Thermal structure and aerosols in Mars' atmosphere from TIRVIM/ACS onboard the ExoMars Trace Gas Orbiter : validation of the retrieval algorithm

S Guerlet, N. Ignatiev, Francois, Forget, T. Fouchet, P. Vlasov, G. Bergeron, R. M. B. Young, E. Millour, S. Fan, H. Tran, et al.

Thermal structure and aerosols in Mars' atmosphere from TIRVIM/ACS onboard the ExoMars Trace Gas Orbiter : validation of the retrieval algorithm

S. Guerlet[1,3*], N. Ignatiev[2], F. Forget[1], T. Fouchet[3], P. Vlasov[2], G. Bergeron[1], R. M. B. Young[1,4], E. Millour[1], S. Fan[1], H. Tran[1], A. Shakun[2], A. Grigoriev[2], A. Trokhimovskiy[2], F. Montmessin[5] and O. Korablev[2]

[1]Laboratoire de Météorologie Dynamique (LMD/IPSL), Sorbonne Université, ENS, PSL Research University, Ecole Polytechnique, Institut Polytechnique de Paris, CNRS, Paris, France

[2]Space Research Institute (IKI), Moscow, Russia

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

[4]Department of Physics & National Space Science and Technology Center, United Arab Emirates University, Al Ain, UAE

[5]LATMOS/IPSL, Guyancourt, France

Key Points:

- We exploit TIRVIM spectra to determine Martian atmospheric, surface temperature, as well as integrated opacity of dust and water ice clouds.
- Different sources of biases are investigated with the help of simulated observations at different local times, latitudes and seasons.
- Atmospheric temperatures retrieved from TIRVIM in April-May 2018 are in excellent agreement with co-located MCS observations.

*4 Place Jussieu, Paris, France

Corresponding author: Sandrine Guerlet, sandrine.guerlet@lmd.jussieu.fr

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2021JE007062.

Abstract

The Atmospheric Chemistry Suite (ACS) onboard the ExoMars Trace Gas Orbiter (TGO) monitors the Martian atmosphere through different spectral intervals in the infrared light. We present a retrieval algorithm tailored to the analysis of spectra acquired in nadir geometry by TIRVIM, the thermal infrared channel of ACS. Our algorithm simultaneously retrieves vertical profile of atmospheric temperature up to 50 km, surface temperature, and integrated optical depth of dust and water ice clouds. The specificity of the TIRVIM dataset lies in its capacity to resolve the diurnal cycle over a 54 sol period. However, it is uncertain to what extent can the desired atmospheric quantities be accurately estimated at different times of day. Here we first present an Observing System Simulation Experiment (OSSE). We produce synthetic observations at various latitudes, seasons and local times and run our retrieval algorithm on these synthetic data, to evaluate its robustness. Different sources of biases are documented, in particular regarding aerosol retrievals. Atmospheric temperature retrievals are found robust even when dust and/or water ice cloud opacities are not well estimated in our OSSE. We then apply our algorithm to TIRVIM observations in April-May, 2018 and perform a cross-validation of retrieved atmospheric temperature and dust integrated opacity by comparisons with thousands of co-located Mars Climate Sounder (MCS) retrievals. Most differences between TIRVIM and MCS atmospheric temperatures can be attributed to differences in vertical sensitivity. Daytime dust opacities agree well with each other, while biases are found in nighttime dust opacity retrieved from TIRVIM at this season.

Plain Language Summary

The Martian surface and atmosphere undergo strong variations in temperature and amount of aerosols (dust or water ice cloud particles). Our knowledge on their variations at diurnal scale is however limited, due to lack of appropriate observations. We present a method to analyze thermal emission spectra of Mars' surface and atmosphere recorded by TIRVIM, a spectrometer onboard the ExoMars Trace Gas Orbiter. We have developed a program to derive surface and atmospheric temperatures from these spectra, along with an estimation of the amount of aerosols. The specificity of the TIRVIM dataset is its capacity to resolve the diurnal cycle over a 54 sol period. However, atmospheric quantities cannot be accurately estimated at all times of day. One of the goals of our paper is to assess the robustness of our algorithm with the help of simulated observations. The

retrieval of aerosol opacity is assessed to be challenging at some times of day, but atmospheric temperature is well determined. We have then applied our algorithm to tens of thousands of TIRVIM observations obtained in April-May 2018 and showed that our derived atmospheric temperatures compare very well with independent measurements obtained from the Mars Climate Sounder, reinforcing our confidence in our method.

1 Introduction

The ESA-Roscosmos ExoMars Trace Gas Orbiter (TGO) was launched on March 14th, 2016 and was successfully inserted into Mars' orbit in November, 2016. After several orbit manoeuvres including a 12-month aerobraking phase it reached its final, 400-km altitude and near-circular orbit on April 13th, 2018. Onboard TGO, the Atmospheric Chemistry Suite (ACS) comprises three spectrometers, each tailored for specific scientific goals (Korablev et al., 2018). We focus here on the thermal infrared spectrometer, named TIRVIM (Thermal InfraRed channel in honor of professor Vassilii Ivanovich Moroz), which was operational from April, 2018 until December, 2019, hence almost a martian year (between mid- martian years 34 and 35). Through nadir-viewing soundings, its main goal is to monitor the atmospheric temperature, surface temperature and integrated aerosol content – dust and water ice clouds – at a great variety of local times. Indeed, TGO's orbit is designed in such a way that nadir observations sample the full diurnal cycle in 54 sols, at all latitudes between 74°N and 74°S. Hence, it is able to capture both seasonal and diurnal variations of these climatological variables and complements other still-operating thermal infrared sounders, such as the Mars Climate Sounder (MCS) onboard NASA's Mars Reconnaissance Orbiter, the Planetary Fourier Spectrometer (PFS) onboard ESA's Mars Express and the Thermal Emission Imaging System (THEMIS) onboard Mars Odyssey.

MCS is a radiometer operating in limb-viewing geometry, allowing the retrieval of vertical profiles of atmospheric temperature from 5 to 80 km as well as dust and water ice vertical profiles, with a vertical resolution of typically 5 km (Kleinböhl et al., 2009). Being on a Sun-synchronous polar orbit, MCS mostly acquires data at local times 3:00 and 15:00 (± 1.5 hours when cross-track observations are performed). PFS is a thermal infrared spectrometer operating in nadir-viewing geometry providing temperature profiles in the range 5–50 km with a vertical resolution of 10 km (Grassi et al., 2005) along with surface temperature and the integrated content of dust and water ice clouds. The

same quantities can be retrieved from TIRVIM observations. The Mars Express orbiter is such that PFS observations sample various local times but with a longer revisit time compared to TIRVIM (150 sols for PFS versus 54 sols for TIRVIM), and a sparser spatial coverage due to the elliptical orbit of Mars Express. Finally, the THEMIS instrument comprises several cameras that image Mars in the visible and thermal infrared. It is sensitive to atmospheric temperature in a broad altitude range centered on 50 Pa, and allows for the retrieval of integrated dust and water ice opacity (Smith, 2009). It is on a Sun-synchronous orbit, but the local time coverage has varied over the past 20 years; it sampled local times 7AM–7PM during Martian Year 34. Another important instrument to add to this list was the Thermal Emission Spectrometer (TES) mounted on the Mars Global Surveyor that operated between 1997 and 2006. It was a spectrometer rather similar to PFS with a coarser spectral resolution of either 5 or 10 cm^{-1}, sensitive to the temperature in the range 5–35 km, dust and water ice cloud opacities in nadir mode (Conrath et al., 2000; Smith et al., 2000). It performed measurements at local times near 2 AM and 2 PM and also operated systematically in limb geometry, allowing the retrieval of atmospheric temperature up to 65 km.

Deriving those atmospheric quantities is of high interest to study the Martian climate at various spatial and temporal scales, from diurnal variations to inter-annual variations. These measurements have helped to broadly characterize the vertical and meridional structure of atmospheric temperature, dust loading and water ice clouds (e.g., Smith et al., 2001; McCleese et al., 2010; Giuranna et al., 2021), which in turn bring insights onto Mars atmospheric dynamics. In particular, the TES and MCS retrieval products have been used in several data assimilation studies (e.g., Steele et al., 2014; Navarro et al., 2014; Greybush et al., 2019). More recently, these thermal infrared measurements have brought new insights on the diurnal variability of dust and water ice clouds (Kleinböhl et al., 2020; Smith, 2019a; Wolkenberg & Giuranna, 2021; Giuranna et al., 2021). However, developing a retrieval algorithm that performs well at all conditions (dust load, local time, etc) is a challenge. Particular care has to be taken regarding the reliability of aerosol retrievals from nadir sounders when the contrast between the surface temperature and atmospheric layer where the aerosols lie is low, as already raised by several previous studies (Pankine et al., 2013; Smith, 2019a, 2019b; Wolkenberg & Giuranna, 2021).

The objective of this paper is to document a retrieval algorithm developed at Laboratoire de Météorologie Dynamique (LMD) for the analysis of TIRVIM observations,

to discuss some challenges identified and validate our retrievals against independent observations from MCS. We first detail characteristics of TIRVIM observations in section 2. We describe our algorithm in section 3 and apply it to synthetic measurements generated for a great variety of scenes (latitudes, local time, surface temperature, aerosol content, topography) in section 4. Challenges arise due to the degeneracy in the inverse problem, especially regarding the combined retrieval of surface temperature, dust and water ice optical depth. This algorithm is then applied to the first 45 sols of TIRVIM observations from mid-March to end of April 2018, as described in section 5. The retrieved temperature profiles are compared to co-located MCS observations near 3 AM and 3 PM to perform a cross-validation of our retrievals and evaluate potential biases. A first assessment of the quality of dust retrievals is also included. Detailed discussion regarding the diurnal temperature variations derived from TIRVIM observations is deferred to another paper. We conclude on the performance of TIRVIM and on our algorithm in section 6.

2 ACS/TIRVIM nadir measurements

2.1 Instrument characteristics

TIRVIM is a double-pendulum Fourier-transform spectrometer sensitive in the spectral range 1.7 to 17 μm (600 – 5200 cm^{-1}). It operates routinely in nadir-viewing geometry and can also operate in solar occultation mode; the former type is the focus of this paper. In nadir geometry, spectra are only exploitable between 620 and 1300 cm^{-1}, as thermal emission from the surface and atmosphere quickly drops beyond 1000 cm^{-1}. In this geometry, the apodized spectral resolution is 1.2 cm^{-1} and the integration time for a single interferogram is 0.4 seconds. The individual projected field of view on Mars from TGO's circular orbit for a single observation is 14 km cross-track × 16 km along-track, accounting for the small (4 km/s) smearing due to the spacecraft's motion during acquisition. Further details on TIRVIM technical characteristics can be found in Korablev et al. (2018).

2.2 Calibration and instrumental issues

Generation of calibrated spectra from raw interferograms is done at the Space Research Institute (IKI) in Moscow, Russia. Absolute radiometric calibration was facili-

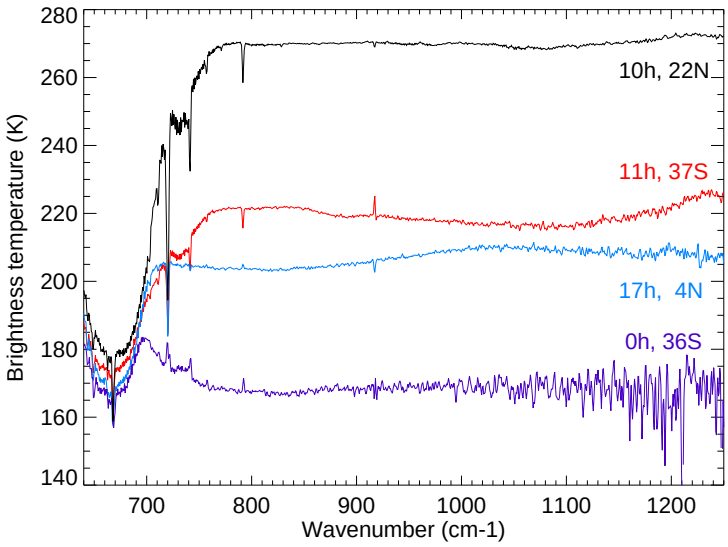

Figure 1. Examples of TIRVIM spectra acquired in March 2018 at four latitudes and local times, with different surface temperatures. They were acquired in the onboard averaging mode, where 8 interferograms were averaged by the instrument software onboard TIRVIM. Absorption by dust centered at 1090 cm^{-1} is particularly visible in the spectrum at 37°S, 11h (in red), while absorption by water ice clouds centered at 820 cm^{-1} is visible in the spectrum acquired at 4°N, 17h (in blue). An electric spike, which is an artifact, can be seen near 920 cm^{-1}.

tated by routine periodical measurements of the internal calibration black body and of the cold space. TIRVIM turned out to be an IR Fourier transform spectrometer with a source-dependent phase function. Radiometric calibration of such an instrument was considered by Revercomb et al. (1988).

Ahead of the beginning of TIRVIM science operations, the chosen strategy to relay the data was to perform an on-board averaging of 8 consecutive interferograms in order to increase the signal-to-noise ratio (SNR). This was justified by the fact that the individual interferograms are recorded at very close locations on Mars, hence surface and atmospheric variables should not vary much amongst 8 sequential acquisitions. This onboard averaging mode was used for the first five months of the mission (except for the very first six days, where individual interferograms were relayed). However, the onboard interferogram averaging appeared to be sometimes incorrect. Interferograms must be aligned in the path difference space before averaging. The onboard TIRVIM software aligned interferograms before averaging according to their maxima. Since the source was chang-

ing, and the instrument had a source dependent phase, maxima of the interferograms did not always correspond to the same path difference. In such cases, the average interferogram and therefore calibrated spectral radiance was incorrect. Those cases were flagged as poor-quality data and were excluded from our analysis. They could represent up to 20% of all measurements. Fortunately, the other ∼80% are of good quality and are exploited in this paper. The corresponding field of view of the averaged interferograms is approximately 25×105 km. To limit data loss, this mode was later abandoned in favor of recording and relaying each individual interferograms. The drawback was the reduced signal-to-noise ratio (down to 3 times) and a higher data rate from the spacecraft. Among the other known issues in the calibrated spectra is an electrical spike near 920 cm^{-1} (that is excluded from our analysis).

Examples of TIRVIM spectra acquired in March 2018 at different local times and latitudes are shown in units of brightness temperature in Figure 1. Apart from the near blackbody surface emission, these spectra are dominated by the atmospheric CO_2 absorption band centered at 667 cm^{-1} (sometimes visible in emission, when the atmosphere is warmer than the surface), a broad water ice feature centered at 820 cm^{-1} and a broad dust feature centered at 1090 cm^{-1}. We emphasize here that the left wing of the broad CO_2 absorption band is almost not captured by TIRVIM (unlike PFS). Apart from a few particular cases (mentioned later), this does not hamper the atmospheric temperature retrievals.

The Noise Equivalent Radiance (NER) of a calibrated spectrum obtained without interferogram averaging is of the order of 0.4×10^{-7} W $cm^{-2}sr^{-1}/cm^{-1}$ for most of the spectral range, and increases near the edge of our spectral domain: it reaches 1.2×10^{-7} W $cm^{-2}sr^{-1}/cm^{-1}$ at 660 cm^{-1} and 0.8×10^{-7} W $cm^{-2}sr^{-1}/cm^{-1}$ at 1300 cm^{-1}. These values are comparable to the NER of PFS spectra at 600 cm^{-1} but are up to 10 times lower than PFS NER at 1300 cm^{-1} (Giuranna et al., 2005). The resulting signal-to-noise ratio strongly depends on the Martian surface (or atmospheric) temperature and wavenumber. Figure 2 illustrates these SNR variations: for a warm surface (eg. 280K) the SNR is in the range 30–200, depending on wavenumber, while for a colder surface (eg. 190K), the SNR is typically ten times lower. An exception is for wavenumbers 650–700 cm^{-1}, where surface emission does not contribute to the measured thermal emission from space and for which the SNR depends on atmospheric temperature, which varies less dramatically than the surface temperature. In this wavenumber range, the SNR is of the order of 30–100.

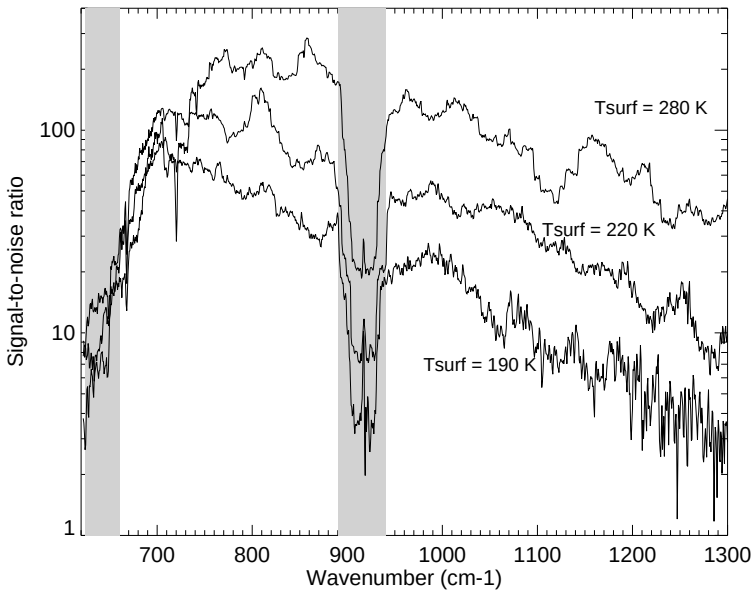

Figure 2. Signal-to-noise ratio for three individual spectra (no onboard averaging) acquired in October 2018, for which the retrieved surface temperature was 190K, 220K and 280K, as labeled. Note that the SNR is similar for the three spectra in the range 660–700 cm^{-1}, as in this spectral range, the outgoing thermal radiation depends on the atmospheric temperature and not on surface temperature. The two grey shaded areas have large noise levels (low SNR) while not being essential to our analysis and are thus excluded. One corresponds to the left edge of the spectrum (620–660 cm^{-1}), the other one to an electric spike centered at 920 cm^{-1}.

If we consider spectra resulting from the onboard averaging of 8 interferograms, the aforementioned SNR values are multiplied by $\sim$3.

In this paper, we will include analysis of spectra acquired in both modes, predominantly the averaging mode. We will see that good performances are achieved with this mode. The calibration version we use is referred to as version 4, where orbit-average black body and space interferograms are used for calibration.

2.3 Spatiotemporal coverage of TIRVIM nadir measurements

The ExoMars Trace Gas Orbiter is set on a near-circular orbit at 400-km altitude with an inclination of 74° (implying that nadir observations cannot be made at latitudes poleward of 74°). TGO executes 12 orbits per (Earth) day, hence sampling 24 different

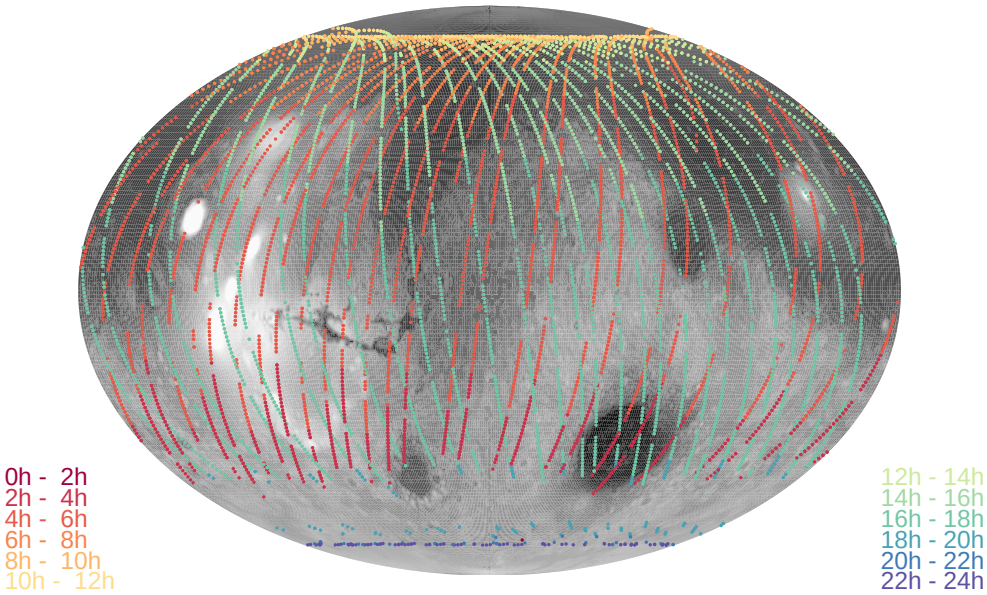

Figure 3. Coverage of TIRVIM nadir observations acquired on the 26th, 27th and 28th of March 2018, showing here only data that passed a set of initial quality filters. Different colors refer to the local time, as labeled. TGO's orbit is such that the local time of TIRVIM observations shifts by ∼13 minutes earlier every sol.

longitudes per day at a given latitude. An example of the coverage obtained after three days (36 orbits of TGO) is shown in Figure 3. At low latitudes, on this short time period, mostly two local times are sampled. However, TGO is not in a Sun-synchronous configuration: rather, the local time coverage slightly drifts earlier each day in such a way that after 54 sols (corresponding to 25°–35° of Ls, depending on season), the diurnal cycle has been evenly sampled in local times over the whole planet – providing that TIRVIM is operating continuously. The revisit time for a given (latitude, longitude, local time) targeted point on Mars is actually 108 sols ; however, if we consider an area of ∼5°×5°, a 54-sol observation period provides coverage at all local times and is relevant for studying diurnal variations (with the caveat that ∼30° of solar longitude has passed).

A Stirling cryo-cooler ensured the stability of the detector operating temperature at 65–75K. Several overheating events occurred in the first months of the mission, which required the cooler to be switched off for several weeks, which meant the absence of meaningful data. TIRVIM started its routine nadir observations on March 13, 2018 but en-

countered an issue with the stability controller of the double pendulum movement on April 28, 2018, which caused dramatic loss of the data quality. It resumed its quality observations on May 26, 2018 until an overheating event occurred on July 15, 2018. After that, overheating events became more frequent. To overcome these issues, from September, 2018 onwards, an automatic switch off of the cryo-cooler was programmed if its temperature exceeded 14°C. Another limitation stemmed from the lifetime of the Stirling cryo-cooler, which was estimated to last 10 000 hours (it actually operated 8,000 hours before failing in December, 2019). If TIRVIM had been switched ON all the time, then the cooler would have stopped functioning after 10 months. In order to mitigate this effect, a duty cycle of ∼50% was undertaken in October, 2018 at which time TIRVIM acquired data roughly 2 days out of 7, except for one 9-day long acquisition every month. This observing strategy largely prevented overheating events from occurring, and allowed TIRVIM to acquire data over almost a full martian year. In this paper, we will report in section 5 on the analysis of the first 45 days of TIRVIM data during March-April, 2018.

3 Retrieval algorithm

TIRVIM spectra contain information on the surface temperature, on the vertical profile of atmospheric temperature between a few kilometers above the surface and ∼50 km (or 2–3 Pa) and on the column integrated opacity of dust and water ice clouds. We aim at retrieving simultaneously these quantities by exploiting TIRVIM spectra between 660 and 1300 cm^{-1}. Our algorithm comprises a forward radiative transfer model used to generate synthetic observations, coupled with a constrained linear inverse model, described below.

3.1 Radiative transfer model

Our forward radiative transfer model computes the spectral radiance I_ν of the outgoing thermal emission of Mars' surface and atmosphere, neglecting scattering contributions, in the plane-parallel approximation :

$$I_\nu(\tau = 0, \mu) = \epsilon_\nu B_\nu(T_{\rm surf}) e^{-\tau_{\rm total}/\mu} + \frac{1}{\mu} \int_0^{\tau_{\rm total}} B_\nu(T(\tau')) e^{-\tau'/\mu} d\tau' \qquad (1)$$

with $\mu = 1/\cos(\theta)$ the airmass at an emission angle θ ; ϵ_ν the surface emissivity; B_ν the Planck function, τ' the partial integrated optical depth from the top of the atmo-

sphere to a given pressure level, $T_{\rm surf}$ the surface temperature and T the atmospheric temperature at a given pressure level. Our model atmosphere is discretized into 45 vertical sigma-levels, with the first level just above the surface. Hence, rather than using a fixed pressure grid, our pressure grid is tuned to local surface pressure to adapt to the topography. Radiances from TIRVIM and those calculated from equation 1 are converted in brightness temperatures $T_{B,\nu}$. Unless stated otherwise, in the following, we work in brightness temperature units and not radiances.

We use a line-by-line approach, where spectra are first computed at a high spectral resolution of 0.01 $\rm cm^{-1}$, and are then convolved at the resolution of the instrument. TIRVIM spectra are apodized with a Hamming function and we use the appropriate instrument function for this convolution. Note that the sampling is 0.645 $\rm cm^{-1}$ and the corresponding spectral resolution (full width at half maximum) is of 1.2 $\rm cm^{-1}$. We take into account opacity from atmospheric CO_2 gas, dust and water ice clouds. Line-by-line, CO_2 absorption coefficients are tabulated offline for a set of 45 reference pressures equally spaced in natural logarithm space (from 1260 Pa to 5×10^{-3} Pa) and 12 temperatures. These sets of 12 reference temperature values themselves depend on pressure and are sampled to encompass climatological conditions over a full martian year, including global dust storm conditions with warmer temperatures. For instance, near the surface (p> 300 Pa), reference temperature values range from 150K to 260K, every 10K while at 10 Pa, these levels comprise ten levels sampled from 100 to 190K (every 10K) plus two extra levels at 210K and 230K. These computations use the HITRAN 2016 spectroscopic database (Gordon et al., 2017). We tested another set of CO_2 coefficients generated from the GEISA 2015 linelist (Jacquinet-Husson et al., 2016) and found that it yielded similar spectra, as far as the CO_2 15 μm band is concerned. The temperature dependence of the line width due to pressure broadening is not taken from the HITRAN database (as it corresponds to air broadening) but instead is adapted for a CO_2 atmosphere. It is computed for each transition, as a function of the rotational quantum number J, based on values tabulated by Lamouroux et al. (2012). Furthermore, pressure broadening is not represented by a standard Lorentz function. Rather, we adopt an asymmetric sub-Lorentzian profile, which empirically takes into account the effects of collisional line mixing and the finite duration of collisions. These line shapes are represented by a Lorentz function multiplied by a factor χ that depends on the distance from the line center and was derived from experimental work relevant for the 4.3-μm CO_2 band (Perrin & Hartmann, 1989). We as-

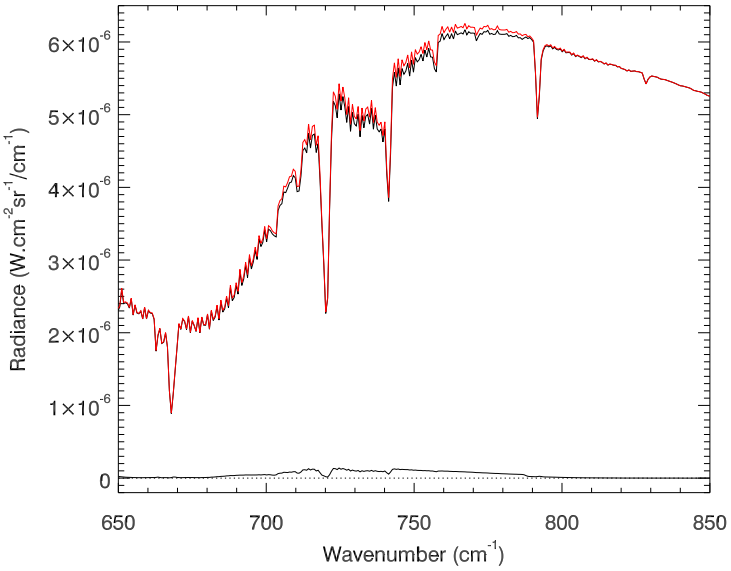

Figure 4. Synthetic TIRVIM spectra focusing here on the CO_2 band, assuming (in black) a Lorentz line profile or (in red) a sub-Lorentzian line profile to empirically take into account the effects of collisional line mixing and the finite duration of collisions, based on experimental work by Perrin and Hartmann (1989). The difference between the two spectra is plotted at the bottom.

sume that these χ factors can also be used for the 15-μm CO_2 band. The final line shape is obtained by a convolution of the sub-Lorentzian profile with a Gaussian profile (that represents Doppler broadening of the lines). An example of synthetic TIRVIM spectra is shown in Figure 4, which compares spectra computed with a Lorentz function or the sub-Lorentzian line profile. In this example, the radiance is increased by 1–2 % (hence, 1 to 2-σ above the noise level) in the range 700–780 cm^{-1} when a sub-Lorentzian line profile is adopted. This effect is noticeable, but its influence is small regarding atmospheric temperature retrievals(typically $\sim$ 1–2 K).

Extinction coefficients Q_{ext} for water ice particles are computed offline using Mie theory, assuming a log-normal size distribution, an effective radius r_{eff} of 1.45 μm, a variance of 0.1 and optical constants from Warren (1984). Particle sizes of typically 1 to 3.5 μm are most frequently observed, as determined from CRISM observations (Guzewich & Smith, 2019) or from MGS TES (Clancy et al., 2003). Assuming a single particle size for water ice clouds in our radiative transfer model should not affect our retrievals of the cloud

integrated optical depth because it is sensitive to the ratio $Q_{\text{ext}}/r_{\text{eff}}$. Indeed, the latter quantity varies by less than 10% for particles sizes in the range 1–3.5 μm. Hence, the uncertainty on total cloud opacity related to a wrong assumption of particle sizes should be on the order of 10%.

Regarding dust particles, their extinction coefficients are computed from a T-matrix code using an effective radius r_{eff} of 1.5 μm, a variance of 0.3 and optical constants derived from CRISM observations by Wolff et al. (2009), which also used a T-matrix code in the analysis. Particle effective sizes of 1–2 μm with a variance of the size distribution of 0.2–0.4 are the most frequently observed (see review by Kahre et al. (2017) and recent results from solar occultations recorded by ACS and analysed by Luginin et al. (2020)). Changing the dust particle radius to 1 μm or 2 μm affects the ratio $Q_{\text{ext}}/r_{\text{eff}}$ by less than 10%, hence the uncertainty on total dust opacity linked to assuming a constant dust particle radius of 1.5 μm remains small as well ($< 10\%$).

Dust is assumed to be well-mixed (ie., with a constant mass mixing ratio) in the first two scale heights above the surface, then its mixing ratio decreases linearly with height (log-pressure, to be precise). The vertical profile of ice mixing ratio is parameterized by a Gaussian profile centred at a condensation level that can either be set arbitrarily, or is computed by combining knowledge on the temperature profile (being retrieved simultaneously) and the water vapor column, taken from the Thermal Emission Spectrometer (TES) climatology (Smith, 2004). More information on the choice for *a priori* parameters for dust and water ice clouds is given in section 3.2.4. As we neglect scattering effects, our retrieved quantities are "effective" dust and water ice absorption. The error induced by neglecting scattering was previously estimated by Smith (2004). The authors ran tests including or not scattering effects and concluded that the actual extinction optical depth was approximately 1.3 and 1.5 times as large as the “effective” absorption optical depth for dust and water ice, respectively. Their work was based on the same dust and ice features as ours, i.e. at 1075 and 825 cm^{-1}, and similar particle sizes. Similar factors between absorption and extinction values were found by Wolff and Clancy (2003).

Additional ancillary data are needed to compute synthetic spectra. For each TIRVIM measurement, surface pressure is extracted at the corresponding season, local time and location from the Mars Climate Database (MCD) version 5.3 (Millour et al., 2018) us-

ing the surface pressure predictor described in Forget et al. (2007). The latter exploit high resolution MOLA (32 pixels/degree) topography. The CO_2 mixing ratio vertical profile is also extracted from the MCD. Spectral surface emissivity for each observation is interpolated from TES spectral emissivity map (Bandfield & Smith, 2003). As information within the CO_2 band is missing, we adopt a simple linear interpolation of surface emissivity in this spectral range. An error in surface emissivity in this region should not impact our results, as the contribution from the surface to the outgoing thermal emission is negligible in this spectral range of strong CO_2 absorption.

3.2 Retrieval algorithm

The overall goal is to find a set of parameters (vertical profile of temperature, surface temperature, integrated opacity of aerosols) that produces a synthetic spectrum (using the aforementioned forward radiative transfer model) in close agreement with a TIRVIM spectrum, within noise levels. As we are facing an ill-posed and underconstrained inverse problem, there exists a strong degeneracy of this set of parameters, including potentially non-physical solutions that could still match the observed spectra. In order to regularize the inverse problem, we choose the widely-used framework of optimal estimation retrieval, described in Rodgers (2000). In this framework, the cost function to be minimized includes not only the evaluation of goodness-of-fit to the data (χ^2), but also an additional regularization term that should contain our best *a priori* physical knowledge on the desired quantities. To be more accurate, we use the same choice of cost function as Conrath et al. (2000) that is slightly different from that described in Rodgers (2000) (see next subsection). The choice of the *a priori* values together with their covariance matrix is not trivial, though, requiring a certain level of trial-and-error process.

We detail below the retrieval scheme for atmospheric temperature on the one hand, surface temperature and aerosols on the other hand, and will lastly describe how they are combined with a detailed description of the different steps of our algorithm.

3.2.1 Atmospheric temperature retrieval scheme

Regarding the retrieval of temperature vertical profiles, we adopt the same approach as Conrath et al. (2000) for the analysis of the TES/MGS data, also successfully applied

to PFS/Mars Express data by Fouchet et al. (2007). A first guess temperature profile T_a is iteratively modified following this equation:

$$\mathbf{T}_{n+1} = \mathbf{T_a} + \mathbf{W}(\mathbf{\Delta T_b} + \mathbf{K}(\mathbf{T}_n - \mathbf{T_a})) \quad (2)$$

$$\mathbf{W} = \mathbf{S_a K^T}(\mathbf{K S_a K^T} + \frac{1}{\alpha}\mathbf{S_e})^{-1} \quad (3)$$

where $\mathbf{T}_n$ is the temperature profile at iteration n, $\mathbf{\Delta T_b}$ the difference between the synthetic spectrum (in brightness temperature) computed with temperature $\mathbf{T}_n$ and the TIRVIM one, $\mathbf{K}$ is the functional derivative matrix, with elements $K_{\rm ij}$ defined as the derivative of the brightness temperature at wavenumber i over the temperature at a pressure level j ($K_{\rm ij} = \frac{dI_i}{dT_j}$), $\mathbf{S_a}$ the covariance matrix for the *a priori* temperature profile, and $\mathbf{S_e}$ the error covariance matrix, whose diagonal elements contain the squared wavenumber-dependent NER of TIRVIM spectra. Finally, α is a parameter used to assign more or less weight to the observations with respect to our *a priori* knowledge. Based on several tests on retrievals on both synthetic and actual TIRVIM data, we find that a value of $\alpha = \beta \times \frac{trace(\mathbf{S_e})}{trace(\mathbf{K S_a K^T})}$, with a nominal value of 3 for β, yields satisfactory results.

3.2.2 Information content for the temperature

Only a portion of a TIRVIM spectrum is exploited in equations 2 and 3, as information on the atmospheric temperature comes from the CO_2 band in the range 660–780 cm^{-1} (recall that we exclude data in the range 620–660 cm^{-1} due to low signal-to-noise ratio). Actually, within this spectral range, there exists a strong redundancy in the information content, where contribution functions at different wavenumbers peak at the same pressure level. While some level of redundancy is desirable (because of noise in the data), all individual spectral measurements within this range are not needed to achieve a satisfactory temperature profile retrieval. Hence, in order to reduce computation time, we select 50 out of the 177 wavenumbers in TIRVIM spectra in the range 660–780 cm^{-1}. This reduces the matrix size of $\mathbf{K}$, $\mathbf{S_e}$ and $\mathbf{\Delta T_b}$ in equations 2 and 3. We keep all 14 points between 665.3 cm^{-1} and 673.7 cm^{-1}, then one every three points until 714.3 cm^{-1}, then one every five points up to 780 cm^{-1}. We emphasize that we have tested retrievals in different configurations (50 or 177 points in the CO_2 band) and confirm that retrieved temperature profiles are almost identical.

We present an example of the information content sampled by these 50 wavenumbers, as a function of pressure, in Figure 5 (examples at other latitudes and local times will be given in section 4, focused on synthetic retrievals). The exact pressure levels at which the functional derivatives peak and their relative amplitudes depend on the temperature profile itself and on the aerosol load ; however, general trends can be drawn. The radiance or brightness temperature at wavenumbers 670–780 cm^{-1} is mostly sensitive to the atmospheric temperature at altitudes 4–35 km (20–400 Pa for a surface pressure of 610 Pa). Despite having selected about a third of the spectral information in the CO_2 band, we note that redundancy in functional derivatives peaking at similar levels is still large. On the other hand, information on the temperature at altitudes 35–55 km (2–20 Pa) exclusively comes from the two spectral points at the core of the CO_2 band, at wavenumbers 667.7 and 668.4 cm^{-1}. We also notice that these two functional derivatives have broader full width at half maximum (being about 15 km) compared to the contribution functions peaking at 4–35 km, which have half-widths of $\sim$10 km. This gives a qualitative estimate of the vertical resolution of the retrieved profile, which varies between 1 scale height (in the lower troposphere) and 1.5 scale heights (in the middle troposphere).

3.2.3 Building the first guess temperature profile

In a Bayesian approach, one possibility would be to compute both the *a priori* temperature profile and its covariance matrix $\mathbf{S_a}$ from a climatology of temperature profiles, such as provided by the Mars Climate Database. The first guess can also be different from the *a priori* temperature profile. This is the methodology employed for instance by Grassi et al. (2005) for PFS data analysis. However, this method has its limits when atmospheric conditions depart significantly from the "climate" scenario of the MCD (in the event of a global dust event, for instance , see Wolkenberg et al., 2018). We choose instead the same approach as Conrath et al. (2000): for each observation, we build the first guess from TIRVIM spectra themselves. Furthermore, this first guess is chosen as our *a priori* profile. Hence, we are not here in a purely Bayesian formalism, as our *a priori* profiles already contain information from the spectra. The method of Conrath et al. (2000) to build this first guess consists of computing a set of brightness temperatures from the radiance at several wavenumbers within the CO_2 band (in our case, we use the set of 50 wavenumbers described above), corrected for atmospheric transmission and surface con-

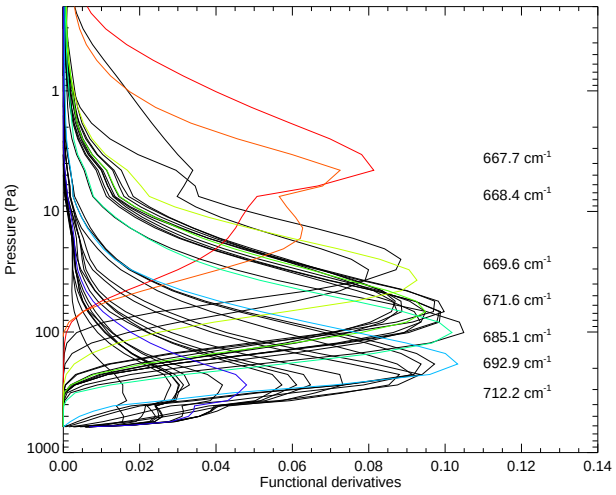

Figure 5. Functional derivatives of the temperature as a function of pressure, for fifty different wavenumbers in the range 665–780 cm^{-1}. Seven out of fifty wavenumbers are highlighted in color and labeled. In general, TIRVIM spectra probe from lower to higher pressure regions starting from the core of the CO_2 band (at 667.7 cm^{-1}) to its wing. This example was generated for atmospheric conditions extracted from the Mars Climate Database for Martian Year 25, Ls=180°, latitude 15°N and local time 9 AM.

tribution. This set of brightness temperatures is then attributed to different atmospheric pressure levels based on the contribution functions (see equations 13 and 14 of Conrath et al., 2000) to build a first rough temperature profile, typically constrained between 10 and 40 km. This profile is then extrapolated upwards and downwards, using a method that differs from Conrath et al. (2000). For the upper part, we force the profile to smoothly go back to a climatology based on four martian years (MY29–MY33) of diurnal averages of MCS temperature profiles, binned every 5° of Ls and 5° of latitude, zonally averaged. Regarding the extrapolation downwards, after several trials, we choose to impose the slope of the temperature profile in the first scale height based on an empirical relationship (determined from simulated synthetic spectra) between the temperature vertical gradient in the first scale height and the difference in brightness temperature between 690 and 702 cm^{-1}. Finally, the entire profile is smoothed vertically by convolution with a Gaussian function with a 0.75 scale height half width. Regarding the covariance matrix $\mathbf{S_a}$, it is defined like in Conrath et al. (2000) as to filter non-realistic vertical oscillations:

$$S_{\mathrm{a_{ij}}} = \exp\left[-\frac{log(p(i)/p(j))^2}{2c^2}\right] \tag{4}$$

with c a correlation length chosen nominally as 0.75 scale heights. Like β, this parameter c is read in an input file.

3.2.4 Surface temperature and aerosol retrieval scheme

We now focus on the range 780–1300 cm^{-1} of the spectra. The general idea is to iteratively update first guess values of the surface temperature T_{surf} and of the dust and water ice vertical mixing ratio profiles q_{d} and q_{i} by exploiting the calculated functional derivatives of the brightness temperature to these parameters (noted K_{Ts}, $K_{\mathrm{q_d}}$ and $K_{\mathrm{q_i}}$). We emphasize that nadir spectra are generally not (or very weakly) sensitive to the vertical distribution of dust and water ice clouds, hence, we only retrieve a scaling factor to *a priori* profiles of their mixing ratio. The functional derivatives $K_{\mathrm{q_d}}$ and $K_{\mathrm{q_i}}$ we use actually relate to the change in brightness temperature associated with a *relative* change in q_{d} and q_{i}, rather than a change in absolute values of q_{d} and q_{i}. Minimization of a cost function leads to the following iterative increment in T_{surf}, q_{d} and q_{i} in the case of the simultaneous retrieval of all three quantities:

$$T_{\mathrm{surf(n+1)}} = T_{\mathrm{surf(n)}} + \sigma_{\mathrm{Ts}} K_{\mathrm{Ts}}^T Y \Delta I \tag{5a}$$

$$q_{\mathrm{d(n+1)}} = q_{\mathrm{d(n)}}(1 + \sigma_{\mathrm{q_d}} K_{\mathrm{q_d}}^T Y \Delta I) \tag{5b}$$

$$q_{\mathrm{i(n+1)}} = q_{\mathrm{i(n)}}(1 + \sigma_{\mathrm{q_i}} K_{\mathrm{q_i}}^T Y \Delta I) \tag{5c}$$

$$Y = \left(\sigma_{\mathrm{Ts}} K_{\mathrm{Ts}} K_{\mathrm{Ts}}^T + \sigma_{\mathrm{d}} K_{\mathrm{q_d}} K_{\mathrm{q_d}}^T + \sigma_{\mathrm{i}} K_{\mathrm{q_i}} K_{\mathrm{q_i}}^T + \frac{1}{\gamma} S_e\right)^{-1} \tag{6}$$

where ΔI is the difference between the synthetic spectrum – computed with surface temperature $T_{\mathrm{surf(n)}}$, aerosol mixing ratio profiles $q_{\mathrm{d(n)}}$ and $q_{\mathrm{i(n)}}$ – and the TIRVIM spectrum in the range 780–1300 cm^{-1}; σ_{Ts}, σ_{d} and σ_{i} are the *a priori* covariance on surface temperature (set to 8K), dust and water ice cloud opacity (both set to a factor of three); and γ is a parameter used to assign more or less weight to the observations with respect to our *a priori* knowledge (determined similarly as α in equation 3). As for the atmospheric temperature retrieval, we do not use all spectral points in the retrieval of

these quantities, especially since dust and water ice exhibit smooth spectral features. A set of 78 out of 805 spectral points in the range 785–1295 cm^{-1} was carefully selected to optimize the information content and gain calculation time. This selection was based on a study of the Jacobians for surface temperature, dust and water ice opacities (derivatives of brightness temperature over these scalar quantities), computed for synthetic data. We retain more points near the peaking values of these Jacobians (one point every ∼3 cm-1). Away from these regions of maximum information content, our sampling is sparser (one point every ∼10 cm-1).

We emphasize that dust and water ice cloud opacities cannot be retrieved at all local times. This issue was already discussed e.g. by Smith (2019a) from the analysis of THEMIS broadband thermal infrared images, and also by Wolkenberg et al. (2018) from PFS/Mars Express spectra. The challenge in retrieving aerosol opacities from these nadir-viewing instruments is that sensitivity to dust and/or water ice (given here by K_{qd} and K_{qi}) tends to be zero in the event when most of the dust (or water ice) mass load lies at altitudes where atmospheric temperature is close to surface temperature. In that case, no absorption nor emission band will be visible in the spectra, irrespective of the aerosol load, and we will not be able to determine the aerosols' opacity. This situation is mostly encountered near 7–8 AM and 7–8 PM, as we will see in section 5.3. Of course, in the event when there are no aerosol absorption or emission bands seen in TIRVIM data, but at the same time the surface – atmosphere temperature contrast is high, even a small amount of dust (or water ice) should have produced an absorption or emission feature, hence we would confidently conclude that the atmosphere is clear.

We have also added the option to not retrieve a given variable among surface temperature, dust and water ice opacity in case a TIRVIM spectrum is too noisy and does not exhibit a clear dust or ice spectral feature. Indeed, we noticed that in such situations, our retrieval algorithm sometimes fitted the data (or rather, fitted the noise) with a combination of values of surface temperature, dust or ice opacities that were unrealistic. Retrieving only a subset of parameters in this situation help regularize the inverse problem. In more detail, we start by computing the standard deviation of TIRVIM brightness temperature in the range 1210–1290 cm^{-1} as a proxy for the noise level (hereafter called noise_{1250}). Indeed, beyond 1200 cm^{-1}, we find that the NER provided by the instrument team often underestimate the actual noise level so that we chose to estimate it from the data itself in this rather flat (in brightness temperature) spectral region. We

then compute the contrast in TIRVIM average brightness temperature between the range 1060–1130 cm^{-1} and the range 1225–1290 cm^{-1} (hereafter called $contrast_{dust}$) and the one between 810–860 cm^{-1} and 1225–1290 cm^{-1} (hereafter called $contrast_{ice}$). Different configurations are considered:

- If spectra are noisy (arbitrary threshold set to $noise_{1250}$ >6K) and $contrast_{dust}$ and $contrast_{ice}$ are smaller than $noise_{1250}$: we retrieve surface temperature only;
- If spectra are noisy and $contrast_{dust} < noise_{1250}$ and $contrast_{ice} > noise_{1250}$: we simultaneously retrieve surface temperature and water ice opacity;
- If spectra are noisy and $contrast_{dust} > noise_{1250}$ and $contrast_{ice} < noise_{1250}$: we simultaneously retrieve surface temperature and dust opacity;
- Else, we simultaneously retrieve surface temperature, dust and water ice opacity.

The retrieval of only one or two parameters among surface temperature, dust and water ice opacity follows the same form as equations 5a–6.

In order to evaluate the level of confidence we have on dust and water ice cloud retrievals, we adopt a similar approach as Smith (2019a) and define the 1-σ relative error on dust opacity $dust_{\mathrm{err}}$ as the ratio between the 1-σ noise level (in brightness temperature; computed from the instrument NER estimates) at 1100 cm^{-1} and the functional derivative $K_{\mathrm{q_d}}$, computed at 1100 cm^{-1} as well. When this ratio equals one, this means that a 100% change in dust opacity results in a change in brightness temperature similar to the 1-σ noise level. In other words, the 1-σ error on dust opacity would be 100% in this case. We arbitrarily define a dust quality flag that considers only dust retrievals with $dust_{\mathrm{err}}$ values lower than one. Similarly, for ice, we define ice_{err} as the ratio between the noise level and $K_{\mathrm{q_i}}$, both estimated at 820 cm^{-1}. Examples of typical errors on dust and water ice cloud opacities will be detailed in section 4.

3.2.5 Choice of a priori values for aerosols and setting of the cloud condensation level

Another challenge is related to the fact that even though the functional derivatives used in the retrieval in equation 5a–5c are updated at each iteration, they are themselves first computed based on prior information on the vertical distribution of dust and water ice clouds that can be erroneous and hamper our retrievals. For instance, if we *a pri-*

ori assume that a cloud lies at an altitude where atmospheric temperature is warmer than the surface temperature, the functional derivative K_{q_i} will be positive in sign (the cloud is assumed to be seen in emission). However, if the cloud is actually at an altitude where atmospheric temperature is colder than the surface (i.e. the cloud is seen in absorption in TIRVIM data), the retrieval scheme will fail to reproduce the observed absorption band as we only retrieve a scaling factor to an initial mixing ratio profile. This actually implies that there are situations where we can say that TIRVIM spectra contain some level of information on the cloud altitude. This piece of information has to be exploited, if possible. More generally speaking, the choice of *a priori* values for the vertical distribution of q_d and q_i can strongly influence our results while in the case of atmospheric temperature retrieval, this choice is less critical.

While dust is rather ubiquitous, water ice clouds can be discrete and exhibit a greater spatio-temporal variability than dust, which makes their choice of *a priori* properties not trivial. It would be tempting to run several retrievals for each observation, starting from different *a priori* values for dust and ice opacities, assuming different cloud condensation levels, and then select *a posteriori* the solution (or solutions) that best match the considered TIRVIM spectrum. However, this would be too costly in computation time. We thus choose an intermediate path: in a first stage of our algorithm, we explore various combinations of dust opacities (τ_{dust}=0.1, 0.3, 0.6), ice opacities (τ_{ice}=0.1, 0.3, 0.7, 1.5 and 4) and cloud altitudes (three different values, among which the condensation level derived from TES water vapor climatology). We exclude combinations that are not realistic, e.g. large values of τ_{ice} at high altitudes. For each combination, we compute a first guess synthetic spectrum in the range 800–1290 cm^{-1}. To compute these spectra, the surface temperature is evaluated directly from the brightness temperature of TIRVIM spectra in the range 1240–1290 cm^{-1}, a portion of the spectrum that is rather transparent. We pick up the combination of parameters that best match the considered TIRVIM spectrum at this stage as *a priori* values, and then start a retrieval. This way, we make sure to start our retrievals with prior values that match basic features of TIRVIM data (dust or ice seen in emission or absorption ; hints for a clear or a high dust load in the atmosphere, for instance) and hence, with sensible values of K_{q_d} and K_{q_i}.

3.3 Algorithm steps

The different steps of our algorithm, including those dealing with the special cases mentioned above, and the link between the retrieval of aerosol opacities, surface and atmospheric temperature, are summarized in this section.

First, a pre-processing step discards observations with known issues and extract ancillary data (surface pressure and CO_2 profile from the MCD, surface emissivity and climatological water vapor from TES, temperatures from MCS climatology, etc). One NetCDF file per TGO orbit is created, which contains TIRVIM spectra with the aforementioned ancillary information. Typically, one file contains hundreds to thousands of spectra. After this pre-processing step, the retrieval algorithm works as follows :

1. Read in the NetCDF file and a list of optional parameters in a text-format input file, begin loop on all observations within this file.
2. Assuming a small dust and water ice opacity (0.03), build first guess temperature profile from the spectrum and compute the condensation level $p_{\mathrm{cond\ TES}}$ of water ice clouds, based on this first guess profile and TES water vapor climatology.
3. Compute noise_{1250}, $\mathrm{contrast}_{\mathrm{dust}}$ and $\mathrm{contrast}_{\mathrm{ice}}$.
4. If both absolute values of $\mathrm{contrast}_{\mathrm{dust}}$ and $\mathrm{contrast}_{\mathrm{ice}}$ are smaller than noise_{1250}, go to step 8 (the spectrum is considered ∼flat and there is no need to explore a range of first guess values), else continue.
5. Build a family of possible and realistic combinations of τ_{dust}, τ_{ice}, p_{cond} (among which the condensation level $p_{\mathrm{cond\ TES}}$ determined above, but also encompassing potential lower and/or upper cloud, depending on the first value $p_{\mathrm{cond\ TES}}$).
6. For each of the above combinations: build the corresponding first guess temperature profile from TIRVIM spectrum in the range 660–740 cm^{-1}; estimate surface temperature from TIRVIM spectrum in the range 1240–1290 cm^{-1}; compute a first synthetic spectrum; compare with the TIRVIM spectrum and store the corresponding χ^2.
7. Select the combination of first guess values for $(\tau_{\mathrm{dust}}, \tau_{\mathrm{ice}}, p_{\mathrm{cond}})$ corresponding to the minimum χ^2 ; keep the corresponding surface and atmospheric temperatures as first guesses as well.
8. Compute a first synthetic spectrum with the first guess values, this time along with the computation of the functional derivatives for all quantities. Initialize χ^2.

9. Start a retrieval loop. Different configurations are considered.

   - If spectra are noisy and $\text{contrast}_{\text{dust}}$ and $\text{contrast}_{\text{ice}}$ are smaller than noise_{1250}: make one step of surface temperature retrieval only;
   - If spectra are noisy and $\text{contrast}_{\text{dust}} < \text{noise}_{1250}$ and $\text{contrast}_{\text{ice}} > \text{noise}_{1250}$: make one step of simultaneous surface temperature and water ice opacity retrieval;
   - If spectra are noisy and $\text{contrast}_{\text{dust}} > \text{noise}_{1250}$ and $\text{contrast}_{\text{ice}} < \text{noise}_{1250}$: make one step of simultaneous surface temperature and dust opacity retrieval;
   - Else: make one step of simultaneous surface temperature, dust and water ice opacity retrieval following eq. 5a–5c.

10. Re-compute a synthetic spectrum with updated quantities (surface temperature and / or τ_{dust} and / or τ_{ice}), calculate the functional derivatives for atmospheric temperature.
11. Make one step of atmospheric temperature retrieval following eq. 2.
12. Re-compute a synthetic spectrum with updated atmospheric temperature, then update χ^2: if the change in χ^2 compared to the previous value is less than 2%, or if the number of iteration is > 9, then: end the retrieval loop, write outputs and go on with the next observation; else go back to step 9.

A solution is reached most of the time in 4 or 5 iterations. Even though only a fraction of the TIRVIM spectrum is exploited in equations 2, 5a–5c and in χ^2 calculations, we do compute a synthetic spectrum on the full spectral range of TIRVIM after convergence, that we keep for quality checks.

Our algorithm is not Bayesian, in particular as the so called *a priori* temperature profile already contains information from the data themselves. We note that it is also the case for the reference algorithm used for MGS/TES data, developed by Conrath et al. (2000). Rapid convergence is achieved mostly because we start the retrievals from first guess values already close to the "true" atmospheric state; however, in the next section, we will also show examples starting from different *a priori* profiles to demonstrate that our code still performs well in these conditions.

4 Synthetic retrievals and error analysis

4.1 Synthetic observations and approach

The precise tuning of the retrieval algorithm described previously results from several phases of trials and development, depending on how well the algorithm performs against synthetic measurements. To build these synthetic observations, we extract surface pressure, surface and atmospheric temperature as well as aerosol mixing ratio profiles from the Mars Climate Database for the scenario corresponding to Martian Year 25. We choose this year as it features a global dust event at Ls 190–240° (Smith et al., 2002). We extract these data at longitude 0° for various conditions: at five seasons (Ls= 0, 90, 180, 210 and 270°), 8 local times (every three hours) and 11 latitudes (from 75S to 75N, every 15°), hence 440 spectra in total. We then generate synthetic TIRVIM spectra with our forward model, then add noise (with realistic spectral dependency of the noise), and run our retrieval algorithm. This exercise is often referred to in the literature as an Observing System Simulation Experiment (OSSE).

Generally speaking, this exercise allows to test the sensitivity to, for instance, assuming a different vertical profile of dust and water ice clouds in the retrieval set-up compared to the forward model used to generate the synthetic data, and its impact on the retrieved quantities. It is also useful to identify flaws in our retrieval pipeline, fine-tune the sampling in wavenumber and refine our error analysis. We do not test the sensitivity of the retrieved quantities to an error in surface pressure, surface emissivity, spectroscopic database or aerosol particle sizes: these parameters are the same in the forward model and retrieval pipeline. Hence, the only forward model error that we consider comes from the different vertical mixing ratio profiles assumed for dust and water ice. Despite this quite favorable setting, we will see that many challenges arose solely due to the intrinsic degeneracy of the inverse problem and/or noise in the spectra.

4.2 Robustness of the synthetic retrievals: aerosols

Overall, synthetic retrievals perform quite as expected: cases with a warm surface temperature (hence high SNR) yield the most robust results for all retrieved parameters. For such cases, the integrated opacities of dust and water ice clouds are well estimated even if the assumed simplified vertical distributions of aerosols differs from the "actual" one (taken from the MCD), thanks to the large surface-atmosphere tempera-

ture contrast. A typical example of such a favorable case, with a surface temperature of 285K, is presented in Figure 6.

For surface temperatures lower than 230K, the fraction of retrievals that fail to correctly retrieve dust and water ice opacity increases. These unsatisfactory cases fall in two categories: either the quality criterion for dust or ice retrieval (defined in section 3.2.4) is not met, or this criterion is met (in the sense that we consider there is enough sensitivity to ice or dust in the spectra to trust the results) but the retrieved integrated ice or dust opacity differs significantly from the "true" (MCD) one nonetheless. This occurs (but not systematically) when the assumed dust mixing ratio at altitudes where the temperature is similar to surface temperature differs significantly from that of the MCD (in other words, TIRVIM spectra are blind to dust loading at some pressure levels but as we integrate the whole dust column opacity, an error in this blind zone impacts the integrated optical depth) ; and/or when there is a high level of degeneracy between surface temperature and dust and/or ice, in particular when spectra exhibit shallow dust and/or ice features.

About 60% of the retrievals pass our quality filter for dust. We present a summary of all these good quality-flagged retrieved dust opacities in Figure 7. Most of the scenes that pass this quality filter also exhibit large (>20K) surface–lower atmosphere temperature contrast, which is consistent as this situation corresponds to a greater sensitivity to dust. Our algorithm performs well even in dusty conditions, ie. with an integrated optical depth greater than one. However, this is probably too optimistic, as we neglect multiple scattering by dust in both the forward model used to generate the synthetic spectra and the retrieval pipeline. The Pearson correlation coefficient between MCD and retrieved dust opacity is 0.94. The standard deviation of the difference between retrieved and "true" dust opacity is 0.12, which is larger than the mean absolute 1-σ error in dust estimated from the ratio between noise level and K_{d}, which yields a value of 0.08. This suggests that our formal error is underestimated, but this is not surprising, given different sources of potential biases. We find that 12% of the dust retrievals that do pass the dust quality filter suffer from a significant bias (compared to the "true" dust opacity), ie. two times larger than our 1-σ error estimate. Some outliers are indeed visible in Figure 7, and we also identify several groups of cases where dust is systematically over or underestimated in the retrievals. The physical reason behind these outliers will be further discussed below for a few representative cases.

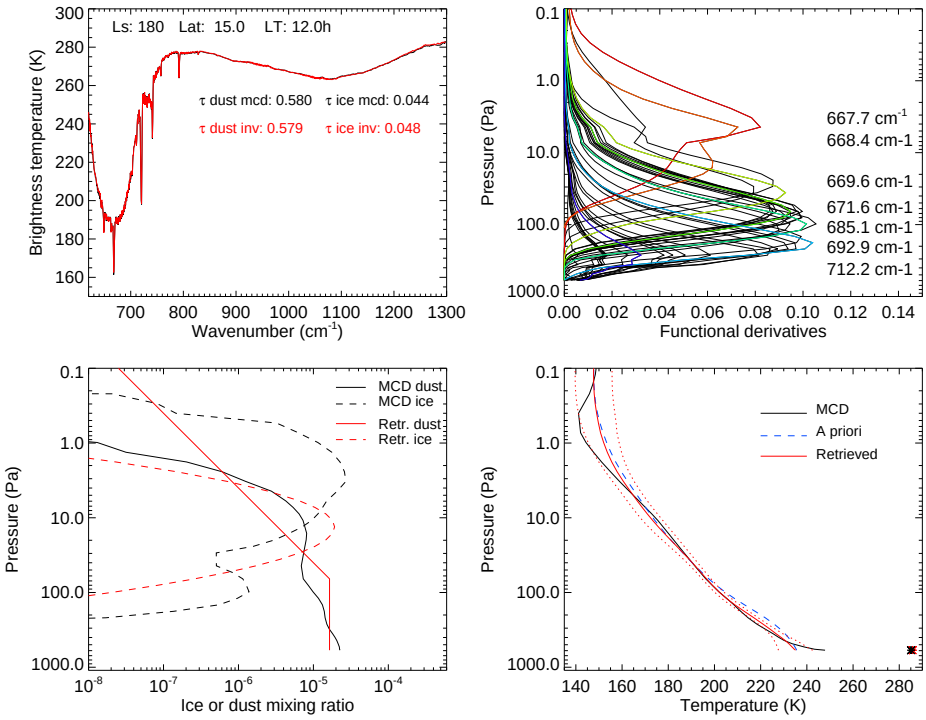

Figure 6. Summary of a synthetic retrieval for a MCD scenario extracted at L_s=180°, latitude 15°N, local time 12 AM (noon). Top, left: Synthetic TIRVIM spectrum (black) along with the best fit (red). Retrieved and MCD dust and water ice cloud integrated opacities are indicated. Top, right: functional derivatives of the temperature as a function of pressure, for 50 different wavenumbers within the CO_2 band. Several wavenumbers are highlighted in different colors and labeled. Bottom, left: Mixing ratio vertical profiles for dust and water ice as taken from the MCD (used to generate the synthetic spectrum, in black) and as derived from the retrieval process (in red ; note that only a scaling factor to a generic *a priori* profile is retrieved). Bottom, right: Temperature vertical profile from the MCD (used to generate the synthetic spectrum, in black), *a priori* profile built from the synthetic spectrum (dashed blue line), and retrieved profile (red), with error bars in red dotted lines. The black and red stars stand for the MCD and retrieved surface temperature, respectively.

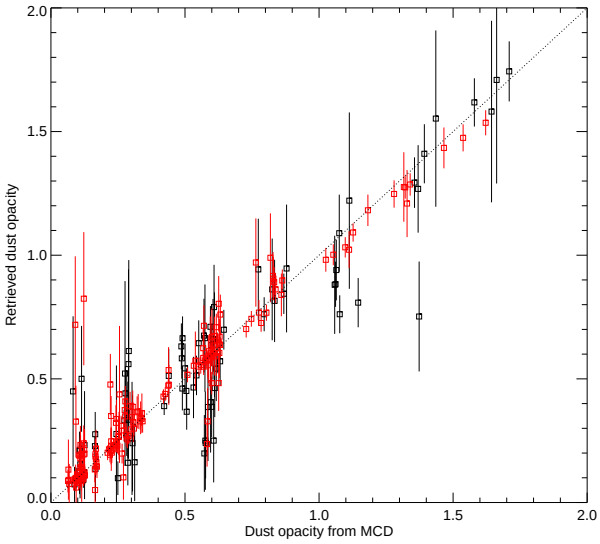

Figure 7. Retrieved versus input (MCD) integrated dust opacity in our OSSE. Only cases that passed the dust quality flag are shown. In red are cases for which the temperature constrast between surface and lower atmosphere (at 4 km) is greater than 20K. The dashed line highlights a one-to-one correspondence.

Retrieved and MCD dust opacities are shown for different seasons, latitudes and local times in Figure 8. A close-up of the Ls=270° case is shown in Figure A1. Daytime retrievals (red squares with error bars in Figures 8 and A1) are in general more robust, with smaller associated retrieval errors. This is mostly explained by a large surface-atmosphere temperature contrast and warmer temperatures (hence higher S/N ratio). Systematic overestimation of dust is visible in particular at Ls=90° at low latitudes (15S–30°N) for nighttime retrievals, where an integrated dust opacity of ∼0.2 is retrieved at night, instead of ∼0.1 in the MCD. Daytime retrievals on the other hand are robust at this season and latitudes. These different behaviors create a spurious daily cycle of dust in the retrievals, not present in the MCD. Two examples of overestimated dust are shown in Figures A2 and A3. In the first case (latitude 15N, midnight), an optically thick cloud centered at ∼100 Pa prevents the correct retrieval of atmospheric temperature in the lower atmosphere (see the functional derivatives of temperature in Figure A2). The retrieved surface-atmosphere temperature contrast is underestimated, and as a result, the dust load is overestimated by our algorithm. Indeed, there is a degeneracy in the solution of the inverse problem between dust opacity and temperature contrast, and we see here an ex-

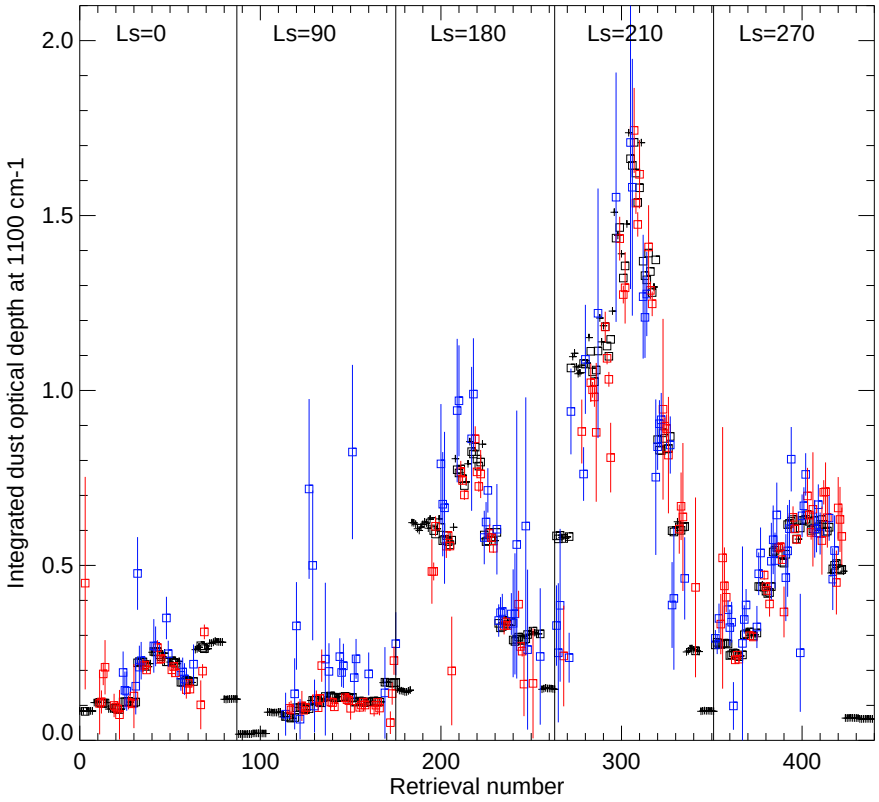

Figure 8. Evaluation of the robustness of dust retrievals sorted by season. Data points are grouped by values of Ls. Within each group, data are sorted by latitude from 75S to 75N, more clearly seen in Figure A1. Black crosses represent integrated dust opacities from the MCD, used to generate synthetic spectra in the OSSE. Blue squares are retrieved dust opacity that passed the dust quality flag for local times 0AM, 3AM, 6AM and 9PM, red squares for local times 9AM, 12AM, 3PM and 6PM. Vertical bars show the $\pm$ 1-σ error as defined in section 3.2.4. When no colored square is associated to a black cross, it means that the retrieval did not pass the dust quality flag, hence they were not plotted.

ample of a compensation of two errors. In the second example in Figure A3 (latitude 30N, 3AM), the atmospheric temperature is well estimated. Here, the issue seems to be the *a priori* dust vertical distribution, assumed well-mixed in the first two scale heights while it quickly decreases with height in the MCD profile. Because atmospheric temperature at 50–200 Pa is close to the surface temperature, we are not sensitive to the dust in this altitude range, but it will be anyhow added up in the calculation of integrated optical depth. Hence, while the amount of dust in the lower part of the atmosphere seems well estimated, the retrieved total integrated dust opacity is mechanically overestimated.

We detail another case study of challenging dust retrieval, this time during the global dust event of MY25 (Ls=210°) and shown in Figure A4. Here, in spite of the high dust opacity (1.37) in the MCD scenario, the synthetic spectrum exhibits a very shallow dust emission feature. This is due to the modest surface-atmosphere temperature contrast in the first scale height, combined to the presence of dust at high (and colder) altitudes in the MCD profile at pressures lower than 60 Pa: the shallow dust feature results from a combination of dust thermal emission at various temperatures. Here, the underestimation of retrieved dust opacity by a factor of two by our algorithm is partly due to the wrong assumption on dust vertical distribution above 60 Pa, and partly due to the fact that the large dust opacity increases the degeneracy between surface temperature and dust retrieval. As a consequence, here, a slightly warmer surface temperature and less dust in the lower atmosphere yields a similar brightness temperature as the MCD scenario, that has a slightly colder surface and higher dust loading but over a greater column, with different emission temperatures. These three examples illustrate well some of the subtle challenges encountered (that we will have to keep in mind when interpreting actual TIRVIM data). In spite of this, we recall that overall, nearly 90% of the dust retrievals that pass our quality flag are found satisfactory.

Regarding cloud retrievals, about 62% of the retrievals pass our water ice quality filter. However, 32% of these cases are significantly biased compared to the MCD values: this fraction is almost three times greater than that for dust retrievals. Furthermore, we obtain a correlation of 0.74 between MCD and retrieved ice integrated optical depth. This confirms the overall worse performance of ice retrievals compared to dust retrievals. We first investigate whether our assumed cloud altitude is realistic or not. Figure 9 shows the MCD cloud "altitude" (pressure level of maximum ice mixing ratio) versus the cloud altitude set during the retrieval. Overall, the correspondence between the two is rather

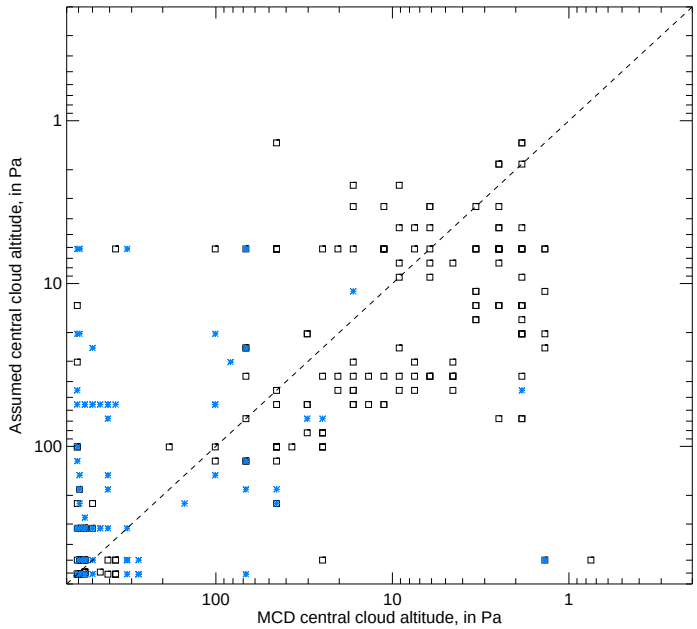

Figure 9. Cloud altitude as set up in the retrieval algorithm (see section 3.3) versus MCD central cloud altitude, in Pa. These values have been normalized to 610 Pa. Only cases that pass our ice quality flag are included. Blue stars highlight cases where ice opacity retrievals were significantly biased with respect to MCD values, despite having been flagged as robust ice retrievals.

good, except for low altitude clouds (in the MCD) whose altitude is not well captured. Actually, most of the biased cloud retrievals correspond to low altitude clouds in the MCD (see Figure 9), with large opacities (>3) in the MCD scenario. In these situations, the retrieval of surface temperature and/or lower atmospheric temperature is challenging, which in turn impacts the cloud opacity retrieval.

We also note that even a moderate error in the assumed cloud altitude can significantly impact the retrieved optical depth. An example of a wrong determination of water ice cloud opacity linked to an error in the assumed water ice vertical mixing ratio profile is shown in Figure A5. Here, the MCD profile features a vertically extended cloud, from 3 to 300 Pa, with an integrated opacity of 0.325, while the retrieval assumes a thinner cloud centered at 10 Pa and yields a much lower ice opacity (0.136). Still, the fit to the synthetic spectrum is satisfactory. This illustrates well a degeneracy frequently observed, typically in the surface temperature range 180–220K, between the ice opacity and the surface-atmosphere (where the cloud resides) temperature contrast. An optically thicker

cloud located at an atmospheric temperature closer to the surface temperature yields a similar spectral signature as a less opaque cloud located at significantly colder (or warmer, if seen in emission) atmospheric temperature, compared to the surface. The chosen example in Figure A5 is actually more subtle, as part of the vertically extended cloud is located in a region of similar atmospheric temperature as the surface, making it invisible to retrievals.

Hence, as for dust retrievals, several examples of over- or underestimation of cloud opacity can be found while still presenting a good fit to the spectra, either linked to a wrong assumption on the ice vertical distribution and/or a degeneracy with surface temperature determination and/or wrong determination of atmospheric temperature near the surface, or a subtle combination of these effects. A few difficult cases of combined degeneracy between surface temperature and both dust and ice opacity are also found.

The degeneracy between surface temperature and aerosol retrievals is further illustrated in Figure 10, which displays the error in surface temperature (the a posteriori difference between retrieved and MCD surface temperature), only for cases that pass the dust or ice quality flags. Cases for which dust or water ice opacities are significantly biased with respect to the MCD values (defined by a difference with the MCD opacity greater than two times our estimated 1-σ error) predominantly exhibit a greater error in surface temperature as well. This illustrates the degeneracy between these quantities in the inverse problem: these errors compensate and a combination of erroneous values of surface temperature, dust and/or ice opacity can yield good fits to the synthetic TIRVIM spectra nonetheless. On the positive side, as mentioned at the beginning of this section, we note that for warm surface temperatures, the retrieval of all quantities remains robust.

4.3 Robustness of the synthetic retrievals: atmospheric temperature

Retrievals of atmospheric temperature profiles are found reliable in a vast majority of cases, even when dust, water ice or surface temperature retrievals are not robust. An exception is for observations featuring low altitude opaque clouds. In this case, we can lose sensitivity to atmospheric temperature in the first scale height, as was shown in Figure A2 during the aphelion cloud belt, or in Figure A6 for a high latitude winter case. In such a situation, the retrievals are more impacted by a wrong *a priori* near-surface

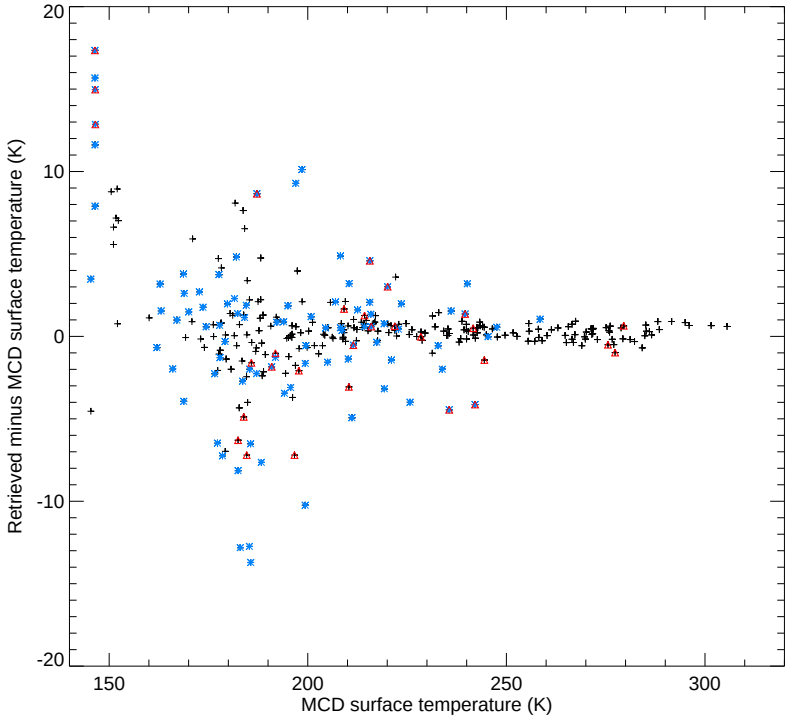

Figure 10. Error in surface temperature (defined as the difference between retrieved and "true" MCD surface temperature) as a function of the MCD surface temperature. We only show data points that pass either or both aerosol quality flags. We highlight as red triangles the dust retrievals that were significantly biased compared to MCD dust opacity, and as blue stars the cases where ice retrievals were significantly biased (greater than two times the 1-σ error).

temperature, which implies a wrong ice emission temperature and in turn, impacts the retrieved ice optical depth. The case in Figure A6 illustrates another challenge for nadir observations that constitutes the presence of a very steep temperature gradient in the lower atmosphere: here, the temperature increases from 163K to 211K between 490 Pa and 230 Pa, which represents a +48K increase over 7 km. Given the coarse vertical sensitivity of TIRVIM retrievals, such a steep gradient cannot be captured properly. We note, however, that this issue is localized and that the retrieved atmospheric temperature behaves well above 10 km as it does depart from the *a priori* profile and satisfactorily reproduces the local temperature maximum seen in the MCD profile at 100–200 Pa, where information content is high enough.

Apart from the caveat described above, we find that the atmospheric temperature retrievals perform well. The difference between the retrieved temperature profiles and that from the MCD, MY25 (used to generate the synthetic observations), along with the mean and 1-σ standard deviation of this difference, are shown in Figure 11. Most of the differences can be explained by the rather coarse vertical resolution of TIRVIM, as will be demonstrated in the following section. In the region of maximum sensitivity (5–35 km), the mean of the difference is close to zero and the 1-σ standard deviation generally lower than 3K. In the 35–55 km (2–20 Pa) region, the vertical resolution is coarser. As a result, the retrieved temperature averaged over 35–55 km may be correct, although the slope may not (as is the case for Ls=180° in our OSSE). This can result in rather large errors at a given pressure level in the range 35–55 km while the fit to the data is correct: this behavior is consistent with the information content of the data and illustrates the degeneracy due to the coarse vertical resolution. Above the altitude of maximum sensitivity (peaking near 2–3 Pa, or 50 km), the temperature profile smoothly returns to the *a priori* profile, and results should not be interpreted. This will have to be kept in mind when actual TIRVIM retrievals are discussed. We report a similar issue in the first atmospheric scale height, in particular when very steep temperature gradient (in the MCD) are not caught by our retrieval, as mentioned previously and illustrated in Figure A6. This is frequently associated to elevated water ice content near the surface and mostly occurs at Ls=270° in the winter high latitudes. Finally, we note that the atmospheric temperature retrieved for global dust storm conditions is very satisfactory and is very similar to the Ls=180° case.

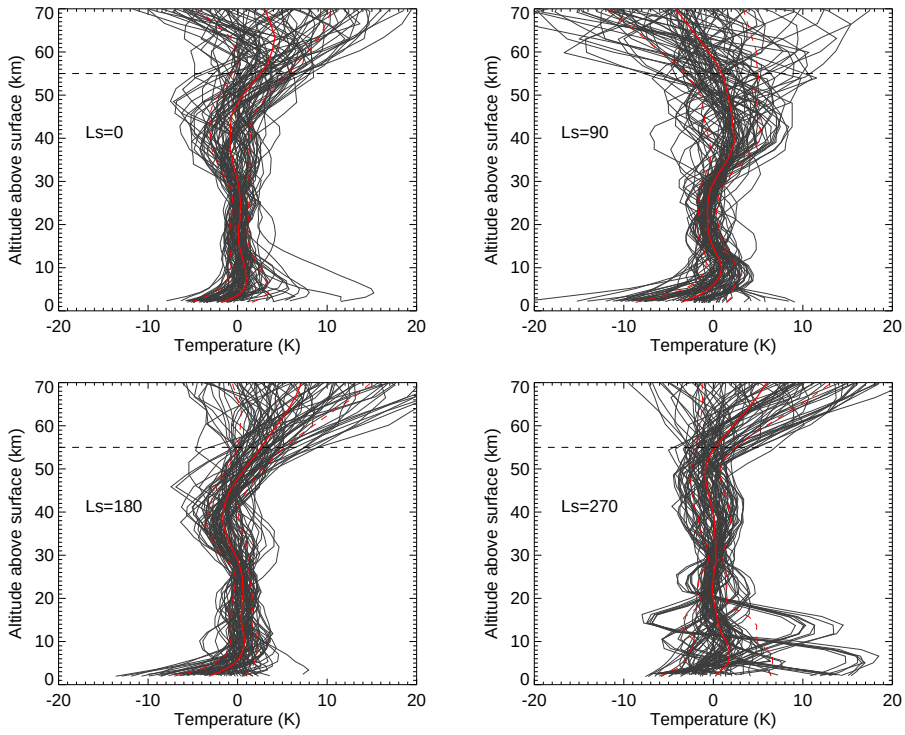

Figure 11. Vertical profiles of the difference between retrieved and MCD temperature profiles, sorted by seasons, as labeled. In each panel, the red solid line is the mean of the difference, and dashed red lines are 1-σ standard deviation added and subtracted from the mean. The case Ls=210° is not shown but is almost identical to the case Ls=180°. Most of the error results from the rather coarse vertical resolution of TIRVIM (see Figure 13 for comparison).

It might seem that our algorithm for atmospheric temperature retrieval performs well mostly because our *a priori* temperatures profiles are well chosen, but it's not fully the case. To demonstrate the limited influence of the *a priori* on our results, we have also run the 440 test cases of our OSSE with *a priori* profiles ±10K warmer or colder than our nominal profiles. Examples are shown in supplementary Figure A7. The retrieved profiles converge towards a very similar solution between a few kilometers above the surface and ∼3 Pa, independently of the choice of the *a priori* profiles. Above 3 Pa, the profiles starts to diverge and go back to their own *a priori* profile, consistently with the information content of the data. The quality of the fit in the CO_2 band is overall a bit better when using our nominal *a priori* profile, which is mostly due to slightly better fits of the core of the CO_2 band. This illustrates well the difficulty for our retrieval algorithm to depart far enough from the prior at 1–3 Pa (where information content is low) solely to fit 1–2 spectral points near 667 cm^{-1}, and justifies our choice of *a priori* profile to mitigate this challenge. Another advantage is a slightly faster convergence of the retrievals (∼4 iterations for our nominal profile instead of ∼5 iterations for the other two profiles). As we will rarely interpret observations above the 2–3 Pa level though, we show here that our results are largely independent of the choice of the *a priori* profile.

Finally, we can also briefly compare the performance of our OSSE with that of Grassi et al. (2005) done for synthetic PFS/Mars Express retrievals. As expected, the authors highlighted a difficulty of their retrievals to capture strong vertical thermal gradients in the lowest levels of the atmosphere. Based on 288 representative simulated spectra (extracted from the MCD at different seasons, local times and latitude), they found that their algorithm performed very well between 5 and 25 km, with systematic errors near zero and random error of the order of 2–3K. However, their retrieval exhibited a systematic error in the retrieved temperature above 30 km that reached 4K at 50 km, while the random error increased from 5 to 8K between 30 and 50 km. Random errors are thus very comparable to ours, while we do not report any systematic errors up to 40 km. Grassi et al. (2005) also emphasized that for altitudes 40–50 km, the contribution of the a priori profile became more important than the weight of the data in constraining the solution. This is consistent with our analysis and is not surprising, as PFS and TIRVIM have similar spectral resolution and hence similar information content.

4.4 Effect of averaging kernels on temperature profiles

We can go one step further in the comparison of retrieved versus "true" profile by emulating the combined effects of TIRVIM coarse vertical resolution and the influence of the *a priori* profile on the MCD profile. Indeed, in theory, TIRVIM retrievals provide our best estimate $\hat{\mathbf{T}}$ of the true atmospheric state $\mathbf{T}_{\text{true}}$, with the caveat that these retrieved profiles represent a smoothed version of the true state and partly contain *a priori* information, in particular in altitude regions where information content is low. This is formally expressed by:

$$\hat{\mathbf{T}} = \mathbf{T_a} + \mathbf{A_k}(\mathbf{T}_{\text{true}} - \mathbf{T_a}) \tag{7}$$

with the averaging kernel matrix $\mathbf{A_k}$ defined by $\mathbf{A_k} = \mathbf{WK}$. This matrix is proportional to the weighting kernels and reflects the fraction of information coming from the *a priori* and from the data. Examples of rows of $\mathbf{A_k}$ are presented in figure 12. The FWHM of a row gives an estimate of the vertical resolution of our retrievals and illustrates well how it is significantly larger (coarser) at higher altitudes. Note that for this example, rows of $\mathbf{A_k}$ at pressures higher than $\sim$4 Pa exhibit a sensitivity peaking at the corresponding pressure level, but that for lower pressures, the peak sensitivity occurs below that level (at higher pressures). In this example, our peak sensitivity does not reach higher than the 3 Pa level. More details on the use of averaging kernels can be found in Rodgers (2000).

By applying equation 7 to the "true" MCD profiles, we can thus run the MCD profiles through the retrieval filter, which degrades their vertical resolution and introduces a fraction of the *a priori* temperature profile. If the retrieval algorithm works without flaw, the retrieved temperature profiles obtained in this OSSE should be almost identical (within the formal retrieval error) to the best estimates $\hat{\mathbf{T}}$ obtained with equation 7. The updated comparison between our retrievals and $\hat{\mathbf{T}}$ profiles is shown in Figure 13. As expected, for most cases, the difference is smaller than 2K at all altitudes, since both profiles now go back to the same *a priori*. This signifies that 2K is a realistic estimate of the formal error linked to noise in the data and the non-linearity of the radiative transfer equation. However, there are a few cases where the difference between profiles significantly exceeds error bars. This happens in particular at Ls=270°, for the challenging scenes with a thick low altitude cloud in the MCD. Here, the unexpected difference

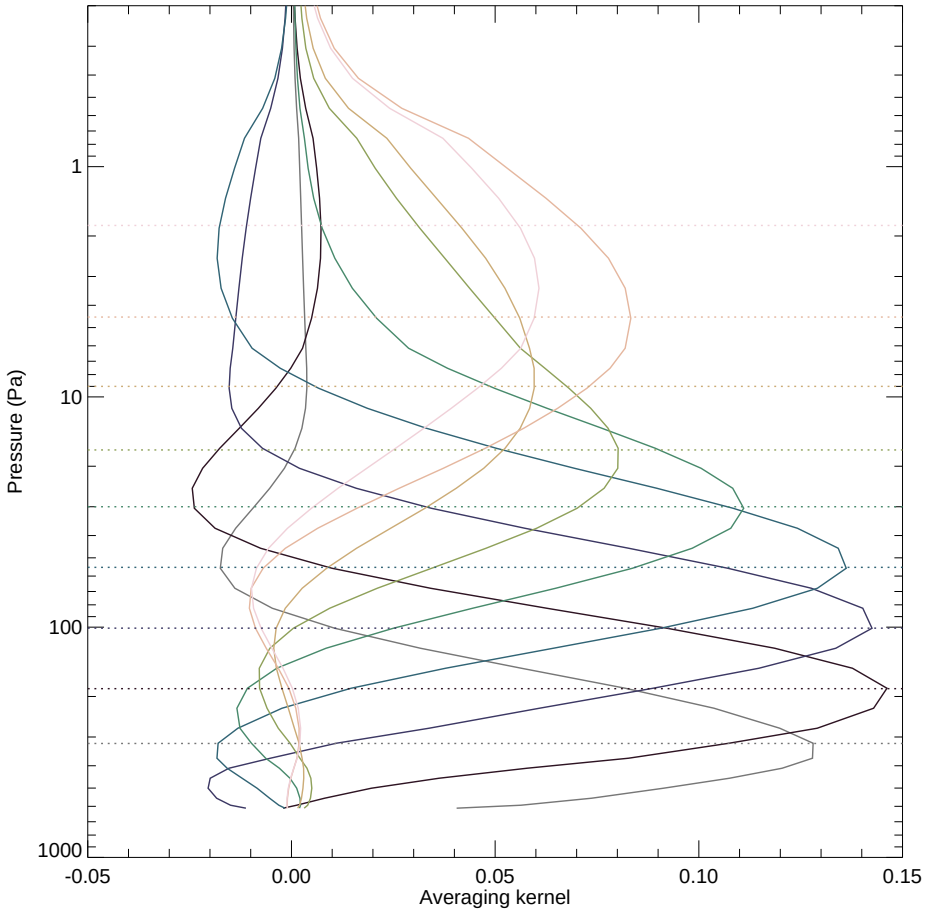

Figure 12. Example of nine rows of the averaging kernel matrix (solid lines), for nine pressure levels, materialized in dotted lines. This example was generated for atmospheric conditions extracted from the MCD for MY25, Ls=180°, latitude 15°N and local time 9 AM (same as Figure 5).

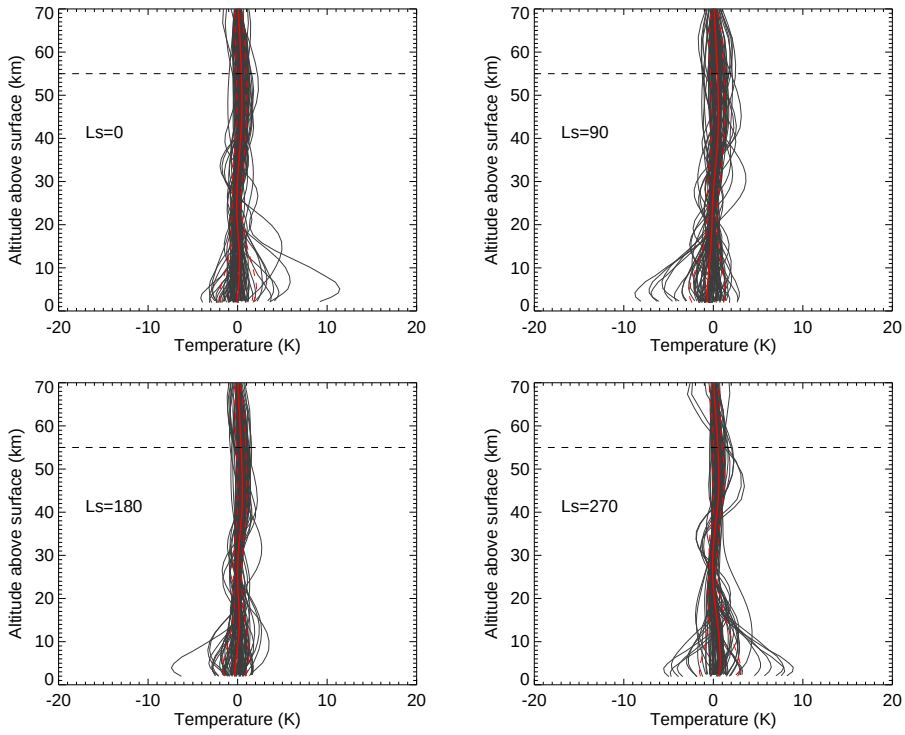

Figure 13. Same as Figure 11 but showing the difference between retrieved and "smoothed" MCD temperature profiles ($\hat{\mathbf{T}}$).

between the retrieval and $\hat{\mathbf{T}}$ stems from the wrong evaluation of the averaging kernel matrix $\mathbf{A_k}$. Indeed, our retrievals often fail to capture the actual high cloud optical depth, which impacts the calculated functional derivatives $\mathbf{K}$ and hence, $\mathbf{A_k}$ as well. In the example shown in Figure A6, the functional derivatives (wrongly) indicate that the sensitivity to the low atmospheric temperature should be rather good, which explains why $\hat{\mathbf{T}}$ is close to the MCD profile. However, the retrieved temperature is hampered here by a wrong estimate of water ice cloud opacity and the functional derivatives should not be as high in the first scale height. We emphasize though that these situations are not frequent, and that overall, the retrieval algorithm behaves as expected. To summarize, except for a few particular cases, based on this OSSE, our confidence in the retrieved temperature profiles is quite high.

5 Application to ACS/TIRVIM spectra and validation of the retrievals

We have applied our retrieval algorithm to the full TIRVIM dataset. The retrieved quantities, along with fits to the data, are distributed on the IPSL data center (Guerlet,

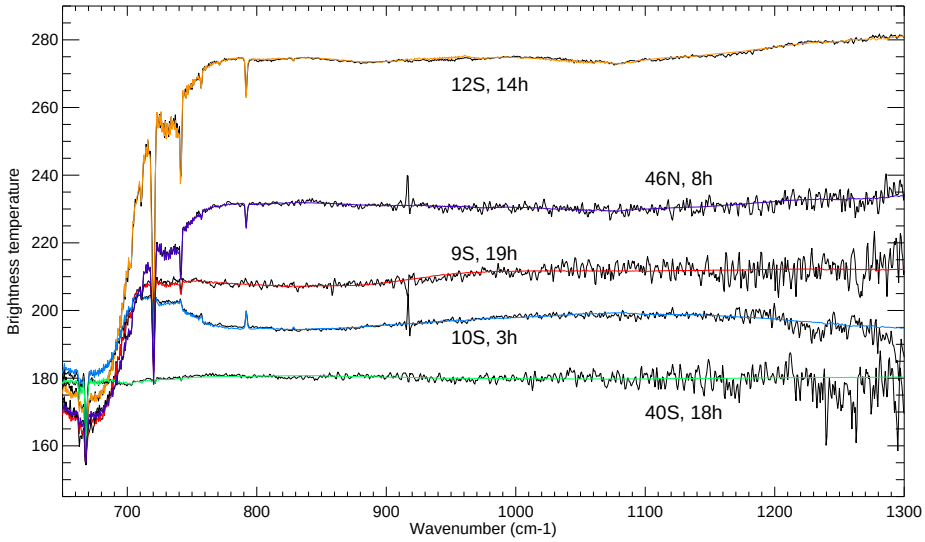

Figure 14. Examples of TIRVIM spectra (in black) and best fits (in color). These five spectra were all acquired close to Ls=150° of Martian Year 34 and near 40°W. Latitudes and true solar local times are labeled; these cases were chosen to sample different surface and atmospheric temperatures.

2021). We illustrate examples of five TIRVIM spectra and their corresponding best fits at different latitudes and local times in Figure 14. In the following, we cross validate the atmospheric temperature profiles retrieved from TIRVIM against Mars Climate Sounder (MCS) profiles and continue with an evaluation of the retrieved dust integrated optical depth. We focus in this paper on the first few weeks of TIRVIM scientific operations. Validation of water ice cloud opacity retrieval is deferred to future work for two reasons. Firstly, there were very few clouds at the season and latitude coverage corresponding to the first weeks of TIRVIM data (at Ls∼150°). Secondly, the OSSE has shown that our cloud retrievals suffered from rather larger uncertainties and biases. We thus plan to improve our ice retrieval algorithm in the future before further validation and scientific exploitation (see discussion in section 6).

5.1 Description of TIRVIM and MCS datasets and co-location method

We focus on the validation of TIRVIM retrievals acquired during the first 45 sols of TGO science phase, ie. between March 13 and April 28, 2018. This corresponds to Martian Year 34 near the northern autumn equinox, from Ls=142° until Ls=167°. This

dataset comprises both individual spectra (acquired during the first 6 days) and spectra resulting from the onboard averaging of 8 consecutive interferograms. Only 3% of the retrievals did not pass the post-processing quality filters (mostly related to the quality of the fits, based on χ^2), which is highly satisfactory. This is actually even more favorable than our exercise on synthetic retrievals (where overall, 9% didn't pass our set of quality filters) and could be explained by the low number of challenging scenes (ie. elevated aerosol load and/or cold surface) at this season and for the latitudinal coverage considered here.

We choose to validate our results against Mars Climate Sounder (MCS) observations. The MCS is a passive radiometer that records the thermal emission of the martian atmosphere in limb viewing geometry, in several thermal infrared channels, including three channels covering the 15-μm CO_2 band. The MCS team provides retrieved vertical profiles of the temperature with a vertical resolution of 5 km, from an altitude of typically ∼10–15 km (as low as 5 km in the event of a clear atmosphere) up to 80–90 km altitude, or from typically ∼300 to 0.02 Pa. Due to the Sun-synchronous orbit of MRO, these profiles are only available near local times 3 AM and 3 PM, with a rather dense coverage in latitude and longitude. The MCS retrieval algorithm includes a single-scattering approximation to account for scattering effects (Kleinböhl et al., 2011) and accounts for temperature and aerosol load variations along the line of sight (Kleinböhl et al., 2017). Temperatures derived from MCS have been used in many scientific studies (e.g., Lee et al., 2009; Kleinböhl et al., 2013). Because MCS samples a greater altitude range, achieves a higher vertical resolution, and its data are known for their high reliability, we choose to validate our TIRVIM temperature retrievals against MCS ones acquired at close locations, dates and local time.

For this exercise we select, for each TIRVIM observation, all MCS data acquired within 6° of longitude, 3° in latitude, and half an hour in local time of the considered TIRVIM profile (although we allow the date of observation to differ by ± one sol between the two data sets). These co-location criterion correspond to a trade-off between a small enough distance between TIRVIM and MCS data (in order to limit the effects of meteorological variability over these spatio-temporal scales), and a large enough amount of co-located data. Given the local time coverage of MCS data, this validation exercise will be based on "near-3 PM" and "near-3 AM" observations. Several MCS profiles can match one TIRVIM profile, and conversely, a MCS profile can be found co-located with several

TIRVIM profiles. As an example, for the near-3 PM observations in the first 45 sols period, this selection process results in 10,386 TIRVIM profiles being co-located with 6,201 unique MCS profiles. Figure 15 shows the coverage, as a function of latitude and local time, of all TIRVIM-MCS co-located observations for these first 45 sols. We note that TIRVIM data south of 55°S are mostly missing at all local times: those have been filtered out at the pre-processing stage and correspond to bad-quality flagged interferograms. Poor quality data at cold surface temperatures seem to be more frequent. This might reflect a poorer quality of the response of the instrument to low signal levels. The longitude–latitude coverage of these two co-located data sets is illustrated in Fig. 15 for the 3 PM data near Ls=150°, revealing a dense and nearly uniform coverage for both instruments.

In the following, we estimate the quality of the match between MCS and TIRVIM retrieved temperature profiles in several ways:

- We compute statistics based on MCS–TIRVIM pairs of co-located individual temperature profiles, for strict co-location criteria.
- We compare latitude-pressure sections of zonally-averaged temperatures, for strict co-location criteria.
- We also compare latitude-longitude sections of temperature at given pressure levels, acquired at similar dates and local times, but without considering spatial co-location.

5.2 Comparison of individual temperature profiles

Comparison of TIRVIM and co-located MCS temperature profiles allows us not only to assess potential biases between the two sets of retrievals, but also to evaluate the impact of the different vertical sensitivity in comparing a nadir with a limb sounder. We present two selected cases of the direct comparison between TIRVIM and MCS temperature profiles in Figure 16. In one case, near 65°N and 3 AM, the comparison is excellent between these two instruments. In the second example, a significant mismatch (up to 10K) is found over the range 1–10 Pa. Here, the difference in vertical resolution and sensitivity is striking, with TIRVIM profiles being much smoother than those from MCS. In particular, TIRVIM does not capture the sharp vertical oscillation seen by MCS in the range 1–10 Pa. As mentioned in section 3, this is consistent with the broad TIRVIM functional derivatives in that pressure range. To further investigate this, we can apply

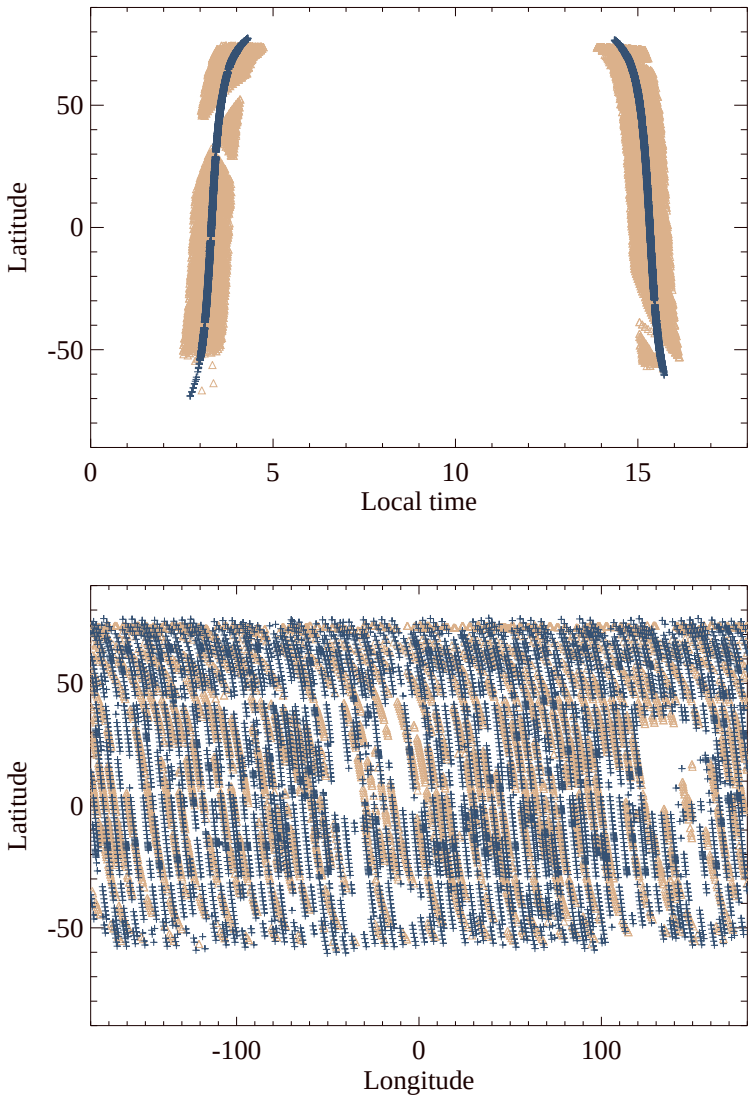

Figure 15. Coverage of TIRVIM (orange triangles) and MCS (dark blue crosses) co-located observations over the course of the first 45 sols of TIRVIM science phase (Ls∼150°). Top: latitude–local time coverage ; bottom: East longitude–latitude coverage for the near 3PM data only. Here we display existing MCS data co-located with TIRVIM, irrespective of whether the MCS profiles extend to near-surface altitudes or not.

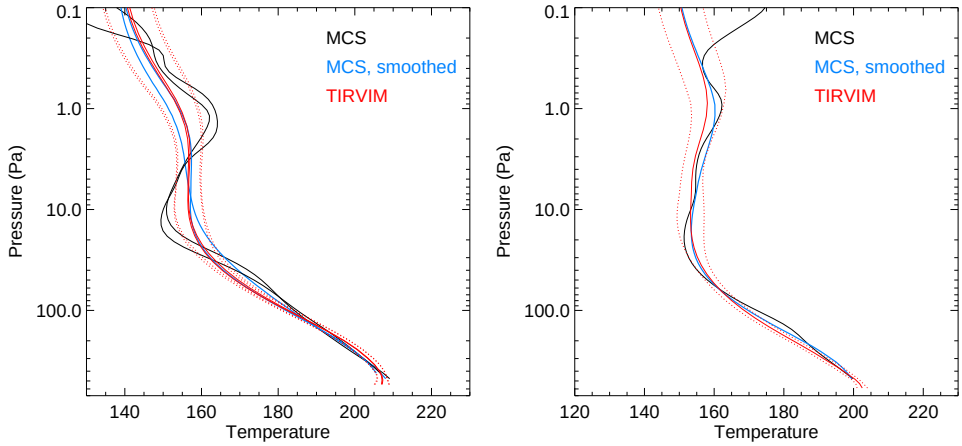

Figure 16. Retrieved temperature profiles from TIRVIM (in red, with error envelop in dotted lines) compared with co-located MCS profiles (in black) and smoothed versions of the MCS profiles (in blue) taking into account TIRVIM averaging kernels and the influence of its *a priori* profile, as described in the text. The example on the left is for latitude 45°N, the one on the right is for 65°N ; both were acquired near 3 AM. Note that the vertical axis ranges till 0.1 Pa for context, to display the behavior of MCS temperature profiles at higher altitudes, but TIRVIM data are sensitive only up to ∼2 Pa.

TIRVIM averaging kernels to MCS profiles to emulate the coarser vertical sensitivity of TIRVIM on MCS profiles and to account for the fact that when TIRVIM information content is low, the temperature smoothly goes back to an *a priori* profile. In a fashion similar to the smoothing of the MCD profiles done in our OSSE described in section 4.3, we thus replace $\mathbf{T}_{\text{true}}$ by $\mathbf{T}_{\text{MCS}}$ in equation 7 to derive a smoothed MCS profile, noted $\hat{\mathbf{T}}_{\text{MCS}}$. This procedure follows the methodology for inter-comparing profiles retrieved from different instruments documented in Rodgers and Connor (2003). By employing equation 7 in such an inter-comparison exercise, we assume that the MCS profiles are not influenced by their own *a priori* (which holds true on the pressure range considered here, 1–500 Pa) and that the MCS vertical resolution significantly exceeds that of TIRVIM, which is also satisfied. When we apply equation 7, the emulated $\hat{\mathbf{T}}_{\text{MCS}}$ profiles are now much closer to the retrieved TIRVIM profiles (see Figure 17), as can be expected. Individual differences that previously reached 10K in the direct comparison are now smaller than 3K. In the following, we thus focus on the comparison between TIRVIM and the smoothed MCS profiles, as this is more relevant to validate our results.

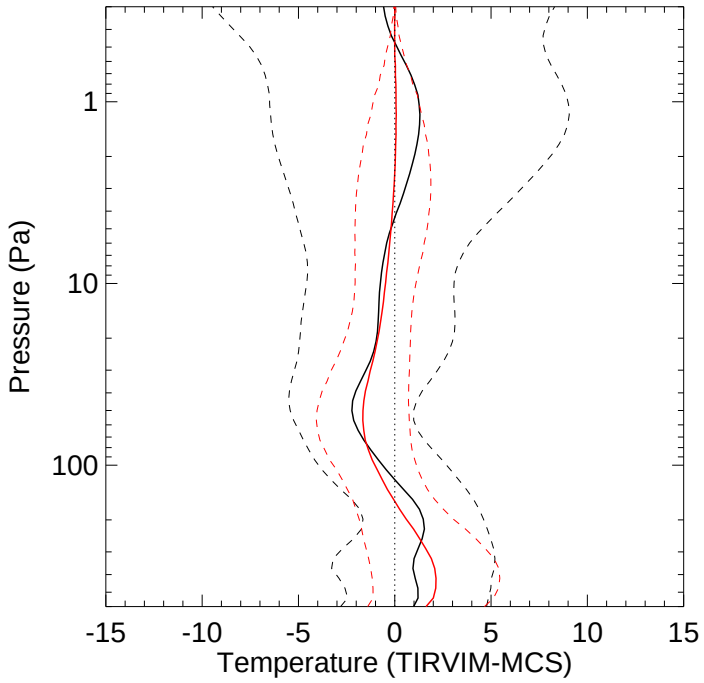

Figure 17. Statistics of the TIRVIM-$\hat{\mathbf{T}}_{\mathrm{MCS}}$ difference based on thousands of co-located temperature profiles near 3 AM, Ls$\sim$150°. The red solid line shows the average of this difference, the red dashed line is the 1-σ standard deviation of this difference. Black lines are the same but considering the difference between TIRVIM and the raw MCS profiles.

We then compute statistics based on all pairs of MCS-TIRVIM co-located profiles. Figure 17 displays the average and 1-σ standard deviation of the TIRVIM–$\hat{\mathbf{T}}_{\mathrm{MCS}}$ difference for all co-located profiles acquired near 3 AM. Both the bias and standard deviation of the difference tend to be zero at high altitudes, which simply results from the fact that the smooth MCS profiles go back to the same *a priori* profile as the one used for TIRVIM retrievals. This tendency starts between 1 and 2 Pa, which is consistent with the information content in nadir observations and remind the fact that TIRVIM retrievals should not be interpreted at pressures lower than 1–2 Pa. Over the range 2–500 Pa, the average of the difference between the two datasets is within ±2K, which we take as an estimate of TIRVIM accuracy. The standard deviation of the TIRVIM–$\hat{\mathbf{T}}_{\mathrm{MCS}}$ difference is in the range to 2–3 K, which is then an estimate of TIRVIM precision. These figures are of the order of magnitude of the typical formal retrieval error of either TIRVIM or MCS temperatures. Hence, this shows that when differences in vertical sensitivities are taken into account, TIRVIM and MCS are in excellent agreement. For completeness, we include in Figure 17 the same statistics should the raw MCS temperature profiles be taken in consideration. This illustrates the precision that would be (wrongly) derived if one were to disregard differences in vertical resolution between instruments. Similar results are obtained for the 3 PM case.

5.2.1 Comparison of the zonally-averaged temperature

We now focus on the comparison of the latitude-pressure thermal structure from TIRVIM and MCS near 3 AM and 3 PM, considering zonal averages of the co-located measurements, in bins of 5° latitude. One word of caution regarding the computation of zonal averages: while TIRVIM (nadir) retrievals are sensitive to almost down to the surface (except in extreme dusty or cloudy conditions not encountered here), MCS profiles often display a limited vertical coverage, extending only down to 80 Pa (20 km), for instance. This occurs when the optical depth of water ice and/or dust was too high at lower tangent altitudes to allow a reliable temperature retrieval from the MCS limb data. We thus have to be careful and avoid comparing a TIRVIM zonal average obtained from a rather uniform longitudinal sampling with a MCS "zonal average" for which the longitudinal coverage could be uneven and vary with altitude. For each 5°-wide latitudinal bin considered, we thus proceed by browsing one by one the MCS pressure grid and search for co-located MCS-TIRVIM data only among a subset of the MCS database, where tem-

perature was indeed retrieved at that level, before averaging the temperatures. In the following, for simplicity, we call zonal average the average over observed longitudes.

The resulting latitude-pressure cross-sections of the zonal-mean temperature are shown in Figure 18 for TIRVIM and $\hat{\mathbf{T}}_{\mathrm{MCS}}$ for 3AM and 3PM, along with the cross-sections of the TIRVIM–$\hat{\mathbf{T}}_{\mathrm{MCS}}$ difference. Similar figures for MCS temperature (on their own retrieval pressure grid, without vertical smoothing) are shown in Figure A8 in Supplementary Material. We also display the day/night temperature difference (3 PM−3 AM) for TIRVIM, $\hat{\mathbf{T}}_{\mathrm{MCS}}$ and MCS in the bottom panels of Figure 18. The agreement between the two data sets is very satisfactory, both qualitatively and quantitatively. Before applying the averaging kernels to MCS profiles, temperature differences between TIRVIM and MCS data sets are already mostly in the range ±4 K except at high altitudes (pressures lower than 10 Pa) where it can reach 12K. These larger discrepancies between TIRVIM and MCS can be explained by differences in their inherent vertical sensitivity, as discussed previously. Indeed, when comparing TIRVIM and $\hat{\mathbf{T}}_{\mathrm{MCS}}$, these differences decrease and become smaller than 4K everywhere, except at 3AM in the lower atmosphere. The TIRVIM–$\hat{\mathbf{T}}_{\mathrm{MCS}}$ temperature differences at pressures greater than 200 Pa at 3 AM is not well understood and may partly reflect an artefact of applying equation 7 on the lower portion of MCS profiles and/or partly result from an actual bias of either TIRVIM or MCS retrievals at these pressures.

The main features of the temperature field seen by MCS are well captured by TIRVIM, both at 3 AM and 3 PM. At this season (end of summer in the northern hemisphere) and in the lower atmosphere (pressures greater than 200 Pa), the meridional temperature structure is characterized by warmer temperatures in the northern hemisphere compared to the southern hemisphere. The temperature decreases rapidly towards high southern (winter) latitudes, from typically 210K at 30°S to 180K at 50°S, at the edge of the cold polar vortex. At lower pressures, the meridional thermal structure is more symmetric about the equator, as the dominant feature is the thermal tide pattern.

The diurnal mode of the thermal tide is well visible in the day/night temperature difference shown in the bottom panels of Figure 18. These figures highlight the well-known pattern of positive and negative temperature anomalies vertically stacked, at the equator and 45° latitude (with an opposite sign). These patterns have been well documented in the past, for instance by Lee et al. (2009) and Kleinböhl et al. (2013) from the anal-

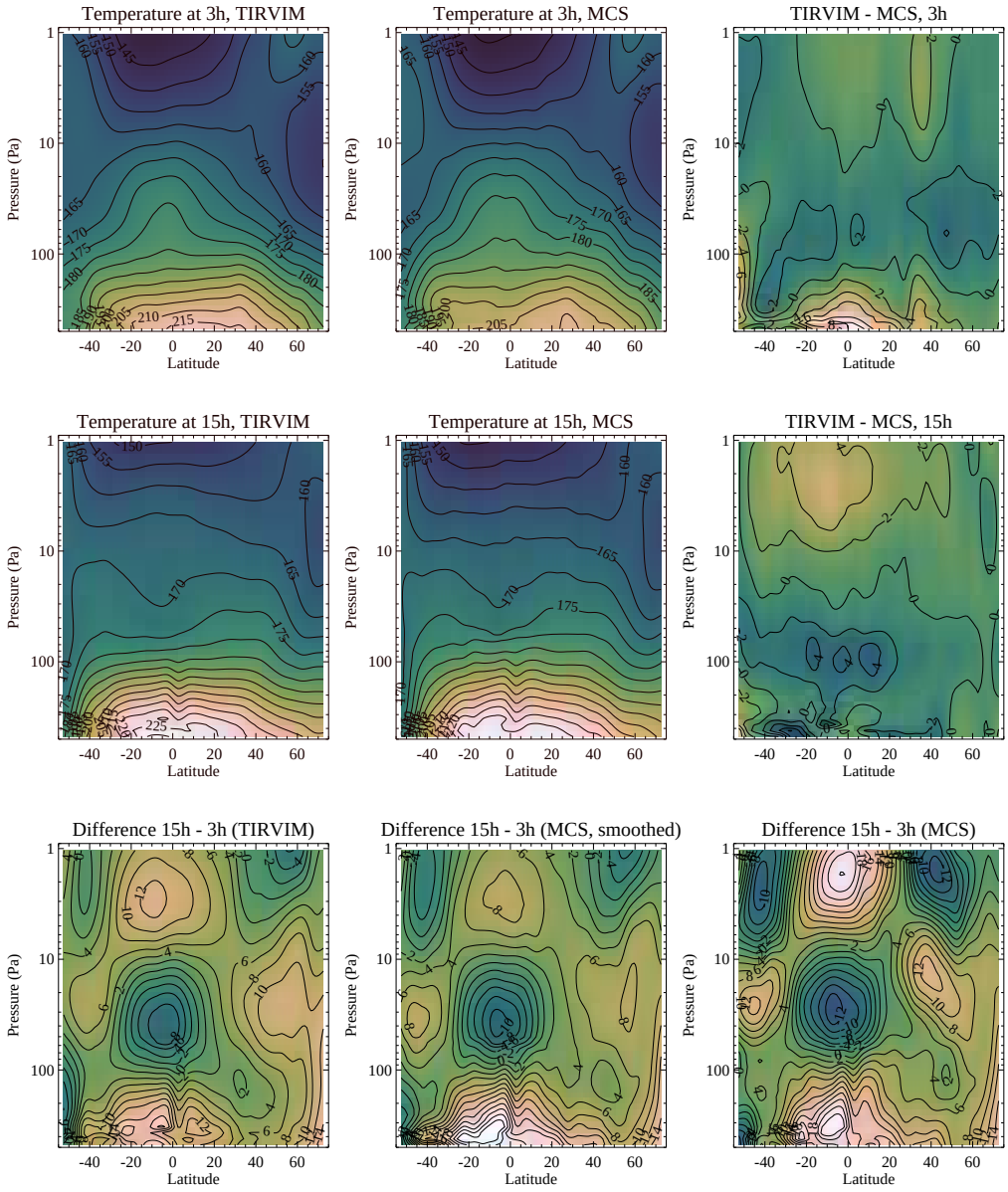

Figure 18. Pressure-latitude sections of the zonally-averaged temperature near 3 AM (top), 3 PM (middle), and the difference between 3 PM and 3 AM (bottom). In the two top rows, left panels correspond to TIRVIM, middle panels to $\hat{\mathbf{T}}_{\mathrm{MCS}}$ and the difference between the two datasets is shown in the right panels. In the bottom row, TIRVIM results are shown on the left, smoothed $\hat{\mathbf{T}}_{\mathrm{MCS}}$ in the middle and raw MCS ones on the right. These figures highlight the diurnal mode of martian thermal tides. Bins of 5° wide in latitude are used. In each panel, zonal averages are performed for a subset of the available TIRVIM and MCS data that meet co-located criteria, as described in the text.

ysis of MCS data. TIRVIM captures well the signs and locations of these temperature extrema, and the amplitude of these extrema are similar to that of the smoothed MCS temperature profiles ($\hat{\mathbf{T}}_{\mathrm{MCS}}$). For instance, near the equator, the 30 Pa afternoon temperature is found by TIRVIM to be ∼8 K colder than at night (10K in the $\hat{\mathbf{T}}_{\mathrm{MCS}}$ field), and the 3 Pa afternoon temperature is found to be ∼12 K warmer than at night (8K in the $\hat{\mathbf{T}}_{\mathrm{MCS}}$ field). The actual amplitudes of the thermal tides are known to increase even more with decreasing density, as can be seen in the raw MCS day-night temperature difference in Figure 18. Indeed, in the raw MCS data, the amplitude of the 3 PM−3 AM difference reaches +20K at the equator at 2 Pa, and +10–12K near 45° latitude at 2 Pa as well. These large extrema value are linked with sharp vertical gradients or oscillations of the temperature profile at these altitudes (see for instance the local temperature minimum seen in the MCS profiles at 3 AM and 45°N in Figure A8) that TIRVIM cannot capture properly. The fact that TIRVIM provides a muted version of the actual thermal tide signal at 1–10 Pa despite a moderate sensitivity in this region and a growing influence of the *a priori* profile (which is not based on any prior knowledge on how the thermal tide should look like) complies with its coarse vertical resolution – as shown from the good comparison with $\hat{\mathbf{T}}_{\mathrm{MCS}}$ – and is thus very satisfactory.

5.2.2 Comparison of longitude–latitude sections

We also investigated the robustness of the retrieved temperature field in terms of longitudinal variations. Here we consider the retrieved temperature binned in latitude and longitude with bin width of 3.75°×5.625°, a resolution typical of GCM simulations or the MCD. We consider all TIRVIM or MCS data acquired near 3 AM or 3 PM (±1 h) and restrict the date between Ls=149° and 155° (sols 316–329). Hence, we do not impose strict co-location between the two datasets, except for similar season and local times, and each of the resulting temperature map reflects the inherent latitude/longitude coverage of each instrument. We detail below two examples: the longitude/latitude cross-section of temperature at 288 Pa, 3 PM and that at 30 Pa, 3 AM, shown in Figure 19.

Firstly, we highlight that MCS coverage features significant gaps in the equatorial region regarding the afternoon temperature at 288 Pa. This is linked to enhanced cloud opacity, as this corresponds to the end of the aphelion belt season, which hampers limb retrievals. TIRVIM has the advantage to provide a more uniform coverage at these altitudes. A prominent wavenumber-3 feature is seen at 30 Pa and 3 AM, well captured

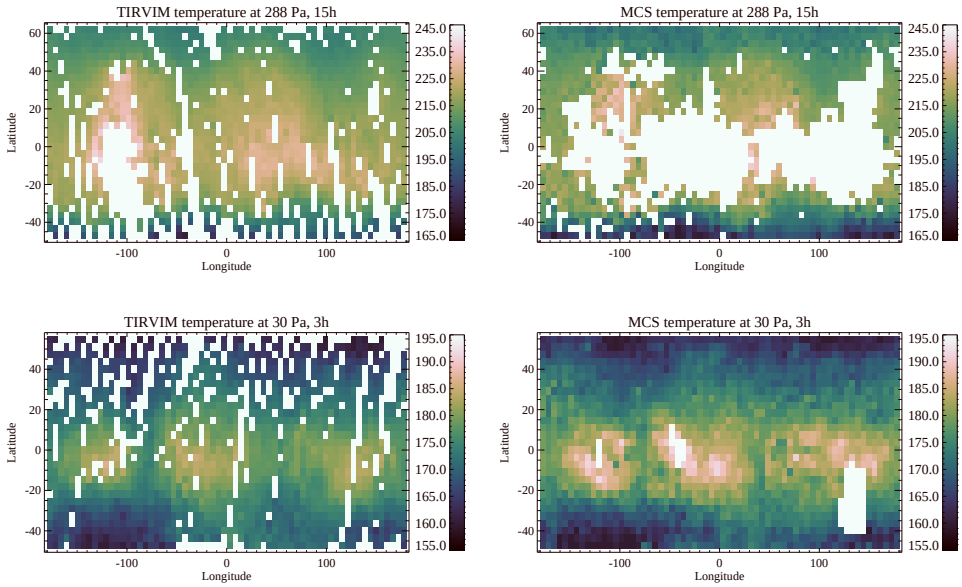

Figure 19. Temperature maps as derived from TIRVIM (left column) and provided by MCS (right column) at Ls∼150°, at 3 PM and 288 Pa (upper row) and at 3 AM and 30 Pa (lower row). Both TIRVIM and MCS feature similar longitudinal wave patterns.

by TIRVIM. The better coverage of TIRVIM data in the lower atmosphere allows a more precise characterization of a similar wave seen at 3 PM, 288 Pa. As these wave signatures are visible in a constant local time framework, this indicates that they are non-migrating thermal tides. An apparent wavenumber 3 can correspond to the diurnal Kelvin wave with zonal wavenumber 2, already reported eg. by Banfield et al. (2003); Wilson (2000) from TES and Guzewich et al. (2012); Forbes et al. (2020) from MCS observations. The analysis of these waves and of the 4-D structure of the thermal tides is deferred to future work. The local time coverage of TIRVIM will be an asset to better characterize, for instance, the amplitudes and phases of different modes of the thermal tides.

Hence, we conclude that TIRVIM temperature retrievals are reliable, whether zonally-averaged temperature fields or individual temperature profiles are considered, as long as their coarse vertical resolution are kept in mind. The main features of the diurnal mode of the thermal tides and longitudinal waves are well reproduced, compared to MCS observations. This validation exercise performed near 3 AM and 3 PM is a necessary step before the scientific exploitation of temperature variations at all local times, which will be addressed in subsequent publications.

5.3 A first assessment of dust total column opacity retrievals

In section 3.2.4, we have defined a quality filter for dust that is satisfied if the derivative of the radiance at 1100 cm^{-1} over a relative change in dust opacity is greater than the 1-σ NER estimated at 1100 cm^{-1}. In other words, we retain dust opacity values for which the estimated error is less than 100%. The fraction of dust retrievals that passed this quality (or sensitivity) filter is shown as a function of local time and latitude in figure 20, for all TIRVIM data acquired between Ls=142° and 167°. As expected and already noted in section 4, two blind zones exist near 7AM and 7PM when retrieving dust is not possible due to the lack of sufficient temperature contrast between the surface and the lower atmosphere, where dust lies. During daytime (10AM – 5PM), 100% of dust retrievals pass our quality filter due to a high sensitivity to dust (a combination of warm surface and cold atmosphere) except at high southern winter latitudes, where there is a low surface/atmosphere thermal contrast even during daytime. At nighttime, most dust retrievals (60–80%) pass our quality filter, however, the uncertainty associated with those nighttime retrievals is quite large, as we will see later.

Validation of the retrieved dust opacity is evaluated near 3 AM and 3 PM by comparison with co-located MCS data, with the same co-location criteria as for the validation of atmospheric temperature. We exploit the data product called *coldust* provided by the MCS team, which is total dust extinction at 21 μm obtained from extrapolation of the dust profile down to the surface. This extrapolation was done under the assumption that dust is well-mixed below the level of the last valid measurement (Montabone et al., 2020). We multiply this value by 2.7 to estimate the MCS dust extinction at 9.3μm and further divide it by 1.3 to get from dust extinction to dust absorption, as recommended by Montabone et al. (2015, 2020) and based on the work by Smith (2004); Wolff and Clancy (2003). Indeed, we recall that we neglect scattering effects and only derive an effective dust absorption. Dust absorption derived from TIRVIM and MCS data are compared in figure 21 for daytime and nighttime observations during the first 45 sols. Overall, TIRVIM daytime retrievals are in excellent agreement with MCS, and the associated 1-σ standard deviation of that difference for individual dust opacity retrievals (not shown) is about 10%. Dust opacity (in absorption, at 9.3μm) is low, of the order of 0.1, which is expected at this season.

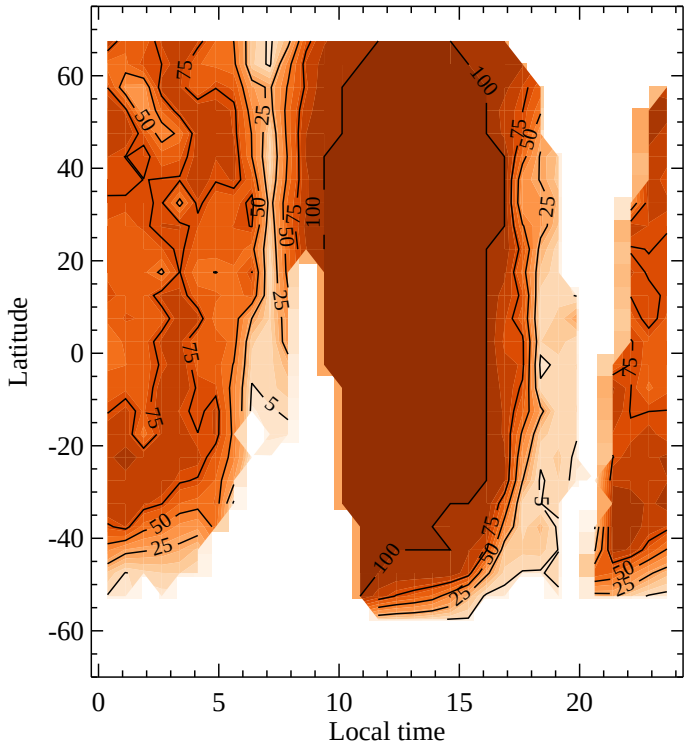

Figure 20. Percentage of retrievals that pass our quality flag for dust retrievals, for all first 45 sols of TIRVIM observations (Ls 142° to 167°), as a function of latitude and local time. White areas are missing data (the latitude - local time coverage after 45 sols is not complete).

Figure 21. Retrieved dust absorption opacity from TIRVIM near 3 PM (top) and 3 AM (bottom) as a function of latitude, for Ls∼150°. Black dots stand for individual retrieved values, while red stars indicate zonal averages for 5°-wide latitude bins. For validation purpose, blue triangles are zonal averages of MCS retrieved dust absorption scaled at 9.3 μm, for data co-located with TIRVIM measurements. Vertical bars illustrate the variance of each data set within a latitudinal bin.

At night, we note a significant scatter in the derived dust opacities from TIRVIM. This is actually consistent with the rather poor S/N ratio and individual errors that are near 40% (hence a 1-σ error of 0.08 for a 0.2 dust opacity). More worrying is a nearly systematic overestimation of nighttime dust opacity by TIRVIM, except at latitudes 15°N–40°N. The reason behind this bias is highly similar to one of the cases described in our OSSE and illustrated in figure A3: our assumed profile for dust vertical distribution overestimates the dust loading above the first scale height, which happens to be at the same atmospheric temperature as the surface. Indeed, we compared the MCS vertical profiles of dust opacity to our assumed one and confirms that the observed dust (by MCS) decreases more steeply with altitude above the first scale height. In this example, we are blind to dust located near 1.5 scale height (as it emits at a similar brightness temperature as the surface) and we are mostly sensitive to dust in the first scale height above the surface (where the surface-atmosphere temperature contrast is high enough). As we only derive a scaling factor to a profile that has too much dust at higher altitudes, this results in an overestimated total opacity.

To overcome this problem, we will have to improve our assumptions about the vertical distribution of dust in a future version of our algorithm. This is not trivial, as previous studies have shown that the dust vertical distribution can vary quite dramatically with latitude, season, and with local time in MCS data (Heavens et al., 2011; Kleinböhl et al., 2020). As MCS only provides constraints near 3 AM and 3 PM, it appears difficult to set up a robust vertical profile that would be relevant for any local time observed by TIRVIM. However, as we have previously shown that our retrieval performs well during the day (even if our assumed vertical distribution is wrong), and because dust retrievals from a nadir sounder are not possible at evening and morning times anyway, it might be sufficient to assume the nighttime MCS profile as a baseline for all our retrievals. This will be investigated in future work.

6 Summary and Conclusions

In this paper, we have documented and evaluated a retrieval algorithm tailored to ACS/TIRVIM thermal infrared spectra acquired in nadir viewing geometry, with the goal of providing the best estimates of Mars' surface temperatures, vertical profiles of its atmospheric temperature (between 5 and 50 km) and the integrated infrared optical depths of dust and water ice clouds.

This algorithm was first tested on a set of synthetic spectra in conditions chosen to be representative of the variability of the martian atmosphere at various seasons, latitudes and local times extracted from the Mars Climate Database. This constitutes our Observing System Simulation Experiment (OSSE). The precision and accuracy of these synthetic retrievals were carefully evaluated, and in parallel, the information content of the data was studied in detail. Regarding surface temperature and aerosol opacity retrievals, we find that our algorithm performs very well for the most favorable scenes, i.e. scenes with a warm surface temperature and featuring a large contrast between the surface and the atmospheric layer where aerosols reside. We caution that there are frequent conditions when this temperature contrast is low, for which the optical depth of aerosols cannot be constrained. This mostly (but not solely) occurs for observations acquired near dawn and dusk. For intermediate cases, for which we have a moderate sensitivity to dust and/or clouds, the retrievals can be significantly biased due to wrong assumptions on the vertical distribution of aerosols and/or a wrong estimate of the temperature in the lower atmosphere. Uncertainty values on dust optical depth are highly variable. They are of the order of 5-20% when the surface-atmosphere temperature contrast is high (typically, daytime observations) and of 20 to 60% when this contrast is moderate (nighttime conditions). Error on retrieved surface temperature is of the order of 1K for warm surfaces (>220K), and 3K for colder surfaces. However, significant biases (up to 10K) are reported, in particular linked with a wrong determination of retrieved dust or ice opacity, as there are situations where there is a high level of degeneracy between these quantities. Future work should thus focus on a more realistic representation of the vertical distribution of dust and water ice clouds in our algorithm, either with the help of models or other independent observations. In order to improve our aerosol estimates, we also envision in the future to co-add spectra acquired at close locations (in particular for low surface temperatures) to increase their signal-to-noise ratio.

The retrieval of atmospheric temperature is found to be very robust. Most of the differences found between retrieved temperatures and the input temperatures used to generate synthetic spectra in our OSSE can be explained by the coarse vertical resolution of TIRVIM, especially in the 2–10 Pa range where the width of the functional derivatives is of the order of two scale heights. These differences can locally reach 10K. However, when TIRVIM vertical resolution is taken into account (through the use of averaging kernel matrices), our retrieved temperatures agrees with the input ones within 2–

3K, which is taken as an estimate of uncertainty in our retrieved profiles and supports the good performance of our algorithm. We report some difficulties in retrieving atmospheric temperatures in the first scale height when there are optically thick clouds at low altitudes. These thick clouds limit our sensitivity to the lower atmosphere as the result of the degeneracy between surface temperature, cloud opacity and/or low atmospheric temperature retrievals. Important biases are also found when there is a very steep temperature inversion in the first scale height, also linked with the rather coarse vertical resolution of TIRVIM. However, these cases remain rare.

We then applied our algorithm to the first 45 days of TIRVIM operations, corresponding to Ls=142–167° of Martian Year 34. Retrieved temperature profiles were validated against co-located measurements by the Mars Climate Sounder near 3 AM and 3 PM. As for the OSSE, we find that most of the differences between the thermal structure derived from TIRVIM and MCS can be attributed to differences in vertical resolution and sensitivity. Local temperature extrema (vertical oscillations of the temperature profile) seen by MCS at pressures lower than 10–20 Pa are not well captured by TIRVIM, consistently with its coarse vertical resolution. As a consequence, the signature of the diurnal mode of the thermal tide, well visible in the day-night temperature difference, appears severely muted in TIRVIM retrievals at these low pressure levels. On the positive side, TIRVIM retrievals are highly comparable to MCS temperatures in the range 300–20 Pa (within 2–3K). Furthermore, they have the advantage of a more homogeneous spatial coverage at higher pressures, where MCS is frequently blind due to larger aerosol opacity along the limb-viewing light path. We also showed that several wave patterns are well visible in longitude-latitude cross-section of the retrieved temperature from TIRVIM.

These results are promising and open the way to future detailed studies of atmospheric dynamics at the diurnal scale in Mars' lower atmosphere. Indeed, the strength of the TIRVIM dataset lies in its sampling of the diurnal cycle, with all local times being sampled on the planet every 54 sols. As an example, full local time coverage is an asset to better characterize the amplitudes and phases of the different modes of the thermal tides.

Appendix A Supplementary figures

A1 Quality of dust retrievals

Figure A1 shows the "true" versus retrieved dust opacity at different latitudes and local times, highlighting a subset of Figure 8 in the main text.

A2 Examples of synthetic retrievals

In this appendix we display five detailed examples (out of the 440 test case retrievals of our OSSE) showing fits to synthetic TIRVIM spectra ; functional derivatives (Jacobian) for the retrieval of atmospheric temperature ; and the comparison between "true" and retrieved quantities. These examples are chosen as they exhibit significant errors on dust opacity, water ice opacity, surface temperature and/or atmospheric temperature despite satisfactory fits to the spectra. These figures are discussed in the main text.

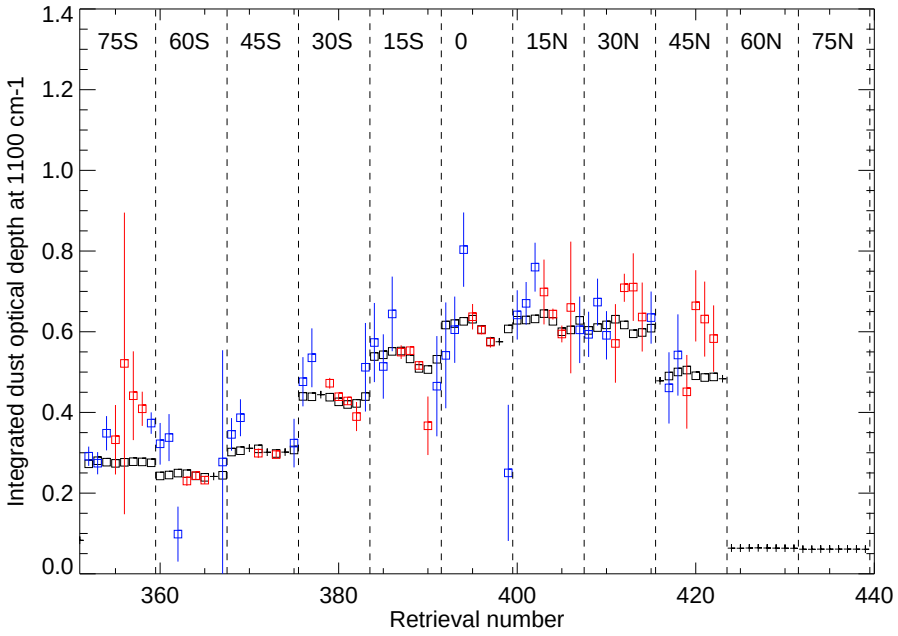

Figure A1. Same as Figure 8 but only for data at Ls=270°, with labels for different latitudes. For a given latitude, data is sorted by local time (from midnight to 9 PM, every three hours). No satisfactory dust retrievals were obtained at latitudes 60N and 75N (winter) due to cold surface and poor signal-to-noise ratio.

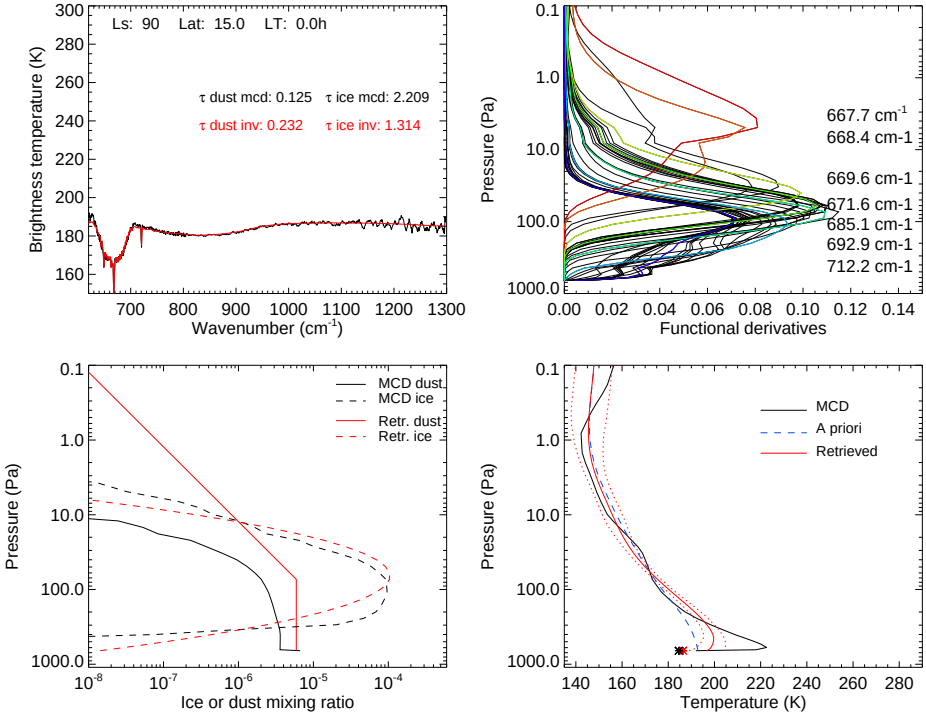

Figure A2. Example of a synthetic retrieval for a MCD scenario extracted at L_s=90°, latitude 15°N, local time 0 AM. Top, left: Synthetic TIRVIM spectrum (black) along with the best fit (red). Retrieved and MCD dust and water ice cloud integrated opacities are indicated. Top, right: functional derivatives of the temperature as a function of pressure, for 50 different wavenumbers within the CO_2 band. Several wavenumbers are highlighted in different colors and labeled. Bottom, left: Mixing ratio vertical profiles for dust and water ice as taken from the MCD (used to generate the synthetic spectrum, in black) and as derived from the retrieval process (in red ; note that only a scaling factor to a generic *a priori* profile is retrieved). Bottom, right: Temperature vertical profile from the MCD (used to generate the synthetic spectrum, in black), *a priori* profile built from the synthetic spectrum (dashed blue line), and retrieved profile (red), with error bars in red dotted lines. The black and red stars stand for the MCD and retrieved surface temperature, respectively.

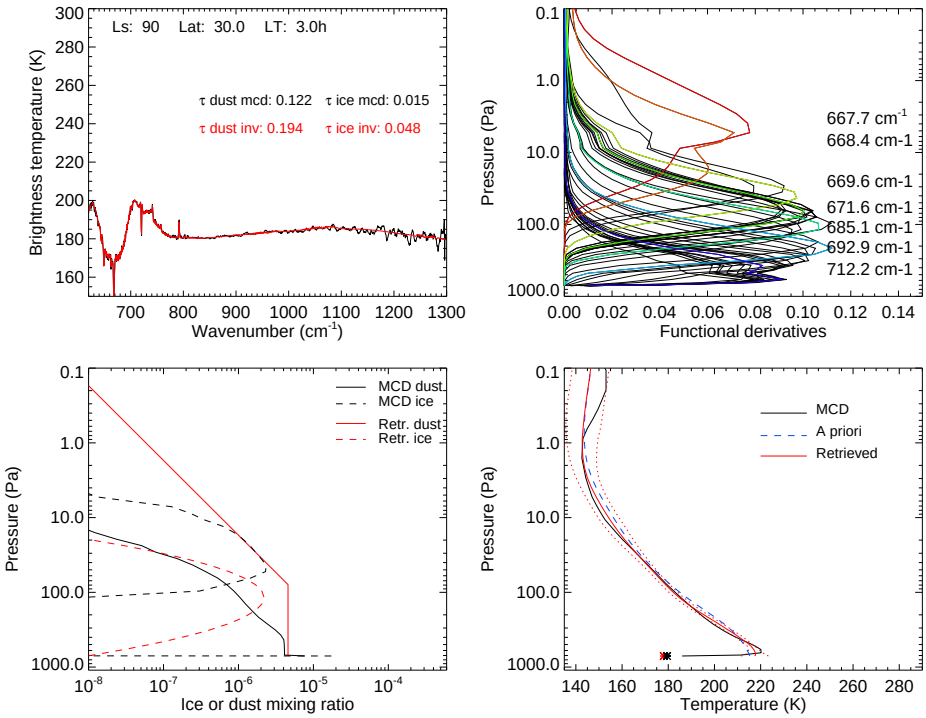

Figure A3. Same as figure A2 for L_s=90°, latitude 30°N, local time 3 AM.

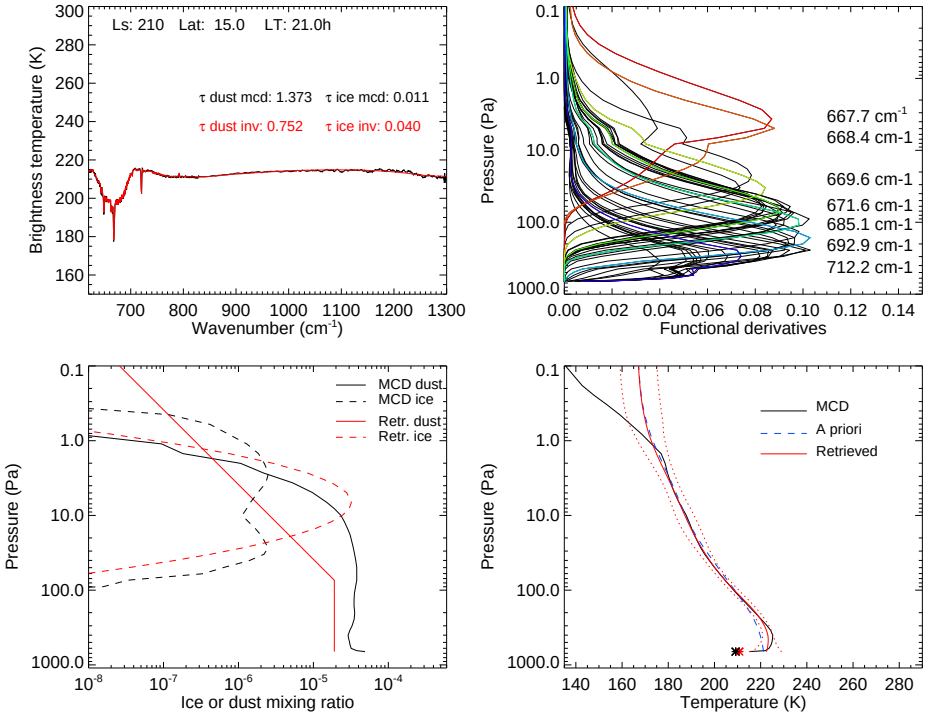

Figure A4. Same as figure A2 for L_s=210°, latitude 15°N, local time 9PM.

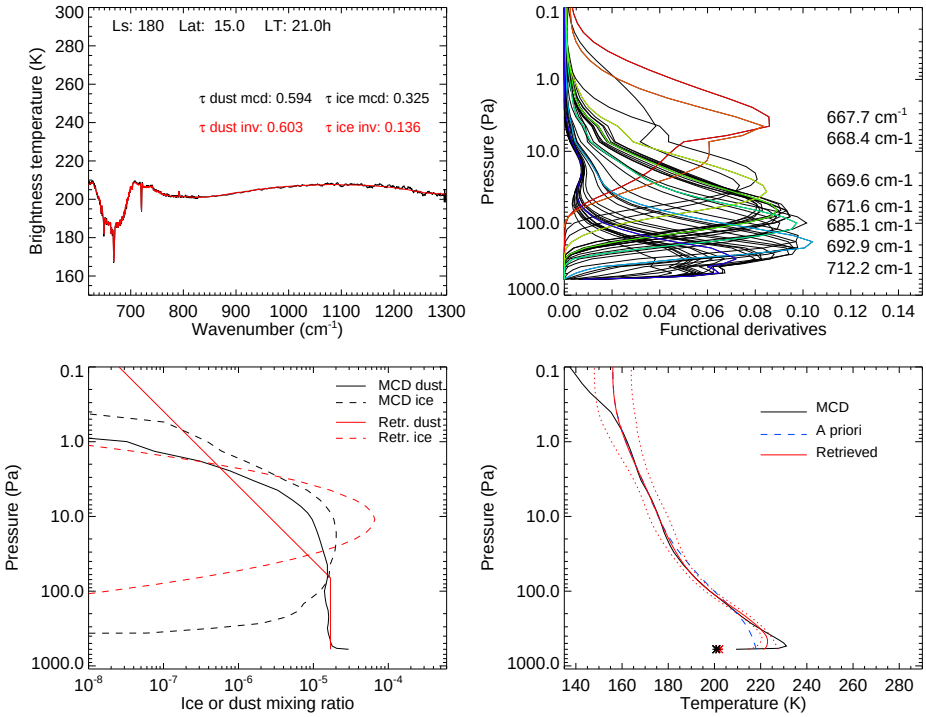

Figure A5. Same as figure A2 for L_s=180°, latitude 15°N, local time 9 PM.

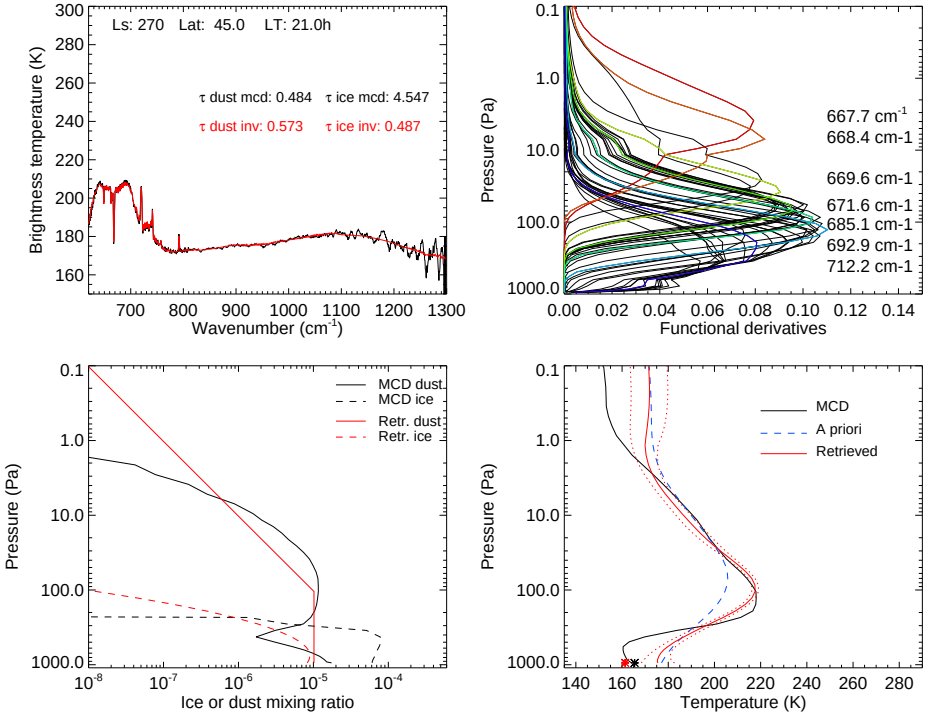

Figure A6. Same as figure A2 for a MCD scenario extracted at L_s=270°, latitude 45°N, local time 9PM.

A3 Impact of using different *a priori* temperature profile

Figure A7 demonstrates the good performance of our atmospheric temperature retrieval starting from different *a priori* profiles (see main text).

A4 TIRVIM-MCS comparison

Figure A8 shows the comparison of the zonally-averaged temperature near 3 AM and 3 PM as retrieved from TIRVIM and MCS, for co-located measurements. No averaging kernels were applied to the MCS temperature fields in that figure.

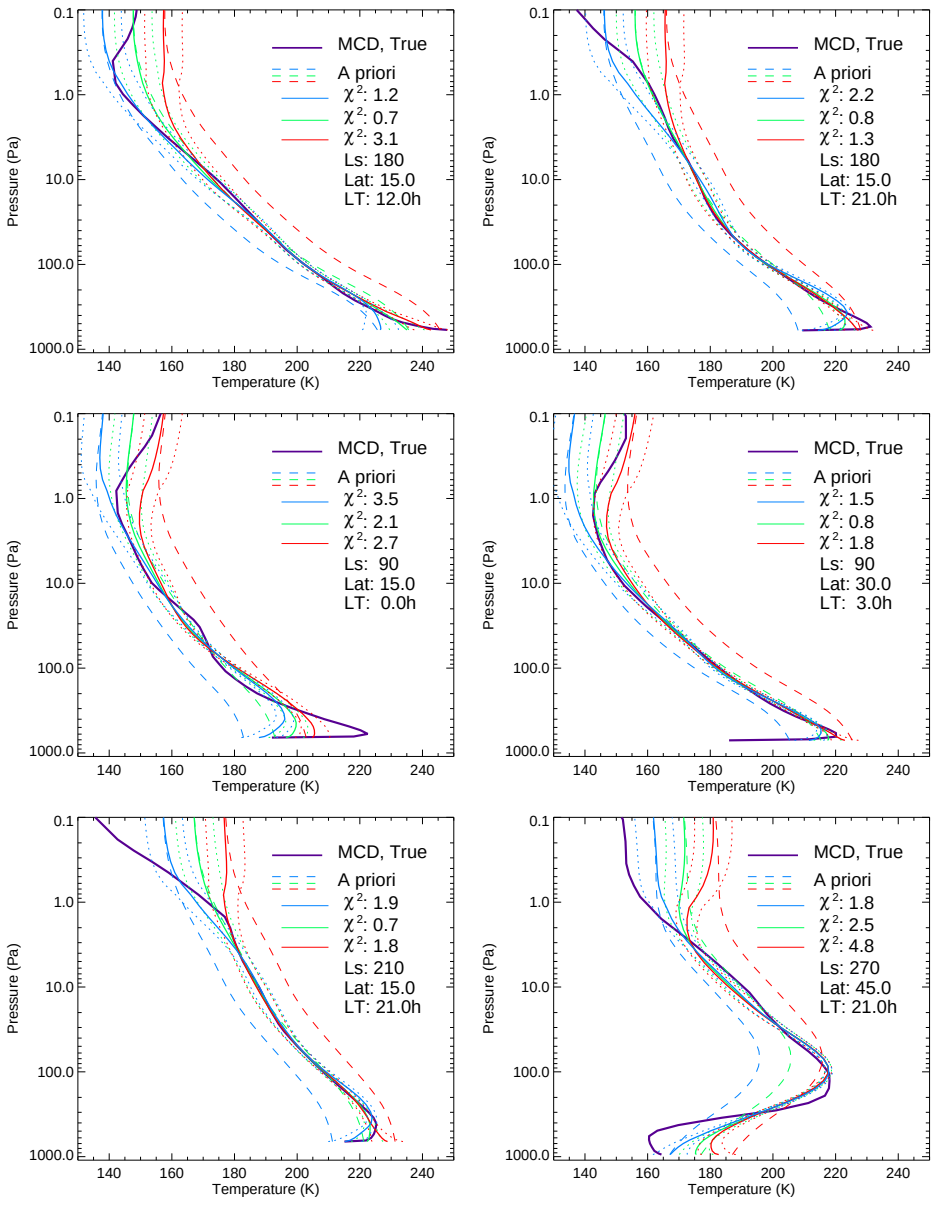

Figure A7. Retrieved temperature profiles (color solid lines) for different *a priori* profiles (dashed lines): in green our nominal *a priori* profile, in blue a profile 10K colder, in red a profile 10K warmer. Quality of the fit in the CO_2 band is given for each retrieval (χ^2). These examples are for 6 out of 440 synthetic retrievals of our OSSE, with the "true" temperature profile in purple. Dotted lines highlight the error envelops. These six examples correspond to local times, season and latitudes of the cases shown in Figures 6 and A2–A6.

Figure A8. Pressure-latitude sections of the zonally-averaged temperature near 3 AM (top) and 3 PM (bottom). Left panels correspond to TIRVIM, middle panels to MCS and the difference between the two datasets is shown in the right panels. Bins of 5° wide in latitude are used. In each panel, these zonal averages are performed for a subset of the available TIRVIM and MCS data that meet co-located criteria, as described in the text.

Acknowledgments

ExoMars is a space mission of ESA and Roscosmos. The ACS experiment is led by IKI, the Space Research Institute in Moscow, assisted by LATMOS in France. This work, exploiting ACS/TIRVIM data, acknowledges funding by CNES. The science operations of ACS are funded by Roscosmos and ESA. RMBY acknowledges support from UAE University grants G00003322 and G00003590. ACS/TIRVIM team at IKI acknowledges the subsidy of the Ministry of Science and High Education of Russia. We warmly thank Michael Smith and another anonymous reviewer for their thorough review of our manuscript. The processed TIRVIM data and retrievals used in this paper are available in NetCDF format on the Institut Pierre Simon Laplace (IPSL) data server, see: https://doi.org/10.14768/ab765eba-0c1d-47b6-97d6-6390c63f0197

References

Bandfield, J. L., & Smith, M. D. (2003, January). Multiple emission angle surface-atmosphere separations of thermal emission spectrometer data. *Icarus*, *161*(1), 47-65. doi: 10.1016/S0019-1035(02)00025-8

Banfield, D., Conrath, B. J., Smith, M. D., Christensen, P. R., & Wilson, R. J. (2003, February). Forced waves in the martian atmosphere from MGS TES nadir data. *Icarus*, *161*(2), 319-345. doi: 10.1016/S0019-1035(02)00044-1

Clancy, R. T., Wolff, M. J., & Christensen, P. R. (2003, September). Mars aerosol studies with the MGS TES emission phase function observations: Optical depths, particle sizes, and ice cloud types versus latitude and solar longitude. *Journal of Geophysical Research (Planets)*, *108*(E9), 5098. doi: 10.1029/2003JE002058

Conrath, B. J., Pearl, J. C., Smith, M. D., Maguire, W. C., Christensen, P. R., Dason, S., & Kaelberer, M. S. (2000, April). Mars Global Surveyor Thermal Emission Spectrometer (TES) observations: Atmospheric temperatures during aerobraking and science phasing. *J. Geophys. Res.*, *105*(E4), 9509-9519. doi: 10.1029/1999JE001095

Forbes, J. M., Zhang, X., Forget, F., Millour, E., & Kleinböhl, A. (2020, September). Solar Tides in the Middle and Upper Atmosphere of Mars. *Journal of Geophysical Research (Space Physics)*, *125*(9), e2020JA028140. doi: 10.1029/2020JA028140

Forget, F., Spiga, A., Dolla, B., Vinatier, S., Melchiorri, R., Drossart, P., ... Gondet, B. (2007, August). Remote sensing of surface pressure on Mars with the Mars Express/OMEGA spectrometer: 1. Retrieval method. *Journal of Geophysical Research (Planets)*, *112*(E8), E08S15. doi: 10.1029/2006JE002871

Fouchet, T., Lellouch, E., Ignatiev, N. I., Forget, F., Titov, D. V., Tschimmel, M., ... Encrenaz, T. (2007, September). Martian water vapor: Mars Express PFS/LW observations. *Icarus*, *190*(1), 32-49. doi: 10.1016/j.icarus.2007.03.003

Giuranna, M., Formisano, V., Biondi, D., Ekonomov, A., Fonti, S., Grassi, D., ... Zasova, L. (2005, August). Calibration of the Planetary Fourier Spectrometer long wavelength channel. *planss*, *53*(10), 993-1007. doi:

10.1016/j.pss.2005.02.007

Giuranna, M., Wolkenberg, P., Grassi, D., Aronica, A., Aoki, S., Scaccabarozzi, D., ... Formisano, V. (2021, January). The current weather and climate of Mars: 12 years of atmospheric monitoring by the Planetary Fourier Spectrometer on Mars Express. *Icarus*, *353*, 113406. doi: 10.1016/j.icarus.2019.113406

Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., ... Zak, E. J. (2017, December). The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, *203*, 3-69. doi: 10.1016/j.jqsrt.2017.06.038

Grassi, D., Ignatiev, N. I., Zasova, L. V., Maturilli, A., Formisano, V., Bianchini, G. A., & Giuranna, M. (2005, August). Methods for the analysis of data from the Planetary Fourier Spectrometer on the Mars Express Mission. *Planet. Space Sci.*, *53*(10), 1017-1034. doi: 10.1016/j.pss.2005.01.006

Greybush, S. J., Gillespie, H. E., & Wilson, R. J. (2019, January). Transient eddies in the TES/MCS Ensemble Mars Atmosphere Reanalysis System (EMARS). *Icarus*, *317*, 158-181. doi: 10.1016/j.icarus.2018.07.001

Guerlet, S. (2021). *ACS/TIRVIM temperature and aerosol retrievals [Dataset].* `https://doi.org/10.14768/ab765eba-0c1d-47b6-97d6-6390c63f0197`. doi: 10.14768/ab765eba-0c1d-47b6-97d6-6390c63f0197

Guzewich, S. D., & Smith, M. D. (2019, February). Seasonal Variation in Martian Water Ice Cloud Particle Size. *J. Geophys. Res.*, *124*(2), 636-643. doi: 10.1029/2018JE005843

Guzewich, S. D., Talaat, E. R., & Waugh, D. W. (2012, March). Observations of planetary waves and nonmigrating tides by the Mars Climate Sounder. *Journal of Geophysical Research (Planets)*, *117*(E3), E03010. doi: 10.1029/2011JE003924

Heavens, N. G., Richardson, M. I., Kleinböhl, A., Kass, D. M., McCleese, D. J., Abdou, W., ... Wolkenberg, P. M. (2011, April). The vertical distribution of dust in the Martian atmosphere during northern spring and summer: Observations by the Mars Climate Sounder and analysis of zonal average vertical dust profiles. *J. Geophys. Res.*, *116*(E4), E04003. doi: 10.1029/2010JE003691

Jacquinet-Husson, N., Armante, R., Scott, N. A., Chédin, A., Crépeau, L., Boutammine, C., ... Makie, A. (2016, September). The 2015 edition of the GEISA

spectroscopic database. *Journal of Molecular Spectroscopy*, *327*, 31-72. doi: 10.1016/j.jms.2016.06.007

Kahre, M. A., Murphy, J. R., Newman, C. E., Wilson, R. J., Cantor, B. A., Lemmon, M. T., & Wolff, M. J. (2017). The Mars Dust Cycle. In R. M. Haberle, R. T. Clancy, F. Forget, M. D. Smith, & R. W. Zurek (Eds.), *The Atmosphere and Climate of Mars* (p. 229-294). doi: 10.1017/9781139060172.010

Kleinböhl, A., Friedson, A. J., & Schofield, J. T. (2017, January). Two-dimensional radiative transfer for the retrieval of limb emission measurements in the martian atmosphere. *J. Quant. Spectrosc. Radiat. Transf.*, *187*, 511-522. doi: 10.1016/j.jqsrt.2016.07.009

Kleinböhl, A., Schofield, J. T., Abdou, W. A., Irwin, P. G. J., & de Kok, R. J. (2011, July). A single-scattering approximation for infrared radiative transfer in limb geometry in the Martian atmosphere. *J. Quant. Spectrosc. Radiat. Transf.*, *112*, 1568-1580. doi: 10.1016/j.jqsrt.2011.03.006

Kleinböhl, A., Schofield, J. T., Kass, D. M., Abdou, W. A., Backus, C. R., Sen, B., ... McCleese, D. J. (2009, October). Mars Climate Sounder limb profile retrieval of atmospheric temperature, pressure, and dust and water ice opacity. *J. Geophys. Res.*, *114*(E10), E10006. doi: 10.1029/2009JE003358

Kleinböhl, A., Spiga, A., Kass, D. M., Shirley, J. H., Millour, E., Montabone, L., & Forget, F. (2020, January). Diurnal Variations of Dust During the 2018 Global Dust Storm Observed by the Mars Climate Sounder. *Journal of Geophysical Research (Planets)*, *125*(1), e2019JE006115. doi: 10.1029/2019JE006115

Kleinböhl, A., Wilson, J. R., Kass, D., Schofield, J. T., & McCleese, D. J. (2013, May). The semidiurnal tide in the middle atmosphere of Mars. *Geophys. Res. Lett.*, *40*(10), 1952-1959. doi: 10.1002/grl.50497

Korablev, O., Montmessin, F., Trokhimovskiy, A., Fedorova, A. A., Shakun, A. V., Grigoriev, A. V., ... Zorzano, M. P. (2018, February). The Atmospheric Chemistry Suite (ACS) of Three Spectrometers for the ExoMars 2016 Trace Gas Orbiter. *Space Sci. Rev.*, *214*(1), 7. doi: 10.1007/s11214-017-0437-6

Lamouroux, J., Gamache, R. R., Laraia, A. L., Hartmann, J.-M., & Boulet, C. (2012, August). Semiclassical calculations of half-widths and line shifts for transitions in the 30012←00001 and 30013←00001 bands of CO_2. III: Self collisions. *J. Quant. Spectrosc. Radiat. Transf.*, *113*(12), 1536-1546. doi:

10.1016/j.jqsrt.2012.03.035

Lee, C., Lawson, W. G., Richardson, M. I., Heavens, N. G., Kleinböhl, A., Banfield, D., ... Toigo, A. D. (2009, March). Thermal tides in the Martian middle atmosphere as seen by the Mars Climate Sounder. *J. Geophys. Res.*, *114*(E3), E03005. doi: 10.1029/2008JE003285

Luginin, M., Fedorova, A., Ignatiev, N., Trokhimovskiy, A., Shakun, A., Grigoriev, A., ... Korablev, O. (2020, November). Properties of Water Ice and Dust Particles in the Atmosphere of Mars During the 2018 Global Dust Storm as Inferred From the Atmospheric Chemistry Suite. *Journal of Geophysical Research (Planets)*, *125*(11), e2020JE006419. doi: 10.1029/2020JE006419

McCleese, D. J., Heavens, N. G., Schofield, J. T., Abdou, W. A., Bandfield, J. L., Calcutt, S. B., ... Zurek, R. W. (2010, December). Structure and dynamics of the Martian lower and middle atmosphere as observed by the Mars Climate Sounder: Seasonal variations in zonal mean temperature, dust, and water ice aerosols. *Journal of Geophysical Research (Planets)*, *115*(E12), E12016. doi: 10.1029/2010JE003677

Millour, E., Forget, F., Spiga, A., Vals, M., Zakharov, V., Montabone, L., ... MCD Development Team (2018, February). The Mars Climate Database (version 5.3). In *From Mars Express to ExoMars* (p. 68).

Montabone, L., Forget, F., Millour, E., Wilson, R. J., Lewis, S. R., Cantor, B., ... Wolff, M. J. (2015, May). Eight-year climatology of dust optical depth on Mars. *Icarus*, *251*, 65-95. doi: 10.1016/j.icarus.2014.12.034

Montabone, L., Spiga, A., Kass, D. M., Kleinböhl, A., Forget, F., & Millour, E. (2020, August). Martian Year 34 Column Dust Climatology from Mars Climate Sounder Observations: Reconstructed Maps and Model Simulations. *Journal of Geophysical Research (Planets)*, *125*(8), e06111. doi: 10.1029/2019JE006111

Navarro, T., Forget, F., Millour, E., & Greybush, S. J. (2014, October). Detection of detached dust layers in the Martian atmosphere from their thermal signature using assimilation. *Geophys. Res. Lett.*, *41*(19), 6620-6626. doi: 10.1002/2014GL061377

Pankine, A. A., Tamppari, L. K., Bandfield, J. L., McConnochie, T. H., & Smith, M. D. (2013, September). Retrievals of martian atmospheric

opacities from MGS TES nighttime data. *Icarus*, *226*(1), 708-722. doi: 10.1016/j.icarus.2013.06.024

Perrin, M. Y., & Hartmann, J. M. (1989, October). Temperature-dependent measurements and modeling of absorption by CO_2-N_2 mixtures in the far line-wings of the 4.3 μm CO_2 band. *J. Quant. Spectrosc. Radiat. Transf.*, *42*, 311-317. doi: 10.1016/0022-4073(89)90077-0

Revercomb, H. E., Buijs, H., Howell, H. B., LaPorte, D. D., Smith, W. L., & Sromovsky, L. A. (1988, August). Radiometric calibration of IR Fourier transform spectrometers: solution to a problem with the High-Resolution Interferometer Sounder. *Applied Opt.*, *27*(15), 3210-3218. doi: 10.1364/AO.27.003210

Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice*. doi: 10.1142/3171

Rodgers, C. D., & Connor, B. J. (2003, February). Intercomparison of remote sounding instruments. *J. Geophys. Res.*, *108*(D3), 4116. doi: 10.1029/2002JD002299

Smith, M. D. (2004, January). Interannual variability in TES atmospheric observations of Mars during 1999-2003. *Icarus*, *167*(1), 148-165. doi: 10.1016/j.icarus.2003.09.010

Smith, M. D. (2009, August). THEMIS observations of Mars aerosol optical depth from 2002-2008. *Icarus*, *202*(2), 444-452. doi: 10.1016/j.icarus.2009.03.027

Smith, M. D. (2019a, November). Local time variation of water ice clouds on Mars as observed by THEMIS. *Icarus*, *333*, 273-282. doi: 10.1016/j.icarus.2019.06.009

Smith, M. D. (2019b, November). THEMIS Observations of the 2018 Mars Global Dust Storm. *Journal of Geophysical Research (Planets)*, *124*(11), 2929-2944. doi: 10.1029/2019JE006107

Smith, M. D., Conrath, B. J., Pearl, J. C., & Christensen, P. R. (2002, May). NOTE: Thermal Emission Spectrometer Observations of Martian Planet-Encircling Dust Storm 2001A. *Icarus*, *157*(1), 259-263. doi: 10.1006/icar.2001.6797

Smith, M. D., Pearl, J. C., Conrath, B. J., & Christensen, P. R. (2000, April). Mars Global Surveyor Thermal Emission Spectrometer (TES) observations of dust opacity during aerobraking and science phasing. *J. Geophys. Res.*, *105*(E4),

9539-9552. doi: 10.1029/1999JE001097

Smith, M. D., Pearl, J. C., Conrath, B. J., & Christensen, P. R. (2001, January). One Martian year of atmospheric observations by the thermal emission spectrometer. *Geophys. Res. Lett.*, *28*(22), 4263-4266. doi: 10.1029/2001GL013608

Steele, L. J., Lewis, S. R., & Patel, M. R. (2014, July). The radiative impact of water ice clouds from a reanalysis of Mars Climate Sounder data. *Geophys. Res. Lett.*, *41*(13), 4471-4478. doi: 10.1002/2014GL060235

Warren, S. G. (1984, April). Optical constants of ice from the ultraviolet to the microwave. *Applied Opt.*, *23*, 1206-1225. doi: 10.1364/AO.23.001206

Wilson, R. J. (2000, December). Evidence for diurnal period Kelvin waves in the Martian atmosphere from Mars Global Surveyor TES data. *Geophys. Res. Lett.*, *27*(23), 3889-3892. doi: 10.1029/2000GL012028

Wolff, M. J., & Clancy, R. T. (2003, September). Constraints on the size of Martian aerosols from Thermal Emission Spectrometer observations. *Journal of Geophysical Research (Planets)*, *108*(E9), 5097. doi: 10.1029/2003JE002057

Wolff, M. J., Smith, M. D., Clancy, R. T., Arvidson, R., Kahre, M., Seelos, F., ... Savijärvi, H. (2009, June). Wavelength dependence of dust aerosol single scattering albedo as observed by the Compact Reconnaissance Imaging Spectrometer. *Journal of Geophysical Research (Planets)*, *114*(E9), E00D04. doi: 10.1029/2009JE003350

Wolkenberg, P., & Giuranna, M. (2021, January). Daily dust variation from the PFS MEx observations. *Icarus*, *353*, 113823. doi: 10.1016/j.icarus.2020.113823

Wolkenberg, P., Giuranna, M., Grassi, D., Aronica, A., Aoki, S., Scaccabarozzi, D., & Saggin, B. (2018, August). Characterization of dust activity on Mars from MY27 to MY32 by PFS-MEX observations. *Icarus*, *310*, 32-47. doi: 10.1016/j.icarus.2017.10.045