The seasonal cycle of water vapour on Mars from assimilation of Thermal Emission Spectrometer data

How to cite:
The seasonal cycle of water vapour on Mars from assimilation of Thermal Emission Spectrometer data

Liam J. Steele\textsuperscript{a,∗}, Stephen R. Lewis\textsuperscript{a}, Manish R. Patel\textsuperscript{a}, Franck Montmessin\textsuperscript{b}, François Forget\textsuperscript{c}, Michael D. Smith\textsuperscript{d}

\textsuperscript{a}Department of Physical Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom
\textsuperscript{b}Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS), CNRS, IPSL, 11 boulevard d’Alembert, F-78280 Guyancourt, France
\textsuperscript{c}Laboratoire de Météorologie Dynamique (LMD), CNRS, IPSL, 4 pl. Jussieu, 75252 Paris, France
\textsuperscript{d}NASA Goddard Space Flight Center, MS 693, Greenbelt, MD 20771, United States

Abstract

We present for the first time an assimilation of Thermal Emission Spectrometer (TES) water vapour column data into a Mars global climate model (MGCM). We discuss the seasonal cycle of water vapour, the processes responsible for the observed water vapour distribution, and the cross-hemispheric water transport. The assimilation scheme is shown to be robust in producing consistent reanalyses, and the global water vapour column error is reduced to around 2–4 pr-µm depending on season. Wave activity is shown to play an important role in the water vapour distribution, with topographically steered flows around the Hellas and Argyre basins acting to increase transport in these regions in all seasons. At high northern latitudes, zonal wavenumber 1 and 2 stationary waves during northern summer are responsible for spreading the sublimed water vapour away from the pole. Transport by the zonal wavenumber 2 waves occurs primarily to the west of Tharsis and Arabia Terra and, combined with the effects of western boundary currents, this leads to peak water vapour column abundances here as observed by numerous spacecraft. A net transport of water to the northern hemisphere over the course of one Mars year is calculated, primarily because of the large northwards flux of water vapour which occurs during the local dust storm around $L_s = 240°–260°$. Finally, outlying frost deposits that surround the north polar cap are shown to be important in creating the peak water vapour column abundances observed during northern summer.

Keywords: Mars, atmosphere, Mars, climate, Atmospheres, dynamics, Meteorology

1. Introduction

In the 50 years since the first detection of water vapour on Mars by Spinrad et al. (1963) there have been numerous studies of the martian water cycle, both observationally (with ground-based and space-based instruments), and more recently with computer modelling. Water vapour is the most variable trace gas observed on Mars, being affected by atmospheric circulation, cloud formation and interactions with the regolith and surface ice deposits (especially the polar caps). Its distribution, particularly vertically, affects other atmospheric processes through its role in photochemical reactions (Lefèvre et al., 2008), and through the radiative effects of the clouds it produces (e.g. Wilson et al., 2008; Madeleine et al., 2012). Our first detailed knowledge of the water vapour distribution came from the MAWD instruments aboard the Viking orbiters (Farmer et al., 1977; Jakosky and Farmer, 1982), with further systematic mapping following from the Thermal Emission Spectrometer (TES) aboard the Mars Global Surveyor (MGS; spacecraft (Smith, 2002, 2004; Pankine et al., 2010), the PFS; OMEGA and SPICAM instruments aboard Mars Express (e.g. Fedorova et al., 2006; Melchiorri et al., 2007; Tschimmel et al., 2008; Sindoni et al., 2011; Maltagliati et al., 2011, 2013) and CRISM aboard Mars Reconnaissance Orbiter (Smith et al., 2009). Observations reveal a water cycle dominated by sublimation and transport from the polar caps, and generally repeatable between years. The observed peak water vapour column abundance over the north pole ranges from around 50–65 pr-µm (Fouchet et al., 2007; Tschimmel et al., 2008; Smith et al., 2009; Maltagliati et al., 2011), while the south pole maximum is observed to be lower than the north at around 25 pr-µm (Smith et al., 2009; Maltagliati et al., 2011). The original TES retrievals showed higher water vapour column maxima over the north and south poles of 75–100 pr-µm and 50 pr-µm respectively (Smith, 2002, 2004). However, after a reanalysis to correct algorithm errors for spectra taken at lower spectral resolution (Smith, 2008) the retrievals were reduced by around 30%, bringing them into agreement with those from other spacecraft (Fouchet et al., 2007; Wolkenberg et al., 2011). The spatial distribution of water vapour shows two regions of increased abundance over Tharsis and Arabia Terra (Smith, 2002; Fouchet et al., 2007; Tschimmel et al., 2008; Smith et al., 2009; Maltagliati et al., 2011), which have been proposed to occur due to local regolith-atmosphere interactions or atmospheric transport.

In order to understand the increasing number of observations, models of the martian water cycle have been developed, progressing from early diffusion models and 2D zonally-symmetric models to fully 3D global models. Such models have performed well in obtaining zonally-averaged water cycles in

---

\*Corresponding author. Tel: +44(0)1908 659811.
Email address: liam.steele@open.ac.uk (Liam J. Steele)
Preprint submitted to Icarus January 18, 2016
agreement with observations, and have highlighted the importance of various processes acting on the water cycle, such as meridional winds caused by CO₂ condensation (James, 1985), eddy mixing (Richardson and Wilson, 2002), transport by water ice clouds (Richardson et al., 2002; Montmessin et al., 2004) and the effects of global dust storms and regolith interactions (Böttger et al., 2004, 2005). Cross-hemispheric transport of water vapour may also be affected by western boundary currents (WBCs), which have been shown in modelling work to occur along the eastward-facing flanks of the Tharsis Plateau and Syrtis Major (Joshi et al., 1994, 1995).

This study combines the benefits of both observations and modelling by coupling a data assimilation scheme (Lorenc, et al., 1991; Lewis and Read, 1995; Lewis et al., 1996) with a Mars global climate model (MGCM), and assimilating TES nadir temperature profiles and water vapour columns. The results of the temperature assimilation allow us to diagnose the impact and importance of the changing model dynamics on the water cycle. By additionally assimilating water vapour columns, we can investigate any remaining deficiencies, and further understand the role different modes of transport play in the water vapour distribution. Data assimilation has previously been used to assimilate martian temperature profiles and dust optical depths (Lewis and Barker, 2005; Montabone et al., 2005; Lewis et al., 2007; Wilson et al., 2008; Lee et al., 2011; Greybush et al., 2012; Hoffman et al., 2012), but this study represents the first assimilation of martian water vapour data. In §2 we give an outline of the MGCM used in the study, and in §3 we discuss the data being assimilated and the assimilation method. In §4 we verify the assimilation procedure before going on to present our results in §5. Our conclusions are given in §6.

2. Mars global climate model

The MGCM used for this study results from a collaboration between the Laboratoire de Météorologie Dynamique (LMD) of the University of Oxford and The Open University. It combines the LMD physical schemes (Forget et al., 1999) with a spectral dynamical core (Hoskins and Simmons, 1975), an energy and angular-momentum conserving vertical finite-difference scheme (Simmons and Burridge, 1981) and a semi-Lagrangian advection scheme for tracers (Newman et al., 2002). The model is run with a triangular spectral truncation at horizontal wavenumber 31, resulting in a physical grid resolution of 5° in latitude and longitude. There are 25 unevenly-spaced vertical levels in sigma coordinates, extending to an altitude of 100 km, with three ‘sponge layers’ at the model top to damp vertically propagating waves.

The physical parameterizations take into account gravity wave and topographic drag, vertical diffusion, convection and CO₂ condensation, as described in Forget et al. (1999). The water cycle is modelled by transporting water vapour and ice mass mixing ratios, and using the cloud scheme of Montmessin et al. (2004) to account for the formation and sedimentation of ice particles. The north polar cap is defined to be an unlimited source of water, and covers all grid points northwards of 85°N, and grid points between 80°–85°N and 0°–180°E. This cap was chosen to give a zonally-averaged latitude vs. time water vapour distribution in the free-running model comparable to that of recent spacecraft observations (e.g. Smith, 2004; Fedorova et al., 2006; Spondini et al., 2011; Maltagliati et al., 2011). Water ice can be deposited onto the surface at any location, with the surface albedo changed from that of the regolith (derived from TES data; Mellon et al., 2000) to that of water ice (0.4) when more than 5 pr-μm is deposited. Surface water ice can only be sublimed if there is no covering layer of CO₂ ice. The modelled water ice clouds are not radiatively active, though their radiative effects will be included via the assimilation of TES temperature profiles (discussed further in §4.3). The dust opacity is prescribed to follow the vertical profile

\[
\frac{\partial \tau}{\partial P} \propto e^{0.007 \left[ 1 - \left( \frac{P_{ref}}{P} \right)^{0.70 \text{km}/\text{mm}} \right]} \tag{1}
\]

for \( P \leq P_{ref} \), with \( P_{ref} = 610 \text{ Pa} \) and \( \tau_{max} \), varying with solar longitude \((L_s)\) and latitude (Forget et al., 1999; Lewis et al., 1999). The dust opacity is constant when \( P > P_{ref} \). The dust opacities in each layer of an atmospheric column are subsequently scaled by a factor \( \tau_{ref}/\tau_{model} \), where \( \tau_{model} \) and \( \tau_{ref} \) are respectively the total column dust optical depths from the model and from a global map of TES nadir dust optical depth retrievals (both referenced to the 610 Pa pressure level).

In order to obtain the daily global dust maps, first a weighted binning procedure was applied to the TES data, which produced incomplete time-space maps due to spacecraft coverage and acceptance criteria. The kriging interpolation method was then used in order to obtain daily maps with complete global coverage (Montabone et al., Eight-year climatology of dust optical depth on Mars, in preparation). Kriging applies a weighting to each of the grid points based on spatial variance and trends among the points. Rescaling the model dust opacities from the global dust maps replaces the direct assimilation of nadir dust retrievals that was performed in previous studies (e.g. Montabone et al., 2005; Lewis et al., 2007). As only daytime (2 p.m.) TES data are available to produce the dust maps, the diurnal variation of dust is not taken into account.

3. Data and assimilation scheme

The data assimilation scheme used is a modified form of the Analysis Correction (AC) scheme, originally developed for use at the UK Met Office (Lorenc et al., 1991) and re-tuned for martian conditions (Lewis and Read, 1995; Lewis et al., 1996, 1997). The AC scheme has previously been used to assimilate TES temperature and dust optical depth retrievals in order to study a variety of topics such as atmospheric tides (Lewis and Barker, 2005), dust storms (Montabone et al., 2005), transient waves and circulation (Lewis et al., 2007) and the radiative influence of water ice clouds (Wilson et al., 2008). The AC scheme has also been verified using independent radio occultation data by Montabone et al. (2006). This study represents the first assimilation of martian water vapour data, which is combined with an assimilation of TES nadir temperature profiles. The following sections give overviews of the data assimilated.
the quality control applied to the data prior to assimilation, and
the assimilation method.

3.1. Thermal Emission Spectrometer data

The data assimilated in this study are from the TES instrument aboard the MGS spacecraft. The spectrometer began its
mapping operations on 1 March 1999 ($L_s = 104^\circ$, MY 24) and ended on 31 August 2004 ($L_s = 81^\circ$, MY 27), where MY represents a Mars year following the designation of Clancy et al. (2000). MGS completed 12 orbits each sol, containing two sets of
12 narrow strips of data, running roughly north-south and separated by $\sim 30^\circ$ in longitude. Due to the sun-synchronous orbit of MGS, the two sets of data contain observations near 02:00 and 14:00 local time.

The temperature and water vapour column retrievals used in the assimilation, along with the retrieval methods, are described in detail by Conrath et al. (2000) and Smith (2002, 2004). Briefly, temperature profiles are retrieved from nadir radiances in the $15\mu$m CO$_2$ band, extending to an altitude of $\sim 40$ km. The vertical resolution is around one scale height $(\sim 10$ km), though the temperature retrievals provided for assimilation have a vertical resolution of one scale height $(\sim 2.5$ km).

Uncertainties for each profile are $\sim 2$ K, but are larger in the lowest scale-height above the ground due to possible errors in estimating the surface pressure (Conrath et al., 2000; Smith, 2004). Systematic errors in excess of 5 K may also be present in retrievals over the winter polar regions due to errors in the absolute radiometric calibration. Conrath et al. (2000) noted that evidence for this was found in retrievals with temperatures lower than the CO$_2$ condensation temperature.

Water vapour column retrievals are obtained by comparing synthetic spectra with TES spectra from rotation bands in the 28–42 $\mu$m region. In this region the TES signal-to-noise ratio is relatively high, and the contribution to the spectrum from water vapour can easily be separated from dust, water ice and surface contributions (Smith, 2002). When calculating the synthetic spectra, water vapour is assumed to be well mixed up to the water vapour condensation level, and zero above. The estimated uncertainties in water vapour column abundance due to random errors are $5$ pr-$\mu$m, while uncertainty in the pressure broadening of water vapour by CO$_2$ could result in a systematic 25% error. This was quoted to be up to 25% (Smith, 2002), though as noted by Pankine et al. (2010), recent analysis of the CO$_2$-broadening of water vapour by Brown et al. (2007) supports the value used in the original retrieval process, and hence the systemic error is likely to be much less than 25%. Due to the restrictions of the retrieval algorithm, water vapour retrievals are only provided for locations where $T_{\text{surf}} > 220$ K, which excludes data from the nighttime orbits, the winter polar regions and over the polar caps in all seasons. The number of TES retrievals varies with time, but each strip of data typically contains $\sim 700$ retrievals, spaced $\sim 0.15^\circ$ (10 km) apart in latitude.

As there are no nighttime water vapour column retrievals available in the TES data set, the assimilation cannot constrain the diurnal variability of water vapour. As such, the model physics (affected by the assimilation of temperature profiles) determines the water vapour evolution outside the assimilation time window. Fig. 4 of Titov (2002) presents the diurnal water vapour column variation from a range of Earth-based, space-based and in situ measurements, which shows diurnal fluctuations of around $10$ pr-$\mu$m. Fluctuations of $15$ pr-$\mu$m have also been observed in CRISM and Phoenix SSI data (Tamppari et al., 2010). However, a large diurnal variability was not detected in comparisons between TES and OMEGA observations (Maltagliati et al., 2011), suggesting that on a global scale the diurnal water vapour column variation is only a few pr-$\mu$m (a similar value was also determined by Melchiorri et al., 2009).
assimilation. As such, the TES ice optical depth retrievals are not assimilated here, and instead the model is allowed to predict clouds independently.

3.2. Quality control and pre-processing of data

The quality control applied to the temperature profiles follows the procedure described in Lewis et al. (2007). Profiles marked as bad by the retrieval algorithm are automatically rejected, with remaining retrievals filtered to remove any profiles with temperatures below 130 K (lower than the CO$_2$ condensation temperature), or temperatures above 300 K in the lower atmosphere (falling linearly to 220 K at 40 km altitude). This additional filtering is applied in order to remove excessively high or low temperatures which may cause problems with the model’s physical schemes, and also removes some of the retrievals affected by possible systematic errors in the radiometric calibration. The filtering removes around 0.5% of the available observations. The remaining profiles are then re-sampled to five standard pressure levels, spaced from -5 km to 45 km. Error ratios for each observation $k$ also need to be supplied to the AC scheme, where the error ratio is defined as $\varepsilon_k^2 = r_k^2/b_k^2$, with $r_k^2$ the observational plus representativeness error variance and $b_k^2$ the a priori model error variance. The value chosen for all the observations assimilated is $\varepsilon_k^2 = 1$, as used by Lewis et al. (2007), which implies that the model and observational errors are of comparable magnitude.

For the water vapour column retrievals, the observations and model values cannot be compared directly as each samples a different thickness of atmosphere (as the model resolution results in a smoothed topography field). Thus, all observed and modelled water vapour columns are rescaled to a reference pressure of 610 Pa before being compared by the assimilation. The water vapour column variation with latitude can be quite large because of differences in topography and temperature, but nearby measurements (scaled to the 610 Pa pressure level to remove the effects of topography) should conform to within a specified standard deviation. In order to investigate possible spurious observations, a buddy-check was applied on each strip of data. For each observation the buddy-check compared the water vapour column value to the mean of all the observations within a specified meridional distance $d$ either side of the observation, and highlighted the observation as an error if its value differed by a set number of standard deviations $\sigma$ from the mean. Numerous tests and investigation by eye showed that comparing observations over a distance $d = 300$ km ($\sim 5^\circ$ in latitude), and excluding any observations differing by more than $3\sigma$ (around 4–10 pr-µm depending on location and season) from the mean, gave the best combination of a large enough sample of data for comparison, the removal of spurious observations, and the retention of observations with natural variation. While there is a possibility that a genuine observation with a water vapour column value much higher or lower than surrounding values could be discarded by the buddy-check, such an observation would likely correspond to a small localized source that could not be represented in a global model. Additionally, as only a small fraction of the observations are discarded (\sim 0.5%) the buddy-check does not negatively impact on the amount of data available for assimilation.

The quality control checks for temperature profiles and water vapour columns are performed independently, so it is possible that if a temperature profile is rejected the associated water vapour column could still be assimilated. However, the number of temperature profiles rejected is only small, and those which fall below the CO$_2$ condensation temperature occur in the winter polar regions where there are no water vapour column retrievals available to assimilate. Additionally, the buddy-check procedure which is applied to the water vapour column data successfully filters out any potentially erroneous retrievals.

In order to determine the error ratios for each water vapour column observation, statistics on the water vapour column variability in the free-running model were compiled after all data had been sampled to 2 p.m. local time to match the observations, and the monthly trend had been removed. Calculation of the monthly standard deviation for each grid point showed model variability ranging from 1–3 pr-µm, with the largest values occurring during the summer periods in each hemisphere. Systematic errors may also be present in both the model and the TES data, though these are difficult to quantify. Since the zonally-averaged water vapour columns in both the TES data and the free-running model are comparable to data published from other spacecraft (e.g. Fedorova et al., 2006; Sindoni et al., 2011; Maltagliati et al., 2011), and the model forecast variations are comparable to the uncertainties in the TES data, the value of $\varepsilon_k^2 = 1$ was chosen for all water vapour column observations.

3.3. Assimilation method

The AC scheme is a form of successive corrections in which analysis steps are interleaved with each model dynamical time step, resulting in 480 analysis steps per sol for the assimilations in this study. The modified successive corrections equation used by the scheme is

$$x_n = x_0 + WQ[y_0 - H(x_0)],$$

(2)

where $x_n$ is the analysis vector, $x_0$ is the model background (forecast), $y_0$ is the observation vector, $H$ is the observation operator, and $W$ and $Q$ are matrices of weights and normalization factors respectively:

$$W = B H^T R^{-1}$$ and $$\hat{Q} = (H W + I)^{-1}.$$  

(3)

Here, $B$ is the background error covariance matrix, $R$ is the observation error covariance matrix, $H$ is the linearization of $H$ and $I$ is the identity matrix.

In each analysis step, Eq. 2 is split into separate vertical and horizontal stages in order to spread the analysis increments from the observation locations to the surrounding model grid points. This is followed by the derivation of multi-variate increment fields for dynamical balance where applicable (e.g. geostrophic wind adjustments are applied after assimilating temperatures). Observations are inserted over a six hour window, from five hours before their valid time until one hour after.
Such an insertion period is valid, as tests show that both temperature and water vapour columns in the model vary monotonically over these insertion periods (21:00–03:00 and 09:00–15:00), and so the assimilation will not unrealistically smooth out any inherent model variation. This asymmetrical time window is also beneficial as it biases the assimilation gains to regions ahead of the satellite ground track, which have not recently been observed. The assimilation scheme first assimilates temperature profiles (see Lewis et al., 2007 for a description of this procedure), before moving on to assimilate water vapour columns. As the water vapour observations sample the whole atmospheric depth, the model water vapour mass mixing ratios \( q_{im} \) in each layer \( m \) above grid box \( i \) are first converted into a water vapour column \( W_i \) at a reference pressure level of 610 Pa via

\[
W_{i \text{model}} = \frac{610 \text{Pa}}{P_{i \text{surf}}} \left( 1 - \frac{1}{g} \sum_m q_{im} \Delta P_{im} \right),
\]

where \( P_{i \text{surf}} \) is the surface pressure, \( \Delta P \) is the layer pressure difference and \( g \) is the mean acceleration due to gravity. The water vapour columns are then interpolated to the observation locations \( k \), and increments \( \delta W_k = W_{k \text{obs}} - W_{k \text{model}} \) are calculated. The horizontal spreading of the increments to model grid points then follows via

\[
\delta W_i = \lambda \sum_j \mu_{ij} \delta W_j / S(\delta t_j),
\]

where the sum is over all observations \( j \) within a specified influence area of grid point \( i \). Here, \( \delta W_j / S(\delta t_j) \) is a normalization factor which takes into account the ratio of observation to first guess model error, as well as the data density of the observations (in order to retain the ‘super-observation’ property of Optimum Interpolation). \( R(\delta t_j) \) is a function of the time difference \( \delta t_j \) between the current time and the observation’s valid time, and \( \alpha_{ij} \) is a correlation function of the horizontal distance \( r_{ij} \) between grid point \( i \) and observation \( j \)

\[
\mu_{ij} = \left( 1 + \frac{r_{ij}}{S(\delta t_j)} \right) \exp \left( - \frac{r_{ij}}{S(\delta t_j)} \right).
\]

where \( S(\delta t_j) \) is a time-dependent correlation scale, with \( S(-5 \text{ hr}) = 540 \text{ km} \) and \( S(0 \text{ hr}) = 340 \text{ km} \). The parameter \( \lambda \) is a relaxation coefficient which determines the fraction of the increment added to the model field. The forms of \( R(\delta t) \) and \( S(\delta t) \) used here are those used by Lorenc et al. (1991), which have been extensively tested with terrestrial assimilations of temperature, pressure, wind and relative humidity. Eq. 6 has a longer tail than a Gaussian curve, which was found by Lorenc et al. (1991) to be of benefit in data-sparse areas on Earth. As the TES data are sparsely distributed (particularly in longitude), this form of the correlation function is also relevant to the assimilation in this study. Additionally, assimilation of the water vapour column field is rescaled to the 610 Pa pressure level before assimilation occurs, many of the large gradients resulting from topographic variations are removed. Thus the field being assimilated has a more homogeneous spatial structure, benefiting the spreading of the assimilation increments which assumes a Gaussian error distribution.

To determine the influence area of the water vapour column observations, a series of tests were performed for three 30-sol periods around \( L_S = 90^\circ \), \( 180^\circ \) and \( 270^\circ \). Influence radii \( r_{\text{inf}} \) were varied from 1–6 correlation scales in steps of 0.5, and the global RMS errors from each test were compared to a free-running model. An influence radius of \( r_{\text{inf}} = 3 \text{S} \) was found to give the best results compared to the free-running model, though the AC scheme was found not to be very sensitive for choices of the correlation scale in the range 2–4. Smaller influence radii had higher RMS errors as the increments were not spread far from the observation locations, and so less corrections were applied to the model. Larger influence radii also had higher RMS errors, mainly caused by the incorrect spreading of data over large areas in regions of strong water vapour gradients. The selected influence radius is slightly smaller than the one used for temperature, where \( r_{\text{inf}} = 3.5 \text{S} \).

For the assimilation of fields strongly deformed by the wind (such as atmospheric water vapour), assimilation schemes often alter the influence area from a circle to a flow-dependent ellipse elongated in the direction of the wind (e.g. Faccani and Ferretti, 2005). While the AC scheme used here is not designed for the shape of the influence area to be modified by the wind, an alteration was made so that the influence area approximated an ellipse with the semi-major and semi-minor axes aligned east–west and north–south respectively. This ellipse shape was tested as the water vapour distribution generally shows a larger variation with latitude than longitude. Tests were performed for values of the semi-minor/semi-major axis ratio \( b/a \) ranging from 0.2–1 in steps of 0.2, and while the global RMS error decreases as the value of \( b/a \) decreases, the difference is only a few tenths of a pr-\( \mu \text{m} \). After 60 sols of assimilation, the resulting water vapour column fields showed little detectable difference between tests, so the circular influence area which the AC scheme was originally designed to use was kept.

As the assimilation increment \( \delta W_i \) is in the form of a column value, a decision had to be made on how to alter the model’s vertical water vapour profile in each column. We follow a similar approach to that used by Nakamura et al. (2004) for water vapour column assimilation on Earth, in which the water vapour mass mixing ratio in each layer is scaled via

\[
d_{\text{new}} = \alpha d_{\text{old}}, \quad \text{with} \quad \alpha_i = (W_i + \delta W_i)/W_i.
\]

Due to the microphysics scheme’s limit on the allowable amount of supersaturation (with water vapour scaled above the saturation curve potentially leading to excessive cloud formation) we do not scale the mixing ratios above saturation. Instead we use an iterative scaling method in which at each iteration the points on the profile below saturation are scaled by the factor \( \alpha \), but not exceeding saturation. If at the end of each iteration the model and observations do not agree, new scaling factors are calculated, and any point below saturation is rescaled. An example of this procedure is shown in Fig. 1. Tests over three different periods around \( L_S = 90^\circ \), \( 180^\circ \) and \( 270^\circ \) show that ~90% of profiles match the assimilation after one iteration, and 99% within 3 iterations.
Additionally, scaling is only applied to model layers with pressures greater than 1 Pa ($z \lesssim 60$ km). This is to avoid inelastic build up of water vapour. While there is a lack of detailed observational evidence of the global vertical water vapour distribution, spacecraft observations by Rodin et al. (1997), and ground-based observations by Clancy et al. (1992, 1996) show that water vapour is mixed up to around 200 km during aphelion season, possibly exceeding 45 km during perihelion season. Recent SPICAM observations by Maltagliati et al. (2013) show evidence for water vapour above 60 km during perihelion, though these high-altitude layers contribute only a small fraction to the water vapour column. Additionally, recent observations of the vertical ice distribution by the Mars Climate Sounder (MCS) instrument (McCleese et al., 2010) show that clouds are found throughout the year below ~1 Pa, which will limit any large abundances of water vapour above this level. Hence, the vertical cut-off in water vapour scaling used in this study is consistent with observations. It should be noted that while the assimilation does not increase the water vapour fields above the 1 Pa pressure level, the model’s physical schemes cannot still transport water vapour to and from these areas.

### 4. Assimilation verification

Before carrying out a multi-year reanalysis in order to study the dynamics of the water cycle, initial tests were performed in order to verify the assimilation procedure. Firstly, the impact of the initial conditions on the assimilation of water vapour columns was investigated, as if the assimilation is sensitive to the initial conditions then the results of the year-long assimilation will be ambiguous. Secondly, the performance of the model during periods where there is no water vapour column data to assimilate was investigated, to see how well the model retained the assimilated data. Such periods occur throughout the TES data set, with the largest data gaps (of the order tens of sols) due to solar conjunction, and smaller gaps caused by periods where the spacecraft was in safing mode. Lastly, assimilations of TES and MCS temperature profiles were compared in order to assess the ability of the assimilation in representing the global temperature structure, considering the TES temperature profiles only extend to ~40 km altitude with a 10 km vertical resolution.

#### 4.1. Sensitivity of assimilation to initial conditions

In order to determine the effect of the initial water vapour field on the assimilation, a series of tests were performed for three 80-sol periods around northern summer, autumn and winter, in which an ensemble of assimilations with different initial conditions were compared. Lewis et al. (2007) performed a similar investigation into the effect of the initial conditions on the large-scale, zonal-mean circulation in a temperature assimilation, and found the circulation to be insensitive to the initial state after around 10 sols. Here we obtain the initial conditions by sampling all fields from a free-running model every two sols in the range ±10 sols around the valid initialization time, and by scaling the water vapour mass mixing ratios in each case by ±20% to obtain a plausible range of values. Selecting model output every two sols allows sampling of the fields during different phases of short period eddies. Combined with a perturbation of the water vapour field by ±20%, this results in peak water vapour column differences between the ensembles of generally around 4–6 pr-µm, though it can vary between 2–15 pr-µm depending on location and season. This variability is comparable to the uncertainty in the TES retrievals, and the resulting water vapour fields are within the range of values observed by other spacecraft. Each test period therefore comprises an ensemble of 22 assimilations of temperature profiles and water vapour columns, each with different initial conditions.

For each period every two hours, the ensemble water vapour column variance $\sigma_g^2$ was calculated for each grid point $k$ via

$$\sigma_g^2 = \frac{1}{N_k} \sum_{m=1}^{M} (W_{k,m} - \overline{W}_k)^2,$$

where $M = 22$ is the ensemble size, $W_{k,m}$ is the water vapour column at grid point $k$ in ensemble member $m$, and $\overline{W}_k$ is the ensemble mean water vapour column at grid point $k$. A global standard deviation $\sigma_g$ was then calculated by averaging the variances over all $N_k$ grid points via

$$\sigma_g = \left( \frac{1}{N_k} \sum_{k=1}^{N_k} \sigma_g^2 \right)^{1/2}.$$ 

The results are shown in Fig. 2, where it can be seen that the standard deviations in each period converge to around 0.3–0.4 pr-µm. Globally-averaged water vapour columns during these periods are around 10–15 pr-µm, though peak values in northern summer, autumn and winter are around 65, 25 and 35 pr-µm respectively. Because of the lower water vapour column abundances during northern autumn and winter, the initial water vapour fields (scaled by ±20%) have less variation in magnitude, and so the ensembles converge more quickly than in the northern summer test.

Fig. 3 shows global maps of the water vapour column variances averaged over sols 70–80 of each test. It can be seen that during both northern autumn and winter (panels b and c) the largest variance in the ensembles occurs over the Arcadia, Acidalia and Utopia Planitias, which have been identified as storm zones in modelling work by Hollingsworth et al. (1996, 1997), and later supported by observations (Banfield et al., 2004; Wang et al., 2005). In these regions the water vapour field is strongly modified by eastward travelling baroclinic waves which rapidly mix the water vapour distribution (particularly latitudinally), causing deviations from the assimilated field and resulting in increased model variation. In northern summer the water vapour is located mostly in the northern hemisphere where the atmosphere is more baroclinically stable. In this case, the variance in the ensembles is confined to latitudes polewards of 60°N (panel a) and is caused by the large initial variability in the water vapour fields of ~15 pr-µm. In all cases the variability is mainly in the magnitude of the water vapour columns rather than their
location, and this variability is only a small fraction of the wa−635
temperature column values. Thus, the assimilation scheme is able to
reach (and retain) a consistent solution to the water vapour column field from a plausible range of initial conditions, and for periods throughout the martian year that experience widely different atmospheric conditions.

4.2. Sensitivity of assimilation to data gaps

In order to assess the performance of the model during periods where there are gaps in the TES water vapour column data, assimilations of temperature profiles and water vapour columns were initially performed over three 30-sol periods around north-summer, autumn and winter, after which the simulations continued to run for a further 60 sols without assimilation. The tests were then compared to a free-running model initialized with the same conditions. Fig. 4 shows the RMS errors for two tests over Ls = 174°–228° (southern spring) and Ls = 268°–332° (southern summer), both MY 24. The fluctuations in RMS errors on a daily basis are related to both the dynamical nature of the water vapour field (with some locations experiencing more diurnal variation than others) and modelling errors at the locations of the observations being assimilated. The number of observations assimilated also affects the RMS error calculation, and this is particularly noticeable around sol 568 where the large peak in RMS error is a statistical artifact resulting from the number of observations assimilated dropping over a short period from ~1000 to ~10.

For the first 30-sol of both tests, during which assimilation of temperature and water vapour columns was occurring, the water vapour column RMS errors are reduced to ~4% pr/mm. Upon cessation of assimilation, both tests begin to revert back to a water vapour field similar to the free-run, but with different relaxation times. The RMS errors in the southern spring tests are comparable to the free-running model less than 10 sols after the assimilation ceased. Conversely, the southern summer tests retains the assimilated data for longer, and still has a slightly lower RMS error than the free-running model 60 sols after the assimilation ceased. Again, these differences can be explained by the atmospheric conditions encountered in each test, with the water vapour field in southern spring mixed away from the as-simulated field by eastward travelling waves, particularly in the storm zones. During southern summer much less meridional mixing occurs close to the pole where the largest water vapour columns occur, and the RMS error only slowly increases as the assimilated water vapour is transported equatorwards. The result of the test around southern winter is not shown in Fig. 4, but falls between the two cases discussed above, with the RMS errors relaxing more quickly towards the free run value than in the southern summer case, but still being lower than the free run after 60 sols. Thus, for the largest gaps in the TES data of around 10 sols, the model will still produce a reliable output around the summer periods in each hemisphere. However, this may not be reflected in the RMS errors, which may show large variations because of the decreased number of observations available for their calculation.

4.3. Temperature assimilation validation

Here we choose to assimilate TES nadir temperature profiles rather than include radiatively-active clouds, as previous modelling work has shown that while the inclusion of radiatively-active clouds improves the agreement between model temperatures and observations in some locations, there are discrepancies in other locations (Madeleine et al., 2012). This will affect the atmospheric circulation, water vapour sublimation and cloud formation, leading to a water vapour distribution different from reality (e.g. Haberle et al., 2011) and invalidating studies of global water transport. The temperature assimilation will also compensate for the simplified vertical dust profiles used in the model. As the TES temperature profiles are assimilated in terms of layer thicknesses, any small-scale features that are apparent in the model temperature profile (which has a higher vertical resolution than the TES retrievals) will not be disrupted. As the resolution of the TES retrievals is around 10 km, any thermal effects from vertically thin features such as ice hazes are unlikely to be captured. The TES temperature profiles also only extend to around 10 Pa (~40 km above the surface), so will not capture the thermal effects of any high-altitude clouds or dust layers which may be present (e.g. McCleese et al., 2010; Heavens et al., 2011; Guzewich et al., 2013a).

For the region the TES temperature profiles do cover, it has been shown previously that the temperature assimilation can capture the thermal signature of the aphelion cloud belt (Wilson et al., 2008), and comparisons between assimilations of TES temperature profiles and radio occultation (RO) profiles using the Ultra Stable Oscillator aboard MGS show that the assimilations capture many of the features seen in the RO profiles (Montabone et al., 2006). The assimilation of temperatures has also been shown to capture thermal tides and transient waves in agreement with observations (Lewis and Barker, 2005; Lewis et al., 2007), providing further validation of the temperature assimilation procedure. Additionally, comparison between TES and PFS temperatures when their observation periods overlapped (between Ls = 330°, MY 26 and Ls = 77°, MY 27) shows that most retrievals are in agreement between instruments (with the temperature differences contained within the retrieval errors), though some larger discrepancies do exist at specific locations such as Hellas, Argyre and Valles Marineris (Wolkenberg et al., 2011). This gives further validation that the TES temperatures being assimilated are representative of the true atmosphere.

In order to assess the ability of the assimilation of TES temperature profiles in capturing the atmospheric thermal structure (particularly at altitudes above the region covered by the TES retrievals), zonally-averaged temperatures were compared between a free-running model (FR), an assimilation of TES temperature profiles (hereafter referred to as TES-TA for ‘TES temperature assimilation’) and an assimilation of MCS temperature profiles (referred to as MCS-TA for ‘MCS temperature assimilation’). Table 1 gives an overview of each simulation. The FR uses the dust profile given by Eq. 1, and does not include radiatively-active clouds. The same model and initial conditions were used for the TES-TA and MCS-TA tests, but in the MCS-TA the number of vertical levels was increased from 25 to...
in order to make the vertical model resolution in the middle
and upper atmosphere comparable to that of the MCS observa-
tions being assimilated (~5 km). The appropriate global dust
maps were also used in order to scale the model dust opacities
in each test (MY 24 and 25 for the FR and TES-TA tests, and
MY 30 for the MCS-TA test). The results of the tests averaged
over 30° of Ls during northern summer, autumn and winter are
given in Fig. 5, while the differences between the temperatures
in the FR and TES-TA are given in Fig. 6. Note that the TES
temperature retrievals extend to around 10 Pa while the MCS
retrievals extend to around $10^{-2}$ Pa (i.e. above the upper plot
boundary in Fig. 5). It is important to note that different Mars
years with are being compared between assimilations, with the
different dust distributions leading to modifications of the tem-
perature structure and global circulation. As such, exact agree-
ment is not to be expected when comparing the temperature
profiles between Mars years. This is particularly true of the
polar warmings, which are sensitive to changes in the dust dis-
tribution (e.g. Guzewich et al., 2013b).

In each season the equatorial middle-atmosphere tempera-
tures in the TES-TA are warmer than those in the FR, with the
largest temperature increase of ~10 K occurring in north-
ern summer when the tropical clouds achieve their largest op-
tical depth (i.e. when the aphelion cloud belt is present - see
Fig. 6a). This results in temperatures below 10 Pa in agree-
ment with those in the MCS-TA, while above this level there
is a tendency for the TES-TA tropical temperatures to be ~5–70 K
too warm. There is little change in the temperature struc-
ture over the northern summer pole between the FR and TES-
TA, though the near-surface temperatures are increased by up
to 5 K in the TES-TA, bringing them in line with those in the
MCS-TA. A larger change is seen over the southern summer
pole (Fig. 6c), where temperatures in the TES-TA (above, be-
low) the 10 Pa level are (decreased, increased) by 5–10 K. This
brings the temperature profile into agreement with that obtained
in the MCS-TA, even in the upper atmosphere where there are
no TES retrievals to assimilate. In the autumn/winter polar re-
gions the TES-TA has slightly warmer polar vortices compared

In each season the equations of state of the model were
integrated with the smoothed topography here. For example, the wind
speed in the near-surface WBC during northern summer ap-
pears weaker than in reality, as the water vapour assimilation
has to remove additional water vapour from this region. Thus,
there may still be errors in the near-surface circulation even
after assimilating temperatures, but it is difficult to quantify this
error because of the lack of direct wind observations.

5. Assimilation results

5.1. Assimilations performed

To study the processes responsible for the observed water
vapour distribution, assimilations were carried out for one mar-
tian year from $L_s = 180°$, MY 24 to $L_s = 180°$, MY 25. This
period was chosen as it avoids the global dust storm of MY 25
and is less affected by data gaps than other periods, and as such
the number of observations remains high throughout. In order
to assess the impact of assimilation on the modelled water cy-
cle, two assimilations were carried out alongside a control run
(CR); see Table 2. The two assimilations consisted of a temper-
ature profile assimilation (hereafter referred to as ‘TA’) and
an assimilation of both water vapour columns and temperature
profiles (referred to as ‘VA’). In order to provide the CR, TA and
VA with the best possible initial conditions at $L_s = 180°$, a free-
running model (FR) was initialized with a north polar cap (but
no atmospheric water), and was run using MY 24 dust condi-
tions until the water cycle reached a quasi-steady state after six
years. A temperature profile and water vapour column assimi-
lation was then performed between

A map showing the latitudinal distribution of the water
vapour column observations with time is given in Fig. 7. As
can be seen, each bin of 5° latitude and 2° $L_s$ typically contains
around 1000–1500 observations to assimilate, though at certain
different dust distributions.

While the temperature differences outlined above will have an
effect on the upper-atmosphere circulation, the bulk of the
water vapour mass is concentrated in the lower atmosphere,
with the model showing that in all seasons >90% of the water
vapour mass in a column is located below 20 km, and >99% is
located below 40 km. The mean meridional circulation shows
slight changes due to the increased upper-atmosphere tempera-
tures, but the differences in mass transport are only a few per-
cent in the regions where water vapour is found. Also, when
water vapour achieves its greatest vertical extent (in the south-
ern hemisphere during southern summer; see Fig. 12) the ver-
tical temperature structure in the TES-TA is in good agreement
with that in the MCS-TA. Thus, while the TES temperature as-
similation may not capture some of the dynamical processes
in the upper atmosphere, this should not negatively impact the
study of the time- and zonal-average meridional water vapour
transport. However, some errors may still remain in modelling
the near-surface circulation, as this is highly dependent upon to-
pographical features which may not necessarily be represented
using the smoothed topography here. For example, the wind
speed in the near-surface WBC during northern summer ap-
pears weaker than in reality, as the water vapour assimilation
has to remove additional water vapour from this region. Thus,
there may still be errors in the near-surface circulation even
after assimilating temperatures, but it is difficult to quantify this
error because of the lack of direct wind observations.

A map showing the latitudinal distribution of the water
vapour column observations with time is given in Fig. 7. As
can be seen, each bin of 5° latitude and 2° $L_s$ typically contains
around 1000–1500 observations to assimilate, though at certain
times this can increase to $>3000$. Also evident is the lack of observations in the winter hemispheres and over the summer poles, as well as the data gaps resulting from solar conjunction and the MGS spacecraft entering safing mode.

5.2. Overview of assimilation results

The zonally-averaged water vapour column field from the VA is shown in Fig. 8a, along with the VA–TA difference (Fig. 8b), and the VA–CR difference (Fig. 8c). For the same period, Fig. 9a shows the number of water vapour column observations assimilated and Figs. 9b and 9c show the RMS and mean global water vapour column errors respectively for both assimilations and the CR. It can be seen in Fig. 8a that there is an abrupt change in the water vapour field around $L_\phi = 85^\circ$–$100^\circ$ and for the same period Fig. 9a shows a distinct change in the number of observations available for assimilation. This is related to a period where the TES instrument changed between 12.5 cm$^{-1}$ and high 62.5 cm$^{-1}$ spectral resolution, and in the water vapour column increase and associated RMS error in the increase around $L_\phi = 90^\circ$ are artifacts of this procedure. Large spikes in RMS error that occur over a short period (most noticeable in Fig. 9b during the period $L_\phi = 300^\circ$–$360^\circ$) are caused by rapid decreases in the number of observations available to calculate the RMS error, rather than model errors.

As shown in Fig. 9b, the TA generally results in small improvements to the water vapour field throughout the year, with larger improvements notable between $L_\phi = 260^\circ$–$285^\circ$ and $L_\phi = 350^\circ$–$70^\circ$ due to improved modelling of the transport of water vapour towards and away from the south pole respectively. The VA results in a reduction of the RMS error to around 2–4 pr-$\mu$m depending on season. This fluctuation is representative of the variability of the water vapour field in the FR, which is at a minimum between $L_\phi = 330^\circ$–$90^\circ$ (northern spring), corresponding to the southern summer peak water vapour abundance and resulting in the northern summer peak values.

5.3. Global water vapour transport

To study the modes of water vapour transport throughout the year, we decompose the time- and zonal-average meridional water vapour transport into its constituents following Peixoto and Oort (1992)

$$[\bar{\nabla}q] = [\bar{q}] [\bar{\nabla}] + [\bar{q}^*\bar{\nabla}^*] + [\bar{q}^\prime\bar{\nabla}^\prime].$$

where $[\bar{q}] [\bar{\nabla}]$ represents transport by the mean meridional circulation, $[\bar{q}^*\bar{\nabla}^*]$ is transport by stationary waves and $[\bar{q}^\prime\bar{\nabla}^\prime]$ is transport by transient eddies (with the star and prime symbols representing the contributions from the zonal and time variations respectively, i.e. $q^* = q - [q]$ and $q^\prime = q - \bar{q}$). To remove any signals associated with the diurnal cycle from the transient eddy term, $q^\prime$ is filtered in order to retain only those variations with periods greater than one sol. Time averages are then taken over 20-sol periods, which is long enough to sample the passing of several transient eddies over each location, which typically have periods less than 10 sols (Hinson and Wilson, 2002; Hinson, 2006). Tests with different averaging periods were also performed, but the results are not particularly sensitive to periods greater than 10 sols, with only slight differences in the magnitudes of the fluxes and no difference in latitudinal location. Eq. 10 can be integrated over the entire atmospheric column to obtain the total meridional flux across each latitude circle

$$[\overline{Q}_\phi] = \frac{2\pi R \cos \phi}{g} \int_0^{\phi_{\text{out}}} [\bar{\nabla}q] \, dP,$$

which again can be decomposed into terms relating to the mean meridional circulation, stationary waves and transient eddies. The locations of stationary waves are determined from the vertically-averaged meridional velocity deviation from the zonal mean (see Fig. 11). Since water vapour is mainly concentrated within around 10–20 km of the surface, its flow (and hence the calculated fluxes) are representative of the general circulation in the lower atmosphere.

As shown by Lewis et al. (2007), the assimilation of TES temperature profiles is capable of capturing the transient wave behaviour seen in observations, rather than just modifying those predicted by the model due to the altered thermal state. However, even if the circulation is well represented, other modelling errors (such as incorrect sublimation from the polar caps) will alter the water vapour distribution and hence the calculated water vapour transport. As such, we study the water vapour transport in the VA, as this combines the assimilation of both TES temperature profiles and water vapour columns, and so gives us our most complete understanding of the role of eddies and waves in the evolving water vapour field.
We now discuss the water vapour assimilation results, consi-

dering each season in turn. Fig. 12 shows the evolution of

the zonally-averaged water vapour field over the assimila-

tion period (as well as the changes to the mean meridional circu-

lation), while Fig. 13 shows evolution of the vertically-integrated

meridional flux \( \mathbf{\Omega}_m \). Note that the peaks in water vapour

mass mixing ratio between 1–10 Pa over each pole during the

equinoxes (Fig. 12, panels a and e) are artifacts resulting from

the meridional spreading of water vapour column data by the as-

similation scheme over regions where the majority of the lower

atmosphere is at saturation. In these cases, the rescaling proce-

dure can only add water vapour above the polar vortices.

5.3.1. Northern hemisphere autumn, \( L_S = 180^\circ–270^\circ \)

Initially the water vapour field is confined below \( \sim 20 \) Pa with

peak water vapour column values around \( 0^\circ–20^\circ N \) correspond-

ing to transport around the Hadley cells. Both stationary waves

and transient eddies generally act to transport water vapour

polewards in each hemisphere, with their relative contribu-

tions again becoming confined between the mid-latitudes, with

wave activity playing a large role in water vapour transport out-

side the \( 0^\circ–20^\circ N \) region (Fig. 13e). Transient eddies account

for almost all the transport polewards of \( 45^\circ N \), with transport

occurring throughout the lower atmosphere (up to \( \sim 50 \) Pa) and

across all longitudes, with slightly increased transport around

Acidalia Planitia.

5.3.2. Northern hemisphere winter, \( L_S = 270^\circ–360^\circ \)

In this season wave activity again begins to play a larger role

in water vapour transport, and by mid-winter waves are the ma-

jor source of northwards transport at mid-latitudes (Fig. 13d).

In the northern hemisphere the main northward transport is to

the west of Tharsis and Arabia Terra via a wavenumber 2 sta-

tionary wave, while in the southern hemisphere it is to the west

of the Hellas and Argyre basins via topographically steered

flows (Fig. 11b). Transport by transient eddies occurs primar-

ily in the Acidalia and Utopia Planitias, regions identified from

modelling and observations as storm zones (Hollingsworth et al.,

1996, 1997; Banfield et al., 2004).

The south pole reaches its peak water vapour column value

of around 30 pr-\( \mu \)m between \( L_S = 270^\circ –330^\circ \), but because of

an intensifying southern hemisphere Hadley cell, the north-

wards transport of water vapour is limited. Combined with

decreasing atmospheric temperatures, the majority of the wa-

ter vapour recondenses back onto the pole between \( L_S = 300^\circ–

360^\circ \) (Figs. 12d and 12e). By \( L_S = 360^\circ \) the water vapour dis-

tribution again becomes confined between the mid-latitudes, with

wave activity playing a large role in water vapour transport out-

side the \( 0^\circ–20^\circ N \) region (Fig. 13c). Transient eddies account

for almost all the transport polewards of \( 45^\circ N \), with transport

occurring throughout the lower atmosphere (up to \( \sim 50 \) Pa)

and across all longitudes, with slightly increased transport around

Acidalia Planitia.

5.3.3. Northern hemisphere spring, \( L_S = 0^\circ–90^\circ \)

As in the southern spring season, the transport of water

vapour polewards from the tropics is entirely caused by wave

activity, with transient eddies initially playing the largest role

(with transport occurring across all longitudes), and station-

ary wave transport becoming more important by mid-spring

(Figs. 13e and 13f). The peak in stationary wave transport

around \( 30^\circ–40^\circ S \) between \( L_S = 200^\circ–220^\circ \), which can be seen in Fig. 8a. The increas-

ing atmospheric temperatures also allow the water vapour to

spread higher into the atmosphere, reaching its largest vertical

extent by \( L_S = 270^\circ \). Northwards of \( 20^\circ N \) the combination of \( \lambda = 3 \)

wavenumber 2 stationary wave and transient eddies increase the

polewards transport of water vapour in the downwelling branch

of the Hadley cell. The seasonal ice deposits around the south-

pole cap also begin to sublimate towards the end of this sea-

son, supplying extra water vapour to the lower atmosphere over

60°–90° S. As there is little meridional transport at this time

(Fig. 13c), the majority of the sublimed water vapour is returned

as water ice to the edge of the retreating ice cap, limiting the at-

mospheric water vapour content until around southern summer

solstice. Note the large northward water vapour fluxes result-

ing from the mean meridional circulation in Figs. 13b and 13c; wa-

ter vapour transport during southern summer is greater than \( \lambda = 3 \)

any other time of year.
### 5.3.4. Northern hemisphere summer, \( L_S = 90^\circ - 180^\circ \)

This season sees the continuation of the north polar cap sublimation until \( L_S = 120^\circ \), with water vapour column values reaching their maximum of \( \sim 50 \, \text{pr}-\text{\mu m} \) during this time (Fig. 8a). The whole season is characterized by a gradual southward transport of the sublimed water vapour from the north pole to equatorial regions. For the first half of the season, transport by wave activity is almost exclusively southwards (Fig. 13g), primarily occurring through the Tharsis and Arabia Terra regions via topographically steered flows (Fig. 11d). South of the equator the southward transport by wave activity opposes the northward transport by the mean meridional circulation, limiting the northward flux of water vapour across the equator.

In the second half of the season, the changing mean meridional circulation generally acts to transport water vapour in the southern hemisphere northwards, and water vapour in the northern hemisphere southwards (Fig. 13b). In the southern hemisphere transport by stationary waves and transient eddies remains southwards, and their combined magnitude is generally larger than that of the mean meridional circulation, resulting in a net southward transport of water vapour. Again, the main transport routes are through Tharsis and Arabia Terra. In the northern hemisphere, the transport polewards of the mid-latitudes changes as autumn equinox approaches, with increasingly northward transport by wave activity. The increase in the prominence of transient waves is the result of the increasing temperature gradient as the north pole begins to substantially cool. The increasing stationary wave flux around 30\(^\circ\)N corresponds to a strengthening zonal wavenumber 2 stationary wave, while transport further south mainly occurs through the Tharsis and Arabia Terra regions via topographically steered flows (Fig. 11d). South of the equator the southward transport by wave activity opposes the northward transport by the mean meridional circulation, limiting the northward flux of water vapour across the equator.

Considering next the mass transport per sol (Fig. 14b) it can be seen that ice is transported southwards in the first half of the year (\( L_S = 0^\circ - 180^\circ \)) and northwards in the second half. The southwards transport is associated with the ACB which forms at altitudes of around 10 km, while the northwards transport is the result of high-altitude clouds which form around 30 km (see Fig. 12). For water vapour, the peak northwards transport around \( L_S = 250^\circ \) corresponds to transport in the upper branch of the Hadley cell. After this time, the water vapour experiences downwelling and eventually enters the return branch of the Hadley cell, causing southwards water vapour transport which peaks around \( L_S = 290^\circ \). For the remainder of the year water vapour generally experiences southwards transport, but this does not compensate for the large northwards transport during southern summer.

### 5.4. Cross-hemispheric water transport

While the previous section discussed the global transport of water vapour, here we consider the cross-hemispheric transport of both water vapour and ice over one Mars year. Fig. 14a shows the total mass transport per year over each model 5\(^\circ\) lon-gitude bin, while Fig. 14b shows the total equatorial mass transport per sol. Considering first the longitudinal variation in mass transport, it can be seen that water vapour transport shows more complex structure than ice transport, which is related to differences in their vertical distributions. As water vapour is located close to the surface, it is affected by complex topographic patterns, in particular stationary waves which are forced by topographic features. For example, between \( L_S = 180^\circ - 360^\circ \) the northwards transport between 0–100\(^\circ\)E and the southwards transport between 100–120\(^\circ\)W are associated with a wavenumber 2 stationary wave. The transport between 100–120\(^\circ\)E is affected by a topographically steered flow to east of Isidis Planitia, while the northwards water vapour peak centred on 130\(^\circ\)W is associated with a topographically steered flow to the west of the Tharsis Montes. Ice shows southwards transport at most longitudes, mainly due to the ACB, with peak transport centred around 120\(^\circ\)W (over the Tharsis region) as this is where the thickest clouds form.

### 5.5. Identifying model biases

The previous sections discussed the global transport of water in the VA, which benefits from the assimilation of TES temperature profiles which can correct errors in the modelled atmospheric circulation. However, while the circulation is improved, errors can still be present when representing physical processes such as the sublimation of surface ice deposits or ice cloud formation, and these can affect the global water vapour distribution. The VA overcomes these problems by directly assimilating water vapour columns, but study of the CR and TA allows for the investigation of biases present in the model, and for an assessment of the impact of the changing dynamics on the water cycle, both of which provide additional information on the processes affecting the water cycle.
As shown in Figs. 8b and 8c, the latitude-time water vapour column distributions in the CR and TA display biases compared to the VA. Firstly, there is a wet bias in the northern mid-latitudes, which is present throughout the northern autumn and winter (Ls = 180°—360°). Secondly, peak water vapour abundances occur over the south pole around Ls = 260°, whereas observations show the peak occurs later around Ls = 290°—310°. Thirdly, there is excess northwards water vapour transport from the southern polar region around the northern spring equinox to the north polar cap during northern summer (Ls = 90°—120°), producing a large dry bias. In this section we discuss the reasons for these biases and what they reveal about the water cycle.

The main source of bias is the excessive northwards transport of water vapour in the CR via the Hadley circulation between Ls = 180°—240°, a period containing a large (but not global) dust storm. This results in a wet bias in northern mid-latitudes and a dry bias southwards of ~30°S. The excess water vapour in the north then remains for much of the year until northern hemisphere summer. Comparing Figs. 8b and 8c it can be seen that the TA results in slightly wetter northern mid-latitudes and slightly drier tropics compared to the CR between Ls = 210°—240°. The reason for this is because of the strengthening of the Hadley cell in the TA, which increases the northwards transport of water vapour.

Fig. 15 shows the VA–TA zonally-averaged water vapour mass mixing ratio difference, and mean meridional circulation from the VA, for a 20-sol period centred on Ls = 225° (a time corresponding to the beginning of the local dust storm). As can be seen, the VA has slightly reduced the water vapour content above around 30 Pa, with a larger reduction northwards of ~30°N corresponding to the downwelling branch of the Hadley cell. Thus, the VA provides information on the vertical distribution of water vapour at this time, showing that the majority of the excess water vapour needs to be confined below ~30 Pa in order for it not to be transported too far north by the Hadley circulation.

Taking the meridional circulation predicted by the TA to be correct (as it is calculated from the best representation of the atmospheric temperature), the remaining wet bias in the TA is therefore the result of an incorrect vertical water vapour distribution, which may result from a combination of incorrect model assumptions of clouds or ignoring the adsorptive properties of the regolith. During the time of the dust storm, the main northward transport of water vapour in the model is around the Tharsis region. While the observed zonal-mean ice optical depths are low at this time (see Fig. 10a), TES observations show relatively high cloud optical depths over the Tharsis Montes, with lower opacity ‘cloud bridge’ connecting to the north polar hood via Alba Patera. A similar feature was seen by Benson et al. (2011) in observations with the MCR instrument during MY 229, but these clouds are not predicted by the model. Thus, these clouds may act to confine the water vapour field to lower levels through sedimentation processes, reducing the northward transport of water vapour in the Hadley cell. Additionally, Böttger et al. (2004) showed that when including an active regolith in the modelled water cycle, the effect of a global dust storm between Ls = 180°—270° is to increase the adsorptive capacity of the regolith, resulting in a water vapour reduction in northern mid-latitudes. While the dust storm of MY 24 is not global, regolith adsorption may still play a role in limiting the water vapour abundance. Thus, both cloud formation and regolith adsorption are possible processes that could explain the remaining deficiencies in the model during this time, though we cannot say which (if any) would play a more dominant role.

The next main source of bias is the sublimation of the south polar seasonal ice deposits. This results in a peak water vapour abundance occurring earlier than observations suggest, producing a wet bias between Ls = 240°—270° followed by a dry bias between Ls = 270°—330°. The ice deposits actually begin to sublimate around Ls = 180°, but up until Ls = 240° the water vapour sublimed is returned as water ice to the edge of retreat, limiting the atmospheric water vapour content. However, the increasing atmospheric temperature after Ls = 240° results in the sublimed water vapour remaining in the atmosphere, hence a peak water vapour abundance occurs. There is no improvement between the CR and TA, as assimilating temperature profiles actually increases temperatures over the south polar region by around 10 K compared to the CR.

Taking the assimilated temperature field to be correct, this implies that either the modelling of the CO2 cycle itself is incorrect (with more CO2 subliming than occurs in reality, allowing the water ice underneath to sublimate), or the CO2 sublimation is correct, and there is some other process acting as a sink for the sublimated water vapour, such as regolith adsorption or cloud formation. Böttger et al. (2004) showed that when including an active regolith, the southern hemisphere water vapour peak was shifted later in time (more in agreement with observations) because of increased adsorption of water vapour during a dust storm. Additionally, TES observations of atmospheric ice around Ls = 240°—270° show that a band of clouds exists polewards of 75°S with absorption-only optical depths at 825 cm−1 up to 0.1 (see Fig. 10a), but these clouds are absent in both the TA and the CR. The correct modelling of these clouds may help to reduce the excess water vapour amounts found at this time, though the quantities of water they hold is likely only a small fraction of the excess in the model.

The third main source of bias is the excess water vapour transport from the south pole during late southern summer and autumn. In the CR, a wet bias of ~5 pr-μm can be seen moving from 40°S at Ls = 330° to the equator at Ls = 50°. The TA is an improvement on the CR, reducing the wet bias by a few pr-μm, and this is due to improved circulation. While the main equator-crossing Hadley cell is broadly similar in both the TA and CR, the TA also has a detectable anticlockwise Hadley cell, located between around 30°—60°S (as in Fig. 12d). While this cell is much weaker than the equator-crossing cell, it impedes the northward transport of water vapour, confining more to the south polar region. The reason for the stronger anticlockwise Hadley cell in the TA is because the TA reduces the temperatures over the south pole more than those over the equator, and also increases the near-surface temperatures by a few K between 30°—60°S during Ls = 270°—315° (which is the region where the warmest temperatures are found at this time). The effect of this increased temperature gradient is to drive a stronger
thermally-direct anticlockwise circulation. While the TA is an improvement, the wet bias still remains. This is likely because of excess water vapour initially transported from the south pole, which in reality will remain in the polar region, turning to ice and falling to the surface as autumn approaches.

The last main source of bias is the low sublimation over the north pole between around $L_\odot = 90^\circ$–$120^\circ$, resulting in dry bias which moves equatorwards as the summer season proceeds. This bias is related to the distribution of the northern polar ice cap in the model, which was chosen to give a zonally-averaged latitude vs. time water vapour distribution in the FR comparable to recent observations (with the peak water vapour amounts differing by less than 5 pr-$\mu$m compared to observations). However, due to the water cycle in the FR being wetter in the northern mid-latitudes than in the assimulations, this is likely to be due to the large northwards flux of water vapour which occurs during northern autumn. This is due to the effect of eastward travelling baroclinic waves, which rapidly mix the model water vapour field away from the assimilated field. The locations of largest variance are the Arcadia, Acidalia and Utopia Planitias, which have previously been identified as storm zones. The model has also been shown to retain the assimilated data for longer during northern and southern summer than during spring and autumn, with global RMS errors in the water vapour field during southern summer remaining lower than in a free-running model up to 60 sols after cessation of assimilation. Again, this is related to the atmospheric conditions, with low wave activity during summer resulting in less horizontal mixing of the assimilated water vapour field.

Results of a year-long assimilation from $L_\odot = 180^\circ$ of MY 24 to $L_\odot = 180^\circ$ of MY 25 show that the assimilation of water vapour columns reduces the model RMS error to 2–4 pr-$\mu$m. The fluctuation of the error with season is related to the variability of the water vapour field in the free-running model, which is at a minimum around northern spring. The assimilation also highlights the importance of transient eddies, stationary waves and WBCs in the evolution of the water vapour distribution.

Stationary waves around the Hellas and Argyre basins act as regions of increased water vapour transport, channeling water vapour southwards along their eastern flanks during northern summer and autumn, and northwards along their western flanks during southern summer. Transport of water vapour away from the polar ice cap during northern summer is via a zonal wavenumber 1 stationary wave, while during the remaining seasons zonal wavenumber 2 waves at high northern latitudes act to transport water vapour northwards, extending the northern edge of the water vapour field to around 50$^\circ$N. This transport occurs primarily to the west of Tharsis and Arabia Terra and, combined with the effects of western boundary currents, this leads to peak water vapour column abundances here as observed by numerous spacecraft (e.g. Smith, 2002; Fouchet et al., 2007; Tschimmel et al., 2008; Smith et al., 2009; Maltagliati et al., 2011). Transient eddy transport is generally in the same direction as that by stationary waves, acting to increase the transport of water vapour. However, apart from at high latitudes around the equinoxes, it has a smaller magnitude than transport by stationary waves.

The assimilation of water vapour columns and temperature profiles suggests a net transport of water to the northern hemisphere over the course of one Mars year. This is mainly caused by the large northwards flux of water vapour which occurs during the local dust storm around $L_\odot = 240^\circ$–$260^\circ$ of MY 24. While different to previous modelling studies which predict a net southwards water transport (e.g. Jakosky, 1983; James, 1990; Houben et al., 1997; Richardson and Wilson, 2002), the results here are in agreement with Clancy et al. (1996) who suggest that water is preferentially transported towards the hemisphere experiencing summer during aphelion. However, the ACB optical depth is lower in the assimilation compared to observations, and so we may be underpredicting the southwards transport of water ice at this time of year.

Comparisons of the water vapour column field between the temperature and water vapour column assimilations highlight four main areas of model bias, relating to excessive northwards
transport of water vapour via the Hadley circulation between $1392^\circ - 240^\circ$ and $1392^\circ - 50^\circ$, as well as errors modulating the sublimation of the north and south polar ice deposits.

The excessive Hadley cell transport during the dust storm period appears to be related to incorrect modelling of the vertical water vapour profile, and suggests water vapour is confined below $\sim 30\,\text{Pa}$ during this time. This confinement may be because of sedimentation by clouds in the upwelling branch of the Hadley cell (which are not as optically thick in the model as in TES and MICS observations), or possibly because increased regolith adsortion during the dust storm period as shown by Böttger et al. (2004). Cloud formation and regolith adsorption could also explain the delayed peak in the south polar water vapour abundance, which occurs around $L_\odot = 270^\circ - 300^\circ$ in observations but around $L_\odot = 240^\circ - 270^\circ$ in the model. The bias over the north polar cap during northern summer shows the importance of the outlying frost deposits in creating the peak water vapour column abundances. When initialized with a realistic water vapour field from the water vapour assimilation, the addition of surface ice deposits between $120^\circ - 210^\circ$E and $75^\circ - 80^\circ$N is required to produce peak water vapour abundances similar to observations.

Acknowledgements

The authors thank Luca Montabone for his help with the assimilation scheme and David Kass for his suggestions which helped improve this paper. This work was funded by the UK Science and Technology Facilities Council.

References


Figure 5: Zonally- and temporally-averaged temperatures for three periods around northern summer (upper panels) autumn (middle panels) and winter (lower panels). Results are shown from a free-run (FR; left panels) an assimilation of TES temperature profiles (TES-TA; middle panels) and an assimilation of MCS temperature profiles (MCS-TA; right panels). White shading represents topography, and the relevant Mars years are labelled at the lower left of each plot.

Figure 6: Zonally- and temporally-averaged temperature differences between the FR and TES-TA around northern hemisphere (a) summer, (b) autumn and (c) winter. Contours are shaded every 2 K from ±1 K, with the (solid, dashed) contours showing where the TES-TA has (increased, decreased) the temperature compared to the FR. Black shading represents topography.
Figure 1: Example of the scaling process, with the water vapour mixing ratio profile scaled each iteration by a factor $\alpha$ (as in Eq. 7) until the column-integrated value agrees with the assimilation. For $\alpha > 1$ a cutoff is applied at 1 Pa, above which no water vapour increase can occur, and points on the profile are not scaled above saturation (dashed line).

Figure 2: Evolution of $\sigma_g$ with time (see Eq. 9) from ensembles of assimilations over northern summer ($L_S = 122^\circ–127^\circ$, MY 25; (b) northern autumn, $L_S = 229^\circ–235^\circ$, MY 24; and (c) northern winter, $L_S = 323^\circ–329^\circ$, MY 24. Contours of surface height are plotted to highlight the areas of largest water vapour column variance, which are the Arcadia, Acidalia and Utopia Planitias, centred at approximately 160°W, 30°W and 120°E.

Figure 3: Water vapour column variances averaged over the last 10 sols from ensembles of assimilations. Periods are; (a) northern summer, $L_S = 122^\circ–127^\circ$, MY 25; (b) northern autumn, $L_S = 229^\circ–235^\circ$, MY 24; and (c) northern winter, $L_S = 323^\circ–329^\circ$, MY 24. Contours of surface height are plotted to highlight the areas of largest water vapour column variance, which are the Arcadia, Acidalia and Utopia Planitias, centred at approximately 160°W, 30°W and 120°E.
Figure 4: Water vapour column global RMS errors over two 90-sol periods covering (a) $L_S = 174^\circ$–$228^\circ$ (southern spring), and (b) $L_S = 268^\circ$–$323^\circ$ (southern summer). RMS errors are shown for a free-running model and a simulation in which the first 30-sols consisted of a temperature and water vapour assimilation. Both using the same initial conditions. The vertical dashed line represents the point assimilation ceased.

Figure 14: The cross-hemispheric transport of water vapour and ice over one Mars year, with (positive, negative) values representing transport (north, south). Panel (a) shows the total yearly mass transport in each 5° longitude bin while panel (b) shows the total equatorial mass transport each sol.

Figure 15: Zonally-averaged VA–TA water vapour mass mixing ratio difference (shading), with contours of mean meridional circulation (solid and dashed lines) and water ice mass mixing ratio (dotted lines) from the VA, as in Fig. 12. Zonal averages are over a 20-sol period centred on $L_S = 225^\circ$. 

18
Table 1: Overview of the simulations performed for the temperature assimilation validation.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>No. model levels</th>
<th>Data assimilated</th>
<th>Vertical extent/resolution of data</th>
<th>Mars year</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR (free run)</td>
<td>25</td>
<td>None</td>
<td>N/A</td>
<td>24/25</td>
</tr>
<tr>
<td>TES-TA</td>
<td>25</td>
<td>TES temperature profiles</td>
<td>~40 km / 10 km</td>
<td>24/25</td>
</tr>
<tr>
<td>MCS-TA</td>
<td>35</td>
<td>MCS temperature profiles</td>
<td>~85 km / 5 km</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2: Overview of the simulations performed for studying the global distribution and transport of water vapour.

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Initial conditions</th>
<th>Data assimilated</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR (free run)</td>
<td>Model spin-up</td>
<td>None</td>
</tr>
<tr>
<td>CR (control run)</td>
<td>From assimilation</td>
<td>None</td>
</tr>
<tr>
<td>TA (temperature assimilation)</td>
<td>From assimilation</td>
<td>TES temperature profiles</td>
</tr>
<tr>
<td>VA (vapour assimilation)</td>
<td>From assimilation</td>
<td>TES water vapour columns and temperature profiles</td>
</tr>
</tbody>
</table>

Figure 7: The number of water vapour column retrievals available for assimilation across each latitude circle, binned by 5° in latitude and 2° in LS. White areas signify times and locations where there are no data to assimilate.
Figure 8: Assimilation results over the range $L_S = 180^\circ$, MY 24 to $L_S = 180^\circ$, MY 25 showing (a) the zonally-averaged water vapour columns from the VA as a function of $L_S$ and latitude, (b) the VA–TA difference, and (c) the VA–CR difference. The difference plots have shaded contours every 2 pr-$\mu$m, with (solid, dashed) lines every 5 pr-$\mu$m showing the (positive, negative) differences. The dotted line represents the zero contour. Data are averaged over 2$^\circ$ of $L_S$. 
Figure 9: (a) The number of water vapour column observations assimilated globally each model time step, with the black line showing the daily average value. (b) Daily average water vapour column global RMS errors. (c) Daily average water vapour column global mean errors. The errors are shown for the CR, TA and VA. The data span the range $L_S = 180^\circ$, MY 24 to $L_S = 180^\circ$, MY 25, with the two main gaps centred around $L_S = 15^\circ$ and $160^\circ$ due to solar conjunction and the MGS spacecraft entering safing mode.
Figure 10: Comparison between (a) the absorption-only TES water ice optical depths at 825 cm$^{-1}$, binned by 5° in latitude and 2° in $L_S$, and (b) the zonally-averaged water ice optical depths resulting from the VA, sampled to 2 p.m. local time and averaged over 2° of $L_S$ (right panel). A mask has been applied to the VA results to make the comparison clearer.

Figure 11: Vertically-averaged meridional wind deviations from the zonal mean between the surface and 100 Pa (≈20 km above the 610 Pa pressure level). Data are averaged over periods spanning 20° of $L_S$, centred on the midpoints of northern hemisphere (a) autumn, (b) winter, (c) spring, and (d) summer.
Figure 12: Zonally-averaged water vapour mass mixing ratios from the VA (shaded), with contours of mean meridional circulation (solid and dashed black lines) and ice mass mixing ratio (white lines) overplotted. Each panel shows zonal averages over a 20-sol period centred on the labelled $L_S$ values. Panels (a)–(d) are for MY 24, and panels (e)–(h) for MY 25. The mean meridional circulation contours have pseudo-logarithmic values of $[0.1, 0.5, 1, 5] \times 10^9$ kg s$^{-1}$, with (solid, dashed) lines representing (clockwise, anticlockwise) circulation. The ice mass mixing ratio contours denote the values $[0.005, 0.01, 0.02, 0.03]$ g kg$^{-1}$.

Figure 13: Vertically-integrated meridional fluxes of water vapour from the VA, with (positive, negative) fluxes representing (northward, southward) transport. The fluxes have been decomposed into their mean-meridional, stationary wave and transient eddy components, and are averaged over 20 sols centred on the labelled $L_S$ values. Note the different vertical scale in panels (b) and (c).
Figure 16: Zonally-averaged water vapour columns over the period $L_S = 30^\circ - 180^\circ$ of MY 25 for (a) the VA, (b) the CR, and (c) the CR with the addition of outlying water ice deposits between $120^\circ - 210^\circ$E and $75^\circ - 80^\circ$N. All used the same initial conditions, obtained from the VA.