RMetS
Royal Meteorological Society

Assessing atmospheric predictability on Mars using numerical weather prediction and data assimilation

P. Rogberg,[a]*P. L. Read,[a] S. R. Lewis[b] and L. Montabone[b]
[a]*Department of Physics, University of Oxford, UK*
[b]*Department of Physics and Astronomy, CEPSAR, The Open University, UK*
*Correspondence to: P. Rogberg, Department of Physics, Parks Road, Oxford OX1 3PU, UK.
E-mail: p.rogberg1@physics.ox.ac.uk

The intrinsic and practical predictability of the Martian atmosphere is determined by use of a comprehensive numerical circulation model and ensemble forecasting techniques. Initial conditions were derived for various seasons of the Martian climate from newly available assimilations of global observations of temperature and dust optical depth in the Martian atmosphere, obtained by the Thermal Emission Spectrometer (TES) instrument on the *Mars Global Surveyor* orbiter. A range of initial value forecasts were obtained and the growth rate of forecast errors measured against the assimilated re-analysis were determined as a function of season during the period from the late Martian year 24 to early 27 (from May 1999 to August 2004). The intra-ensemble growth rates confirm earlier stand-alone model results, indicating that internal baroclinic transients lead to significant chaotic growth of model 'error' on time-scales of 3–10 sols only during Northern Hemisphere autumn and winter seasons. At other seasons, ensembles retain coherence at least over several tens of sols. Imperfections in dust distribution and coupling with other aerosols can lead to large errors in the dust field when it evolves rapidly. Owing to the very short time-scale on Mars, these errors can lead to rapid divergence of the entire ensemble from the observed atmospheric state. These results demonstrate the importance of even relatively small errors in the aerosol fields and so in the computed radiative balance of the Martian atmosphere for deterministic forecasts throughout the Martian year, including times at which baroclinic transients are inactive.

Key Words: ensemble forecasting; Martian atmosphere; *Global Surveyor*; Thermal Emission Spectrometer

Received 09 February 2009; Revised 10 June 2010; Accepted 18 June 2010; Published online in Wiley Online Library 31 August 2010

Citation: Rogberg P, Read PL, Lewis SR, Montabone L. 2010. Assessing atmospheric predictability on Mars using numerical weather prediction and data assimilation. *Q. J. R. Meteorol. Soc.* **136**: 1614–1635. DOI:10.1002/qj.677

1. Introduction

Since early observations of the atmosphere of Mars from ground-based telescopes, it has been clear that Martian meteorology is highly variable on a wide range of space and time scales. The range and extent of this variability has come into sharp focus in the past decade with the arrival of a flotilla of spacecraft, including NASA's *Mars Global Surveyor* (MGS), *Mars Odyssey* and (most recently) *Mars Reconnaissance Orbiter*, and ESA's *Mars Express* orbiter. Surface observations, too, from the *Viking* (VL), *Pathfinder* and the *Spirit* and *Opportunity* lander spacecraft (Barnes, 1980, 1981; Schofield *et al.*, 1997; Smith, 2004; Smith *et al.*, 2006) have allowed detailed insights into the nature of Mars' variable atmosphere close to the surface. With earlier missions aimed towards low latitudes

(VL1 and *Pathfinder*) and at midlatitudes (VL2), there is now a focus on the meteorology nearer the poles with the latest landing at 68°N (*Phoenix*).

Together with studies using detailed numerical models of the Martian atmosphere, it is clear that this variability on large spatial scales is dominated by the seasonal cycle, and by components of the general atmospheric circulation, such as the thermal tides driven by the diurnal cycle or insolation, baroclinic transient weather systems and local, regional and planet-encircling dust storms (e.g. Read and Lewis, 2004).

Some of these modes of variability are externally forced in a periodic manner (by the seasonal and diurnal cycles), and are hence highly predictable (at least in principle, though details of the atmospheric response may depend on complex nonlinear feedbacks, interactions with topography, etc). Other modes, however, are highly nonlinear and complex, leading to chaotic and/or turbulent behaviour which poses substantial challenges to atmospheric models to capture and predict realistically. While there is (as yet) no operational imperative for atmospheric modellers to provide deterministic forecasts of atmospheric conditions on Mars, a comparison of the intrinsic and practical predictability of Martian and terrestrial meteorology is of significant interest from a scientific viewpoint. Such predictions may also eventually be of practical significance for future robotic and even manned operations at the Martian surface.

Studies of the time series of surface measurements of wind, pressure and temperature at the two *Viking* landers (Barnes, 1980, 1981) revealed that baroclinic transient travelling waves on Mars occur mostly during Northern Hemisphere (NH) autumn, winter and early spring, and typically take the form of highly coherent patterns with planetary wavenumbers 1–3 that can persist for intervals of up to 30–60 sols before changing erratically. Such behaviour is almost unknown on Earth, where individual baroclinic weather systems typically persist for no longer than 5–10 d and seldom remain coherent around entire latitude circles. This occurrence of planetary-scale coherent baroclinic wave-like weather systems on Mars led to suggestions (Leovy, 1985) that Mars' atmospheric circulation operates in a quite different dynamical regime to that of the Earth, one that tends to favour regular, symmetrical baroclinic wave activity in a manner reminiscent of the regular wave regimes found in laboratory fluid dynamics experiments on sloping convection in a rotating, thermally driven fluid annulus (e.g. Hide and Mason, 1975; Read *et al.*, 1998). In its extreme form, this hypothetical comparison would suggest the possibility of a highly regular atmospheric circulation on Mars, though subsequent modelling work (Collins *et al.*, 1996) indicated that perturbations due to the thermal tide can lead to chaotic transitions back and forth between different intransitive wave states. This form of chaotic mode-flipping appeared to be consistent with the *Viking* observations of Mars, suggesting nevertheless that the intrinsic predictability of Mars' midlatitude meteorology was qualitatively and quantitatively quite different from that of the Earth.

Most recently, Newman *et al.* (2004) have attempted to quantify the intrinsic predictability of the large-scale atmospheric circulation on Mars (at least in the absence of significant dust storms), and its seasonal variability, by applying methods of ensemble weather prediction and 'breeding vectors' (Toth and Kalnay, 1993) to numerical simulations of Martian meteorology. Such methods directly address the sensitivity of chaotic atmospheric behaviour to initial conditions by running a set of simulations in parallel from a closely similar (though non-identical) set of initial conditions, typically derived by perturbing a fully three-dimensional atmospheric state obtained from a previous simulation. Breeding vectors (Toth and Kalnay, 1993) were used as a means of optimising the form of perturbation applied to the initial conditions in order to find those perturbations that would grow most rapidly and lead to the strongest divergence of the set of parallel forecasts. Newman *et al.* (2004) found that forecast ensembles would typically only diverge during NH autumn and winter, with e-folding time-scales of 3–8 sols. The patterns of divergence were consistent with chaotic behaviour dominated by the action of planetary-scale baroclinic waves at middle and high northern latitudes with wavenumbers 1–3, much as are found to dominate the actual circulation during these seasons. At other seasons, perturbations to the forecasts were typically found not to grow, provided the perturbation was sufficiently large, suggesting that, at least on large scales, the Martian atmosphere was non-chaotic (even in the presence of diurnal tides, topography and other complex features). During the latter seasons, however, baroclinic activity in the NH was virtually absent, although some weaker baroclinic transients were found during the Southern Hemisphere (SH) autumn and winter seasons. These did not, however, appear to lead to sustained chaotic divergence of model forecasts.

A major limitation of the previous study by Newman *et al.* (2004) was that it took no account of modes of variability other than internal nonlinear interactions and external periodic forcing. In particular, the dust distribution in the model simulations was fixed, so that, in the absence of baroclinic wave activity, the simulated meteorology was relatively quiescent in response to the highly predictable (seasonal and diurnal) external forcing. In the present study, therefore, we have extended this ensemble prediction approach to investigate the intrinsic *a*nd practical predictability of a more realistic system, as represented in assimilated analyses of observations of Mars itself (Lewis *et al.*, 1996; Montabone *et al.*, 2006; Lewis *et al.*, 2007). This allows us to obtain more complete and realistic atmospheric states from which to initialize deterministic forecasts, for which variability of dust and other aspects not included in the baseline version of the model used at Oxford University and the Open University in the UK (UK-MGCM) is taken into account. In this paper, therefore, we examine the typical growth rates of perturbations in ensemble predictions of Martian meteorology, initialised from observed states and predicted using more realistic distributions of atmospheric dust, again derived from observations. Error growth is considered both with respect to the intra-ensemble spread as the forecast progresses, and with respect to observations subsequent to the initial state act as an observational verification of the forecast. The latter turns out to be a particularly stringent test of model performance, and serves to expose a number of systematic errors in the stand-alone model climatology.

In section 2 we describe the essential features of the Mars general circulation model, UK-MGCM, and briefly review the assimilated analysis dataset based on the observations of temperature and dust from the Thermal Emission Spectrometer (TES) on board the MGS spacecraft. Section 3 is an examination of the growth of perturbations within ensembles of forecasts, relative to a reference forecast,

and its seasonal variability, while section 4 discusses the performance of the model forecasts relative to observational verifications from the assimilated record. Some conclusions and further discussion are presented in section 5.

2. Methodology

2.1. Mars general circulation model

The model, UK-MGCM, consists of a dynamical core which uses a spectral decomposition into spherical harmonics in the horizontal direction and a finite-difference discretization in the vertical, together with a comprehensive set of parametrizations of subgrid-scale physical processes, and is essentially the same as described by Newman *et al.* (2004) and Forget *et al.* (1999). The model resolution was fixed to T31 (triangular truncation to total horizontal wavenumber 31), corresponding approximately to a dynamical grid resolution of 96×48 in longitude and latitude, and for linear physical processes to grid resolution of 72×36. The vertical resolution used 25 levels with a terrain-following σ-coordinate, arranged on a non-uniform grid over the approximate height range 5–100 km. The physical processes parametrized include radiative transfer, gravity-wave and low-level form drag, a ten-layer soil scheme, turbulent vertical diffusion and convective adjustment, CO_2 condensation and sublimation. The model can also include dust lifting, atmospheric transport and deposition, though this was not used in the present study. Detailed descriptions of these can be found in Forget *et al.* (1999) and Newman *et al.* (2002), together with studies of their verification against past observations. The boundary conditions are based on Mars Orbiter Laser Altimeter (MOLA) topography (Smith *et al.*, 1999) and a map of surface thermal inertia and albedo was adapted from Paige *et al.* (1994) and Palluconi and Kieffer (1981).

With the aim of minimizing model-induced error sources, the free-standing model simulations were run with parameter values closely following those used during the assimilation (see below). These were based on the 'MGS scenario', designed for recent versions of the European Mars Climate Database (Forget *et al.*, 2001). Since the TES observations provide only profiles of temperature and a measure of vertically-integrated total dust optical depth, without invoking any dust lifting or transport, we chose to employ in each experimental forecast a prescribed dust distribution in which the horizontal distribution was adapted to be consistent with the observed variations of total optical depth at the beginning of the forecast interval, with an explicitly formulated distribution prescribed in the vertical. The dust distribution was not 'frozen' but evolved in time. The vertical profile was based on the profile used by Forget *et al.* (1999) and Montabone *et al.* (2006), according to the following equation, which is used to describe the vertical distribution of the dust at a given latitude and time in terms of volume mixing ratio q (cf. Forget *et al.*, 1999):

$$q = q_0 \exp\left(a\left[1 - (p_0/p)^{\left(\frac{b}{z_{max}}\right)}\right]\right) \qquad (1)$$

for pressure $p \leq p_0$, where p_0 is a reference pressure, taken to be $p_0 = 700$ Pa, and with $q = q_0$ for $p > p_0$. q_0 may be prescribed for each longitude, latitude and time, and a and b are empirically determined parameters with the typical values $a = 0.007$, $b = 70$ km. $z_{\max}$ varies with aerocentric longitude L_s and latitude ϕ according to the equation:

$$z_{\max} = 60 + 18\sin(L_s - 160^\circ) - \sin^4\phi\,[32 + 18\sin(L_s - 160^\circ) - 8\sin(L_s - 160^\circ)\sin\phi]. \qquad (2)$$

This dust distribution was found to give a good statistical agreement with both radio occultation profiles and TES retrievals from MGS (Forget *et al.*, 2001; Montabone *et al.*, 2006). The time step for the dynamical part of the model was 3 min while the physical processes were updated every 30 min. However, to match the re-analysis (see below), the model data were stored with a 2 h time step.

2.2. Assimilation

For this study, a detailed synoptic analysis of observations of Mars was used to define a set of initial states, from which an ensemble of deterministic forecasts was run using the UK-MGCM described above. The analysis was obtained by assimilation of thermal profiles and total dust opacities which were retrieved from nadir soundings of the Martian atmosphere between the surface and about 40 km altitude, obtained by TES on board MGS during the scientific mapping phase Conrath *et al.* (2000) and Smith *et al.* (2000) provide details of the retrieval technique, and Lewis *et al.* (2007) details of the assimilation scheme.

Assimilation of TES nadir retrievals into the UK-MGCM (Forget *et al.*, 1999) has been carried out for the entire MGS period of almost three complete Martian years (MYs), as described above, with global, three-dimensional, multivariate fields being stored at two-hourly intervals throughout this period. This extensive dataset is now being used to provide regular analyses of the state of the Martian atmosphere for scientific studies, such as of the interannual variability of dust storms (Montabone *et al.*, 2006) and the atmospheric thermal tides (Lewis and Barker, 2005). Such analyses are fully three-dimensional, dynamically balanced states that are consistent with the parent model dynamics and physics. As such, they are well suited directly to provide initial states from which to launch a deterministic forecast, much as is done operationally for the Earth at many weather forecasting centres.

In the present study, initial states were selected from the full assimilated dataset from MY 25 to 27, which include a year with a dust storm event as well as 'non-storm' years. (By a non-storm year, we mean a year with no planet-encircling storm.) To compare with intra-model runs, we focused on two years 'MY 25 (dusty) and MY 26 (non-storm)' where we employed ensembles of forecasts which were then run from each of these initial states, as described below.

2.3. Model initialisation and ensemble generation

With the aim of assessing the model's ability to predict the atmospheric state at different times of the year, we employed the same division as used in the re-analysed data. This means that we use data subsets each extending over a period of 30 sols. This also forms a natural upper limit on the time-scale of these experiments. Other studies, e.g. Wilson *et al.* (2002), have found the longest period of the travelling waves to be $\sim$ 20 sols, which suggests that a time window of 30 sols

will encompass most transient atmospheric features short of seasonal time-scales. Following Newman *et al.* (2004), each year is divided into 12 equal sized periods (i.e. they extend over intervals of $L_s = 30^\circ$). At the beginning of each period, we use data from the re-analyses to initiate the model, which then simulated the atmospheric evolution for 30 sols. This forms our control run. To generate an ensemble of initial perturbed states we used the approach of Houtekamer and Derome (1994) in constructing pairwise sets of two-member ensembles, i.e. initial states were produced in pairs by respectively adding and subtracting a given perturbation pattern to the basic analysed field. Houtekamer and Derome (1994) showed that using symmetrically distributed initial states is an efficient way of creating initial states for ensemble members in that they tend to enclose the 'true' solution and thus minimise the ensemble size.

Newman *et al.* (2004) used a different method, by employing a perturbed ensemble forecasting technique with a global circulation model. In their study, a reference initial state was obtained as the end-point of a previous stand-alone model simulation. The initial state for each ensemble member was then modified from the reference state by a bred vector perturbation, obtained as the result of a series of short model runs initialised from the reference state following a perturbation with random noise of a given r.m.s. amplitude ($\sim$ 5 K in the temperature field). Bred vectors are obtained by successively re-normalising the amplitude of growing perturbations over a prescribed time interval (Toth and Kalnay, 1993, 1997).

In the present study, we used an ensemble of perturbed initial states, though we did not find it necessary to use the bred vector technique to re-normalize the error growth. Various tests using simpler random-noise perturbations were found to be almost as effective as full breeding vectors in producing perturbed initial states that would diverge in the presence of chaotic dynamics. Thus, initial states within an ensemble were created by adding a uniform distribution of random noise of r.m.s. amplitude 5 K and zero mean to the atmospheric temperature. In order to obtain symmetrically distributed initial states, we used the approach of Houtekamer and Derome (1994) in constructing pairwise sets of two-member ensembles, i.e. initial states are produced in pairs by respectively adding and subtracting a given perturbation pattern to the basic analysed field.

In quantifying the ensuing forecast, the major diagnostic variables used to quantify the divergence of the model trajectories were the temperature at field at 30 Pa, and the mean surface level pressure $p_{\rm msl}$. The latter is defined as the pressure hydrostatically interpolated (or extrapolated) to the reference areoid height of 0 km (much as mean sea level pressure is computed on Earth). In practice, this can be computed approximately as

$$p_{\rm msl} = p_{\rm s} \exp\left(\frac{g\, z_{\rm s}}{R T_r}\right), \tag{3}$$

where $p_{\rm s}$ is the surface pressure, $g\, z_{\rm s}$ the areopotential height of the surface, R is the gas constant and $T_{\rm r}$ a representative atmospheric temperature, taken to be that applicable at a height of $\sim$ 600 m above the actual surface (Spiga *et al.*, 2007). This has the effect of removing the first-order effects of topographic variations in the surface pressure field, leaving a component dominated by the dynamically varying meteorology.

Figure 1 shows the time-evolving variations of a set of variables at a particular location (close to that of the *Viking* Lander 2 (VL2) spacecraft: 48°N, 226°W) in an ensemble of simulations, initialized from a single assimilated analysis of the atmospheric state in northern mid-winter ($L_s = 300^\circ$ in Mars year 26) and randomly perturbed in their temperature fields. From these variables, the growth of perturbations is clearly apparent over a time-scale of 5–20 sols. The relative growth of perturbations is largely similar in each variable, though appears to be somewhat more rapid in $p_{\rm msl}$. This may reflect the superposition in the surface pressure field of effects due to both deep and shallow weather systems as they pass over the VL2 location. Deep disturbances in the atmosphere are less apparent at the surface in variables other than $p_{\rm msl}$, possibly reflecting differing growth rates or vertical structures (i.e. 'external' or 'internal' modes; Wilson and Hamilton, 1995). Accordingly, in the following discussions we analyze time series of both $p_{\rm msl}$ and temperature obtained at 30 Pa (T_{30}), corresponding to an altitude of around 30 km above the surface. In order to determine the spread of ensembles of initial states during the course of parallel forecast runs, we compute the r.m.s. difference e.g. in temperature

$$\sigma^{\rm I}_{\Delta T_{30}} = \sqrt{< c_j (\Delta T_j)^2 >}\,, \tag{4}$$

where $\Delta T = T^m(t) - T^i(t)$, T^m is the ensemble prediction and T^i represents either the assimilated analysis $T^{\rm a}$ or 'control' model simulation $T^{\rm c}$, and c_j is the area weight at grid point j. In most cases this was computed as an r.m.s. average over either the whole surface of Mars or within a given latitude band.

3. Results: Intra-ensemble error growth

Two years, MY 25 and MY 26 (which were the second and third full years of MGS operations) are considered in this part of the study. The two years are different in that a series of dust storms occurred in MY 25, beginning around $L_s = 180^\circ$ and for some time a veil of suspended dust covered all but the polar regions. The next year, MY 26, on the other hand did not exhibit any major dust storm event.

We first consider the seasonal variation in forecast sensitivity to initial conditions, using ensembles of initial conditions perturbed from the assimilated analyses of MGS TES data. To examine the seasonal variability, the year was divided into 12 periods of roughly the same size. Figures 2 and 3 show the evolution of $\overline{\sigma}^{\rm I}_{\Delta T_{30}}$ from ensemble-averaged forecasts, with each ensemble member being perturbed about an unperturbed reference forecast started at the beginning of each period. As mentioned above, each ensemble was integrated forward in time for up to 30 sols and the divergence of the ensemble about the (unperturbed) baseline was monitored. Each figure shows the global-averaged divergence and the three latitude bands corresponding to 30°–90°N (NH), 30°S–30°N (EQ) and 30°–90°S (SH). In some cases, especially in the later half of the year, there is a clear difference between the three bands, in which the NH band displays a rapid growth rate compared to the equatorial and southern bands. This reflects the extent to which localized baroclinic transients dominate error growth in their respective storm zones.

Figure 4 shows two examples of these r.m.s. statistics in the sample ensemble of forecasts illustrated in Figure 1, just

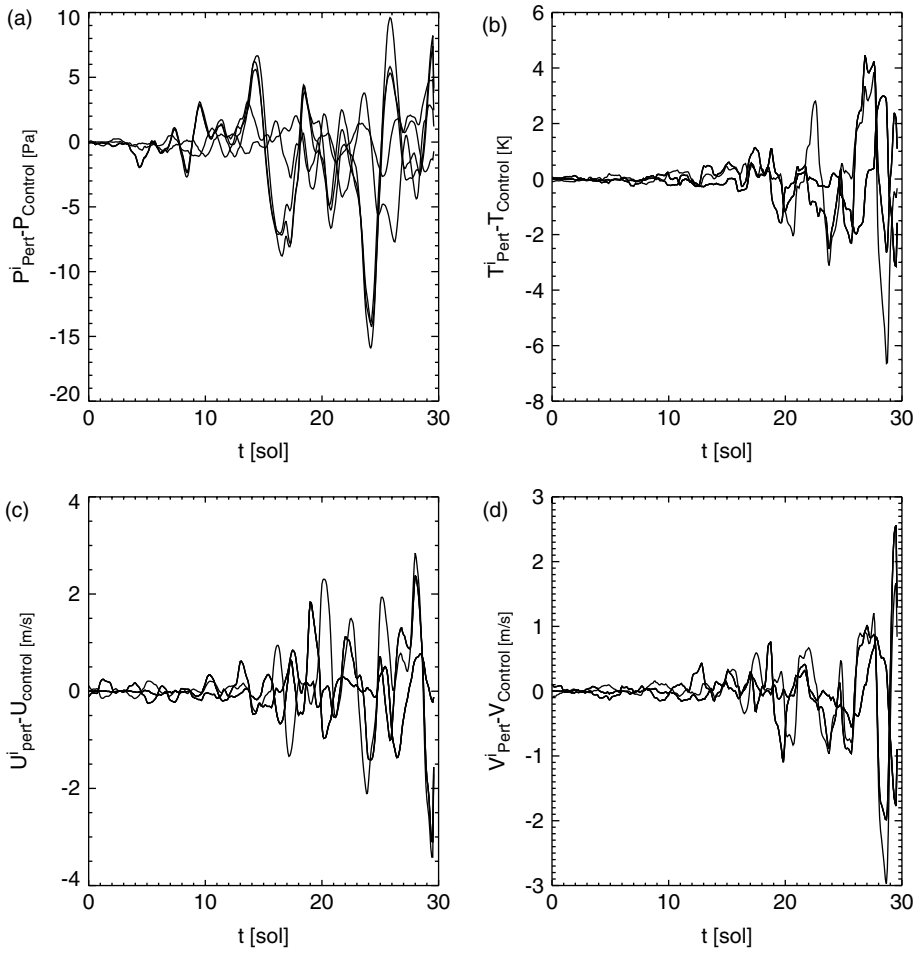

Figure 1. The variation of meteorological variables interpolated to the surface and at $p = 30$ Pa at the location of the *Viking* Lander 2 spacecraft (48°N, 226°W), from an ensemble of six model simulations, perturbed from an initial state interpolated from an assimilated analysis of MGS TES profiles of temperature and dust optical depth in northern mid-winter. The initial states correspond to $L_s = 300°$ in MY 26, and were randomly perturbed in their temperature fields (see text): (a) p_{msl}, (b) T (at $p = 30$ Pa), (c) u at $z \sim 5$ m and (d) v at $z \sim 5$ m.

after the spring equinox and before the NH winter solstice. The envelope of this variability again exhibits clear growth, at least for the first 5–10 sols, after which the growth slows and begins to saturate. This appears to be typical for most cases where perturbations are found to grow in the first 10 sols or so, indicating nonlinear saturation of the initial instability that leads to sensitivity to initial conditions, much as found on Earth.

Only a very weak divergence between ensemble members and the baseline simulation was found to take place during the interval $L_s \approx 20°$ to $160°$ (Figure 2), corresponding to NH summer. The divergence rate only begins to become significant from $L_s \approx 180°$ onwards, reaching a maximum shortly before the winter solstice, after which a decline begins. Figure 5 shows intense wave activity during $L_s = 237°$ to $257°$ of temperature amplitudes in the meridional plane. The strongest amplitudes correspond to modes with zonal wave numbers $m = 1$ and $m = 2$ with periods 4.29 and 15 sols. These waves have earlier been observed by Barnes (1981) and Newman *et al.* (2004), where the waves with longer periods were identified as Rossby waves (Hollingsworth and Barnes, 1995). Examining results from MY 25 and 26 reveals a consistent pattern of eastward travelling waves with high amplitudes during periods with rapid linear divergence while periods with slow divergence correlate with only relatively weak wave activity.

Table I. Time-averaged growth rate, $\overline{\alpha_A}$, averaged over the entire period.

Latitude band	$\overline{\alpha_A}$ (sol^{-1})
NH	0.47 ± 0.19
EQ	0.61 ± 0.18
SH	0.58 ± 0.17
Global	0.57 ± 0.13

This suggests that the rate of divergence in model simulations is strongly correlated with baroclinic wave activity.

In contrast, during NH late spring and summer the ensemble growth is virtually negligible, at least over 30 sols. This seasonal behaviour corresponds well qualitatively with the results found by Newman *et al.* (2004, their Table I) for stand-alone model simulations with prescribed dust distributions, which also found positive global growth rates in r.m.s. streamfunction perturbations during late NH autumn.

Figure 3 shows the divergence of the ensemble during MY 25. The overall pattern closely resembles MY 26, in that the divergence grows slowly during the first half of the year compared with the second half. The initial divergence during sols 1–4 during NH autumn and winter, $L_s = 210°$ to $300°$,

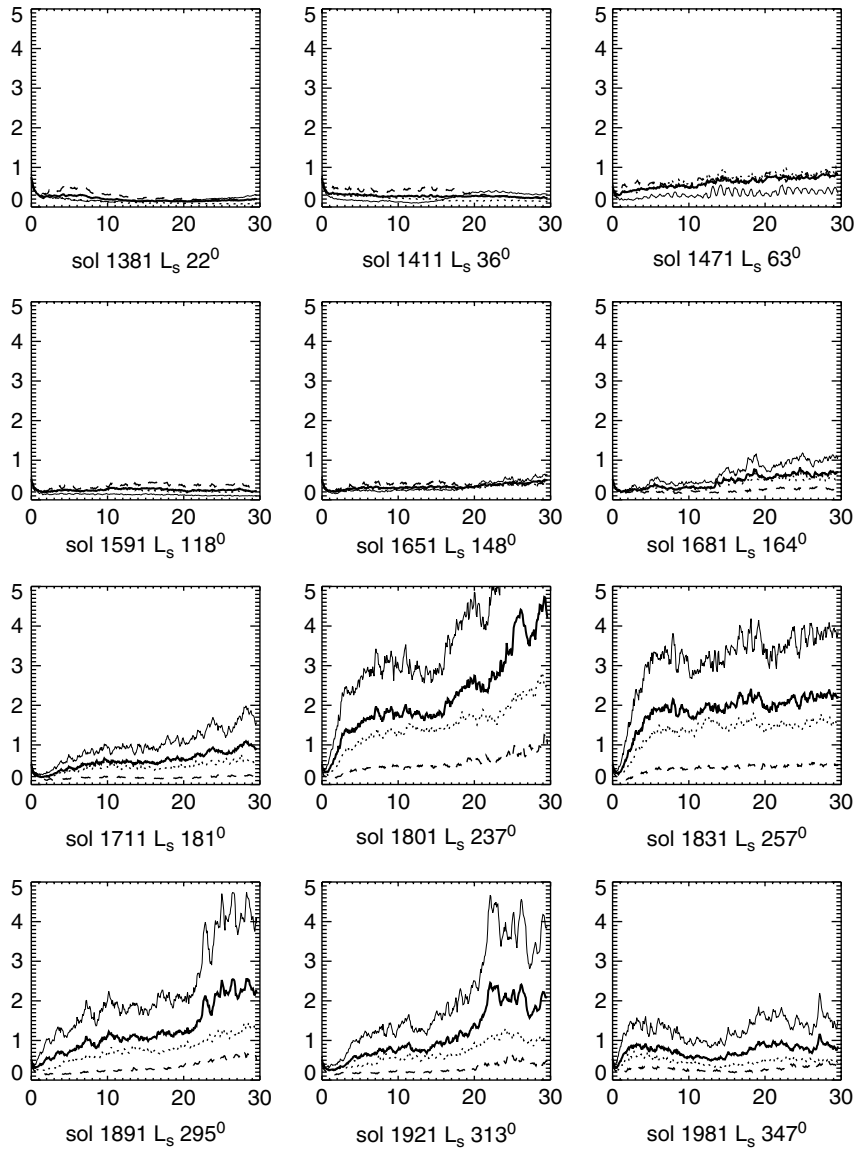

Figure 2. The variation of $\overline{\sigma}_{\Delta T_{30}}$ (K), averaged over the globe (bold solid line), NH (thin solid line), EQ (dotted) and SH (dashed) during twelve periods in MY 26, from an ensemble of six model simulations, each perturbed from an initial state obtained from an assimilated analysis of MGS TES profiles of temperature and dust optical depth. The start date of each simulation is shown beneath each frame.

are similar for the two years. The most notable difference is seen during NH spring and summer, which shows a larger divergence. MY 25 was characterized by a series of large dust storms that begun around $L_s = 180^\circ$ and persisted to about $L_s = 200^\circ$. The model response in terms of ensemble divergence is hardly detectable.

3.1. *Intra-ensemble growth rates*

The early rate of divergence was estimated by calculating the daily growth rate of $\sigma^I_{\Delta T_{30}}$, approximated by

$$\sigma^{\mathrm{I}}_{\Delta T_{30}} = A \exp(\alpha_{\mathrm{I}}\, t)\,, \tag{5}$$

where α_{I} is estimated by linearly regressing $\ln\{\sigma^{\mathrm{I}}_{\Delta T_{30}}(t)\}$ over 1 sol.

The mean daily growth rate, $< \alpha_{\mathrm{I}} >$, for each of the 12 'seasons' of MY 25 and MY 26 is shown in Figure 6, together with the corresponding values obtained by Newman *et al.* (2004). The values from the forecasts initialised from the re-analysis dataset (discussed in the Introduction) are significantly larger than those obtained by Newman *et al.* (2004); a rapid increase of the growth rates takes place after the NH autumn equinox and the rate reaches its maximum just before the NH winter solstice. Negative growth rates are found after the spring equinox. In contrast to Newman *et al.*, we found a local maximum around the summer solstice; a possible explanation is that the choice of diagnostic parameter, $\sigma_{\Delta T_{30}}$ is sensitive to perturbations in the temperature field and the increase at times coincides with wave activities seen in Figure 7.

It is clear from Figures 3 and 2 that periods such as the NH spring and summer are less difficult to predict than other times of the year. Earlier model studies have found regular baroclinic transients occurs during most of the year, with the exception of NH summer (Collins *et al.*, 1996). Figure 7 shows the wave amplitude around $L_s = 26^\circ$ to 40° in year 25 for model and re-analysis; the largest differences are seen for waves with wave numbers $k = 2$ and 3 with a period of $\approx$4.29 sols.

The seasonal variability of predictability is not confined to atmospheric temperature, as illustrated in Figure 6(a).

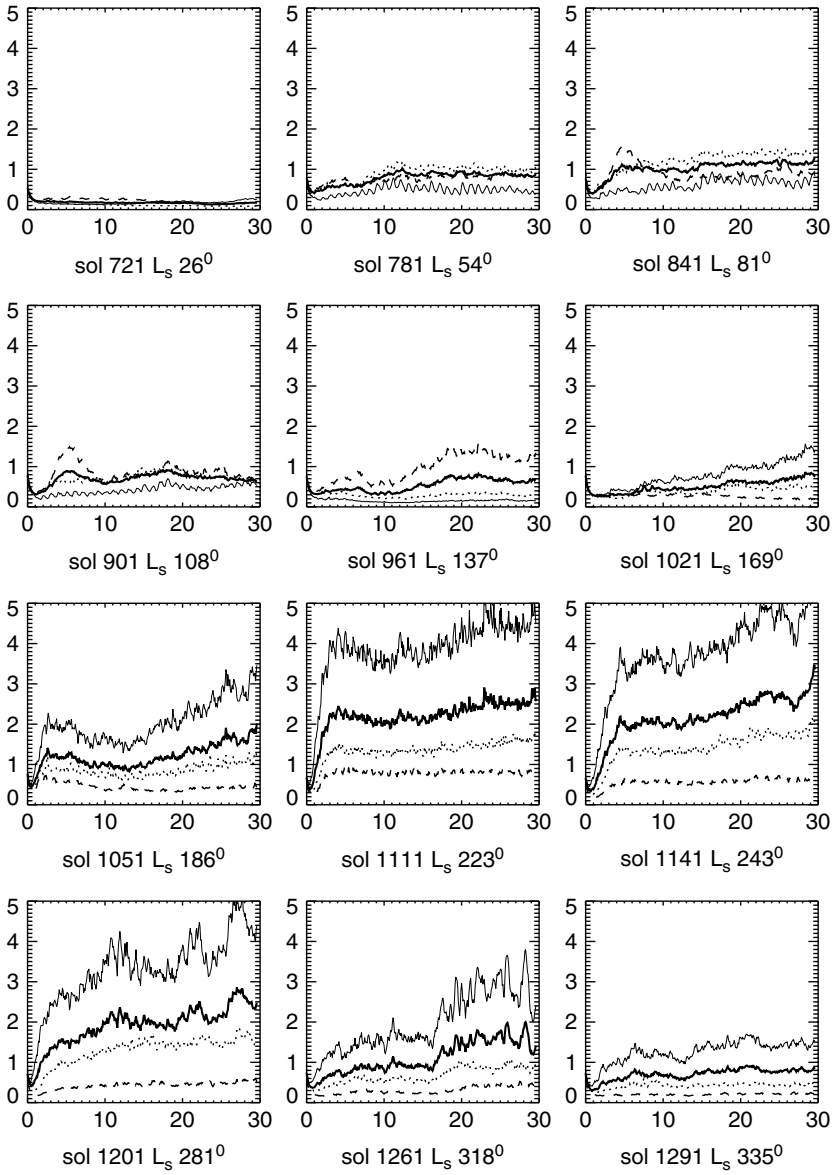

Figure 3. As Figure 2, but for MY 25.

This figure shows the growth rate of the surface pressure, $\alpha^{\mathrm{I}}_{\Delta p_{\mathrm{msl}}}$, where the pressure is hydrostatically corrected to the mean Martian aeroid. The grow rates have been calculated similarly to $\alpha^{\mathrm{I}}_{\Delta T_{30}}$. The divergence pattern (not shown here) shows that individual ensemble members separate earlier, which leads to higher uncertainty. Another difference is the period around the spring equinox in which $\sigma^{\mathrm{I}}_{\Delta p_{\mathrm{msl}}}$ is less predictable than $\sigma^{\mathrm{I}}_{\Delta T_{30}}$.

4. Results: Model assimilation forecast verification

To assess the practical atmospheric predictability with reference to the *observed* behaviour of Mars, a series of reference forecasts was compared with subsequent assimilated analyses, covering the entire re-analysed period. In each forecast, a simulation was generated with the stand-alone UK-MGCM, starting from an unperturbed initial state, taken from the re-analyzed dataset (essentially the unperturbed forecasts from the previous section). The model was then integrated forward in time for 30 sols. During this interval, the prescribed dust distribution was taken only from the observed initial state using assimilated TES measurements of dust optical depth. However, unlike the assimilation itself, this dust distribution was held constant and equal to the observed initial state throughout the subsequent 30 sol forecast, conservatively representing the state of knowledge of the atmosphere at the beginning of the forecast period.

In quantifying the ensuing forecast, the diagnostic variable used here is the r.m.s. temperature difference between a single model forecast and an assimilated analysis at 30 Pa:

$$\sigma^{\mathrm{A}}_{\Delta T_{30}} = \sqrt{< c_j (\Delta T_j)^2 >} \qquad (6)$$

where $\Delta T = T^{\mathrm{a}}(t) - T^{\mathrm{f}}(t)$, T^{f} is the standalone model prediction and T^{a} represents the assimilated analysis; c_j is the area weight at grid point j. In most cases this was computed as an r.m.s. average over either the whole surface of Mars or within a given latitude band.

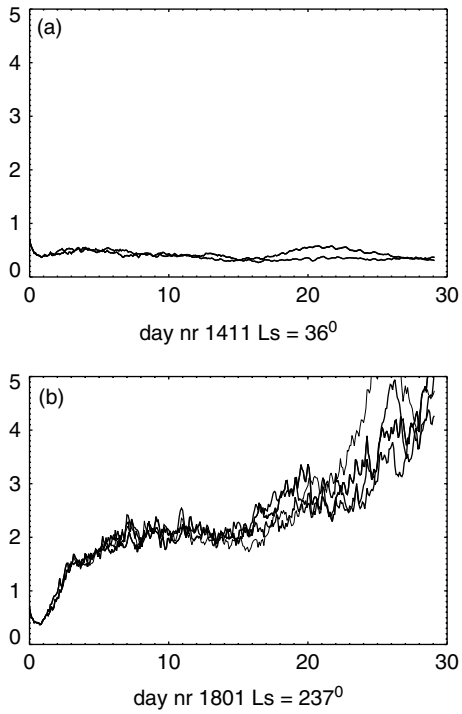

Figure 4. Ensembles of $\sigma^{\mathrm{I}}_{\Delta T_{30}}$ from intra-model simulations at two different times in MY 26.

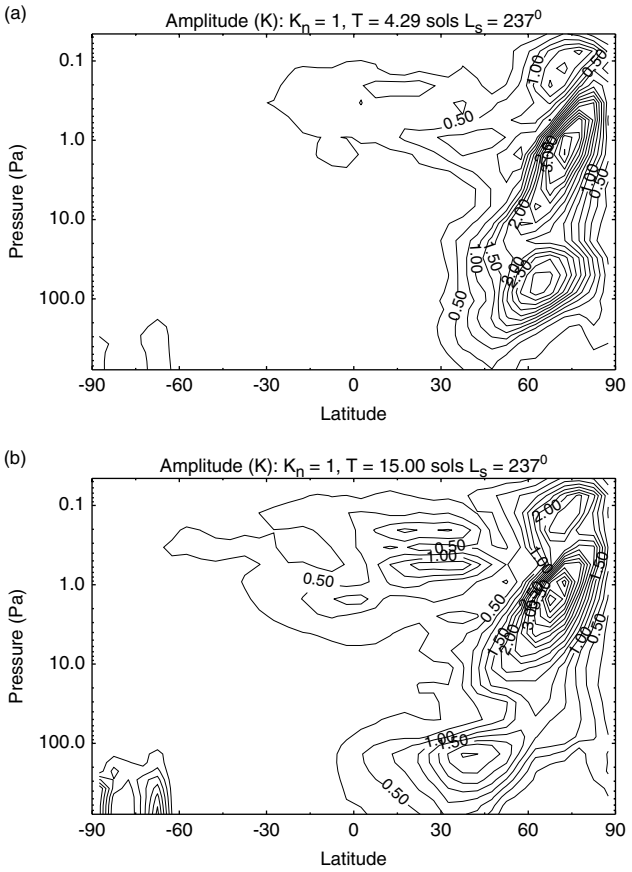

Figure 5. Temperature wave amplitudes (K) for zonal wave number 1 and periods (a) 4.29 sols and (b) 15 sols during the period $L_s = 237^\circ$ to 257° in MY 26 in the model simulation.

4.1. *Seasonal variability in error growth*

Figures 8 and 9 show $\sigma^{\mathrm{A}}_{\Delta T_{30}}$ for a representative set of cases at each of 12 'seasons' during MY 25 and MY 26 respectively. $\sigma^{\mathrm{A}}_{\Delta T_{30}}$ evidently grows rapidly during the first couple of sols in each case, but subsequent behaviour is quite variable and does not always increase monotonically. Overall, the figures indicate the presence of two main time-scales that characterize the divergence between model and re-analysis, one rapid on the order of 1–2 sols and one rather longer of about 4–5 sols. While the rate of divergence on short time-scales only shows a weak latitudinal dependence, there is a significant variation at longer times where the divergence rate in the high-latitude bands is significantly higher than in the equatorial region. There is also a seasonal signature in which the SH growth has an increased amplitude during $L_s = 80^\circ$ to 180°, while the NH band dominates during the remaining part of the year. The global mean divergence tends to level off at most times to ∼6–7 K, but there are periods during NH autumn and winter when the divergence either settles down at a much higher level or continues to increase.

Of particular interest is the period from $L_s = 186^\circ$ in MY 25 and onwards, which shows a particularly large divergence. This is partly due to a strong phase shift in the temperature fields between assimilation and model forecast. The amplitude of the diurnal component also increases dramatically during this period, which coincides with the onset of the planet-encircling dust storm of 2001 (Lewis and Barker, 2005). Such changes are directly attributable to the rapid increase in dust loading during the growth of the dust storm, leading to a clear breakdown in forecasts initiated before the start of the storm.

The 'jumps' that appear in the growth curves from time to time during various seasons (e.g. at sols 8 and 24 at $L_s = 26^\circ$ in Figure 3) are related to the number of available observations for assimilation. When few observations (or none) are available to constrain the process, the assimilated variables will tend to converge towards the control run. Equally, the divergence of the ensemble may increase rapidly as more observations become available during an assimilation period following a data gap.

4.2. *Origin of model assimilation divergence*

There are several possible sources for the more rapid model assimilation divergence than for the intra-ensemble model-only cases. Because of the short radiative time-scale in the Martian atmosphere, the effect of an imperfect representation of the distribution of radiatively active aerosols in the atmosphere rapidly becomes apparent and is perhaps the most likely candidate. The actual vertical distribution of mineral dust in the Martian atmosphere was not observed directly and so is almost certainly incorrectly represented in both model and assimilated analysis. Moreover, the distribution assumed at the start of each model forecast was not evolved realistically in response to the winds in these experiments, further compounding any initial errors in the assumed dust distribution.

In considering the divergence of forecasts from the assimilated record, two circulation components appear to be the main contributors to the error growth, discussed hereafter in terms of α_A, a slowly evolving zonal mean flow and the strength of diurnal tidal forcing. Figure 10 displays latitude-height sections of the zonal mean temperature difference ΔT_{AM} at three times: after the summer solstice,

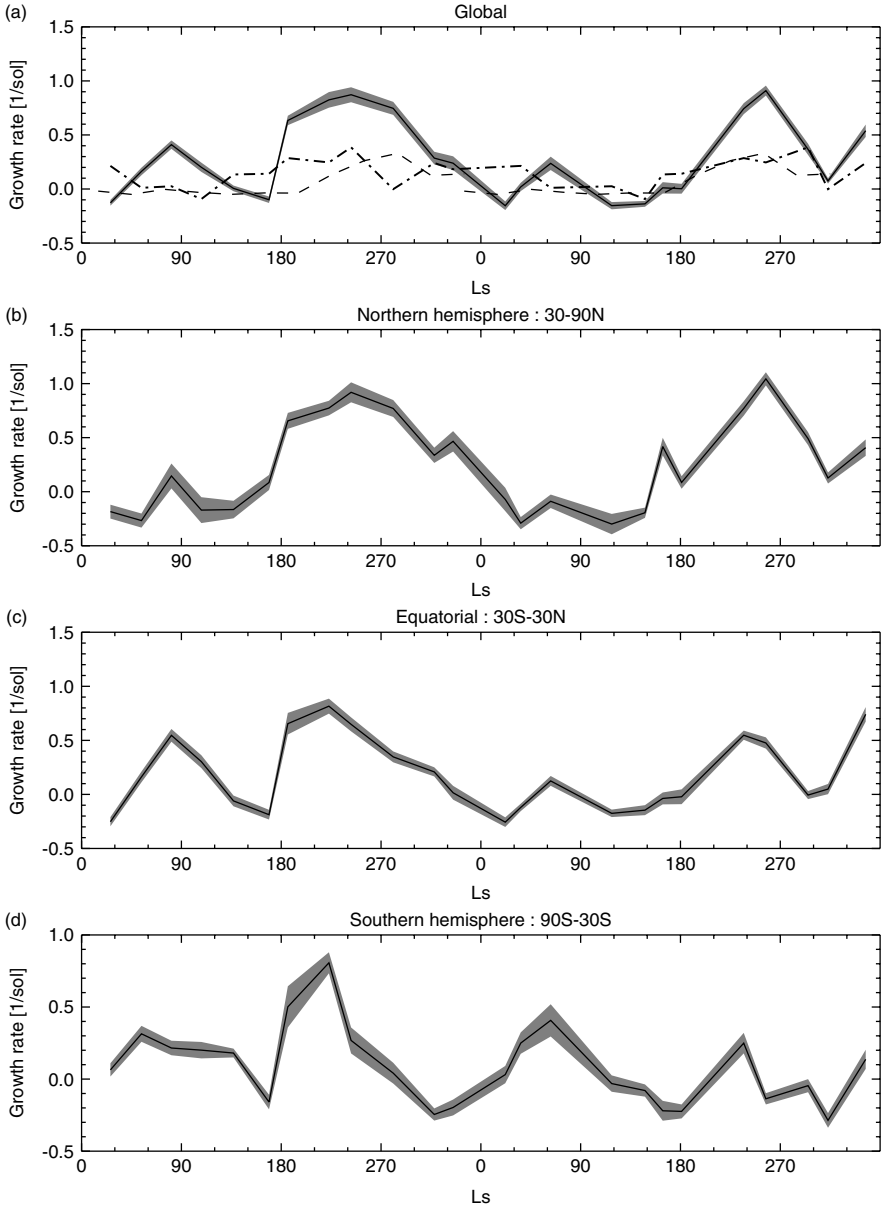

Figure 6. Variation of growth rate, α_1, for $\sigma^{\mathrm{I}}_{\Delta T_{30}}$, (sol^{-1}) during MY 25 and MY 26 for (a) the globe, (b) NH, (c) EQ, and (d) SH. The solid line is the mean, and the shaded area is the standard deviation of the ensemble. In (a), the dashed curve is based on the data in Table I (from Newman *et al.*, 2004), and the dash-dot line shows the global growth rate of $\sigma^{\mathrm{I}}_{p_{\mathrm{msl}}}$ over the same period.

mid NH autumn and late NH winter. The figures are representative of the whole MGS period in the sense that the model to a large extent predicts lower temperatures than observed during the period around the NH summer solstice ($L_{\mathrm{s}} = 118^{\circ}$), while later on in the year the model becomes warmer than observations ($L_{\mathrm{s}} = 164^{\circ}$). The polar regions show a tendency to reverse the sign of this anomaly, particularly in the band 70°–87.5°N during NH autumn and winter. The contour levels in Figure 10 match the amplitude of α_A at corresponding periods in Figure 9, particularly at $L_{\mathrm{s}} = 313^{\circ}$, which exhibits the largest values of α_A and in which the difference amplitude increases to more than 15 K after $\approx$ 12 sols over the South Pole. This matches well with the corresponding value of α_A, which reaches its maximum at the same time (Figure 9).

The tidal amplitude in the model forecast tends to be significantly smaller than observed at most times of the year, and is on the order of $\approx$1–2 K with the exception of NH autumn, particularly during the dust storm in MY 25. A recurring pattern, which can be seen in Figure 11, is a phase locking of the tidal component between the hemispheres. This becomes most prominent during autumn and spring equinoxes but also appears during other times as well. These results suggest that the major source of the discrepancy between the model forecast and the assimilated observations often is linked to the zonal mean component.

The difference fields of zonal mean temperatures around autumn equinox in MY 25 and 26 are displayed in Figure 12. The fields here express the temperature difference at: 4, 24 and 72 h respectively after initialisation. The figure sequence suggests that (i) an initial bi-modal temperature structure seems to evolve at high latitudes in which the model forecast tends to overestimate temperatures at low altitudes but underestimates them at high levels; and (ii) the structures of temperature differences tend to become

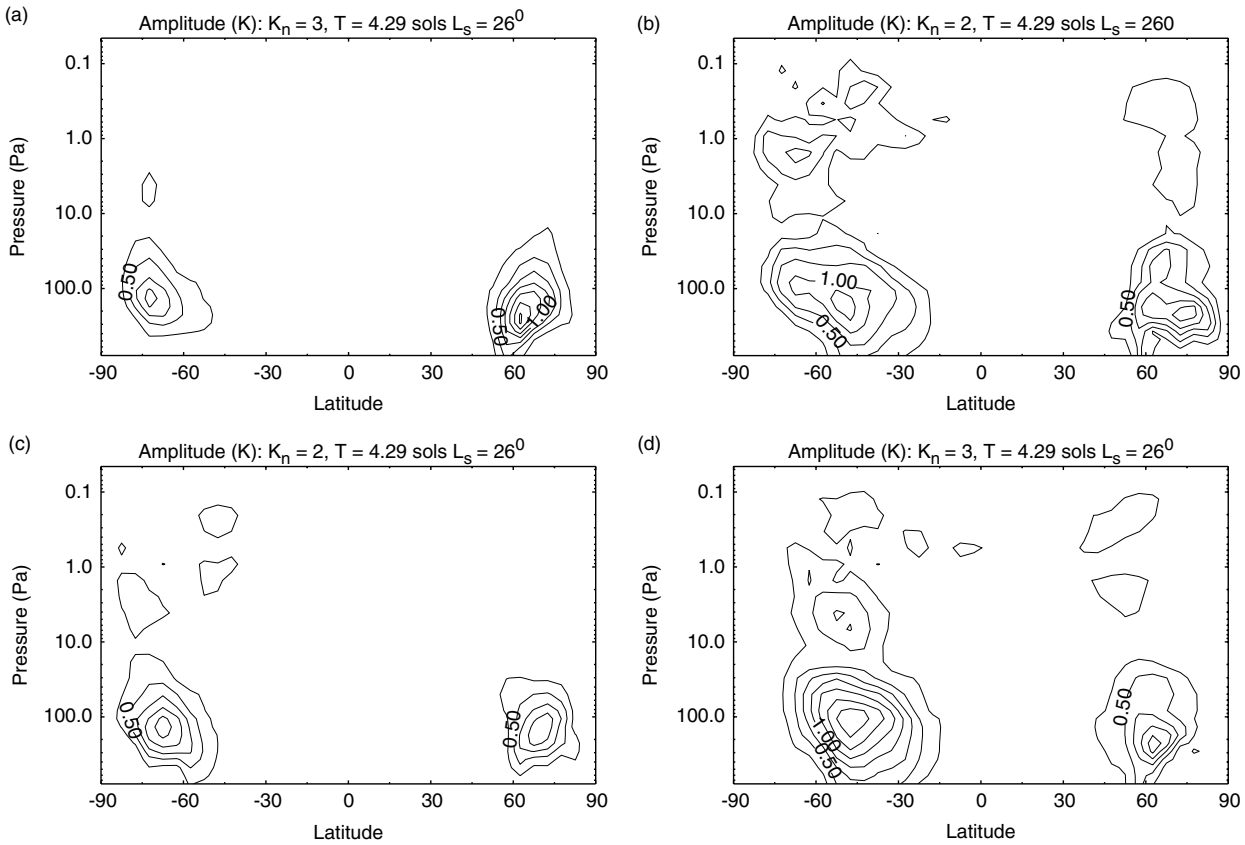

Figure 7. Wave amplitudes for wave number $K_n = 2$ with period $T = 4.285$ sol from (a) the stand-alone model and from (b) the re-analysed data, in Mars year 25. (c, d) are as (a, b), but for $K_n = 3$.

established at an early stage and then simply intensify with time.

Figure 13 shows the zonal velocity difference fields at the same time as in Figure 12. A characteristic feature is that the difference patterns in $\overline{u}$ are deeper in the vertical than for temperature. Figure 11 shows the time sequence of temperatures at the 30 Pa level for the re-analysis and the difference at the same times as in previous figures. The wave number 1 pattern that is seen at high latitudes in both hemispheres around the equinox evidently disappears when the solstice approaches, while a weak wave number 2 pattern appears to develop in both fields above the polar region in the NH. The temperature difference in Figure 11 is relatively symmetric along the Equator, though slightly more intense in the SH.

4.3. Spatial variations in error growth

Mars is in several respects an asymmetric planet, particularly with regard to its topography. This has an impact on the atmospheric circulation, for example in the significantly asymmetric strength of the Hadley circulation (Richardson *et al.*, 2002). The asymmetry leads to the question whether the predictability, in terms of a spatially averaged growth rate, depends on location. Figures 8 and 9 show $\sigma^A_{\Delta T_{30}}$ between the assimilated T_{30} and the model forecast for the three latitude bands (NH, EQ and SH) defined above, together with the global mean divergence during MY 25 and MY 26. The figures clearly show the seasonal impact on the divergence, where the general trend is for the highest growth rate to be found in the winter hemisphere. The inter-annual variability also plays a role here, as can be seen at $L_s = 316^\circ$ in MY 26 (Figure 9) which displays a stronger divergence than the previous year. This takes place in the aftermath of the 2001 planet-encircling dust storm. Comparing these two MGS years indicates similar behaviour with a rapid rate of divergence during the first couple of sols.

Estimates of the divergence growth rates at each season are shown in Figure 14, The panels in the figure show the growth rate coefficient, α_A, separately for the entire planet and for each of the three latitude bands. The growth rate seems to vary only weakly with latitude band, the mean rate (averaged over the entire period) fluctuates around $\overline{\alpha_A} \sim 0.5\,\mathrm{sol}^{-1}$ (Table I); the hypothesised stronger dependence does not appear. The maximal growth rate peaks at $\sim 1.5\,\mathrm{sol}^{-1}$ around $L_s = 180^\circ$ in MY 26, implying a typical e-folding time-scale ranging from 1–2 sols to $\sim$ 0.8 sol. Moreover, the difference between latitude bands is hardly detectable, much in contrast to the behaviour of α_I in the intra-ensemble case (Figure 6). The reason for this may be that model imperfections dominate, leading to a 'climate drift'. Figure 15 shows zonal temperature differences around the NH summer solstice. The differences evolve at a much slower rate than observed later in the year. A similar pattern where the assimilated temperatures are higher in the middle atmosphere evolves during both years.

5. Correcting for zonal mean model error: A sensitivity test

The previous section showed evidence that, at least at some times of year, the dominant contribution to the departure of the model forecast from the assimilated analysis was a systematic error in the zonal mean temperature

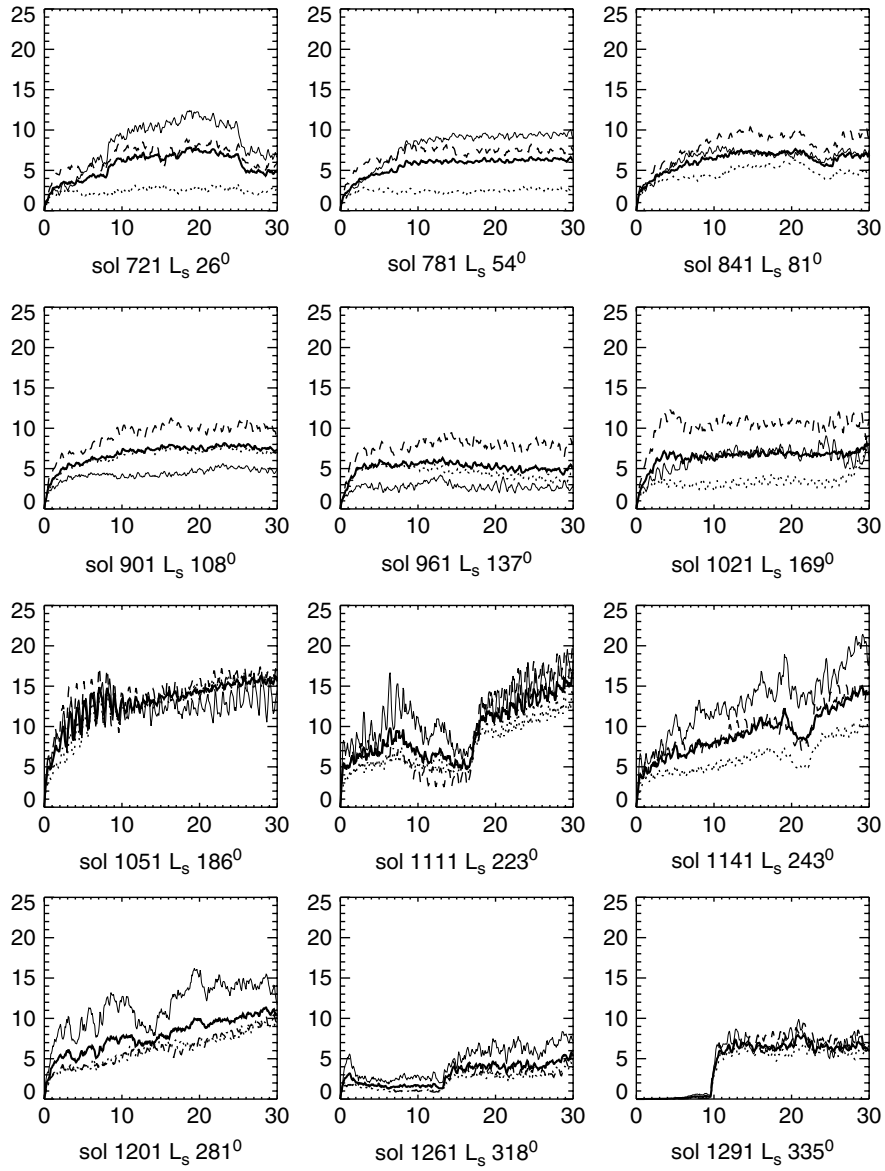

Figure 8. The evolution of $\sigma^{\mathrm{A}}_{\Delta T_{30}}$ (K), at twelve times in MY 25 (at same times as in Figure 2) averaged over the globe (bold solid line), NH (dashed), EQ (dash-dotted), and SH (dotted).

field. This was particularly noticeable around NH summer ($L_\mathrm{s} =\sim 50^\circ$–$150^\circ$), corresponding to the seasons around aphelion. As is apparent in Figure 17, around this time of year the difference pattern in ΔT_{AM} adopts a fairly constant structure with a positive band of temperature anomalies (corresponding to the analysis being warmer than the stand-alone model) between $\pm 60^\circ$ latitude and $\sim$ 10–100 Pa in the vertical.

Recent work by Wilson *et al.* (2008) has identified a similar anomaly, occurring around $L_\mathrm{s} = 120^\circ$ in the zonal mean temperature, with the presence of tenuous water ice clouds that form ubiquitously at low latitudes around aphelion (e.g. Smith, 2004). Most MGCMs at present, however take little account of such water ice clouds in their radiative transfer schemes. It is being increasingly recognized, however, that their contribution to solar heating in the Martian atmosphere may not be negligible, even though the abundance of water ice is typically significantly less than dust (the molar ratio is $\approx 0.03\%$; Kieffer *et al.*, 1992). A cold bias in the UK-MGCM was identified by Wilson *et al.* (2008) in their control simulations of tropical temperature between 10 and 30 km altitude developing around the NH summer solstice. The bias found in the model simulation could be corrected when radiatively active water ice clouds are taken into account.

5.1. *An empirical correction for model error*

A hypothetical reason for the rapidly growing systematic divergence between the control forecasts and the corresponding re-analyses is the model's radiation scheme not adequately accounting for the presence of radiatively active aerosols such dust or water ice. To examine whether the omission of dust and water ice might have an effect on some of our forecasts, we have conducted some test forecasts in which we introduce an empirical zonally symmetric relaxation correction term for the temperature increment in the model,

$$T_{i+1} = T_i + \left(F_{\mathrm{ex}} + \frac{\overline{T}_\mathrm{A} - \overline{T}_\mathrm{M}}{\tau_\mathrm{r}} \right) \delta t, \qquad (7)$$

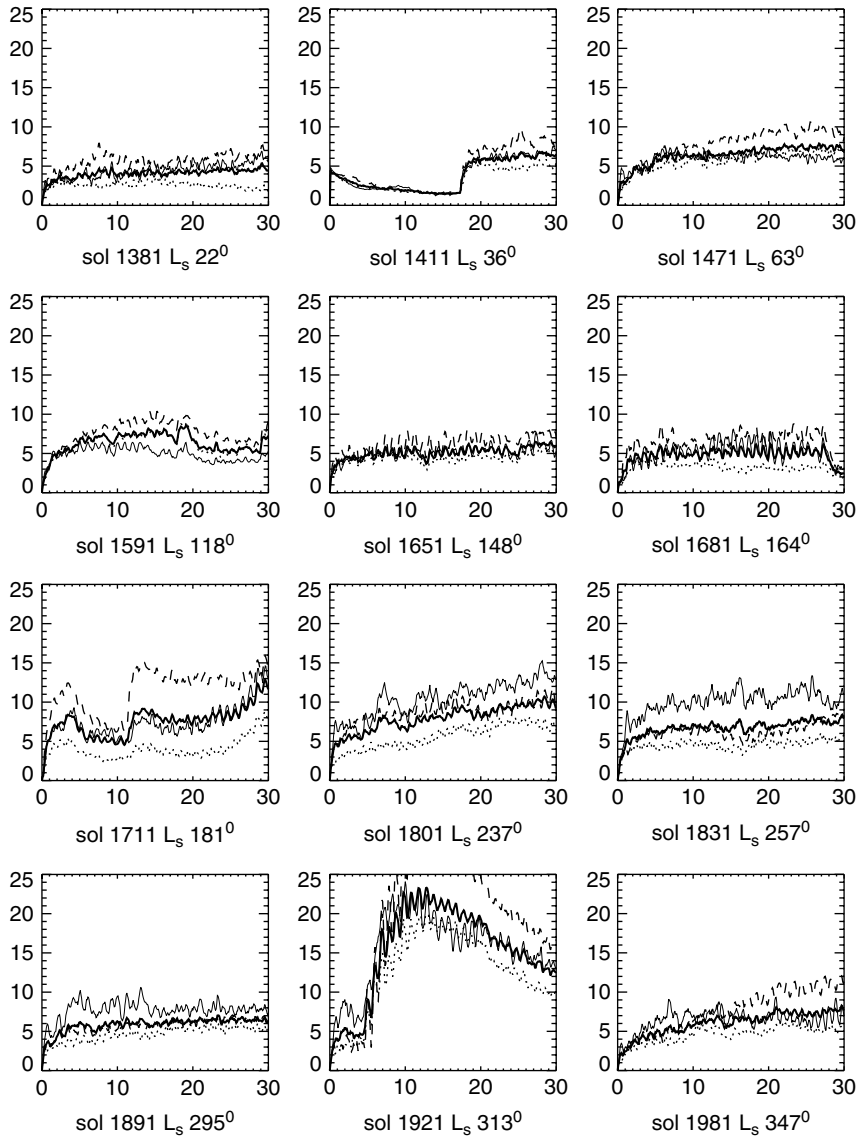

Figure 9. As Figure 8, but for MY 26.

where T_i is the atmospheric temperature at model time step i, $F_{\rm ex}$ is an explicit forcing, $\overline{T}_{\rm A} - \overline{T}_{\rm M}$ is the difference between the zonal averages of assimilated and model fields, δt is the model time step and $\tau_{\rm r}$ is a free parameter representing the strength of relaxation of the model towards the assimilated analysis. For the present purpose, $\overline{T}_{\rm A} - \overline{T}_{\rm M}$ was computed as a time-average quantity over the forecast interval, so was held constant in time during the modified model forecast. This kind of adjustment, therefore, is not so much intended as a practical model 'improvement' but as a sensitivity test, where the purpose is to examine the model's response to an artificial, zonally symmetric forcing term and its impact on forecast error.

5.2. *Results with error correction*

Figure 16 shows the time variation of $\alpha_A(t)$ with the adjustment term in Eq. (7) included for several different values of $\tau_{\rm r}$. The figure shows results for forecasts initialised at $L_{\rm s} = 186^\circ$ during NH autumn and at $L_{\rm s} = 80^\circ$ in NH summer, with α_A computed for the entire globe, and separately in each of the main NH, EQ and SH latitude bands.

The impact of the empirical zonal mean relaxation correction varies significantly with season. During the NH autumn equinox, for example, the effect is relatively weak with the possible exception of the equatorial region, where equilibrated values of α_A may be reduced by a factor of ~ 5 given a suitable choice of $\tau_{\rm r}$. However, small values of $\tau_{\rm r}$ ($\tau_{\rm r} < 1$ sol) appear to significantly overcompensate for the model error during the first 5 sols, although this leads to a much smaller equilibrated error for $t > 5$ sols in the EQ band. In the NH band, however, $\tau_{\rm r} < 1$ sol evidently leads to an *increased* model error by up to a factor of 2. With $\tau_{\rm r} > 1$ sol, there is hardly any impact at all throughout each 30 sol forecast in any of the latitude bands. This would seem to suggest that the systematic model error at this time of year involves not just the zonal mean component but has a more complex spatial structure associated with large-scale stationary waves and tides.

At the summer solstice (Figure 16), there is still a tendency for overcompensation with $\tau_{\rm r} < 1$ sol, such that the model error increases more strongly with the relaxation correction than without it at all forecast times and latitudes. However,

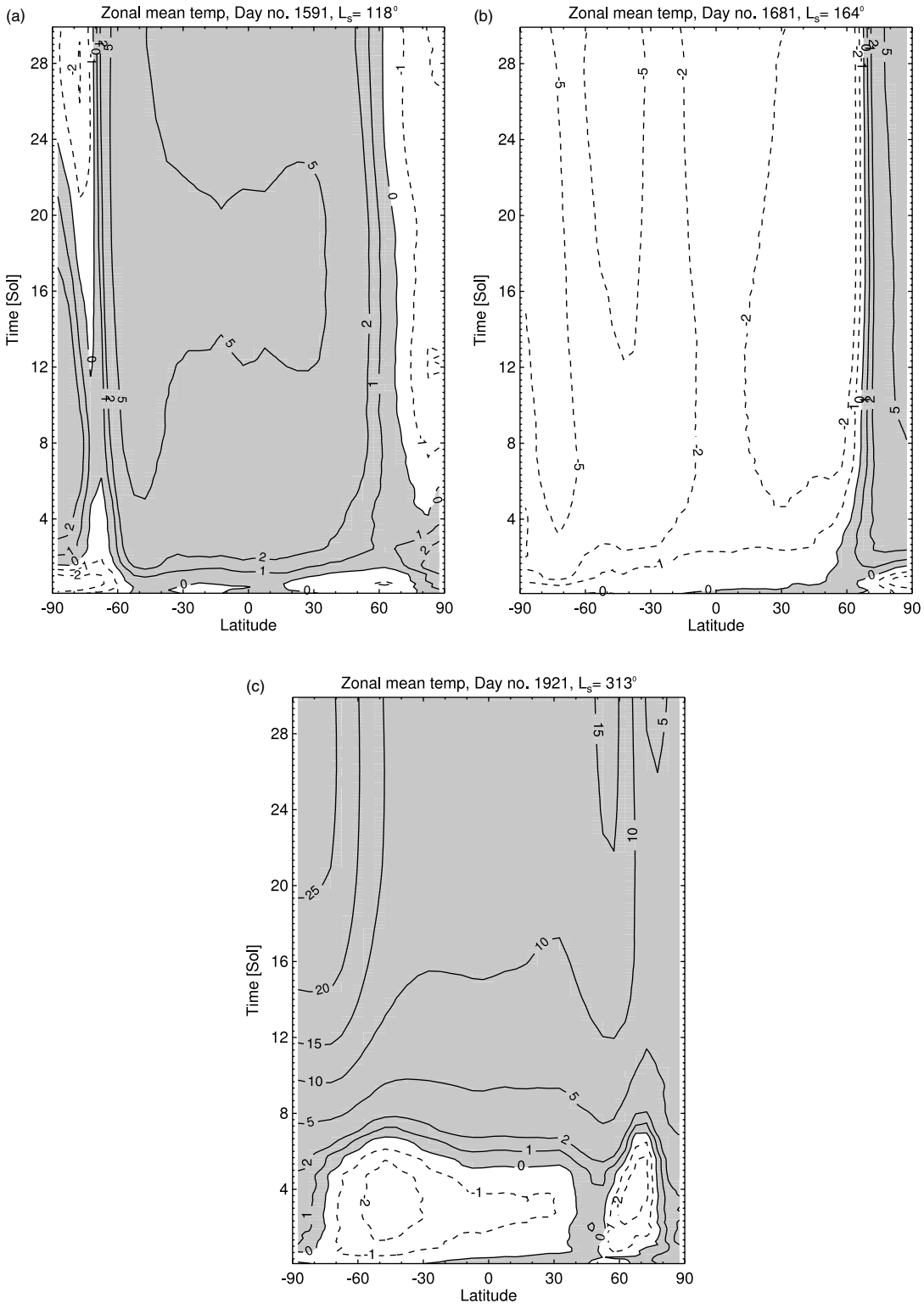

Figure 10. Evolution of ΔT_{AM} at 30 Pa, in MY 26 beginning at (a) $L_{\mathrm{s}} = 118^{\circ}$, (b) 164° and (c) 313°. Areas where the assimilation is warmer than the control run are shaded grey and have solid contour lines; regions where the assimilated temperatures are lower are shown as white areas with dashed contours.

when $\tau_{\mathrm{r}} > 1$ sol at this time of year there is a significant decrease of divergence in all three latitude bands. Indeed with $\tau_{\mathrm{r}} \approx 40$ sols, $\sigma_{\Delta T_{30}}(t) \approx 1$ K, which is well within the expected r.m.s. measurement error (Montabone *et al.*, 2006) and comparable with the equilibrated intra-ensemble spread of stand-alone model simulations found in section 3 (cf. Figure 2). This clearly demonstrates that, at least close to aphelion, the dominant contribution to systematic model error appears to be due to a divergence in the zonal mean temperature structure predicted by the model away from the structure observed on Mars itself. In this case, we suggest that such a systematic error most probably arises from the absence of a realistic representation of the radiative effects of the aphelion water ice clouds that are known to be persistently present all around the planet at low latitudes at this time of year.

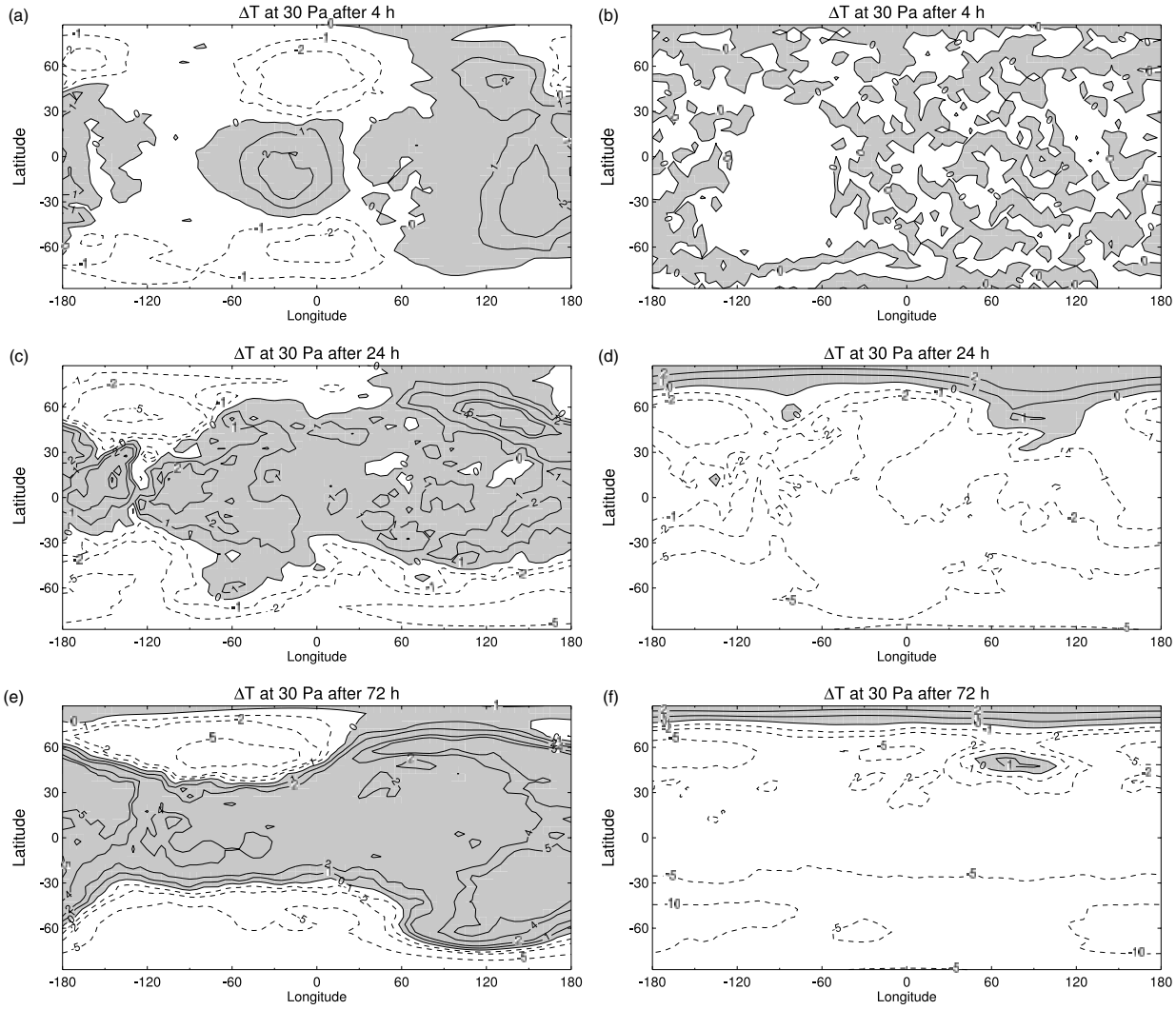

Figure 11. ΔT_{AM} at the 30 Pa surface and at $L_{\mathrm{s}} = 186^{\circ}$ in MY 25 from re-analysis after (a) 4 h, (c) 24 h and (e) 72 h. (b, d, f) are as (a, c, e), but at $L_{\mathrm{s}} = 181^{\circ}$ in MY 26. Shading is as Figure 10.

To illustrate this further, Figure 17 shows latitude–height maps of the effect of the empirical relaxation adjustment term for $\tau_{\mathrm{r}} = 40$ on the zonal mean temperature field during the first 3 sols of the forecast. The left-hand column of Figure 17 shows the growth of model assimilation zonal mean temperature differences without the addition of the relaxation term, in which the characteristic warming pattern at low latitudes between 10 and 100 Pa is seen to build up coherently over 3 sols to an amplitude of ~ 2 K. With the addition of the relaxation correction, however, with $\tau_{\mathrm{r}} = 40$ sols, the divergence of the model away from the assimilation is much reduced, so that after 3 sols the amplitude of the difference field is only around 0.1 K. The fact that the total model error α then remains less than 1–2 K throughout the whole corrected forecast indicates that the correction of the zonal mean temperature tendency is sufficient to reduce the model errors to a level comparable with the intrinsic nonlinear variability of the model dynamics. The latitude–height structure of the zonal mean correction required to reduce the model error to such small limits is closely similar to the pattern identified by Wilson *et al.* (2008) as originating from the absence of water ice cloud radiative heating in the Geophysical Fluid Dynamics Laboratory (GFDL) model, suggesting that the introduction of a suitable parametrization of such clouds into the model should lead to significant improvements in the climatology and predictive capability of this, and similar, models.

6. Discussion

This study has sought to investigate the respective roles of internal (chaotic) dynamics and systematic model error in determining the intrinsic and practical predictability of the Martian atmosphere. Our intention has been to build on the earlier stand-alone model study of Newman *et al.* (2004), which concentrated on investigating seasonal variations in atmospheric predictability in a MGCM using ensemble methods, and in which the dust distribution was prescribed and kept identical for each ensemble member. The resulting divergence of model trajectories within each ensemble was then due primarily to internal chaotic dynamics, dominated largely by the behaviour of baroclinic transient waves and their interactions with the zonal mean flow, topography and thermal tides. In each case considered by Newman *et al.* (2004), the 'forecast' was initialised from a previous stand-alone model simulation, under conditions that were a reasonable climatological representation of a 'typical' Mars year consistent with observations from MGS. However, no

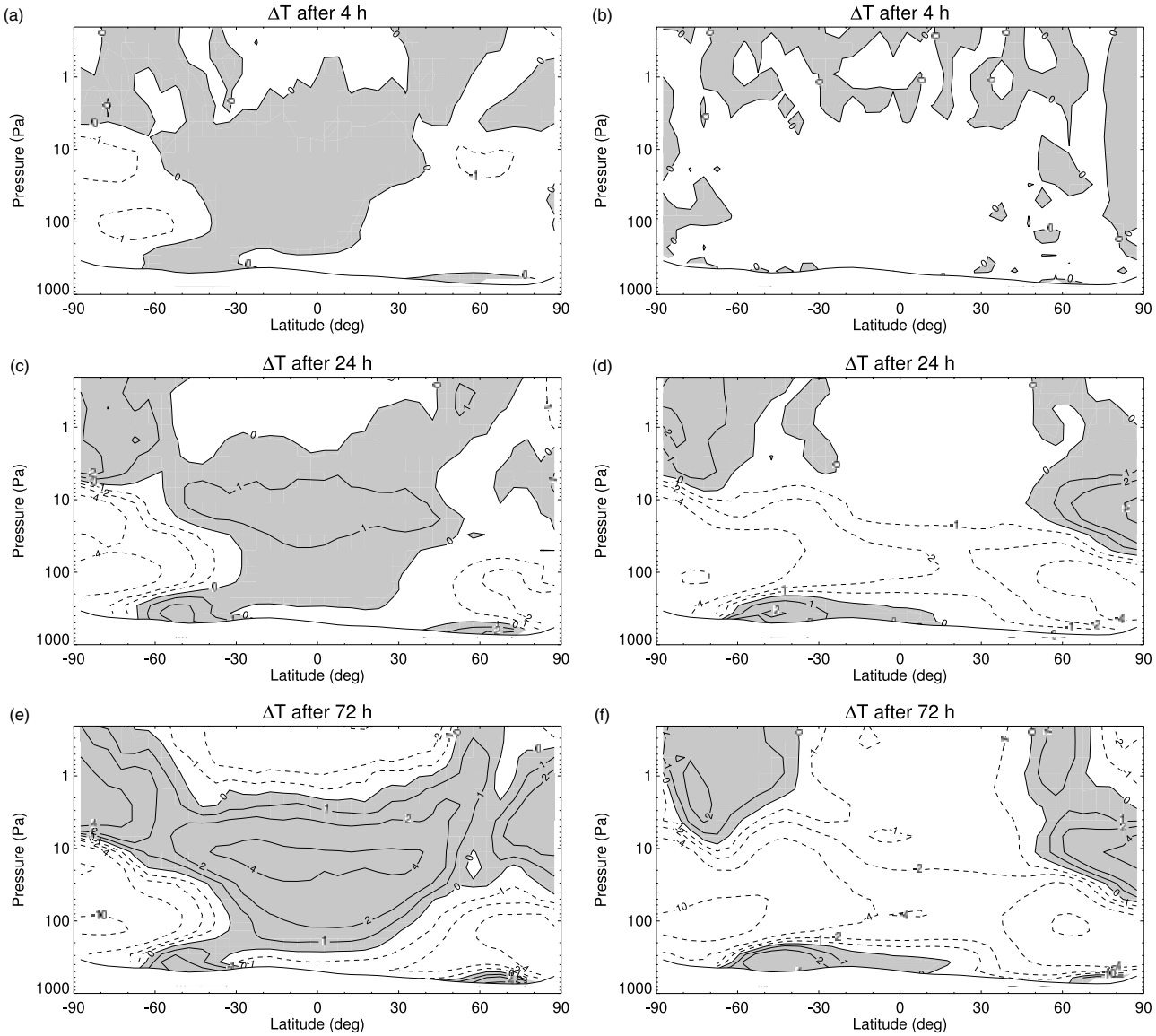

Figure 12. Evolution of zonal mean temperature differences at $L_s = 186^\circ$ in MY 25 from re-analysis after (a) 4 h, (c) 24 h, and (e) 72 h. (b, d, f) are as (a, c, e), but for $L_s = 181^\circ$ in MY 26. Shading is as Figure 10.

attempt was made precisely to reproduce observed synoptic events.

In the new work presented here, we have directly addressed the assessment of both intrinsic and practical predictability with reference to the best available global, synoptic observations of the Martian atmosphere to date, using the UK re-analysis dataset derived from the MGS/TES measurements of temperature and dust. This has allowed us to initialise model ensemble forecasts directly from observed initial states, sampling a wide range of seasonal conditions and covering two 'typical' Mars years, one with and one without a major planet-encircling dust storm. Moreover, the availability of the re-analysis dataset, with its continuous 3D multivariate record covering the entire globe and sampled every 2 h, also enables the possibility of direct observational verification of model forecasts, by analogy with the practice of terrestrial operational weather forecasting, to make the first investigation of the practical predictability of Martian weather on time-scales of hours up to 30 sols. It should be noted that the assimilated data constrain the model effectively at mid-altitudes but less so closer to the surface and that the impact of dynamics in the boundary layer on the predictability is yet to be established.

In this study we have used the UK-MGCM and some aspects of the results may reflect particular biases and features due to the exact parametrization schemes used in this model. However, this model is very similar in formulation to a number of other contemporary MGCMs which do not include interactive aerosol transport schemes, and would therefore be expected to produce broadly similar results.

6.1. *Intra-ensemble predictability*

The intra-ensemble part of the present study confirms in large part the results obtained by Newman *et al.* (2004), in that the intrinsic predictability of the Martian atmosphere, at least within the limitations of the current generation of models, appears to vary with season and has a distinct seasonal asymmetry that seems to be broadly reproducible from year to year. This asymmetry is characterized by low (or even negative) growth rates

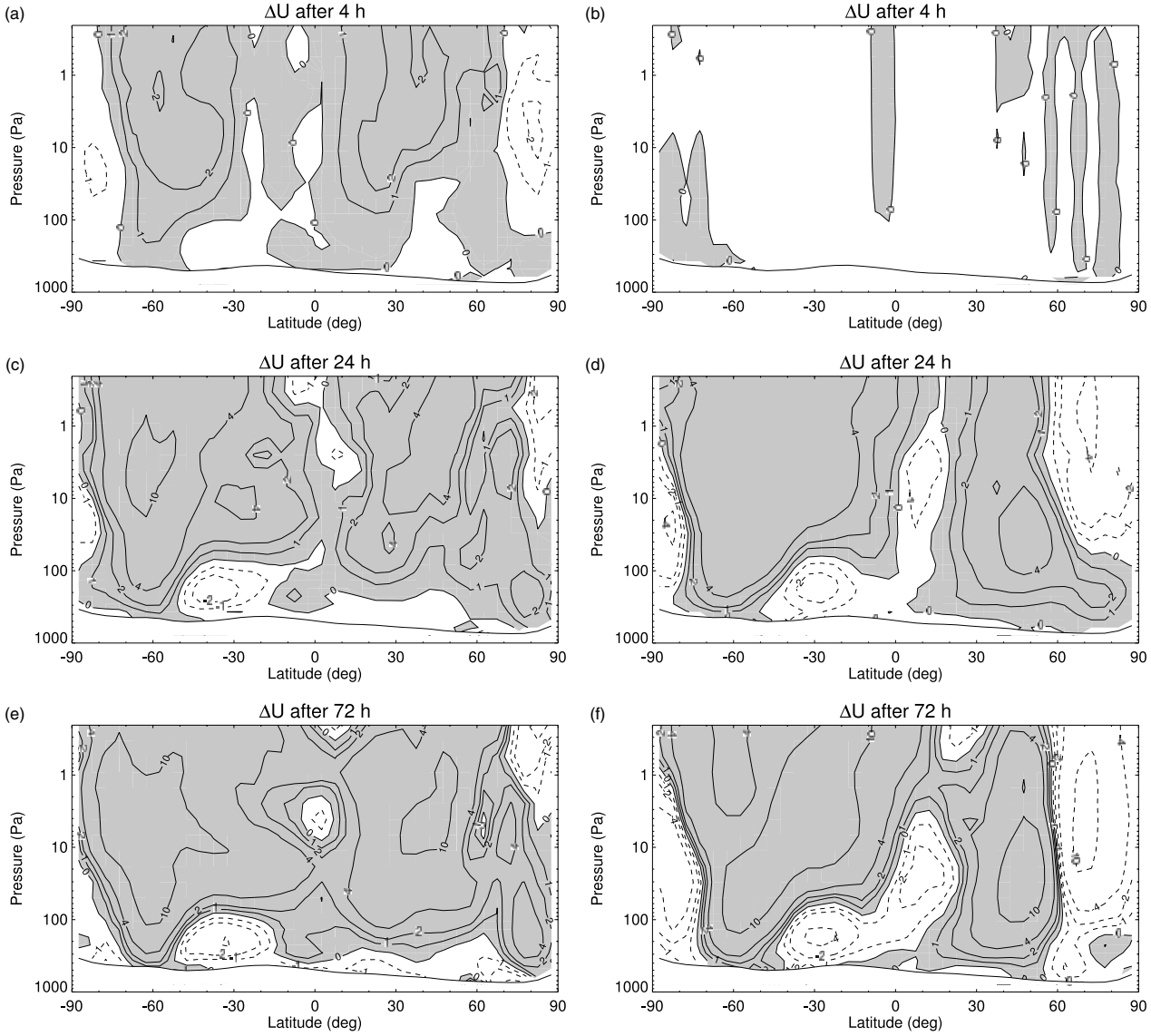

Figure 13. As Figure 12, but for zonal velocity differences.

during NH late spring and summer, while during NH autumn and winter the growth rates increase significantly. Such behaviour is largely associated with the ebb and flow of intense baroclinically unstable storm systems in the midlatitude storm zones, especially in the NH. These systems are typically at least somewhat chaotic, in the sense that their behaviour is significantly sensitive to initial conditions.

A noticeable difference between the present study and the results of Newman *et al.* (2004) is that positive growth rates in our work were found at all times of year, not just during NH autumn–spring as found by Newman *et al.* (2004). This may be because the stand-alone model tends to underestimate baroclinic activity in the SH compared to observations, so that baroclinic activity tended to be weak during SH winter. This led to a very weak or even zero growth, based on an initial amplitude of ~ 5 K for the ensemble perturbations used by Newman *et al.* (2004). In the present work, initialising forecasts from assimilated analyses of observed atmospheric states, the baroclinic activity in the south remains reasonably strong during southern winter, leading to positive growth rates, albeit weaker than during the corresponding seasons in the north.

6.2. *Nonlinear saturation*

Figures 2, 3, 8 and 9 showed that both the model assimilation and the intra-ensemble error growth saturates after a few sols for long periods of the year. This behaviour is a characteristic signature for systems where the rate of divergence is not effectively dominated by the largest non-negative exponent but by exponents whose magnitudes themselves depend nonlinearly on initial location and separation between the two fields (Chen *et al.*, 2006). A possible generalization from the classic linear formulation (which assumes the initial separation to be infinitesimal and independent of position) to a nonlinear local Lyapunov exponent is outlined and used to examine the predictability of atmospheric data for the Earth by Ding and Li (2007) and Chen *et al.* (2006), based on the well-known ERA-40 (ECMWF) re-analysis dataset. To explore the potential effect of saturation on the estimate of growth rate, and subsequently, the time of predictability,

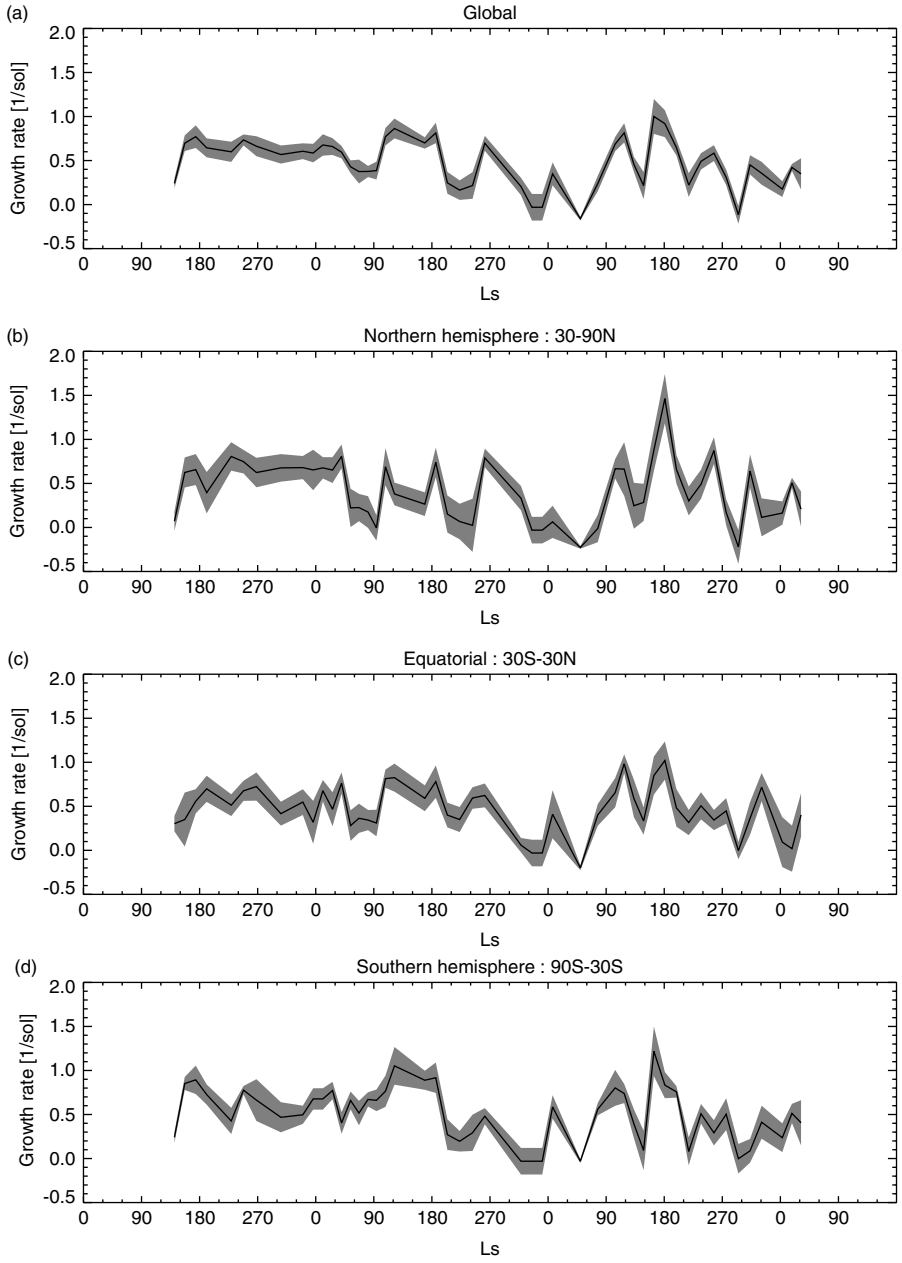

Figure 14. The mean growth rate α_A (solid line) for (a) the globe, (b), NH, (c) EQ, and (d) SH. The shaded area is the standard deviation of α_A.

an area-averaged nonlinear growth rate was calculated as

$$\lambda(\tau, K) = \frac{1}{\tau} \log \frac{\sqrt{< c_j \Delta T^2(t+\tau, \phi_j) >}}{\sqrt{< c_j \Delta T^2(t, \phi_j) >}}, \qquad (8)$$

where ϕ_j is the jth latitude band of grid points, τ is the time lag, $\Delta T^2(t, \phi_j)$ is the square of the temperature difference between assimilated data and model, and c_j the latitude-dependent area weighting, with angle-brackets denoting ensemble averaging either over the whole globe or the Kth latitude band as in Eq. (4).

The decay of λ as a function of delay time τ in MY 26 (sol 1501 ($L_s = 77^\circ$) to sol 2071 ($L_s = 32^\circ$)) is shown in Figure 18(a), showing profiles for both the global atmosphere and the NH, EQ and SH latitude bands used above. The exponential decay of λ with τ follows qualitatively what has been observed on Earth (Figure 1 in Chen *et al.*, 2006). The initial decay rate appears to be slower in the Martian atmosphere than for the Earth. Figure 18(b) displays $\log(\lambda t)$ for the corresponding latitude bands. Chen *et al.* (2006) used the value $0.99 \log(\lambda t)$ as $t \to \infty$ (i.e. when $\log(\lambda \tau)$ reaches a stable value) to define a predictability time-scale for the variable in question, while Ding and Li (2007) used 0.98 of the saturation value (following Dalcher and Kalnay, 1987), for comparison with α_A^{-1} or other measures of linearised growth time-scales. In our example here, shown in Figure 18(b), this leads to a predictability horizon at around 10–13 sols, which is comparable with the maximum predictability time-scales for Earth found using this method (Chen *et al.*, 2006). It is also a somewhat longer estimated time-scale than suggested by the other methods we use herein, but this is not unduly surprising given the different definitions of predictability time-scale. Our estimate for the rate of model assimilation divergence, based on linearised growth, indicates an error growth e-folding time-scale of around 0.5 sol. At least at face value, this would correspond to a nonlinear error growth to around

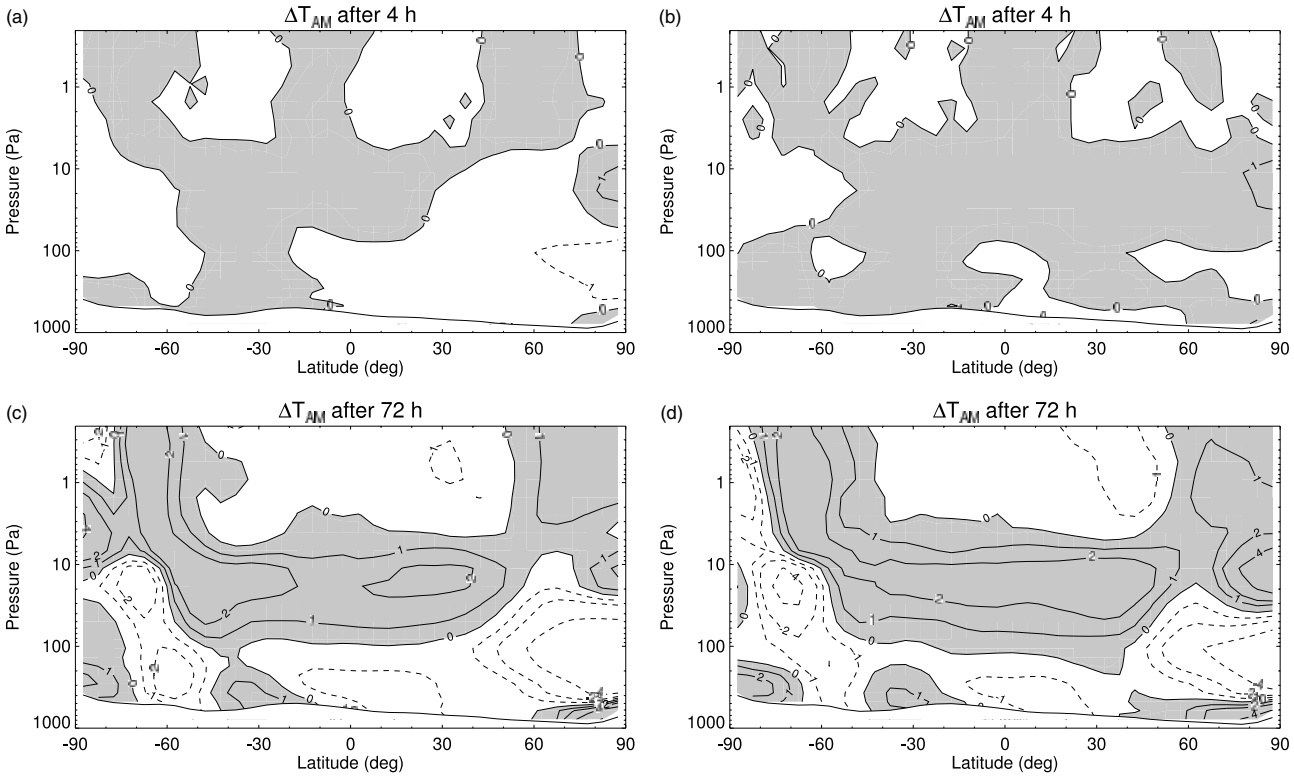

Figure 15. Evolution of zonal mean temperature differences at $L_s = 81^\circ$ in MY 25 from re-analysis after (a) 4 h and (c) 72 h. (b, d) are as (a, c), but for $L_s = 77^\circ$ in MY 26. The shading convention is as Figure 10.

0.5–0.6 of the saturated limit, based on the curves shown in Figure 18(b).

This approach would seem to be promising for systems exhibiting significant saturation of divergence error, of which the Martian atmosphere appears to be a good example. The approach suggested by Chen *et al.* (2006) and Ding and Li (2007), however, requires a long observational record from which to obtain close dynamical analogues to generate model perturbations. Such a method requires much longer time series than are available for Mars, or likely to become available for some time to come. However, similar nonlinear methods may well reward further investigation in future predicability studies.

6.3. *Forecast verification with observations*

From comparisons between model forecasts, initialised from observed atmospheric states, and subsequent observations from the re-analysis dataset, it is apparent that the model diverges rapidly away from the observed state of the atmosphere. Determinations of forecast error indicate that, after several sols, it typically becomes much larger than found within the stand-alone model ensembles. This seems to be a manifestation of the growth of a systematic model error that causes the model trajectory to drift rapidly away from the observed evolution of the atmosphere, presumably as a result of deficiencies in the model formulation (notably in the radiative transfer parametrizations).

An examination of the divergence between data from a typical stand-alone model run and the assimilated dataset suggests that there may be several separate sources contributing to model error, and that the relative importance of these sources may vary throughout the year. The rate of divergence seems typically to be characterized by two time-scales, a shorter scale ≈O(1–2 sol), apparently caused mainly by incompleteness and imperfections in the model's radiative scheme, and a longer time-scale around 10–20 sols.

In relation to the shorter time-scale, the model does not currently parametrize the effects of water ice clouds, which have a significant impact on the radiative balance, particularly around NH summer solstice at low latitudes. In addition, the transport and evolution of the distribution of radiatively active dust is also not taken into account in the present work, which may also lead to the growth of errors in the numerical forecast as the dust distribution evolves.

6.4. *Radiative and dynamical time-scales*

To further investigate possible causes of the initial rapid divergence between model forecasts and assimilation, we can estimate the radiative and dynamical time-scales by calculating respective characteristic relaxation times τ_{rad} and τ_{dyn}, in which τ_{rad} quantifies the rate at which radiative energy exchanges can dissipate or modify a thermal perturbation while τ_{dyn} measures the relaxation time due to advection by winds.

Assuming that the Martian atmosphere is close to radiative equilibrium, i.e. incoming short-wave radiation is balanced by emission of long-wave infrared radiation, this balance can then be expressed as in Eq. (9). In this case, a first-order estimate can be obtained of τ_{rad} for a slab of atmosphere of thickness h with uniform density ρ which radiates as a black body at a temperature of T. Such a slab will undergo a small temperature change ΔT according to

$$c_p \rho h \frac{\mathrm{d}\Delta T}{\mathrm{d}t} = 8\sigma T^3 \Delta T, \qquad (9)$$

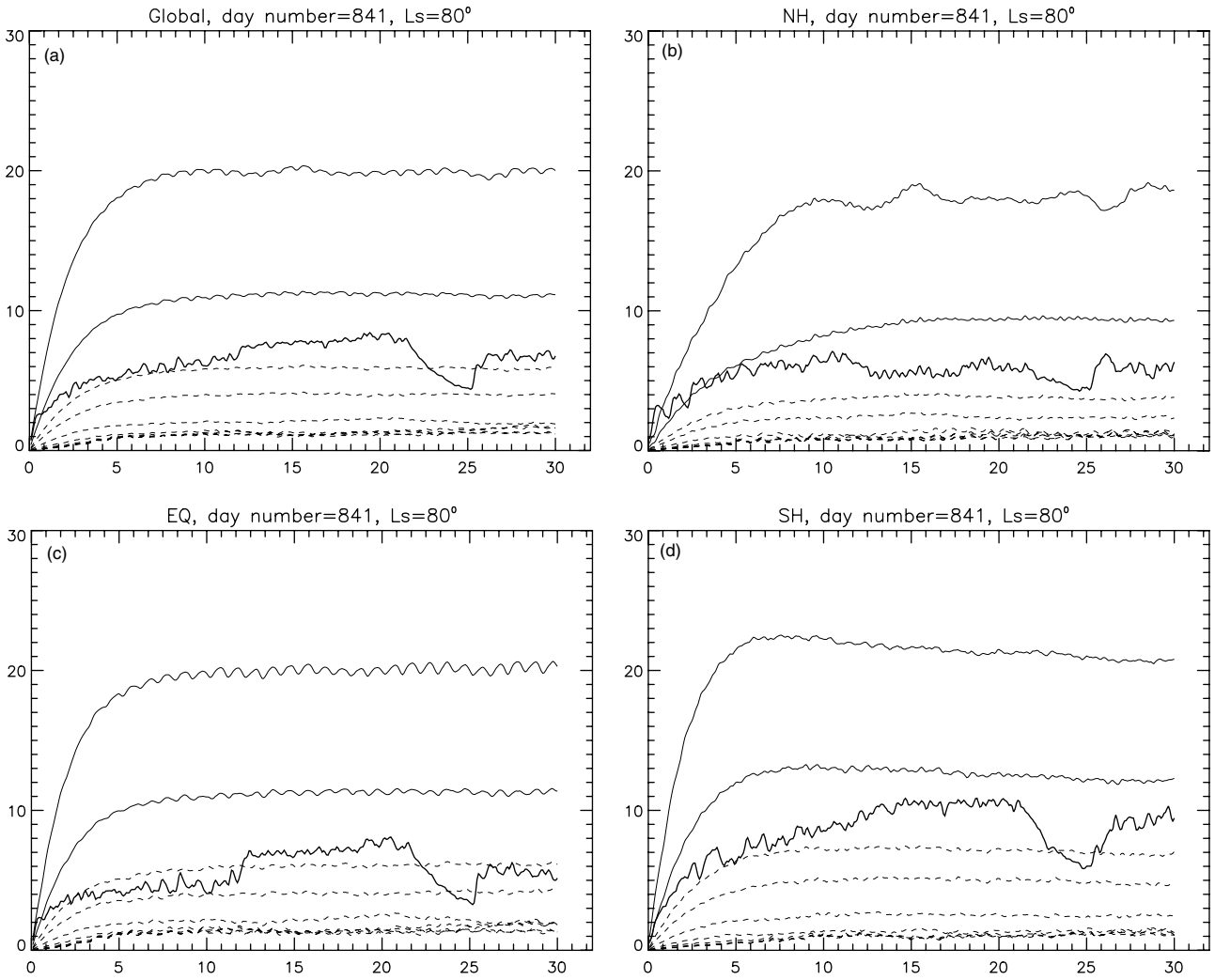

Figure 16. $\alpha_A(t)$ at 30 Pa as a function of the relaxation time constant, τ_r, from the period around the autumn equinox for (a) the globe, (b) NH, (c) EQ and (d) SH. The bold solid line is with no adjustment, thin solid lines mark $\tau_r < 1$ sol, and dashed lines mark $\tau_r > 1$ sol.

where $\sigma = 5.67 \times 10^{-8}\,\mathrm{J\,K^{-4}m^{-2}s^{-1}}$ is the Stefan–Boltzmann constant and $C_p = 860\,\mathrm{J\,K^{-1}kg^{-1}}$. The radiative relaxation time-scale is then given by

$$\tau_{\mathrm{rad}} = \frac{C_p \rho H}{8\sigma T^3}, \tag{10}$$

(Houghton , 1986) where H is the density scale height ($\simeq 10.8$ km on Mars). Values of τ_{rad} for the lower part of the atmosphere are then shown in Table II.

The last column in Table II shows τ_{rad} from Goody and Belton (1967), in which they calculated the relaxation time analytically in a homogeneous CO_2-rich atmosphere. Dynamical time-scales are likely to be on the order of an advective time-scale, L/U, where L is a typical horizontal scale for synoptic weather systems ($\gtrsim 1000$ km) and U is a typical horizontal velocity scale ($\sim$ a few tens of $\mathrm{m\,s^{-1}}$). On this basis, $\tau_{\mathrm{dyn}} \gtrsim$ a few sols, so dynamical time-scales are likely to be O(5–10) times longer than radiative time-scales, at least in Mars' lower atmosphere. A rough estimate of the critical dynamical relaxation time at 30 Pa of $\tau_{\mathrm{dyn}} \sim \pi R/u_{\mathrm{zon}} \approx 2$ sol (where $u_{\mathrm{zon}} \approx 60\,\mathrm{m\,s^{-1}}$) suggests that relaxation due to radiative exchanges is the dominant factor governing thermal adjustment during the first few sols of the model forecast. Small errors in the radiative balance computed in the model compared with the real atmosphere will manifest themselves within a few τ_{rad}, which are O(hours) to a few sols. However, a major factor that comes into play in the difference between assimilated and stand-alone model data is the lack of a self-consistent and realistic description of the dust, both with respect to distribution as well as interaction with other aerosols. The close link between aerosols and radiation indicates that the effect of an imperfect dust scheme will be seen as deficiencies in the radiative forcing. An improved dust transport scheme is likely, therefore, to decrease the differences between model forecasts and assimilated observations. Whether this will improve the time-scale over which useful practical forecasts can be obtained remains to be seen, however.

We see, therefore, that the initial rapid growth commonly occurs on a radiative time-scale, most likely associated with radiative imbalances between the modelled and actual Martian atmosphere in the initial state of each forecast. This is particularly evident in the examples shown above around $L_s = 80^\circ$, in association with the lack of solar heating due to water ice clouds around aphelion. Later in the year, however, there appears to be a shift in dominance from radiative processes to dynamical, so that the baroclinic storm track which appears around 60°N during NH autumn–winter contributes to the divergence between forecast and analysis. The divergence on longer time-scales, $\approx$ O(10–20) sols, seems

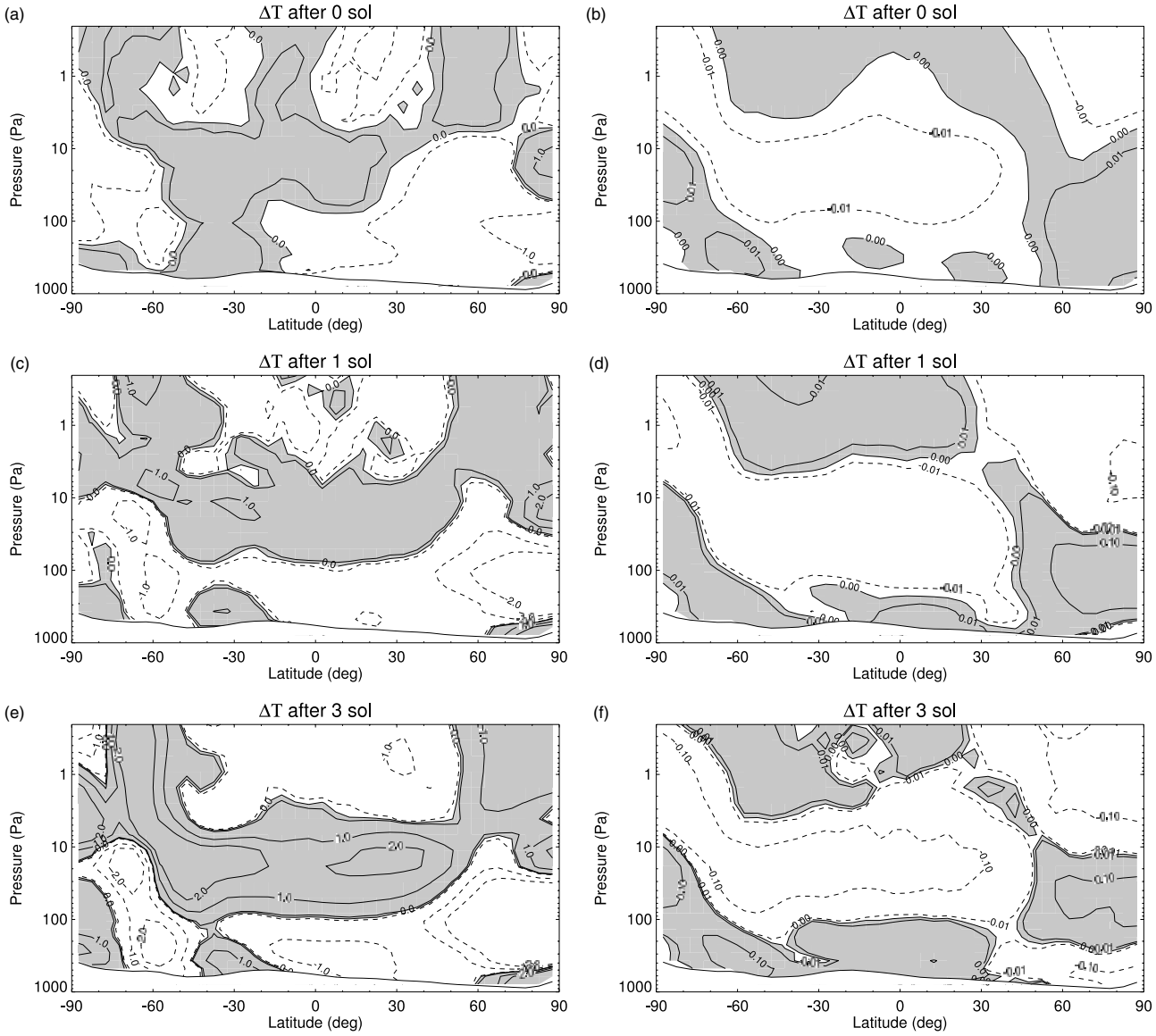

Figure 17. The zonal temperature difference after (a), 2 h, (c) 1 sol and (e) 3 sols at $L_s = 80^\circ$ MY 25 with no adjustment. (b, d, f) are as (a, c, e), but with adjustment and $\tau_r = 40$. The contour levels differ between the columns.

Table II. Radiative relaxation times.

Pressure (Pa)	Height (km)	Temp (K)	τ_{rad} (h)	τ_{rad} (Goody)
610	0	200	11.3	–
500	2.2	200	9.3	–
300	7.7	185	7.6	7.5
200	12.0	180	5.6	–
30	32.5	170	1.1	–

more linked to dynamical effects associated with dust transport and the gradual growth of a phase shift of the baroclinic waves between forecast and assimilation. Given a radiatively induced change in the zonal mean temperature structure, corresponding changes in the zonal mean flow would be expected to lead to different advection rates and propagation speeds of synoptic weather systems, consistent with the rate of divergence between forecast and assimilation found here.

Considering the high thermal and orographic variability of the Martian surface, it was hypothesized that the predictability has a spatial dependency. However, the verification of model forecast assimilation suggests that the differences of growth rates between latitude bands are small on time-scales of $\tau \lesssim 2$ sols. At longer times, however, there is a strong correlation between divergence and season, with the northern and southern bands displaying a periodicity of maximal divergence, while the divergence in the equatorial band is less influenced by season. The biggest differences seem to be found near the edge of the polar caps, though the relatively coarse model resolution has to be taken into account. Overall, however, this suggests that the differences between model and observations

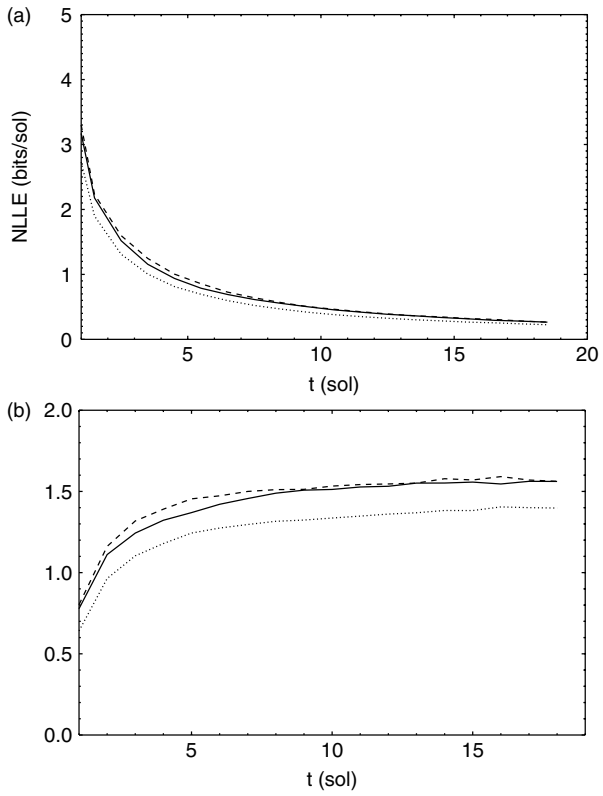

Figure 18. (a) λ and (b) log(λ*t*) at 30 Pa in MY 26, both beginning at sol 1: NH (dotted), EQ (dashed), SH (solid).

are less due to topography but are largely thermal in origin.

6.5. Outlook

This first attempt at a full-scale numerical weather forecast for Mars has clearly revealed that such a task is extremely demanding on the accuracy of the model, despite the circulation of the Martian atmosphere being apparently somewhat less complex and chaotic than its terrestrial counterpart. The time-scale for the growth of forecast divergence in the model assimilation forecast study suggests that the radiative time-scale of the Martian atmosphere might set a practical limit for the time-scale of a forecast, due to the rapid onset of 'climate drift'. Such a conclusion suggests a relatively pessimistic outlook for practical weather forecasting on Mars, unless systematic errors in atmospheric models, and observational uncertainties in the distributions of atmospheric dust and water ice clouds, can be reduced to an acceptable level.

What is clear, however, is that this approach to the direct comparison of model simulations of synoptic meteorology on Mars with detailed observations via data assimilation is an extremely powerful and sensitive means of identifying and characterizing systematic model errors and biases. Such an approach should lead to major improvements in the formulation of numerical models and their parametrization schemes in future work.

References

Barnes JR. 1980. Time spectral analysis of midlatitude disturbances in the Martian atmosphere. *J. Atmos. Sci.* **37**: 2002–2015.

Barnes JR. 1981. Midlatitude disturbances in the Martian atmosphere: A second Mars year. *J. Atmos. Sci.* **38**: 225–234.

Collins M, Lewis SR, Read PL, Hourdin F. 1996. Baroclinic wave transitions in the Martian atmosphere. *Icarus* **120**: 344–357.

Conrath BJ, Pearl JC, Smith MD, Maguire WC, Dason S, Kaelberer MS, Christensen PR. 2000. Mars Global Surveyor Thermal Emission Spectrometer (TES) observations: Atmospheric temperatures during aerobraking and science phasing. *J. Geophys. Res.* **105**: (E4) 9509–9519.

Ding R, Li J. 2007. Nonlinear finite-time Lyapunov exponent and predictability. *Phys. Lett. A* **364**: 396–400.

Chen B, Li J, Ding R. 2006. Nonlinear local Lyapunov exponent and atmospheric predictability research. *Science in China series D: Earth Sciences* **49**: 1111–1120.

Dalcher A, Kalnay E. 1987. Error growth and predictability in operational ECMWF forecasts. *Tellus A* **39**: 474–491.

Forget F, Hourdin F, Fournier R, Hourdin C, Talagrand O, Collins M, Lewis SR, Read PL, Huot J-P. 1999. Improved general circulation models of the Martian atmosphere from the surface to above 80 km. *J. Geophys. Res.* **104**: (E10) 24155–24176.

Forget F, Wanherdrick Y, Lewis SR. 2001. 'Validation of the Mars General Circulation Model and Climate Database with new spacecraft observations'. Technical Report: Contract 11369/95/NL/JG. Work Package 7, ESA, LMD: Paris and AOPP: Oxford.

Goody R, Belton MJS. 1967. Radiative relaxation times for Mars. *Planet. Space Sci.* **15**: 247–256.

Hide R, Mason PJ. 1975. Sloping convection in a rotating fluid. *Adv. in Phys.* **24**: 47–99.

Hollingsworth JL, Barnes JR. 1995. Forced stationary planetary waves in Mars's winter atmosphere. *J. Atmos. Sci.* **53**: 428–448.

Houghton JT. 1986. *The Physics of Atmospheres*. (2nd ed.) Cambridge University Press: Cambridge, UK.

Houtekamer PL, Derome J Prediction experiments with two-member ensembles. 1994. *Mon. Weather Rev.* **122**: 2179–2191.

Kieffer HH, Jakovsky BM, Snyder CW, Matthews MS. (eds.) 1992. *Mars*. The University of Arizona Press: USA.

Leovy CB. 1985. The general circulation of Mars: Models and observations. *Adv. Geophys.* **28a**: 327–346.

Lewis SR, Barker PR. 2005. Atmospheric tides in a Mars general circulation model with data assimilation. *Adv. in Space Res.* **36**: 2162–2168.

Lewis SR, Collins M, Read PL. 1996. Martian atmospheric data assimilation with a simplified general circulation model: Orbiter and lander networks. *Plan. Space Sci.* **44**: 1395–1409.

Lewis SR, Read PL, Conrath BJ, Pearl JC, Smith MD. 2007. Assimilation of Thermal Emission Spectrometer atmospheric data during the Mars Global Surveyor aerobraking period. *Icarus* **192**: 327–347.

Montabone L, Lewis SR, Read PL, Hinson DP. 2006. Validation of martian meteorological data assimilation for MGS/TES using radio occultation measurements. *Icarus* **185**: 113–132.

Newman CE, Lewis SR, Read PL, Forget F. 2002. Modeling the Martian dust cycle, 1. Representations of dust transport processes. *J. Geophys. Res.* **107**: (E12: DOI: 10.1029/2002JE001910.

Newman CE, Lewis SR, Read PL. 2004. Investigating atmospheric predictability on Mars using breeding vectors in a general-circulation model. *Q. J. R. Meteorol. Soc.* **130**: 2971D2989.

Paige DA, Bachman JE, Keegan KD. 1994. Thermal and albedo mapping of the polar regions of Mars using Viking thermal mapper observations, 1, North polar region. *J. Geophys. Res.* **99**: 25959–25991.

Palluconi FD, Kieffer HH. 1981. Thermal inertial mapping of Mars from 60^{circ}S to 60^{circ}N. *Icarus* **45**: 415–426.

Read PL, Lewis SR. 2004. *The Mars Climate Revisited: Atmosphere and Environment of a Desert Planet*. Springer-Praxis: Berlin, Heidelberg, New York.

Read PL, Collins M, Früh W-G, Lewis SR, Lovegrove AF. 1998. Wave interactions and baroclinic chaos: a paradigm for long timescale variability in planetary atmospheres. *Chaos, Solitons and Fractals* **9**: 231–249.

Richardson MI, Wilson RJ. 2002. A topographic forced asymmetry in the Martian circulation and Climate. *Nature* **416**: 298–301.

Schofield JT, Crisp D, Barnes JR, Haberle RM, Magalh aes JA, Murphy JR, Seiff A, Larsen S, Wilson G. 1997. The Mars Pathfinder Atmospheric Structure Investigation/Meteorology (ASI/MET) experiment. *Science* **278**: 1752–1757.

Smith MD. 2004. Interannual variability in TES atmospheric observations of Mars during 1999-2003. *Icarus* **167**: 148–165.

Smith MD, Pearl JC, Conrath BJ, Christensen PR. 2000. Mars Global Surveyor Thermal Emission Spectrometer (TES) observations of dust opacity during aerobraking and science phasing. *J. Geophys. Res.* **105**: (E4): 9539–9552.

Smith DE, Zuber MT, Solomon S, Phillips R, Head J, Garvin J, Banerdt WB, Muhleman DO, Pettengill GH, Neumann GA, Lemoine FG, Abshire JB, Aharonson O, Brown CD, Hauck SA, Ivanov AB, McGovern PJ, Zwally HJ, Duxbury TC. 1999. The global topography of Mars and implications for surface evolution. *Science* **284**: 1495–1503.

Smith MD, Wolff MJ, Spanovich N, Ghosh A, Banfield D, Christensen PR, Landis GA, Squyres SW. 2006. One Martian year of atmospheric observations using MER Mini-TES. *J. Geophys. Res.* **111**: E12S13: ISI:000242945300002.

Spiga A, Forget F, Dolla B, Vinatier S, Melchiorri R, Drossart P, Gendrin A, Bibring JP, Langevin Y, Gondet B. 2007. Remote sensing of surface pressure on Mars with the Mars Express/OMEGA spectrometer: 2. Meteorological maps. *J. Geophys. Res.* **112**: E08S16, DOI: 10.1029/2006JE002870.

Toth Z, Kalnay E. 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bull. Am. Meteorol. Soc.* **74**: 2317–2330.

Wilson RJ, Hamilton RJ. 1995. Comprehensive model simulation of thermal tides in the Martian atmosphere. *J. Atmos. Sci.* **53**: 1290–1326.

Wilson RJ, Smith MD, Banfield D, Conrath BJ. 2002. Travelling waves in the northern hemisphere of Mars. *Geophys. Res. Lett.* **29**(:): DOI: 10.1029/2002GL014866.

Wilson RJ, Lewis SR, Montabone L, Smith MD. 2008. Influence of water ice clouds on Martian tropical atmospheric temperatures. *Geophys. Res. Lett.* **35**(:): DOI: 10.1029/2002GL032405.

