



HAL
open science

Assimilation of temperatures and column dust opacities measured by ExoMars TGO-ACS-TIRVIM during the MY34 Global Dust Storm

Roland M. B. Young, Ehouarn Millour, Sandrine Guerlet, François Forget, Nikolay Ignatiev, Alexey Grigoriev, Alexey Shakun, Alexander Trokhimovskiy, Franck Montmessin, Oleg Korablev

► To cite this version:

Roland M. B. Young, Ehouarn Millour, Sandrine Guerlet, François Forget, Nikolay Ignatiev, et al.. Assimilation of temperatures and column dust opacities measured by ExoMars TGO-ACS-TIRVIM during the MY34 Global Dust Storm. *Journal of Geophysical Research. Planets*, 2022, e2022JE007312 (in press). 10.1029/2022JE007312 . insu-03768361v1

HAL Id: insu-03768361

<https://insu.hal.science/insu-03768361v1>

Submitted on 3 Sep 2022 (v1), last revised 26 Oct 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Assimilation of temperatures and column dust opacities**
2 **measured by ExoMars TGO-ACS-TIRVIM during the**
3 **MY34 Global Dust Storm**

4 **Roland M. B. Young^{1,2}, Ehouarn Millour², Sandrine Guerlet^{2,3}, François**
5 **Forget², Nikolay Ignatiev⁴, Alexey V. Grigoriev⁴, Alexey V. Shakun⁴,**
6 **Alexander Trokhimovskiy⁴, Franck Montmessin⁵, Oleg Korablev⁴**

7 ¹Department of Physics & National Space Science and Technology Center, UAE University, Al Ain,
8 United Arab Emirates

9 ²Laboratoire de Météorologie Dynamique (LMD/IPSL), Sorbonne Université, ENS, PSL Research
10 University, École Polytechnique, Institut Polytechnique de Paris, CNRS, Paris, France

11 ³LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place
12 Jules Janssen, 92195 Meudon, France

13 ⁴Space Research Institute (IKI), 84/32 Profsoyuznaya, 117997 Moscow, Russia

14 ⁵LATMOS/IPSL, UVSQ Université Paris-Saclay, UPMC Univ. Paris 06, CNRS, Guyancourt, France

15 **Key Points:**

- 16 • We assimilate temperature and dust observations from ExoMars TGO-ACS-TIRVIM
17 during the MY34 global dust storm.
- 18 • The analysis verifies well against independent Mars Climate Sounder temperature
19 and dust profiles, and Curiosity pressure measurements.
- 20 • At the peak of the global dust storm the winds strengthen, the diurnal tide mi-
21 grates poleward, and the semi-diurnal tide strengthens.

Corresponding author: Roland M. B. Young, roland.young@uaeu.ac.ae

22 **Abstract**

23 We assimilate atmospheric temperature profiles and column dust optical depth obser-
 24 vations from the ExoMars Trace Gas Orbiter Atmospheric Chemistry Suite thermal in-
 25 frared channel (TIRVIM) into the LMD Mars Global Climate Model. The assimilation
 26 period is Mars Year 34 $L_s = 182.3 - 211.4^\circ$, covering the onset and peak of the 2018
 27 global dust storm. We assimilated observations using the Local Ensemble Transform Kalman
 28 Filter with 36 ensemble members and adaptive inflation; our nominal configuration as-
 29 similated TIRVIM temperature profiles to update temperature and dust profiles, followed
 30 by dust column optical depths to update the total column dust abundance. The obser-
 31 vation operator for temperature used the averaging kernels and prior profile from the TIRVIM
 32 retrievals.

33 We verified our analyses against in-sample TIRVIM observations and independent
 34 Mars Climate Sounder (MCS) temperature and dust density-scaled opacity profiles. When
 35 dust observations were assimilated, the root-mean-square temperature error verified against
 36 MCS fell by 50% during the onset period of the storm, compared with assimilating tem-
 37 perature alone. At the peak of the storm the analysis reproduced the location and mag-
 38 nitude of the peak in the nighttime MCS dust distribution, along with the surface pres-
 39 sure diurnal cycle measured by Curiosity with a bias of less than 10 Pa. The analysis
 40 winds showed that, at the peak of the storm, the meridional circulation strengthened,
 41 a 125 m s^{-1} asymmetry developed in the midlatitude zonal jets, the diurnal tide weak-
 42 ened near the equator and strengthened to 10–15 K at midlatitudes, and the semi-diurnal
 43 tide strengthened almost everywhere, particularly in the equatorial lower atmosphere.

44 **Plain Language Summary**

45 The ExoMars Trace Gas Orbiter (TGO) has been in a low orbit around Mars since
 46 early 2018. Halfway through 2018 (Mars Year 34) a large amount of dust was lifted from
 47 Mars’ surface into the lower atmosphere. This dust eventually obscured most of the planet,
 48 called a global dust storm. These storms have a large effect on the temperature and weather
 49 in Mars’ atmosphere.

50 In this study we have used observations from TGO’s thermal infrared instrument
 51 TIRVIM. This measures how much light comes from the Martian atmosphere and sur-
 52 face at wavelengths where heat is emitted; these measurements can be used to work out

53 the atmospheric temperature and how much dust it contains. We combined these ob-
 54 servations with a numerical model of Mars’ climate in a way that takes into account how
 55 uncertain we are about the atmospheric properties measured by TIRVIM and the cli-
 56 mate predicted by the model.

57 We compared our results with separate temperature and dust observations made
 58 by the Mars Climate Sounder instrument on board another Mars orbiting satellite, Mars
 59 Reconnaissance Orbiter. We also studied how the atmospheric wind, the day-night cy-
 60 cle, and the surface pressure all changed while the dust storm was raging.

61 **1 Introduction**

62 The ExoMars Trace Gas Orbiter (TGO) reached its final orbit around Mars in April
 63 2018. It carries the Atmospheric Chemistry Suite (ACS), which comprises three infrared
 64 spectrometers covering different wavelength ranges (Korablev et al., 2018). The thermal
 65 infrared spectrometer TIRVIM continuously monitors the atmospheric column in nadir,
 66 while the near- and mid-infrared channels NIR and MIR (and occasionally TIRVIM) per-
 67 form solar occultation measurements when the spacecraft-Mars-Sun geometry permits.
 68 TGO’s orbit made ACS-TIRVIM’s (TIRVIM hereafter) observations of Mars unique at
 69 the time among existing Mars spacecraft, because it systematically measures the atmo-
 70 spheric structure over all local times of day. Nearly all spacecraft that have orbited Mars
 71 since Mars Global Surveyor at the start of the modern era have taken observations at
 72 two fixed local times of day separated by 12 hours (Mars Express has an orbit that means
 73 it does take observations at varying local times of day (Giuranna et al., 2021), but not
 74 systematically, and since May 2021 the Emirates Mars Mission has observed Mars from
 75 high orbit with full local time coverage (Amiri et al., 2022)). TGO takes about 55 sols
 76 to cycle through the full 24 hours of sub-spacecraft local times, which is close to the op-
 77 timal repeat period to sample the diurnal cycle (Capderou & Forget, 2004).

78 TIRVIM took observations at Mars from March 2018, shortly before TGO reached
 79 its final orbit, until the end of 2019, when its cryocooler failed, rendering it impossible
 80 to cool the instrument to the required temperature to make thermal infrared measure-
 81 ments. At the start of this ~ 20 -month period, about halfway through Mars Year (MY)
 82 34, a Global Dust Storm (GDS) took place. This started around $L_s = 186^\circ$ in Acidalia,
 83 reached a peak in visible dust optical depth between $L_s = 200\text{--}205^\circ$, and faded away

84 by about $L_s = 300^\circ$ (Kass et al., 2019; Montabone et al., 2020; Kleinböhl et al., 2020).
 85 The storm was severe enough to end the solar-powered Opportunity rover’s 15-year mis-
 86 sion. The start of the TIRVIM dataset covers the onset and peak of this GDS; fortuitously,
 87 as these events happen in only about one Mars year out of three (Kahre et al., 2017).

88 Over the last 20–25 years, scientists have studied Mars using data assimilation, which
 89 combines observations with our best scientific understanding encoded within numerical
 90 simulations (Kalnay, 2003). A variety of methods have been used, with most work us-
 91 ing Analysis Correction (AC) (Lorenc et al., 1991; Lewis & Read, 1995; Lewis et al., 2007;
 92 Steele, Lewis, & Patel, 2014) or the Ensemble Kalman Filter (Evensen, 2003; M. J. Hoff-
 93 man et al., 2010; Greybush et al., 2012; Navarro et al., 2014). These have assimilated
 94 various observations, primarily atmospheric temperature profiles, but also dust column
 95 optical depths (Lewis & Barker, 2005; Montabone et al., 2014; Ruan et al., 2021), dust
 96 profiles (Navarro et al., 2017; Ruan et al., 2021), and column abundances of water ice
 97 (Steele, Lewis, & Patel, 2014), water vapour (Steele, Lewis, Patel, Montmessin, et al.,
 98 2014), ozone (Holmes et al., 2018), and carbon monoxide (Holmes et al., 2019). The pri-
 99 mary assimilated quantity is temperature, because it is readily available from thermal
 100 infrared measurements from multiple spacecraft, changes in other atmospheric quanti-
 101 ties such as pressure and winds are related to changes in temperature via known phys-
 102 ical laws, and while wind observations are a vital part of Earth data assimilation (Hersbach
 103 et al., 2019), they are not available for Mars.

104 Most relevant to the current work, Navarro et al. (2014, 2017) built a data assim-
 105 ilation scheme for Mars’ atmosphere using the Local Ensemble Transform Kalman Fil-
 106 ter (LETKF) (Hunt et al., 2007) and the Laboratoire de Météorologie Dynamique Mars
 107 Global Climate Model (LMD Mars GCM) (Forget et al., 1999; Pottier et al., 2017), as-
 108 simulating observations from the Mars Climate Sounder (MCS) instrument on board NASA’s
 109 Mars Reconnaissance Orbiter (MRO). MCS retrieval products include temperature, dust,
 110 and water ice profiles taken from a Sun-synchronous polar orbit at ~ 3 AM / 3 PM lo-
 111 cal mean solar time (Kleinböhl et al., 2009).

112 In this paper we report on assimilation of observations from the early part of the
 113 ExoMars TGO ACS-TIRVIM observation period, covering the onset and peak of the MY34
 114 GDS. We make two methodological improvements to the existing scheme documented
 115 by Navarro et al. (2017). First, we improve the way temperature forecasts are compared

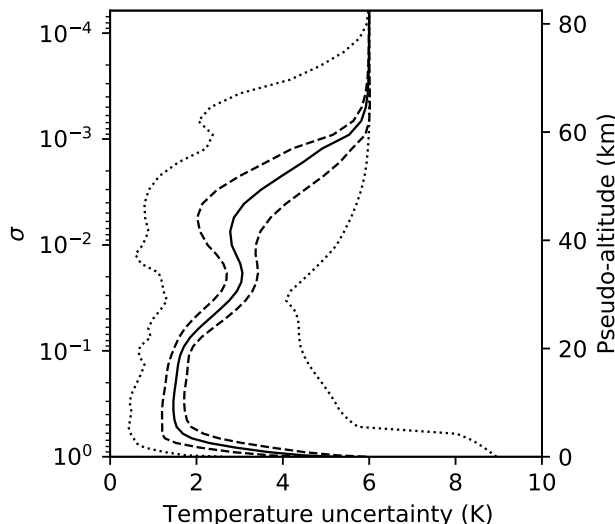
116 with observations by using the averaging kernel matrix and prior from the TIRVIM tem-
 117 perature retrievals. This is the correct way to compare forecast with observations when
 118 assimilating retrieval products (R. N. Hoffman, 2011). Second, the scheme can now as-
 119 simulate column dust optical depths (CDOD); previous work with this particular scheme
 120 did not assimilate CDOD because such data are not available from MCS (while dust pro-
 121 files are available, they do not reach the ground, and published MCS CDODs interpo-
 122 late the dust profiles to the ground in the lowest part of the atmosphere, typically the
 123 lowest 10 km). We describe both of these improvements in detail in Sect. 3.

124 The aims of this paper are as follows. First, to demonstrate assimilation of obser-
 125 vations from TIRVIM on board ExoMars TGO. Second, to validate the assimilation against
 126 independent datasets such as MCS temperatures and dust opacities, and surface pres-
 127 sure measurements from the Curiosity rover. Finally, to investigate the dynamics of Mars’
 128 atmosphere during the MY34 GDS, particularly those properties that cannot be mea-
 129 sured directly but which can be retrieved by the assimilation process, such as wind. Sec-
 130 tion 2 describes the TIRVIM observations, and in Section 3 we describe the model, as-
 131 simulation scheme, and observation operator. Section 4 describes sensitivity tests done
 132 to optimise the assimilation parameters. Section 5 describes the main results of the pa-
 133 per, and in Section 6 we conclude.

134 2 Observations

135 TIRVIM is a thermal infrared spectrometer that observes Mars at wavelengths be-
 136 tween 1.7–17 μm (590–5900 cm^{-1}) (Korablev et al., 2018). It was designed to be con-
 137 tinuously operating during TGO’s mission (although in practice this was not possible
 138 due to the limited lifetime of the instrument’s cryocooler). TIRVIM measures radiance
 139 spectra in nadir, from which have been retrieved atmospheric temperature profiles, sur-
 140 face temperatures, dust column optical depths (at 1090 cm^{-1}) and water ice column op-
 141 tical depths (at 820 cm^{-1}). These retrievals are fully described by Guerlet et al. (2022).
 142 In this paper we focus on atmospheric temperature and column dust optical depth ob-
 143 servations only. TIRVIM operated from MY34 $L_s = 142.79^\circ$ (13 March 2018) until MY35
 144 $L_s = 115.16^\circ$ (2 December 2019), when its cryocooler failed.

145 TIRVIM takes observations at all local times over a 55-sol period. TGO’s inclina-
 146 tion is 74° , so there are no nadir observations poleward of this latitude in either hemi-



156

157 **Figure 1.** Temperature uncertainty in the TIRVIM dataset between MY34 $L_s = 182.288 -$
 158 211.388° as a function of $\sigma = p/p_{\text{surf}}$. The right y -axis is pseudo-altitude $z = -H \ln \sigma$, assuming
 159 a scale height of $H = 8.5$ km, as per the Mars standard atmosphere from the surface to 85 km
 160 altitude in Table 4.7 of M. D. Smith et al. (2017) (an underestimate in the lower atmosphere and
 161 an overestimate in the middle atmosphere, but it is sufficiently accurate for plotting). We use
 162 $H = 8.5$ km for all pseudo-altitudes in this paper. The solid line is the mean, dashed lines are
 163 mean ± 1 standard deviation, and the dotted lines show the minimum and maximum.

147 sphere. The temperature profiles have a maximum sensitivity between 5 and 50 km above
 148 the ground, and a vertical resolution of about 10 km (Guerlet et al., 2022). The instru-
 149 ment is only sensitive to the true atmospheric temperatures at altitudes of 3–55 km, based
 150 on retrieval averaging kernel statistics (see Section 3.3). From the trace of the averag-
 151 ing kernel matrix, the number of degrees of freedom in each temperature retrieval is 3.3
 152 ± 0.2 . The measurement uncertainty, shown in Fig. 1, is a function of altitude. It is typ-
 153 ically 2–4 K over most of the altitude range, increasing at the top and the bottom of the
 154 profile. Column dust optical depths typically have an uncertainty of 0.06–0.20 (the 25–
 155 75% range over the assimilated observations, with mean 0.16).

164 In this paper we assimilate observations from just before the onset of the MY34
 165 global dust storm to just after its peak, from $L_s = 182.288 - 211.388^\circ$. Figure 2 shows
 166 the data availability and distribution in latitude and local time for this period. These
 167 specific start and end points were chosen because there are long gaps in the data imme-

168 diately before and afterwards. There are 312,741 temperature profiles and 214,806 col-
 169 umn dust retrievals during this period. There is a gap in the data between $L_s = 191.567$ –
 170 194.420°), with significantly higher observation density afterwards. The local time of day
 171 for the majority of observations moves backwards over time; almost all 24 hours of lo-
 172 cal time are sampled during this period. Exceptions are the northern hemisphere around
 173 4 AM, and the southern hemisphere around 4 PM. A significant fraction of TIRVIM ob-
 174 servations have missing values for CDOD, as it cannot be retrieved in conditions when
 175 the temperature contrast between surface and lower atmosphere is low, which typically
 176 occurs near dawn and dusk in non-stormy conditions (see Guerlet et al. (2022)). Most
 177 of the missing dust retrievals are poleward of 60° , where almost all dust retrievals are
 178 removed.

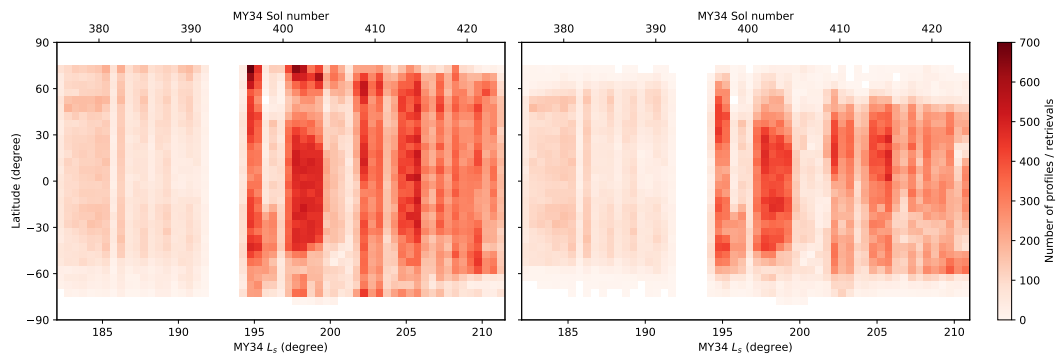
187 **3 Methods**

188 **3.1 Model**

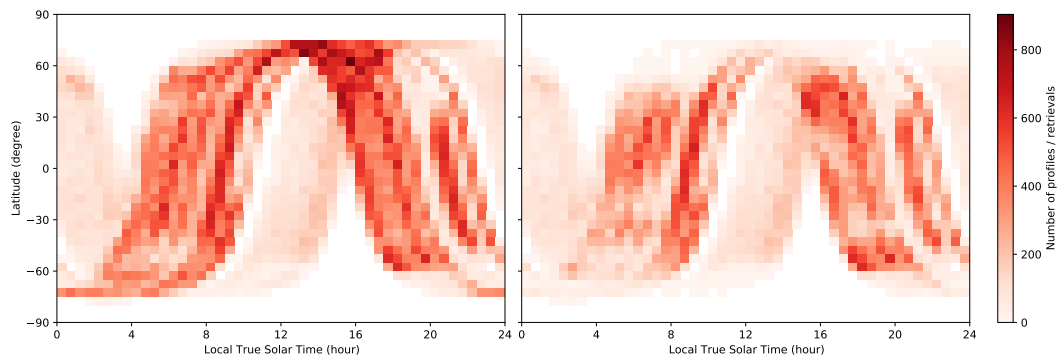
189 The LMD Mars GCM solves the hydrostatic primitive equations on a rotating sphere
 190 alongside parametrizations of many physical processes relevant to Mars (Forget et al.,
 191 1999). Most pertinent to the current work is the treatment of dust, which is transported
 192 by the model in a two-moment scheme which separately transports a mass mixing ra-
 193 tio and the number of dust particles within each grid cell (Madeleine et al., 2011). By
 194 transporting these two quantities separately one can deduce the dust particle size dis-
 195 tribution by assuming a log-normal particle size distribution. The LMD Mars GCM has
 196 been validated against many observational data sets since the Viking landers in the 1970s
 197 (Forget et al., 1999).

198 We run the model at the standard climate model resolution of 64×48 points in
 199 longitude and latitude, corresponding to $5.625^\circ \times 3.75^\circ$ horizontal resolution. The ver-
 200 tical grid is stretched so there are more points near the ground; we use 32 hybrid (sigma-
 201 pressure) levels with the lowest level about 4 m above the ground, and the highest level
 202 at about 100 km altitude, depending on local conditions. The model’s dynamical time
 203 step δt is 1/960 of a sol (~ 94.5 s), and the physical parametrisations are called every 10
 204 dynamical timesteps (i.e. every physics timestep, 1/96 of a sol, or four times per Mar-
 205 tian hour). Supporting Text S1 describes the time axis used in the GCM and how it re-
 206 lates to observation times.

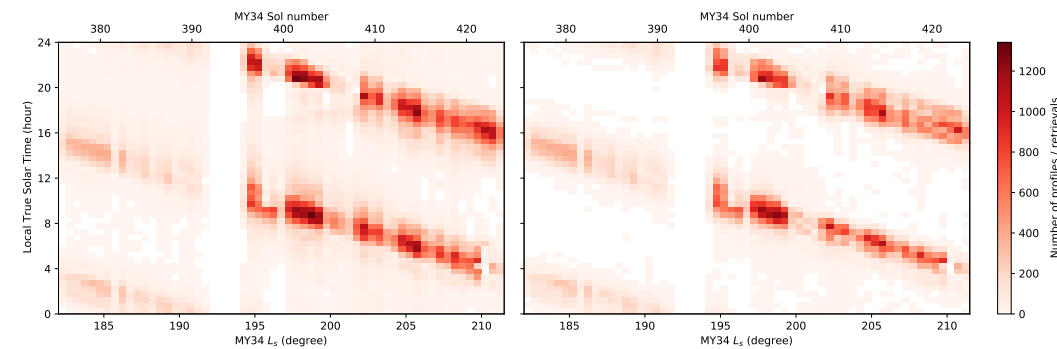
(a) vs MY34 L_s and latitude.



(b) vs local true solar time and latitude.



(c) vs MY34 L_s and local true solar time.



179

180

181 **Figure 2.** Number of TIRVIM temperature profiles (left) and column dust optical depth re-
 182 trievals (right) as a function of MY34 L_s , latitude, and local true solar time during the period
 183 assimilated in this work. (a) vs MY34 L_s and latitude; (b) vs local true solar time and latitude;
 184 (c) vs MY34 L_s and local true solar time. Observations are counted in bins with width 5.0° in
 185 latitude, 0.5 hours in local time, and 0.5° in L_s ; bins with no observations are white. The upper
 186 x -axis shows the equivalent model sol numbers; see Supporting Text S1 for details.

207 Dust sediments and can fall out of the bottom of the model onto the surface. When
 208 run without assimilation, dust column abundances in the GCM are forced by rescaling
 209 at each physics timestep to a global “dust scenario” map based on observed dust opac-
 210 ities, with gaps filled in by kriging (Montabone et al., 2015, 2020). The GCM also lifts
 211 dust from the surface at a constant rate, to preserve the dust particle size distribution
 212 by seeding the atmosphere with larger particles, which preferentially sediment out. Wa-
 213 ter is transported by the model (Pottier et al., 2017), and can condense into water ice
 214 clouds, whose particles have an effective radius that varies with local conditions, and which
 215 are treated as radiatively active. Dust acts as a source of cloud condensation nuclei, and
 216 precipitating water ice can remove dust from the atmosphere by scavenging.

217 **3.2 Assimilation Scheme**

218 Assimilation is based on the LETKF, which is a standard method for ensemble data
 219 assimilation (Evensen, 2003; Hunt et al., 2007). Full details of the LETKF itself as ap-
 220 plied in this context are described in Navarro et al. (2014) and Navarro et al. (2017); here
 221 we give a summary and focus on the changes to the scheme since Navarro et al. (2017).

222 Before starting the assimilation we spin up an ensemble of N GCM simulations.
 223 This ensemble is used as the forecast for the first assimilation step. Its ensemble mean
 224 $\bar{\mathbf{x}}^b$ is used as the *background* atmospheric state in the assimilation, and the purpose of
 225 the ensemble is to estimate the uncertainty in the background state. We use the spread
 226 of values to set the background error covariance matrix \mathbf{B} in the assimilation. For the
 227 spin up stage in this case we ran an ensemble of GCM simulations starting at MY34 sol
 228 318 ($L_s = 150.2^\circ$), i.e. about 30 L_s before the start of the assimilation period. This
 229 gave the ensemble members enough time to spin up to an equilibrium state individually,
 230 and to spread out over the model state space such that the standard deviation of the en-
 231 semble at each point quantified the uncertainty in the background state. As dust is the
 232 major unknown in Mars’ atmosphere, and the atmospheric state is so sensitive to it, to
 233 span the range of possibilities we ran each spin-up simulation with a different globally-
 234 fixed visible CDOD, log-uniformly distributed between 0.05 and 1.5. This spans the typ-
 235 ical range of realistic dust column optical depths in non-GDS years (Montabone et al.,
 236 2015).

237 The assimilation step computes an *analysis* ensemble $\mathbf{x}^{a,i}$, $i = 1 \dots N$, at the anal-
 238 ysis time t_a . Most of the results are presented using the mean of this ensemble, which
 239 we shall simply refer to as the *analysis*. The assimilation uses the forecast ensemble as
 240 the background state $\mathbf{x}^{b,i}$, and observations \mathbf{y}_o within an assimilation window stretch-
 241 ing from $t_a - t_b$ to $t_a + t_f$, where t_b is the window length backwards in time and t_f is
 242 the window length forward in time. To ensure each observation is used in exactly one
 243 assimilation step, $t_b = t_f = \Delta t/2$, where Δt is the assimilation cycle length, the time
 244 between consecutive assimilation steps. Adaptive covariance inflation is applied to the
 245 background ensemble. The observation operator $H(\dots)$ is applied to the background en-
 246 semble, to interpolate the forecasts in space and time to the observation points, and to
 247 ensure that the forecast and observations are subsequently compared as if TIRVIM were
 248 observing the background state; Sect. 3.3 provides full details.

249 To perform the assimilation, at each model grid point we identify observations within
 250 $L_{eh} = 900$ km horizontally (along a great circle), and within $L_{ev} = 0.2\sqrt{6}$ vertically
 251 in $\log(p)$, a process called localisation (Hamill et al., 2001), which significantly speeds
 252 up the calculation and also reduces the effects of random noise in the observations act-
 253 ing over large distances. The measurement uncertainty for each observation is modified
 254 according to its horizontal (d_h) and vertical distance (d_v) from the grid point, where d_v
 255 is measured in $\log p$. For observation l , this modified observational uncertainty is

$$\sigma_l^o = \sigma^o \exp\left(\frac{d_h^2}{2L_{ih}^2}\right) \exp\left(\frac{d_v^2}{2L_{iv}^2}\right) \quad (1)$$

256 where σ^o is the unmodified observational uncertainty, $L_{ih} = 600$ km is the internal hor-
 257 izontal localisation length, and $L_{iv} = 0.2$ is the internal vertical localisation length.

258 The Local Ensemble Transform Kalman Filter equation for the analysis mean $\bar{\mathbf{x}}^a$
 259 is (Hunt et al., 2007; Navarro et al., 2017)

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b \tilde{\mathbf{P}}^a (\mathbf{Y}^b)^\top \mathbf{R}^{-1} (\mathbf{y}^o - H(\bar{\mathbf{x}}^b)) \quad (2)$$

260 where $\bar{\mathbf{x}}^b$ is the background mean, \mathbf{X}^b is the matrix whose columns are individual fore-
 261 cast ensemble members minus the ensemble mean, in model space, \mathbf{Y}^b is the matrix whose
 262 columns are individual forecast ensemble members mapped to the observation points mi-

263 nus the ensemble mean, in observation space, \mathbf{R} is the matrix of observation uncertain-
 264 ties (diagonal) modified by the localisation weights above, $\mathbf{y}^o - H(\bar{\mathbf{x}}^b)$ is the difference
 265 between the observations, in observation space, called the “innovation”, and

$$\tilde{\mathbf{P}}^a = [(N - 1)\mathbf{I} + (\mathbf{Y}^b)^\top \mathbf{R}^{-1} \mathbf{Y}^b]^{-1} \quad (3)$$

266 is a normalization factor. The terms before the innovation are essentially the ratio of the
 267 background error covariance to the sum of the background and observational error co-
 268 variances, but for vectors of background forecasts and observations.

269 Equation 2 can be written as

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w}^a \quad (4)$$

270 where \mathbf{w}^a is a matrix of weights assigned to each ensemble member based on the differ-
 271 ences between the forecast and observations, the observational error covariance, and the
 272 background error covariance. We are not restricted to using \mathbf{w}^a to update solely the quan-
 273 tity that is observed; \mathbf{w}^a can be used to update any prognostic variable. This is only phys-
 274 ically sensible where there are correlations between changes in the observed quantity and
 275 changes in the other prognostic variable. This can be justified by noting that some quan-
 276 tities are related by known physical laws; for example the atmospheric temperature and
 277 wind structure are related, as are the atmospheric temperature and surface pressure. This
 278 assimilation of such *indirectly observed* variables is discussed by Navarro et al. (2017) in
 279 detail.

280 In this work we use atmospheric temperature observations to update the GCM at-
 281 mospheric temperature field (temperature updates temperature, TuT). We then use \mathbf{w}^a
 282 computed for temperature to update the zonal and meridional velocities (TuW) and the
 283 surface pressure (TuPs, using \mathbf{w}^a at the lowest model grid level).

284 In most configurations (see below) we also use temperature to update the dust mass
 285 mixing ratio (TuD). Navarro et al. (2014) showed that Mars’ observed detached dust lay-
 286 ers (Heavens et al., 2011) are reproduced by the assimilation when TuD is employed, as
 287 dust modifies the temperature field. TuD is applied only where the correspondence be-
 288 tween changes in dust and changes in temperature are strong, i.e. where the shortwave

289 heating rate is above 0.2 K hr^{-1} and the total insolation is above 100 W m^{-2} (this is
 290 essentially a condition on time of day and latitude as a function of time of year). It is
 291 also constrained to occur only above pseudo-altitude $\sim 11 \text{ km}$, because the correspon-
 292 dence between dust and heating is based on MCS observations, which have systematic
 293 biases close to the ground (Navarro et al., 2014). Whenever the dust mass mixing ra-
 294 tio (MMR) is updated by the assimilation the dust number is also updated, so that any
 295 new dust has a particle size distribution with fixed dust effective radius; this radius is
 296 a free parameter of the assimilation scheme.

297 In the forecast step we step forward the GCM in time, starting from the analysis
 298 ensemble, from analysis time t_a up to $t_a + \Delta t + t_f$, which is the next analysis time plus
 299 the window length forward in time. The cycle then begins anew at the new analysis time
 300 $t_a + \Delta t$. During the forecast step the GCM runs freely with no constraints on the dust
 301 field. In the assimilation experiments the only processes that can add or remove dust
 302 from the atmosphere are the assimilation step and sedimentation onto the surface.

303 **3.3 Observation Operator Using Averaging Kernels**

304 The comparison between forecast and observations $\mathbf{y}^o - H(\mathbf{x}^b)$ is a crucial part
 305 of the data assimilation process. This comparison must compare like-with-like to min-
 306 imise representation errors and ensure that forecast errors are calculated accurately. The
 307 observation operator must not only interpolate the forecast to the observation locations,
 308 but also reproduce what TIRVIM would observe if it were to observe the forecast atmo-
 309 sphere. When retrievals are used, the correct way to do this is to use the same averag-
 310 ing kernels and prior profile that were used in the retrieval (Rodgers & Connor, 2003;
 311 R. N. Hoffman, 2011).

312 We have improved on earlier versions of this assimilation scheme by basing the ob-
 313 servation operator for atmospheric temperature on the TIRVIM retrieval averaging ker-
 314 nels. Earlier work using this scheme by Navarro et al. (2017) had no vertical smooth-
 315 ing in the observation operator. This was (reasonably well) justified as they assimilated
 316 limb observations from MCS, which have a vertical resolution of about 5 km (Kleinböhl
 317 et al., 2009), similar to the model’s vertical resolution throughout most of the atmosphere.
 318 Nadir measurements have a coarser vertical resolution, so some smoothing of the fore-

319 cast profile is required. We believe this is the first time averaging kernels have been ap-
 320 plied directly in Mars data assimilation.

321 Before the averaging kernels are applied, we interpolate from the forecast grid to
 322 each observation location linearly in longitude, latitude, and time, and then vertically
 323 to the observed σ coordinates, linearly in $\log \sigma$.

324 The averaging kernels quantify the relative contributions to each point in a retrieved
 325 profile from the true and prior atmospheric states at each altitude. The vertical width
 326 of the averaging kernel sets the true vertical resolution of the instrument. The retrieval
 327 \mathbf{y}^o is related to the true atmospheric state $\hat{\mathbf{x}}$ and the prior profile \mathbf{x}^p by the averaging
 328 kernels \mathbf{A} (Rodgers & Connor, 2003) such that

$$\mathbf{y}^o = \mathbf{x}^p + \mathbf{A}(\hat{\mathbf{x}} - \mathbf{x}^p) \quad (5)$$

329 A perfect retrieval has $\mathbf{A} = \mathbf{I}$, in which case $\mathbf{y}^o = \hat{\mathbf{x}}$, with no prior information retained.
 330 A retrieval that contains no information from the true atmospheric state has $\mathbf{A} = \mathbf{0}$,
 331 in which case $\mathbf{y}^o = \mathbf{x}^p$. Real instruments are somewhere in between, and \mathbf{A} is a strong
 332 function of altitude.

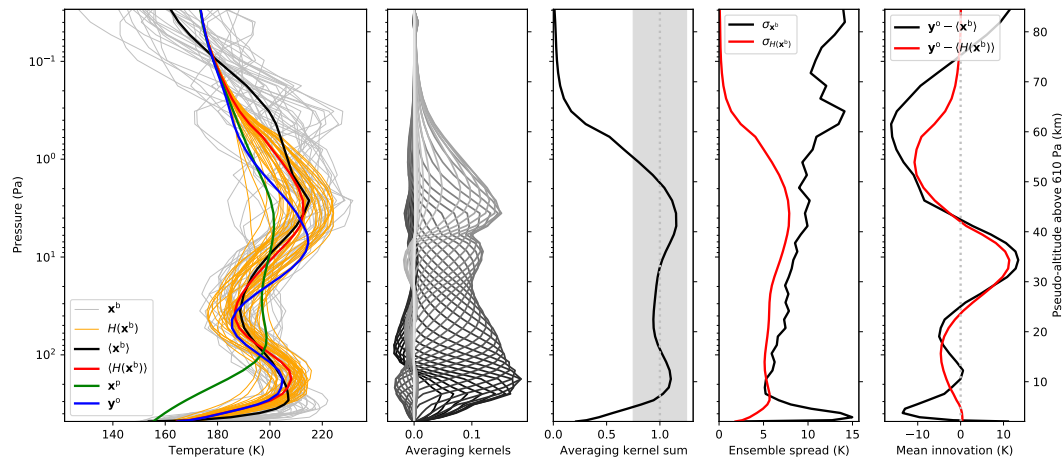
333 To assimilate a retrieval, the observation operator $H(\mathbf{x}^b)$ is a slightly modified form
 334 of Eq. 5:

$$H(\mathbf{x}^b) = \mathbf{x}^p + \mathbf{A}(\mathbf{x}^b - \mathbf{x}^p) \quad (6)$$

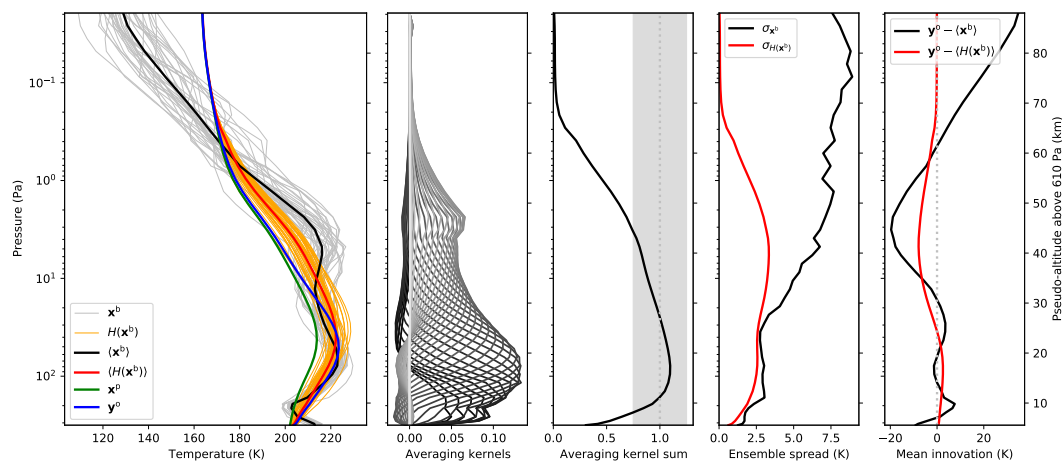
335 where \mathbf{x}^b is the background profile interpolated to the retrieved profile levels. The re-
 336 sult is a “retrieved forecast”, i.e. what a retrieval of a hypothetical TIRVIM spectrum
 337 would look like if it observed the forecast atmosphere.

354 Figure 3 shows two examples of this observation operator applied to a real fore-
 355 cast of TIRVIM retrievals. Figure 3a is a “good” case near the top of the range of re-
 356 trieval degrees of freedom (3.97), and Fig. 3b is a “bad” case at the low end of the range
 357 (2.62). As TIRVIM is a nadir sounder, the averaging kernels (second column) are quite
 358 broad, and the “retrieved forecast” ensemble and ensemble mean are smoother than the
 359 raw forecast ensemble and ensemble mean.

(a) $L_s = 208.86^\circ$, longitude 34.53°E , latitude 64.94°S , LTST 2.24.



(b) $L_s = 208.84^\circ$, longitude 108.94°W , latitude 40.82°N , LTST 15.87.



338

339 **Figure 3.** Demonstration of the observation operator acting on the ensemble forecast for
 340 two temperature profiles retrieved from TIRVIM radiance spectra. (a) $L_s = 208.86^\circ$, longi-
 341 tude 34.53°E , latitude 64.94°S , local true solar time 2.24, retrieval degrees of freedom
 342 $L_s = 208.84^\circ$, longitude 108.94°W , latitude 40.82°N , local true solar time 15.87, retrieval degrees
 343 of freedom 2.62. The left panel shows the raw forecast ensemble members and ensemble mean
 344 (thin grey lines and thick black line, respectively), forecast ensemble members and ensemble
 345 mean after the observation operator is applied (thin orange lines and thick red line, respectively),
 346 retrieval prior (thick green line), and the retrieved temperature profile, i.e. the “observations”
 347 (thick blue line). From left to right, the other panels show: (1) full set of averaging kernel func-
 348 tions, indicated by different greyscale shades; (2) sum over the averaging kernels, where the
 349 shaded region indicates the range 0.75–1.25, and the dotted line indicates 1; (3) ensemble spread
 350 for the raw forecast ensemble (black) and the forecast ensemble after the observation operator
 351 is applied (red); and (4) the mean innovation, i.e. the retrieved temperature profile minus the
 352 ensemble mean, showing this for the raw ensemble mean (black) and the ensemble mean after the
 353 observation operator is applied (red).

360 The retrieved temperature profile and the retrieval prior converge at the top of the
 361 profiles, indicating that almost all the information in the retrieval comes from the prior
 362 there, and so we should avoid assimilating this part of the profile. After the observation
 363 operator is applied to the forecast ensemble (Eq. 6), at the top of profile the ensemble
 364 collapses onto the retrieval prior, so differences between the forecast and retrieved tem-
 365 perature profiles are artificially small in this region (note also the change in ensemble
 366 spread in the 4th column of Fig. 3 before and after Eq. 6 is applied). Conversely, lower
 367 in the profiles, most or all of the spread in the forecast ensemble is preserved by the ob-
 368 servation operator, hence in this region the difference between the forecast temperature
 369 profile and the retrieved temperature profile is real, and so we should include such ob-
 370 servations in the assimilation.

371 To distinguish between these two cases, and quantify how much information comes
 372 from the prior and the true atmospheric state, we use the sum over the averaging ker-
 373 nels at each level (Fig. 3, 3rd column). Where the averaging kernel sum is near zero, all
 374 the information in the retrieval comes from the prior. Where the averaging kernel sum
 375 is close to one (e.g. at 10 Pa), almost all the information in the retrieval comes from the
 376 observed radiances. Therefore a suitable way to distinguish between these two cases is
 377 to impose a filter on the averaging kernel sum at each level of the temperature profile,
 378 and only assimilate observations whose averaging kernel sum falls within a particular range.
 379 This ensures that observations which are assimilated contain as much information from
 380 the observed radiances as possible, and as little information from the retrieval prior. The
 381 grey band in the 3rd column of Fig. 3 identifies the levels that would be retained if a fil-
 382 ter of 0.75–1.25 were applied. The main practical difference between the two cases in Figs 3a
 383 and 3b is that the vertical range of retained observations extends higher in the atmo-
 384 sphere when the retrieval degrees of freedom are higher.

393 Figure 4 shows averaging kernel sum statistics for TIRVIM temperature retrievals
 394 between MY34 sols 376–424. Between 5–40 km a large majority of the profiles have av-
 395 eraging kernel sums close to one, and even between 1–50 km in many cases. Figure 4c
 396 shows how much of the dataset remains when various filters are applied. Between 10–
 397 30 km almost all observations are kept, whatever the condition. The largest differences
 398 are between 40–60 km, where imposing strict limits of 0.85–1.15 removes at least 50%
 399 of the observations, and also in a narrow band of altitudes around 5–10 km. If the fil-

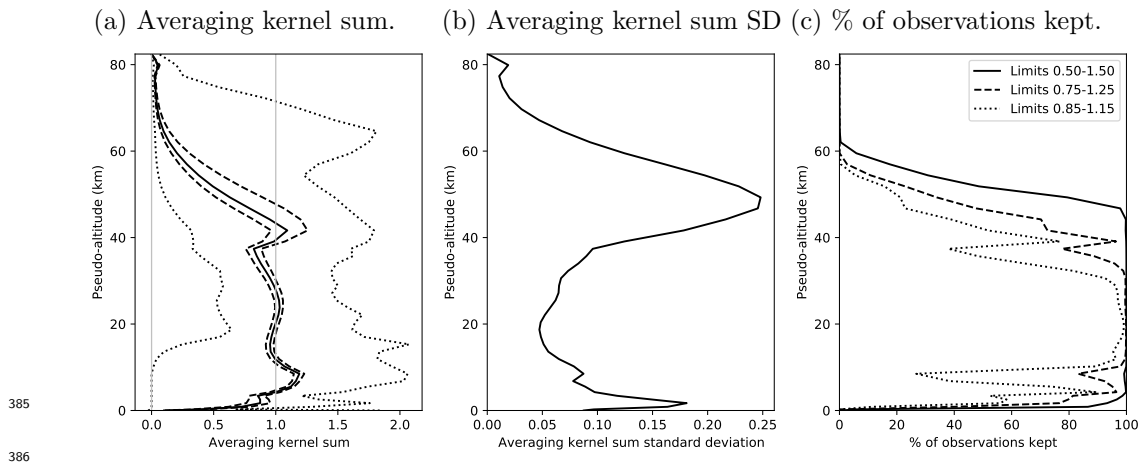


Figure 4. Averaging kernel sum statistics for TIRVIM temperature retrievals between MY34 $L_s = 182.288 - 211.388^\circ$, as a function of pseudo-altitude. (a) Sum over averaging kernels. The black dotted lines are the minimum and maximum, the solid line is the median, the dashed lines are the 25th and 75th percentiles, and the vertical grey lines indicate 0 (all weight on the prior) and 1 (all weight on the measured radiances). (b) Standard deviation of averaging kernel sums. (c) Percentage of observations kept when an averaging kernel sum filter is applied.

ter is relaxed slightly to 0.75–1.25, then most of the observations that were rejected for the strict limits are included.

3.4 Column Dust Optical Depth Assimilation

The second major update to our scheme is to assimilate column dust optical depths (denoted CuD, i.e column-updates-dust). Column dust optical depths have been assimilated by others previously (Lewis & Barker, 2005; Montabone et al., 2014; Ruan et al., 2021), but were not used by Navarro et al. (2014, 2017) in previous work with this scheme. They used MCS observations, and MCS is a limb scanner that does not measure CDOD directly as it rarely sounds the dust profile all the way to the ground. TIRVIM dust retrievals do have to make some assumptions about the vertical dust distribution, which is a limitation of TIRVIM data, but as a nadir instrument it does view the full atmospheric column.

Guerlet et al. (2022) assume dust is well-mixed in the lowest two scale heights and that its mixing ratio decreases linearly with $\log(p)$ above. A quality flag is assigned to each CDOD observation, related to the sensitivity of the radiance to changes in dust load.

415 Above warm surfaces, dust retrievals are quite reliable, but above colder surfaces, there
 416 can be systematic biases and non-unique solutions depending on the assumed dust ver-
 417 tical distribution. Only retrievals with a good quality flag are assimilated; this excludes
 418 about a third of the CDOD retrievals (Fig. 2).

419 We forecast the column dust optical depth at the TIRVIM frequency of 1090 cm^{-1}
 420 using the forecast dust mass mixing ratio profile. For each available column dust obser-
 421 vation we interpolate the forecast dust MMR field to the observation longitude, latitude,
 422 and time. We then integrate the total dust column opacity in the vertical based on Eq. 1
 423 of Madeleine et al. (2011):

$$\tau_{\text{dust}} = \sum_{k=1}^K \frac{3}{4} \frac{Q_{\text{ref}} q_k}{\rho_p r_{\text{eff}} g} \Delta p_k \quad (7)$$

424 where τ_{dust} is the column dust optical depth; K is the total number of model levels; $Q_{\text{ref}} =$
 425 1.53112 is the effective dust extinction coefficient at the TIRVIM wavelength, that ac-
 426 counts for integration over a dust particle size distribution with dust effective radius $r_{\text{eff}} =$
 427 $1.5 \mu\text{m}$ and effective variance $\nu_{\text{eff}} = 0.3$; q_k is the dust mass mixing ratio at level k ; $\rho_p =$
 428 2500 kg m^{-3} is the dust density; and Δp_k is the pressure thickness of model level k . We
 429 made some of the same assumptions as in the retrieval, such as using the same fixed Q_{ref}
 430 and r_{eff} , but in the GCM r_{eff} is a function of location, and the GCM does not assume
 431 any particular vertical dust distribution, so these restrictions could be relaxed in future.

432 Once CDOD is forecast at each observation location, observations are assimilated
 433 into the GCM CDOD field using the LETKF in the same way as for atmospheric tem-
 434 perature (Eq. 2), except vertical weighting is not required. As CDOD is not a prognos-
 435 tic variable in the LMD Mars GCM, once we have the CDOD analysis we adjust the dust
 436 MMR profile so that its column dust optical depth is the same as the CDOD analysis.
 437 As CDOD varies linearly with dust MMR this is a simple linear rescaling of the atmo-
 438 spheric dust profile. The dust number is also updated, as described at the end of Sect. 3.2.

439 We ran two versions of the dust assimilation. In the first case (CuD) we just used
 440 the column dust optical depth observation to update the dust column abundance. In this
 441 case we might expect the vertical dust distribution not to reflect the real vertical dust
 442 distribution, as the model does not generate detached dust layers spontaneously. We try
 443 to avoid this problem with our second version (TuD-CuD), which combines the direct

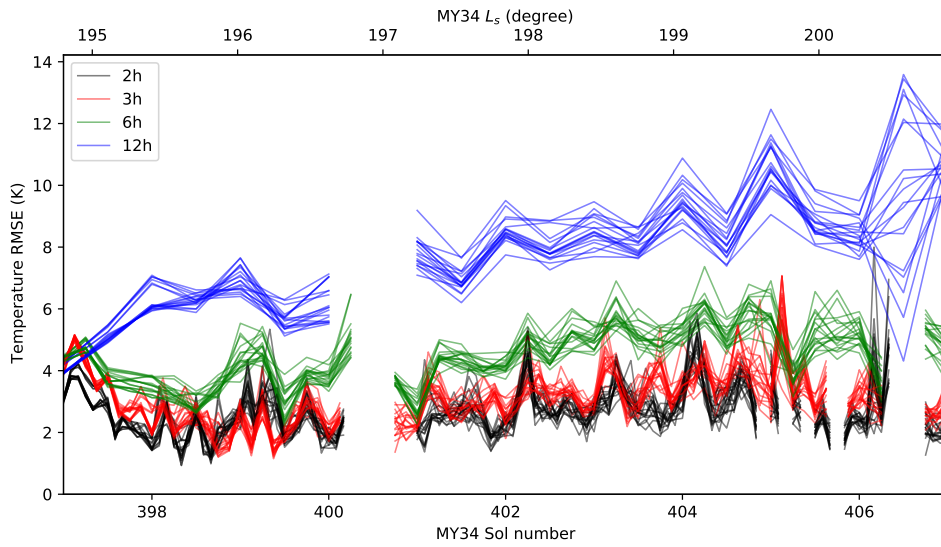
444 assimilation of CDOD with an indirect update of the dust profile using temperature (TuD).
 445 First we perform the TuD step, and then assimilate CDOD. The TuD step should pro-
 446 duce a more accurate vertical dust distribution than the model alone, and the column
 447 dust assimilation constrains the total amount of dust in each column using real obser-
 448 vations. The order of operations is important here: TuD must be done first and CuD
 449 second, because otherwise the total amount of dust in the column will not be the amount
 450 that is assimilated from observations; if we assimilate CDOD first there is no guaran-
 451 tee that the TuD step will conserve the total amount of dust in the column. This is sim-
 452 ilar to Ruan et al. (2021), who assimilated MCS dust profiles alongside TES and THEMIS
 453 dust column abundances, but note we do not assimilate the dust vertical distribution di-
 454 rectly, but only infer it from changes in temperature.

455 **4 Sensitivity Analysis**

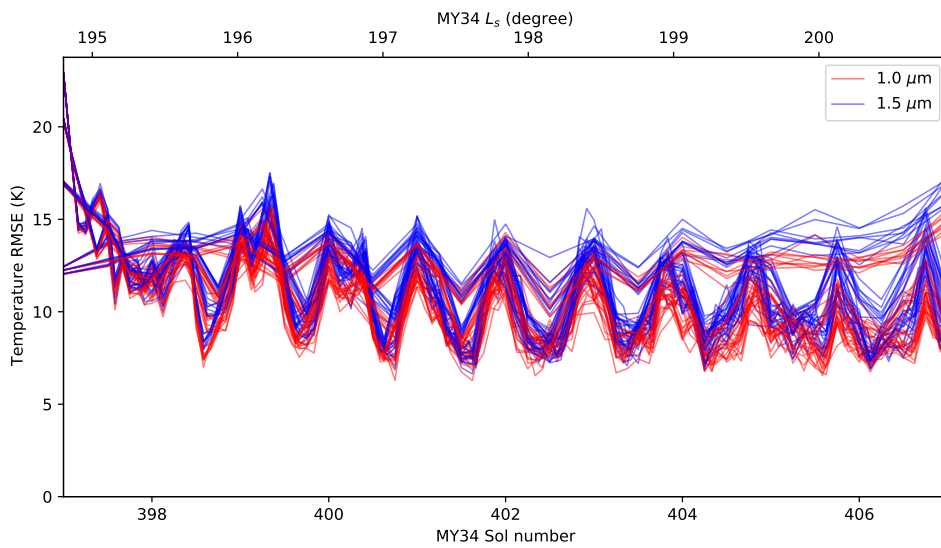
456 The assimilation scheme has several tuneable parameters. To optimise the temper-
 457 ature assimilation we performed sensitivity tests running short assimilations during the
 458 onset of the MY34 GDS (i.e. the period during which the dust concentrations are chang-
 459 ing most rapidly). We varied four parameters, and ran an assimilation for each combi-
 460 nation of values, 72 in total. We varied the number of ensemble members N (18, 36, and
 461 72), the length of the assimilation cycle Δt (2, 3, 6, and 12 hours), the threshold on the
 462 averaging kernel sum filter (0.50–1.50, 0.75–1.25, and 0.85–1.15), and the effective ra-
 463 dius of new dust added to the model r_{eff} (1.0 μm and 1.5 μm).

464 The goal of the ensemble is to sample the range of uncertainties in the forecast model.
 465 However, the ensemble is finite, and the mean and covariance of a finite ensemble will
 466 differ from the mean and covariance of an infinitely large ensemble. If an infinitely-large
 467 ensemble spanning all possible atmospheric states has mean \bar{x} and variance σ^2 , then an
 468 ensemble of size N will have an expected error in the ensemble mean of σ/\sqrt{N} , and an
 469 expected error in the ensemble variance of $\sigma^2\sqrt{2/(N-1)}$ (Ahn & Fessler, 2003). Hence
 470 if the ensemble is too small then outlying values and noise will have a disproportionately
 471 large influence on the ensemble mean and variance, so we expect the assimilation to be
 472 more accurate for larger ensembles. This is an important part of the rationale for using
 473 localization in LETKF assimilation (Hamill et al., 2001). Conversely, larger ensembles
 474 consume more resources, and we might expect the forecast step to be more unstable, as
 475 the simulations span a larger (and hence more extreme) range of atmospheric states.

(a) Verification against TIRVIM at 10–30 km pseudo- z , coloured by cycle length.



(b) Verification against MCS at 10–100 Pa, coloured by new dust effective radius.



476

477 **Figure 5.** Time series showing RMS error for short test assimilations verified against (a)
 478 TIRVIM temperature observations at 10–30 km pseudo-altitude, and (b) MCS temperature ob-
 479 servations at 10–100 Pa. In (a) the lines are colour-coded by the length of the assimilation cycle,
 480 and in (b) they are colour-coded by the dust effective radius used for new dust. The different
 481 lines for a given colour correspond to variations in the other three parameters.

482 There is a trade-off when choosing the assimilation cycle length between the fore-
 483 cast length and the number of grid points that are updated during each cycle. When the
 484 cycle is short, the model has less time to diverge from reality. For Mars the forecast er-
 485 ror can saturate in 3–6 hours (Navarro et al. (2017), Fig. 2). But when the analysis cy-
 486 cle is longer a larger fraction of the model grid points are updated during each analy-
 487 sis. This is particularly important for assimilation of tracer quantities such as dust, be-
 488 cause the only way that new dust is added to the model state is during the assimilation
 489 step.

490 Figure 4c shows different averaging kernel sum filter limits. The very narrow range
 491 (0.85–1.15) excludes many observations near 10 km and 40 km pseudo-altitude, the mod-
 492 erate range (0.75–1.25) encompasses those observations, and the wide range (0.5–1.5)
 493 includes a lot of observations, but some will have 50% of their information content from
 494 the prior profile.

495 The dust effective radius r_{eff} for new dust added to the analysis state is tested for
 496 $1.0 \mu\text{m}$, because Navarro et al. (2017) use this value, and $1.5 \mu\text{m}$, because the dust re-
 497 trievals we assimilate use that value (Guerlet et al., 2022). Dust with smaller radius takes
 498 longer to sediment and fall out onto the surface.

499 We ran each test assimilation from sols 397–407 ($L_s = 194.8\text{--}200.9^\circ$), using the
 500 TuTD configuration (i.e. no CDOD assimilation), and we compared the results with TIRVIM
 501 and MCS temperature profiles. We interpolated the analysis in time and space to the
 502 TIRVIM (on σ levels) and MCS (on pressure levels) observations. To compare with TIRVIM
 503 we applied the averaging kernel operator and the averaging kernel sum filter. We did not
 504 apply any vertical smoothing before making the comparison with MCS observations, as
 505 the vertical resolution of those observations is comparable with the grid level spacing.
 506 We then computed the bias and the root-mean squared error (RMSE) over various pseudo-
 507 altitude ranges for TIRVIM (0–10, 10–30, and 30–50 km) and various pressure ranges
 508 for MCS (100–1000, 10–100, 1–10, 0.1–1, and 0.01–0.1 Pa). Time series for each of the
 509 72 test assimilations are shown in Fig. 5. Because we ran every possible combination of
 510 parameters, we were also able to plot individual comparisons between cases varying just
 511 one parameter with all others held constant. These are shown in Fig. 6. In both figures
 512 we only show select cases where there were significant differences (gauged by eye) be-
 513 tween parameter values.

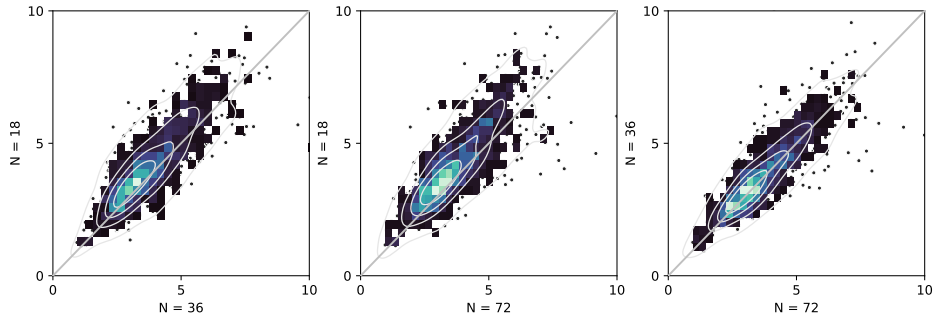
514 The clearest difference is between the assimilation cycle lengths. Over altitude ranges
 515 where TIRVIM is sensitive (i.e. up to 50 km), the shorter the cycle the better the match
 516 to observations. Figure 5a shows a representative set of RMSE time series, and Fig. 6b
 517 shows scatter plots, both verifying against TIRVIM. This trend remains when compared
 518 against MCS, although it is not as strong. A clearer trend when verifying against MCS
 519 is that $r_{\text{eff}} = 1.0 \mu\text{m}$ for new dust is generally better than $1.5 \mu\text{m}$, particularly for 1–
 520 100 Pa (i.e. where the TIRVIM observations are), as illustrated in Figs 5b and 6d.

521 The ensemble size makes a smaller difference, but overall both $N = 36$ and $N =$
 522 72 produce lower RMSE than $N = 18$ (Figure 6a). The differences between the aver-
 523 aging kernel sum filters are smaller still (Fig. 6c); 0.50–1.50 clearly performs the worst,
 524 but there is no visible difference between 0.75–1.25 and 0.85–1.15.

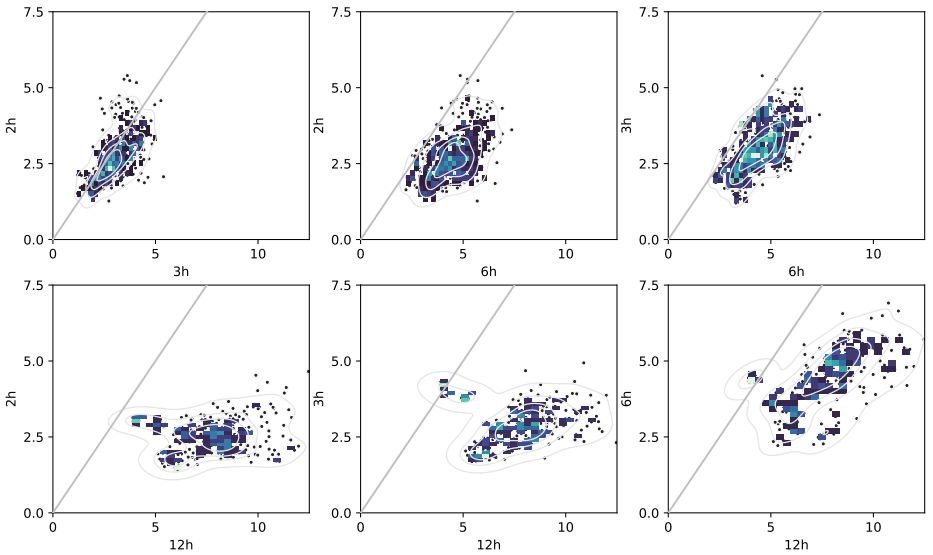
534 The only parameter that had any effect on the practical aspects of running the as-
 535 simulation was the ensemble size. The clock time to assimilate 1 sol of data was strongly
 536 correlated with ensemble size, with $N = 72$ taking 38–58 min (5–95th percentiles over
 537 all $N = 72$ cases) while $N = 36$ took 25–39 min and $N = 18$ took 22–34 min. The
 538 forecast step run time depends on the maximum time for any one ensemble member to
 539 finish, so the more ensemble members there are, the longer this is. The number of times
 540 individual forecast members became unstable was significant only for $N = 72$ (1.2–13.6
 541 times per sol, compared with 0.1–1.2 for $N = 36$). We had planned to run a series of
 542 tests with $N = 108$, but chose not to do so as a result.

543 In conclusion, $r_{\text{eff}} = 1.0 \mu\text{m}$ for new dust was generally better than $1.5 \mu\text{m}$. An
 544 averaging kernel sum filter of 0.75–1.25 was better than 0.5–1.5 in most cases, and marginally
 545 better than 0.85–1.15 in some cases. As it allows more observations to be assimilated
 546 over a wider range of altitudes, we chose the more inclusive 0.75–1.25 condition. An en-
 547 semble size of $N = 36$ was better than $N = 18$ in most cases. $N = 36$ and $N = 72$
 548 were very similar, but $N = 72$ required lots of manual restarts and had more model in-
 549 stability in the forecast step, so we chose $N = 36$ ensemble members. Finally, a 2-hour
 550 cycle length was better than 6- and 12-hours in most cases. It was also better than 3 hours
 551 in some cases, when compared with TIRVIM observations, but 2- and 3-hour cycles were
 552 generally similar. As a 2-hour cycle provides higher time resolution in the final data prod-
 553 uct, we chose $\Delta t = 2$ hrs.

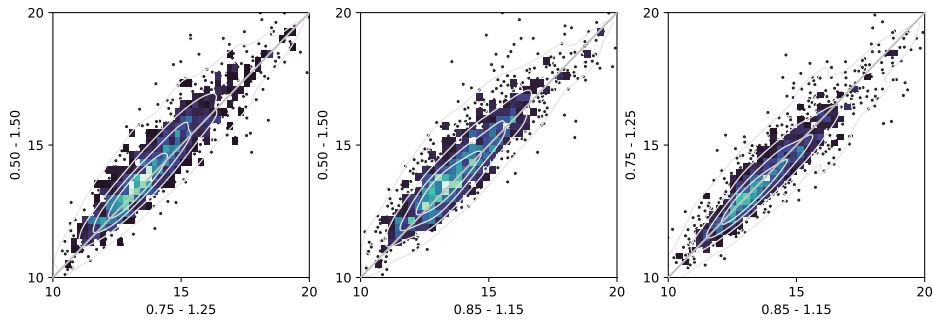
(a) Varying ensemble size N , verifying against TIRVIM at 30–50 km.



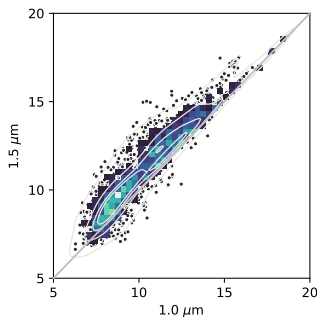
(b) Varying cycle length Δt , verifying against TIRVIM at 10–30 km.



(c) Varying averaging kernel sum filter, verifying against MCS at 10–100 Pa.



(d) Varying new dust effective radius r_{eff} , verifying against MCS at 1–10 Pa.



525

526

527

528

529

530

531

532

533

Figure 6. Scatter plots showing temperature RMS error comparisons between 10-sol test assimilations, varying one parameter at a time. Each dot corresponds to a single analysis. Filled and line contours indicate point density, and individual points are outliers. Equality is the diagonal grey line, and all axes are in K (note not all go to zero). In (a) ensemble size N is varied, verifying against TIRVIM at 30–50 km; (b) varies the assimilation cycle length Δt , verifying against TIRVIM at 10–30 km; (c) varies the averaging kernel sum filter range, verifying against MCS at 10–100 Pa; and (d) varies the effective radius for new dust r_{eff} , verifying against MCS at 1–10 Pa.

554 **5 Results**

555 We ran three versions of the assimilation between sols 376–424 of MY34, encom-
 556 passing the onset and peak of the Global Dust Storm. These were TuTD (temperature
 557 updates temperature and dust profile), TuT-CuD (temperature updates temperature,
 558 and then column dust updates the dust column and hence the dust profile), and TuTD-
 559 CuD (temperature updates temperature and dust profile, and then column dust updates
 560 the dust column and hence the dust profile). The assimilation period splits nicely into
 561 three distinct segments: before the storm (sols 376–385, $L_s = 182.3 - 187.6^\circ$); during
 562 the onset of the storm (sols 385–410, $L_s = 187.6-202.7^\circ$); and at the peak of the storm
 563 (sols 410–424, $L_s = 202.7 - 211.4^\circ$).

564 In assimilation experiments it is usual to compare the analysis with a free running
 565 model as a control. Alongside the three assimilations we ran a “GCM ensemble” of GCM
 566 simulations without assimilation, where the dust column abundances are constrained to
 567 match the MY34 dust scenario (Montabone et al., 2015, 2020). This is not a true free-
 568 running model, but it is the standard configuration for this GCM; if the LMD Mars GCM
 569 is run without such constraints then dust will sediment out and the model will crash af-
 570 ter O(10 sols) as the lack of airborne dust leads to extremely low atmospheric temper-
 571 atures. When the GCM is run with a dust scenario, dust is added and removed from the
 572 atmosphere by rescaling to the column abundances, by dust lifting from the surface (which
 573 occurs at a constant rate), and by sedimentation onto the surface. The results below use
 574 the mean of this ensemble when referring to the “GCM ensemble”.

575 One additional complication is that the dust scenario used in the GCM ensemble
 576 is based on MCS observations. So when our analyses are verified against MCS observa-
 577 tions and the assimilation’s performance compared against the GCM ensemble, the com-
 578 parison between the GCM ensemble and MCS observations is more like an “in-sample
 579 comparison”, while the comparison between the analyses and MCS observations is a com-
 580 pletely independent verification. When we assimilate column dust optical depths directly,
 581 the comparison between our analyses and the GCM ensemble is a fairer comparison than
 582 when they are not assimilated.

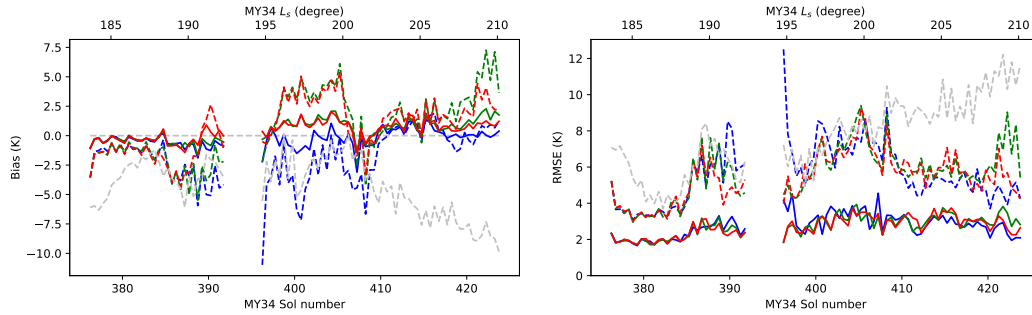
583 **5.1 In-sample Verification against TIRVIM Observations**

584 We computed the forecast and analysis bias and RMSE with respect to the TIRVIM
 585 observations that were assimilated, as described in Sect. 4, and over the same vertical
 586 ranges. We also computed CDOD at 1090 cm^{-1} from the forecasts and analysis using
 587 Eq. 7, and verified them against TIRVIM CDOD observations (i.e. with fixed Q_{ref} and
 588 r_{eff}).

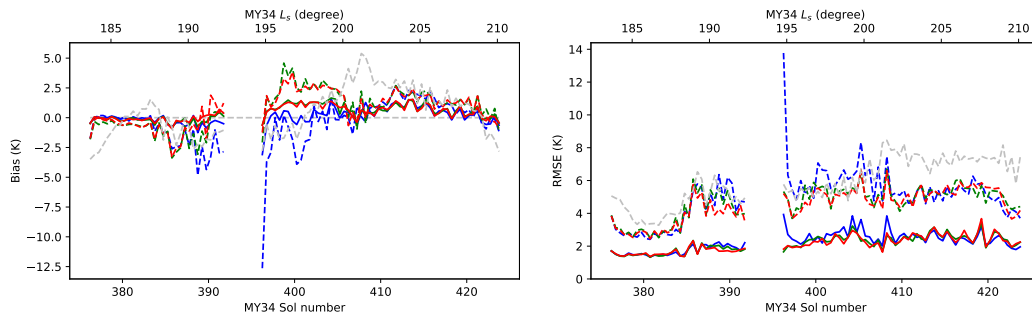
596 Figure 7 shows the temperature bias and RMSE, and Fig. 8 shows the CDOD bias
 597 and RMSE. In both fields the analysis RMSE is typically 50% of the forecast RMSE, and
 598 the analysis bias is close to zero, showing that the assimilation reduces the error by adding
 599 observational information to the system. The temperature RMSE is close to the observed
 600 uncertainty (Fig. 1). There is no systematic bias in the analyses at any altitude. The
 601 forecasts typically have a cool bias below 10 km altitude, and between 30–50 km in the
 602 TuTD case there is a cool bias while in the other two cases there is a warm bias from
 603 about sol 398 onwards. The GCM ensemble is generally biased cool before the dust storm
 604 begins, and once the dust storm begins it is biased warm at 0-10 km and cold at 30-50
 605 km. There is a gap in the data between sols 391.7126–396.4636, after which the forecast
 606 RMSE is temporarily high, but returns to an equilibrium after about 1 sol (note this also
 607 quantifies the convergence time for the assimilation).

608 For the CDOD observations, when dust is assimilated the error reduction by the
 609 assimilation is about 50% before the GDS begins, and 20–40% once the GDS has begun.
 610 In the TuTD case (where no dust observations are assimilated, so this is an out-of-sample
 611 verification), the error reduction is close to zero, except at the peak of the storm. There
 612 is a significant negative bias in the TuTD case (i.e. significantly less dust in the anal-
 613 ysis than in the TIRVIM observations). When dust is assimilated directly the analysis
 614 error is significantly less, leaving a small negative bias of about 0.1–0.2 in each case. This
 615 is comparable with the typical uncertainty in the CDOD measurements (mean 0.16 over
 616 this study period) The GCM ensemble has a strong positive bias (about 1.0) compared
 617 with the TIRVIM CDOD observations at the peak of the storm.

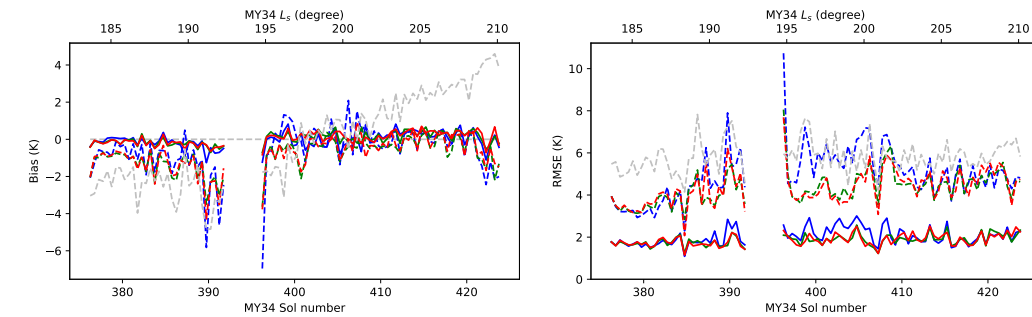
(a) 30–50 km.



(b) 10–30 km.

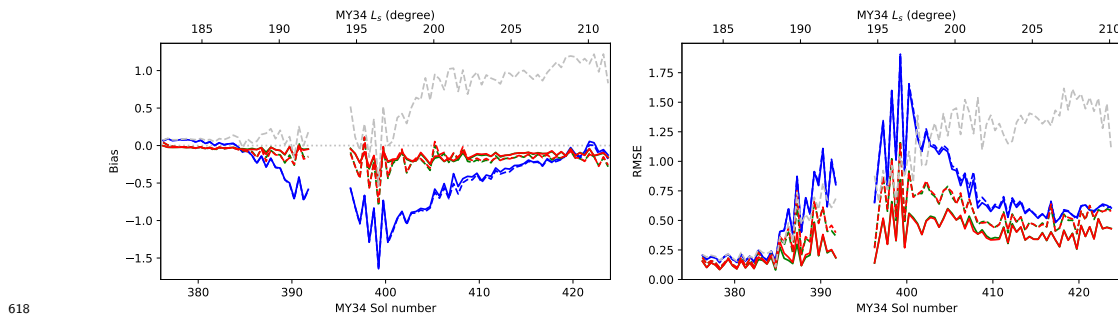


(c) 0–10 km.



589
590

591 **Figure 7.** In-sample verification against TIRVIM temperature profiles. The left column shows
592 the bias (analysis/forecast – observations) and the right column shows the RMS error. Data
593 are split by pseudo-altitude: (a) 30–50 km; (b) 10–30 km; and (c) 0–10 km. Solid lines show
594 the analysis mean, and dashed lines show the forecast mean. TuTD is blue, TuT-CuD is green,
595 TuTD-CuD is red, and the GCM ensemble is grey. Data are averaged over 0.5-sol bins.



618

619 **Figure 8.** As Fig. 7, but showing in-sample verification against TIRVIM extinction column
 620 dust optical depths.

621

5.2 Out-of-sample Verification against MCS Observations

622

623

624

We verified the analysis and forecast means against MCS temperature observations as described in Sect. 4. We also computed the dust density-scaled opacity (DSO) in extinction at $21.6 \mu\text{m}$ (Eq. 7 from Madeleine et al. (2011)):

$$\tau_{\text{DSO}} = \frac{3 Q_{\text{ref}}(r_{\text{eff}})q}{4 \rho_{\text{dust}}r_{\text{eff}}} \quad (8)$$

625

where in this case both r_{eff} and Q_{ref} take the local dust effective radius into account.

626

627

628

629

630

Figures 9 and 10 show bias and RMSE time series for temperature and dust DSO, over pressure ranges where the assimilated TIRVIM observations typically lie. These results are clearly poorer than the in-sample verification against TIRVIM observations, but that is not surprising. However, it is important to identify and understand the degree to which they are poorer.

638

639

640

641

642

643

The GCM ensemble, which is constrained by CDOD derived from MCS dust opacities, has a bias and RMSE that falls within the range of values from the three cases assimilating TIRVIM observations. This is encouraging because it implies that, while the absolute errors are larger than when compared with TIRVIM observations, the out-of-sample comparison is comparable with what is effectively an in-sample comparison between the GCM ensemble and MCS observations.

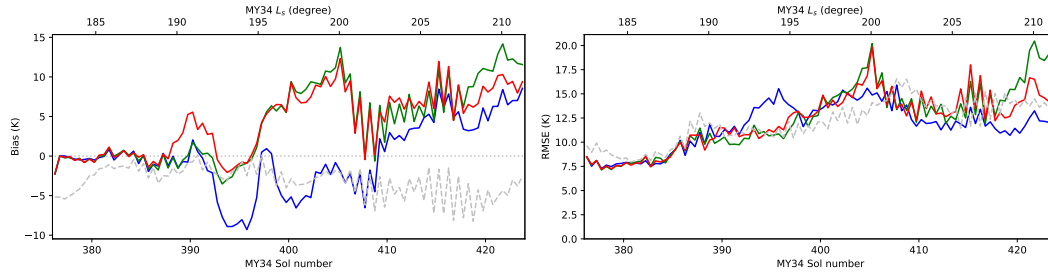
644

645

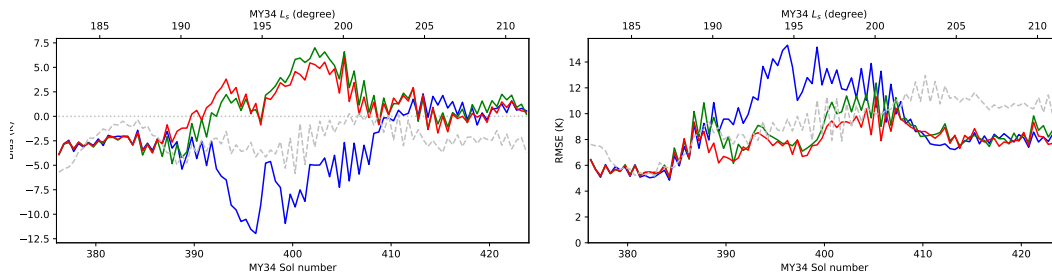
646

Before the storm there is little difference between the bias and RMSE in the three cases where TIRVIM observations are assimilated. Lower in the atmosphere than 100 Pa, the temperature analysis (as well as the GCM ensemble) has a cold bias of 2–8 K

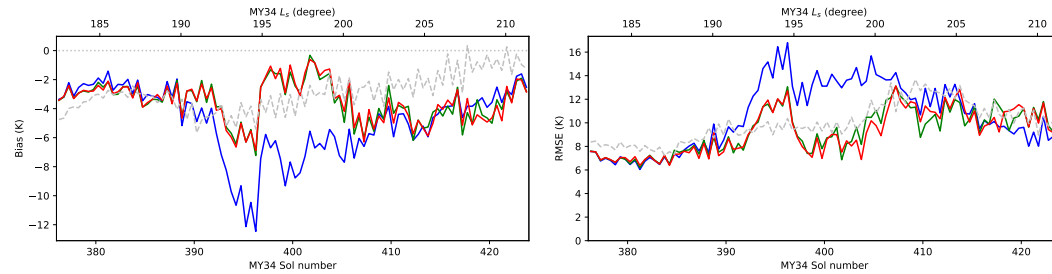
(a) 1–10 Pa



(b) 10–100 Pa



(c) 100–1000 Pa



631
632

633 **Figure 9.** Independent verification against MCS temperature profiles. The left column shows
634 the bias (analysis/forecast – observations) and the right column shows the RMSE. Data are split
635 by pressure ranges: (a) 1–10 Pa; (b) 10–100 Pa; and (c) 100–1000 Pa. Solid lines show analysis
636 means, and the dashed line shows the mean of the GCM ensemble. TuTD is blue, TuT-CuD is
637 green, TuTD-CuD is red, and the GCM ensemble is grey. Data are averaged over 0.5-sol bins.

647 in all cases; we note that when dust observations are assimilated this bias is closer to zero.
 648 Higher in the atmosphere than 100 Pa, the TuTD case has a cold bias during the onset
 649 of the storm, while the TuT-CuD and TuTD-CuD cases have a warm bias during that
 650 period, with the three cases converging at the peak of the storm. Similarly, during the
 651 onset of the storm the RMSE in the TuTD case is significantly poorer than in the two
 652 other assimilation cases, showing that it is crucial to assimilate some information about
 653 dust when it is changing quickly.

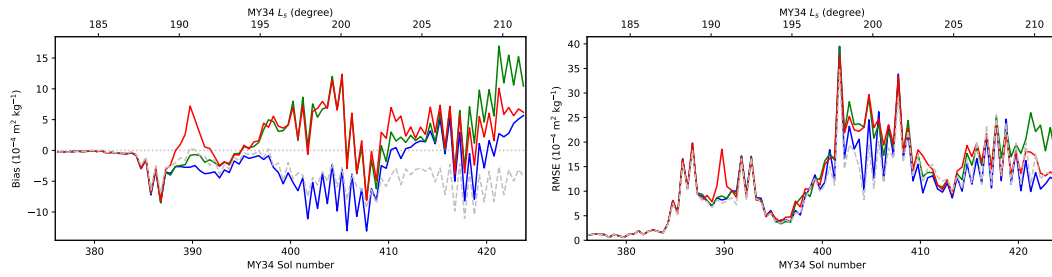
658 There is significantly lower dust DSO during the onset of the storm in the TuTD
 659 case than in the MCS observations, at all pressure ranges shown (Fig. 10). For TuT-CuD
 660 and TuTD-CuD there is, on average, a higher dust DSO than in the MCS observations
 661 during the onset of the storm, and by the peak of the storm the bias is close to zero be-
 662 tween 10-100 Pa and positive in the other two ranges. The GCM ensemble overestimates
 663 the amount of dust below the 100 Pa pressure level, and underestimates it above. At all
 664 altitudes the RMSE increases from a low level before the onset of the storm, and increases
 665 thereafter. As for temperature, we note that the RMSE in the three analyses is compa-
 666 rable with the RMSE in the GCM ensemble.

667 The amount of dust in the atmosphere is considerably more when dust observa-
 668 tions are assimilated, and so the heating effect of more dust helps to explain why, on av-
 669 erage, the cases assimilating dust are warmer than the TuTD case. At the peak of the
 670 dust storm the analysis temperature bias in all three cases between 10-100 Pa is close
 671 to zero, while the GCM ensemble has a cold bias. This is possibly because the GCM en-
 672 semble does not exhibit detached dust layers, with its dust distribution monotonically
 673 decreasing upwards, so the dust is more concentrated at the surface, heating the low-
 674 est atmospheric layers and cooling the higher layers, compared with when dust is assim-
 675 ilated. This is analysed further in the next section.

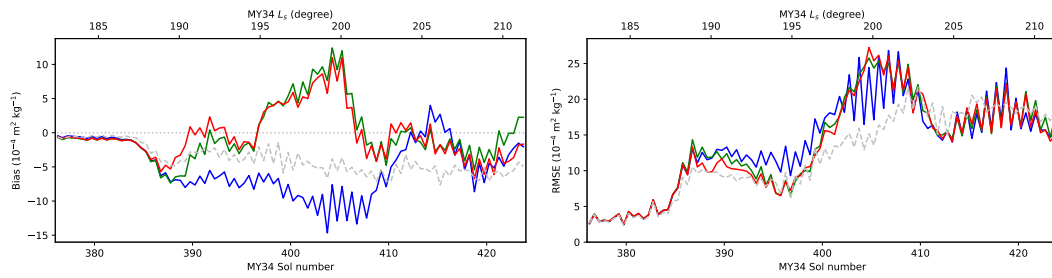
676 **5.3 Temperature and Dust Structure during the Onset of the MY34 GDS**

677 We now compare the atmospheric temperature and dust structure between the anal-
 678 yses, the GCM ensemble, and MCS observations. Figures 11 and 12 show Hovmöller di-
 679 agrams for temperature and dust DSO at 21.6 μm , at 30 Pa (where MCS dust concen-
 680 trations are highest during the storm). These cover all latitudes starting before the storm
 681 and ending just after its peak. Figure 13 shows vertical cross-sections of temperature at

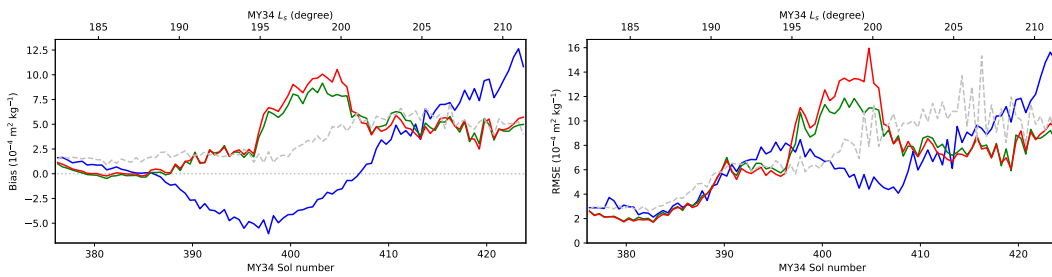
(a) 1–10 Pa.



(b) 10–100 Pa.



(c) 100–1000 Pa.



654

655

656 **Figure 10.** As Fig. 9, but showing independent verification against MCS extinction dust

657 density-scaled opacity profiles.

682 3 PM and 3 AM local mean solar time (chosen because that is the observation time for
 683 MCS), and Fig. 14 shows the same for dust density-scaled opacity.

684 There is significant heating associated with the global dust storm, with equatorial
 685 temperatures rising from 180 K to 225 K over this period (Fig. 11). Assimilating dust
 686 observations significantly improves the temperature analysis with respect to the inde-
 687 pendent MCS observations. When no dust observations are assimilated (Fig. 11a) the
 688 temperature change due to the storm over this L_s range is somewhat slower than observed
 689 by MCS (although 225 K is reached by the end). But when column dust optical depths
 690 are assimilated (Fig. 11 b-c) the change in temperature better matches the MCS obser-
 691 vations. Other details also match better, such as the hemispheric asymmetry at the peak
 692 of the storm. The differences between the case with and without the TuD step are small,
 693 but TuTD-CuD gives marginally better results, particularly for sols 391–396 (Fig. 11 c).

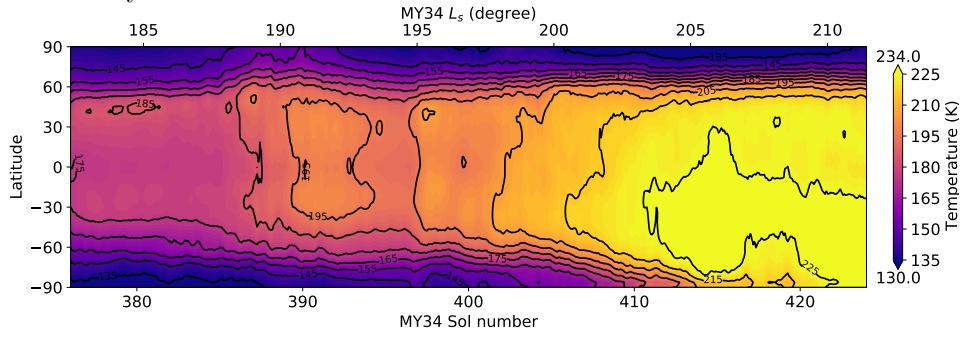
694 The GCM ensemble captures the minimum at the equator around sol 395 better
 695 than the analyses, but overall the analyses match the MCS observations better. The for-
 696 mer underestimates the atmospheric temperature at most latitudes, and does not fully
 697 reproduce the observed hemispheric asymmetry at the peak of the storm.

698 The temperature results at 3 AM are similar, although the differences between the
 699 GCM ensemble and the analyses are smaller. Hovmöller diagrams for 3 AM are shown
 700 in Supporting Fig. S1.

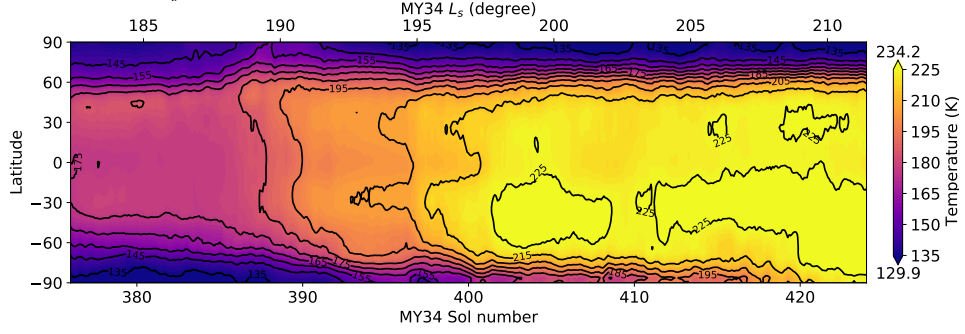
711 The differences in dust DSO between the various configurations (Fig. 12) are larger
 712 than for temperature. MCS observations show a peak in the dust DSO around 60×10^{-4}
 713 $\text{m}^2 \text{kg}^{-1}$. The GCM ensemble (Fig. 12d), whose column abundance is constrained by
 714 MCS observations, matches the dust DSO reasonably well, at least in terms of its evo-
 715 lution during the onset of the storm up to the peak, although the distribution is more
 716 strongly concentrated at the equator than for the MCS observations.

717 The three analyses each contain overall less dust than do the MCS observations.
 718 The largest difference is between TuTD and the other two cases. For TuTD (Fig. 12a),
 719 dust opacity increases slowly during the onset of the storm, and only begins to change
 720 rapidly close to its peak. The other two cases (Figs 12 b-c) follow the increase in dust
 721 during the onset of the storm better, although neither reach the peak opacity measured
 722 by MCS at this pressure by the peak of the storm. None of the three analyses reproduce

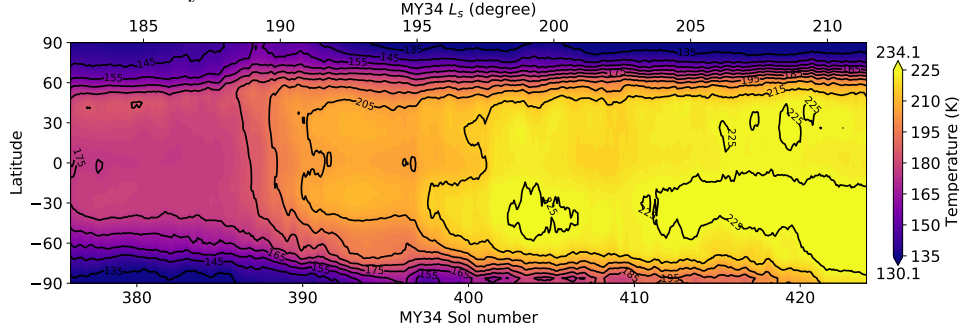
(a) TuTD analysis.



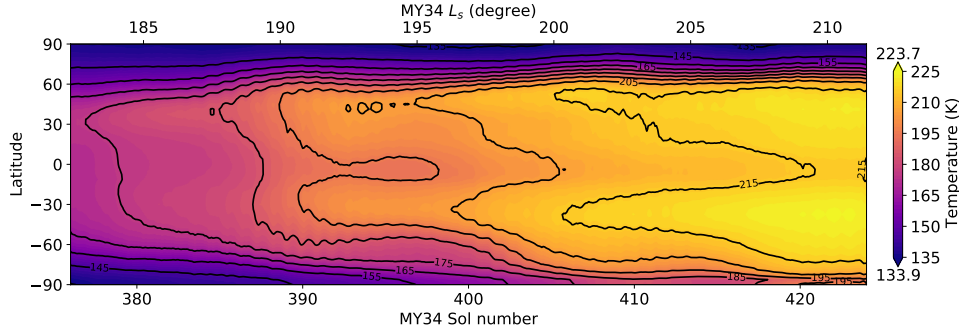
(b) TuT-CuD analysis.



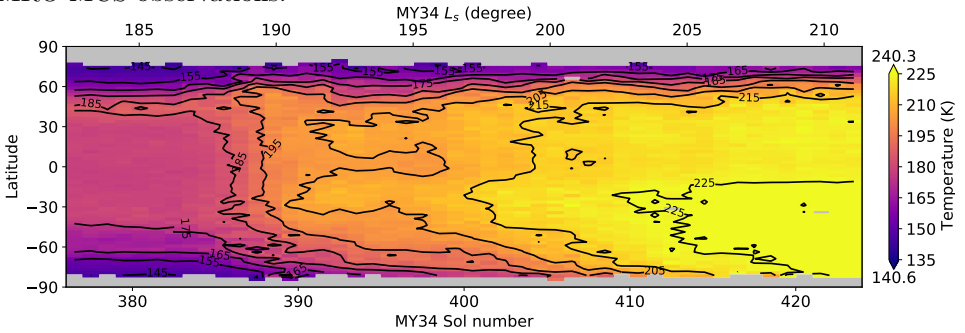
(c) TuTD-CuD analysis.



(d) GCM ensemble.



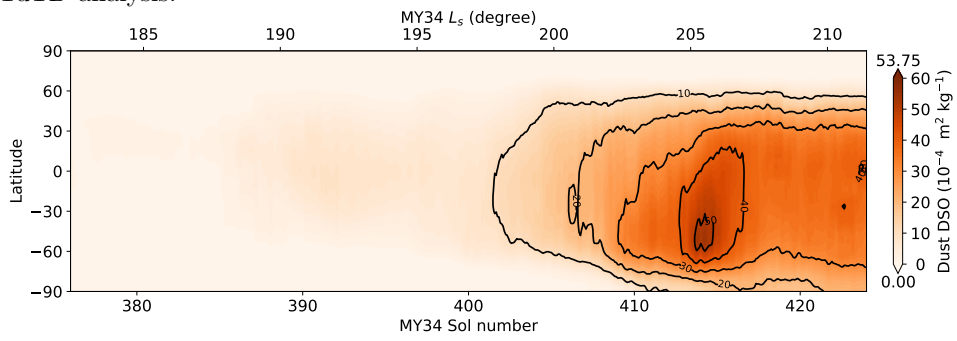
(e) MRO-MCS observations.



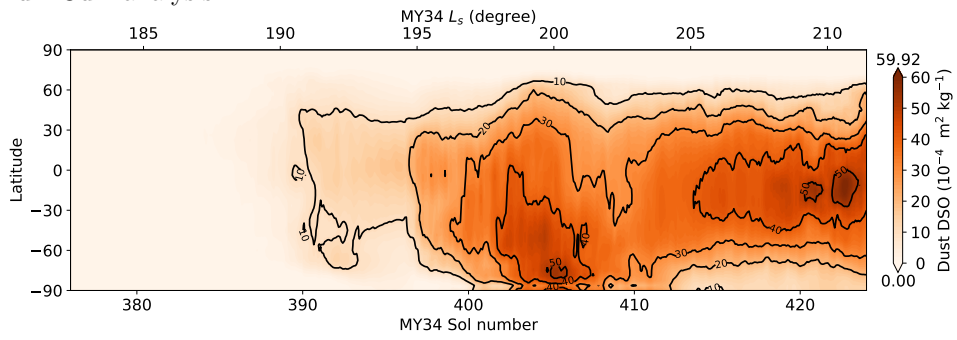
701

702 **Figure 11.** Hovmöller diagrams showing analysis temperature at 30 Pa at 3 PM local mean
 703 solar time (LMST). (a) TuTD analysis; (b) TuT-CuD analysis; (c) TuTD-CuD analysis; (d) GCM
 704 ensemble; (e) MRO-MCS observations. Data are smoothed using a 1-sol running mean. Grey
 705 indicates missing data. Line contours are every 10 K. MCS observations between 2-4 PM LMST
 706 are used; analyses are interpolated to 3 PM LMST. Maxima and minima are shown above and
 707 below the colour bar.

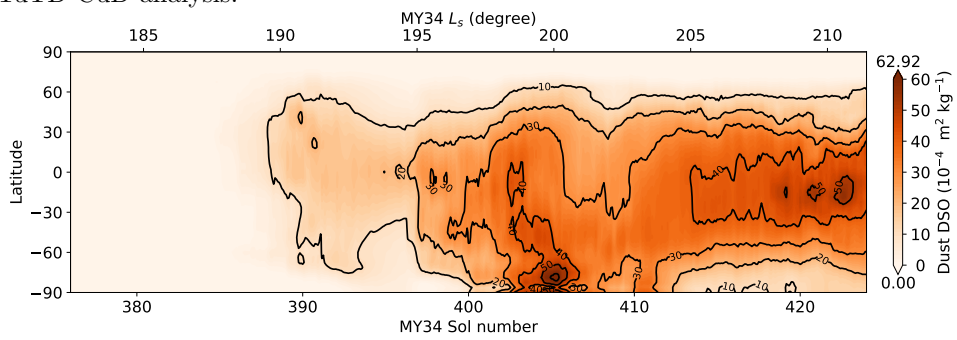
(a) TuTD analysis.



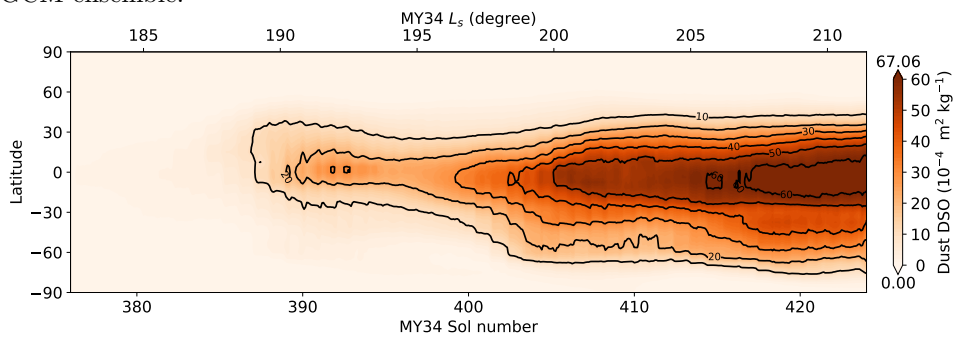
(b) TuT-CuD analysis.



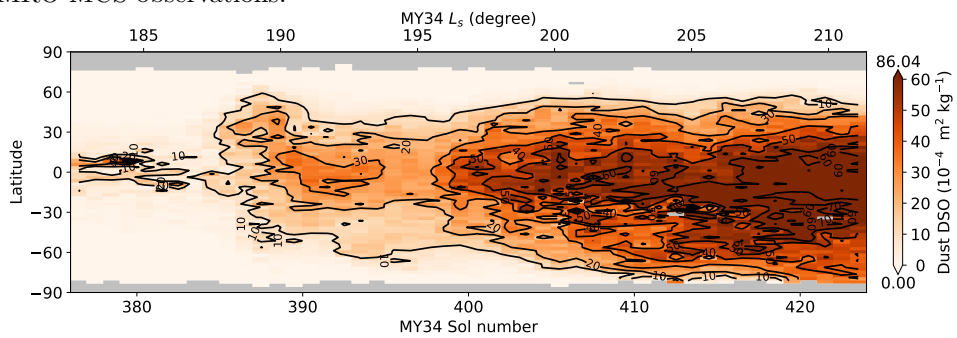
(c) TuTD-CuD analysis.



(d) GCM ensemble.



(e) MRO-MCS observations.



708

709

Figure 12. As Fig. 11 but showing Hovmöller diagrams for dust density-scaled opacity at 21.6

710

μm at 30 Pa. Line contours are every $10 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$.

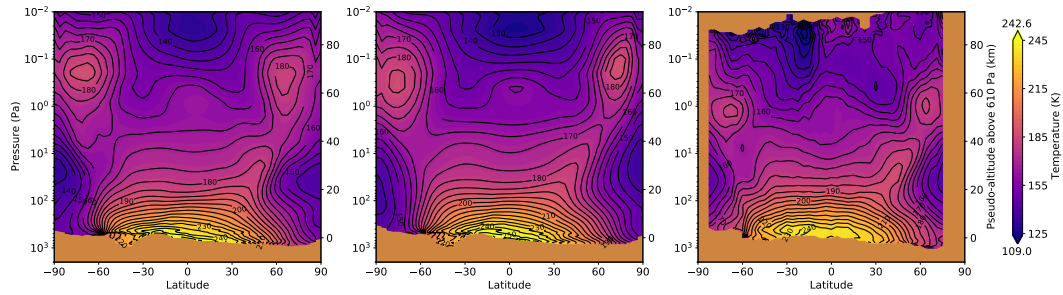
723 the short-lived peak at the equator between sols 389–393, but we note this partially co-
 724 incides with a gap in the TIRVIM data. Nevertheless, it is clearly better to have assim-
 725 ilated the dust information directly, both in terms of the temperature analysis and the
 726 dust analysis.

727 The equivalent plots at 3 AM, shown in Supporting Fig. S2, show a more favourable
 728 comparison between the TuT-CuD / TuTD-CuD analyses and MCS observations. Both
 729 reach the peak dust opacity seen in MCS observations by the peak of the storm (which
 730 we note the GCM ensemble does not do at 3 AM). They also reasonably reproduce the
 731 growth in dust concentration from sol 397 onwards. Kleinböhl et al. (2020) discuss the
 732 significant diurnal variations of temperature and dust opacity in the MCS observations
 733 during the dust storm in some detail.

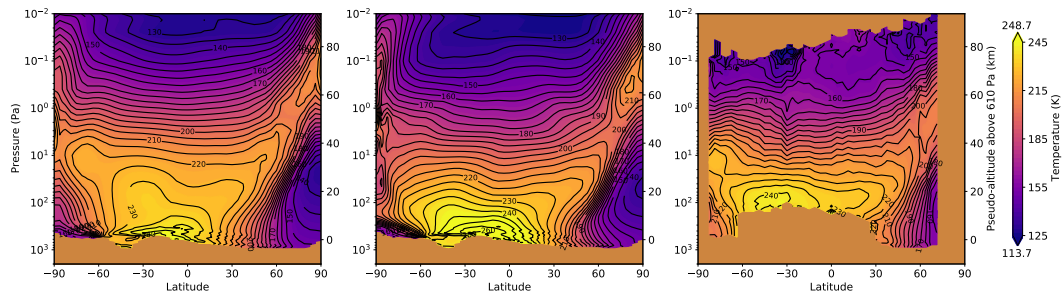
743 Figures 13 and 14 show vertical cross-sections of the temperature and dust density-
 744 scaled opacity before and at the peak of the global dust storm. These figures show how
 745 well the analysis reproduces the vertical structure, compared with the GCM ensemble
 746 and MCS observed profiles. In these figures and hereafter we only use our “best case”
 747 assimilation TuTD-CuD, alongside the GCM ensemble.

748 Both the temperature and dust opacity increase significantly during the storm, dur-
 749 ing daytime and nighttime. Unfortunately it is not possible to compare directly with day-
 750 time MCS observations at the peak of the storm near the surface, as the high dust con-
 751 centration prevents a reliable retrieval of both temperature and dust (Fig. 13b, right).
 752 Before the storm, both the analysis and the GCM ensemble match the MCS observa-
 753 tions well during nighttime. During the daytime the temperatures below 100 Pa are gen-
 754 erally lower than the MCS observations. This may be because at low altitudes during
 755 daytime the MCS observations are biased warm compared with TIRVIM due to a lack
 756 of retrievals in cloudy regions; see Guerlet et al. (2022), Fig. 19. The analysis and GCM
 757 ensemble both have a temperature maximum of 250–260 K close to the surface during
 758 the daytime at the peak of the storm. The maximum MCS temperature is at about 100
 759 Pa, and appears to be a local maximum. In northern midlatitudes the observed profiles
 760 do reach the ground, and the analysis reasonably reproduces the observed meridional tem-
 761 perature gradient from about 30° poleward (Fig. 13b). Both the analysis and the GCM
 762 ensemble are cooler than MCS at its maximum temperature point around 100 Pa, 30–
 763 50°S, with the analysis about 10–20 K cooler, and the GCM ensemble about 5–10 K cooler

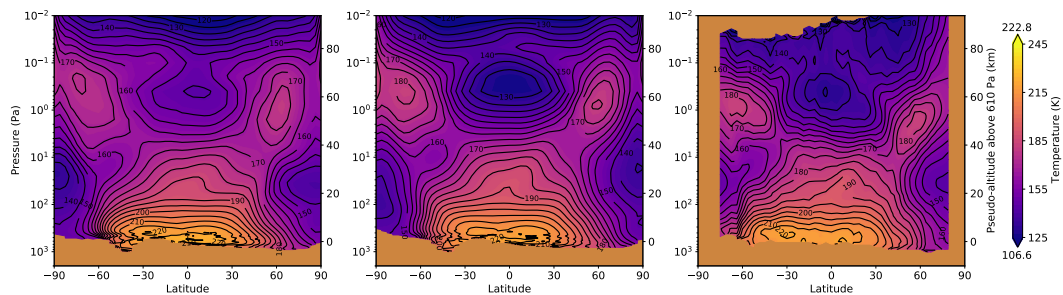
(a) 3 PM LMST, before the storm (sols 380–384).



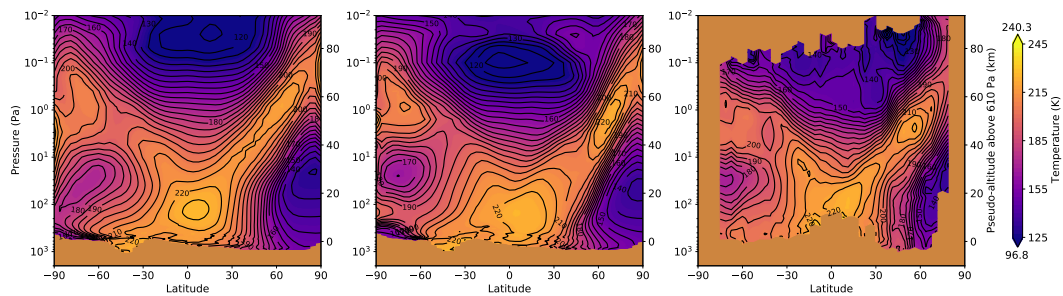
(b) 3 PM LMST, at the peak of the storm (sols 416–420).



(c) 3 AM LMST, before the storm (sols 380–384).



(d) 3 AM LMST, at the peak of the storm (sols 416–420).



734

735 **Figure 13.** Temperature cross-sections at 3 PM and 3 AM local mean solar time. The left
 736 column shows the TuTD-CuD analysis, the middle column shows the GCM ensemble, and the
 737 right column shows MCS observations (up to one hour either side of the target local time). (a-b)
 738 show 3 PM, and (c-d) show 3 AM. (a) and (c) show before the storm, averaged over sols 380–384,
 739 and (b) and (d) show the peak of the storm, averaged over sols 416–420. The right vertical axis
 740 shows the pseudo-altitude above 610 Pa. The numbers at the end of each colour bar indicate the
 741 maximum and minimum values in the MCS panels. Line contours are spaced every 5 K. Flat
 742 brown areas show missing data (either below the surface, or where there are no observations).

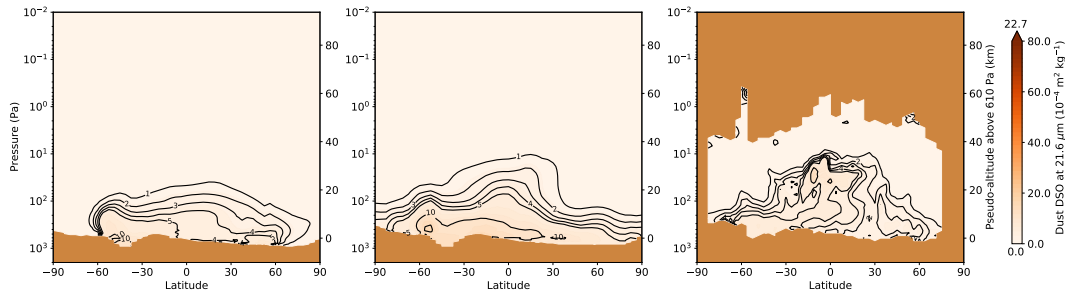
764 than observed. Both the analysis and the GCM ensemble reproduce the tongue of warm
 765 air in the middle atmosphere at northern hemisphere midlatitudes, which appears at the
 766 peak of the storm.

767 At 3 AM the analysis and GCM ensemble reproduce the MCS observations bet-
 768 ter at the peak of the storm than during the daytime (Fig. 13d). The analysis reproduces
 769 the local temperature maximum around 100 Pa that appears in the MCS observations
 770 better than does the GCM ensemble. The analysis also better reproduces the meridional
 771 temperature variation below 10 Pa. The analysis and GCM ensemble have quite differ-
 772 ent temperature structures near the top of the domain near the equator, with the local
 773 temperature minimum in the GCM ensemble being about a decade in pressure below where
 774 it is in the analysis. However, the MCS observations do not allow us to distinguish be-
 775 tween these two, as the observations cease just above 0.1 Pa, but there does not appear
 776 to be an obvious temperature minimum as the GCM ensemble predicts. Note also that
 777 the assimilated TIRVIM observations typically go up to about 1 Pa, so above this any
 778 differences are due to the GCM’s response to assimilated data rather than being con-
 779 strained directly by observations.

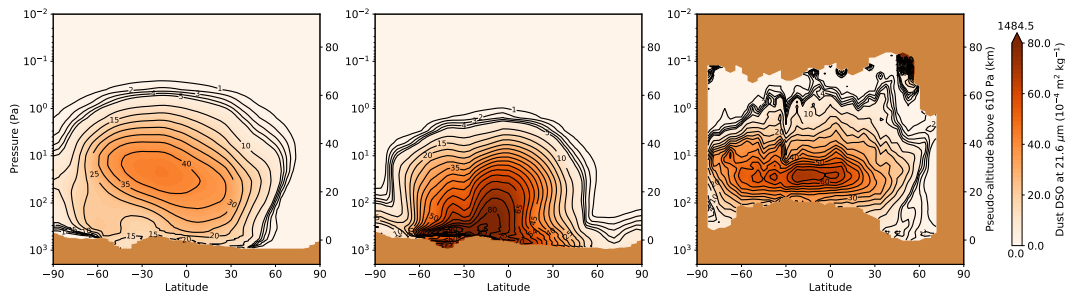
784 As in the Hovmöller diagrams, there is a larger difference between the analysis and
 785 the GCM ensemble in the dust opacity cross-sections (Fig. 14) than in temperature (Fig. 13).
 786 Before the storm, the amount of dust peaks around $5 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ in the analysis and
 787 $10 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ in the GCM ensemble, with little diurnal variation (Fig. 14 a, c). The
 788 MCS observations peak at about $5 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$, but the main difference is that the
 789 dust is more vertically extended in the observations, with a clear “detached dust layer”
 790 peak near the equator during the night-time around 100 Pa of about $5 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$,
 791 which is not visible in either the analysis or the GCM ensemble.

792 At the peak of the storm there are larger differences between daytime (Fig. 14b)
 793 and nighttime (Fig. 14d), as well as between the analysis, GCM ensemble, and MCS ob-
 794 servations (Fig. 14, different columns). In general, the GCM ensemble contains signif-
 795 icantly more dust than either the analysis or the MCS observations. The peak in the GCM
 796 ensemble is close to the ground in the southern hemisphere, about $90 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ at
 797 both 3 PM and 3 AM, while the MCS observations peak at about 30 Pa, with maxima
 798 around $60 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ at 3 PM and a concentrated peak of $50 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ at 3
 799 AM.

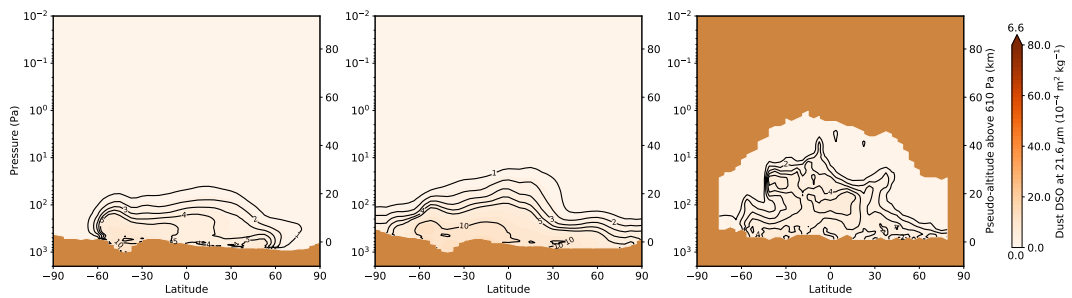
(a) 3 PM LMST, before the storm (sols 380–384).



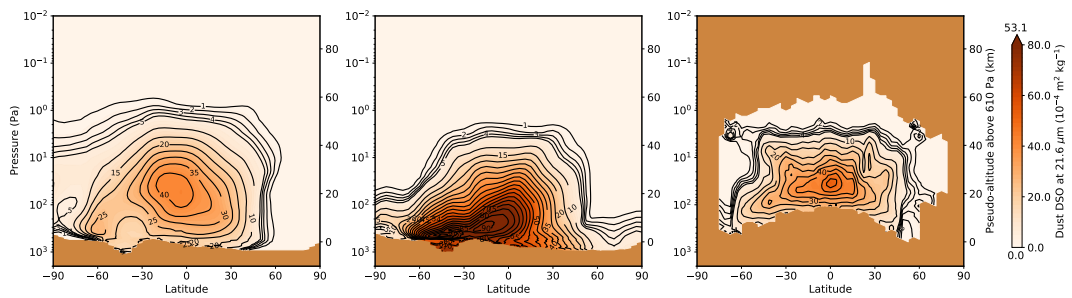
(b) 3 PM LMST, at the peak of the storm (sols 416–420).



(c) 3 AM LMST, before the storm (sols 380–384).



(d) 3 AM LMST, at the peak of the storm (sols 416–420).



780

781 **Figure 14.** As Fig. 13, but showing dust density-scaled opacity at 21.6 μm at 3 PM and 3
 782 AM local mean solar time. Line contours are spaced every $10^{-4} \text{ m}^2 \text{ kg}^{-1}$ from $1-4 \times 10^{-4} \text{ m}^2$
 783 kg^{-1} , and then every $5 \times 10^{-4} \text{ m}^2 \text{ kg}^{-1}$ thereafter.

800 The analysis reproduces the vertical and latitudinal distribution of dust opacities
 801 significantly better than does the GCM ensemble. In particular, it very clearly contains
 802 a detached dust layer at both 3 PM and 3 AM, with a broad peak between 10–100 Pa,
 803 as seen in the MCS observations, while the GCM ensemble peaks near the surface (al-
 804 though note not exactly at the surface). The peak dust DSO in the analysis is about 40×10^{-4}
 805 $\text{m}^2 \text{kg}^{-1}$ at both 3 PM and 3 AM. During the daytime this is an underestimate, as the
 806 dust opacity peak in the MCS observations is above $50 \times 10^{-4} \text{m}^2 \text{kg}^{-1}$ between 60°S and
 807 20°N . During nighttime the MCS observations peak at $50 \times 10^{-4} \text{m}^2 \text{kg}^{-1}$ near the equa-
 808 tor. However, there is significantly less dust than during the daytime, with the opacity
 809 above $40 \times 10^{-4} \text{m}^2 \text{kg}^{-1}$ between 20°S and 20°N . This distribution is matched well by
 810 the analysis dust opacity at this time, which is above $40 \times 10^{-4} \text{m}^2 \text{kg}^{-1}$ between 20°S
 811 and the equator.

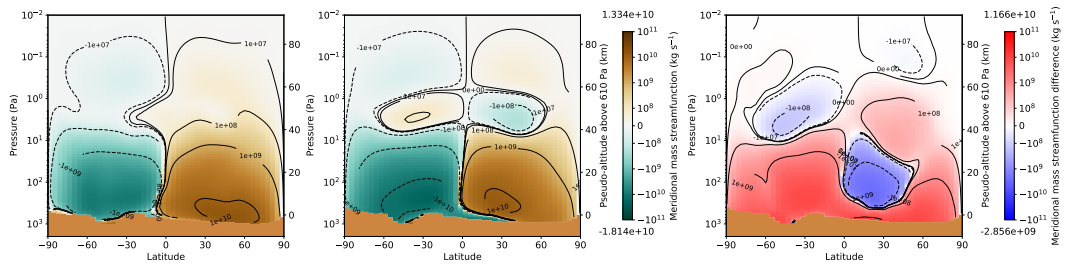
812 **5.4 Wind Structure Retrieved by Assimilation**

813 Having demonstrated that our reanalysis state fits the independent MCS observa-
 814 tions reasonably well, we now turn to wind, a quantity that is not observed but which
 815 is retrieved by the assimilation. This is one of the main strengths of the data assimi-
 816 lation process.

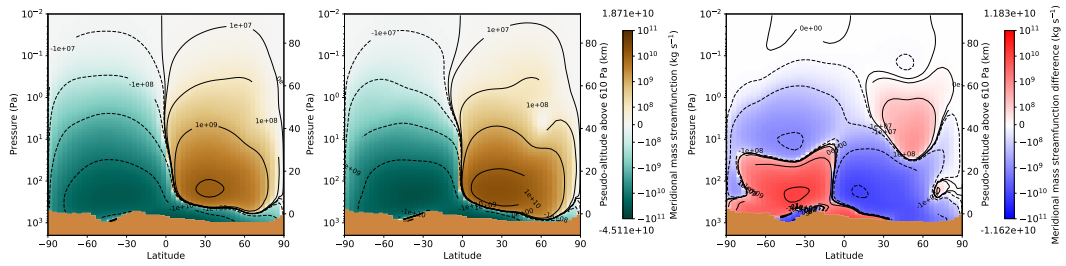
823 Figures 15 and 16 show the meridional mass stream function and zonal velocity at
 824 3 PM and 3 AM local mean solar time averaged over two time periods before the storm
 825 and at its peak. Before the storm, the daytime overturning circulation (Fig. 15 a-b) shows
 826 a Hadley cell covering the whole of each hemisphere. There is a weak counter-rotating
 827 cell higher than 10 Pa, which only appears in the GCM ensemble. The study period is
 828 close to the autumnal equinox, so the Hadley cell is approximately symmetric about the
 829 equator. During nighttime (Fig. 15 c-d) the overturning circulation is reversed, with the
 830 descending branch of the main overturning cell at the equator, with a weaker thermally
 831 direct cell between 1–10 Pa. The meridional overturning circulation is generally weaker
 832 in the analysis than in the GCM ensemble.

833 The overturning circulation strengthens during the daytime as the storm progresses,
 834 and extends higher into the atmosphere, significantly strengthening the overturning cir-
 835 culation aloft. The thermally-indirect cell present in the GCM ensemble disappears, leav-
 836 ing a single Hadley cell spanning the whole model domain from the surface to 100 km

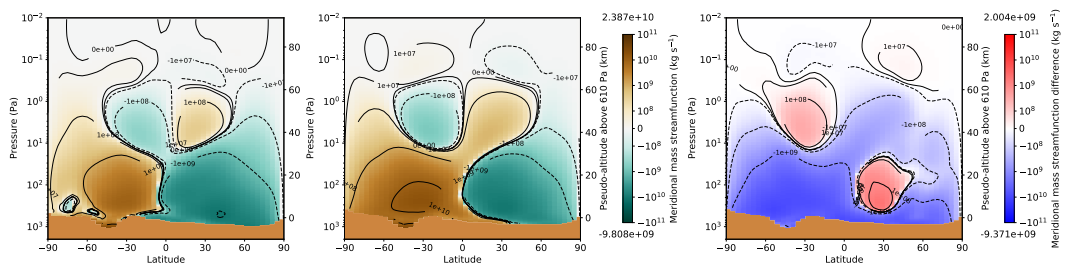
(a) 3 PM LMST, before the storm (sols 380–384).



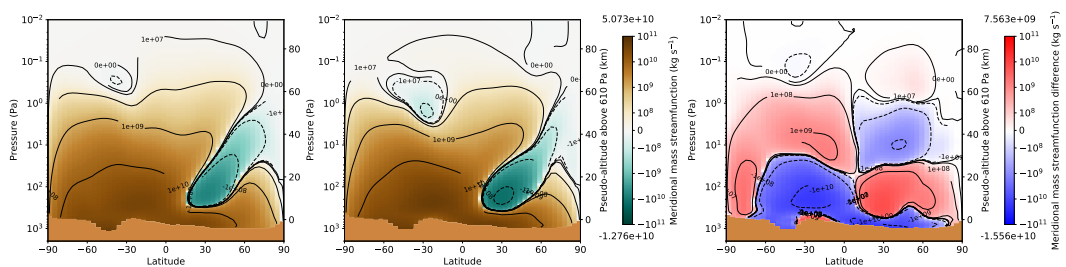
(b) 3 PM LMST, at the peak of the storm (sols 416–420).



(c) 3 AM LMST, before the storm (sols 380–384).



(d) 3 AM LMST, at the peak of the storm (sols 416–420).



817

818 **Figure 15.** As Fig. 13, but showing the meridional mass streamfunction at 3 PM and 3 AM
 819 local mean solar time. The right hand column shows the difference between the analysis and
 820 GCM ensemble (analysis — GCM). In the left two columns brown is clockwise motion and blue
 821 is anticlockwise motion. In the right column red means more clockwise, and blue means more
 822 anticlockwise. Line contours are spaced every factor of 10 between 10^8 and 10^{11} kg s^{-1} .

837 pseudo-altitude. Close to the surface, the southern hemisphere cell extends about 40°
 838 into the northern hemisphere. At nighttime the circulation changes considerably. The
 839 symmetric Hadley cells are severely disrupted, and a cross-equatorial cell develops which
 840 covers most latitudes with a weak Hadley circulation in the northern polar region.

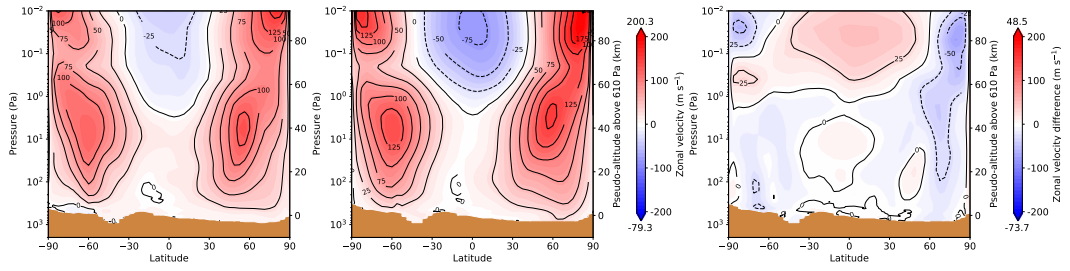
844 The zonal velocity response to the storm (Fig. 16) is noticeably asymmetric about
 845 the equator, both during daytime and nighttime. Like the meridional overturning cir-
 846 culation, this change is more profound during nighttime. Before the storm there are pro-
 847 grade jets in both hemispheres with a retrograde jet above 1 Pa at low latitudes, and
 848 these jets are slightly weaker at 3 AM. By the peak of the storm, the equatorial retro-
 849 grade jet and the prograde northern hemisphere jet increase in speed by about 25 m s^{-1}
 850 and 75 m s^{-1} respectively, while the southern hemisphere prograde jet remains approx-
 851 imately the same speed but its vertical extent shrinks, becoming a more focused jet around
 852 $75^\circ\text{S} / 10 \text{ Pa}$. During nighttime, the northern prograde jet strengthens by about 50 m
 853 s^{-1} , but the southern prograde jet weakens significantly by about 75 m s^{-1} , almost dis-
 854 appearing. At the peak of the storm, the two prograde jets differ in speed by 125 m s^{-1} .
 855 The strong strengthening of the northern hemisphere jet at both 3 PM and 3 AM is con-
 856 sistent with the steepening of the latitudinal temperature gradient near 60°N at the peak
 857 of the storm (Fig. 13 b-d), via thermal wind balance. A similar argument explains the
 858 weakening of the southern prograde jet: the temperature contours are flattened in the
 859 southern hemisphere around 10 Pa (Fig. 13 a-b)

860 Differences between the analysis and the GCM ensemble are relatively small com-
 861 pared with the magnitude of the jets, with typical differences up to $\pm 25 \text{ m s}^{-1}$ over most
 862 of the domain. The analysis winds are typically more eastward at lower latitudes than
 863 the GCM ensemble, in particular above 10 Pa, and they are typically more westward at
 864 higher latitudes.

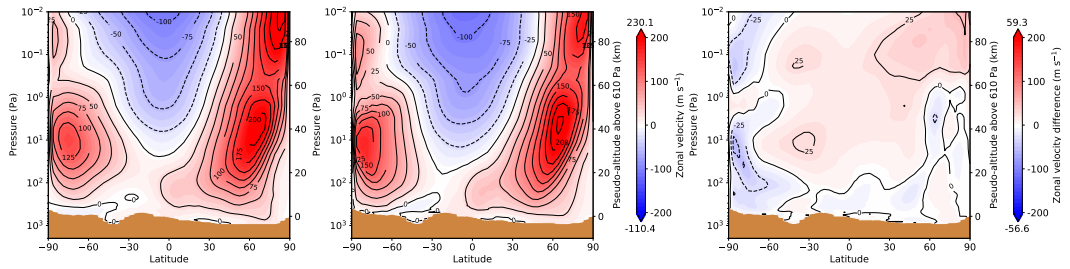
865 5.5 Diurnal Cycle and Thermal Tides

866 One of the reasons for the design of the ExoMars TGO orbit is that it samples the
 867 Martian atmosphere at different local times of day. In this section we present how this
 868 affects the analysis, and analyse how the diurnal cycle changes during the global dust
 869 storm. Assimilation constrains the atmosphere at particular local times, and then the
 870 forecast model transfers that information to other local times during model integration,

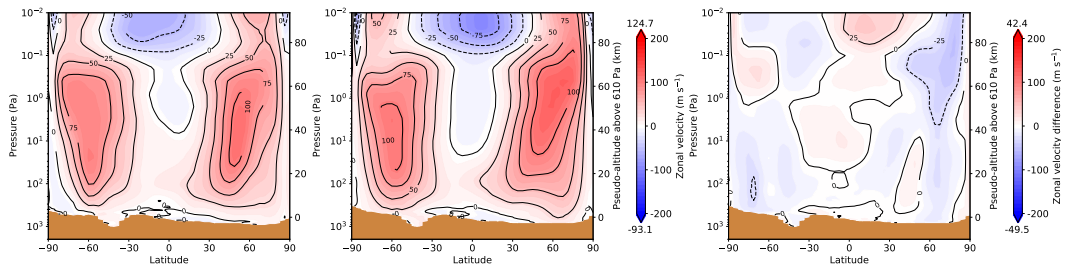
(a) 3 PM LMST, before the storm (sols 380–384).



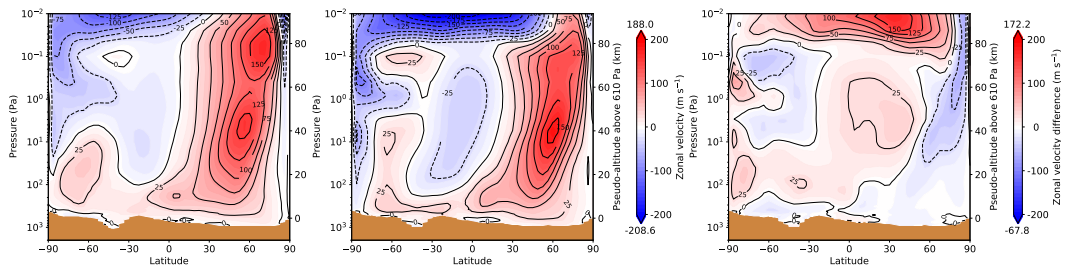
(b) 3 PM LMST, at the peak of the storm (sols 416–420).



(c) 3 AM LMST, before the storm (sols 380–384).



(d) 3 AM LMST, at the peak of the storm (sols 416–420)



841

842 **Figure 16.** As Fig. 15, but showing the zonal velocity at 3 PM and 3 AM local mean solar

843 time. Line contours are spaced every 25 m s^{-1} .

871 further constraining the atmospheric structure at local times that were not observed. For
 872 sols 380–384 the bulk of the TIRVIM observations are at 1–2 AM / PM, and for sols 416–
 873 420 they are at 5–6 AM / PM (Fig. 2c). First we look at the diurnal cycle in two dif-
 874 ferent ways (Figs 17–18) and then analyse the migrating Sun-synchronous tide modes
 875 (Figs 19–20). Note that where MCS observations are not part of the analysis, we do our
 876 computations in LTST, as that is more dynamically relevant than LMST, and it is the
 877 local time used in the GCM.

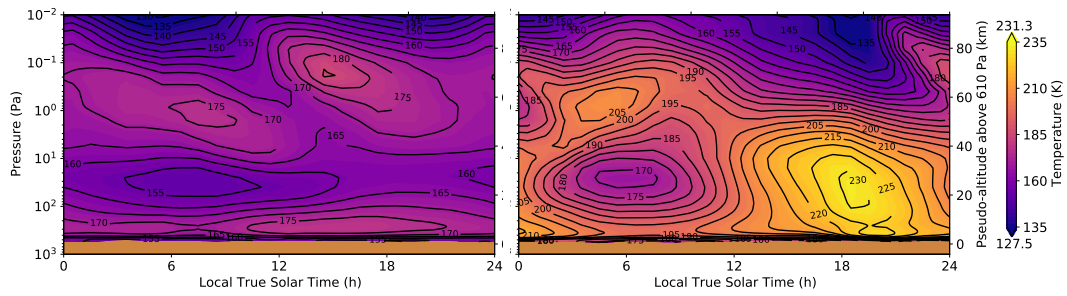
878 Figure 17 shows the temperature diurnal cycle averaged over two different latitude
 879 bands before and at the peak of the global dust storm. We interpolated the analysis and
 880 the GCM ensemble to fixed pressure levels, and then interpolated to each local time in
 881 1-hour steps. We then averaged the temperature at each pressure and local time over
 882 latitude and sol ranges. This was done at 60–65° S (to match Fig. 4.13 of M. D. Smith
 883 et al. (2017)), and over a latitude band at the equator.

892 There is a significant difference between the diurnal cycle before and at the peak
 893 of the storm. At 60–65° S (Fig. 17 a-b) the range of temperatures at all altitudes sig-
 894 nificantly increases with the peak to peak diurnal variability increasing from about 10
 895 K at 30 Pa before the storm to about 60 K at the peak of the storm. In the middle at-
 896 mosphere (0.1–1 Pa), the 12-hour oscillation is overwhelmed by a 24-hour oscillation that
 897 peaks between 3-6 AM.

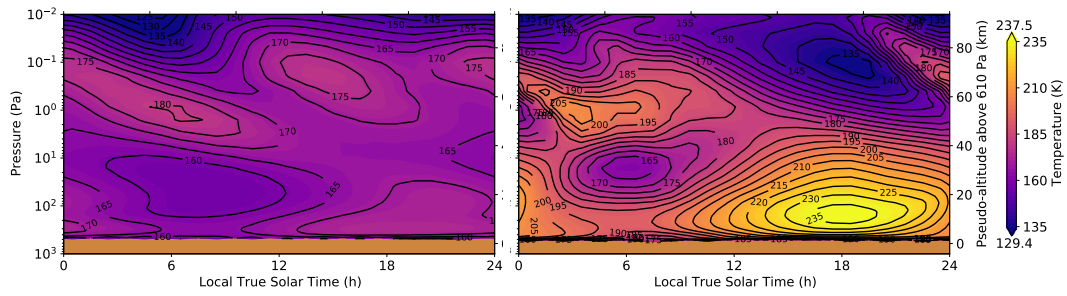
898 The main difference between the analysis and GCM ensemble is in the lower at-
 899 mosphere temperature maximum around 6 PM. In the analysis this is concentrated in
 900 the local time direction between 6-8 PM, and is extended in the pressure direction up
 901 to about 10 Pa. In the GCM ensemble the peak is concentrated closer to the ground,
 902 up to about 100 Pa, but spans a longer range of local times. The peak during the night
 903 time in the lower atmosphere is quite similar in the analysis and the GCM ensemble.

904 At the equator (Fig. 17 c-d) we also see the 12-hour oscillation in the upper part
 905 of the domain weaken. In the analysis the peak around 12 PM between 0.1–1 Pa almost
 906 completely disappears during the peak of the storm, while in the GCM ensemble it is
 907 still distinguishable. Lower in the atmosphere the 12-hour oscillation is strengthened both
 908 in the the analysis and the GCM ensemble at the peak of the storm, and the 24-hour
 909 oscillation is less clear, although in the analysis the 12-hour oscillation is dominant while
 910 in the GCM ensemble the 24-hour oscillation is dominant.

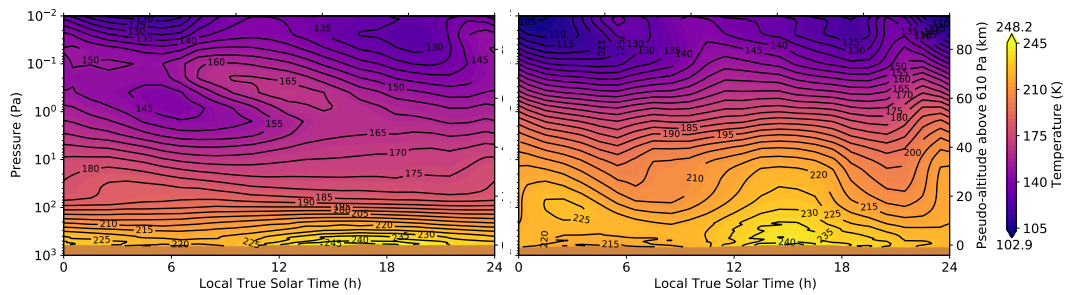
(a) TuTD-CuD analysis averaged over 60–65°S.



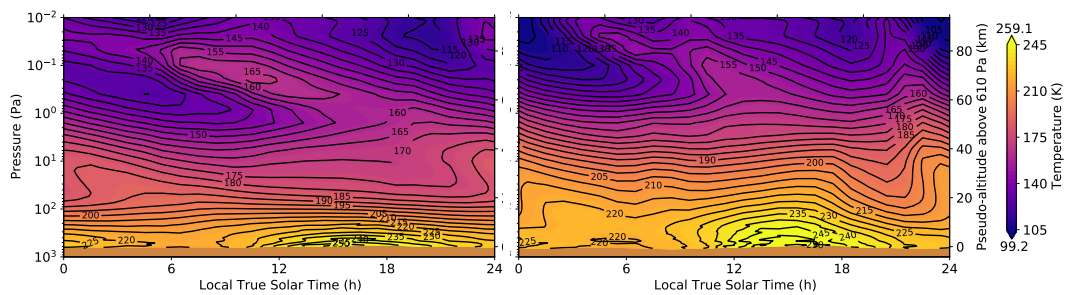
(b) GCM ensemble averaged over 60–65°S.



(c) TuTD-CuD analysis averaged over 10°S–10°N.



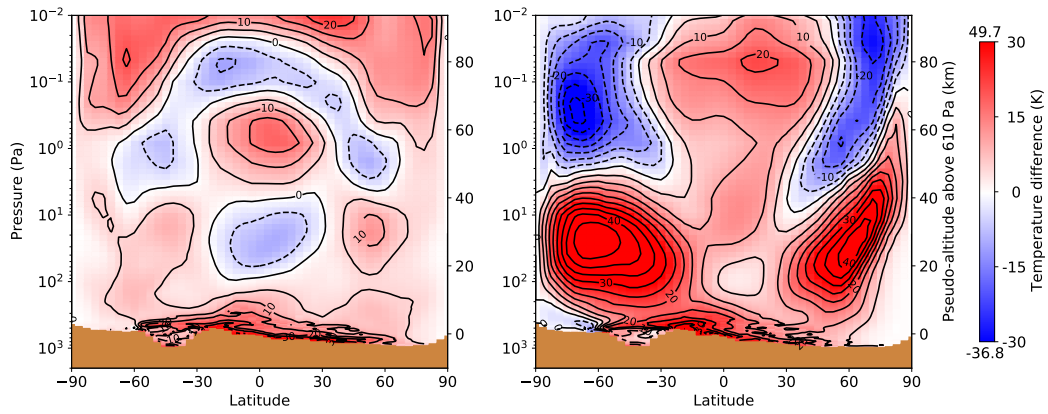
(d) GCM ensemble averaged over 10°S–10°N.



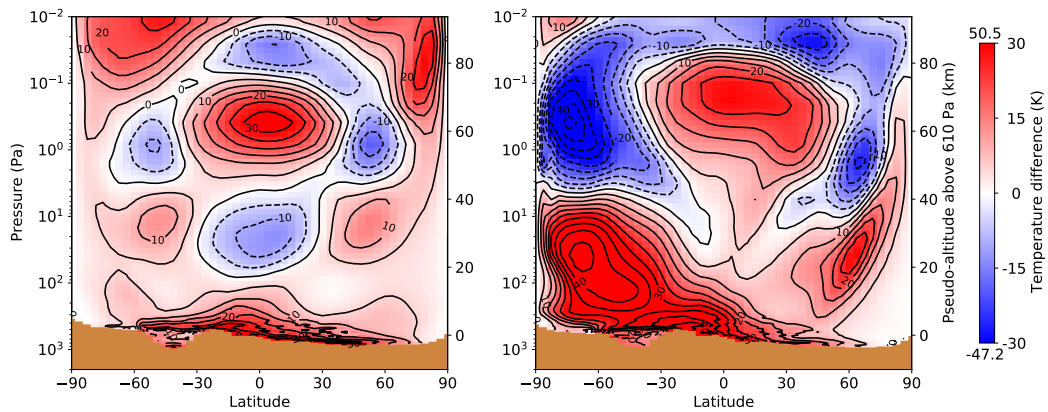
884

885 **Figure 17.** Vertical temperature structure against local true solar time. The left column
 886 shows a mean over MY34 sols 380–384 (before the storm) and the right column shows sols 416–
 887 420 (at the peak of the storm). The colour scale is the same for both columns, but note the
 888 colour scales for (a–b) and (c–d) are slightly different. From top: (a) TuTD-CuD analysis aver-
 889 aged over 60–65°S; (b) GCM ensemble for 60–65°S; (c) TuTD-CuD analysis for 10°S–10°N; (d)
 890 GCM ensemble for 10°S–10°N. Numbers above and below the colour bars indicate the minimum
 891 and maximum temperature for sols 416–420, and line contours are every 5 K.

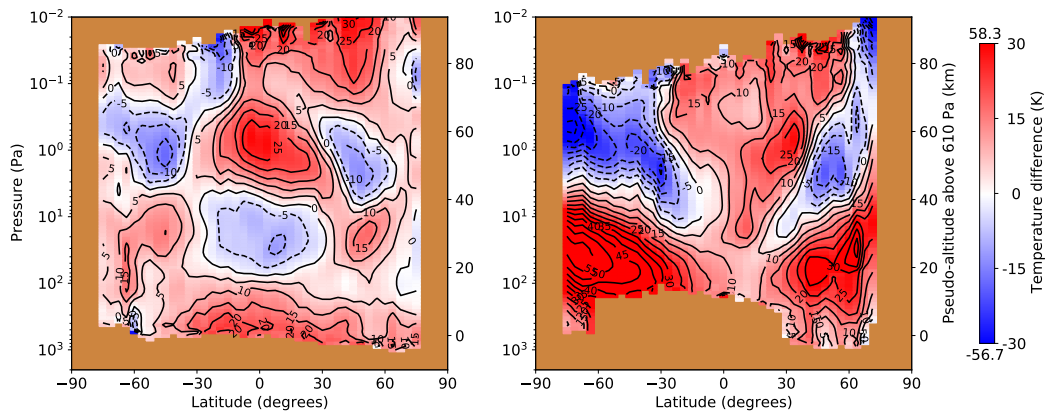
(a) TuTD-CuD analysis.



(b) GCM ensemble.



(c) MCS observations.



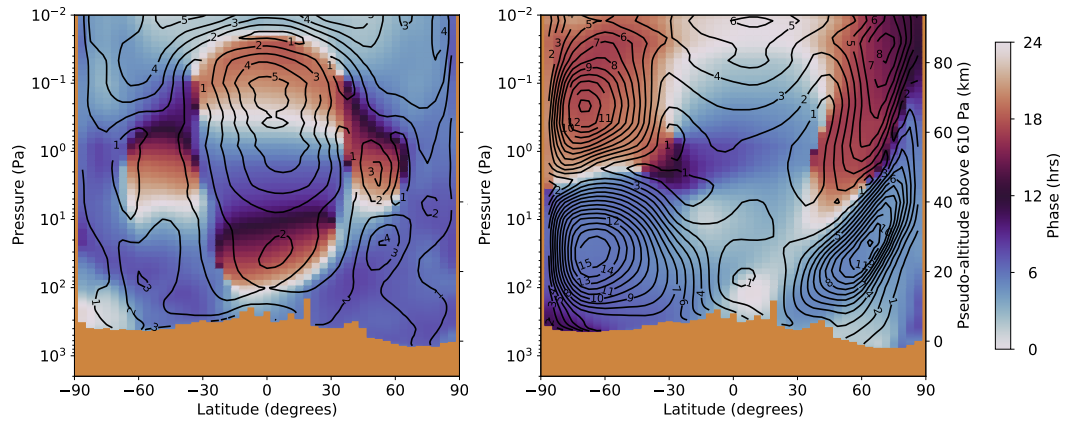
911 **Figure 18.** Difference between 3 PM and 3 AM temperatures (local mean solar time) for (a)
 912 TuTD-CuD analysis; (b) GCM ensemble; and (c) MCS observations (using observations between
 913 LMST 2-4 AM/PM). The left column shows sols 380–384, before the storm, and the right column
 914 shows sols 416–420, at the peak of the storm. Brown is missing data (either below the surface, or
 915 not observed). Numbers above and below the colour bars indicate the minimum and maximum
 916 values for sols 416–420. Line contours are at 5 K intervals.

917 We can compare diurnal variations in our analysis against independent MCS ob-
 918 servations by measuring the temperature differences between 3 PM and 3 AM. This com-
 919 parison can only be made between those two local times, but it is one aspect of the di-
 920 urnal cycle that can be compared directly with observations. This is shown in Fig. 18.
 921 Both the analysis and the GCM ensemble reproduce most basic aspects of the MCS ob-
 922 servations, both before the storm and at its peak. Differences between 3 PM and 3 AM
 923 temperatures are generally smaller in the analysis compared with the MCS observations.
 924 The GCM ensemble matches the MCS observations reasonably well before the storm,
 925 while at the peak of the storm it tends to overestimate differences in the upper atmo-
 926 sphere and underestimate them in the lower atmosphere. The general position of the peaks
 927 and troughs in the MCS observations are reproduced well by both the analysis and the
 928 GCM ensemble.

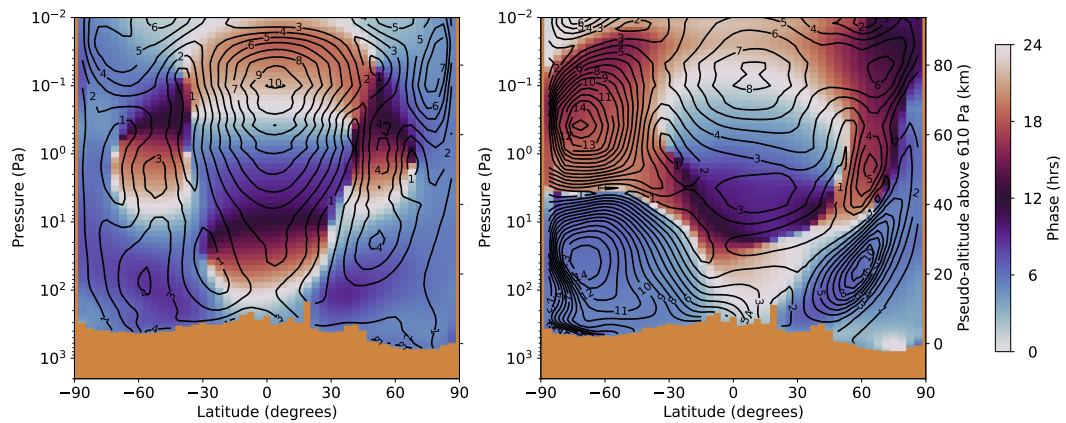
929 At the peak of the storm the quadrupole structure disappears and is replaced at
 930 mid-to-high latitudes by a strong positive anomaly in the lower atmosphere and a strong
 931 negative anomaly in the middle atmosphere. Near the equator the temperature differ-
 932 ence is small in the lower atmosphere and increases significantly in the middle atmosphere.
 933 In most areas the analysis is closer than the GCM ensemble to the MCS observations.
 934 In the northern hemisphere the analysis better matches the anomaly magnitudes and ver-
 935 tical structure. In the equatorial region the positive anomaly in the GCM ensemble is
 936 more compact than in the MCS observations. Because the MCS observations don't reach
 937 the ground, we cannot tell whether the analysis or GCM ensemble better matches re-
 938 ality near the ground in the southern hemisphere.

939 Figures 19 and 20 show the amplitude and phase of the westward-propagating Sun-
 940 synchronous migrating diurnal and semi-diurnal tides, respectively, before and at the peak
 941 of the storm. We interpolated the analysis and the GCM ensemble temperature fields
 942 to a fixed set of pressure coordinates. At each pressure and latitude point this gave tem-
 943 perature as a function of longitude and absolute time (over the 4-sol periods we have used
 944 throughout), which we then decomposed into its Fourier modes to give a two-dimensional
 945 spectrum as a function of longitudinal wavenumber and frequency. We then extracted
 946 the amplitude and phase of the period-(1 sol), wavenumber-(-1) mode, which is the westward-
 947 propagating Sun-synchronous migrating diurnal tide S_1 ; and the period-(0.5 sol), wavenumber-
 948 (-2) mode, which is the equivalent semi-diurnal tide mode S_2 . The phase is expressed
 949 in terms of hours (0–24 for the diurnal tide, 0–12 for the semi-diurnal tide).

(a) TuTD-CuD analysis.



(b) GCM ensemble.



950

951 **Figure 19.** Amplitude and phase of the sun-synchronous migrating diurnal tide S_1 . (a)
 952 TuTD-CuD analysis; (b) GCM ensemble. The left column is for sols 380–384, and the right col-
 953 umn is for sols 416–420. Filled colour contours show the phase (scaled to local true solar time
 954 between 0–24 hours), and line contours show the amplitude (spaced at 1 K intervals).

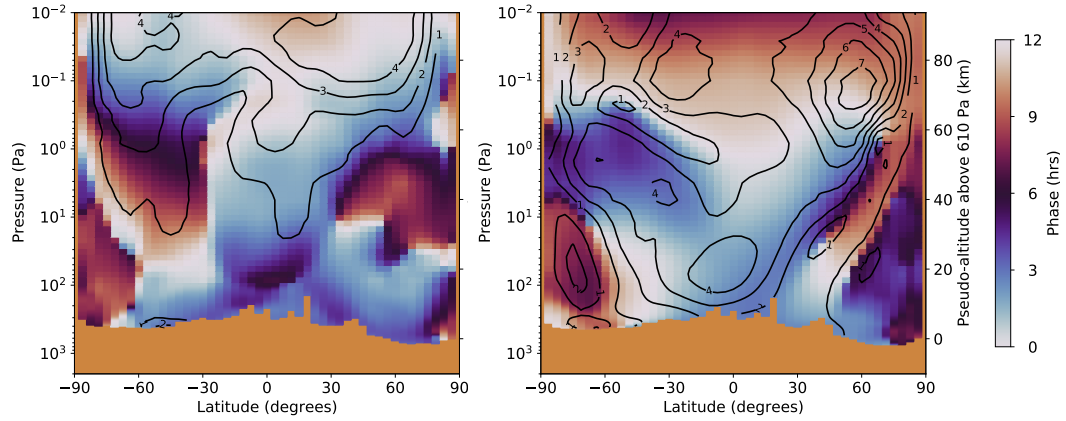
955 Before the storm the the spatial patterns of the diurnal tide (Fig. 19) are similar
 956 to those in Fig. 18. The phase data exhibit a quadrupole structure before the storm, and
 957 at the peak of the storm the vast majority of the power in the diurnal tide is at mid- and
 958 high latitudes, with a weaker diurnal tide at the equator. This is consistent with what
 959 we showed in Fig. 17. Before the storm, the diurnal tide has a maximum amplitude near
 960 the equator of about 5 K in the analysis and 10 K in the GCM ensemble (note that Fig. 18
 961 shows the difference between two local times separated by 12 hours, not a wave ampli-
 962 tude). This difference suggests that the amplitude of the diurnal tide is overestimated
 963 in the GCM ensemble.

964 At the peak of the storm the diurnal tide amplitudes in the GCM ensemble and
 965 analysis are closer, although near the equator the analysis is again weaker than the GCM
 966 ensemble, particularly at high altitudes. The analysis peaks at about 6 K right at the
 967 top of the domain, while the GCM ensemble peaks at 8K around 0.1 Pa. At mid and
 968 high latitudes, however, the analysis and the GCM ensemble agree within about 1 K.
 969 One difference between the analysis and the GCM ensemble at the peak of the storm is
 970 the phase of the diurnal tide near the equator below 1 Pa. In the analysis this phase is
 971 3-9 hours, whereas in the GCM ensemble it is 9-24 hours, a significant phase shift.

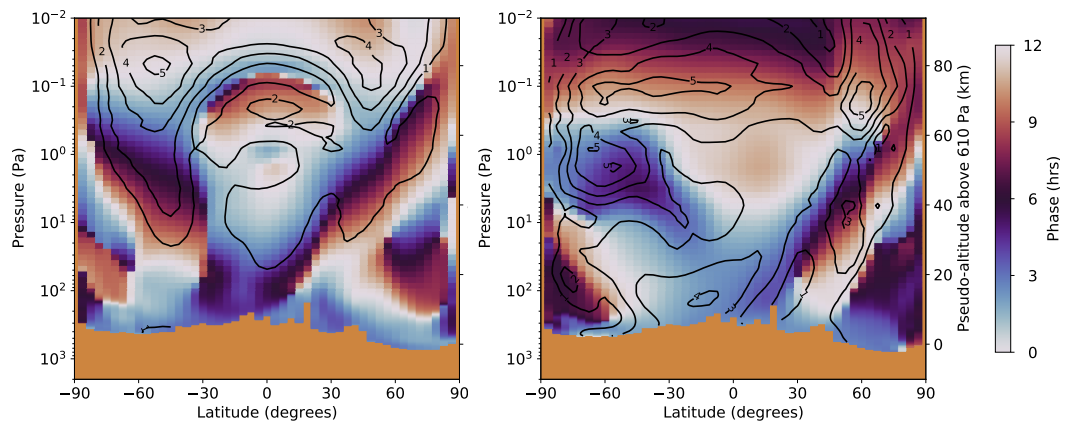
976 Before the storm the amplitude of the semi-diurnal tide (Fig. 20) is small in the
 977 lower atmosphere, increasing to about 4 K in the upper parts of the domain. The am-
 978 plitude is largely independent of latitude in the analysis, but there is a slightly larger am-
 979 plitude and one peak in each hemisphere around 50° N/S in the GCM ensemble. We may
 980 compare these figures before the storm directly with equivalent figures in Kleinböhl et
 981 al. (2013). Before the storm the general pattern is similar except for the phase near the
 982 equator. The phase generally increases downwards, starting around 0 hours at the top
 983 of the domain, and reaching 6 hours around 1 Pa. In Kleinböhl et al. (2013) this is only
 984 a weak function of latitude, but in our analysis (Fig. 20a, left) the phase of the semi-diurnal
 985 tide is 0-3 hours throughout most of the column, while in the GCM ensemble it displays
 986 the full range of phases, but goes though 6 hours much higher in the atmosphere than
 987 at higher latitudes.

988 At the peak of the storm the amplitude of the semi-diurnal tide increases at all al-
 989 titude and latitudes, particularly near the surface, which increases in the analysis from
 990 less than 1 K to about 4 K. In the upper part of the domain the amplitude increases from

(a) TuTD-CuD analysis.



(b) GCM ensemble.



972

973 **Figure 20.** Same as Fig. 19, but showing the amplitude and phase of the sun-synchronous
 974 migrating semi-diurnal tide S_2 . In this case the filled colour contours show the phase scaled to
 975 local true solar time between 0–12 hours.

991 4 K to about 7 K at 0.1 Pa around 60° N. Again the amplitudes in the analysis are gener-
 992 ally smaller than in the GCM ensemble, but not by much.

993 **5.6 Verification against Independent Surface Pressure Observations from**
 994 **Curiosity-REMS**

995 The final verification of our analysis is to compare surface pressures with high time
 996 cadence measurements taken by the Mars Science Laboratory (MSL) Curiosity rover Rover
 997 Environmental Monitoring Station (REMS). We interpolated the analysis surface pres-
 998 sure field from the GCM grid to Curiosity’s longitude and latitude before the storm (sols
 999 380–384) and at the peak of the storm (sols 416–420). We then corrected the surface pres-
 1000 sure to account for the difference in surface elevation between the GCM grid and Cu-
 1001 riosity. We used Mars Orbital Laser Altimeter (MOLA) data with 32 points per degree
 1002 (D. E. Smith et al., 2001) to obtain the altitude of the Curiosity rover, interpolating the
 1003 MOLA surface elevation to Curiosity’s location at each time it made a pressure measure-
 1004 ment during these two periods. We corrected the analysis surface pressure assuming hy-
 1005 drostatic balance and a pressure scale height based on the temperature at 1 km pseudo-
 1006 altitude (Spiga et al., 2007). The correction factor is

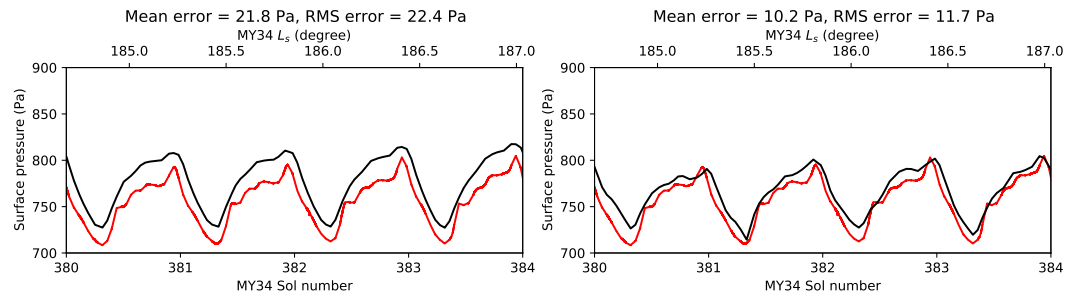
$$p_{\text{corrected}} = p_{\text{original}} \exp\left(-\frac{z_{\text{Curiosity}} - z_{\text{GCM}}}{H}\right) \quad (9)$$

1007 where $H = RT/g$ is the scale height. Because Curiosity is deep within Gale Crater, which
 1008 is not resolved by the LMD Mars GCM, the typical change in elevation from the GCM
 1009 grid to Curiosity’s location was an enormous -2,720 m, typically +160 Pa.

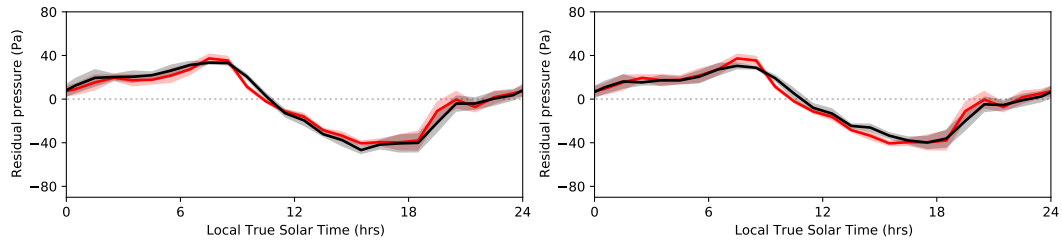
1010 There are uncertainties associated with the finite resolution of the MOLA grid (about
 1011 2 km), and we also assume that errors due to using Curiosity’s location in the rover land-
 1012 ing frame of reference (which is inclined to the direction of gravity by about 3°) are small.
 1013 However, our estimated uncertainty due to these assumptions is smaller than the uncer-
 1014 tainty in the pressure measurements themselves (5.8 Pa).

1027 Figure 21 shows surface pressure time series and diurnal cycles at the Curiosity rover
 1028 location over two 4-sol periods before and at the peak of the storm, for the GCM ensem-
 1029 ble and our best-case assimilation (TuTD-CuD). Both the analysis and the GCM ensem-
 1030 ble generally fit the measured Curiosity observations well. For the analysis, the match

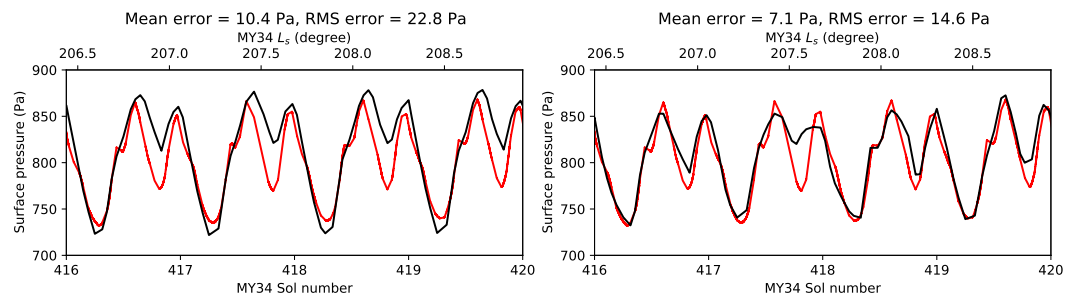
(a) Time series for sols 380–384.



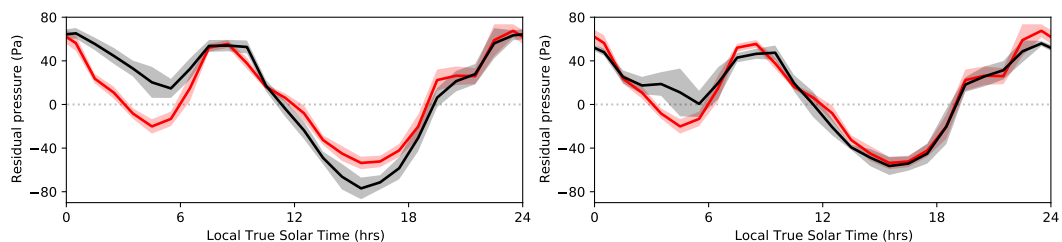
(b) Diurnal cycle for sols 380–384.



(c) Time series for sols 416–420.



(d) Diurnal cycle for sols 416–420.



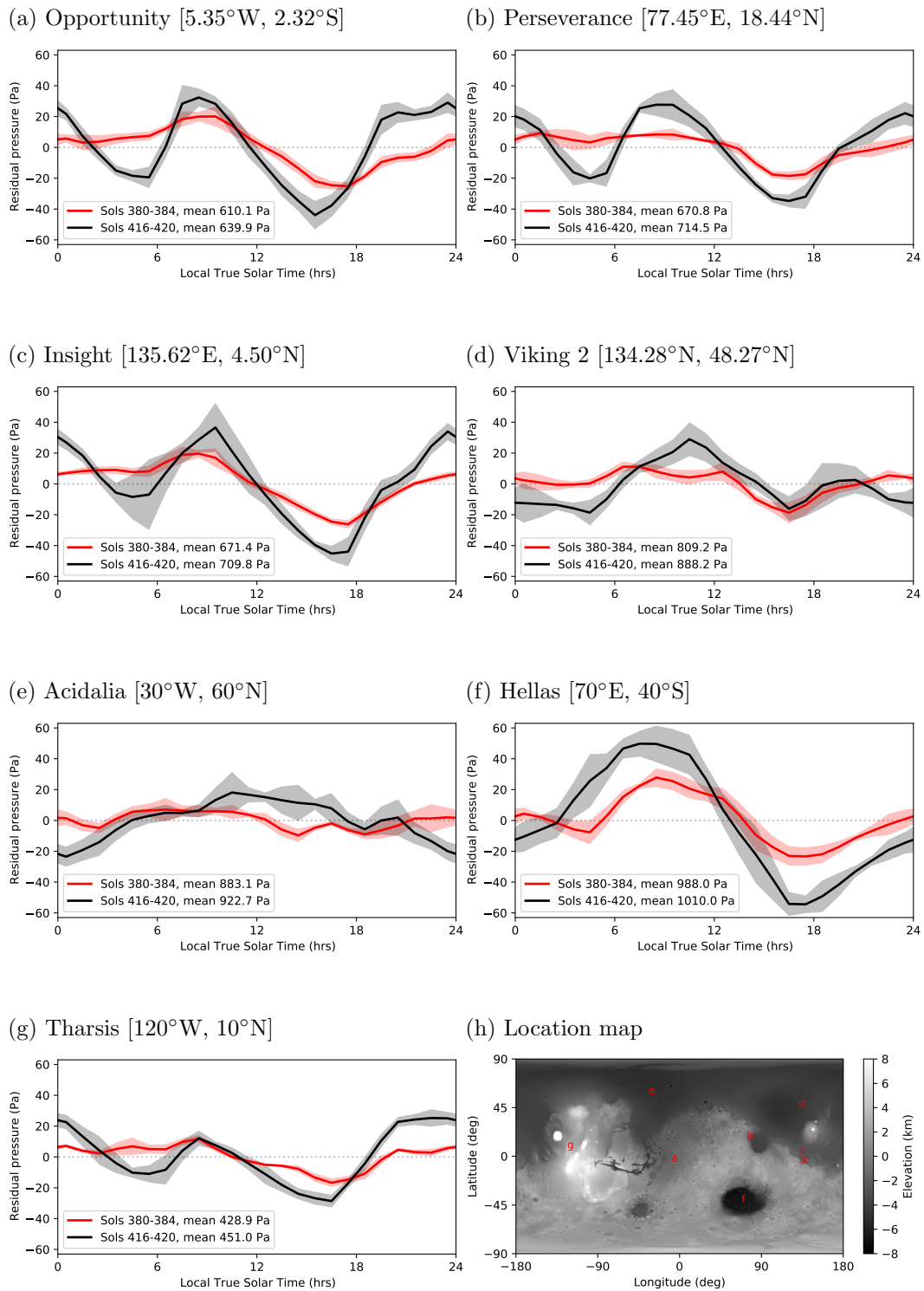
1015

1016 **Figure 21.** Surface pressure time series and diurnal cycle at the Curiosity rover location. The
 1017 left column shows the GCM ensemble, and the right column shows assimilation TuTD-CuD. In
 1018 each case the red line shows the Curiosity data, and the black line shows the analysis or GCM
 1019 ensemble. (a-b) are before the storm, sols 380–384, and (c-d) are at the peak of the storm, sols
 1020 416–420. (a) and (c) show surface pressure time series, corrected for the difference in surface
 1021 elevation between the GCM grid and the rover location on the surface, but otherwise unmodified.
 1022 (b) and (d) show the same data binned into 30-minute local true solar time bins, where the mean
 1023 in each bin is the thick line, and the shaded region shows +/- the standard deviation. Before bin-
 1024 ning the data the diurnal mean pressure was removed from the time series using a 1-sol moving
 1025 average. The mean and RMS surface pressure error (analysis/GCM – Curiosity) are listed above
 1026 each time series.

1031 to the time series at the peak of the storm is particularly good, falling within the Cu-
 1032 riosity measurement uncertainty. Given that we are down-sampling a climate resolution
 1033 model (which does not resolve Gale Crater) with grid boxes approximately 300×200
 1034 km to a single location within the crater, and then correcting the surface pressure for
 1035 a difference in surface elevation close to 3 km, a mean error close to 10 Pa is remarkable.

1036 The surface pressure rises during the storm by about 50 Pa, and the diurnal mode
 1037 becomes stronger, with the peak-to-peak diurnal variation increasing from 80 Pa to 120
 1038 Pa. The analysis is closer to the Curiosity observations than the GCM ensemble, reflect-
 1039 ing the additional constraints on surface pressure from the assimilated temperatures. The
 1040 analysis is better both in its absolute match to the Curiosity time series, and in how well
 1041 the diurnal cycle is reproduced, both before and at the peak of the storm. The GCM en-
 1042 semble overestimates the absolute pressure, and overestimates the peak-to-peak diurnal
 1043 variability, compared with the analysis and the Curiosity measurements. This is reflected
 1044 in the mean and RMS errors, which are significantly smaller for the analysis than for the
 1045 GCM ensemble: about 50% before the storm and about 70% at the peak of the storm.
 1046 Only between 2–7 AM does the analysis poorly track the diurnal surface pressure cy-
 1047 cle at the Curiosity location at the peak of the storm, where it overestimates the obser-
 1048 vations. The same is seen in the GCM ensemble, although the error is larger. This is most
 1049 likely due to mesoscale meteorology specific to Gale Crater, which is not accounted for
 1050 in the model (see e.g. Rafkin et al. (2017) for a review). Our analysis' good agreement
 1051 with observations, despite the difficulties associated with interpolating into a crater, is
 1052 encouraging for possible future assimilation of surface pressure observations.

1063 As the analysis verifies well against Curiosity observations, Fig. 22 shows surface
 1064 pressure diurnal cycles at several other locations; we may reasonably believe the anal-
 1065 ysis reproduces the surface pressure diurnal cycle there. These points of interest are: the
 1066 final location of the Opportunity rover (which succumbed to dust loading during the study
 1067 period); the landing sites for Perseverance and Insight, both currently operational; the
 1068 Viking 2 landing site, chosen as it is one of the locations used to tune the pressure cy-
 1069 cle in the LMD Mars GCM; Acidalia, where flushing dust storms often originate, and
 1070 within the northern hemisphere baroclinic zone; Hellas, where global dust events have
 1071 originated (Mulholland et al., 2013), and the location of the highest surface pressures
 1072 on Mars; and Tharsis, the location of the lowest surface pressures on Mars.



1053

1054

1055

1056

1057

1058

1059

1060

1061

1062

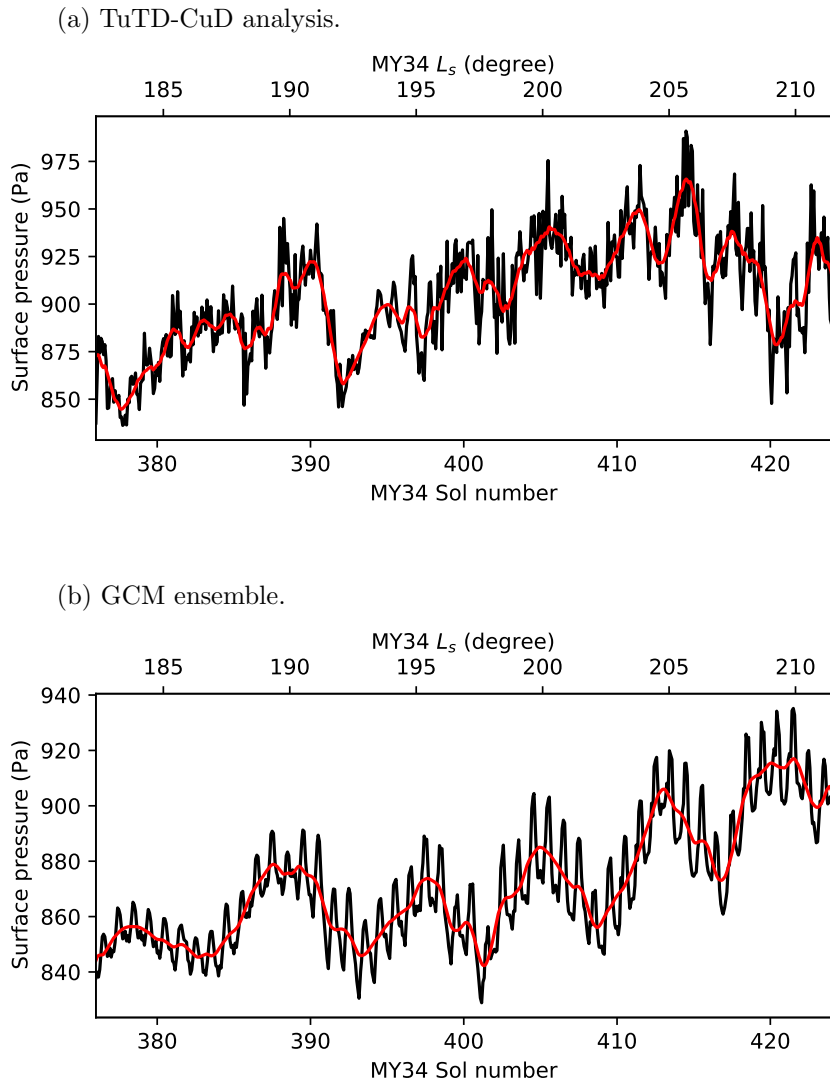
Figure 22. Surface pressure diurnal cycles at several locations during assimilation TuTD-CuD. Red lines are from the period before the storm (sols 380–384) and black lines are from the peak of the storm (sols 416–420). The diurnal cycle plots are otherwise the same as in Fig. 21. The mean pressure over each 4-sol period is given in each legend. The locations are (a) Opportunity final location; (b) Perseverance landing site; (c) Insight landing site; (d) Viking 2 landing site (these locations were all obtained using NASA GISS’ Mars24 software); (e) Acidalia; (f) Hellas; and (g) Tharsis. (h) shows these seven locations with an additional X at the Curiosity landing site, for reference [137.44°E, 4.59°S]; topography is from MOLA (D. E. Smith et al., 2001).

1073 In all cases the surface pressure increases from before the storm to the peak of the
 1074 storm, ranging from about +20 Pa in Hellas to about +80 Pa at the Viking 2 landing
 1075 site. The magnitude of the diurnal mode also increases at the peak of the storm, and at
 1076 most locations the magnitude of the semi-diurnal mode also increases. The day-to-day
 1077 variability (quantified by the width of the shaded region accompanying each line) is no-
 1078 ticeably higher in the analysis than in the GCM ensemble (an equivalent figure to Fig. 22
 1079 for the GCM ensemble is included as Supporting Fig. S3).

1080 The mid-afternoon low at each location is typically weaker in the analysis than in
 1081 the GCM ensemble, as it was at the Curiosity location, and the mean pressure differ-
 1082 ence between the analysis and the GCM ensemble is typically 5–15 Pa. One exception
 1083 is Acidalia (Fig. 22e), where the analysis pressure is about 30 Pa higher than the GCM
 1084 ensemble, and the diurnal cycle is qualitatively different. The diurnal mode is relatively
 1085 weak compared with other locations, about 20 Pa peak-to-peak before the storm, and
 1086 40 Pa at the peak of the storm, and the phase of the diurnal pressure mode is shifted
 1087 in the analysis by about 3–6 hours relative to the GCM ensemble. Figure 23 shows the
 1088 pressure time series in both cases for the whole assimilation period. In the GCM ensem-
 1089 ble the baroclinic waves are clear, with a period of about 10 sols. In the analysis, how-
 1090 ever, while some signatures of the baroclinic wave remain (peaks at sols 390 and 405, for
 1091 example), there are shorter-period oscillations that do not appear in the GCM ensem-
 1092 ble. Variability on diurnal timescales is also noisier, better reflecting real conditions; we
 1093 saw this in the Curiosity dataset, where the day-to-day variability in the analysis diur-
 1094 nal cycle (Fig. 21d, right) is similar to the day-to-day variability in the Curiosity obser-
 1095 vations.

1100 **6 Summary and Conclusion**

1101 In this paper we have assimilated observations from ACS-TIRVIM, the thermal in-
 1102 frared spectrometer on board ExoMars Trace Gas Orbiter, into the LMD Mars Global
 1103 Climate Model during the onset and peak of the MY34 global dust storm. We assim-
 1104 ilated both temperature profiles and dust column optical depth measurements using the
 1105 LETKF in three configurations, with our nominal case TuTD-CuD assimilating temper-
 1106 ature profiles to update temperature and dust profiles, followed by column dust optical
 1107 depths to update the total dust column abundance. We ran an ensemble of GCM sim-
 1108 ulations alongside the assimilations, constrained by MY34 MCS dust observations. Our



1096 **Figure 23.** Surface pressure time series in Acidalia [30°W, 60°N] throughout the full as-
 1097 simulation period for (a) analysis TuTD-CuD and (b) the GCM ensemble. Black shows the
 1098 instantaneous surface pressure (every two hours) and red shows a running diurnal mean. Note
 1099 the different y -axis scales.

1109 other improvement over previous work was to compare forecast with observations using
 1110 the retrieval averaging kernels. This allowed a like-by-like comparison to be made be-
 1111 tween forecast and observation, and by application of a filter to the sum over the aver-
 1112 aging kernels at each vertical level, we ensured that information assimilated from the re-
 1113 trievals was strongly weighted towards information from the real atmospheric state rather
 1114 than the retrieval prior. This is important for assimilating such temperature profiles de-
 1115 rived from nadir-viewing geometry, which have rather coarse vertical resolution.

1116 Our nominal assimilation verified well against in-sample observations from TIRVIM,
 1117 and was at least as good as the GCM ensemble when verified against independent MCS
 1118 observations. This was encouraging because the GCM ensemble was itself constrained
 1119 by dust opacities measured by MCS. There was a significant reduction in RMS error when
 1120 verifying temperatures against MCS once we assimilated column dust optical depth ob-
 1121 servations.

1122 The atmospheric temperature structure at 30 Pa followed the MCS observations
 1123 closely in our TuTD-CuD case, and was noticeably closer to the MCS values than the
 1124 GCM ensemble (Fig. 11 c-e). While the 3PM analysis dust density-scaled opacity did
 1125 not reach the maximum observed by MCS at the peak of the storm, there was partic-
 1126 ularly good correspondence with MCS at 3AM. The analysis reproduced detached dust
 1127 layers, while the GCM ensemble did not, and the latter also overestimated the dust DSO
 1128 at the storm’s peak (Fig. 14). The assimilation retrieved the (unobserved) atmospheric
 1129 wind structure before and during the storm, which showed that the meridional overturn-
 1130 ing circulation significantly strengthened during the storm (Fig. 15), as well as the de-
 1131 velopment of a 125 m s^{-1} asymmetry in the midlatitude jets between the two hemispheres
 1132 (Fig. 16).

1133 Assimilating observations over a range of local times gives us confidence in our re-
 1134 sults related to the diurnal cycle. The diurnal tide strengthened considerably at mid-
 1135 latitudes during the global dust storm, at the expense of the diurnal tide near the equa-
 1136 tor (Fig. 19), and the semi-diurnal tide increased in strength everywhere, particularly
 1137 in the lower atmosphere (Figs 17c, 20). The diurnal cycle was also verified against in-
 1138 dependent Curiosity-REMS surface pressure measurements. Despite the large difference
 1139 in surface elevation between the GCM topography and the rover, the analysis surface

1140 pressures at the peak of the storm accurately reproduced the diurnal cycle measured by
1141 Curiosity (Fig. 21).

1142 Assimilation during the onset of a Mars global dust storm is a particularly hard
1143 stress test of our assimilation scheme, so it is encouraging that we are able to reproduce
1144 many features of independent observational datasets. But one shortcoming of the assim-
1145 ilation has been made particularly clear when the total amount of dust in the atmosphere
1146 changes quickly over time. There is a lag in changes to the global amount of dust (and
1147 hence the onset of the global dust storm) by 6–8 sols (Fig. 12c). This is also present in
1148 temperature, but it is only 1–2 sols in that case. We expect this is because not all model
1149 grid points are updated during each assimilation cycle. If only a small fraction of grid
1150 points are updated at any one time, trends in the total amount of dust and heat in the
1151 analysis will lag behind such trends in the real atmosphere, because in reality the dust
1152 and heat content change at both unobserved and observed locations at the same time.
1153 This is a fundamental limitation in the observations, particularly when assimilating data
1154 from a single polar-orbiting satellite where the number of model grid columns updated
1155 in a single 2-hour assimilation cycle is typically 20–30% of the total.

1156 This could be overcome in a number of ways. First, by simultaneously assimilating
1157 more observations at different locations, such as by assimilating MCS in addition to
1158 ACS observations. Second, by increasing the length of the assimilation window to include
1159 more observations, but this increases the forecast time. Third, by assimilating single ob-
1160 servations multiple times, as in Analysis Correction (Lorenz et al., 1991). Fourth, one
1161 could use a longer assimilation window for dust than for temperature. Finally, one could
1162 update non-observed locations in some ad hoc way, for example by assuming that dust
1163 lingers in place over some timescale, or by assuming some spatial distribution of dust link-
1164 ing observed locations. This is already done in a sophisticated way for the GCM dust
1165 scenarios, which are complete maps but which use incomplete observations with gaps filled
1166 in by kriging (Montabone et al., 2015). Only the first of these options avoids some ad
1167 hoc intervention or a substantial change to the assimilation method, so the next step will
1168 be to focus on joint assimilation of multiple instruments' observations.

1169 **Acknowledgments**

1170 ExoMars is a space mission of ESA and Roscosmos. The ACS experiment is led
 1171 by IKI, the Space Research Institute in Moscow, Russia, assisted by LATMOS in France.
 1172 This work, exploiting ACS/TIRVIM data, acknowledges funding by CNES. The science
 1173 operations of ACS are funded by Roscosmos and ESA. The ACS/TIRVIM team at IKI
 1174 acknowledges the subsidy of the Ministry of Science and Higher Education of Russia.

1175 RMBY acknowledges funding from UAE University grants G00003322 and G00003407.
 1176 Supercomputing resources were provided by UAE University High Performance Com-
 1177 puting, with technical support from Anil Thomas and Asma Alneyadi, and at LMD by
 1178 the IPSL mesocentre. The authors thank Luca Montabone for access to processed ver-
 1179 sions of Mars Climate Sounder temperature and dust observations, and Thomas Navarro
 1180 and Claus Gebhardt for useful discussions.

1181 **Data Availability Statement**

1182 The processed TIRVIM data and retrievals used in this paper are available in NetCDF
 1183 format on the Institut Pierre Simon Laplace (IPSL) data server, see:
 1184 <https://doi.org/10.14768/ab765eba-0c1d-47b6-97d6-6390c63f0197>. The authors acknowl-
 1185 edge Sandrine Guerlet and the ACS/TGO team for supplying the data and the data cen-
 1186 ter ESPRI/IPSL for their help in accessing the data. The observations use version 7 of
 1187 the retrieval algorithm, and version 4 of the IKI calibration algorithm. Mars Climate Sounder
 1188 observations are available on the NASA Planetary Data System Atmospheres node at
 1189 [https://pds-atmospheres.nmsu.edu/data_and_services/atmospheres_data/](https://pds-atmospheres.nmsu.edu/data_and_services/atmospheres_data/MARS/mars_reconnaissance_orbiter.html)
 1190 [mars_reconnaissance_orbiter.html](https://pds-atmospheres.nmsu.edu/data_and_services/atmospheres_data/MARS/mars_reconnaissance_orbiter.html). Mars Science Laboratory Curiosity rover Rover En-
 1191 vironmental Monitoring Station observations are available on the NASA Planetary Data
 1192 System Atmospheres node at [https://atmos.nmsu.edu/data_and_services/atmospheres_data/](https://atmos.nmsu.edu/data_and_services/atmospheres_data/MARS/curiosity/rem.html)
 1193 [MARS/curiosity/rem.html](https://atmos.nmsu.edu/data_and_services/atmospheres_data/MARS/curiosity/rem.html). MOLA topography data are available on the NASA Plan-
 1194 etary Data System Geosciences node at <https://pds-geosciences.wustl.edu/missions/mgs/megdr.html>.
 1195 The LMD Mars GCM is available from <http://www-mars.lmd.jussieu.fr/>; we used GCM
 1196 subversion revision r2533. The assimilation uses letkfmars git revision
 1197 [f28935f6472a280394a98ac64f99213d1cbb2e01](https://github.com/letkfmars/letkfmars). Mars24 is available from NASA GISS at
 1198 <https://www.giss.nasa.gov/tools/mars24/>. The data produced by our assimilation will
 1199 be deposited after review using Zenodo or an equivalent repository.

1200 **References**

- 1201 Ahn, S., & Fessler, J. A. (2003). *Standard Errors of Mean, Variance, and Standard*
 1202 *Deviation Estimators* (Technical Report 413 No. 48109-2122). Comm. and
 1203 Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI.
- 1204 Amiri, S., Brain, D., Sharaf, O., Withnell, P., McGrath, M., Alloghani, M., ...
 1205 Yousuf, M. (2022). The Emirates Mars Mission. *Space Science Reviews*, 218,
 1206 4. doi: 10.1007/s11214-021-00868-x
- 1207 Capderou, M., & Forget, F. (2004). Optimal orbits for Mars atmosphere remote
 1208 sensing. *Planetary and Space Science*, 52, 789–798. doi: 10.1016/j.pss.2004.03
 1209 .006
- 1210 Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and
 1211 practical implementation. *Ocean Dynamics*, 53(4), 343–367. doi: 10.1007/
 1212 s10236-003-0036-9
- 1213 Forget, F., Hourdin, F., Fournier, R., Hourdin, C., Talagrand, O., Collins, M., ...
 1214 Huot, J.-P. (1999). Improved general circulation models of the Martian at-
 1215 mosphere from the surface to above 80 km. *Journal of Geophysical Research*,
 1216 104(E10), 24155–24175. doi: 10.1029/1999JE001025
- 1217 Giuranna, M., Wolkenberg, P., Grassi, D., Aronica, A., Aoki, S., Scaccabarozzi, D.,
 1218 ... Formisano, V. (2021). The current weather and climate of Mars: 12 years
 1219 of atmospheric monitoring by the Planetary Fourier Spectrometer on Mars
 1220 Express. *Icarus*, 353, 113406. doi: 10.1016/j.icarus.2019.113406
- 1221 Greybush, S. J., Wilson, R. J., Hoffman, R. N., Hoffman, M. J., Miyoshi, T., Ide, K.,
 1222 ... Kalnay, E. (2012). Ensemble Kalman filter data assimilation of Thermal
 1223 Emission Spectrometer temperature retrievals into a Mars GCM. *Journal of*
 1224 *Geophysical Research*, 117(11), E11008. doi: 10.1029/2012JE004097
- 1225 Guerlet, S., Ignatiev, N., Forget, F., Fouchet, T., Vlasov, P., Bergeron, G., ... Ko-
 1226 rablev, O. (2022). Thermal Structure and Aerosols in Mars’ Atmosphere
 1227 From TIRVIM/ACS Onboard the ExoMars Trace Gas Orbiter: Validation
 1228 of the Retrieval Algorithm. *Journal of Geophysical Research - Planets*, 127,
 1229 e2021JE007062. doi: 10.1029/2021JE007062
- 1230 Hamill, T. M., Whitaker, J. S., & Snyder, C. (2001). Distance-Dependent Fil-
 1231 tering of Background Error Covariance Estimates in an Ensemble Kalman
 1232 Filter. *Monthly Weather Review*, 129(11), 2776–2790. doi: 10.1175/

- 1233 1520-0493(2001)129(2776:DDFOBE)2.0.CO;2
- 1234 Heavens, N. G., Richardson, M. I., Kleinböhl, A., Kass, D. M., McCleese, D. J.,
 1235 Abdou, W., ... Wolkenberg, P. M. (2011). Vertical distribution of dust in
 1236 the Martian atmosphere during northern spring and summer: High-altitude
 1237 tropical dust maximum at northern summer solstice. *Journal of Geophysical*
 1238 *Research*, 116(1), E01007. doi: 10.1029/2010JE003692
- 1239 Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., ...
 1240 Dee, D. (2019). Global reanalysis: Goodbye ERA-Interim, hello ERA5.
 1241 *ECMWF Newsletter*, 159, 17–24. doi: 10.21957/vf291hehd7
- 1242 Hoffman, M. J., Greybush, S. J., Wilson, R. J., Gyarmati, G., Hoffman, R. N.,
 1243 Kalnay, E., ... Szunyogh, I. (2010). An Ensemble Kalman Filter Data Assim-
 1244 ilation System for the Martian Atmosphere: Implementation and Simulation
 1245 Experiments. *Icarus*, 209(2), 470–481. doi: 10.1016/j.icarus.2010.03.034
- 1246 Hoffman, R. N. (2011). A retrieval strategy for interactive ensemble data assimi-
 1247 lation. *arXiv*, 1009.1561.
- 1248 Holmes, J. A., Lewis, S. R., Patel, M. R., & Lefèvre, F. (2018). A reanalysis of
 1249 ozone on Mars from assimilation of SPICAM observations. *Icarus*, 302, 308–
 1250 318. doi: 10.1016/j.icarus.2017.11.026
- 1251 Holmes, J. A., Lewis, S. R., Patel, M. R., & Smith, M. D. (2019). Global analy-
 1252 sis and forecasts of carbon monoxide on Mars. *Icarus*, 328, 232–245. doi: 10
 1253 .1016/j.icarus.2019.03.016
- 1254 Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for
 1255 spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*,
 1256 230(1-2), 112–126. doi: 10.1016/j.physd.2006.11.008
- 1257 Kahre, M. A., Murphy, J. R., Newman, C. E., Wilson, R. J., Cantor, B. A., Lem-
 1258 mon, M. T., & Wolff, M. J. (2017). The Mars Dust Cycle. In *The Atmosphere*
 1259 *and Climate of Mars* (pp. 295–337). Cambridge University Press.
- 1260 Kalnay, E. (2003). *Atmospheric Modeling, Data Assimilation and Predictability*.
 1261 Cambridge University Press.
- 1262 Kass, D. M., Schofield, J. T., Kleinböhl, A., McCleese, D. J., Heavens, N. G.,
 1263 Shirley, J. H., & Steele, L. J. (2019). Mars Climate Sounder Observa-
 1264 tion of Mars’ 2018 Global Dust Storm. *Geophysical Research Letters*, 46,
 1265 e2019GL083931. doi: 10.1029/2019GL083931

- 1266 Kleinböhl, A., Schofield, J. T., Kass, D. M., Abdou, W. A., Backus, C. R., Sen, B.,
 1267 ... McCleese, D. J. (2009). Mars Climate Sounder limb profile retrieval of
 1268 atmospheric temperature, pressure, and dust and water ice opacity. *Journal of*
 1269 *Geophysical Research*, *114*(10), E10006. doi: 10.1029/2009JE003358
- 1270 Kleinböhl, A., Spiga, A., Kass, D. M., Shirley, J. H., Millour, E., Montabone, L.,
 1271 & Forget, F. (2020). Diurnal Variations of Dust During the 2018 Global
 1272 Dust Storm Observed by the Mars Climate Sounder. *Journal of Geophysical*
 1273 *Research - Planets*, *125*, e2019JE006115. doi: 10.1029/2019JE006115
- 1274 Kleinböhl, A., Wilson, R. J., Kass, D., Schofield, J. T., & McCleese, D. J. (2013).
 1275 The semidiurnal tide in the middle atmosphere of Mars. *Geophysical Research*
 1276 *Letters*, *40*, 1952–1959. doi: 10.1002/grl.50497
- 1277 Korablev, O., Montmessin, F., Trokhimovskiy, A., Fedorova, A. A., Shakun, A. V.,
 1278 Grigoriev, A. V., ... Zorzano, M. P. (2018). The Atmospheric Chemistry Suite
 1279 (ACS) of Three Spectrometers for the ExoMars 2016 Trace Gas Orbiter. *Space*
 1280 *Science Reviews*, *214*(1), 7. doi: 10.1007/s11214-017-0437-6
- 1281 Lewis, S. R., & Barker, P. R. (2005). Atmospheric tides in a Mars general circula-
 1282 tion model with data assimilation. *Advances in Space Research*, *36*(11), 2162–
 1283 2168. doi: 10.1016/j.asr.2005.05.122
- 1284 Lewis, S. R., & Read, P. L. (1995). An operational data assimilation scheme for the
 1285 Martian atmosphere. *Advances in Space Research*, *16*(6), 9–13. doi: 10.1016/
 1286 0273-1177(95)00244-9
- 1287 Lewis, S. R., Read, P. L., Conrath, B. J., Pearl, J. C., & Smith, M. D. (2007).
 1288 Assimilation of thermal emission spectrometer atmospheric data during the
 1289 Mars Global Surveyor aerobraking period. *Icarus*, *192*(August), 327–347. doi:
 1290 10.1016/j.icarus.2007.08.009
- 1291 Lorenc, A. C., Bell, R. S., & Macpherson, B. (1991). The Meteorological Office anal-
 1292 ysis correction data assimilation scheme. *Quarterly Journal of the Royal Mete-*
 1293 *orological Society*, *117*(497), 59–89. doi: 10.1002/qj.49711749704
- 1294 Madeleine, J.-B., Forget, F., Millour, E., Montabone, L., & Wolff, M. J. (2011).
 1295 Revisiting the radiative impact of dust on Mars using the LMD Global
 1296 Climate Model. *Journal of Geophysical Research*, *116*(11), E11010. doi:
 1297 10.1029/2011JE003855
- 1298 Montabone, L., Forget, F., Millour, E., Wilson, R. J., Lewis, S. R., Cantor, B., ...

- 1299 Wolff, M. J. (2015). Eight-year climatology of dust optical depth on Mars.
 1300 *Icarus*, 251, 65–95. doi: 10.1016/j.icarus.2014.12.034
- 1301 Montabone, L., Marsh, K., Lewis, S. R., Read, P. L., Smith, M. D., Holmes, J.,
 1302 ... Pamment, A. (2014). The Mars Analysis Correction Data Assimilation
 1303 (MACDA) Dataset V1.0. *Geoscience Data Journal*, 1(2), 129–139. doi:
 1304 10.1002/gdj3.13
- 1305 Montabone, L., Spiga, A., Kass, D. M., Kleinböhl, A., Forget, F., & Millour,
 1306 E. (2020). Martian Year 34 Column Dust Climatology from Mars Cli-
 1307 mate Sounder Observations: Reconstructed Maps and Model Simulations.
 1308 *Journal of Geophysical Research - Planets*, 125, e2019JE006111. doi:
 1309 10.1029/2019JE006111
- 1310 Mulholland, D. P., Read, P. L., & Lewis, S. R. (2013). Simulating the interan-
 1311 nual variability of major dust storms on Mars using variable lifting thresholds.
 1312 *Icarus*, 223(1), 344–358. doi: 10.1016/j.icarus.2012.12.003
- 1313 Navarro, T., Forget, F., Millour, E., & Greybush, S. J. (2014). Detection of de-
 1314 tached dust layers in the Martian atmosphere from their thermal signature
 1315 using assimilation. *Geophysical Research Letters*, 41(19), 6620–6626. doi:
 1316 10.1002/2014GL061377
- 1317 Navarro, T., Forget, F., Millour, E., Greybush, S. J., Kalnay, E., & Miyoshi, T.
 1318 (2017). The challenge of atmospheric data assimilation on Mars. *Earth and*
 1319 *Space Science*, 4, 690–722. doi: 10.1002/2017EA000274
- 1320 Pottier, A., Forget, F., Montmessin, F., Navarro, T., Spiga, A., Millour, E.,
 1321 ... Madeleine, J.-B. (2017). Unraveling the martian water cycle with
 1322 high-resolution global climate simulations. *Icarus*, 291, 82–106. doi:
 1323 10.1016/j.icarus.2017.02.016
- 1324 Rafkin, S. C. R., Spiga, A., & Michaels, T. I. (2017). Mesoscale Meteorology. In *The*
 1325 *Atmosphere and Climate of Mars* (pp. 203–228). Cambridge University Press.
- 1326 Rodgers, C. D., & Connor, B. J. (2003). Intercomparison of remote sounding in-
 1327 struments. *Journal of Geophysical Research*, 108(D3), 4116. doi: 10.1029/
 1328 2002JD002299
- 1329 Ruan, T., Young, R. M. B., Lewis, S. R., Montabone, L., Valeanu, A., & Read,
 1330 P. L. (2021). Assimilation of Both Column- and Layer-Integrated Dust Opac-
 1331 ity Observations in the Martian Atmosphere. *Earth and Space Science*, 8,

1332 e2021EA001869. doi: 10.1029/2021EA001869

1333 Smith, D. E., Zuber, M. T., Frey, H. V., Garvin, J. B., Head, J. W., Muhleman,
 1334 D. O., . . . Sun, X. (2001). Mars Orbiter Laser Altimeter: Experiment sum-
 1335 mary after the first year of global mapping of Mars. *Journal of Geophysical*
 1336 *Research*, 106(E10), 23689–23722. doi: 10.1029/2000JE001364

1337 Smith, M. D., Bougher, S. W., Encrenaz, T., Forget, F., & Kleinböhl, A. (2017).
 1338 Thermal Structure and Composition. In *The Atmosphere and Climate of Mars*
 1339 (pp. 42–75). Cambridge University Press.

1340 Spiga, A., Forget, F., Dolla, B., Vinatier, S., Melchiorri, R., Drossart, P., . . .
 1341 Gondet, B. (2007). Remote sensing of surface pressure on Mars with the
 1342 Mars Express/OMEGA spectrometer: 2. Meteorological maps. *Journal of*
 1343 *Geophysical Research: Planets*, 112(E8), E08S16. doi: 10.1029/2006JE002870

1344 Steele, L. J., Lewis, S. R., & Patel, M. R. (2014). The radiative impact of water ice
 1345 clouds from a reanalysis of Mars Climate Sounder data. *Geophysical Research*
 1346 *Letters*, 41(13), 4471–4478. doi: 10.1002/2014GL060235

1347 Steele, L. J., Lewis, S. R., Patel, M. R., Montmessin, F., Forget, F., & Smith,
 1348 M. D. (2014). The seasonal cycle of water vapour on Mars from assimi-
 1349 lation of Thermal Emission Spectrometer data. *Icarus*, 237, 97–115. doi:
 1350 10.1016/j.icarus.2014.04.017