

Multi-spectrum retrieval of Venus IR surface emissivity maps from VIRTIS/VEX nightside measurements at Themis Regio

David Kappel^{a,b,*}, Gabriele Arnold^a, Rainer Haus^c

^aGerman Aerospace Center (DLR), Institute of Planetary Research, Rutherfordstrasse 2, 12489 Berlin, Germany

^bUniversity of Potsdam, Institute of Physics and Astronomy, Karl-Liebknecht-Straße 24/25, 14476 Potsdam-Golm, Germany

^cWestfälische Wilhelms-Universität Münster, Institute for Planetology, Wilhelm-Klemm-Straße 10, 48149 Münster, Germany

Abstract

Surface emissivity maps in the infrared can contribute to explore Venus' geology. Nightside radiance spectra at Themis Regio acquired by the IR mapping channel of the Visible and InfraRed Thermal Imaging Spectrometer (VIRTIS-M-IR) aboard Venus EXpress (VEX) are used to derive emissivity data from the three accessible spectral surface windows at 1.02, 1.10, and 1.18 μm . The measured spectra are simulated by applying a full radiative transfer model. Neglecting geologic activity, a multi-spectrum retrieval algorithm is utilized to determine the emissivity maps of the surface target as parameter vectors that are common to many spectrally resolved images that cover this target.

Absolute emissivity values are difficult to obtain due to strong interferences from other parameters. The true emissivity mean of the target cannot be retrieved, nor can the emissivity mean of a retrieved map be strictly preset. The retrieved map can exhibit trends with latitude and topography that are probably artificial. Once the trends have been removed in a post-processing step, it can be observed that the magnitude of the resulting spatial emissivity fluctuations around their mean value increases with increasing mean value. A linear transformation is applied that converts the detrended map to exhibit a defined emissivity mean value called reference emissivity, here 0.5, yielding the 'renormalized emissivity map' with accordingly transformed fluctuations. It is verified that renormalized emissivity maps are largely independent of the emissivity mean before renormalization, of modifications to interfering atmospheric, surface, and instrumental parameters, and of selected details of the retrieval pipeline and data calibration and preprocessing. Extremely large emissivity retrieval errors due to imperfect or unconsidered forward model parameters are effectively avoided. If the absolute emissivity at a given bin of the target were known, the absolute emissivity map of the entire target could be computed according to the mentioned transformation, assuming absent true trends with latitude and topography. Until then, the renormalized emissivities are interpreted as spatial variations relative to the reference emissivity. They represent an important step toward the retrieval of absolute emissivities.

Renormalized emissivity maps of Themis Regio at the three surface windows are determined from 64 measurement repetitions. Retrieval errors are estimated by a statistical evaluation of maps derived from various disjoint selections of spectra and using different assumptions on the interfering parameters. Double standard deviation errors for the three surface windows amount to 3%, 8%, and 4%, respectively, allowing geologic interpretation. A comparison to results from an earlier error analysis based on synthetic spectra shows that unconsidered time variations of interfering atmospheric parameters are a major error source. Spatial variations of the 1.02 μm surface emissivity of 20% that correspond to the difference between unweathered granitic and basaltic rocks would be easily detectable, but such variations are ruled out for the studied target area. Emissivity anomalies of up to 8% are detected at both 1.02 and 1.18 μm . At present sensitivity, no anomalies are identified at 1.10 μm , but anomalies exceeding the determined error level can be excluded. With single standard deviation significance, all three maps show interesting spatial emissivity variations.

Keywords: Venus, surface; Infrared observations; Radiative transfer

*Corresponding Author

Email address: David.Kappel@dlr.de, dkappel@gmx.net (David Kappel)

1. Introduction

Near-infrared (NIR) spectroscopy is a powerful tool to study composition, texture, and grain size of the surface material of planetary objects. Venus is a geologically interesting target, but little is known about the global surface composition, since the hot dense atmosphere complicates NIR remote sensing. On Venus' dayside, the surface is not observable at NIR wavelengths. The available radar data acquired by the Magellan mission (global maps of topography, slope, reflectivity, and emissivity, [Pettengill et al., 1991](#); [Ford and Pettengill, 1992](#)) are less diagnostic for surface composition analyses.

Due to the high surface temperatures, thermal radiation from the surface and low atmosphere dominates NIR spectra of Venus' nightside. Compositional information about the surface is given by its emissivity. Numerous diagnostic absorption bands for minerals and rocks appear within the NIR. Spectra of rock-forming iron-rich mafic minerals like olivine and pyroxenes display characteristic bands at 1 and between 2 and 2.5 μm . Generally, these features tend to increase the spectral emissivity at these wavelengths compared to the spectral continuum. Bright felsic minerals have mostly featureless spectra and low NIR emissivities. In contrast, weathering products like iron oxides or iron sulfides are characterized by featureless spectra of high emissivity. Thus, maps of NIR emissivity data can provide useful information about the nature of Venusian surface material on a global scale. The present work discusses the remote sensing of such maps.

Radiance in the infrared originating from Venus' surface on the dayside is far outweighed (typically by more than three orders of magnitude) by sunlight scattered back from the cloud tops, and surface information cannot be obtained there. The present work focuses on the nightside, therefore. The surface emits radiation that depends on surface temperature (and thus topography) and emissivity. As it propagates upward, it gets mixed with radiation resulting from thermal emissions by the atmospheric gases. In the dense atmosphere of Venus, radiation is multiply scattered and mostly absorbed by atmospheric gases and clouds, and partly reflected back from the surface. Much of the radiation's surface information is destroyed by these processes. Multiple scattering also leads to a spatial blurring of the surface information on a horizontal scale of 100 km ([Moroz, 2002](#)). Only a few narrow spectral transparency windows allow strongly modulated surface radiation to escape into space. This radiation was recorded by the InfraRed Mapping channel of the Visible and InfraRed Thermal Imaging Spectrometer (VIRTIS-M-IR) aboard Venus EXpress (VEX) ([Drossart et al., 2007](#); [Piccioni et al., 2007](#); [Arnold et al., 2012](#)). VIRTIS-M-IR acquired spectrally resolved images in the range 1.02–5.1 μm . Each spatial pixel provided a radiance spectrum sampling this range with 432 spectral bands. These measurements are currently the only hyperspectral data that allow for a reliable derivation of IR surface emissivity information on global scales. In the VIRTIS-M-IR spectral range, the only parts of the acquired radiance spectra that carry surface information are the spectral peaks at 1.02, 1.10, and 1.18 μm , therefore called the surface window peaks. The fraction of radiance that originates from the surface is smaller in the latter two peaks ([Haus and Arnold, 2010](#), Fig. 5), which are consequently less sensitive to surface information than the first one.

It has been demonstrated by [Helbert et al. \(2008\)](#) and [Müller et al. \(2008\)](#) that surface information at 1.02 μm can be extracted from VIRTIS-M-IR measurements to some extent in the form of the time-averaged so-called thermal flux anomaly. This is the thermal flux from Venus' surface that is not due to topography-related surface temperature. It is correlated to surface emissivity. The cited authors applied semi-empirical approaches to de-cloud the 1.02 μm VIRTIS-M-IR radiances, with the small surface-unaffected 1.31 μm radiance peak as reference for cloud absorption, and to remove the influence of surface topography. This method utilizes only two spectral bands and is computationally very fast. The authors were therefore able to process all suitable VIRTIS-M-IR nightside spectra. Based on these results and on gravity field data, [Smrekar et al. \(2010\)](#) were able to identify candidates for recent hotspot volcanism. By correlating the thermal flux anomalies and topography data, [Müller et al. \(2012b\)](#) have also identified a possible variation in Venus' rotation period that may be due to long-term changes in the atmospheric circulation. Furthermore, constraints on the planet's volcanism rate ([Müller et al., 2012a](#)) were derived based on the semi-empirical approach to evaluate VIRTIS-M-IR data.

A disadvantage of the method of [Helbert et al. \(2008\)](#) and [Müller et al. \(2008\)](#) is that the small 1.31 μm radiance is very susceptible to measurement noise and data calibration and preprocessing errors. Moreover, the varying characteristics of the spectral instrumental response function and spectral band-to-wavelength mapping cannot be taken into account, and the meaning of the flux anomaly map with respect to emissivity and its reliability are not entirely clear in this otherwise very successful approach. Based on these difficulties, it was an open question, whether surface emissivity data can be obtained from the 1.02 μm window on the basis of a full radiative transfer simulation that takes into account

the extreme environmental conditions in Venus' deep atmosphere in combination with a retrieval algorithm that respects available *a priori* information. The present work demonstrates that this is possible. In addition, emissivity data from the other two surface windows are determined. Moreover, the reliability of the determined emissivity data from all three windows is quantified.

Starting with an initial parameter set describing defined states of atmosphere, surface, and instrument, it is possible to simulate VIRTIS-M-IR spectra by numerically solving the radiative transfer equation, taking into account thermal emissions by surface and atmosphere, and absorption and multiple scattering by gases and clouds (Haus and Arnold, 2010). Some of the initial parameters can be iteratively modified to improve the fit of the simulated to a given measured spectrum. The parameters that result in the best fit (i.e. the 'retrieved parameters') can be interpreted to describe the actual atmospheric, surface, and instrumental states that led to the measured spectrum. In the present work, the spectral range 1.02–2.3 μm is utilized to determine in addition to emissivity a number of auxiliary parameters like the cloud abundance that are required to obtain meaningful emissivity values.

This inversion of the radiative transfer equation is mathematically ill-posed, since different surface and atmospheric parameters can describe the same spectrum equally well. The retrieval has to be regularized, therefore. Moreover, the emissivity retrieval error analysis by Kappel et al. (2015) that was based on synthetic spectra showed that the retrieval of emissivity from a single spectrum must fail, even when taking into account Bayesian *a priori* information in the form of expected mean values and standard deviations for the parameters to be retrieved, along with measurement and simulation error information. The reason is that there are interfering parameters like the cloud bottom altitude that cannot be sufficiently reliably derived from the measured spectrum itself but have to be set to certain assumed values to enable the radiative transfer simulation.

Since the single-spectrum information content is thus obviously too low, several spectra have to be utilized to derive surface data. This can be done by applying the recently developed Multi-Spectrum Retrieval algorithm MSR (Kappel, 2014). MSR allows for the additional incorporation of expected correlation lengths and times as *a priori* information for the retrieval parameters of several spectra. It can also retrieve surface emissivity as parameter that is common to several spectra that cover the same surface spot. This way, the probability of retrieving unreasonable parameter values and unlikely spatial-temporal parameter distributions is decreased, and the reliability of retrieved surface emissivity can be enhanced.

Kappel et al. (2015) also showed that the retrieval of absolute emissivity values produces extremely large errors even when applying MSR. The authors proposed to rather determine emissivity maps relative to a certain initial emissivity value. This way, large emissivity errors, for example due to differences between the spatial distributions of time-averaged assumed and true interfering parameters, can be avoided. However, it must then be clarified, how the relative emissivity maps have to be interpreted. The present work refines the concept of relative emissivity maps.

The question now is, whether the magnitude of the relative emissivities' errors is finally small enough to allow for a geologic interpretation of the derived maps. Or in other words, how well do the retrieved maps represent the true ones? No ground truth within the surface coverage suitable for MSR is available so far. To answer this question, the first main objective of the present work is the exemplary determination of emissivity data maps at 1.02, 1.10, and 1.18 μm of a selected surface target, and the second one is an estimate of the reliability of these maps. It will be shown that the relative maps are reasonably independent of the selection of the measurement data set, the initial emissivity value as well as of other interfering parameters. A geologic interpretation of the determined maps is beyond the scope of the present work. It would require additional information to be taken into account like the Magellan radar reflectivity, emissivity, and slope data, geologic maps derived from the Magellan data, as well as context from relative infrared emissivity maps of further target areas to study large-scale trends. A detailed geologic discussion will be given in a follow-up paper.

In contrast to the very fast semi-empirical method by Müller et al. (2008), the present approach, as any retrieval based on full radiative transfer simulations, requires considerable computer resources and is first selectively applied to targets that are beforehand identified to be of special geological interest. Since the estimate of the reliability of the determined maps in this work is partly based on the comparison of maps derived from disjoint measurement data sets, it is also important that the target is located in an area that exhibits as many as possible suitable VIRTIS-M-IR measurements. Both conditions are fulfilled at Themis Regio, the western part of which is selected as target area (35–47 °S, 270–288 °E) to demonstrate the present approach.

Themis Regio is a highland with a diameter of about 2000 km, located in the southern hemisphere of Venus. It is covered by the USGS quadrangle V-53 extending from 25–50 °S and from 270–300 °E (Stofan and Brian, 2012).

Themis Regio has an average height of about 500 m. It is characterized by a corona-dominated hotspot rise related to a long-duration, non-simultaneous, small-scale upwelling (Stofan et al., 1995). Themis Regio contains significant extensional deformation (Stofan and Brian, 2012). Gravity data and the topographic swell suggest that this region is likely underlain by an active plume with ongoing surface deformation due to growth of the rise (Smrekar and Stofan, 1999). The target area used for this paper includes geologic structures like the Shiwanokia Corona (42.0 °S, 279.8 °E, see Hayward (2015) for the nomenclature and Fig. 3d for a topography map of the target area) with steep sided domes, Shulamite Corona (38.8 °S, 284.3 °E), Abeona Mons (44.8 °S, 273.1 °E), Mertseger Mons (38.1 °S, 270.3 °E), impact craters like Kenny (44.4 °S, 271.1 °E), Aksentyeva (42.0 °S, 271.9 °E), and Sabin (38.5 °S, 274.7 °E) with a dark parabola, and wrinkle ridges. Former work suggests that some edifice fields in Themis Regio correlate with high thermal flux anomalies at 1.02 μm extracted from VIRTIS-M-IR/VEX measurements (Müller et al., 2008) and attributed to relatively stratigraphically young units of volcanic origin (Stofan et al., 2010). Therefore, Themis Regio is a geologically interesting target area, and the cited thermal flux data can be compared to real emissivity values derived here at 1.02, 1.10, and 1.18 μm .

The present paper is a direct continuation of a research line that aims at the determination of emissivity data from VIRTIS-M-IR measurements based on radiative transfer simulations. This line was started by Arnold et al. (2008) who developed a quick-look method to identify emissivity anomalies. This method was corroborated by first radiative transfer simulations. The next contribution was a detailed description of a radiative transfer simulation model developed for Venus studies (Haus and Arnold, 2010). A number of refinements in the VIRTIS-M-IR data calibration and preprocessing required for the determination of surface emissivity were presented by Kappel et al. (2012). Since the derivation of surface information from single spectra is prone to extremely large errors, the multi-spectrum retrieval algorithm MSR was developed that allows for taking the spectral information content of many VIRTIS-M-IR measurements into account. An early version of MSR and corresponding impacts on the data analysis were already presented by Kappel et al. (2012). The full version of MSR that allows for the retrieval of the emissivity map of a surface target as a parameter vector that is common to several spectrally resolved images covering this target was described at a later time (Kappel, 2014). The expectable single- and multi-spectrum emissivity retrieval errors were then estimated by Kappel et al. (2015) on the basis of synthetic spectra. Even when applying MSR, the retrieval errors were predicted to be too large to obtain useful emissivity data from real VIRTIS-M-IR measurements. It was proposed to rather retrieve relative emissivity maps to achieve a reasonable reliability.

Section 2 of the present paper describes important prerequisites for the retrieval of emissivity maps. First, the VIRTIS-M-IR measurements are selected that are the most suitable for the derivation of emissivity data from the target area (high signal-to-noise ratio, thin to moderate cloud cover, few straylight, high measurement repetition number), and the necessary calibration and preprocessing steps are cited (Section 2.1). Next, the radiative transfer forward model required to simulate VIRTIS-M-IR spectra is explained (Section 2.2). Then, Section 2.3 summarizes the capabilities of the multi-spectrum retrieval algorithm MSR, and it is discussed, why and how the raw retrieved emissivity maps have to be renormalized to obtain well-defined relative emissivity maps. The full MSR retrieval pipeline, which has to anticipate this post-processing, is described in Section 2.4, along with the retrieval parameters and their *a priori* data. Section 3 presents the raw retrieved emissivity maps of Themis Regio, illustrates the post-processing, and resolves the first major objective of this work by showing the final renormalized emissivity maps at 1.02, 1.10, and 1.18 μm . The second objective, the determination of the maps' reliabilities, is addressed in Section 4. For this purpose, results from an earlier retrieval error analysis based on synthetic spectra are first recited (Section 4.1). It is then explained that a statistical analysis of maps retrieved under different conditions from real measurements yields an improved error estimate (Section 4.2). The 'Measurement Selection Tests' (MSTs) detailed in Section 4.3 alter these conditions by retrieving maps from different disjoint selections of measurements, whereas for the 'Parameter Modification Tests' (PMTs, Section 4.4), maps are determined from always the same measurement data set but many important interfering parameters are modified. Based on results from these tests, it is discussed, how retrieved maps can be referred to a given reference emissivity (Section 4.5), and how retrieval errors correspondingly transform. Section 5 discusses the results of the previous sections, provides the double standard deviation uncertainties of the renormalized emissivity maps, and gives a preliminary discussion of the geologic relevance of the determined maps. Summary and conclusion are presented in Section 6.

Appendix A, published as online supplementary material, provides additional information. The utilized spectrally resolved VIRTIS-M-IR images are listed in Appendix A.1. Next, Appendix A.2 presents the mathematical background of the renormalization, its effect on the retrieval errors, and how the MSTs and PMTs can be used to estimate the latter.

Detector-related trends of the VIRTIS-M-IR spectral registration retrieved from the measurements themselves are reported in Appendix A.3. Finally, Appendix A.4 shows figures that illustrate at 1.10 and 1.18 μm the post-processing of the retrieved maps and the MSTs and PMTs, complementing the corresponding figures at 1.02 μm given in the main text.

2. Data selection, forward model, and multi-spectrum retrieval

This section discusses important prerequisites for the retrieval of emissivity maps. First, the selection of the utilized VIRTIS-M-IR measurements is described (Section 2.1). The forward model (Section 2.2) is needed to simulate VIRTIS-M-IR spectra. Parameters can be retrieved from the measurements by applying a multi-spectrum retrieval algorithm (Section 2.4), but the algorithm has to anticipate a suitable post-processing of retrieved emissivity maps to enable meaningful results (Section 2.3).

2.1. Data selection, calibration, and preprocessing

Each Venus Express orbit corresponds to 24 (Earth-)hours (when VIRTIS-M-IR still acquired data). VIRTIS-M-IR data from each orbit are divided into a number of sessions, depending on the science objectives. The data acquired during a session are stored as 'data cube'. Depending on the acquisition mode, the instrument recorded at each exposure at 256 spatial samples a spectrum of 432 bands uniformly dividing the approximate range 1.02–5.1 μm , forming a 'frame'. A data cube is composed of a series of several frames scanning a certain target and thus provides a spectrally resolved two-dimensional spatial image.

Multi-spectrum retrieval of emissivity as parameter common to N_r measurements that cover the same surface bin relies on a self-consistent calibration and data preprocessing. The VIRTIS-M-IR data are calibrated according to Cardesin-Moinelo et al. (2010). Refinements of the calibration in view of the multi-spectrum emissivity retrieval and based on a detailed detector responsivity analysis are applied according to Kappel et al. (2012). They include an improvement of the spatial homogeneity of the detector responsivity leading to a wavelength dependent change of the calibrated radiance in the outer region of the detector in sample direction in the order of about 15%. The refinements also include an adjustment of the responsivity in the spectral dimension resulting in an increase of the calibrated radiance between 0% at 1.3 μm and 35% at 1.0 μm . The detector responsivity analysis also revealed a temperature and spatial sample dependent possible nonlinearity of the detector-count-to-radiance mapping that differently affects the even and the odd spectral bands. This leads to the so-called even-odd effect where a sawtooth-like pattern seems to be superimposed on the spectra. For night spectra in the range 1.0–2.6 μm , this effect can be partially removed. At 3.3 s (≥ 8 s) detector exposure time, this correction can amount to 5% ($\leq 1\%$) of the radiance in the nightside peaks. Note that some but not all of these issues are by now (partially) resolved in the most recent data archive as of 2015 at PSA (Drossart et al., 2013).

Kappel et al. (2012) also described a data preprocessing of the newly calibrated spectra that is required for emissivity retrieval and includes the removal of straylight and the blurring of the topography data. The latter shall emulate the atmospheric blurring of observable surface information caused by the multiple scattering of radiation by gases and clouds. Note that not topography itself is blurred but the fundamental carrier of the observed information, the radiance. An effective radiance in dependence on topography is therefore computed and blurred with respect to a Gaussian with 100 km FWHM. This corresponds to the expected achievable spatial surface resolution in the infrared (Moroz, 2002). The effective blurred topography is computed from the blurred effective radiance (Müller et al., 2008). For a description of the current implementation see Kappel et al. (2012). In order to reflect this limited achievable surface resolution in the emissivity retrieval, to increase the signal-to-noise ratio, to decrease the amount of data to be analyzed by the retrieval procedure, and to obtain spectra that are referred to always the same surface portion, Venus' surface is binned with respect to Lambert equal-area map projections (Bugayevskiy and Snyder, 1995), and the data of each cube are binned accordingly. The surface bins' aspect ratios shall not deviate too much from unity, and thus Lambert cylindrical projection (with standard parallel 30°; no local distortion due to the map projection at this latitude) is used equatorward of 38.68° latitude, and Lambert north/south pole azimuthal projection elsewhere. To match the mentioned surface resolution, each bin covers a surface area of about $100 \times 100 \text{ km}^2$ ($91.5 \times 91.5 \text{ km}^2$ to be exact, allowing for 360 bins in longitude direction in the cylindrical projection). In the following, the term 'measurement' refers to the binned data.

In order to exemplarily demonstrate the intended emissivity retrieval, a target area is selected where as many as possible usable VIRTIS-M-IR measurement repetitions are available. A measurement is said to be usable for this work when the following conditions are fulfilled:

1. surface elevation data is available to avoid the gaps in the Magellan radar topography data (Ford and Pettengill, 1992; Rappaport et al., 1999),
2. the zenith-sun angle is $\geq 98^\circ$ to exclude dayside and twilight measurements,
3. the observation angle is $\leq 60^\circ$ to avoid significant spherical effects in the radiative transfer,
4. the detector exposure time is ≥ 3 s to exclude spectra with very low signal-to-noise ratios in the surface windows,
5. the number of individual spectra contributing to a binned measurement is ≥ 10 to ensure a sufficient statistical weight and high signal-to-noise ratio,
6. the full spatial resolution of 256 spatial samples per image is available to avoid a coarse assignment of the individual spectra to the bins,
7. the spectrally averaged radiance in the range 2.15–2.30 μm is ≥ 3 $\text{mW}/(\text{m}^2 \text{sr } \mu\text{m})$ to exclude measurements with very thick clouds,
8. the removed straylight plus the even-odd-effect correction are ≤ 4 $\text{mW}/(\text{m}^2 \text{sr } \mu\text{m})$ on spectral average in the range 1.0–2.6 μm to avoid spectra with very large corrections and thus possibly large uncertainties.

Themis Regio (western part) is selected as target area, since it is geologically interesting and is located in an area that exhibits the largest number of usable VIRTIS-M-IR measurements on the planet. The target comprises all surface bins whose bin center coordinates are located between 35 and 47 $^\circ\text{S}$ and 270 and 288 $^\circ\text{E}$, resulting in $B = 219$ surface bins with ≥ 80 usable measurements each. A measurement data set with 64 repetitions for each bin is selected to be used in most cases in this work. 14,016 binned spectra constitute this selection that will be referred to as $N_r^0 = 64$ repetition data set in the following. Appendix A.1 lists the VIRTIS-M-IR data cubes utilized for the present work.

2.2. Forward model

In order to numerically simulate observable nightside radiances, a plane-parallel, LTE (local thermodynamic equilibrium), line-by-line radiative transfer simulation forward model is utilized that takes into account thermal emissions by surface and atmosphere, and absorption and multiple scattering by gases and clouds. It is similar to the forward model described by Haus and Arnold (2010) with updates by Haus et al. (2013), but the underlying radiative transfer equation solver DISORT (Stamnes et al., 1988) is replaced by LIDORT (Spurr, 2001, 2008). This way, the forward model is capable of providing analytic derivatives of the simulated radiances with respect to a number of atmospheric, surface, and instrumental parameters. The forward model was also used by Kappel (2014) and Kappel et al. (2015) where the following summary is based on.

From the total VIRTIS-M-IR spectral range (≈ 1.0 – 5.1 μm), only the interval 1.0–2.3 μm is utilized here excluding the 1.28 μm window range (1.225–1.295 μm) that is often contaminated by O_2 nightglow (see below). The line-forming altitude region for the specified range is located below 40 km (Haus and Arnold, 2010, Fig. 4). Venus' nightside emissions in this spectral range are thus nearly unaffected by strong mesospheric temperature variations above 58 km that were observed by Tellmann et al. (2009) and Haus et al. (2013, 2014). The temperature below 40 km is quite stable with time (Seiff et al., 1985). Furthermore, details of the cloud altitude distribution have a minimal impact in this range, since the main cloud deck resides above 48 km (Marov et al., 1980). In contrast, spectral signatures longward of 3 μm are, according to the cited figure, strongly influenced by variations of temperature and cloud altitude distributions above 48 km. The range 2.3–2.6 μm is excluded, since the number of interfering parameters for emissivity retrieval shall be kept as low as possible, and in the VIRTIS-M-IR spectral range, most minor gases are active only there (Haus and Arnold, 2010, Fig. 19). Nightside radiances in the range 2.6–3.0 μm are too small to be useful (Haus and Arnold, 2010, Fig. 2). Spatially and temporally varying non-LTE O_2 emissions (' O_2 nightglow') at 1.27 μm from an altitude region around 100 km (Piccioni et al., 2009) contaminate the 1.28 μm window. For retrievals, this window will thus be blacked out, and O_2 nightglow need not be considered. Note that calibration refinements and preprocessing for VIRTIS-M-IR data as performed by Kappel et al. (2012) had mainly concentrated on the range 1.0–2.6 μm .

In the considered spectral range, spectral surface emissivity is observable in the three 'surface window peaks' (1.02, 1.10, and 1.18 μm). The three parameters $e_{1.02}$, $e_{1.10}$, and $e_{1.18}$ parameterize the spectral emissivity such that it is constant over the ranges of the respective windows (1.000–1.055 μm , 1.055–1.125 μm , 1.125–1.225 μm , comprising

about four, eight, and ten spectral bands, respectively). Knowledge of all three emissivities can suffice to discriminate between a number of potential surface materials and grain sizes (Haus and Arnold, 2010). Emissivities must lie in the interval $[0, 1]$, and initial values are all set to 0.5.

Temperature altitude profiles are taken from the Venus International Reference Atmosphere (VIRA, Seiff et al. (1985)). For studies of nightside spectra in the utilized spectral range, essentially only the tropospheric temperature field is important. Below 32 km, it is usually considered to be independent of latitude as described by VIRA. In that range, remote sensing has not yet been performed, and *in situ* measurements were mainly restricted to low latitudes up to about 30° . But altitudinal or latitudinal deviations from VIRA may occur there, and in fact are suggested by a General Circulation Model (GCM, Lebonnois et al. (2010, personal communication)). However, *in situ* measurements and the GCM suggest that the deep atmospheric temperature field is quite constant in time and local time, a consequence of high thermal inertia and thermodynamic stable layering. Deviations from the VIRA temperature profile shall be parameterized by two parameters. ΔT_0 leads to a quadratic offset to VIRA below 25 km that is zero and has zero derivative at 25 km and is ΔT_0 at 0 km, regardless of the actual surface elevation. ΔT_{25} leads to a quadratic offset below 50 km that is zero at 0 and at 50 km and is ΔT_{25} at 25 km. This allows for a modification of the temperature in the bulk of the deep atmosphere, and temperature and temperature lapse rate near the surface (see Kappel et al. (2015, Fig. 3)). Surface temperature is set to the atmospheric temperature at the respective surface elevation. Pressure altitude profiles follow from hydrostaticity and the ideal gas law with VIRA pressure at the respective surface elevation as initial value. The elevation of a surface bin is provided according to Kappel et al. (2012), who use Magellan radar data (Ford and Pettengill, 1992; Rappaport et al., 1999) as basis. The elevation uncertainty of a $91.5 \times 91.5 \text{ km}^2$ surface bin (see Section 2.1) is assumed as 0.2 km (see Kappel et al. (2015)).

The CO_2 volume mixing ratio is 96.5%. The only minor gases that provide significant opacity in the studied spectral range are H_2O (including HDO, which is treated such that the initial D/H ratio corresponds to 150 times the corresponding ratio for Earth) and HCl (Haus and Arnold, 2010, Fig. 19). A small CO influence at the interval boundary (2.28–2.30 μm) is neglected here. Initial altitude profiles of volume mixing ratios of minor gases are taken from Haus and Arnold (2010) and are based on the profiles described by Pollack et al. (1993). Note that in contrast to the present work, Kappel et al. (2015) scaled the H_2O column such that the concentration at the surface was 25 instead of 32.5 ppmv. Both H_2O and HCl affect the peak at 1.74 μm (probed altitudes 20–30 km), and H_2O is in addition observable between 1.10 and 1.20 μm ($\leq 26 \text{ km}$). H_2O (excluding HDO) and HCl model variations shall be parameterized by total column factors $t_{\text{H}_2\text{O}}$ and t_{HCl} that scale the respective initial altitude profiles which translate to 32.5 or 0.5 ppmv at 25 km, respectively. Recent studies showed that H_2O concentration in the deep atmosphere amounts to a value of 30^{+10}_{-5} ppmv (Bézar et al., 2011), corresponding to $t_{\text{H}_2\text{O}} \sim 0.92^{+0.31}_{-0.15}$. Monochromatic absorption cross-sections of gases due to allowed molecular transitions are computed from the spectral line databases CDSD (CO_2 , 0.4–0.9 μm and 1.25–1.33 μm , Tashkun et al. (2003)), HITEMP (CO_2 , Pollack et al. (1993), H_2O isotopologues 161, 181, 171, Rothman et al. (1995)), and HITRAN08 (H_2O isotopologues 162, 182, 172, HCl, Rothman et al. (2009)) by using spectral line shapes listed by Haus and Arnold (2010). To be in accordance with the composite line database used by Haus and Arnold (2010) that was extensively validated for radiative transfer in Venus' atmosphere, the more recent HITEMP2010 (Rothman et al., 2010) is not yet considered here. Molecular Rayleigh scattering is treated according to Hansen and Travis (1974).

The opacity of the main gaseous constituent CO_2 for the high-pressure and high-temperature environment of Venus' deep atmosphere is not sufficiently well predictable to provide reliable simulations of observable spectra. The line databases of the allowed transitions are not based on laboratory measurements, but mostly on theoretical models (Tashkun et al., 2003) and contain errors (Haus and Arnold, 2010; Bézar et al., 2011). The line shapes are still under discussion, as are additional opacity effects (continuum, collisional induced absorption, line mixing). The CO_2 opacity due to allowed transitions requires a wavelength dependent correction to produce good fits to real data at all. This correction is in the following shortly referred to as 'continuum'. For this work, it does not only include the actual continuum contribution, but opacity corrections in general. It is independent of the measurement but depends on the utilized line databases and line shapes and is thus not directly comparable to continua given by authors who made other choices. The continuum is treated here as a spectrally constant value throughout the range of an atmospheric transparency window, except for the 1.10 and the 2.3 μm windows. The spectral fine structure fluctuations of the latter around the value $c_{2.3} = 160 \cdot 10^{-29} \text{ cm}^2$ are relatively easy to retrieve and can be found in Kappel et al. (2012, Fig. 11). To allow for satisfactory fits to the 1.10 μm peak, its continuum is modeled as a (wavenumber-) linear function that connects the values $c_{1.056} = 22 \cdot 10^{-29} \text{ cm}^2$ and $c_{1.124} = -6.9 \cdot 10^{-29} \text{ cm}^2$ at wavelengths 1.056 and 1.124 μm .

The other continua are denoted by $c_{1.02}$, $c_{1.18}$, $c_{1.31}$, and $c_{1.74}$, where the subscripts indicate the wavelengths of the respective windows, and are set to 0.39, 1.9, 9.6, and 26 in units of 10^{-29} cm^2 , respectively. These values are the most recent continuum results of the present authors and were retrieved with MSR (Kappel, 2014) as parameters that are common to a large set of highly diverse VIRTIS-M-IR nightside spectra and assuming a fixed surface emissivity of 0.5 (the initial emissivity). Modifications to the retrieval pipeline and the choice of the set of utilized spectra and the initial emissivity can affect the results (order of 10%–20%, depending on the window). However, a retrieval error analysis analogous to the one performed by Kappel et al. (2015) for surface emissivity would be needed to understand their significance, but these values are not the primary target of this work. Therefore, values stated in this and in earlier publications of the present authors should be, as mentioned in those works, regarded as working hypotheses and preliminary values. Note that in this work, the surface window continua are treated as fine-tuning parameters that shall prevent the retrieved emissivities to systematically approach the boundaries 0 or 1, as will be explained in Section 2.3. It was shown by Kappel et al. (2015) that uncertainties in the continua are important sources of retrieval errors for absolute emissivities. To avoid high emissivity retrieval errors due to continuum uncertainties, the strategy for the emissivity retrieval is modified by rather retrieving relative emissivities and showing the results to be reasonably independent of the continua, the initial emissivity, and other parameters.

The clouds are assumed to comprise four modes (1, 2, 2', and 3) and are modeled according to Haus et al. (2013). Each mode consists of spherical droplets of 75% sulfuric acid. Refractive indices are taken from Palmer and Williams (1975) with supplements from Carlson and Anderson (2011). Cloud particle radii are log-normally distributed with modal radii of 0.3, 1.0, 1.4, 3.65 μm and unitless dispersions of 1.56, 1.29, 1.23, 1.28 for the four modes, respectively (Pollack et al., 1993). Initial profiles of particle number densities N at altitudes h are taken from Haus et al. (2013) and are given by

$$N(h) = \begin{cases} N_0 \exp(- (h - h_t)/s_t) & h > h_t \\ N_0 & h_b \leq h \leq h_t \\ N_0 \exp(- (h_b - h)/s_b) & h < h_b \end{cases} \quad (1)$$

Mode-specific values are listed in Table 1. $N(h)$ is set to zero above 85 km. Actual cloud modal abundances are defined

Cloud mode	1	2	2'	3
Top of peak density, h_t in km	65	66	60	57
Base of peak density, h_b in km	49	65	49	49
Upper scale height, s_t in km	5	3.5	1	1
Lower scale height, s_b in km	1	3	0.1	0.5
PND at h_b , N_0 in cm^{-3}	181	100	50	14

Table 1: Initial cloud parameters for Eq. (1). PND is the particle number density.

by cloud mode factors m_1 , m_2 , $m_{2'}$, and m_3 that scale the number densities of the initial altitude profiles. Variations of h_b , h_t , s_b , or s_t for cloud mode i rescale N_0 for mode i such that the total number of mode i particles per horizontal unit area remains constant. This is done in order to disentangle variations of altitude distribution and variations of particle abundance. Deviations of the h_b for modes 1, 2', and 3 from their initial value 49 km are treated as a joint parameter, since below that altitude, the temperature and pressure conditions are thought to force cloud particle numbers to sharply decrease except for possible haze (see *in situ* data described by Knollenberg and Hunten (1980) and Marov et al. (1980)). The cloud top altitude is defined as the altitude where the cloud optical depth at 1.0 μm attains unity. Haus et al. (2013, 2014) described a slow decrease of the local-time- and time-averaged nightside cloud top altitudes from ~ 71 km at the equator to ~ 70 km at mid latitudes and then to ~ 61 km at polar latitudes. Here, this is implemented in an abstracted form as a linear decrease of h_b and h_t for cloud mode 2 and h_t for cloud mode 1 by 10 km (referred to Table 1) between 55° and 80° latitude. Wavelength dependent scattering and absorption properties of the clouds are computed according to Mie theory (Wiscombe, 1980).

2.3. General remarks on multi-spectrum retrieval and renormalization

Since spectral information content is limited, there are always atmospheric and surface parameters required for the numerical simulation of spectra that cannot be uniquely derived from the measured spectra. These parameters interfere

with emissivity data and are usually set to certain reasonable values that are compatible with current knowledge on atmospheric and surface conditions and allow for the generation of realistic synthetic spectra. When these assumptions deviate from the true physical values, this may cause surface emissivity retrieval errors, even though the fits well match the measured spectra. Compared to this effect, the impact of measurement noise is of secondary importance. It was shown by Kappel et al. (2015) that these errors are too large to obtain reliable absolute emissivities from single VIRTIS-M-IR measurements. However, it was proposed to apply the multi-spectrum retrieval algorithm MSR developed by Kappel (2014) to retrieve emissivity as a parameter that is common to several measured spectra that cover the same surface bin. This assumes that geologic activity (and thus time variations of emissivity) is negligible. This seems to be indeed the case, as Müller et al. (2012a) have argued on the basis of the VIRTIS-M-IR data. MSR allows for the retrieval of parameters that are common to a certain selection of measurements and to incorporate expected mean values, standard deviations, and correlation lengths and times as Bayesian *a priori* information for the retrieval parameters, along with measurement and simulation error information. This decreases the probability of retrieving unreasonable values and spatial-temporal distributions of the parameters, allows for taking the context of adjacent measurements into account, and thus enhances the reliability of the retrieved results. All considered measurements can be parameterized by a fully consistent set of atmospheric, surface, and instrumental parameters.

Even with MSR, absolute emissivity values are difficult to obtain. Given a certain surface bin that is covered by N_r measurements, the impact of fluctuations around the long-term average of a temporally varying parameter at that surface bin should statistically average out with $1/\sqrt{N_r}$. Examples of such temporally varying interfering parameters are the column abundances m_i of the cloud modes, top and bottom altitudes of the vertical cloud mode particle distributions, the H_2SO_4 concentration of the cloud droplets, and atmospheric trace gas abundances. Since the spatial distributions of the long-term averages of these interfering parameters are not exactly known, they cannot be considered in the radiative transfer simulations, and retrieved emissivity maps may therefore systematically deviate from the true emissivity map. However, as long as the measurements of the surface target are not in the vicinity of the terminators or the polar regions, the long-term average of the interfering parameters should not spatially vary on scales smaller than a few thousand km. This can be motivated by the symmetries and spatial scales of the global atmospheric circulation. Uncertainties of constant interfering parameters like the continuum parameters can lead to additional large emissivity errors.

It was proposed by Kappel et al. (2015) to remove these errors as follows. First, to minimize emissivity errors that are due to continuum parameter uncertainties, wavelength dependent continua are pre-determined with MSR as common parameters that are compatible with as many and diverse as possible observed spectra. An initial emissivity value has to be defined for this purpose to complete the parameter set required to solve this retrieval problem without too strongly underconstraining it. The pre-determined continua do not necessarily coincide with the true ones but have proved to be compatible with almost all VIRTIS-M-IR nightside spectra analyzed so far. Using them, a surface emissivity map, quasi 'relative to the initial emissivity', can be retrieved with MSR from a suitable measurement data set that repeatedly covers the actual surface target. This map is henceforth also called the 'raw retrieved emissivity map'. The mean value of this map can differ from the initial emissivity, since the measurement data sets utilized to determine the continua and the map (and therefore the corresponding time-averaged interfering true parameters) do not necessarily coincide. It was proposed to add a constant value to the retrieved relative emissivity map such that its new median value corresponds to the initial emissivity. This way, the impacts of uncertainties in the mentioned interfering parameters were predicted to be largely removed. In the present work, this addition is not performed at this point. Note that the present work does not use median values but mean values, because additional statistical parameters like standard deviation have to be analyzed here. The median's advantage of robustness against statistical outliers was verified to be not really necessary here.

Additional effects have to be taken into account. For instance, the relative emissivity maps can exhibit spatially slowly varying trends in latitude direction that are probably unrealistic from the geologic point of view. Such trends were assumed to be negligible by Kappel et al. (2015) who utilized synthetic spectra to develop the retrieval of relative emissivities. The relevance of these trends was only recognized when real measurements were analyzed in the frame of the present work. The occurrence of these latitudinal trends can be explained with the spatial distributions of the time-averaged true values of interfering parameters. Again motivated by the symmetries and spatial scales of the global atmospheric circulation, and provided that the measurements of the surface target are not in the vicinity of the terminators or the polar regions, these parameter distributions are expected to be spatially slowly varying, and only in latitude direction. Moreover, they can depend on the utilized measurement data set. The respective parameters used in the forward model simulation are, in absence of detailed knowledge, mostly set to constant values. These

spatially slowly varying deviations between time-averaged true and assumed parameters lead to corresponding artificial latitudinal trends in the retrieved emissivity map.

It also turned out that, for not too high surface elevations, the retrieved emissivities can exhibit a trend with surface elevation that is approximately linear. This trend seems to be unrealistic from a geologic point of view, at least if it should turn out as a global phenomenon that occurs regardless of geologic context. If not real, it might be explained by unexpected extinction properties close to the surface in the low-lying regions of the planet, for instance due to errors in the CO₂ absorption or to a kind of haze close to the surface, compare also [Grieger et al. \(2004\)](#); [Erard et al. \(2009\)](#). But such a conclusion would require a detailed retrieval analysis of a much larger portion of Venus' surface than is currently possible with the available computer resources. Therefore, the origins of such trends are not further discussed here. It is assumed until the emergence of convincing arguments to the contrary that these are not real emissivity trends, but retrieval artifacts due to imperfect or unconsidered forward model parameters. Note that this is just an assumption at the moment, and it should be kept in mind that actual trends with surface elevation (due to trends in the distribution of dust, surface composition, or weathering) might exist.

At each surface window, the trends of the raw retrieved emissivity map with both latitude and surface elevation are jointly removed ('de-trended') by a multiple linear regression (see Appendix A.2.1 for details), but in a way that does not change the emissivity mean value. For an illustration, see Fig. 1a that depicts a raw retrieved emissivity map at 1.02 μm , and Fig. 2 that shows the resulting trends with latitude and topography. In Fig. 1b, the topography trend is removed, and in Fig. 1c both the topography and the latitude trends. These figures are explained later in Section 3, since the retrieval pipeline required to derive these maps can only be introduced after the discussions in the present subsection.

However, the spatial fluctuations of the de-trended emissivities around their mean tend to increase in magnitude for increasing emissivity mean, because the radiance response to small emissivity perturbations around an emissivity base value decreases for increasing emissivity base value. The fluctuations are small compared to the emissivity mean. This effect is detailed in Section 4.5. It is shown there that it can be quantified according to an (affinely) linear transformation with respect to the mean values of the raw retrieved emissivities. To remove the impact on the derived emissivity maps, this linear transformation is used to refer the de-trended maps to a certain reference emissivity value, here 0.5. The result is called the 'renormalized emissivity map referred to the reference emissivity 0.5'. In the following, the maps will always be referred to 0.5, except when stated otherwise, and the result is shortly called 'renormalized emissivity map'. The emissivity mean of this renormalized map coincides with the reference emissivity, i.e. the emissivities now spatially fluctuate around the reference emissivity. If the reference emissivity coincided with the (unknown) true absolute mean emissivity of the target area, the amplitude of the derived fluctuations would coincide with the amplitude of the true fluctuations, assuming that there are no real emissivity trends with latitude and topography. The described linear transformation replaces the plain addition operation proposed by [Kappel et al. \(2015\)](#) to make the emissivity mean value coincide with the initial emissivity. Using the linear transformation, the renormalized emissivity map can easily be referred to other reference emissivities if needed. The dependence of the fluctuation amplitude on the assumed mean emissivity of the target area is an important effect for the quantification of spatial emissivity variations and cannot be neglected. This effect is most pronounced at 1.02 μm .

By applying the entire renormalization procedure described in the last paragraphs, the extremely high emissivity retrieval errors due to systematically wrong assumptions on interfering parameters can be largely avoided. In return, it has to be verified that the renormalized retrieved maps do not significantly change when the initial emissivity, the continua, slowly varying latitudinal deviations of the utilized from the actual deep atmospheric temperature field, and other interfering parameters are modified. The selected surface target should not exceed spatial dimensions of a few thousand km, a scale on which the interfering parameter trends and thus the systematic deviations of the retrieved from the true emissivity map can be thought to be approximately linear functions of latitude, allowing for a corresponding renormalization. In the same way, the surface elevation range of the target should not be exceedingly wide to ensure approximate linearity of the trend with topography and its successful renormalization.

The regression coefficients for each map are listed in addition to the mean emissivity before renormalization (Table 4). The interplay between unexpected extinction properties close to the surface, an unconsidered latitudinal trend of the deep atmospheric temperature field, the utilized continua, and mean and latitudinal trends of time-averaged cloud abundances, cloud bottom altitudes, and H₂O abundances, as well as other discrepancies between utilized and true parameters, depending on the selection of spectra, can lead to different emissivity means and regression coefficients.

In practice, this means that the raw retrieved emissivities may approach their domain boundaries, i.e. the boundaries

of the set of physically valid values attainable by the respective parameter (0 or 1 for emissivity), or are even cut off at these boundaries. This may occur even for fixed continua and deep atmospheric temperature field, in particular when comparing target areas at different latitudes. Note that violations of these boundaries are prevented in the retrieval process by a modification (Kappel, 2014, Appendix A.1) of the logarithmic barrier method (Nocedal and Wright, 1999, Section 17.2). In the extreme case, the raw retrieved emissivities of all surface bins may thus become 1, for instance, in which case renormalization would lead to a featureless map where all emissivities would correspond to the reference emissivity. On the one hand, this is the reason, in particular in presence of trends of the raw retrieved emissivities with topography or latitude, for choosing the initial emissivity 0.5 as the value that is farthest away from the emissivity domain boundaries. On the other hand, renormalization alone turns out to be insufficient to make the determination of useful emissivity maps fail-safe. As a workaround, the current approach for operational retrievals is to retrieve the surface window continua each time anew along with the emissivity maps. The resulting continua differ only slightly, but now they take on the character of fine-tuning parameters that shall prevent the emissivity maps to approach the boundaries 0 or 1 from the outset. Renormalization is still needed to ensure an exact coincidence of the emissivity mean with the reference emissivity and to remove the mentioned latitude and topography trends. Note that after renormalization, the emissivity mean corresponds to the reference emissivity 0.5, and the slopes of linear trends with latitude and surface elevation are zero, such that an additional subsequent renormalization would not change results anymore. In the following, the term 'renormalized emissivity' is to be understood to include the fine-tuning step. It will be verified that no further modifications to this approach are required to obtain reasonably reliable emissivity data maps.

Note that the retrieval error analysis by Kappel et al. (2015) did not consider measurement errors. Moreover, the shapes of simulated and measured spectral peaks can slightly differ, in which case the parameters can more easily converge to subsidiary solutions. For instance, when the spatially slowly varying time-averaged cloud abundance for the utilized set of measurements exceeds the long-term average, the retrieval algorithm may converge to a solution where the decreased radiance average is explained by a modified cloud mode abundance in conjunction with a decreased emissivity. This may lead to retrieval errors that exceed those listed by Kappel et al. (2015) and amplify the emissivity domain boundary problem. Application of the approach described in this section, however, should largely remove such errors.

All these subtleties necessitate a verification of the applicability of the synthetic retrieval error analysis results by Kappel et al. (2015) to real renormalized emissivity maps. This is discussed in Section 4.

2.4. Retrieval pipeline

Studying synthetic spectra, Kappel et al. (2015) have compared several single-spectrum retrieval pipelines for surface emissivity retrieval. For the pipeline that yielded the smallest retrieval errors, the auxiliary parameters that have to be retrieved to allow for useful fits at all comprise the cloud mode 2, 2', and 3 factors. Along with the emissivities, these parameters are determined from the spectral range 1.0–2.3 μm excluding the blacked out 1.28 μm window. This pipeline is used as a baseline for a suitable multi-spectrum retrieval pipeline. To be applicable to real, not synthetic measurements, the new pipeline has to be more elaborate. The main reason is that three parameters that describe the spectral registration and that are crucial for emissivity retrieval are not sufficiently predictable by the calibration pipeline: the spectral sampling $\delta\lambda^s$, the wavelength λ_1^s of first spectral band, and the full-width-at-half-maximum $FWHM^s$ of the spectral instrumental response function (Kappel et al., 2012). An $FWHM^s$ that is wavelength independent can be checked to allow for satisfactory fits, but $FWHM^s$ and λ_1^s turn out to vary with spatial sample on the detector and detector temperature. $\delta\lambda^s$ seems to be generally constant, and an optimal value of 9.49 nm per spectral band can thus be determined with MSR as parameter that is common to many (thousands) and diverse VIRTIS-M-IR spectra. $FWHM^s$ and λ_1^s , on the other hand, must be individually determined for each spectrum. It was shown by Kappel (2014) that $FWHM^s$ can be retrieved quite reliably already for single-spectrum retrieval, although MSR results taking into account the context of adjacent measurements are more accurate. λ_1^s can be retrieved very reliably, since it well disentangles from any other considered retrieval parameter due to its distinct Jacobian. Appendix A.3 depicts and describes the detector-related trends for retrieved λ_1^s and $FWHM^s$.

In the following, it is to be understood that the retrieval parameters are always determined for the entire set of utilized measurements by applying MSR and using the *a priori* data listed in Table 2. The initial guess for the retrieval iterations is always set to the corresponding *a priori* mean value a , except for λ_1^s where it is set to the value predicted by the calibration pipeline. The *a priori* standard deviations are here set to values that are larger than is realistic, to not constrain the effective parameter ranges too much, thereby focusing more on the spatial-temporal relations. In the future,

Parameter	a	σ	λ	τ
λ_1^s	1.02 μm	0.03 μm	0.3×256	0.2 h
$FWHM^s$	17 nm	5 nm	0.5×256	0.2 h
$m_1, m_2, m_{2'}, m_3$	1	5	500 km	3.6 h
H_2O^b	1	0.5	2000 km	8 h
Continua	10^{-29} cm^2	10^{-27} cm^2	/	/
$e_{1.02}, e_{1.10}, e_{1.18}$	0.5	2	50 km	/

Table 2: *A priori* mean values a , standard deviations σ , correlation lengths λ , and correlation times τ for retrieval parameters. λ_1^s : wavelength of first spectral band, $FWHM^s$: FWHM of spectral instrumental response function, m_i : column scaling factor for cloud mode i , H_2O^b : factor for H_2O column below 26 km, e_j : emissivity at surface window j . λ for λ_1^s ($FWHM^s$) is set to 30% (50%) of the spatial width of the detector of 256 samples. Retrieval of common parameters (infinite correlation lengths or times) is indicated by '/'.

it will be tested, whether it is safe to set them to smaller values. Single-spectrum retrieval results show that both $FWHM^s$ and λ_1^s have a quite smooth detector sample and temperature dependence. Due to the properties of the VEX orbit, detector temperature with the spacecraft close to an eclipse boundary can vary with time on a scale of a few minutes. It improves disentanglement from other parameters, in particular for $FWHM^s$, when correlation lengths and times are set as *a priori* data for MSR according to Table 2. As discussed by Kappel et al. (2015, Section 3.5), the expected correlation lengths for the cloud mode factors are set to $\lambda = 500$ km (with a few hours correlation time), which can be motivated by computing the auto-correlation function of observation movies of 2.3 μm nightside radiances as proxy. But there is also evidence for smaller scale cloud structures (100–200 km for mottled clouds observed by VMC/VEX (Titov et al., 2012)), and small-scale convection in the middle cloud region causes large optical thickness variations (Imamura and Hashimoto, 2001). It can be shown by computing the corresponding Jacobian that the parameter H_2O^b , which is defined to scale the H_2O column only below 26 km, mostly affects the range 1.10–1.20 μm and only slightly influences the 1.74 μm peak. In contrast (see Fig. 2 by Kappel et al. (2015) for a selection of relevant Jacobians), a factor for the entire H_2O column significantly affects both ranges and thus leads to a stronger interference with clouds than H_2O^b does. Since the number of interferences between retrieval parameters should be kept as low as possible, usage of H_2O^b is preferred here over a factor for the entire H_2O column. Note that, according to the respective Jacobians, H_2O variations at higher altitudes can be compensated in the 1.74 μm peak by slightly different cloud mode abundances. The correlation length for H_2O^b is chosen to be compatible with data for minor gases by Tsang et al. (2009) and Haus et al. (2015), with the atmospheric convection and its spatial scales as motivation. Correlation time is such that a distance corresponding to half of the correlation length is covered by gases at wind speeds of ≈ 35 m/s that prevail at 26 km altitude (Schubert, 1983, Fig. 9). For surface emissivities, the *a priori* correlation length can be set to 50 km, which corresponds to the half FWHM of the expected achievable spatial surface resolution in the IR (100 km, Moroz (2002)).

It can be seen from deep space observations that noise equivalent spectral radiance for actual VIRTIS-M-IR measurements is in the order of $4 \cdot 10^{-4} \text{ W}/(\text{m}^2 \text{ sr } \mu\text{m})$ for wavelengths shortward of 2.6 μm (256-sample-mode, unbinned, 3 s exposure time). As discussed in Section 2.1, each VIRTIS-M-IR cube is binned to match the 100 km FWHM of the surface data blurring. For a given data cube, typically 30 spectra contribute to each bin for southern hemisphere mapping data at 45 $^\circ\text{S}$ (Kappel et al., 2015). It is thus reasonable to assume a Gaussian with standard deviation $\sigma_N = 10^{-4} \text{ W}/(\text{m}^2 \text{ sr } \mu\text{m})$ as first guess for the measurement and simulation error distribution for binned data.

The retrieval pipeline is arranged into several stages that shall ensure a steady convergence and a minimal use of resources. If all retrieval parameters were determined in a single stage from the entire wavelength range, not only unnecessarily large computer memory usage and processing time had to be used (faster than linearly increasing with problem size (Kappel, 2014)), but the parameters could more easily converge to subsidiary minima far away from the global minimum of the retrieval cost function.

The first stage retrieves λ_1^s from the 1.74 μm peak. Note that the result is not final, since later stages will refine this value, based on the ever-increasing improvement of the fit for an ever-wider wavelength range. The second stage determines λ_1^s and $FWHM^s$ in addition to cloud mode factors $m_{2'}$ and m_3 from the 1.74 and 2.3 μm peaks. In the later stages, λ_1^s and $FWHM^s$ are both determined with ever-improving refinements, but this will not be mentioned anymore. The third stage estimates the surface window continua (as fine-tuning parameters as explained), while the fourth stage repeats the second stage but also determines m_2 and utilizes the surface windows in addition. At this point, the retrieved values quite well parameterize the measured spectra for most of the relevant wavelength range, and it is

unlikely that in the later stages, the retrieval parameters lock to subsidiary solutions that are farther away from the global minimum of the retrieval cost function than at this stage. The emissivities are technically only needed to refine the fit and are determined later. Note that the $1.31\ \mu\text{m}$ peak is not included yet, since it does not yield significant additional information (Kappel et al., 2015) and that the $1.31\ \mu\text{m}$ continuum is not determined yet. The fifth stage repeats the third to refine the surface windows continua, and the sixth stage repeats the fourth, but (optionally omitted) retrieves m_1 and H_2O^b in addition. The seventh stage retrieves a first version of the surface emissivities, and the eighth and ninth stages repeat stages six and seven for a final refinement. Optionally, a last stage retrieves the $1.31\ \mu\text{m}$ continuum for completeness. The resulting maps are renormalized as described above.

3. Renormalized emissivity maps at Themis Regio

The raw emissivity map at $1.02\ \mu\text{m}$ retrieved according to Section 2.4 from the $N_r^0 = 64$ repetition data set introduced in Section 2.1 is shown in Fig. 1a. The emissivity average over all bins of the target area is 0.48 (compare Table 4, row ' $N_r^0 = 64$ ', column 'Mean', first value). To enable a better comparison to all subsequent map representations, a constant value of $(0.5 - 0.48)$ is added to the raw emissivity of each bin leading to a 'shifted' map with average 0.5. This map is interpolated (using thin plate splines (Williams et al., 2014)) between the shifted emissivity values at the bin centers and represented according to the depicted color bar. An equirectangular map projection is used with the center of the shown latitude range as standard parallel, i.e. there is no local distortion due to the map projection at this latitude. The longitude range $270.5\text{--}287.8^\circ\text{E}$ covered by the bin centers corresponds to a distance of 1373 km measured along the standard parallel, the latitudinal extension -46.99 to -35.63°N corresponds to 1200 km. The '+' symbols indicate the centers of the surface bins. The switch between Lambert cylindrical and Lambert south pole azimuthal projections at -38.68°N for the equal area surface binning (Section 2.1) is clearly discernible from the distribution of the bin centers. Fig. 1d depicts the blurred Magellan topography (see Section 2.1, for the unblurred map see Fig. 3d) of the target area.

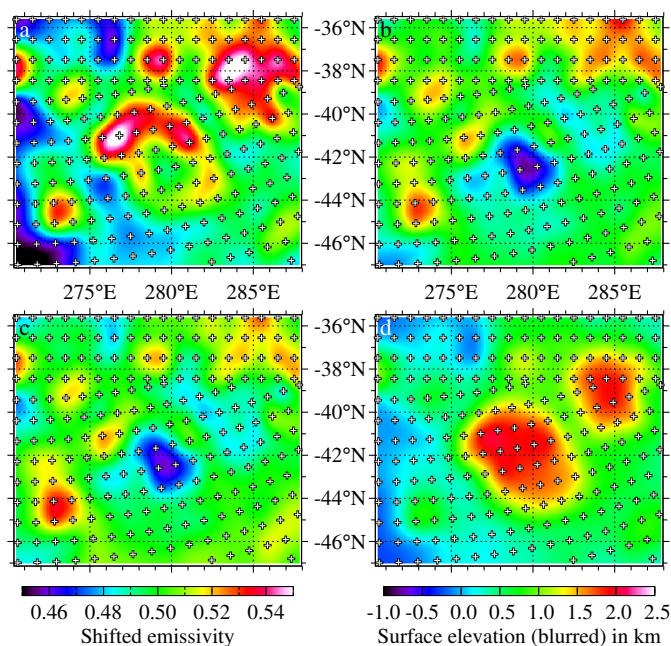


Figure 1: Removal of trends with topography and latitude for the $1.02\ \mu\text{m}$ emissivity map of Themis Regio retrieved from the $N_r^0 = 64$ measurement repetition data set. (a) Raw map, (b) trend with topography removed, (c) trends with topography and latitude removed, (d) utilized surface elevation (blurred Magellan topography). Suitable constants were added to the emissivity maps such that the map averages are always 0.5 for better representation. The corresponding results at 1.10 and $1.18\ \mu\text{m}$ are given in Figs. A.3 and A.4 in the supplementary material. Horizontal axes: longitudes, vertical axes: latitudes. 1° latitude corresponds to 105.62 km at the surface. For color representations of the figures in this work, the reader is referred to the online version of this article.

As it was explained in Section 2.3, the raw emissivity maps can exhibit (probably artificial) trends with topography

and latitude. This is illustrated in Fig. 2 for the raw emissivities at 1.02, 1.10, and 1.18 μm derived from the $N_r^0 = 64$ data set. For a better representation, the trends are shown separately (trend with topography left, trend with latitude right). There is an obvious correlation between raw emissivity and topography. The latitudinal trend is less pronounced, but its removal is still important for quantitative comparisons between different maps of the same target. Moreover, the latitude range is quite small here, and additional maps with wider latitude ranges are needed to assess this trend's global relevance. The slopes of both trends are listed in row ' $N_r^0 = 64$ ' of Table 4. It will be shown in Section 4 that for the same target area, the slopes depend on interfering parameters and the selection of the measurements (compare Table 4), which leads to large differences between raw maps of the same target. On the other hand, the renormalized maps, where both trends have been removed, are very similar among each other as will be shown. This indicates that the trends are artificial indeed. The unrealistic systematic divergence between the raw emissivities at different surface windows for increasing surface elevation points to the same conclusion. It was already explained that the trends are probably due to imperfect forward modeling. If the trends are artificial, their magnitudes are so large that their occurrence cannot be neglected and they have to be removed to obtain useful maps. Fig. 1b depicts the map from Fig. 1a with the topography trend removed, and in Fig. 1c, both the trends with topography and latitude are removed.

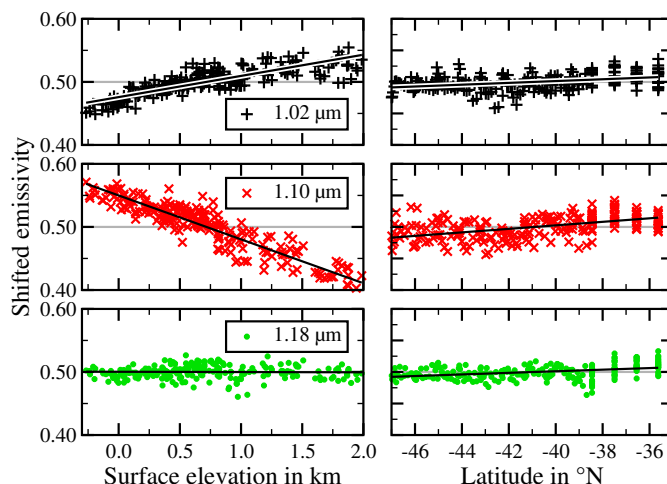


Figure 2: Scatterplots (symbols) and linear trends (black lines, for slopes see row ' $N_r^0 = 64$ ' of Table 4) illustrating the trends with topography and latitude in the raw retrieved emissivity maps of Themis Regio. Left: trend with blurred Magellan topography when trend with latitude is removed. Right: trend with latitude when trend with topography is removed. Suitable constants were added to the emissivities such that the map averages are always 0.5 for better representation.

It was discussed in Section 2.3 that the spatial fluctuations of the de-trended emissivities around their mean value tend to increase in magnitude for increasing emissivity mean. To finally obtain the renormalized maps, the de-trended maps have to be referred to the reference emissivity 0.5 according to a linear transformation that is detailed in Section 4.5. The renormalized emissivity maps derived from the $N_r^0 = 64$ repetition data set and referred to the reference emissivity 0.5 are shown in Fig. 3. This is the first main result of this work. The second main objective, the reliability of these results, is addressed in Section 4. Recall that the maps represent not absolute but relative emissivity values.

The 1.02 μm renormalized emissivity map shows distinct spatial variations. They indicate heterogeneous surface properties. Pronounced anomalies occur relative to the average 0.5 (green) of the renormalized emissivity map. Two major trends characterize the studied area. Low emissivity is centered over Shiwanokia Corona (42.0 °S, 279.8 °E, see Section 1), whereas a positive emissivity anomaly appears in association with Abeona Mons (44.8 °S, 273.1 °E). These results cannot be explained by spatial topography variations, since both these areas exhibit higher surface elevations than their respective surroundings. Rather, the emissivity variations clearly correspond to relevant geologic surface patterns. A detailed discussion of emissivity correlation at all three wavelengths with geomorphologic surface features will be given in a follow-up paper.

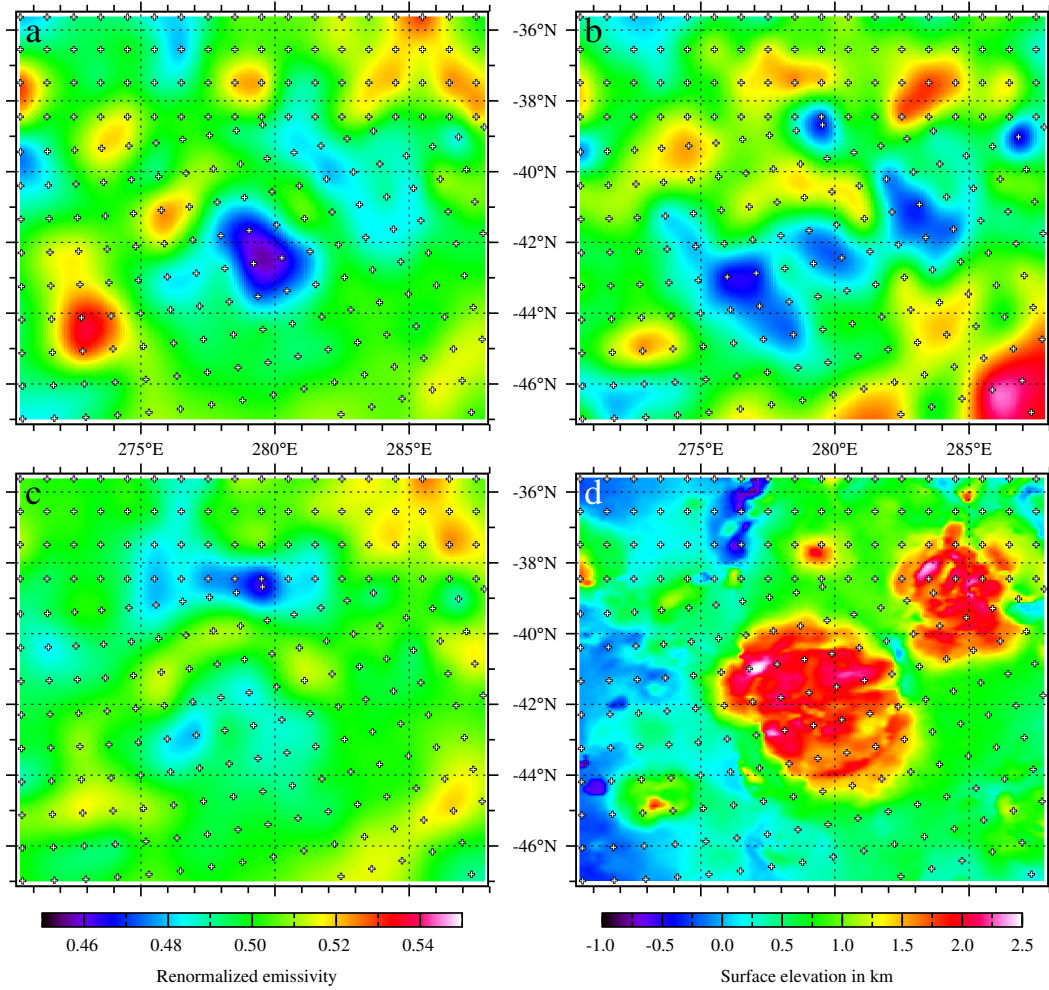


Figure 3: Renormalized emissivity maps ($N_r^0 = 64$ measurement repetition data set) of Themis Regio referred to reference emissivity 0.5 at (a) 1.02 μm , (b) 1.10 μm , and (c) 1.18 μm , and (d) surface elevation (unblurred Magellan topography). Axes as in Fig. 1.

4. Retrieval errors and reference emissivity

No measured value makes sense without knowledge about its reliability. This is particularly true for the emissivity maps derived in this work. Section 4.1 recites predictions regarding their reliability gained from an earlier study with synthetic spectra. A statistical analysis of maps retrieved from real measurements is used for an improved error estimation (Sections 4.2, 4.3, and 4.4). Based on results from Section 4.4, it is discussed, how retrieved maps can be referred to a given reference emissivity (Section 4.5).

4.1. Retrieval errors determined from synthetic spectra

Kappel et al. (2015) have performed an error analysis for the single-spectrum retrieval of surface emissivity from synthetic VIRTIS-M-IR spectra. Corresponding multi-spectrum retrieval errors for a renormalized emissivity map were estimated from the respective single-spectrum errors by plausible scaling rules that emulate the statistical impact when each surface bin is covered by N_r repeated measurements. Appendix A.2 of the present work justifies that these rules are approximately valid, also discussing the assumptions required to derive them. The case $N_r = 25$ was used by Kappel et al. (2015) to illustrate the order of magnitude of the various contributions to the retrieval errors of the renormalized emissivity maps. The present work focuses on the case of $N_r = 64 =: N_r^0$, and in the following, retrieval

errors are given that are scaled to N_r^0 , based on the values given by Kappel et al. (2015, Table 5) where five different reference spectra were studied.

Errors due to uncertainties of temporally varying interfering parameters (cloud parameters, minor gas abundances) scale with $1/\sqrt{N_r}$, compare Eq. (A.3). Double standard deviation errors are in the order of 2%–6% of the true emissivity, depending on the surface window and the utilized spectrum, and assuming statistical independence of the interferences. The main contributors for this case are uncertainties of the cloud bottom altitude and the H₂O content of the atmosphere.

Time independent interfering parameters that are spatially constant (continua) or vary on a scale $\gg 100$ km (deep atmospheric temperature field uncertainties) theoretically do not contribute to retrieval errors (Eq. (A.3)). Those parameters that spatially vary on a scale of 100 km (surface elevation, interfering emissivities) have been excluded from the discussions in Appendix A.2. They cannot be removed by the renormalization procedure since they are not spatially slowly varying. Largest error contributor in this category is the surface elevation uncertainty. For smooth surface regions, surface elevation uncertainties for bins with an area of roughly 100×100 km² are assumed to be in the order of 100 m, and for more rugged mountainous areas even higher. Generally, an uncertainty of 200 m is assumed here. Corresponding emissivity errors do not improve for an increased N_r , because uncertainties do not average out over time. The errors at a given surface window due to (25%-) uncertainties of the emissivity parameters of the respective other surface windows are assumed to statistically scale with $1/\sqrt{N_r}$ (Kappel et al., 2015, Section 2.3). Together, the temporally constant and spatially fast varying interfering parameters lead to corresponding errors in the order of 9–13% for the three windows.

Measurement noise with a standard deviation of 10^{-4} W/(m² sr μ m) leads to additional 1%–2%. Reasonable modifications of *a priori* mean values have negligible impacts. Retrieved maps are most reliable at 1.02 μ m. There is an overall tendency for better results for cases with small cloud opacity, high surface elevation, high true emissivity, and small observation angle, but this depends on the emissivity window, retrieval pipeline, and measurement repetition number. Spatial fine structure is more reliable than large-scale spatial trends are.

Table 3 summarizes the renormalized emissivity errors referred to a typical synthetic VIRTIS-M-IR spectrum, the default reference spectrum from Kappel et al. (2015, Section 3 and Table 4). Note that these values slightly differ from the errors just stated that are based on five different reference spectra. Also note that these errors are referred to a surface emissivity of 0.65. The column R_1 shows the (not renormalized) single-spectrum retrieval errors. Note that the statistically scaled errors for the case $N_r = 64$ (column R_{64}) only apply when the emissivity maps are renormalized according to Section 2.3. The rows with ‘/’ in the R_{64} -column show the error contributions that are largely removed when using renormalized multi-spectrum retrieval and that would not decrease for increasing N_r without renormalization.

Uncertain parameters	2σ	R_1			R_{64}		
		t	17	35	49	2	4
Cloud bottom	3 km	5	20	14	1	3	2
H ₂ SO ₄	10%	2	3	12	0	0	2
H ₂ O	25%	15	26	43	2	3	5
HCl	25%	2	3	8	0	0	1
Measurement noise	$2\sigma_N$	4	11	8	0	1	1
Interfering emissivities	25%	9	34	27	1	4	3
Surface elevation	200 m	9	12	12	9	12	12
ΔT_0	3 K	18	35	35	/	/	/
ΔT_{25}	3 K	12	21	23	/	/	/
Continua	25%	38	145	130	/	/	/
$c_{1.74}$	25%	24	83	94	/	/	/
$c_{2.3}$	25%	17	71	73	/	/	/

Table 3: Quadratically added errors R_{N_r} of 1.02, 1.10, and 1.18 μ m emissivities (renormalized in case of $N_r = 64$) that were retrieved from a synthetic spectrum, in %, for 2σ -uncertainties of the listed parameters and N_r - measurement repetitions. The row ‘ t ’ represents quadratically added (root of sum of squares) errors due to many temporally and spatially varying parameters and measurement noise with $\sigma_N = 10^{-4}$ W/(m² sr μ m), see Kappel et al. (2015, Table 4), but only the largest contributors are listed here in the sub-rows. ‘Cloud bottom’: base altitude for cloud modes 1, 2’, 3; ‘H₂SO₄’: H₂SO₄ concentration of cloud droplets; ‘H₂O’ and ‘HCl’: column factors for these minor gases; ‘ ΔT_0 ’, ‘ ΔT_{25} ’: deep atmospheric temperature profile near 0 or 25 km; ‘Continua’: quadratically added contributions from the continua of all considered windows, among them from the 1.74 and the 2.3 μ m continua. Impact of reasonable *a priori* mean value modifications: $\leq 1\%$.

Calibration, preprocessing, and simulation errors can lead to additional emissivity errors. But calibration and preprocessing were improved by Kappel et al. (2012), and the forward model is capable of simulating spectra that extremely well fit the measurements for a large variety of physical and observational conditions. Calibration, preprocessing, and functional descriptions of the simulated spectra are not expected to be severely wrong, therefore. Even if they were wrong, renormalized emissivities would not be strongly affected by that. Interfering parameters that were not studied by Kappel et al. (2015) can also introduce further errors, for instance varying cloud mode particle size distributions or chemical compositions, further details of the cloud mode altitude distributions, unknown absorbers like possible haze close to the surface (Grieger et al., 2004; Erard et al., 2009), etc. But impacts of systematically wrong assumptions on these parameters should be greatly diminished through the renormalized emissivity retrieval, and impacts of quasi-random time variations around the long-term averages should disappear with the inverse square root of the repetition number.

One can try to reduce the emissivity errors by retrieving one of the interfering parameters whose uncertainty causes large emissivity errors (like H₂O) as a further parameter. But this can only succeed, when the parameter can be reliably determined. A retrieval error analysis like that reported by Kappel et al. (2015) would be required to decide this. Haus et al. (2015) performed such an analysis for the retrieval of minor gas abundances from VIRTIS-M-IR measurements at the long wavelength flank of the 2.3 μm nightside peak with the result that H₂O can be determined with a moderate accuracy. This was a different wavelength range than used for the present work, and results cannot be carried over. But nevertheless, this is used as an opportunity to optionally include the H₂O abundance as a retrieval parameter here (see Section 2.4) in contrast to Kappel et al. (2015), with the aim to slightly improve emissivity results.

4.2. Measurement Selection Tests and Parameter Modification Tests

According to Table 3, the renormalization approach is thus predicted to enable much more useful results at $N_r = 64$ than at $N_r = 1$, but it involves a number of subtleties, and errors are still not negligible. Moreover, the errors listed in Table 3 were estimated from synthetic spectra. Thus, it is advisable to verify their applicability to real data. To examine the reliability of actual renormalized emissivity maps, two types of tests are designed, in the following called 'Measurement Selection Test' (MST) and 'Parameter Modification Test' (PMT).

The VIRTIS-M-IR data can be regarded as measurements of surface emissivity maps. The MSTs shall explore, whether these measurements are repeatable within the frame of a certain (acceptable) measurement uncertainty, provided that geologic activity as would be observable by VIRTIS-M-IR is negligible. For this purpose, maps are determined from S disjoint data sets ('batches') with N_r repetitions each, using the same retrieval pipeline as for N_r^0 . These sets should lead to maps that are reasonably similar among each other. At a fixed bin, the scatter of the renormalized emissivities around their mean value provides a measure of the emissivity accuracy that is achievable at N_r repetitions. This measure can be scaled to estimate the uncertainty of the N_r^0 map. This is discussed in detail in Appendix A.2.3. Since the resulting raw emissivity mean values may differ for the different maps, it is particularly important at this point to refer all the maps to the same reference emissivity in view of the mean value dependence of the amplitude of spatial emissivity fluctuations around their mean value. To facilitate a direct comparison of the various error measures used in the present work, all of them are represented in terms of double standard deviations in percent of the reference emissivity 0.5 (which is at the same time the mean renormalized emissivity), referred to $N_r^0 = 64$ repetitions. For the MSTs, surface elevation and the other temporally constant interfering parameters and their uncertainty impacts are identical (at least up to first order in perturbations by the remaining interfering parameters). Only temporally varying parameters can lead to significant differences. Disagreements between the maps that are too large to be explained by the values in Table 3 may be due to calibration and preprocessing errors, may indicate that there are temporally varying interfering parameters that have not been considered here, may hint at an underestimation of the probable range the interfering parameters are assumed to cover, may be due to actual time variations of the emissivities, etc. In order to verify the statistical properties, MSTs are performed with the combinations $N_r = 32$ & $S = 2$, $N_r = 16$ & $S = 4$, and $N_r = 8$ & $S = 8$, such that always $N_r S = N_r^0$.

With the PMTs, it shall be checked, whether the renormalized emissivity maps are reasonably independent with respect to modifications of the most important interfering parameters. These include parameters that have to be set to assumed values to allow for the generation of realistic synthetic spectra and that yield the majority of the errors listed in Table 3. The dependence of the maps on the choice of the initial emissivity and the continua is also examined. In addition, a number of modifications to the details of the retrieval pipeline as well as the calibration and preprocessing will be studied. Since even the computation of a single map requires considerable computer resources, not the full 64

repetition data set is utilized, but a data set leading to 25 repetitions. The results are statistically scaled to allow for uncertainty interpretations of the $N_r = 64$ case. The mathematical background of the PMTs and their relation to the MSTs are discussed in Appendix A.2.4.

Both reliability tests are complementary in the sense that for the MSTs, the parameters that are not retrieved stay constant between batches in the forward model simulations whereas their true underlying values vary, and vice versa for the PMTs. The MSTs can be regarded to yield a measure for the random errors, and the PMTs for the systematic errors. Results of both tests are presented in Sections 4.3 and 4.4.

Each of these tests requires one or more full retrieval runs ('scenarios'). Each scenario is based on a certain measurement data set comprising N_r measurement repetitions, the configuration of the retrieval pipeline, the choice of the forward model parameters, and the details of the calibration and preprocessing. An abbreviation (in **bold**) and a description are given in the following for each scenario. The abbreviations are used in Table 4 (column 'Scenario'). This table lists for each of the three surface windows the renormalization coefficients of the emissivity maps, as well as a statistical comparison measure. The latter can be found in the column 'SRMSD' (scaled root-mean-square-deviation, see comment before Eq. (A.3)) and quantifies the deviation of the respective scenario's renormalized emissivity map from a reference map, in a reasonable sense scaled to be directly comparable to the N_r^0 scenario, and represented as double standard deviations in percent of the reference emissivity 0.5. In case of an MST with S batches, the reference map is the average over the batches, and for the PMTs, it is scenario $N_r = \mathbf{25\ base}$ which is completely analogous to scenario ' $N_r^0 = 64$ ', except that the data set has only 25 repetitions. For details, see Appendix A.2. The VIRTIS-M-IR cubes utilized for these tests are listed in Appendix A.1.

4.3. Measurement Selection Tests

Three MSTs are performed and are used to estimate the retrieval uncertainty with respect to the selection of the measurement data set, see Appendix A.2.3.

$N_r = \mathbf{32, batch\ 1 / batch\ 2}$: Renormalized emissivity maps are determined for two disjoint data sets (batches 1 and 2) with 32 repetitions each. The maps at $1.02\ \mu\text{m}$ are depicted in Figs. 4a & b. Although there are differences, the magnitudes and spatial distributions of the deviations from the respective map averages 0.5 (green) are very similar between both maps. As it was discussed, the difference between the maps is a measure for the uncertainty of the results. Also compare to Fig. 3a showing the more reliable corresponding map determined from the $N_r^0 = 64$ data set, which is very similar to both Figs. 4a & b. $N_r = \mathbf{16, batches\ 1-4}$: The $N_r^0 = 64$ data set is divided into four disjoint $N_r = 16$ data sets. $N_r = \mathbf{8, batches\ 1-8}$: The $N_r^0 = 64$ data set is divided into eight disjoint $N_r = 8$ data sets.

Results of the three tests are divided by single horizontal lines in Table 4. The quadratically averaged SRMSDs of the S batches of a given MST yield an estimate of the uncertainty of the renormalized emissivity map at N_r^0 (see discussion below Eq. (A.4)). Table 5 summarizes these estimates for the three MSTs and gives a rough estimate of these measures' uncertainties (according to Appendix A.2.1).

4.4. Parameter Modification Tests

Based on $N_r = \mathbf{25\ base}$ (see Fig. 4c showing the renormalized $1.02\ \mu\text{m}$ map), it is now investigated, whether renormalized emissivity maps are reasonably independent with respect to modifications of the most important interfering parameters that are not sufficiently well known. For each PMT, one of the interfering parameters is modified by setting it to a value that differs by a reasonable amount from the value utilized for the retrieval of the map $N_r = \mathbf{25\ base}$. The corresponding renormalized emissivity map is determined and compared to the base map (root-mean-square-deviation, scaled to N_r^0 and represented as double standard deviations in percent of the reference emissivity, see Appendix A.2.4).

The studied parameters can be in principle temporally varying (indicated by uppercase letter ' T ' on the right of the 'Scenario' column in Table 4), spatially varying but time independent (' L '), or constant in time and space (' C '). However, in the tests, the modifications can have a different character. For instance, the cloud bottom altitude can in principle vary with time (' T '), but the modification in the test is constant in time and space. This is indicated by a lowercase ' c ' in the table. In the same way, parameter modifications that vary with time are indicated by ' t ', and those that are time independent but vary in space by ' l '. The lowercase letters coincide with the categorization into t -, l -, and c -parameters in Appendix A.2 and decide the scaling behavior of the renormalized emissivity errors. Parameter modifications that theoretically (see Eq. (A.6) and also Eq. (A.3)) do not lead to renormalized emissivity retrieval errors, because their impact is annihilated by the renormalization procedure (c - and spatially slowly varying l -parameters), are indicated by '0' on the right of the 'Scenario' column, the others by '1'.

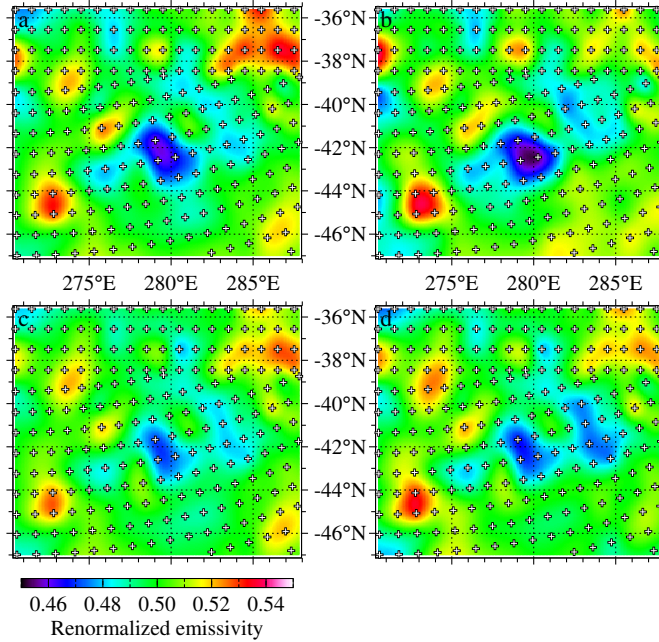


Figure 4: Examples for MSTs and PMTs showing renormalized $1.02\ \mu\text{m}$ emissivity maps of Themis Regio referred to reference emissivity 0.5. (a) MST ' $N_r = 32$, batch 1', (b) MST ' $N_r = 32$, batch 2', (c) PMT ' $N_r = 25$ base', (d) PMT 'Continuum $1.74\ \mu\text{m}$ '. Axes as in Fig. 1. The corresponding results at 1.10 and $1.18\ \mu\text{m}$ are given in Figs. A.5 and A.6 in the supplementary material.

The scaled PMT error results can be used to roughly estimate the impact that such modifications would have on the N_r^0 -case. This can be done by quadratically adding them to the MST estimates, see Appendix A.2.4. However, some of the PMTs study the same parameter but in different ways, other PMTs have been performed just for general interest. Thus, only the PMT errors indicated by superscript '**' in the 'Scenario'-column of Table 4 are added to the MST error estimate. The quadratically added errors of the flagged PMTs are shown in Table 5 in the 'PMTs'-row. The quadratic sum of the latter and the largest of the three MST estimates is represented for each of the three surface windows in the 'Total'-row of this table. These values provide an estimate for the total error of the N_r^0 -maps in Fig. 3, represented as double standard deviations in percent of the reference emissivity.

The PMT results are listed below the double horizontal line in Table 4. The modified parameters fall into three groups, forward model parameters, modifications to the characteristics of the retrieval pipeline, and modifications to calibration and preprocessing. Results for these groups are divided by single horizontal lines in the table.

Forward model parameters

Init. emiss. 0.625 / 0.750 / 0.875: The initial emissivity e_{init} is set to 0.625 / 0.750 / 0.875, respectively, instead of 0.500. The results listed in Table 4 demonstrate for the PMTs 'Init. emiss. 0.625' and 'Init. emiss. 0.750' that the choice of the initial emissivity does not significantly affect renormalized emissivity maps referred to 0.5, SRMSDs are smaller than 0.4%. Note that this would not be the case without reference emissivity transformation. The best choice for the initial emissivity (or rather the raw emissivity mean value) would be the surface emissivity average of the planet. Fegley et al. (1997) reviewed that weathering products like iron oxides could be widely present on Venus' surface. This corresponds with the argumentation by Smrekar et al. (2010) that the background emissivity of Venus' southern hemisphere is 0.5-0.6 at $1\ \mu\text{m}$, a value consistent with fine-grained hematite (Baldrige et al., 2009). Moreover, it is important to use an initial emissivity value as far away from the emissivity domain boundaries 0 and 1 as possible, in particular when the presence of trends with latitude and topography in the raw retrieved emissivities cannot be excluded, as it was explained in Section 2.3. The results for the value 0.875 illustrate this point, as the SRMSD at $1.10\ \mu\text{m}$ is the largest of all PMT results. A detailed inspection reveals that the raw emissivities for many bins are cut off at the upper emissivity domain boundary 1 (raw emissivity mean value 0.99 !). The initial emissivity that otherwise leads to the largest SRMSDs (0.750) is selected to contribute to the total error in Table 5.

Scenario	Mean	100×Slope/°lat	10×Slope/km	SRMSD
$N_r^0 = 64$	0.48 0.67 0.62	0.14 0.28 0.13	0.34 -0.70 -0.00	
$N_r = 32$, batch 1	0.49 0.64 0.66	0.10 0.65 0.14	0.29 -0.62 -0.08	1.56 4.87 2.55
$N_r = 32$, batch 2	0.46 0.70 0.59	0.20 -0.15 0.23	0.29 -0.70 -0.01	1.56 4.87 2.55
$N_r = 16$, batch 1	0.49 0.65 0.61	0.06 1.07 0.04	0.36 -0.75 0.04	1.91 6.27 2.02
$N_r = 16$, batch 2	0.48 0.67 0.60	0.08 0.34 0.09	0.37 -0.48 0.04	1.92 7.35 2.56
$N_r = 16$, batch 3	0.49 0.71 0.66	0.09 -0.09 -0.19	0.35 -0.77 0.03	3.15 9.08 3.61
$N_r = 16$, batch 4	0.48 0.72 0.65	0.22 0.48 0.51	0.38 -0.60 -0.00	2.56 8.13 2.85
$N_r = 8$, batch 1	0.49 0.67 0.62	-0.04 0.71 0.10	0.39 -0.70 -0.07	1.75 3.38 2.39
$N_r = 8$, batch 2	0.49 0.64 0.59	0.11 1.50 -0.06	0.29 -0.82 0.13	1.52 8.09 2.57
$N_r = 8$, batch 3	0.49 0.64 0.59	-0.09 0.35 0.23	0.24 -0.72 0.01	1.83 6.19 2.14
$N_r = 8$, batch 4	0.48 0.68 0.61	0.18 0.29 -0.04	0.46 -0.27 0.06	2.55 6.89 2.21
$N_r = 8$, batch 5	0.49 0.72 0.66	-0.21 -1.03 -0.74	0.43 -0.54 0.16	3.22 9.10 4.65
$N_r = 8$, batch 6	0.49 0.70 0.66	0.34 0.91 0.37	0.26 -0.99 -0.09	2.47 7.81 2.90
$N_r = 8$, batch 7	0.48 0.70 0.65	0.28 1.12 0.68	0.35 -0.89 -0.24	3.00 9.72 3.68
$N_r = 8$, batch 8	0.48 0.72 0.64	0.14 -0.16 0.28	0.38 -0.35 0.22	2.40 7.95 4.24
$N_r = 25$ base	0.48 0.64 0.59	0.08 0.43 0.02	0.22 -0.74 -0.02	0 0 0
Init. emiss. 0.625 <i>Cc0</i>	0.60 0.77 0.72	0.10 0.48 0.03	0.23 -0.81 -0.08	0.11 0.15 0.10
Init. emiss. 0.750 * <i>Cc0</i>	0.71 0.90 0.85	0.10 0.52 0.04	0.25 -0.88 -0.15	0.19 0.39 0.17
Init. emiss. 0.875 <i>Cc0</i>	0.83 0.99 0.97	0.11 0.12 0.04	0.27 -0.34 -0.20	0.25 3.65 0.40
Continuum 1.74 μm * <i>Cc0</i>	0.48 0.65 0.59	0.16 0.71 0.17	0.23 -0.98 -0.23	0.73 1.17 1.49
Continuum 2.3 μm * <i>Cc0</i>	0.48 0.62 0.57	0.08 0.40 0.02	0.23 -0.62 0.06	0.14 0.39 0.26
Temp. 0 km <i>Lc0</i>	0.48 0.63 0.59	0.08 0.46 0.01	0.24 -0.66 0.01	0.17 0.24 0.11
Temp. 25 km <i>Lc0</i>	0.48 0.63 0.59	0.08 0.50 0.02	0.24 -0.71 0.01	0.08 0.16 0.11
Temp. 25 km -0.3 K <i>Lc0</i>	0.48 0.64 0.59	0.07 0.49 0.00	0.22 -0.75 -0.02	0.04 0.10 0.06
Temp. 0 km trend * <i>Ll0</i>	0.47 0.65 0.60	-0.23 0.56 -0.34	0.20 -0.81 -0.05	0.34 0.52 0.36
Temp. 25 km trend * <i>Ll0</i>	0.48 0.65 0.60	0.17 0.60 0.07	0.20 -0.78 -0.04	0.08 0.15 0.11
Cloud bottom * <i>Tc0</i>	0.47 0.67 0.61	0.05 0.48 -0.05	0.21 -0.82 -0.07	0.10 0.26 0.24
H ₂ SO ₄ * <i>Tc0</i>	0.48 0.65 0.60	0.14 0.63 0.11	0.28 -0.70 -0.03	0.55 0.70 0.93
H ₂ SO ₄ alt. gradient <i>Tc0</i>	0.48 0.68 0.61	0.08 0.58 0.01	0.23 -0.88 -0.16	0.22 0.64 0.76
H ₂ O * <i>Tc0</i>	0.52 0.60 0.58	0.05 0.97 0.31	0.38 -0.86 -0.04	0.89 1.07 0.64
HCl * <i>Tc0</i>	0.48 0.63 0.59	0.07 0.47 0.00	0.22 -0.70 0.02	0.08 0.18 0.15
Surface haze <i>Tc0</i>	0.46 0.69 0.62	0.14 0.63 0.04	0.57 -1.28 -0.23	1.53 1.75 0.69
No m_1 retrieval <i>Tt1</i>	0.50 0.72 0.72	0.14 0.65 0.21	0.28 -0.81 -0.15	0.36 0.71 0.78
No H ₂ O ^b retrieval <i>Tt1</i>	0.50 0.62 0.65	0.10 0.55 0.17	0.24 -0.51 -0.00	0.33 1.03 0.82
Pipeline II <i>Tt1</i>	0.52 0.55 0.53	0.08 0.99 0.37	0.31 -0.58 0.05	0.52 0.83 0.55
Pipeline II, normal noise <i>Tt1</i>	0.52 0.55 0.53	0.08 0.99 0.37	0.31 -0.58 0.04	0.51 0.82 0.54
Pipeline II+ m_1 +H ₂ O ^b * <i>Cc0</i>	0.48 0.63 0.58	0.07 0.55 0.03	0.22 -0.82 -0.02	0.16 0.44 0.25
Wavel. dep. FWHM * <i>Cc0</i>	0.49 0.59 0.56	0.05 0.48 0.06	0.21 -0.67 0.02	0.15 0.53 0.32
Detector responsivity * <i>Cc0</i>	0.47 0.70 0.62	0.08 0.54 0.04	0.16 -1.13 -0.06	0.28 0.88 0.41
Unsmoothed straylight * <i>Tt1</i>	0.48 0.64 0.59	0.07 0.26 0.10	0.24 -0.55 -0.08	0.65 1.86 0.61
Topo. 65 km <i>Ll1</i>	0.48 0.64 0.59	0.06 0.43 0.02	0.25 -0.72 -0.00	0.58 0.52 0.33
Topo. 65 km R100 <i>Ll1</i>	0.48 0.64 0.59	0.06 0.43 0.02	0.25 -0.75 -0.01	0.80 0.34 0.32
Topo. 30 km <i>Ll1</i>	0.48 0.64 0.59	0.07 0.51 0.01	0.29 -0.68 0.02	1.16 1.29 0.65
Topo. 30 km R100 * <i>Ll1</i>	0.48 0.64 0.59	0.08 0.51 0.02	0.28 -0.75 0.01	1.60 0.49 0.67

Table 4: Results of MSTs and PMTs. Detailed explanations can be found in Section 4. The entries in column ‘Scenario’ correspond to the bold-typed shortcuts used in the descriptions in Sections 4.3 and 4.4. Renormalization coefficients and SRMSDs are given in the next four blocks of columns. Each block shows the data for the surface windows at 1.02, 1.10, and 1.18 μm next to each other. Mean values of the retrieved emissivity maps before renormalization are listed in the ‘Mean’ columns, the slopes of linear trends with latitude, scaled with a factor 100 for better representation, in the columns ‘100 × Slope/°lat’, and the scaled slopes of linear trends with surface elevation in the columns ‘10 × Slope/km’. The SRMSDs are comparison measures referred to $N_r^0 = 64$ repetitions and are represented as double standard deviations in percent of the reference emissivity 0.5.

Continuum 1.74 μm / 2.3 μm : The continuum parameters $c_{1.74}$ and $c_{2.3}$, respectively, are increased by 25%. SRMSDs contribute to the total error. Uncertainties of these continua potentially cause large errors without renormalization (Table 3), but these PMTs demonstrate that renormalized emissivities are only moderately affected. This is illustrated for ‘Continuum 1.74 μm ’ when comparing the resulting renormalized map at 1.02 μm represented in Fig. 4d to Fig. 4c showing the map for ‘ $N_r = 25$ base’. The respective SRMSD is among the largest of those that contribute to the total error at 1.02 μm . Recall that the surface window continua are retrieved as fine-tuning parameters (Section 2.3), and that the 1.31 μm peak is not used here for emissivity retrieval.

Temp. 0 / 25 km: ΔT_0 and ΔT_{25} , respectively, are decreased by 3 K to modify the deep atmospheric temperature

profile close to the surface and around 25 km altitude. These modifications hardly affect the renormalized maps. To estimate the impact that an even smaller modification has, **Temp. 25 km -0.3 K** decreases the temperature around 25 km by only 0.3 K. Even though the magnitude of this modification is a tenth of the one for the preceding test, the SRMSDs are only about halved. The difference to the base case seems not to fast converge to zero. This indicates a limited numerical repeatability of the retrieval under slightly altered conditions, but the effect is not large compared to the other PMTs. **Temp. 0 / 25 km trend:** ΔT_0 and ΔT_{25} , respectively, are set to exhibit a $\cos(\text{latitude})$ -trend with a range of 3 K over the latitude range $[\theta_1, \theta_2] = [-46.99^\circ, -35.63^\circ]$ of the target area, $(\cos \theta - \cos \theta_1) / (\cos \theta_2 - \cos \theta_1) \cdot 3$ K. The magnitude of this trend is exaggerated as it would correspond to a range of 23 K between equator and pole, whereas the order of a few K seems to be more realistic. Note the strong impact on the latitude trend renormalization coefficient for 1.02 μm in case of the ΔT_0 test. Nevertheless, the effect on the renormalized emissivity maps is surprisingly small, in particular for ΔT_{25} , even though the renormalization not exactly annihilates this only approximately linear trend of the parameters (see Appendix A.2.2). SRMSDs contribute to the total error.

Cloud bottom: The h_b for cloud modes 1, 2', 3 are jointly decreased by 3 km. According to Kappel et al. (2015) (see Table 3 of the present work), this is expected to be one of the largest error contributors for parameters that can temporally vary, along with the H_2SO_4 concentration of the cloud droplets and the abundances of H_2O and HCl . But this test demonstrates that a constant misplacement of the cloud bottom hardly affects the renormalized emissivities. SRMSDs contribute to the total error.

H_2SO_4 / H_2SO_4 alt. gradient: The H_2SO_4 concentration of the cloud droplets is set in the first test to 87.5% instead of 75.0%. For the other test, the concentration is set to change with altitude according to an abstracted form of the profile shown by James et al. (1997) (linear function with slope -2.5 %/km, attaining 88% at 55 km, cutoff at 50% and 95.6%). These modifications yield much larger errors than the cloud bottom modifications. SRMSDs of the H_2SO_4 test with the larger errors (the first one) are selected to contribute to the total error.

H_2O / HCl : The H_2O and the HCl column factors, respectively, are decreased by 25%. While the HCl concentration change has almost no impact, H_2O has a larger response. Note that H_2O^b is not retrieved for the H_2O test (see Section 2.4), so it really constitutes a change of the retrieval pipeline as well. SRMSDs of both tests contribute to the total error.

Surface haze: An additional absorber is considered below 2000 m altitude, for simplicity modeled as gas (0.1 ppmv) with wavelength independent molecular absorption cross-section 10^{-20} cm^2 . This test is performed to illustrate the impact of this hypothetical parameter that is normally not included in the forward model. It has a quite strong effect on the topography trend renormalization coefficients, and the error contribution is among the largest of all PMTs. Note that the bins whose surface elevation reaches above this haze yield the largest errors. But as a parameter whose role is not clear yet, its SRMSDs are excluded from the total error. Since this haze can in principle change with time, it is categorized as a T -parameter, but the PMT only changes the abundance by a constant value and has c -character, therefore.

Retrieval pipeline

No m_1 retrieval: The retrieval of cloud mode factor 1 in stages six and eight of the pipeline is omitted. The character of this PMT is difficult to specify. On the one hand, the difference of m_1 to the retrieval result from the base case is naturally time dependent. On the other hand, this is the case for all the other PMTs, too, since the retrieved m_1 values are not likely to always exactly agree. However, for this PMT, m_1 differs as a matter of principle. This PMT is thus set to have T - and t -character. This is corroborated by the consideration that the renormalized emissivity retrieval errors due to m_1 uncertainties should average out more efficiently with increasing N_r . **No H_2O^b retrieval:** Here, the retrieval of H_2O^b in stages six and eight of the pipeline is omitted. Again, this leads to a T - and t -character. The error contribution is quite large in the 1.10 and 1.18 μm windows.

Pipeline II: This is a test, where several characteristics of the retrieval pipeline are modified. The *a priori* correlation data is slightly modified, the standard deviation σ_N of the measurement and simulation error distribution is doubled, and, as in the two preceding tests, m_1 and H_2O^b are not retrieved. Thus, it has T - and t -character. The error contribution is smaller than could be expected from all these changes. **Pipeline II, normal noise:** This PMT is the same as the one before, except that σ_N is set to the usual value. The impact compared to 'Pipeline II' is marginal. **Pipeline II+ m_1 + H_2O^b :** Here, the retrieval of m_1 and H_2O^b is reintroduced compared to 'Pipeline II'. The PMT has therefore C - and c -character. The only differences to the base case are the modified *a priori* correlation data and the doubled σ_N . Since operational retrievals in the present work always include m_1 and H_2O^b retrieval, but the correlation data and σ_N were chosen somewhat arbitrarily, the resulting SRMSDs contribute to the total error. However, this PMT demonstrates

that certain details of the retrieval pipeline can be less important than the choice of the value for some of the interfering forward model parameters.

Calibration and preprocessing

Wavel. dep. FWHM: The FWHM of the spectral instrumental response function is not sufficiently well predictable by the calibration pipeline. In this work, it is normally retrieved for each spectrum as a value that is assumed to be constant with wavelength. For this PMT, it is modeled to change in dependence on wavelength according to a quadratic function of the spectral bands b . The offset to the retrieved wavelength constant value amounts to 0 nm at the first ($b = 1, \approx 1.02 \mu\text{m}$) and the last ($b = 432, \approx 5.1 \mu\text{m}$) spectral bands and to +3 nm in the middle ($b = 216.5, \approx 3.1 \mu\text{m}$). This is a PMT that is parameterized by a single constant value and has thus C - and c -character. The error impact is not large. SRMSDs contribute to the total error.

Detector responsivity: One of the calibration refinement issues discussed by Kappel et al. (2012) constitutes an adjustment of the detector responsivity in the spectral dimension resulting in an increase of the calibrated radiance between 0% at $1.3 \mu\text{m}$ and 35% at $1.0 \mu\text{m}$. To examine the effect of possible errors of this adjustment, the amount of the correction is halved for this C - and c -PMT. Surprisingly, this has only a small effect on the SRMSD for the surface window with the largest responsivity modification, the one at $1.02 \mu\text{m}$. At 1.10 and $1.18 \mu\text{m}$, the impact is larger. SRMSDs contribute to the total error. Note that the impact of a spatial detector inhomogeneity as it was present before the calibration refinement is not studied here. But error contributions are expected to be large and not well reproducible, because they would depend on details like which part of the surface target is probed by which part of the detector. However, they would not be included in the total errors, since the inhomogeneity is corrected quite well.

Unsmoothed straylight: Normally, the straylight is assumed in this work to show no spatial fine structure over the spatial dimensions, but only instrument-related general trends that vary slowly compared to the spatial extension of a single VIRTIS-M-IR image. This can be justified by the nature of the most relevant straylight source where Venus' dayside is outside but close to the VIRTIS-M-IR field of view (Kappel et al., 2012, Section 4.1.1, issue 3). Therefore, the straylight, which is computed for each (unbinned) spectrum of a VIRTIS-M-IR cube, is normally median-smoothed over the cube's spatial dimensions before being subtracted from the measured spectra. The size of the two-dimensional neighborhood used for smoothing is 64×64 pixels for the utilized 256-sample-mode cubes. For this PMT, the smoothing is not performed. The result is not very realistic but can serve to estimate the impact of errors in the determination of the straylight spectra. The difference between smoothed and unsmoothed straylight statistically varies with time and should more efficiently average out for increased N_T . This PMT has T - and t -character, therefore. Note that the amount of straylight is clearly correlated with the distance of the field of view to the terminator. However, most of this correlation is removed for the difference between smoothed and unsmoothed straylight. SRMSDs contribute to the total error and are very large at $1.10 \mu\text{m}$ ($\approx 2\%$). This is because this radiance peak is the smallest of the surface window peaks and thus the most affected by errors in the straylight removal.

Topo. 65 km: The emulation of the atmospheric blurring of the surface topography normally assumes a FWHM of 100 km. Now, the FWHM is set to 65 km. This can still be regarded as a realistic value and is probably valid for observations with few clouds. This test can also serve to emulate the effect of slight errors in the underlying Magellan topography and the blurring approach just in the most critical areas, the mountainous terrains, where the topography is spatially particularly fast varying and topography errors are more likely. For the target area, the topographies for the 65 km- and the 100 km-blurring differ by up to about 125 m (root-mean-square-deviation 35 m). This PMT has L - and l -character, but the impact is not annihilated by the renormalization, because the topography difference is spatially fast varying. However, this PMT is explicitly excluded in the discussions in Appendix A.2 and must be treated separately, because the renormalization operator itself depends on topography. For this PMT, the renormalization is performed using the 65 km-blurring, whereas the base case renormalization naturally relies on the 100 km-blurring. For comparison, **Topo. 65 km R100**, which otherwise coincides with 'Topo. 65 km', is renormalized based on the 100 km-blurring. While the SRMSD increases at $1.02 \mu\text{m}$ because of the modified renormalization, it surprisingly decreases at the other two surface windows. **Topo. 30 km / Topo. 30 km R100:** These PMTs are completely analogous to the two preceding cases, but the FWHM is even reduced to 30 km. This is probably not realistic anymore but helps to further quantify emissivity errors due to topography errors. The topographies for the 30 km- and the 100 km-blurring differ by up to about 250 m (RMSD 63 m). Of the topography PMTs, the largest SRMSDs for each window are selected to contribute to the total error ('Topo. 30 km' at $1.10 \mu\text{m}$, 'Topo. 30 km R100' otherwise, but only the latter is indicated with superscript '**' in the 'Scenario'-column of Table 4).

N_r	S	R_{64}			Δ
32	2	1.6	4.9	2.6	10
16	4	2.4	7.8	2.8	6
8	8	2.4	7.6	3.2	4
PMTs		2.2	3.2	2.2	
Total		3.3	8.4	4.0	
Synthetic		2.4	6.1	7.0	
Synth. rev.		1.9	5.8	6.2	

Table 5: Summary of MST and PMT results. The three MST rows below the first horizontal line show for the surface windows at 1.02, 1.10, and 1.18 μm the quadratically averaged SRMSDs from Table 4 and represent the uncertainty of renormalized emissivities with respect to the selection of the measurement data set, determined from the MSTs ' $N_r = 32$, batches 1–2', ' $N_r = 16$, batches 1–4', and ' $N_r = 8$, batches 1–8'. They coincide with the statistical error estimate for the deviation of the N_r^0 -map from the true renormalized emissivity map (see $\delta^{N_r, S}$ from Appendix A.2.3). 'PMTs': quadratic sum of SRMSDs from '*'-indicated PMTs in Table 4. 'Total': quadratic sum of PMT estimate and maximum of the MST estimates. 'Synthetic': quadratic sum of errors from Table 3, but excluding those due to surface elevation uncertainties. 'Synth. rev.': revised synthetic errors, see Section 5. Values are double standard deviation errors in percent of the reference emissivity referred to 64 repetitions. Δ : double standard deviation errors of MST estimates in percent according to Appendix A.2.1.

4.5. Reference emissivity

As discussed in Section 2.3, when trends of a raw retrieved emissivity map with latitude and topography are removed such that the emissivity mean value M is not changed, the thus de-trended emissivities spatially fluctuate around M with a higher amplitude for cases with higher M . This effect can be observed by comparing emissivities computed with respect to different initial emissivities e_{init} as it was done in the frame of the PMTs 'Init. emiss. 0.625 / 0.750 / 0.875' in Section 4.4. M can differ from e_{init} as explained earlier and is listed for different cases in the 'Mean' columns in Table 4 in the rows ' $N_r = 25$ base' ($e_{\text{init}} = 0.500$), 'Init. emiss. 0.625', 'Init. emiss. 0.750', and 'Init. emiss. 0.875'. M_0 refers to the case with $e_{\text{init}} = 0.500$ (e.g. $M_0 = 0.48$ at 1.02 μm according to the table). To avoid the dependence of fluctuation amplitudes on an uncontrollable M when performing quantitative comparisons between different emissivity maps, for instance in the frame of the retrieval error estimations, the emissivities have to be referred to a certain fixed reference emissivity, here 0.5, with then directly comparable fluctuation amplitudes. The results can be later transformed to other reference emissivities if needed.

The transformation between different reference emissivities, and therefore different values of M , will now be derived. For this purpose, the trends of a given raw retrieved emissivity map at a fixed surface window with latitude and topography are removed by multiple linear regression in such a way that the mean emissivity M is unchanged. The resulting values at the different surface bins are compiled into the vector \mathbf{r}_M .

The scatterplots of $\mathbf{r}_M - M$ for the case $e_{\text{init}} = 0.750$ in dependence on $\mathbf{r}_{M_0} - M_0$ are shown in Fig. 5 (left). Note that the mean value of $\mathbf{r}_M - M$ is zero for each surface window, as it is the case with $\mathbf{r}_{M_0} - M_0$. The fluctuations are small compared to the emissivity mean, giving rise to a first order model. The scatterplots approximately follow the regression function $(\mathbf{r}_M - M) = \lambda_{M_0 \rightarrow M} \cdot (\mathbf{r}_{M_0} - M_0)$ with the slope $\lambda_{M_0 \rightarrow M}$ that transforms results for cases with mean value M_0 into those with mean value M . This slope depends on M_0 and M (see Fig. 5, right) with regression function $\lambda_{M_0 \rightarrow M} = 0.137 + 1.802 \cdot M$ at 1.02 μm , $\lambda_{M_0 \rightarrow M} = 0.649 + 0.556 \cdot M$ at 1.10 μm , and $\lambda_{M_0 \rightarrow M} = 0.840 + 0.283 \cdot M$ at 1.18 μm . These regressions are regarded as valid for M larger than ≈ 0.5 and not close to 1 to avoid extrapolation and stay in the linear regime. Note that the case $e_{\text{init}} = 0.875$ at 1.10 μm is unreliable as explained, because $M = 0.99$ is too close to 1. It is therefore excluded from the regression. Since $(\mathbf{r}_{0.5} - 0.5) = \lambda_{M_0 \rightarrow 0.5} \cdot (\mathbf{r}_{M_0} - M_0)$, the 'renormalized emissivity map referred to the reference emissivity 0.5' as it was introduced in Section 2.3 follows for a given surface window as $\mathbf{r}_{0.5} = 0.5 + \lambda_{M_0 \rightarrow 0.5} / \lambda_{M_0 \rightarrow M} \cdot (\mathbf{r}_M - M)$. As already mentioned, the maps will here always be referred to 0.5, except when stated otherwise, and $\mathbf{r}_{0.5}$ is shortly called 'renormalized emissivity map'.

A renormalized emissivity map already referred to 0.5 can be referred to another reference emissivity M according to $\mathbf{r}_M = M + \lambda_{0.5 \rightarrow M} \cdot (\mathbf{r}_{0.5} - 0.5)$, where $\lambda_{M' \rightarrow M} := \lambda_{M_0 \rightarrow M} / \lambda_{M_0 \rightarrow M'}$. Note that the dividing of $\lambda_{M_0 \rightarrow M}$ by $\lambda_{M_0 \rightarrow 0.5}$ ensures the scaling factor to attain the value 1 at $M = 0.5$. The impact of such a reference emissivity change on the first main result of this work represented in Fig. 3 can be estimated by simply scaling the tick labels of the color bar for the renormalized emissivities according to this transformation. As an example, for the reference emissivity $M = 0.75$, the color bar range is transformed from $[0.450, 0.550]$ to the interval $0.250 + [0.428, 0.572]$ at 1.02 μm , to $0.250 + [0.443, 0.557]$ at 1.10 μm , and to $0.250 + [0.446, 0.554]$ at 1.18 μm . When the true absolute emissivity \hat{A} of just

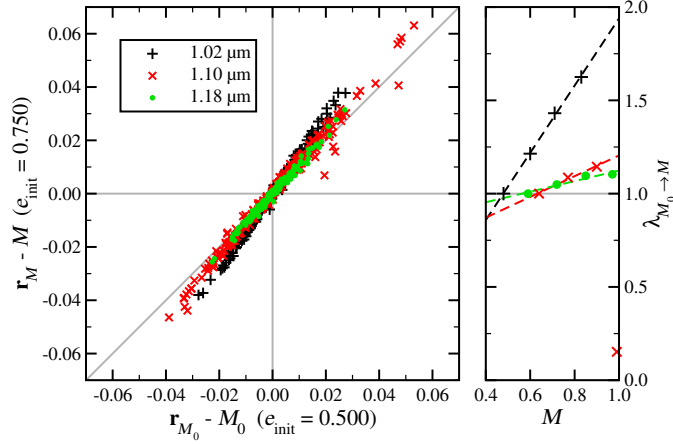


Figure 5: Dependence of emissivity fluctuations on their raw mean value M for the three surface windows. Left: Scatterplots between $r_M - M$ for the case $e_{\text{init}} = 0.750$ and $r_{M_0} - M_0$ ($e_{\text{init}} = 0.500$). Right: Slopes $\lambda_{M_0 \rightarrow M}$ of the scatterplots for different M (symbols) and corresponding linear regression functions (dashed), see text for explanations.

a single surface bin b of the target with renormalized retrieved emissivity $r_{0.5}(b)$ were known, for instance by ground truth, the absolute emissivity map of the entire target could be computed according to the mentioned transformation, assuming absent trends of the true emissivity map with latitude and topography. For this purpose, only a suitable M has to be found. Writing $\lambda_{0.5 \rightarrow M} = p + q \cdot M$, it follows $M = (\hat{A} - p \cdot (r_{0.5}(b) - 0.5)) / (1 + q \cdot (r_{0.5}(b) - 0.5))$.

As a consequence of the mentioned linear transformation, a small emissivity perturbation $\delta_{0.5}$ in the renormalized emissivity map referred to the reference emissivity 0.5 scales according to $\delta_M = \lambda_{0.5 \rightarrow M} \cdot \delta_{0.5}$ when the reference emissivity is changed from 0.5 to M . Hence, a relative emissivity error $\varepsilon_{0.5}$ (e.g. $\delta_{0.5}$ divided by the reference emissivity) scales according to $\varepsilon_M = \lambda_{0.5 \rightarrow M} \cdot 0.5 / M \cdot \varepsilon_{0.5}$. In the example for the reference emissivity 0.750, the relative errors are multiplied by 0.96 at 1.02 μm , by 0.77 at 1.10 μm , and by 0.71 at 1.18 μm . This does not mean that the errors decrease, but rather that the reference value increases. The color relations in Fig. 3 with respect to (absolute, i.e. not relative) emissivity retrieval errors are not affected by a change of the reference emissivity, i.e. an emissivity anomaly exceeding the error levels stays an anomaly regardless of the reference emissivity transformation (this will later be illustrated by Fig. 6). In this respect, the renormalized emissivity maps and absolute errors are well-defined with respect to the reference emissivity.

5. Discussion

The variation of the renormalization coefficients for the different scenarios is not unexpected (Section 2.3) and necessitates the renormalization procedure. According to the MSTs and PMTs, the renormalized maps statistically agree within tolerances listed in Table 4, taking into account the different N_r . The spatially slowly (or not) varying time independent interfering parameters studied by Kappel et al. (2015) (continuum parameters, deep atmospheric temperature field modifications) lead to very large unrenormalized emissivity retrieval errors (up to an order of 100%) that do not statistically scale with $1/\sqrt{N_r}$, because they do not average out over time. However, it was argued that these errors are largely removed by the renormalization. The PMTs of the present work support this conclusion. The same applies with respect to unconsidered systematic latitude trends of the cloud or minor gas parameters (Appendix A.2). Furthermore, renormalized emissivity maps were shown to be reasonably independent of the choice of the initial emissivity, as long as the latter is not close to 0 or 1. Slight modification to the retrieval pipeline and calibration and preprocessing are not critical as well.

The (unscaled with respect to N_r) retrieval errors for the MST with $N_r = 32$ (2.2%, 6.9%, 3.6%) are clearly smaller than for $N_r = 16$ results (4.9%, 15.6%, 5.6%), and those are clearly smaller than for $N_r = 8$ (6.8%, 21.6%, 9.1%), which can be found by backscaling the MST errors from Table 5 from N_r^0 to the respective N_r . Moreover, the N_r^0 -referred MST errors in Table 5 do not differ much for different N_r and even have the tendency to decrease with increasing N_r . The error scaling rules thus seem to apply, and the estimation of the N_r^0 -errors using these rules seems

to be valid. However, it should be expected that even the N_r^0 -referred MST errors decrease with increasing N_r due to MSR's usage of more context data from measurements adjacent in time or space. This is the case for $N_r = 32$, but $N_r = 16$ -errors are not consistently smaller than those for $N_r = 8$. Also note the high errors Δ for the MST estimates of 4%–10%, so no conclusion regarding decreasing N_r^0 -errors with increasing N_r is possible yet based on present results. This must be studied with further surface targets and, if possible, with higher N_r^0 by relaxing the demands on 'usable measurements' (Section 2.1).

Next, the compatibility of present results with the retrieval errors estimated by Kappel et al. (2015) from synthetic spectra ('synthetic errors') is examined, see comparison in Table 5. Since the MST estimates are not strictly monotonically decreasing with increasing N_r , not the result from the MST with $N_r = 32$ that is closest to N_r^0 is used as estimate for N_r^0 , but the maximum value from the three MSTs. MST errors are interpreted to be mainly due to actual variations of those interfering parameters that can vary with time. This is just what the synthetic errors describe, and thus, these two estimates are compared. The synthetic errors well agree with the MST estimate at 1.02 μm but are underestimated at 1.10 μm and overestimated at 1.18 μm . The synthetic errors are based on assumptions for the effective range (double standard deviation) the interfering parameters are thought to statistically cover, while the MSTs do not rely on such assumptions. These parameter ranges are obvious candidates for the disagreements, therefore. Table 3 shows that H_2O might be responsible for the synthetic overestimate at 1.18 μm . Indeed, the utilized spectral range is mainly sensitive to tropospheric properties, which are thought to be quite stable with time and spatially slowly varying. The assumed 25%-range for the H_2O abundance is therefore probably too high. Moreover, the H_2O^b abundance is retrieved here in contrast to Kappel et al. (2015), admittedly probably with low accuracy, but this could also help to reduce the impact of H_2O variations. The double standard deviation range is now assumed as 5% instead of 25%. The response of the retrieved emissivity to H_2O is quite linear in the discussed range (Kappel et al., 2015, Fig. 5), and thus the error contribution approximately scales with the assumed range. In the same way, the contribution from 25%-uncertainties of emissivities from the other respective surface windows should be reduced to 9%-uncertainties, which is more realistic in view of the 'Total'-row in Table 5. The effective variability of the cloud bottom altitude is likely smaller than the assumed 3 km, since this altitude is related to the probably spatially slowly varying (thereby annihilated by renormalization) and temporally almost not varying tropospheric temperature and pressure conditions, see discussion by Kappel et al. (2015, Section 3.5) based on data by Knollenberg and Hunten (1980) and Marov et al. (1980). It is now set to 1 km. Although the response to the retrieved emissivity is not linear (Kappel et al., 2015, Fig. 5), a linear downscaling leads to a valid upper boundary for the error contribution in this case. Next, even though it is a PMT result, errors in the straylight removal have to be taken into account by quadratically adding the 'Unsmoothed straylight' SRMSDs, to make the synthetic errors better comparable to what the MST errors represent. Finally, the errors listed in Table 3 were based on a reference spectrum with surface emissivity 0.65 and have to be referred to the reference emissivity 0.5 to be comparable to the error measures in the present work (see Section 4.5 for a discussion on the reference emissivity transformation of relative errors). In conclusion, this yields 1.9%, 5.8%, and 6.2% as revised synthetic estimate for the three surface windows, see Table 5. Even though the information content of the 1.10 μm peak is comparable to that of the 1.18 μm peak (Haus and Arnold, 2010, Fig. 5), the 1.10 μm MST errors are probably so large, because this radiance peak is much smaller than the 1.18 μm peak (Haus and Arnold, 2010, Fig. 21) and real measurement errors and simulation errors thus have a much larger impact there. Also note that errors in the straylight removal have the largest impact there. The remaining disagreements might be explained by impacts of calibration, preprocessing, measurement, and simulation errors, unconsidered interfering parameters, subsidiary retrieval solutions, or higher order terms and side effects from other assumptions in Appendix A.2.2 that were used to justify the error adding and scaling rules that led to Table 3. Moreover, this table shows errors based on just a single reference spectrum. Kappel et al. (2015, Table 5) have shown that the errors for other reference spectra can be quite different. But in general, it seems plausible that the MST errors mainly originate from interfering parameters that vary with time. MSTs do not suggest time variations of the true surface emissivities of the target, but note that for some reason, the MST errors are slightly higher for the chronological second half of the utilized VIRTIS-M-IR data than for the first half (Table 4). Maybe atmospheric motions were increased or the straylight situation was different.

The choice of smaller FWHMs for the topography blurring does not lead to as large impacts as could be expected. The renormalized emissivity errors determined by Kappel et al. (2015) due to topography uncertainties might be overestimated, therefore. Spatially small-scale topography errors average out due to the blurring. The 65 km- and the 30 km- differ from the 100 km-blurred topography of the bins by up to about 125 m and 250 m, respectively, with double root-mean-square-deviations of 70 m and 126 m. As mentioned, these differences can serve to study the effect of

slight errors in the underlying Magellan topography just in the areas where the topography is spatially fastest varying and errors are most likely. They also illustrate that the assumed uncertainty of the surface elevation of 200 m is not completely wrong. However, the statistical impact is rather connected to the double root-mean-square-deviations between the topography maps. The synthetic emissivity errors due to topography uncertainties should therefore be scaled from 9%, 12%, and 12% (Table 3) for the three surface windows down to 6%, 8%, and 8%, based on the double root-mean-square-deviation between the 100 km- and the 30 km-blurred topography. But this can only partly explain the small SRMSDs in Table 4. The main reason is probably the renormalization, which itself depends on topography and naturally attenuates its influences, even for the PMTs 'Topo. 65 km R100' and 'Topo. 30 km R100'. Therefore, the synthetic error estimates in Table 5 exclude topography impacts, which are difficult to determine in view of the renormalization, and which are not present in the MST estimates anyway. However, they are included in the PMT error contributions as mentioned.

Apart from the PMT 'Init. emiss. 0.875' (unrenormalized emissivities are clearly cut off at the upper emissivity domain boundary 1 as noted), the largest SRMSDs occur for the PMTs, 'Continuum 1.74 μm ', 'H₂O', 'Surface haze', 'Unsmoothed straylight', and for topography uncertainties. Special care must be given to these parameters, in order to improve retrieval results.

Fig. 6 shows for the three surface windows the renormalized retrieved emissivities from the $N_r^0 = 64$ repetition case (see Fig. 3) for the bin with center coordinates at 280.26 °E, 42.44 °S (located at Shiwanokia Corona). The error bars

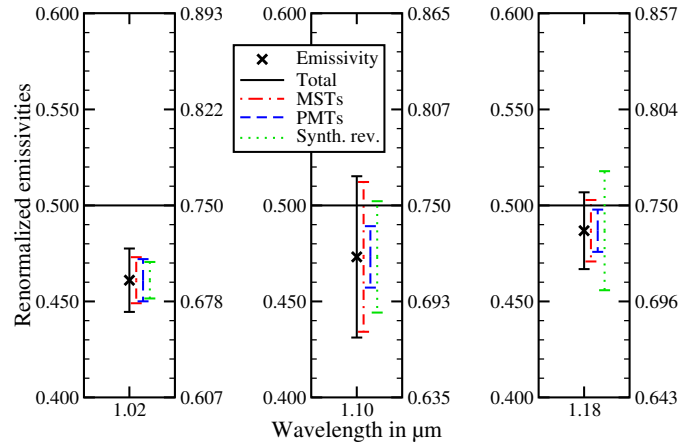


Figure 6: Renormalized emissivities from Fig. 3 at 280.26 °E, 42.44 °S for the 1.02, 1.10, and 1.18 μm windows (0.461, 0.473, and 0.487 referred to reference emissivity 0.500). The error bars (wavelength offsets for better representation) illustrate the double standard deviation errors listed in Table 5. Respective left axes: reference emissivity 0.500, right: 0.750.

represent the double standard deviations summarized in Table 5. Two cases for the reference emissivity are illustrated, 0.5 and 0.75. 'MSTs' corresponds to the maximum value of the three MST rows in this table. 'Total' is now the valid estimate of the multi-spectrum retrieval errors at $N_r^0 = 64$ measurement repetitions and is the quadratic sum of 'MSTs' and 'PMTs' in the figure. For the represented surface bin, the deviation of the renormalized emissivity at 1.02 μm from the reference emissivity is statistically significant (double standard deviation) according to both 'Total' and synthetic error measures. At 1.10 and 1.18 μm , neither of both errors indicates a significant deviation. From now on, only the 'Total' errors will be used. Recall that the shown emissivities are not absolute emissivities. Rather, they represent deviations from the reference emissivity and are only valid as absolute emissivities when the reference emissivity coincides with the true emissivity mean of the target area. A true emissivity mean different from the reference emissivity requires a transformation as discussed. This figure illustrates that these deviations can be statistically double-standard-deviation-significant or not. However, true emissivities may exhibit real trends with topography or latitude, which are not reflected by renormalized retrieved emissivities. According to the 'Total' errors, only bins where the renormalized emissivity deviations from the spatial mean value 0.5 exceed 0.017, 0.042, and 0.020, respectively, for the three surface windows, have significant emissivity anomalies. Referred to the 1.02 μm -map in Fig. 3, this is the case for areas with deviations from green that are at least orange or light blue. For 1.10 μm , only white and black are significant, and for 1.18 μm red and blue. These color criteria still apply, when the reference emissivity is changed and

the color bar is accordingly transformed as discussed in Section 4.5 (compare also the two reference emissivity cases illustrated in Fig. 6). When the demand on the statistical significance is relaxed to one standard deviation, the criteria change correspondingly.

Based on the MST and PMT results, it is possible to roughly extrapolate the expectable renormalized emissivity errors referred to the reference emissivity 0.5 for larger N_r . About 400 repetitions would be needed for MST errors at 1.02 μm to fall below 1%, in which case the errors for the other windows would amount to 3.1% and 1.3%, respectively. At $N_r = 1000$, MST errors would read 0.6%, 2.0%, and 0.8%, and it would take about 4000 repetitions for the errors in all surface windows to fall below 1% (0.3%, 0.99%, 0.4%). PMT errors, on the other hand, hardly decrease with increasing N_r . Even at 10000 repetitions, they would still amount to 2.1%, 2.6%, and 2.2%. This is because many PMT errors do not scale with $1/\sqrt{N_r}$, since the parameter modifications are constant with time. However, the extrapolation from $N_r = 25$ to such large repetition numbers is certainly not valid. In addition, the resources needed for MSR retrieval would be enormous, and VIRTIS-M-IR did not acquire so many usable measurement repetitions for any target on Venus.

On the other hand, spatial variations of surface emissivity of 20% that would correspond to the difference between unweathered granitic and basaltic rocks (Hashimoto and Sugita, 2003) would be easily detectable already at $N_r = 8$ in case of the 1.02 μm window. The present work suggests that such large variations do not occur in the studied target area, but note that contrasts for weathered material might be smaller, see discussions by Haus and Arnold (2010) and Arnold et al. (2012).

Present results show relevant variations that indicate the heterogeneity of Venus' surface emissivity. High emissivity signatures correlate with edifices like Abeona Mons and other volcanic structures. Shivanokia Corona has areas of lower emissivity. The complexity in this region might suggest long histories and an apparently older corona (Stofan and Smrekar, 1998). In general, these anomalies can be caused by local changes in composition, age, and/or textural differences of the surface (Arnold et al., 2015). Compared to the thermal flux data at 1.02 μm obtained by Müller et al. (2008), the new relative emissivity maps at 1.02, 1.10, and 1.18 μm enable a more direct and more quantitative interpretation. Moreover, the removal of topographic, atmospheric, and other interferences is improved. Thus, the new maps will allow for a more precise disentanglement of the different sources causing the observed surface emissivity variations. A detailed geologic discussion of these results is under preparation for a follow-up publication.

6. Summary and conclusion

Maps of surface emissivity data at 1.02, 1.10, and 1.18 μm are determined from VIRTIS-M-IR/VEX measurements of Venus' nightside emissions. The surface target at Themis Regio is located between 35 and 47 °S and 270 and 288 °E. The usable measurements most suitable for emissivity retrieval are selected (high signal-to-noise ratio, thin to moderate cloud cover, few straylight, etc.). The radiance spectra are carefully calibrated and preprocessed (Kappel et al., 2012). A full radiative transfer forward model with special consideration of the extreme environmental conditions in Venus' deep atmosphere is used for the simulation of radiance spectra, taking into account thermal emissions by surface and atmosphere, and absorption and multiple scattering by gases and clouds (Haus and Arnold, 2010). Since spectral information content is limited, there are always parameters that cannot be uniquely derived from the measured spectra. They are set to reasonable values that allow for the generation of realistic synthetic spectra. When these assumptions deviate from the true physical values, this may cause surface emissivity retrieval errors, even though the fits well match the measured spectra. To improve this, a multi-spectrum retrieval algorithm (MSR, Kappel (2014)) is applied to the measured spectra. MSR can retrieve parameters that are common to a set of spectra and takes into account expected mean values, standard deviations, and correlation lengths and times as Bayesian *a priori* information for the retrieval parameters, along with measurement and simulation error information. This decreases the probability of retrieving unreasonable values and spatial-temporal distributions of the parameters, allows for taking the context of adjacent measurements into account, and thus enhances the reliability of the retrieved results. All considered measurements can be parameterized by a fully consistent set of atmospheric, surface, and instrumental parameters. Assuming geologic activity as would be observable by VIRTIS-M-IR to be negligible, the emissivity map of a surface target can be retrieved at each surface window as a parameter vector that is common to spectrally resolved VIRTIS-M-IR images that cover this target N_r times.

In spite of MSR's advantages, absolute emissivity values are difficult to obtain, and an emissivity map has to be retrieved by first setting a certain fixed initial emissivity value and retrieving the map relative to this value. However,

the mean value of the retrieved emissivity map does not necessarily coincide with the initial emissivity due to technical reasons. Moreover, the retrieved map can exhibit approximately linear trends with latitude and topography that are likely due to imperfect forward model properties. When these trends are removed for a sufficiently small target by multiple linear regression, it can be observed that the magnitude of the resulting spatial emissivity fluctuations around their mean value increases with increasing mean value. This is because the radiance response to small emissivity perturbations around an emissivity base value decreases for increasing emissivity base value. The fluctuations are small compared to the emissivity mean. To still obtain a well-defined quantitative emissivity map, a linear transformation is applied that converts the map to have an emissivity mean value that attains a fixed reference emissivity, here 0.5. The accordingly changed fluctuation amplitudes are then directly comparable to those of other maps that are referred to the same reference emissivity. The process of retrieving the map relative to an initial emissivity, removing the trends with latitude and topography, and transforming the map to the reference emissivity is called renormalized retrieval, and the result is the renormalized emissivity map. The results can be later transformed to other reference emissivities if needed. This is particularly useful, since the true emissivity mean value of the target is not known and cannot be determined from the VIRTIS-M-IR spectra with current techniques, and geologic interpretations may have to take the magnitude of the fluctuations into account. If the absolute emissivity at a given bin of the target were known, for instance in the form of ground truth data, the absolute emissivity map of the entire target could be computed according to the mentioned transformation, even if the emissivity of the given bin were not the mean emissivity of the target.

An appendix explains the mathematical background of error origination and renormalization and justifies approximate error adding and scaling rules. Kappel et al. (2015) have estimated renormalized emissivity errors based on synthetic spectra. Independently from that, the present work estimates errors based on a statistical evaluation of retrieval results derived from real measurements and compares results from both analyses.

The maps are retrieved and renormalized from a data set with $N_r^0 = 64$ measurement repetitions, which is close to the maximum of the available usable data. Measurement Selection Tests (MSTs) retrieve renormalized emissivity maps from different disjoint selections of spectra with $N_r \in \{8, 16, 32\}$. Double standard deviation error estimates for N_r^0 are derived from these results and amount to 2.4%, 7.8%, and 3.2% for the three surface windows, respectively. This is compared to the synthetic error analysis, but a new evaluation is performed that better matches the conditions of the present work and predicts synthetic errors of 1.9%, 5.8%, and 6.2%. This is compatible with the hypothesis that MST errors are mainly caused by unconsidered time variations of interfering atmospheric parameters. Disagreements might be explained by impacts of calibration, preprocessing, measurement, and simulation errors, unconsidered interfering parameters, subsidiary retrieval solutions, higher order terms and side effects from other assumptions that were used to justify the error adding and scaling rules, and the dependence of the synthetic errors on their reference spectrum. MSTs do not suggest time variations of the true surface emissivities of the target.

Parameter Modification Tests (PMTs) at $N_r = 25$ verify that renormalized emissivity maps are reasonably independent of the choice of the initial emissivity, of modifications to interfering atmospheric, surface, and instrumental parameters, and of some details of the retrieval pipeline and data calibration and preprocessing. Scaled PMT results are used to further quantify the retrieval errors for the N_r^0 -maps. The PMT errors that are taken to contribute to the total renormalized emissivity errors at 64 repetitions amount to 2.2%, 3.2%, and 2.2%, respectively, for the three surface windows. The combined scaled MST and PMT results estimate total double standard deviation errors for the three surface windows of 3.3%, 8.4%, and 4.0%. The results of the MSTs and PMTs suggest that no further modifications to the present approach of the renormalized retrieval are required to obtain reasonably reliable emissivity data maps. MSTs and PMTs also verify that the error scaling rules approximately apply, and that emissivity errors due to unconsidered time dependent latitudinal trends or offsets of underlying atmospheric parameters are effectively removed by the renormalization. The synthetic errors (9%, 12%, 12%) determined by Kappel et al. (2015) due to topography uncertainties are probably overestimated, and the PMTs show that they are more likely below 2%. On the one hand, the assumed 200 m uncertainty was chosen too high. Effects from spatially small-scale topography errors largely average out due to atmospheric blurring on the scale of 100 km. On the other hand, the renormalization itself depends on topography and naturally attenuates its influences. The PMTs demonstrate that the initial emissivity must be chosen far from 0 and 1, because otherwise, unrenormalized emissivities can be cut off at these emissivity domain boundaries. Significant error sources (1%–2% for at least one surface window) are uncertainties in the 1.74 μm CO₂ continuum absorption, H₂O abundance, topography, and in the removal of straylight. Special care must be given to these parameters for optimal retrieval results.

Significant errors can also be produced by an additional absorber close to the surface (surface haze) that could explain

certain (unrenormalized) emissivity trends with surface elevation in interplay with unconsidered latitudinal trends of the deep atmospheric temperature field, the utilized continua, and latitudinal trends of time-averaged atmospheric parameters. Although actual trends with surface elevation (due to trends in the distribution of dust, surface composition, or weathering) may exist, they seem to be unrealistic. It is reasonable, therefore, to regard the trends in the retrieved emissivities, in particular if they should turn out as a global phenomenon that occurs regardless of geologic context, as artifacts due to imperfect or unconsidered forward model parameters, and to remove them by the renormalization. The planned analysis of further surface targets can help to find the correct procedure. Until then, the renormalized emissivities must be interpreted as spatial variations relative to the reference emissivity. However, possible real trends of the true emissivities with topography or latitude cannot be reflected by renormalized retrieved emissivities. Nevertheless, the latter have a well-defined quantitative meaning and represent an important step toward the retrieval of absolute emissivities. Future analyses will aim at the systematic avoidance of trends with latitude and topography and of deviations between the mean of the raw retrieved emissivity map and the initial emissivity from the outset by a suitable regularization of the retrieval, such that trends in the interfering atmospheric parameters and surface haze can be retrieved as parameters common to many diverse spectra with a wide latitude and topography coverage. Then it may be feasible to estimate absolute emissivity maps by looking for the initial emissivity that leads to the smallest residuals between measured and simulated spectra, which is currently not possible due to the trends with latitude and topography in the raw retrieved maps.

Fig. 3 depicts the renormalized emissivity maps determined from 64 measurement repetitions. At 1.02 μm , areas with deviations from green that are at least orange or light blue differ statistically significantly (double standard deviation) from the mean of the target area. For 1.10 μm , only white and black are significant, and for 1.18 μm red and blue. Referred to the reference emissivity 0.5, emissivity deviations from the mean of the target of up to 8.1%, 8.4%, and 7.6% are detected at 1.02, 1.10, and 1.18 μm , respectively. When statistically significant emissivity anomalies are defined as deviations exceeding the determined double standard deviation error levels at present $N_r = 64$ -sensitivity (3.3%, 8.4%, 4.0%), anomalies were found at both 1.02 and 1.18 μm . Anomalies $\geq 9\%$ can be ruled out at 1.10 μm . With single standard deviation significance, all three maps show interesting spatial emissivity variations. The retrieval errors of the determined relative emissivity maps are sufficiently small to allow geologic interpretation. A detailed geologic discussion of the results is under preparation for a follow-up publication.

Spatial variations of surface emissivity of 20% that would correspond to the difference between unweathered granitic and basaltic rocks would be easily detectable already at $N_r = 8$ in case of the 1.02 μm window. The present work suggests that such large variations do not occur in the studied target area, but note that contrasts for weathered material might be smaller. According to present results, about 400 repetitions would be needed for MST errors at 1.02 μm to fall below 1%. For all surface windows this would happen only at 4000 repetitions. The combined PMT errors only slightly improve with increasing N_r .

Acknowledgements

This work is funded by the German Aerospace Center e.V. (DLR e.V.). We acknowledge the work of the VIRTIS/VEX PIs Pierre Drossart (CNRS/LESIA, Paris) and Giuseppe Piccioni (INAF-IAPS, Rome), the VIRTIS/VEX team, and also the entire Venus Express team of ESA and Astrium, who made the measurement data available that were used in this study. The authors would like to thank the reviewers for their valuable suggestions to improve this paper.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at ...

References

- Arnold G., Kappel D., Haus R., Ivanov M., and Pedroza L. T. Retrieval and study of near-infrared surface emissivity maps of Themis Regio on Venus with VIRTIS-M (Venus Express). In *EPSC Abstracts Vol. 10, European Planetary Science Congress 2015*, 2015. EPSC2015-160.
- Arnold G., Haus R., Kappel D., Drossart P., and Piccioni G. Venus surface data extraction from VIRTIS/Venus Express measurements: Estimation of a quantitative approach. *J. Geophys. Res.*, 113(E5):E00B10, 2008. ISSN 0148-0227. doi:10.1029/2008JE003087.

- Arnold G. E., Haus R., Kappel D., Piccioni G., and Drossart P. VIRTIS/VEX observations of Venus: overview of selected scientific results. *J. Appl. Remote Sens.*, 6(1):063580–1–063580–20, 2012. doi:10.1117/1.JRS.6.063580.
- Baldrige A. M., Hook S. J., Grove C. I., and Rivera G. The ASTER spectral library version 2.0. *Remote Sens. Environ.*, 113(4):711–715, 2009. ISSN 0034-4257. doi:10.1016/j.rse.2008.11.007.
- Bézar B., Fedorova A., Bertaux J.-L., Rodin A., and Korabev O. The 1.10- and 1.18- μm nightside windows of Venus observed by SPICAV-IR aboard Venus Express. *Icarus*, 216(1):173–183, 2011. ISSN 0019-1035. doi:10.1016/j.icarus.2011.08.025.
- Bugayevskiy L. M. and Snyder J. *Map Projections: A Reference Manual*. Taylor & Francis, 1995. ISBN 9780748403042.
- Cardesin-Moinelo A., Piccioni G., Ammannito E., Filacchione G., and Drossart P. Calibration of hyperspectral imaging data: VIRTIS-M onboard Venus Express. *IEEE Trans. Geosci. Remote*, 48:3941–3950, 2010. doi:10.1109/TGRS.2010.2064325.
- Carlson R. and Anderson M. Absorption properties of sulfuric acid in Venus's infrared spectral windows region. In *EPSC-DPS Joint Meeting 2011*, page 1171, 2011.
- Drossart P., Piccioni G., Adriani A., Angrilli F., Arnold G., Baines K. H., Bellucci G., Benkhoff J., Bézar B., Bibring J.-P., Blanco A., Blecka M. I., Carlson R. W., Coradini A., Lellis A. D., Encrenaz T., Erard S., Fonti S., Formisano V., Fouchet T., Garcia R., Haus R., Helbert J., Ignatiev N. I., Irwin P. G. J., Langevin Y., Lebonnois S., Lopez-Valverde M. A., Luz D., Marinangeli L., Orofino V., Rodin A. V., Roos-Serote M. C., Saggin B., Sanchez-Lavega A., Stam D. M., Taylor F. W., Titov D., Visconti G., Zambelli M., Hueso R., Tsang C. C. C., Wilson C. F., and Afanasenko T. Z. Scientific goals for the observation of Venus by VIRTIS on ESA/Venus Express mission. *Planet. Space Sci.*, 55(12):1653–1672, 2007. ISSN 0032-0633. doi:10.1016/j.pss.2007.01.003.
- Drossart P., Piccioni G., and ESA. FTP access to data by Visible and InfraRed Thermal Imaging Spectrometer (VIRTIS) / Venus Express at the Planetary Science Archive (PSA), 2013. <ftp://psa.esac.esa.int/pub/mirror/VENUS-EXPRESS/VIRTIS/>.
- Erard S., Drossart P., and Piccioni G. Multivariate analysis of Visible and Infrared Thermal Imaging Spectrometer (VIRTIS) Venus Express nightside and limb observations. *J. Geophys. Res.*, 114(E9):1–20, 2009. ISSN 2156-2202. doi:10.1029/2008JE003116. E00B27.
- Fegley B., Klingelhöfer G., Lodders K., and Widemann T. Geochemistry of surface–atmosphere interactions on Venus. In Bougher S. W., Hunten D. M., and Phillips R. J., editors, *Venus II: Geology, Geophysics, Atmosphere, and Solar Wind Environment*, Space science series, pages 591–636. University of Arizona Press, Tucson, 1997. ISBN 9780816518302.
- Ford P. G. and Pettengill G. H. Venus topography and kilometer-scale slopes. *J. Geophys. Res.: Planets*, 97(E8):13103–13114, 1992. ISSN 2156-2202. doi:10.1029/92JE01085.
- Grieger B., Ignatiev N. I., Hoekzema N. M., and Keller H. U. Indication of a near surface cloud layer on Venus from reanalysis of Venera 13/14 spectrophotometer data. In Wilson A., editor, *Planetary Probe Atmospheric Entry and Descent Trajectory Analysis and Science*, volume 544 of *ESA Special Publication*, pages 63–70, 2004. ISBN 92-9092-855-7.
- Hansen J. E. and Travis L. D. Light scattering in planetary atmospheres. *Space Sci. Rev.*, 16(4):527–610, 1974. ISSN 0038-6308. doi:10.1007/BF00168069.
- Hashimoto G. L. and Sugita S. On observing the compositional variability of the surface of Venus using nightside near-infrared thermal radiation. *J. Geophys. Res.: Planets*, 108(E9):5109, 2003. ISSN 2156-2202. doi:10.1029/2003JE002082.
- Haus R. and Arnold G. Radiative transfer in the atmosphere of Venus and application to surface emissivity retrieval from VIRTIS/VEX measurements. *Planet. Space Sci.*, 58(12):1578–1598, 2010. ISSN 0032-0633. doi:10.1016/j.pss.2010.08.001.
- Haus R., Kappel D., and Arnold G. Self-consistent retrieval of temperature profiles and cloud structure in the northern hemisphere of Venus using VIRTIS/VEX and PMV/VENERA-15 radiation measurements. *Planet. Space Sci.*, 89:77–101, 2013. ISSN 0032-0633. doi:10.1016/j.pss.2013.09.020.
- Haus R., Kappel D., and Arnold G. Atmospheric thermal structure and cloud features in the southern hemisphere of Venus as retrieved from VIRTIS/VEX radiation measurements. *Icarus*, 232:232–248, 2014. ISSN 0019-1035. doi:10.1016/j.icarus.2014.01.020.
- Haus R., Kappel D., and Arnold G. Lower atmosphere minor gas abundances as retrieved from Venus Express VIRTIS-M-IR data at 2.3 μm . *Planet. Space Sci.*, 105:159–174, 2015. ISSN 0032-0633. doi:10.1016/j.pss.2014.11.020.
- Hayward R. Gazetteer of planetary nomenclature, International Astronomical Union (IAU) working group for planetary system nomenclature (WGPSN), 2015. <http://planetarynames.wr.usgs.gov/Page/VENUS/target>.
- Helbert J., Müller N., Kostama P., Marinangeli L., Piccioni G., and Drossart P. Surface brightness variations seen by VIRTIS on Venus Express and implications for the evolution of the Lada Terra region, Venus. *Geophys. Res. Lett.*, 35(11):1–5, 2008. ISSN 1944-8007. doi:10.1029/2008GL033609. L11201.
- Imamura T. and Hashimoto G. L. Microphysics of Venusian clouds in rising tropical air. *J. Atmos. Sci.*, 58(23):3597–3612, 2001. ISSN 0022-4928. doi:10.1175/1520-0469(2001)058<3597:MOVCIR>2.0.CO;2.
- James E. P., Toon O. B., and Schubert G. A numerical microphysical model of the condensational Venus cloud. *Icarus*, 129(1):147–171, 1997. ISSN 0019-1035. doi:10.1006/icar.1997.5763.
- Kappel D. MSR, a multi-spectrum retrieval technique for spatially-temporally correlated or common Venus surface and atmosphere parameters. *J. Quant. Spectrosc. Rad.*, 133:153–176, 2014. ISSN 0022-4073. doi:10.1016/j.jqsrt.2013.07.025.
- Kappel D., Arnold G., Haus R., Piccioni G., and Drossart P. Refinements in the data analysis of VIRTIS-M-IR Venus nightside spectra. *Adv. Space Res.*, 50(2):228–255, 2012. ISSN 0273-1177. doi:10.1016/j.asr.2012.03.029.
- Kappel D., Haus R., and Arnold G. Error analysis for retrieval of Venus' IR surface emissivity from VIRTIS/VEX measurements. *Planet. Space Sci.*, 113–114:49–65, 2015. ISSN 0032-0633. doi:10.1016/j.pss.2015.01.014. SI:Exploration of Venus.
- Knollenberg R. G. and Hunten D. M. The microphysics of the clouds of Venus: Results of the Pioneer Venus particle size spectrometer experiment. *J. Geophys. Res.: Space Physics*, 85(A13):8039–8058, 1980. ISSN 2156-2202. doi:10.1029/JA085iA13p08039.
- Lebonnois S., Hourdin F., Eymet V., Crespin A., Fournier R., and Forget F. Superrotation of Venus' atmosphere analyzed with a full general circulation model. *J. Geophys. Res.*, 115(E6):1–23, 2010. ISSN 2156-2202. doi:10.1029/2009JE003458. E06006.
- Marov M. Y., Lystsev V. E., Lebedev V. N., Lukashevich N. L., and Shari V. P. The structure and microphysical properties of the Venus clouds: Venera 9, 10, and 11 data. *Icarus*, 44(3):608–639, 1980. ISSN 0019-1035. doi:10.1016/0019-1035(80)90131-1.
- Moroz V. I. Estimates of visibility of the surface of Venus from descent probes and balloons. *Planet. Space Sci.*, 50(3):287–297, 2002. ISSN 0032-0633. doi:10.1016/S0032-0633(01)00128-3.
- Müller N., Helbert J., Hashimoto G. L., Tsang C. C. C., Erard S., Piccioni G., and Drossart P. Venus surface thermal emission at 1 μm in VIRTIS

- imaging observations: Evidence for variation of crust and mantle differentiation conditions. *J. Geophys. Res.*, 113(E5):E00B17, 2008. ISSN 0148-0227. doi:10.1029/2008JE003118.
- Müller N., Helbert J., Stofan E. R., Smrekar S. E., Piccioni G., and Drossart P. Search for active lava flows with VIRTIS on Venus Express. In *American Geophysical Union, 45th Fall Meeting*, pages P24B–02, 2012a.
- Müller N. T., Helbert J., Erard S., Piccioni G., and Drossart P. Rotation period of Venus estimated from Venus Express VIRTIS images and Magellan altimetry. *Icarus*, 217(2):474–483, 2012b. ISSN 0019-1035. doi:10.1016/j.icarus.2011.09.026. In: Advances in Venus Science.
- Nocedal J. and Wright S. J. *Numerical Optimization*. Springer Series in Operations Research. Springer, 1999. ISBN 9780387987934.
- Palmer K. F. and Williams D. Optical constants of sulfuric acid; application to the clouds of Venus? *Appl. Opt.*, 14(1):208–219, 1975. doi:10.1364/AO.14.000208.
- Pettengill G. H., Ford P. G., Johnson W. T. K., Raney R. K., and Soderblom L. A. Magellan: Radar performance and data products. *Science*, 252(5003):260–265, 1991. doi:10.1126/science.252.5003.260.
- Piccioni G., Drossart P., Sanchez-Lavega A., Hueso R., Taylor F. W., Wilson C. F., Grassi D., Zasova L., Moriconi M., Adriani A., Lebonnois S., Coradini A., Bézard B., Angrilli F., Arnold G., Baines K. H., Bellucci G., Benkhoff J., Bibring J. P., Blanco A., Blecka M. I., Carlson R. W., Di Lellis A., Encrenaz T., Erard S., Fonti S., Formisano V., Fouchet T., Garcia R., Haus R., Helbert J., Ignatiev N. I., Irwin P. G. J., Langevin Y., Lopez-Valverde M. A., Luz D., Marinangeli L., Orofino V., Rodin A. V., Roos-Serote M. C., Saggin B., Stam D. M., Titov D., Visconti G., and Zambelli M. South-polar features on Venus similar to those near the north pole. *Nature*, 450(7170):637–640, 2007. ISSN 0028-0836. doi:10.1038/nature06209.
- Piccioni G., Zasova L., Migliorini A., Drossart P., Shakun A., Garcia Munoz A., Mills F. P., and Cardesin-Moinelo A. Near-IR oxygen nightglow observed by VIRTIS in the Venus upper atmosphere. *J. Geophys. Res.: Planets*, 114(E5), 2009. ISSN 2156-2202. doi:10.1029/2008JE003133.
- Pollack J. B., Dalton J. B., Grinspoon D., Watson R. B., Freedman R., Crisp D., Allen D. A., Bézard B., DeBergh C., Giver L. P., Ma Q., and Tipping R. Near-infrared light from Venus' nightside: A spectroscopic analysis. *Icarus*, 103(1):1–42, 1993. ISSN 0019-1035. doi:10.1006/icar.1993.1055.
- Rappaport N. J., Konopliv A. S., Kucinskis A. B., and Ford P. G. An improved 360 degree and order model of Venus topography. *Icarus*, 139(1):19–31, 1999. ISSN 0019-1035. doi:10.1006/icar.1999.6081.
- Rothman L. S., Gordon I. E., Barbe A., Benner D. C., Bernath P. F., Birk M., Boudon V., Brown L. R., Campargue A., Champion J.-P., Chance K., Coudert L. H., Dana V., Devi V. M., Fally S., Flaud J.-M., Gamache R. R., Goldman A., Jacquemart D., Kleiner I., Lacombe N., Lafferty W. J., Mandin J.-Y., Massie S. T., Mikhailenko S. N., Müller C. E., Moazzen-Ahmadi N., Naumenko O. V., Nikitin A. V., Orphal J., Perevalov V. I., Perrin A., Predoi-Cross A., Rinsland C. P., Rotger M., Simeckova M., Smith M. A. H., Sung K., Tashkun S. A., Tennyson J., Toth R. A., Vandaele A. C., and Auwera J. V. The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Rad.*, 110(9–10):533–572, 2009. ISSN 0022-4073. doi:10.1016/j.jqsrt.2009.02.013.
- Rothman L. S., Gordon I. E., Barber R. J., Dothe H., Gamache R. R., Goldman A., Perevalov V. I., Tashkun S. A., and Tennyson J. HITRAN, the high-temperature molecular spectroscopic database. *J. Quant. Spectrosc. Rad.*, 111(15):2139–2150, 2010. ISSN 0022-4073. doi:10.1016/j.jqsrt.2010.05.001.
- Rothman L. S., Watson R. B., Gamache R., Schroeder J. W., and McCann A. HITRAN HAWKS and HITRAN: high-temperature molecular database. In *Proc. SPIE 2471, Atmospheric Propagation and Remote Sensing IV, 105 (June 15, 1995)*, pages 105–111, 1995. doi:10.1117/12.211919.
- Schubert G. General circulation and the dynamical state of the Venus atmosphere. In Hunten D. M., Colin L., Donahue T. M., and Moroz V. I., editors, *Venus*, Space science series, pages 681–765. University of Arizona Press, Tucson, 1983. ISBN 9780816507887.
- Seiff A., Schofield J. T., Kliore A. J., Taylor F. W., Limaye S. S., Revercomb H. E., Sromovsky L. A., Kerzhanovich V. V., Moroz V. I., and Marov M. Y. Models of the structure of the atmosphere of Venus from the surface to 100 kilometers altitude. *Adv. Space Res.*, 5(11):3–58, 1985. ISSN 0273-1177. doi:10.1016/0273-1177(85)90197-8.
- Smrekar S. E. and Stofan E. R. Origin of corona-dominated topographic rises on Venus. *Icarus*, 139(1):100–115, 1999. ISSN 0019-1035. doi:10.1006/icar.1999.6090.
- Smrekar S. E., Stofan E. R., Müller N., Treiman A., Elkins-Tanton L., Helbert J., Piccioni G., and Drossart P. Recent hotspot volcanism on Venus from VIRTIS emissivity data. *Science*, 328(5978):605–608, 2010. doi:10.1126/science.1186785.
- Spurr R. *Linearized radiative transfer theory: a general discrete ordinate approach to the calculation of radiances and analytic weighting functions, with application to atmospheric remote sensing*. PhD thesis, Technische Universiteit Eindhoven, 2001. doi:10.6100/IR545442.
- Spurr R. LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In Kokhanovsky A. A., editor, *Light Scattering Reviews 3*, Springer Praxis Books, pages 229–275. Springer Berlin Heidelberg, 2008. ISBN 978-3-540-48305-2. doi:10.1007/978-3-540-48546-9_7.
- Stammes K., Tsay S.-C., Wiscombe W., and Jayaweera K. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.*, 27(12):2502–2509, 1988. doi:10.1364/AO.27.002502.
- Stofan E. R. and Brian A. W. *Geologic map of the Themis Regio quadrangle (V-53), Venus*. US Geological Survey Scientific Investigations Map 3165, 2012. <http://pubs.usgs.gov/sim/3165/>.
- Stofan E. R. and Smrekar S. E. Coronae at large topographic rises on Venus: Implications for plume evolution and delamination. In *Lunar and Planetary Science Conference LPSC XXIX*, volume 29, page 1502, 1998.
- Stofan E. R., Smrekar S. E., Helbert J., and Müller N. Themis Regio, Venus: Emissivity signatures in VIRTIS-Venus Express data. In *2010 VEXAG International Workshop, Madison, Wisconsin, USA*, 2010. http://www.ssec.wisc.edu/venus-workshop-submission/files/stofan_ellen-22.pdf.
- Stofan E. R., Smrekar S. E., Bindschadler D. L., and Senske D. A. Large topographic rises on Venus: Implications for mantle upwelling. *J. Geophys. Res.: Planets*, 100(E11):23317–23327, 1995. ISSN 2156-2202. doi:10.1029/95JE01834.
- Tashkun S. A., Perevalov V. I., Teffo J.-L., Bykov A. D., and Lavrentieva N. N. CDSD-1000, the high-temperature carbon dioxide spectroscopic databank. *J. Quant. Spectrosc. Rad.*, 82(1-4):165–196, 2003. ISSN 0022-4073. doi:10.1016/S0022-4073(03)00152-3.
- Tellmann S., Pätzold M., Häusler B., Bird M. K., and Tyler G. L. Structure of the Venus neutral atmosphere as observed by the radio science experiment VeRa on Venus Express. *J. Geophys. Res.*, 114(E9):E00B36, 2009. ISSN 0148-0227. doi:10.1029/2008JE003204.
- Titov D. V., Markiewicz W. J., Ignatiev N. I., Song L., Limaye S. S., Sanchez-Lavega A., Hesemann J., Almeida M., Roatsch T., Matz K.-D., Scholten F., Crisp D., Esposito L. W., Hviid S. F., Jaumann R., Keller H. U., and Moissl R. Morphology of the cloud tops as observed by the Venus

Express Monitoring Camera. *Icarus*, 217(2):682–701, 2012. ISSN 0019-1035. doi:10.1016/j.icarus.2011.06.020. Advances in Venus Science.

Tsang C. C. C., Taylor F. W., Wilson C. F., Liddell S. J., Irwin P. G. J., Piccioni G., Drossart P., and Calcutt S. B. Variability of CO concentrations in the Venus troposphere from Venus Express/VIRTIS using a band ratio technique. *Icarus*, 201(2):432–443, 2009. ISSN 0019-1035. doi:10.1016/j.icarus.2009.01.001.

Williams T., Kelley C., and many others. Gnuplot 4.6: an interactive plotting program. <http://gnuplot.sourceforge.net/>, 2014.

Wiscombe W. J. Improved Mie scattering algorithms. *Appl. Opt.*, 19(9):1505–1509, 1980. doi:10.1364/AO.19.001505.