more prominent group. In order to include 90% TE in transient waves it is necessary to retain 76 EOFs. Other extracted waves are the diurnal non-migrating mode and the westward propagating diurnal wavenumber-3 wave. These are also distributed among a broad range of EOFs. Thus, to account for 90% TE in the diurnal non-migrating mode, 71 modes are required, with EOF24 (10.4%), EOF25 (11.2%) and EOF3 (7.1%) being the most significant. For explaining 90% TE in the diurnal wavenumber-3 wave, 86 EOFs are necessary with EOF36. EOF19 and EOF20 being the most important (jointly accounting for 16.5% TE in this wave).

In the following sections we shall compare these results to those obtained by the analysis of partial intervals in order to establish connections between the initial stage of the planet-encircling dust storm and the changes to wave modes in the atmosphere.
5.2 Energy evolution and dynamical EOF indices

The onset of a planet-encircling dust storm is a highly transient process which brings the atmosphere from a relatively clear atmospheric state, where much solar radiation reaches the surface of Mars, to a state of enhanced global dust loading with strong internal heating of the atmosphere through absorption of short-wave solar radiation.

Following the evolution of the energy distribution we notice that the energy associated to each EOF is far from constant. For example, Fig. 4 shows the energy in the pair EOF6–7 as a function of time, computed as $E_{6–7} = a_k^2 + a_k^2$, where $a_k$ is the $k$-th principal component. It is evident from this figure the dramatic change undergone by this EOF pair, which primarily represents the semidiurnal tidal signal, as the dust storm sets in (which is indicated by the vertical line B). As a result, the EOFs are not expected to have the same significance throughout the time interval spanned by the complete dataset.

Since our interest is centred on the onset of the dust storm, we analysed the interval from $L_s = 182^\circ$ to $L_s = 190^\circ$ which covers the time when the dust storm in the plains to the northeast of Hellas Planitia took place ($L_s \simeq 186^\circ$), followed by the activation of secondary lifting centres on the Tharsis Plateau. This interval is indicated in Fig. 4 by the dashed vertical lines A and C. Before this period, the system was in a relatively steady state. Therefore, we shall define the interval from $L_s = 170^\circ$ to $L_s = 182^\circ$ as our base state which will serve as a point of comparison to understand the changes that originate due to the initial dust outburst.

In order to understand the EOF dynamics during the interval of interest we define a time-dependent EOF index. Just as the whole-dataset EOFs are indexed by the average energy they explain (with the first EOF being the one that explains the largest amount of energy), the dynamical index is determined by re-ordering the EOFs according to the average energy they explain within a short time-window. In this study we have taken a time-window spanning five sols, or 60 sample intervals, sliding in steps of two Martian hours (24 Martian hours = 1 sol). Index 1 is then given to the pattern associated with the largest amount of energy within the time-window. The rest of the EOFs are re-ordered with increasing index as associated energy decreases. Note that for identification purposes the EOFs are still labelled by their index arising from the whole-dataset POD.

According to the evolution of their relative energy significance, the whole-dataset EOFs can be classified in three groups. The first group (low-order EOFs) comprises those modes with an index always below a given threshold. The second group (high-order EOFs) is composed of those modes that always remain above the threshold. A third group (transitional EOFs) is formed by those modes that cross the threshold at some point during the analysed interval. One additional group comprising the seasonal EOFs 1, 2 and 5 was also defined. As mentioned before, these three modes are characterised by very long periods in the order of the seasonal trend. By defining this fourth group we can differentiate these three modes from other modes characterised by shorter periods in the order of a few sols. Fixing the threshold at $I = 14$ there are six low-order EOFs and 16 transitional EOFs. The particular modes in each group can be seen in Table 3.

In the next section we shall investigate the physical interpretation of these various EOF groups. In particular, we will focus on the low-order and transitional EOFs as these are the modes that seem to play a critical role during the onset of the 2001 dust storm.

5.3 Physical interpretation of transitional EOFs

Figure 5a shows the energy percentage distribution in each analysed wave shortly before the dust storm took place. Figure 5b shows this same distribution during the period of interest, i.e. the onset of the planet-encircling dust storm. These figures show that one of the main effects of the initiation of the dust storm was to concentrate wave energy in a few EOFs. For example, before the dust storm onset, the diurnal tide (blue line) was distributed between the pair EOF3–4, accounting for approximately 47% TE and other modes such as EOFs 9 to 13, which jointly contribute with another 45% TE. Once the dust storm was initiated the diurnal tide was mainly
Montabone et al. summarises the distribution of the various atmospheric motions under analysis across the four EOF groups, namely low-order, transitional, seasonal and high-order, before and during the initial phase of the 2001 dust storm. It confirms the observations just described as some atmospheric motions tend to concentrate in a specific group once the initial stage of the dust storm takes place. For example, the semidiurnal tide that was distributed between transitional and high-order modes in an approximate proportion of 7 to 3, finds itself almost completely located in the transitional EOFs during the onset of the dust storm. The diurnal tide experiences a similar effect although this is less noticeable as it is clearly concentrated within the low-order EOFs even before the dust storm has initiated. Thus, it decreases from 30% TE to 5% TE in the transitional EOFs while increasing from 65% TE to 92% TE in the low-order modes.

The case of the diurnal non-migrating wave is interesting as it experiences a slight increase in the low-order modes and a decrease in transitional modes while high-order modes become more important. However, recalling that there are 78 high-order modes while there are only 3 low-order modes and 16 transitional modes, we find that on average the energy represented by a low-order (9.4/3=3.1% TE) or a transitional mode (40.3/16=2.5% TE) is much larger than that represented by a high-order mode (50.1/78=0.6% TE). Thus, the change in the energy distribution for this mode can be interpreted as a transition where low-order modes are enhanced. This gives an indication of the central role the diurnal non-migrating wave plays during the initiation of the planet-encircling dust storm. This is consistent with the results by Montabone et al. (2008) who found that the teleconnection between the dust outburst over Hesperia Planum and the subsequent outbursts over the Tharsis Plateau can be described in terms of constructive interference between the diurnal tide, characterised by a constant phase, and the diurnal non-migrating wave, which sees its phase altered as a consequence of anomalous heating due to enhanced dust loading.

The westward propagating wavenumber-3 wave approximately maintains its structure during this transition with perhaps an enhanced concentration on high-order modes (following the same argument as for the diurnal non-migrating wave). This indicates that this wave plays only a secondary role during the initial phase of the dust storm.

Transients waves are distributed primarily between transitional and high-order modes before the onset of the dust storm. This situation slightly changes during the onset, with transitional modes approximately retaining the same level of significance while seasonal EOFs grow in importance. On the other hand, high-order modes go down from 77% TE before the onset of dust storm to 58% during this process. This reflects the coincidence between the initial stage of the dust storm and the baroclinic activity seasonal change. Before the dust storm onset, remnant baroclinic activity in the Southern Hemisphere was still present at high latitudes around the retreating southern polar cap as spring sets in while activity in the Northern Hemisphere at high latitudes becomes stronger with autumn’s arrival.

Baroclinic activity, though, should not play any role at the latitudes we are considering in this paper, namely the subtropical latitudes.

Table 4 summarises the distribution of the various atmospheric waves (a) before \( L_s < 182^\circ \) and (b) during the onset of the planet-encircling dust storm \( 182^\circ < L_s < 190^\circ \). Low-order modes and transitional-modes are marked with cross and circle, respectively. Grey vertical lines represent high-order modes.

Fig. 5. Energy percentage represented by each EOF for various atmospheric waves (a) before \( L_s < 182^\circ \) and (b) during the onset of the planet-encircling dust storm \( 182^\circ < L_s < 190^\circ \).
Table 4. Energy percentage per EOF group before and at the onset of the 2001 dust storm in the various atmospheric motions under analysis.

<table>
<thead>
<tr>
<th>EOF Group</th>
<th>Before</th>
<th>Onset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diurnal tide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-order</td>
<td>65.2</td>
<td>92.3</td>
</tr>
<tr>
<td>Transitional</td>
<td>29.6</td>
<td>5.3</td>
</tr>
<tr>
<td>Seasonal</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>High-order</td>
<td>5.0</td>
<td>2.1</td>
</tr>
<tr>
<td>Semidiurnal tide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-order</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Transitional</td>
<td>63.9</td>
<td>89.8</td>
</tr>
<tr>
<td>Seasonal</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>High-order</td>
<td>35.9</td>
<td>10.0</td>
</tr>
<tr>
<td>Diurnal non-migrating tide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-order</td>
<td>7.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Transitional</td>
<td>48.8</td>
<td>40.3</td>
</tr>
<tr>
<td>Seasonal</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>High-order</td>
<td>43.6</td>
<td>50.1</td>
</tr>
<tr>
<td>Diurnal wavenumber-3 wave</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-order</td>
<td>5.9</td>
<td>4.9</td>
</tr>
<tr>
<td>Transitional</td>
<td>17.6</td>
<td>15.9</td>
</tr>
<tr>
<td>Seasonal</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>High-order</td>
<td>76.3</td>
<td>79.0</td>
</tr>
<tr>
<td>Transient waves</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-order</td>
<td>1.2</td>
<td>2.7</td>
</tr>
<tr>
<td>Transitional</td>
<td>21.3</td>
<td>35.1</td>
</tr>
<tr>
<td>Seasonal</td>
<td>0.6</td>
<td>3.9</td>
</tr>
<tr>
<td>High-order</td>
<td>76.9</td>
<td>58.3</td>
</tr>
</tbody>
</table>

In the next section we shall investigate the interaction between these various groups of modes. In particular we will study the interaction between the low-order EOFs and the transitional ones, trying to identify the effects from the enhanced heating from the initial phase of the dust storm.

5.4 Physical space evolution

In order to have a physical understanding of the changes in the EOF energies induced by the additional heat source (enhanced dust loading over Tyrrhena Terra and Hesperia Planum), we reconstructed the fields represented by the various EOF groups. The reconstruction method can be found in Martínez-Alvarado et al. (2009). First, however, we present a view of the full dataset (i.e. including every EOF) in order to have a point for comparison. Figure 6 shows a Hovmöller diagram of temperature anomalies at 200 Pa in a latitudinal band between 39° S and the equator. The temperature anomaly field has been reconstructed using the first 100 whole-dataset EOFs. The lower part of the diagram (from \( L_s = 175° \) to \( L_s = 182° \)) corresponds to the period before the dust storm starts. During this interval the behaviour seems fairly periodic, strongly driven by the diurnal cycle. There are persistent alternating extrema around 60° E. There are other persistent alternating extrema around 250° E as well. Maxima and minima at this latter location are stronger than those around 60° E during this preliminary stage. The position of temperature extrema in this (and the next two) Hovmöller diagram might be affected by large-scale topographic features.

Towards the start of the dust storm this situation changes. The first signs of perturbations appear in fact around \( L_s = 182° \) on the three sites with a slight transient growth in the amplitude of alternating extrema. Hence, maxima and minima get stronger and deeper, respectively. However, the most prominent perturbation takes place around \( L_s = 185° \) very close, in time and location, to the occurrence of the maximum in dust lifting over Tyrrhena Terra and Hesperia Planum. This perturbation grows in amplitude and expands preferentially towards the east as indicated by the black solid line in Fig. 6. This expansion seems to enhance the thermal wave activity on the other side of the planet around 250° E.
until it fully merges with this second site at approximately $L_s = 188^\circ$. This is in agreement with the teleconnection event discovered in a global Fourier analysis of surface pressure (Montabone et al., 2008). Part of this effect is due to the dust travelling eastward. Further work is needed in order to determine what amount can be attributed solely to the perturbation generated by the initial dust outburst and to what extent the teleconnection event was due to the direction of transported dust.

Having recognised the presence of the teleconnection event not only at the surface (Montabone et al., 2008) but also at upper levels (in this case at 200 Pa), we can now proceed to investigate how it is reflected in the behaviour of the three-dimensional multi-variable EOFs. For this purpose we analysed the changes undergone by low-order and transitional EOFs separately. Figure 7 is a Hovmöller diagram of temperature anomalies at 200 Pa, where this field has been reconstructed considering low-order EOFs only. A comparison with Fig. 6 shows that the characteristic pattern of the teleconnection is also clearly present in the low-order reconstruction. As in the previous case the lower part of the diagram shows a nearly periodic pattern. Around $L_s = 185^\circ$ a temperature perturbation sets in around the position of Hesperia Planum. This perturbation then starts growing and travelling eastwards until it merges with the oscillating region around $250^\circ$ E.

The features and behaviour just described lead to the conclusion that the teleconnection can be accurately explained in terms of low-order EOFs. However, transitional modes still carry important information. Figure 8 is a Hovmöller diagram of temperature anomalies reconstructed by considering transitional modes only. As in the case of low-order EOFs, the mark of the perturbation appears approximately at $L_s = 185^\circ$. However, in this case the perturbation does not expand eastwards as such. Instead, it appears as though the waves do not change their phase but exclusively their amplitude. This is seemingly important to obtain the correct amplitude of the perturbation as illustrated by Fig. 6.

At this point we can hypothesise the following mechanism for the observed EOF group dynamics.

1. Interference between the diurnal tide and the unperturbed diurnal non-migrating wave gives low-order EOFs a nearly periodic behaviour before the onset of the dust storm. This interference is systematic and independent from the dust storm.

2. The heating anomaly due to enhanced dust load over Tyrrhena Terra and Hesperia Planum produces a perturbation to the phase of the diurnal non-migrating wave and the amplitude of both the diurnal tide and the diurnal non-migrating wave.

3. As a result, the action centres in the pattern generated by the interference between the diurnal tide and the diurnal non-migrating wave (originally located around $60^\circ$ E) experience an effective displacement towards the east.

4. The combined response of low-order and transitional EOFs to the initial dust outburst around $60^\circ$ E yields a new state where the original centre of action at this position has expanded eastwards while the centre of action at approximately $250^\circ$ E is intensified favouring the existence of secondary dust lifting points around this position.

Future work will address the origin of the excitation of the transitional EOFs and their connection to the dynamics of the event.
6 Discussion and final remarks

We have shown how the inherently transient process that lead to the development of a planet-encircling dust storm on Mars can be described by a small number of modes according to the average TE explained by these modes throughout the process. These modes were computed using a fully-three-dimensional multi-variable version of POD. Moreover, we have defined time-dependent EOF indices to account for changes in time in the energy explained by each EOF relative to the rest of the modes. We have focused our analysis on the initial phase of the development of the 2001 planet-encircling dust storm.

The main conclusions of this work are two-fold. First, the response of the atmosphere to a localised source of heating is restricted to a small set of global three-dimensional modes. This can be taken as a supporting statement regarding the hypothesis of a low-dimensional Martian climate attractor explored by Martínez-Alvarado et al. (2009). In that work, 20 EOFs were necessary to explain 90% TE during the analysed period. However, one standing question was to what extent high-order modes have influence on the dynamics of low-order modes. In this work, we have shown that the classification of EOFs can be extended by including a third class of modes, namely transitional EOFs. These are excited or suppressed in the presence of an anomalous localised forcing. For completeness, a fourth class would be given by seasonal EOFs. These are a few modes explaining a large amount of energy whose variation is linked to the seasonal cycle. Low-order EOFs seem to be enough to account for the essential features of the transient teleconnection event. The diurnal migrating and non-migrating modes are included to a large extent in this group of EOFs. This results confirm that these are the main components that explain the transient teleconnection as previously seen in surface pressure analysis by Montabone et al. (2008). Transitional EOFs constitute a correction to the description given by low-order modes. By including the information explained by transitional EOFs the accuracy in the estimation of the magnitude and position of the disturbance is significantly improved. High-order EOFs remain at the lower end of the energy spectrum and appear to have little influence during the development of the phenomenon under analysis. Low-order and transitional EOFs form together a set of 20 modes, a number that is comparable with that previously found by Martínez-Alvarado et al. (2009).

Similar results have been found for the Earth in the work by Achatz and Branstator (1999); Achatz and Opsteegh (2003). These authors found that the number of EOFs required to explain 90% TE in a quasi-geostrophic, perennial January simulation was approximately the same as that required to explain the same amount of TE in a primitive-equation simulation including the seasonal cycle. The number, however, was much larger than for Mars, being in the order of a few hundreds (Achatz and Opsteegh, 2003).

A second conclusion from the present study is that EOFs are capable of accounting for localised effects despite their inherent global character. The joint evolution of the low-order EOFs clearly show two centres of action around 60°E and 250°E. These two centres of action are perturbed once the first major regional dust storm sets on. The effect of these perturbation is an increase in amplitude and a change in the position of at least one of the action centres (the one around 60°E). When joining this information to the results from spatio-temporal STFA, the response to the perturbation can be understood as a modification in the behaviour of certain important thermal tides.

These results are robust in the sense that analyses of shorter and longer time series around the time of the onset of the dust storm have produced the same qualitative results. Small differences have been found in the position and distribution of the different atmospheric waves. However, the classification of EOFs in low-order, high-order, seasonal and transitional modes can still be done. Moreover, the behaviour
and role of each group remains essentially the same.

As reported in Sect. 3, the kind of teleconnection described in this article is extremely efficient on Mars because of its thin atmosphere, the crucial role played by the mineral dust in the atmospheric forcing, the strong thermal tides and the lack of water vapour that could limit the growth of dust clouds. Nevertheless, this study is not meant to be exclusively related to Mars. On the contrary it is meant to use Mars as an alternative natural laboratory to develop concepts and techniques that can be applied to the study of dust storms and related phenomena on our planet.

On the Earth, the thermal tides are not good candidates to propagate localized thermal anomalies in the troposphere. Their impact with respect to other components of the climate system is not as strong as on Mars given the large thermal inertia of the lower-atmospheric and oceanic reservoirs. Nevertheless, equatorial waves, such as Kelvin waves, and baroclinic waves at mid-latitudes can constitute good examples of waves whose structures might be modified by strong thermal anomalies, favouring the propagation of teleconnection signals.

Given the possibility of progressive climate change on the Earth and the observed increasing number of large dust storms, the effects of strong thermal anomalies due to the presence of dust or other radiation-absorbing aerosols might play an even larger role on our planet in the future, as they might have done in drier climates in the distant past. Even though on the Earth, dust storms are not the phenomena which produce the strongest, localised thermal anomalies, there are other processes for which this analysis might be relevant. For example, super-volcanic eruptions in the distant past of our planet have possibly injected in the atmosphere amounts of ashes that might have caused thermal anomalies comparable to those observed on Mars during a planet-encircling dust storm at the present time (Harris, 2008). More generally, this analysis of transient teleconnection events could be applied to the study of short time scale, long-range effects of any phenomenon that produces an anomalous, localized heating or cooling, where “short time scale” refers to a suitable scale for the planet and region of the atmosphere under consideration (since, for example, on the Earth the radiative time constant is much shorter in the stratosphere than in the troposphere). Finally, this analysis might also be a useful tool to identify links between teleconnection events and teleconnection patterns, possibly helping to establish causal relationships for the latter.

Acknowledgements. O.M.-A. thanks the Mathematical Institute, University of Oxford, for providing support for the development of this work. L.M. acknowledges support from the NASA contract “Mars Critical Data Products Initiative”. The authors gratefully acknowledge the use of MGS/TES temperature and dust opacity retrievals provided by B. J. Conrath, J. C. Pearl and M. D. Smith of Goddard Space Flight Center, Greenbelt, Maryland (USA). The authors also acknowledge support for the publication of this work by the NASA Grant: NNX06AF30G – Global Environmental Change-hazards and regional impacts, and support by the technical officers Don Anderson and Lucia Tsousi.

Topical Editor F. D’Andrea thanks two anonymous referees for their help in evaluating this paper.

References

O. Martínez-Alvarado et al.: Transient teleconnection event during a Martian dust storm


1975.


