

# Error analysis for retrieval of Venus' IR surface emissivity from VIRTIS/VEX measurements

David Kappel<sup>a,\*</sup>, Rainer Haus<sup>b</sup>, Gabriele Arnold<sup>a</sup>

<sup>a</sup>German Aerospace Center (DLR), Institute of Planetary Research, Rutherfordstrasse 2, 12489 Berlin, Germany

<sup>b</sup>Institute for Planetology, Westfälische Wilhelms-Universität Münster, Wilhelm-Klemm-Straße 10, 48149 Münster, Germany

---

## Abstract

Venus' surface emissivity data in the infrared can serve to explore the planet's geology. The only global data with high spectral, spatial, and temporal resolution and coverage at present is supplied by nightside emission measurements acquired by the Visible and InfraRed Thermal Imaging Spectrometer VIRTIS-M-IR (1.0–5.1  $\mu\text{m}$ ) aboard ESA's Venus Express. A radiative transfer simulation and a retrieval algorithm can be used to determine surface emissivity in the nightside spectral transparency windows located at 1.02, 1.10, and 1.18  $\mu\text{m}$ . To obtain satisfactory fits to measured spectra, the retrieval pipeline also determines auxiliary parameters describing cloud properties from a certain spectral range. But spectral information content is limited, and emissivity is difficult to retrieve due to strong interferences from other parameters.

Based on a selection of representative synthetic VIRTIS-M-IR spectra in the range 1.0–2.3  $\mu\text{m}$ , this paper investigates emissivity retrieval errors that can be caused by interferences of atmospheric and surface parameters, by measurement noise, and by *a priori* data, and which retrieval pipeline leads to minimal errors.

Retrieval of emissivity from a single spectrum is shown to fail due to extremely large errors, although the fits to the reference spectra are very good. Neglecting geologic activity, it is suggested to apply a multi-spectrum retrieval technique to retrieve emissivity relative to an initial value as a parameter that is common to several measured spectra that cover the same surface bin. Retrieved emissivity maps of targets with limited extension (a few thousand km) are then additively renormalized to remove spatially large scale deviations from the true emissivity map that are due to spatially slowly varying interfering parameters. Corresponding multi-spectrum retrieval errors are estimated by a statistical scaling of the single-spectrum retrieval errors and are listed for 25 measurement repetitions. For the best of the studied retrieval pipelines, temporally varying interfering atmospheric parameters (cloud parameters, minor gas abundances) contribute errors in the order of 3%–10% of the true emissivity, depending on the surface window, the reference spectrum, and assuming statistical independence of the parameters. Temporally constant interfering parameters that spatially vary on a scale of 100 km (surface elevation, interfering emissivities) add 9%–16%. Measurement noise with a standard deviation of  $10^{-4} \text{ W}/(\text{m}^2 \text{ sr } \mu\text{m})$  leads to additional 1%–4%. Reasonable modifications of *a priori* mean values have negligible impacts. Retrieved maps are most reliable at 1.02  $\mu\text{m}$ . There is an overall tendency for better results for cases with small cloud opacity, high surface elevation, high emissivity, and small observation angle, but this depends on the emissivity window, retrieval pipeline, and measurement repetition number. Calibration, preprocessing, and simulation errors can lead to additional errors. Based on the presented results, a subsequent paper will discuss emissivity data retrieval for a selected surface target.

**Keywords:** Venus, Surface emissivity, Retrieval error, VIRTIS

---

---

\*Corresponding Author

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

## 1. Introduction

Insight into Venus' surface geology is currently quite limited. It is mainly based on radar data (e.g. topography, slope, reflectivity, emissivity acquired by the Magellan probe (Pettengill et al., 1991)), see Ivanov and Head (2011), ground based infrared observations (Meadows and Crisp, 1996), gravitational data (Smrekar et al., 2010), as well as on a few *in situ* measurements (Abdrakhimov and Basilevsky, 2002). Knowledge of wavelength dependent surface emissivity in the infrared can help to better categorize the surface material, texture, and weathering. Such data can be used to study the surface geology, and thus to learn more about the planet's geologic history, for instance about the global resurfacing event (Basilevsky et al., 1997).

Neglecting geologic activity and assuming thermodynamic equilibrium between surface and bottom of the atmosphere, surface temperature decreases with increasing surface elevation according to the Venus International Reference Atmosphere VIRA (Seiff et al., 1985). The hot surface (735 K at 0 km) emits radiation that depends on both surface temperature and surface emissivity. Measured upwelling radiation carries information on surface as well as on emissions from the hot deep atmosphere. But the signals are absorbed and multiply scattered by atmospheric gases and clouds. This also leads to a partial loss of spatial information according to an approximate Gaussian blurring with full-width-at-half-maximum (FWHM) in the order of 100 km (Moroz, 2002). Surface emissions in the infrared are far outweighed by scattered sunlight and can thus only be evaluated on Venus' nightside. Broad regions of nightside infrared spectra are completely blacked out by the high-density hot carbon dioxide atmosphere and the thick sulfuric acid clouds. In the infrared, only a few narrow atmospheric transparency windows between 0.8 and 1.3  $\mu\text{m}$  remain to sound the surface. But measurable top-of-atmosphere radiances are strongly attenuated and distorted by the overlying atmospheric layers (Pollack et al., 1993; Meadows and Crisp, 1996). Non-LTE  $\text{O}_2$  emissions (' $\text{O}_2$  nightglow') at 1.27  $\mu\text{m}$  from an altitude region around 100 km (Piccioni et al., 2009) strongly affect the 1.28  $\mu\text{m}$  radiance peak. Additional windows between 1.3 and 2.6  $\mu\text{m}$  probe the deep atmosphere.

The only suitable data on a global scale with high spectral, spatial, and temporal resolution and coverage at present is supplied by Venus nightside emission measurements acquired by the InfraRed Mapping channel of the Visible and InfraRed Thermal Imaging Spectrometer (VIRTIS-M-IR) aboard ESA's Venus Express (VEX) space probe (Drossart et al., 2007; Piccioni et al., 2007; Arnold et al., 2012). The instrument maps spectrally resolved (432 spectral bands dividing the range 1.0 – 5.1  $\mu\text{m}$  equidistantly with wavelength) two-dimensional images of targets on Venus. The carefully calibrated and preprocessed measurements (Cardesin-Moinelo et al., 2010; Kappel et al., 2012c) provide the data base where surface information shall be extracted from. Radiances in the 1.28  $\mu\text{m}$  window are not very sensitive against surface emissivity changes. Therefore, and because the instrument's limited spectral resolution does not allow to disentangle interfering  $\text{O}_2$  nightglow from emissivity responses in this window, surface emissivity cannot be determined there. Only three radiance peaks between 1.0 and 1.2  $\mu\text{m}$  remain, where surface emissivity can be derived from (1.02, 1.10, 1.18  $\mu\text{m}$ ). Emissivity is not easy to obtain from that data, since spectral information content is comparatively low, and atmospheric influences strongly interfere with surface information.

The radiative transfer in the atmosphere of Venus can be modeled by suitable algorithms (e.g. Haus and Arnold (2010)). Such a radiative transfer simulation forward model computes the synthetic radiance spectrum for a given state of the surface, atmosphere, measuring instrument, and for given optical properties of the gases and clouds, given observational geometry, and if necessary illuminational geometry and solar spectrum, etc. Starting from an initial guess, a retrieval algorithm iteratively varies the parameters that shall be retrieved from a measured spectrum until the corresponding simulation well fits the measurement in the least-squares sense. The parameter values that yield the best-fit spectrum ('retrieved parameters') then adequately parameterize the measurement and can be regarded as estimations of the corresponding actual states of atmosphere, surface, and instrument.

However, different parameter combinations may describe the same spectrum equally well. Thus, retrieval is mathematically an ill-posed problem. A usual way to treat such problems is by Bayesian regularization (Rodgers, 2000), where the parameters themselves have to fulfill additional conditions. This is achieved by assuming that, prior to the knowledge of the measurement outcome, the parameters follow an *a priori* probability distribution, a Gaussian with certain mean value vector and standard deviations. Measurement and simulation errors are assumed to follow a Gaussian with zero mean value vector and certain standard

deviations. The location  $\hat{\mathbf{x}}$  of the global maximum of the Bayesian *a posteriori* probability distribution  $\exp(-F_c(\mathbf{x})/2)/N$  of the parameter vector  $\mathbf{x}$ , given the measured spectrum, is then the state vector that is the most consistent with the measurement and the *a priori* and error information.  $F_c$  is the retrieval cost function that has to be minimized to find that global maximum,  $N$  is a normalization factor. An iterative algorithm like the trust region formulation of the Levenberg-Marquardt algorithm (Moré, 1978) can be used to identify local minima of  $F_c$ , but note that the algorithm possibly converges to just a subsidiary minimum, especially in the presence of measurement noise. The *a posteriori* covariance matrix (essentially the inverse of the Hessian of  $F_c/2$  at  $\hat{\mathbf{x}}$ ) can be regarded as a first measure for statistical uncertainties of, and correlations between the retrieved parameters. Basically, the incorporation of *a priori* data decreases the probability to find unreasonable parameter values. As this regularization considers only independent single spectra, it will be called 'single-spectrum regularization'. However, as will be shown in the present paper, retrieved emissivities turn out to have very high uncertainty margins.

But contiguous measurements are not likely to originate from completely unrelated state vectors. For atmospheric parameters, this follows from the inertia of matter and the balancing of thermodynamic disequilibria. If contiguous measurements are nevertheless treated as independent from each other, then spatial or temporal continuity in the measurements may not translate to a certain expected continuity in retrieved parameters. This is due to the existence of different parameters with very similar impacts on the spectra, the ill-posed nature of the retrieval problem, and the existence of subsidiary minima of the cost function. Therefore, a more sophisticated retrieval regularization was developed that in addition incorporates *a priori* spatial-temporal correlations between state vectors and also allows to retrieve parameters that are common to certain selections of measurements (Kappel, 2014). This 'multi-spectrum retrieval algorithm' (MSR) essentially decreases the probability to find unreasonable spatial-temporal state vector distributions. Although MSR decreases the number of potential solutions and could be expected to produce larger residuals between measurements and fits, the residuals in fact decrease on average, which was demonstrated by Kappel et al. (2012a,c) for real spectra, and by Kappel (2014) for synthetic spectra. This indicates that MSR helps to avoid subsidiary solutions. It was also shown that parameters with more different *a priori* correlation lengths or times can be better disentangled. In particular, it was demonstrated that for repeated measurements of a surface target, emissivity can be reasonably well disentangled from cloud parameters, since emissivity can be regarded as parameter with infinite correlation time. This is, because it is a parameter that is common to several measurements that cover the same surface target, when neglecting geologic activity.

It is always important to know the reliability of retrieved results when physical values are determined from measured data. This is especially true for Venus' surface emissivity that is difficult and error prone to retrieve from VIRTIS-M-IR measurements as will be shown. To begin with, the *a posteriori* covariance matrix encodes, how well the surface emissivity could be determined from the utilized measurements, and how well it could be disentangled from the other retrieved parameters, all within the frame of the utilized forward model and under prerequisite of the assumptions on *a priori* data and measurement noise and the choice of the retrieval parameters.

But there are errors that are not reflected in the *a posteriori* covariance matrix. For instance, the latter is completely local. It does not reflect errors that are caused by finding subsidiary solutions and that are due to the possibly complicated *global* dependency of the least-squares residual between measurement and simulation on the model parameters, especially in the presence of measurement noise. Also, the choice of the retrieval parameters, of the spectral ranges that shall be utilized in the retrieval, and of the *a priori* data can lead to different emissivity retrieval results. Next, there are always atmospheric and surface parameters that are not sufficiently well known to be accurately set in the simulation of a synthetic spectrum and that cannot be derived from the spectrum itself. They are usually set to reasonable values that are compatible with current knowledge on atmospheric and surface conditions and allow to generate realistic synthetic spectra. But when the true underlying parameter values deviate from these assumptions, this can lead to emissivity retrieval errors. This work will estimate the impacts of all error sources listed in this paragraph. It will turn out that these errors can be large compared with the errors encoded in the *a posteriori* covariance matrix.

Finally, forward model simulation errors can lead to errors in the retrieved emissivities, but this effect is difficult to estimate, because of the insufficient knowledge of the optical properties in Venus' hot dense atmosphere and the lack of *in situ* data to compare with. But when realistic input data are employed, the

utilized forward model is capable to generate synthetic spectra that very well match the measured spectra for very different kinds of atmospheric and measurement conditions (Haus and Arnold, 2010). Measurement, calibration, and preprocessing errors and their impacts on retrieved emissivities were already partially discussed by Kappel et al. (2012c).

Section 2 of the present paper defines the terminology and the algorithm that allows to determine the above listed surface emissivity retrieval errors. A number of error measures are defined that are used to monitor the success of the emissivity retrieval. For the case when the multi-spectrum retrieval algorithm MSR is applied, it is discussed, how multi-spectrum retrieval errors can be roughly estimated by plausibly scaling single-spectrum retrieval errors. Section 3 recites the properties of the forward model, discusses the various forward model parameters involved in this work, and defines several synthetic reference spectra that are representative for a number of typical key measurement situations. Section 4 investigates, which spectral ranges should be utilized in the retrieval procedures and which auxiliary parameters should be retrieved in order to minimize the emissivity retrieval errors and the noise impacts. It is also explored, how results depend on the measurement situation. The single-spectrum retrieval errors are scaled to obtain rough measures for the multi-spectrum retrieval errors. First results have been presented by Kappel et al. (2012b).

## 2. Notation and strategy of retrieval error analysis

### 2.1. Basic terminology

The 'primary state vector'  $\mathbf{e} \in \mathbb{R}^p$  shall comprise the parameters of primary interest, like surface emissivity (giving rise to the notation  $\mathbf{e}$ ), that shall be retrieved from a measured spectrum  $\mathbf{y} \in \mathbb{R}^m$ . The  $i$ -th entry of  $\mathbf{e}$ , where  $i \in \{1, \dots, p\}$ , is denoted as  $(\mathbf{e})_i =: e_i$ , and similarly, the entries of  $\mathbf{y}$  are the measured radiances  $y_k$  at wavelengths  $\lambda_k$ , where  $k \in \{1, \dots, m\}$ . It may be necessary to retrieve parameters that are not of primary interest but must be retrieved along with  $\mathbf{e}$  to obtain useful fits to  $\mathbf{y}$  at all, like cloud particle abundances. They shall be called 'auxiliary parameters' and are compiled into the 'auxiliary state vector'  $\mathbf{z} \in \mathbb{R}^q$ . Together,  $\mathbf{e}$  and  $\mathbf{z}$  comprise the 'state vector'  $\mathbf{x} := (\mathbf{e}^T, \mathbf{z}^T)^T \in \mathbb{R}^{p+q}$ . In the following, the concatenation of column vectors shall be shortly written without the superscript sign for transposition, such that  $\mathbf{x} := (\mathbf{e}, \mathbf{z})$ .

Not every relevant parameter can be retrieved from a measured spectrum due to the limited spectral information content. For instance, parameters with small or unspecific spectral signatures cannot be determined. However, they still may affect the spectrum and must be set to certain values to enable the radiative transfer simulation at all. Deviations of the true from these assumed values can lead to errors in the retrieval of  $\mathbf{x}$ . Those parameters shall be called 'external parameters' and are combined into the 'external parameter vector'  $\mathbf{t} \in \mathbb{R}^r$ . Their impact on the retrieved  $\mathbf{e}$  is one of the main subjects of this paper.

Finally, there are parameters that are needed for the radiative transfer simulation but whose possible uncertainties are not considered in this paper, like viewing geometry.

The forward model simulation for the parameter combination  $(\mathbf{x}, \mathbf{t})$  shall be denoted as  $\mathbf{f}(\mathbf{x}, \mathbf{t}) \in \mathbb{R}^m$  and is the synthetic spectrum that has to be fitted to  $\mathbf{y}$  by iterative variations of  $\mathbf{x}$ . The partial derivative of  $\mathbf{f}$  with respect to a scalar parameter  $s$  is denoted as  $(\mathbf{J})_s := \partial \mathbf{f} / \partial s \in \mathbb{R}^m$ , and the matrix of all first-order partial derivatives, i.e. the Jacobian, as  $\mathbf{J} \in \mathbb{R}^{m \times (p+q+r)}$ . It is common to refer to  $(\mathbf{J})_s$  also as 'Jacobian'.

For a given spectrum, the 'true' values that actually led to the spectrum, shall be denoted as  $\underline{\mathbf{e}}$  with entries  $\underline{e}_i$ , and similarly with  $\underline{\mathbf{z}}$ ,  $\underline{\mathbf{x}}$ , and  $\underline{\mathbf{t}}$ . When retrieving  $\mathbf{x} = (\mathbf{e}, \mathbf{z})$  from the spectrum, the best estimate for  $\mathbf{e}$ , i.e. the retrieved primary state vector, shall be denoted as  $\hat{\mathbf{e}}$  with entries  $\hat{e}_i$ , and similarly with  $\hat{\mathbf{z}}$ .

In order to determine the impact that wrong assumptions on the external parameters  $\mathbf{t}$  may have on  $\hat{\mathbf{e}}$ , the true parameter values must be known for a studied spectrum. There are only few *in situ* data of surface properties. Also, there are no *in situ* measurements of the interfering atmospheric parameters at all that are simultaneous with VIRTIS-M-IR measurements of Venus' nightside emissions. Thus, only synthetic spectra can serve as spectra with known true underlying parameters. Therefore, a 'true' spectrum  $\mathbf{y}$  is now defined as the synthetic spectrum  $\mathbf{y} = \mathbf{f}(\underline{\mathbf{x}}, \underline{\mathbf{t}})$  using a certain parameter set  $(\underline{\mathbf{x}}, \underline{\mathbf{t}})$  (the 'true' parameters for  $\mathbf{y}$ ).

Now let the primary state vector  $\mathbf{e}$  comprise surface emissivities  $e_i$  at wavelengths  $\lambda_i$  ( $i \in I$ ) that are sensitive to surface emissivity. Focusing on VIRTIS-M-IR measurements, only three surface windows (at 1.02, 1.10, and 1.18  $\mu\text{m}$ ) can be used in practice for emissivity retrieval. The set  $I$  may thus be written  $I = \{1, 2, 3\}$ ,

or more suggestively  $I = \{1.02, 1.10, 1.18\}$ , which then leads to the notation  $e_{1.02}$  for the emissivity in the 1.02  $\mu\text{m}$  window, etc. Now vary a single external parameter  $t$ . To study the impact of variations of  $t$  on the retrieved  $\hat{e}_i$ , a series of  $w$  'wrong' values  $t_v$  ( $v \in \{1, \dots, w\}$ ) is defined. The  $t_v$  are chosen to cover the range where  $t$  is expected to lie in. Along with  $\mathbf{z}$ ,  $\mathbf{e}$  is then retrieved from always the same  $\mathbf{y}$  for each of the  $t_v$ , yielding  $\hat{\mathbf{e}}(t_v)$ . While the  $\hat{e}_i(t_v)$  are monitored for further evaluation, the auxiliary parameters are not considered further on. The ranges of the  $\hat{e}_i(t_v)$  are then measures for the retrieval errors introduced by wrong assumptions on  $t$ . These error measures are determined for the most important external parameters and yield an estimate of the uncertainty of  $\hat{\mathbf{e}}$  that is due to the uncertainty of these parameters.

The retrieval error analysis procedure is repeated for several different combinations of auxiliary parameters and spectral ranges utilized for the retrieval (Section 4). For each of these retrieval pipelines, the retrieval errors are compared, in order to determine an optimal choice for operational retrievals that yields small retrieval errors. Also, the robustness of retrieved results against modifications of *a priori* data is checked. For the best retrieval pipeline, a number of different reference spectra are studied that represent typical measurement situations (Section 4.6).

## 2.2. Error measures for single-spectrum retrieval

Now let the true emissivity  $\underline{e}$  be the same for all  $\lambda_i$ . The retrieved surface emissivities  $\hat{e}_i(t)$  may depend on the external parameter  $t$  and on  $\lambda_i$ . Note that for simplicity, the parameter name is notationally not distinguished from the value it attains. The corresponding surface emissivity retrieval errors shall be summarized by a few characteristic values. For this purpose, define the characteristic range  $CR$  the external parameter  $t$  is expected to typically (95% probability) vary in.

A first measure  $RG_i$  ('Global eRror') for emissivity retrieval errors for each  $i \in I$  is then the maximum relative deviation of the retrieved emissivities from the true emissivity when  $t$  varies over  $CR$ .

$$RG_i := \max_{t \in CR} \frac{|\hat{e}_i(t) - \underline{e}|}{\underline{e}} \quad (1)$$

In addition, define a measure  $CD$  for the worst case 'Color Dispersion' of the retrieved emissivities, i.e. compute at fixed  $t$  the maximum difference between the  $\hat{e}_i(t)$ , relative to the center of the range of the  $\hat{e}_i(t)$ , and find the maximum of that value when  $t$  varies over  $CR$ .

$$CD := \max_{t \in CR} \frac{|\max_{i \in I} [\hat{e}_i(t)] - \min_{i \in I} [\hat{e}_i(t)]|}{|\max_{i \in I} [\hat{e}_i(t)] + \min_{i \in I} [\hat{e}_i(t)]|/2} \quad (2)$$

In practice, it may be difficult to retrieve the absolute emissivities, and only relative spatial variations may be accessible. To have a quantitative estimation to describe this situation for simultaneously measured spectra, let the retrieved emissivity at wavelength  $\lambda_i$  and location  $A$  be  $\hat{e}_i(t_A)$ , and analogously for location  $B$ . Both  $t_A$  and  $t_B$  may be uncertain, but when  $B$  is close to  $A$ , then  $t_B$  may not be very different from  $t_A$ . The expected spatial resolution of retrieved surface data is in the order of 100 km (Moroz, 2002), so it is useful to consider spatially relative surface emissivity retrieval errors that are due to possible variations between  $t_A$  and  $t_B$ , when  $A$  and  $B$  are separated by not more than 100 km.

In order to estimate the typical variation between  $t_A$  and  $t_B$  on the length scale of 100 km, the correlation function  $f_d$  for parameter  $t$  is introduced. It describes the statistical correlation between  $t_A$  and  $t_B$  in dependence on the Euclidean distance  $d(A, B)$  between  $A$  and  $B$ . It is closely related to the *a priori* correlation function that describes statistical correlations of retrieval parameters before knowledge from their measurement is available, see Kappel (2014, Section 3.2.1) for a discussion of the latter in context of a multi-spectrum retrieval algorithm. A reasonable correlation function  $f_d$  was given by Gaspari and Cohn (1999, Eq. 4.10) and used by Kappel (2014, Eq. (6)), see Fig. 1.

$$f_d(x) = \begin{cases} -\frac{x^5}{4} + \frac{x^4}{2} + \frac{5x^3}{8} - \frac{5x^2}{3} + 1 & 0 \leq x < 1 \\ \frac{x^5}{12} - \frac{x^4}{2} + \frac{5x^3}{8} + \frac{5x^2}{3} - 5x + 4 - \frac{2}{3x} & 1 \leq x < 2 \\ 0 & x \geq 2 \end{cases} \quad (3)$$

Here,  $x = 0.808768 \cdot d(A, B)/\lambda_t$ . The correlation length  $\lambda_t$  for parameter  $t$  is the distance where the correlation function attains the value  $e^{-1}$ .  $f_d$  is continuous and twice continuously differentiable on  $\mathbb{R}_+$ . It has derivative 0 at  $d = 0$  and thus reflects inertial properties of atmospheric molecules and the fast balancing of thermodynamic disequilibria, because observations that are only slightly separated, are modeled to perceive slowly changing correlations, see discussions by Daley (1991); Gelb (1974); Gaspari and Cohn (1999); Balgovind et al. (1983).  $f_d$  has compact support (it vanishes outside of a closed bounded set) and thus reflects the effective vanishing of correlations for widely separated  $A$  and  $B$ , see discussion by Gaspari and Cohn (1999, Section 4.d) and references therein on forecast error correlations for Earth's troposphere.

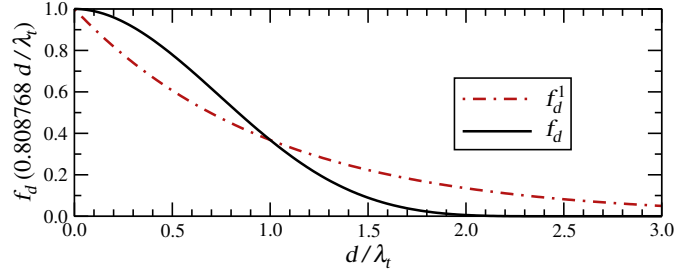


Figure 1: Correlation function  $f_d$  in comparison to a simple exponential correlation function  $f_d^1$  with the same correlation length.  $f_d$  vanishes for  $d \geq 2.47\lambda_t$ . It has zero derivative at  $d = 0$  and, compared to  $f_d^1$ , it has a stronger relative weighting of the correlation of nearby measurements ( $d < \lambda_t$ ), which better represents the inertial properties of the physical system.

$f_d$  gives rise to a  $2 \times 2$  matrix  $\mathbf{C}$  with entries  $C_{ij} := f_d(0.808768 \cdot d_{ij}/\lambda_t)$ , where the  $d_{ij}$  are the Euclidean distances that arise for the locations  $A$  and  $B$ :  $d(A, A) =: d_{11} = 0 = d_{22} := d(B, B)$ ,  $d_{21} := d(A, B) =: d_{12}$ . Since  $f_d$  is a correlation function,  $\mathbf{C}$  is positive definite for distinct  $A$  and  $B$  (Gaspari and Cohn, 1999).  $\mathbf{C}$  is also real, symmetric, and all diagonal entries are 1. Hence, it is a correlation matrix (Anderson, 1958, Theorem 2.3.1). It describes the statistical correlation between  $t_A$  and  $t_B$ . Let the probability distributions for  $t_A$  and  $t_B$  both be Gaussian with mean value  $\bar{t}$  and standard deviation  $\sigma_t$ .  $CR$  can thus be set to

$$CR := [\bar{t} - 2\sigma_t, \bar{t} + 2\sigma_t]. \quad (4)$$

Then the corresponding joint probability distribution for  $t_A$  and  $t_B$  is  $p(t_A, t_B) = \frac{1}{NF} \exp(-\frac{1}{2}\mathbf{v}^T \mathbf{S}^{-1} \mathbf{v})$  with normalization factor  $NF$ . Here,  $\mathbf{S} := \sigma_t^2 \mathbf{C}$  is the covariance matrix that arises from the correlation matrix  $\mathbf{C}$  and the standard deviation  $\sigma_t$ , and  $\mathbf{v} := (t_A - \bar{t}, t_B - \bar{t})^T$ . Hence,  $p(t_A, t_B) = \frac{1}{NF} \exp(-[(t_A - \bar{t})^2 + (t_B - \bar{t})^2 - 2C_{12}(t_A - \bar{t})(t_B - \bar{t})]/[2\sigma_t^2(1 - C_{12}^2)])$ . Then the (conditional) probability distribution for  $t_B$  for given fixed  $t_A$  is a Gaussian with maximum  $\frac{1}{NF} \exp(-(t_A - \bar{t})^2/(2\sigma_t^2))$ , which is located at  $t_B = \bar{t} + C_{12}(t_A - \bar{t})$ , the corresponding mean value. The corresponding double standard deviation is  $2\sigma_t\sqrt{1 - C_{12}^2}$ . Thus, for given  $t_A$ , the probability is 95% to find  $t_B$  in interval

$$I_{t_A} := \bar{t} + C_{12}(t_A - \bar{t}) + \sqrt{1 - C_{12}^2} \cdot [-2\sigma_t, +2\sigma_t]. \quad (5)$$

This defines a measure for relative spatial errors  $RL_i$  ('reLative eRror') for the retrieved emissivity  $\hat{e}_i$ ,

$$RL_i := \frac{1}{2} \max_{t_A \in CR} \left( \max_{t \in I_{t_A} \cap CR} \frac{|\hat{e}_i(t) - \hat{e}_i(t_A)|}{\hat{e}_i(t_A)} \right), \quad (6)$$

i.e. the maximum of the ratio is computed for  $t$  in a sliding interval  $I_{t_A} \cap CR$  around a  $t_A$  that varies in  $CR$ . The factor  $1/2$  is needed, because it is assumed that retrieved emissivities at both locations  $A$  and  $B$  are affected by errors, and without this factor,  $RL$  would indicate the error at  $B$  relative to that at  $A$ . For strong correlations  $C_{12}$  close to 1, the interval  $I_{t_A}$  can be approximated by  $I_{t_A} = t_A + \sqrt{1 - C_{12}^2} \cdot [-2\sigma_t, +2\sigma_t]$ . This means that over a sliding interval of length  $2I_t$  with

$$I_t := 2\sigma_t\sqrt{1 - C_{12}^2}, \quad (7)$$

half the maximum relative difference of  $\hat{e}_i(t)$  to its value at the center of the interval is wanted. For the computation of  $RL_i$ , it is thus sufficient for parameters with  $C_{12}$  close to 1 to only state  $I_t$  instead of  $I_{t_A} = t_A + [-I_t, I_t]$  in the following. On the other hand, when  $C_{12} \ll 1$ , then  $I_{t_A}$  can be approximated by  $CR$  (Eq. (4)). In this case,  $RL_i$  is just the half of the maximal relative difference between any two  $\hat{e}_i(t)$  that is possible over the interval  $CR$ . This information may be encoded by again using the approximate interval  $I_{t_A} = t_A + [-I_t, I_t]$  but now setting  $I_t := 4\sigma_t$ .

To compute  $RG_i$ ,  $CD$ , and  $RL_i$ , the expected mean value  $\bar{t}$ , double standard deviation  $2\sigma_t$ , and correlation length  $\lambda_t$  are given in Section 3 for each external parameter  $t$ . The ideal case would be  $RG_i = 0$  for all  $i$ , which would imply  $CD = 0$  and  $RL_i = 0$  for all  $i$ . Note that  $RL_i = 0$  or  $CD = 0$  do not imply  $RG_i = 0$ .

Random measurement errors destroy some of the spectrum's information content and can lead to a more complicated shape of the graph of the retrieval cost function and thus to an increased number of subsidiary minima and therefore wrong potential retrieval results. To estimate the impact on retrieved emissivities, the retrieval of the  $e_i(t)$  at fixed  $t$  is repeated  $N$  times, where each time a different noise  $\epsilon_n$  ( $n \in \{1, \dots, N\}$ ) is added to  $\mathbf{y}$ , yielding 'noisy true spectra' or 'synthetic measurements'  $\mathbf{y}_{\epsilon n} := \mathbf{y} + \epsilon_n$ . Only a simple noise model shall be used here, Gaussian noise with zero mean value and a certain wavelength independent standard deviation  $\sigma$ . This way, also the numerical repeatability of the retrieval under slightly altered conditions can be studied. Denoting the emissivity retrieved at wavelength  $\lambda_i$  and at fixed external parameter  $t$  from spectrum  $\mathbf{y}_{\epsilon n}$  by  $\hat{e}_{in}(t)$ , then  $\bar{e}_i(t)$  shall be the mean value of the  $\hat{e}_{in}(t)$  over all  $N$  noise repetitions and  $\sigma_{N_i}(t)$  the corresponding standard deviation. The corresponding error measures are denoted by  $\overline{RG}_i$ ,  $\overline{CD}$ , and  $\overline{RL}_i$  and are computed according to Eqs. (1), (2), and (6), but with  $\hat{e}_i(t)$  replaced by  $\bar{e}_i(t)$ .

When the number  $N$  of noise repetitions is large,  $\bar{e}_i(t)$  is expected to be very similar to the noiseless result  $\hat{e}_i(t)$ . Differences are due to both remaining statistical fluctuations and subsidiary solutions. To minimize the impact of statistical fluctuations on  $\overline{RG}_i$ ,  $\overline{CD}$ , and  $\overline{RL}_i$ , the number  $N$  must be set to a reasonably high value. Note that fluctuation impacts scale with  $1/\sqrt{N}$ . But to make computational resources easier to handle, the similarity of  $\bar{e}_i(t)$  and  $\hat{e}_i(t)$  shall be verified with  $N := 100$  for only one typical synthetic measurement. In the remaining cases, just the noiseless  $RG_i$ ,  $CD$ , and  $RL_i$  shall be determined (Section 4.6).

It will turn out that the double standard deviations  $2\sigma_{N_i}(t)$  of the retrieved emissivities for the  $N$  noise repetitions vary only slightly with the value  $t$  (see for instance Fig. 5). They are also quite independent from the choice of which external parameter is studied. For each of the wavelengths  $\lambda_i$ , the median  $2\sigma_i$  over all those double standard deviations is computed. For better comparison to the error measures, the  $2\sigma_i$  are divided by the true emissivity (by analogy with Eq. (1)), yielding the 'noise measures'

$$N_i := \frac{2\sigma_i}{e}. \quad (8)$$

As a measure  $QF$  for monitoring the quality of the fits, define the (unitless) relative deviation of the fit  $\mathbf{f}_n(t)$  from the synthetic measurement  $\mathbf{y}_{\epsilon n}$  as the Euclidean norm (denoted by  $\|\cdot\|_2$ ) of their difference, divided by the Euclidean norm of the synthetic measurement,  $QF_n(t) := \|\mathbf{y}_{\epsilon n} - \mathbf{f}_n(t)\|_2 / \|\mathbf{y}_{\epsilon n}\|_2$ . A good fit can be characterized by a small  $QF$  in the order of  $\|\epsilon_n\|_2 / \|\mathbf{y}_{\epsilon n}\|_2$ . For  $\mathbf{y} \in \mathbb{R}^m$ , this can be approximated by the value  $QF_0 := \sigma\sqrt{m} / \|\mathbf{y}\|_2$ , because the standard deviation  $\sigma$  of the synthetic measurement noise with zero mean value  $\mu$  can be estimated by  $\sqrt{\frac{1}{m} \sum_{k=1}^m [(\epsilon_n)_k - \mu]^2} = \frac{1}{\sqrt{m}} \|\epsilon_n\|_2$ . When  $QF$  exceeds a certain threshold value, signaling a failed retrieval run due to a bad fit, the retrieved emissivities will not be taken into account for the error evaluation, and the maximal valid interval for the external parameter is stated. The threshold value is set to 3%, which exceeds  $QF_0$  for the reference spectrum with the smallest  $\|\mathbf{y}\|_2$  (0.137 W/(m<sup>2</sup> sr  $\mu$ m)) for scenario 2 from Section 3.8,  $\sigma=10^{-4}$  W/(m<sup>2</sup> sr  $\mu$ m),  $m = 129$  for the range 1.0–2.3  $\mu$ m excluding the 1.28  $\mu$ m peak, see Section 3.1) by half an order of magnitude as safety margin. For real measurements,  $QF$  is always monitored and rarely exceeds 6%, and it is typically between 3%–5%. These values even include all measurement and simulation errors, which are not investigated in the present study.

There are three kinds of external parameters that will be studied here. The first kind can vary spatially and temporally and is denoted by 'Atm' as it comprises only atmospheric parameters. The second kind can vary only spatially ('Spa'), and the third kind varies neither spatially nor temporally and is denoted by 'Cnt', since it comprises only so-called 'continuum' parameters, see Section 3.6. Note that Table 1

indicates the kinds of all external parameters discussed in Section 3. Sums of error measures over different external parameters are most useful when these parameters are of the same kind. Assuming the spatially and temporally varying external parameters to vary statistically independently of each other, the global error measures of those parameters can be summarized for each  $i$  by a single value  $\langle \overline{RG}_i \rangle^{Atm}$ .  $\langle \overline{RG}_i \rangle^{Atm}$  is the square root of the sum of the squares of those global error measures at wavelength  $i$  over the parameters of the first kind (*Atm*), i.e. these error measures are quadratically added.  $\langle \overline{CD} \rangle^{Atm}$  and  $\langle \overline{RL}_i \rangle^{Atm}$  are defined analogously. Together, they are denoted as  $\langle Atm \rangle$ . The summarized values for the parameter kinds *Spa* and *Cnt* are similarly defined. All together,  $\langle Atm \rangle$ ,  $\langle Spa \rangle$ , and  $\langle Cnt \rangle$  are called the 'summarized error measures'.

### 2.3. Scaling properties of error measures for multi-spectrum retrieval

It will turn out (Section 4) that retrieval errors can be very large when emissivity is retrieved from just one single spectrum. The multi-spectrum retrieval algorithm (MSR) developed by Kappel (2014) allows to retrieve the emissivity of a surface bin as a parameter that is common to  $N_r$  measurements that all cover that bin. This approach assumes that geologic activity as observable by VIRTIS-M-IR is negligible, and indeed, corresponding hints on its presence have not been found in the data so far (Müller et al., 2012). In this sense, emissivity itself has the character of a *Spa* parameter. Emissivities retrieved with MSR are more reliable. In part, this is due to the fact that for a high measurement repetition number  $N_r$ , the impact of fluctuations around the long-term average of an *Atm* parameter at that surface bin should average out with  $1/\sqrt{N_r}$ . While a multi- is more complex than a single-spectrum retrieval error analysis, a scaling by this statistical factor yields a rough measure for the order of magnitude of all averaged-out single-spectrum *Atm* error measures  $\overline{RG}_i$  and  $\overline{CD}$ , including the corresponding 'summarized' error measures. Note that this scaling approach is based on the assumption that the mean values of the external parameters at that surface bin coincide with the 'true' parameter values discussed in Section 3. Deviations from this assumption will lead to retrieval errors that may systematically differ from the statistically scaled errors, but these differences can be largely removed as will be discussed in the following paragraphs. A scaling of relative errors in dependence on the repetition number is not considered here. Relative errors were only defined for simultaneous measurements in order to simplify their definitions by only considering correlation lengths and not correlation times. The noise measures  $N_i$  (Eq. (8)) should scale with  $1/\sqrt{N_r}$  as well. In contrast, errors for *Spa* and *Cnt* parameters will not average out, since they do not statistically vary with time.

*Cnt* parameters, the largest sources of error, are fixed for all measurements and should mostly lead to errors that affect retrieved emissivities in the same way. When the ambition to retrieve absolute emissivities is dropped and only spatial emissivity variations shall be studied, one can set an initial emissivity value (e.g. 0.65) and determine compatible *Cnt* parameters by retrieving them with MSR as parameters that are common to as many and as diverse as possible measurements. For the actual target area, an emissivity map relative to the initial value can then be retrieved with MSR by using these pre-determined *Cnt* parameters. The *Cnt* error measures are then assumed to be negligible.

But even then, there may remain spatially large scale deviations of the retrieved from the true emissivity map. This is due to spatially slowly (on a scale of thousands of km) varying systematic interferences from other than *Cnt* parameters. These deviations can be removed as follows. Averages of *Atm* parameters over time are not expected to show distinct spatial features on scales below  $10^\circ$  in latitude direction and a few hours in local time direction, respectively (when sufficiently far away from the poles or the terminators), compare Haus et al. (2013, 2014) for cloud parameters. For a surface target not exceeding such dimensions and where only the spatial emissivity fine structure is of interest, the relative emissivity map can be additively 'renormalized' by a constant value, such that the map's new median corresponds to the initial emissivity. The above mentioned systematic differences from the statistically scaled retrieval errors for the cases where the actual mean values of the *Atm* parameters deviate from the assumed values can be largely removed this way. The corresponding summarized *Atm* error measures can then be scaled with  $1/\sqrt{N_r}$  (except for the relative errors) and are denoted by  ${}^s\langle Atm \rangle$ . The scaled noise measures are similarly denoted by  ${}^sN_i := N_i/\sqrt{N_r}$ . Errors due to uncertainties of spatially slowly varying *Spa* parameters (deep atmospheric temperature field, Section 3.3) are spatially slowly varying and are denoted by  ${}^s\langle Spa \rangle_S$ , even though they are not scaled. For a renormalized relative emissivity map of a target of the mentioned dimensions, they can be neglected as spatially almost constant emissivity offsets that have been removed by the renormalization.



There are *Spa* parameters that are spatially fast varying (on a scale of 100 km), for instance surface elevation (Section 3.7). Also, Section 4.5 investigates a retrieval pipeline where emissivity directly interferes with the auxiliary parameters, because the latter are retrieved from a spectral range that includes the surface windows. For a given surface window, the emissivities from the other windows are then external parameters for the emissivity in the given window and will lead to additional *Spa* errors. In practice, however, clouds have *a priori* correlation lengths (500 km, see Section 3.5) and times (a few hours) different from those for emissivity (50 km,  $\infty$  h, see Section 3.2). As has been shown by Kappel (2014), incorporation of *a priori* correlation lengths and times improves disentanglement of retrieval parameters, especially when they have very different *a priori* correlation lengths or times. The impact of the better disentanglement of clouds and emissivities on the *Spa* errors may be roughly estimated by scaling the external emissivity errors with the statistical factor  $1/\sqrt{N_r}$ . This can be motivated by the consideration that more repetitions should result in lesser impacts of interfering external emissivities, since clouds vary in time but the interfering emissivities do not. Even the relative *Spa* errors caused by interfering emissivities can be scaled in the same way, because their *a priori* correlation time is infinite. Note that such a scaling does not apply to surface elevation, because this is not a retrieval parameter. Scaled errors due to spatially fast varying parameters are summarized by quadratically adding the scaled external emissivity contributions and the (unscaled) contributions from other spatially fast varying *Spa* parameters. This 'spatial fine structure' part is denoted by  ${}^s\langle Spa \rangle_F$ . Note that it is not sufficient to renormalize an emissivity map that is not retrieved relative to an initial emissivity. In this case, the retrieved map can easily become saturated by approaching one of the emissivity bounds 0 or 1, and renormalization yields only a featureless map. A subsequent paper will exemplarily show that a renormalized relative emissivity map retrieved from actual measurements is reasonably independent of the *Cnt* parameters, the initial emissivity, and spatially and temporally constant modifications to *Atm* parameters.

This work studies retrieval errors only. Calibration and preprocessing errors (see Kappel et al. (2012c)) of the measured spectra as well as simulation errors and unconsidered external parameters can lead to additional errors for the retrieved emissivities.

### 3. Parameters and reference spectra

Section 3.1 recites some aspects of the radiative transfer forward model. Then, forward model parameters are discussed that may affect quantitative results of emissivity retrieval from VIRTIS-M-IR spectra of Venus' nightside emissions. For a certain 'default reference spectrum', the 'true' model parameter values are given. Also, for each discussed parameter, the range *CR* (see Eq. (4)) is specified over which it is expected to vary. These values are given in form of the expected mean value  $\bar{t}$  and double standard deviation  $2\sigma_t$ . Also, the interval length  $I_t$  (Eq. (7)) is given, based on a reasonable correlation length  $\lambda_t$ . The stated information that is discussed in Sections 3.2 to 3.7 and needed in Section 4 is summarized in Table 1. For the noiseless default reference spectrum, Fig. 2 displays radiance and Jacobians with respect to the discussed parameters in the range 1.0–2.3  $\mu\text{m}$ . Section 3.8 introduces a number of additional synthetic reference spectra that cover several important measurement situations.

#### 3.1. Forward model

In order to numerically simulate observable nightside radiances, a plane-parallel, LTE, line-by-line radiative transfer 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) and was also used by Kappel (2014), where the following summary is based on.

Temperature altitude profiles are taken from the Venus International Reference Atmosphere (VIRA, Seiff et al. (1985)) but are optionally modified according to Section 3.3. 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 surface elevation as initial value. CO<sub>2</sub> volume mixing ratio is 96.5%. Considered minor gaseous constituents are H<sub>2</sub>O, CO, SO<sub>2</sub>, OCS, HCl, and HF. HDO is treated such that for the default reference model, the D/H ratio corresponds to 150 times the

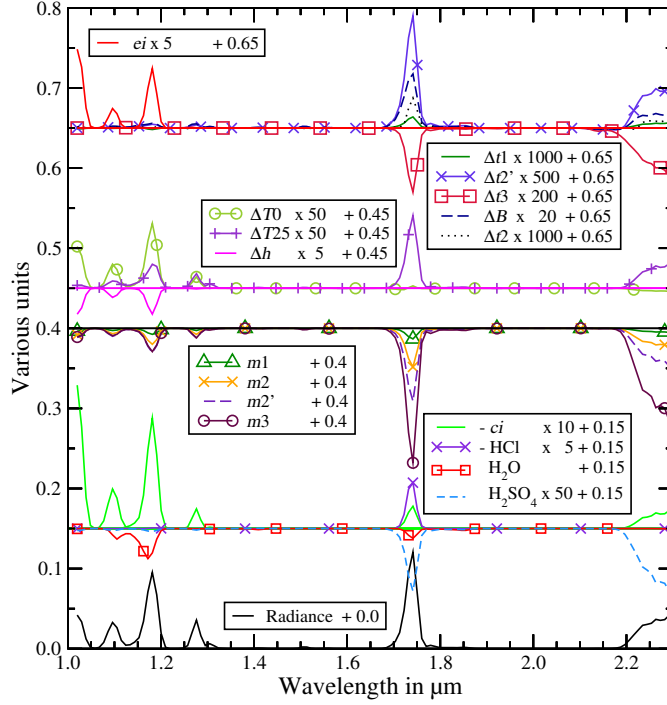


Figure 2: (From bottom to top) Offset +0.0: synthetic radiance spectrum in units of  $R:=W/(m^2 \text{ sr } \mu\text{m})$ . Various positive offsets: Jacobians with respect to:  $\text{H}_2\text{SO}_4$  concentration of cloud droplets ( $'\text{H}_2\text{SO}_4'$ , unit  $R/100\%$ ),  $\text{H}_2\text{O}$  and  $\text{HCl}$  column factors (units  $R$ ), continuum factor in window  $i$  ( $'ci'$ , between 2.1 and 2.5  $\mu\text{m}$  scaled by factor of 10 relative to remaining range for better representation, unit  $R \cdot 10^{29} \text{ cm}^{-2}$ ), cloud mode  $i$  column factor ( $'mi'$ , unit  $R$ ), variation of surface elevation ( $'\Delta h'$ , unit  $R/\text{km}$ ), variation of temperature profile near 0 and 25 km, resp. ( $'\Delta T25'$ ,  $'\Delta T0'$ , unit  $R/\text{K}$ ), cloud altitude profile modifications ( $'\Delta B'$ ,  $'\Delta ti'$ , units  $R/\text{km}$ ), and surface emissivity in window  $i$  ( $'ei'$ , unit  $R$ ), compare Table 1. The notation  $' \times 50'$  etc. in the legends indicates by which factor, if any, a Jacobian was scaled for better representation.

corresponding ratio for Earth, compare with literature values of  $95 \pm 15$  at 74 km (Krasnopolsky et al., 2013);  $120 \pm 40$  below the clouds (de Bergh et al., 1991);  $157 \pm 15$  at 72 km (Bjoraker et al., 1992). A  $\text{HDO}/\text{H}_2\text{O}$  ratio of  $240 \pm 25$  between 70 and 95 km was reported by Fedorova et al. (2008). Altitude profiles of volume mixing ratios of minor gases are given by Haus and Arnold (2010) and are based on the profiles described by Pollack et al. (1993). Monochromatic absorption cross-sections of gases due to allowed molecular transitions are computed from the spectral line databases CDS (CO<sub>2</sub>, 0.4–0.9  $\mu\text{m}$  and 1.25–1.33  $\mu\text{m}$ , Tashkun et al. (2003)), HITEMP (CO<sub>2</sub>, Pollack et al. (1993), CO, H<sub>2</sub>O isotopologues 161, 181, 171, Rothman et al. (1995)), and HITRAN08 (H<sub>2</sub>O isotopologues 162, 182, 172, SO<sub>2</sub>, OCS, HCl, HF, Rothman et al. (2009)) by using spectral line shapes listed by Haus and Arnold (2010). To be in line with Haus and Arnold (2010), the more recent HITEMP2010 (Rothman et al., 2010) is not yet considered here. Molecular Rayleigh scattering is treated according to Hansen and Travis (1974). Spatially and temporally varying non-LTE O<sub>2</sub> emissions ( $'\text{O}_2$  nightglow') at 1.27  $\mu\text{m}$  from an altitude region around 100 km (Piccioni et al., 2009) contaminate the 1.28  $\mu\text{m}$  window (1.225–1.295  $\mu\text{m}$ ). For retrievals, this window will thus be blacked out, and O<sub>2</sub> nightglow need not be considered. The clouds are modeled according to Haus et al. (2013). As they form a complex source of emissivity retrieval errors, details are presented in Section 3.5. Wavelength dependent scattering and absorption properties of the clouds are computed according to Mie theory (Wiscombe, 1980).

Resembling an argumentation by Kappel (2014) for a similar situation, from the total VIRTIS-M-IR spectral range (1.0–5.1  $\mu\text{m}$ ), only 1.0–2.3  $\mu\text{m}$  shall be utilized for this study. Venus' nightside emissions in this range mainly originate from altitudes below 40 km (Haus and Arnold, 2010, Fig. 4), where temperature is quite stable with time, and they are thus nearly unaffected by the strong mesospheric temperature variations above 58 km that were observed by Haus et al. (2013, 2014); Tellmann et al. (2009). Also, details of the

$t$	Description	$\bar{t}_t$	$2\sigma_t$	$I_t$
$mi$	Cloud mode $i$ column factor	1.0	0.5	0.14
$\Delta t1$	Var. of mode 1 top alt.	0 km	3 km	0.85 km
$\Delta t2'$	Var. of mode 2' top alt.	0 km	3 km	0.85 km
$\Delta t3$	Var. of mode 3 top alt.	0 km	3 km	0.85 km
$\Delta B$	Var. of modes 1+2'+3 base alt.	0 km	3 km	0.11 km
$\Delta t2$	Translation of mode 2 profile	0 km	3 km	0.85 km
$H_2SO_4$	$H_2SO_4$ conc. of cloud droplets	80.0%	7.5%	2.1%
$H_2O$	$H_2O$ column factor	1.0	0.25	0.018
$HCl$	$HCl$ column factor	1.0	0.25	0.018
$\Delta h$	Variation of surface elevation	0 km	0.2 km	0.4 km
$ei$	Surface emissivity, window $i$	0.65	0.15	0.3
$\Delta T0$	Var. of temp. profile near 0 km	0 K	3 K	0.11 K
$\Delta T25$	Var. of temp. profile near 25 km	0 K	3 K	0.11 K
$ci$	Continuum factor, window $i$	1.0	0.25	0

Table 1: Summary of discussed external parameters  $t$ . Listed are each parameter’s expected mean value  $\bar{t}_t$ , double standard deviation  $2\sigma_t$ , and half length of the sliding interval  $I_t$  for the determination of relative error measures, see Section 2. The three kinds of parameters are grouped by the horizontal lines. The first group comprises *Atm* parameters, the second *Spa*, and the third *Cnt* parameters. Abbreviations: Var. - Variation, alt. - altitude, conc. - concentration, temp. - temperature.

cloud altitude distribution have a minimal impact here, since the main cloud deck ( $\geq 48$  km, Marov et al. (1980)) resides above the line forming altitude region. In contrast, spectral signatures longward of  $3 \mu\text{m}$  are strongly influenced by variations of temperature and cloud altitude distributions above 48 km. The range  $2.3\text{--}2.6 \mu\text{m}$  is excluded in contrast to Kappel (2014), since the number of interfering parameters shall be kept as low as possible for this study, 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  $2.6\text{--}3 \mu\text{m}$  are too small to be useful (Haus and Arnold, 2010, Fig. 2). Note that calibration refinements and preprocessing for VIRTIS-M-IR data as performed by Kappel et al. (2012c) had mainly concentrated on the interval  $1.0\text{--}2.6 \mu\text{m}$ .

To be in line with the typical characteristics of VIRTIS-M-IR, the full-width-at-half-maximum (FWHM) of the instrumental response function is set to 17 nm, spectral sampling to 9.49 nm, first spectral band is at  $1.02 \mu\text{m}$ , corresponding to  $m = 129$  wavelength bands for the range  $1.0\text{--}2.3 \mu\text{m}$  excluding the  $1.28 \mu\text{m}$  peak. For the default reference spectrum, the observation angle is set to zero (nadir-looking).

### 3.2. Surface emissivity

The Jacobians in Fig. 2 illustrate that in the considered spectral range, spectral surface emissivity is observable in three peaks ( $1.02$ ,  $1.10$ , and  $1.18 \mu\text{m}$ ), which shall be called ‘surface window peaks’. The  $1.28 \mu\text{m}$  peak is also slightly affected by surface emissivity but is blacked out here. This gives rise to three parameters of primary interest,  $\mathbf{e} =: (e_{1.02}, e_{1.10}, e_{1.18})$ . They parameterize the spectral emissivity such that it is constant over the ranges of the respective windows ( $1.02\text{--}1.055 \mu\text{m}$ ,  $1.055\text{--}1.125 \mu\text{m}$ ,  $1.125\text{--}1.225 \mu\text{m}$ , 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 default reference values are all set to 0.65.

In Section 4.5, emissivities themselves are partially treated as external parameters. For this purpose, it is assumed that for most surface bins, the three corresponding emissivities should not differ much among each other, and the expected double standard deviations are set to  $2\sigma_{ei} = 0.15$ . With the expected mean values  $\bar{t}_{ei} = 0.65$ , the *CR* follow as  $[0.5, 0.8]$ , corresponding to a variation of 23%. Neglecting geologic activity, emissivities have the character of a *Spa* parameter (and thus have infinite correlation times), and the *a priori* correlation lengths can be set to 50 km (half FWHM of the surface data blurring). According to Eq. (3), the correlation coefficient at 100 km footprint separation follows as  $C_{12} = 0.0059$ . Thus, the approximation of  $I_{t_A}$  (Eq. (5)) for the case  $C_{12} \ll 1$  can be applied, and  $I_{ei}$  can be set to 0.3, see paragraph subsequent to Eq. (7). Recall that the stated information is summarized in Table 1.

### 3.3. Deep atmospheric temperature field

A potential source of emissivity retrieval errors is the deep atmospheric temperature field. 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 (Lebonnois et al., 2010, GCM, 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.

Only the equatorial VIRA profile shall be considered here as default reference, and modifications to the VIRA deep atmospheric temperature altitude profile below 50 km are parameterized by the two external parameters  $t_{\Delta T0}$  and  $t_{\Delta T25}$ , see Fig. 3.  $t_{\Delta T0}$  leads to a quadratic offset to VIRA that is zero and has zero

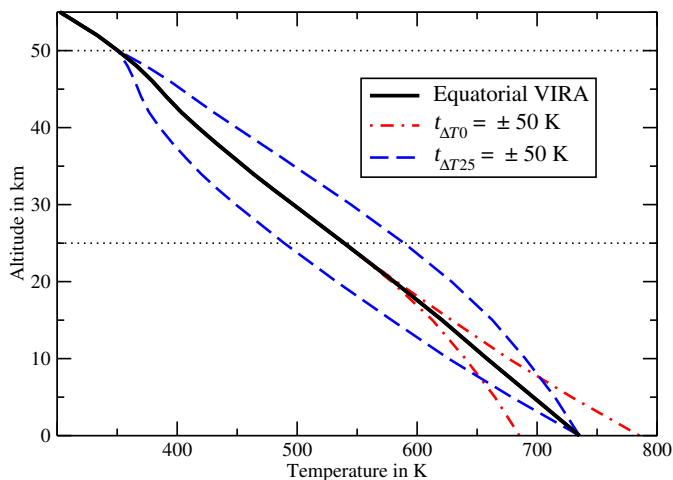


Figure 3: Modified deep atmospheric temperature altitude profiles for values of  $t_{\Delta T0}$  and  $t_{\Delta T25}$  that are set to exaggerated values for better representation.

derivative at 25 km and is  $t_{\Delta T0}$  at 0 km, regardless of the actual surface elevation.  $t_{\Delta T25}$  leads to a quadratic offset that is zero at 0 and at 50 km and is  $t_{\Delta T25}$  at 25 km. These crude modifications allow to change the temperature in the bulk of the deep atmosphere, and temperature and temperature lapse rate near the surface.  $t_{\Delta T0}$  and  $t_{\Delta T25}$  are assumed to have expected mean values  $\bar{t}_{\Delta T0} = 0 \text{ K} = \bar{t}_{\Delta T25}$ . Due to the extremely high atmospheric thermal inertia close to the surface, expected double standard deviations are conceivable to be in the order of  $2\sigma_{\Delta T0} = 3 \text{ K} = 2\sigma_{\Delta T25}$  and correlation lengths in the order of  $\lambda_{\Delta T0} = 4000 \text{ km} = \lambda_{\Delta T25}$ , compare also tropospheric VIRA data above 32 km. The stated mean values correspond to the default reference VIRA temperatures 735.3, 539.2, and 337.0 K at 0, 25, and 50 km, respectively. The stated correlation length corresponds to about 38° latitude difference at the surface and translates to a correlation coefficient  $C_{12} = 0.99932$  at 100 km footprint separation. Thus, the approximation of  $I_{t_A}$  (Eq. (5)) for the case  $C_{12}$  close to 1 can be applied, and it follows that  $I_{\Delta T0} = 0.11 \text{ K} = I_{\Delta T25}$  (Eq. (7)).

### 3.4. Minor gases

The only minor gases that provide significant opacity in the studied spectral range (1.0–2.3  $\mu\text{m}$  excluding the 1.28  $\mu\text{m}$  peak) are  $\text{H}_2\text{O}$  (including HDO) and HCl (Haus and Arnold, 2010, Fig. 19). CO influence at the interval boundary (2.28–2.3  $\mu\text{m}$ ) is neglected here. Both  $\text{H}_2\text{O}$  and HCl affect the peak at 1.74  $\mu\text{m}$ , and  $\text{H}_2\text{O}$  is in addition observable between 1.10 and 1.20  $\mu\text{m}$  (Fig. 2). The profiles from Haus and Arnold (2010, Fig. 6) are used as default reference altitude profiles, but the  $\text{H}_2\text{O}$  column is here scaled such that the concentration at the surface is 25 instead of 32.5 ppmv. The probed altitude regions are located below 25 km (1.10–1.20  $\mu\text{m}$ ) and between 20 and 30 km (1.74  $\mu\text{m}$ ).  $\text{H}_2\text{O}$  (excluding HDO) and HCl variations shall be parameterized by just total column factors  $t_{\text{H}_2\text{O}}$  and  $t_{\text{HCl}}$  with respect to their default reference altitude

profiles that consequently correspond to factors of  $t_{\text{H}_2\text{O}} = 1.0 = t_{\text{HCl}}$ , translating to 25 or 0.5 ppmv at 25 km, respectively. Recent studies showed that  $\text{H}_2\text{O}$  concentration in the deep atmosphere amounts to a value of 30 ppmv (Bézar *et al.*, 2011), corresponding to  $t_{\text{H}_2\text{O}} = 1.2$ . This study considers  $t_{\text{H}_2\text{O}}$  and  $t_{\text{HCl}}$  as external parameters with expected mean values  $\bar{t}_{\text{H}_2\text{O}} = 1.0 = \bar{t}_{\text{HCl}}$ . Expected double standard deviations are set to  $2\sigma_{\text{H}_2\text{O}} = 0.25 = 2\sigma_{\text{HCl}}$  and correlation lengths to  $\lambda_{\text{H}_2\text{O}} = 2000 \text{ km} = \lambda_{\text{HCl}}$ . This can be estimated by observing the spatial variation in the results by Tsang *et al.* (2009) (for CO) and Haus *et al.* (2015) (for CO and  $\text{H}_2\text{O}$ ) and is here carried over to other minor gases as first estimate, with the atmospheric convection and its spatial scales as motivation. This leads to  $C_{12} = 0.9973$  (Eq. (3)) and to  $I_{\text{H}_2\text{O}} = 0.018 = I_{\text{HCl}}$  (Eq. (7)).

### 3.5. Clouds

The clouds of Venus are usually assumed to comprise four modes (1, 2, 2', and 3). Here, the modes 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  $\mu\text{m}$  and unitless dispersions of 1.56, 1.29, 1.23, 1.28 for the four modes, respectively (Pollack *et al.*, 1993). Default reference profiles of particle number densities  $N$  at altitudes  $h$  are taken from Haus *et al.* (2013) and are given by

$$N(h) = \begin{cases} N_0 \exp(- (h - h_t)/s_t) & h > h_t \\ N_0 & h_b \leq h \leq h_t \\ N_0 \exp(- (h_b - h)/s_b) & h < h_b \end{cases} \quad (9)$$

Mode-specific values are listed in Table 2.  $N(h)$  is set to zero above 85 km. Actual cloud modal abundances

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 $\text{cm}^{-3}$	181	100	50	14

Table 2: Cloud parameters for default reference model as given by Eq. (9). PND is the particle number density.

are defined by cloud mode factors  $m_1$ ,  $m_2$ ,  $m_{2'}$ , and  $m_3$  that scale the number densities of the default reference altitude profiles that consequently correspond to cloud mode factors of  $m_i = 1.0$ . Some of the  $m_i$  will be parameterized through auxiliary parameters  $z_{mi}$ , while the remaining will be treated as external parameters  $t_{mi}$ . Variations of  $h_b$ ,  $h_t$ ,  $s_b$ , or  $s_t$  for cloud mode  $i$  rescales  $N_0$  for mode  $i$  such that the total number of mode  $i$  particles remains constant. This is done in order to disentangle variations of altitude distribution and variations of particle abundance.

The influences of variations of the mode-common  $\text{H}_2\text{SO}_4$  concentration  $t_{\text{H}_2\text{SO}_4}$  of the cloud droplets and of the mode-specific cloud mode factors  $m_i$  on surface emissivity retrieval results are investigated in this work. There are studies like that by Barstow *et al.* (2012) that use or derive values of  $t_{\text{H}_2\text{SO}_4}$  of locally up to 95%, which is much higher than the default reference value 75% and values used by other authors. However, the expected mean value  $\bar{t}_{\text{H}_2\text{SO}_4}$  is set to 80% and double standard deviation to  $2\sigma_{\text{H}_2\text{SO}_4} = 7.5\%$ . The  $m_i$  may vary strongly, but no detailed knowledge for all modes is available. Expected mean values are all set to 1.0 and double standard deviations to 0.5. Also, some crucial altitude profile variations are studied as external parameters. For physical reasons, deviations of the  $h_b$  for modes 1, 2', and 3 from their default reference value 49 km are treated as a joint parameter  $t_{\Delta B}$ , 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 sole considered mode 2 profile variation is modeled to be just a vertical translation by the value  $t_{\Delta t_2}$ . This is sufficient, because mode 2 particles are located well above the line forming altitude region (surface to 40 km) in the considered spectral range. Finally,  $h_t$  variations for modes 1, 2', and 3 are studied individually. Deviations

from their default reference values are denoted by  $t_{\Delta t1}$ ,  $t_{\Delta t2}$ , and  $t_{\Delta t3}$ , respectively.  $t_{\Delta B}$  and the  $t_{\Delta ti}$  are all assumed to have expected mean values 0 km and double standard deviations 3 km, compare (Barstow et al., 2012; Haus et al., 2013, 2014; Ignatiev et al., 2009) to get an impression of expectable variations of nightside cloud bases, tops, and dayside cloud tops, respectively. Any scale height variations are not considered, as they are of minor importance in the utilized spectral range, compared to  $h_b$  and  $h_t$  variations.

The expected correlation length for the total cloud opacity is 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  $2.3 \mu\text{m}$  nightside radiances as proxy. But note that there is also evidence for smaller scale cloud structures (100–200 km for mottled clouds observed by Venus Express VMC (Titov et al., 2012)), and that small-scale convection in the middle cloud region causes large optical thickness variations (Imamura and Hashimoto, 2001). The correlation length 500 km is carried over to all other cloud parameters, except for  $t_{\Delta B}$ , where it is set to a value that is in accordance with the deep atmospheric temperature correlation length of 4000 km. The reason for this is the mentioned temperature related sharp decrease of cloud particle numbers at  $t_{\Delta B} = 0$  km.

For  $t_{\Delta B}$ , it follows that  $C_{12} = 0.99932$  (Eq. (3)) and  $I_{\Delta B} = 0.11$  km (Eq. (7)).  $t_{\Delta ti}$ ,  $t_{mi}$ , and  $t_{\text{H}_2\text{SO}_4}$  all have  $C_{12} = 0.9594$  and thus  $I_{\Delta ti} = 0.85$  km,  $I_{mi} = 0.14$ , and  $I_{\text{H}_2\text{SO}_4} = 2.1\%$ .

### 3.6. $\text{CO}_2$ opacity correction

The atmospheric composition and the physical properties of the atmospheric gases and clouds are not sufficiently well constrained to provide reliable simulations of observable spectra. Especially, the opacity of Venus’ atmospheric main gaseous constituent  $\text{CO}_2$  is not well predictable for the high-pressure and high-temperature environment of Venus’ deep atmosphere. 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; Bézard et al., 2011). The line shapes are still under discussion, as are additional opacity effects (continuum, collisional induced absorption, line mixing). The  $\text{CO}_2$  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’. It depends on the utilized line databases and line shapes but is independent of the measurement. For this study, the continuum is treated as spectrally constant throughout the range of an atmospheric transparency window, but it can depend on the window. These window-specific values are denoted by  $c_{1.02}$ ,  $c_{1.10}$ ,  $c_{1.18}$ ,  $c_{1.28}$ ,  $c_{1.31}$ ,  $c_{1.74}$ , and  $c_{2.3}$ , and their default reference values in this study are set to 0.5, 3, 2, 7, 13, 27, 160 in units of  $10^{-29} \text{ cm}^2$ , respectively.

These values are inspired by preliminary results of retrieving them with MSR (Kappel, 2014) as parameters that are common to a large set of highly diverse VIRTIS-M-IR nightside spectra, but they differ from the values used in the cited paper. With the exception of  $c_{2.3}$ , the continua are difficult to retrieve, even with MSR, since modifications to the retrieval pipeline and the choice of the set of utilized spectra and an initial surface emissivity can (moderately) affect the results. Recall that the here used continuum parameters do not only include the actual continuum contribution, but opacity corrections in general. In addition, they are not well-defined without at the same time stating the utilized line databases and line shapes and are thus not directly comparable to values given by authors who made other choices. Furthermore, a retrieval error analysis analogous to the present one would be needed to understand the meaning of such retrieved opacity corrections. Therefore, values stated in earlier publications of the present authors should be, as mentioned in those works, regarded as working hypotheses and preliminary values. For better handling in this work, the above stated values are only abstractions of the latest retrieved continua, the latter allowing for extremely well fits to real spectra for a very high diversity of atmospheric and observational conditions. The stated values are rounded and do not include contributions that are not constant throughout the respective window’s range. Note that  $c_{2.3}$  is here the optimal value for the range  $2.1\text{--}2.3 \mu\text{m}$  and differs from the optimal value for the range of the entire  $2.3 \mu\text{m}$  window. It will be shown in Section 4 that uncertainties in the continua are important sources of retrieval errors when absolute emissivities are retrieved. To avoid the high emissivity retrieval errors implied by the mentioned difficulties, the strategy for the emissivity retrieval will be modified by rather retrieving emissivities relative to an initial emissivity and showing the results to be reasonably independent of the continua (and the initial emissivity).

For the emissivity retrieval analysis, the stated continua are varied according to the external parameters  $t_{c1.02}$ ,  $t_{c1.10}$ , etc., such that  $c_{1.02} \cdot t_{c1.02}$  is the perturbed continuum in the  $1.02 \mu\text{m}$  window, and so on. The

expected mean values and double standard deviations are  $\bar{t}_{ci} = 1.0$  and  $2\sigma_{ci} = 0.25$ , respectively. The correlation length of any continuum parameter is necessarily infinite, because it does not spatially vary, see also discussion by Kappel (2014). Thus, there are no relative spatial errors ( $RL_i = 0$ ). This is compatible with Eqs. (5) and (6) in the limit where  $C_{12}$  approaches 1 (Eq. (3)).

### 3.7. Surface elevation

The observed radiance depends on the surface elevation of a measurement’s footprint. With increasing surface elevation, on the one hand, surface temperature decreases, which leads to lower emissions from the surface. On the other hand, the atmospheric path the emitted radiance has to travel, shortens, leading to less absorption and scattering. For each measurement footprint, surface elevation as fixed input to the radiance simulation is provided according to Kappel et al. (2012c), who use Magellan radar data (Ford and Pettengill, 1992; Rappaport et al., 1999) as basis. Magellan data itself is not perfect and a source of error, as unrealistic small-scale discontinuities in the raw data clearly show. Multiple scattering by gases and clouds cause atmospheric blurring of observable surface information (FWHM 100 km, (Moroz, 2002)). This is taken into account, as is motion blurring caused by the detector’s movement relative to the footprint, but uncertainty in the knowledge of the exact blurring function for instance in dependence on cloud thickness is an additional source of error. Since not the topography itself has to be blurred, but the observed fundamental carrier of the data, the top-of-atmosphere radiance, an effective radiance computed from topographic data is blurred and transformed back to get the blurred surface topography. This procedure introduces further topography uncertainties, since it slightly depends for instance on assumptions on atmospheric gases and clouds, observation angle, and wavelength. Finally, the exact location and size of a measurement’s surface footprint depend on the cloud top altitude, as this is roughly the effective source of observed emissions, because scattering strongly decreases only at even higher altitudes. This is also an error source, because this is taken into account for just an initial guess of the cloud top altitude in dependence on latitude based on results by Ignatiev et al. (2009).

The default reference surface elevation is set to 0 km. The deviation of the surface elevation utilized in the radiance simulations from the ‘true’ surface elevation is parameterized by the external parameter  $t_{\Delta h}$  and is expected to amount to the order of 100 m in absolute value for planetary plains and even more for mountainous regions. This is compatible with the order of magnitude of the difference between the 100 and the 65 km FWHM atmospheric blurring of the topography as simple proxy for surface elevation uncertainties. Expected mean value and double standard deviation are set to  $\bar{t}_{\Delta h} = 0$  km and  $2\sigma_{\Delta h} = 0.2$  km, respectively. The *a priori* correlation length is 50 km (half FWHM of atmospheric blurring), such that  $C_{12} = 0.0059$ . Thus, the approximation of  $I_{t_A}$  (Eq. (5)) for the case  $C_{12} \ll 1$  can be applied, and  $I_{\Delta h}$  can be set to 0.4 km. Recall that Table 1 summarizes the data for all discussed external parameters.

### 3.8. Reference spectra

Retrieval errors may depend on environmental and observational conditions. For instance, for all other parameters unchanged, observations of targets with few clouds should lead to smaller surface emissivity retrieval errors than observations of very cloudy regions. Thus, along with the default reference model, additional scenarios shall be investigated, which cover important key situations that may affect the retrieval errors. The scenarios are as follows (also see Fig. 4).

1. Default reference model.
2. 50% larger abundances of cloud modes 2’ and 3.
3. Surface elevation is set to 5000 m.
4. Surface emissivity is set to 0.9.
5. Observation angle is set to 60 degrees.

These descriptions always state the difference to the default reference model. The Euclidean norms  $\|y\|_2$  for the five scenarios in the studied spectral range (1.0–2.3  $\mu\text{m}$  excluding the 1.28  $\mu\text{m}$  peak) are 0.257, 0.137, 0.225, 0.261, and 0.173  $\text{W}/(\text{m}^2 \text{sr} \mu\text{m})$ , compare paragraph following Eq. (8) in Section 2. Information losses due to random measurement errors are emulated by adding Gaussian noise with standard deviation  $\sigma=10^{-4} \text{W}/(\text{m}^2$

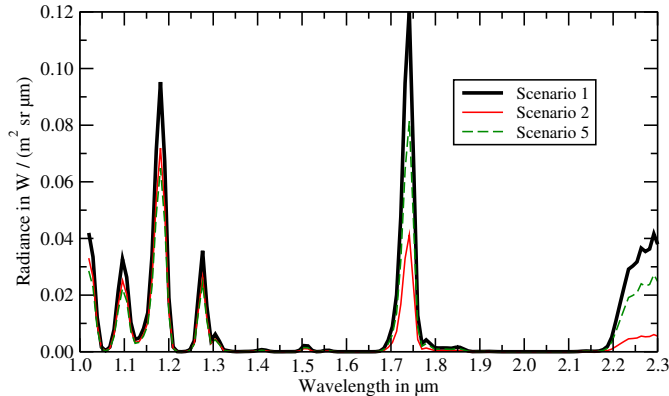


Figure 4: Synthetic reference spectra for scenarios 1, 2, and 5. In comparison with scenario 1 (peak radiances at 1.02, 1.10, 1.18, and 1.28  $\mu\text{m}$  are 0.0420, 0.0327, 0.0952, 0.0357 in units of  $\text{W}/(\text{m}^2 \text{sr } \mu\text{m})$ ), radiances for scenarios 3 (0.0176, 0.0204, 0.0605, 0.0332) and 4 (0.0462, 0.0339, 0.0989, 0.0357) significantly differ only shortward of 1.3  $\mu\text{m}$  and are not depicted.

$\text{sr } \mu\text{m})$  to the respective reference spectrum. For each retrieval run, noise is newly generated. As it can be seen from deep space observations, 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  $\mu\text{m}$  (256-sample-mode, unbinned, 3 s exposure time). Thus,  $\sigma$  is rather low, but it is a good description for measurements that are binned to match the expected achievable surface resolution (100 km FWHM, equal-area binning (Kappel et al., 2012c) leading to typically 30 spectra per bin for southern hemisphere mapping data at 45  $^{\circ}\text{S}$ ). Finally, note that for the emissivity retrieval error analysis, *a priori* mean values and initial guesses for emissivities and cloud mode factors are intentionally set to values (0.5) that are different from their 'true' values, and their double *a priori* standard deviations are set to values (5 and 10, respectively) that are higher than that for operational retrievals, in order to not bind the parameters to their true values from the outset. Dependence on the *a priori* mean values is studied in Section 4.

## 4. Results

### 4.1. Spectral ranges and auxiliary parameters

Modal abundances of Venus' clouds are parameterized by the cloud mode factors  $m_1$ ,  $m_2$ ,  $m_2'$ , and  $m_3$  (Section 3.5). Due to their similar Jacobians, they are difficult to disentangle, and there is almost no *a priori* information known. Also, they are strongly varying with large overall impacts on nightside spectra (Figs. 2 and 4). While some of the  $m_i$  must thus be retrieved in any case to allow for reasonable fits at all, this already suffices to produce quite good overall fits between 1.0 and 2.3  $\mu\text{m}$ , whereas, technically speaking, surface emissivity is needed to just refine the quality of fits. The auxiliary parameters are selected from the  $m_i$ , therefore. Since they are difficult to disentangle, and it may not be possible to retrieve all of them, they will be just treated as auxiliary parameterization of how the cloud modes can possibly affect the shape of the spectra in the frame of a reasonable cloud model. They are not regarded as parameters with well-defined physical meaning, but may still indicate certain trends in the effective cloud particle size (Haus et al., 2013).

It was explained in Section 3.1 that this study is restricted to the range 1.0–2.3  $\mu\text{m}$  excluding the 1.28  $\mu\text{m}$  window. This range comprises the surface windows at 1.02  $\mu\text{m}$  (1.02–1.055  $\mu\text{m}$ ), 1.10  $\mu\text{m}$  (1.055–1.125  $\mu\text{m}$ ), and 1.18  $\mu\text{m}$  (1.125–1.225  $\mu\text{m}$ ) and the deep atmospheric windows at 1.31  $\mu\text{m}$  (1.295–1.32  $\mu\text{m}$ ), 1.74  $\mu\text{m}$  (1.65–1.9  $\mu\text{m}$ ), and 2.3  $\mu\text{m}$  (short wavelength flank, 2.15–2.30  $\mu\text{m}$ ), see Figs. 2 and 4. The parameters of primary interest, the emissivities, can only be retrieved from the surface windows. Different retrieval pipelines can be applied to retrieve the auxiliary parameters along with the emissivities. Each pipeline is specified by a choice of  $m_i$  parameters and the spectral range where these  $m_i$  are retrieved from. For several different pipelines it is investigated, which one leads to the smallest emissivity retrieval errors. Only results for this



pipeline are discussed in detail. The choice of the best pipeline should not strongly depend on details of the reference spectrum. Thus, only the default reference spectrum is considered for this purpose.

#### 4.2. Total cloud factor from 1.31 $\mu\text{m}$ peak

The first pipeline retrieves the total cloud factor  $z_{tc}$  from the 1.31  $\mu\text{m}$  radiance peak and then the surface emissivities.  $z_{tc}$  simultaneously changes all four cloud mode factors in the same way, i.e.  $m_i := z_{tc}t_{mi}$  for all external cloud mode scaling parameters  $t_{mi}$ . This pipeline emulates the approach by Müller et al. (2008) on the basis of a full radiative transfer simulation. Müller et al. (2008) developed a semi-empirical method to estimate the so-called thermal flux anomaly at 1.02  $\mu\text{m}$ . This is the thermal flux from Venus’ surface that is not due to topography related surface temperature. It is correlated to surface emissivity. Along with other semi-empirical data preprocessings, the authors de-cloud the 1.02  $\mu\text{m}$  data with the surface-unaffected 1.31  $\mu\text{m}$  peak as reference for cloud thickness. In doing so, they implicitly assume that there is no cloud induced color dispersion between 1.02 and 1.31  $\mu\text{m}$  (‘clouds are gray’). Their method has the advantage of being computationally very fast, and the authors can utilize all suitable VIRTIS-M-IR nightside spectra to generate a southern-hemisphere map of the time-averaged thermal flux anomaly. However, they utilize only two spectral bands, and especially the small 1.31  $\mu\text{m}$  radiance is very susceptible to measurement noise and data calibration and preprocessing errors for instance due to imperfect straylight removal (Kappel et al., 2012c). Note that the 1.31  $\mu\text{m}$  peak is sometimes affected by 1.27  $\mu\text{m}$  O<sub>2</sub> nightglow. In contrast, Haus and Arnold (2010); Kappel et al. (2012c); Kappel (2014) apply a detailed radiative transfer simulation model and a retrieval algorithm and utilize the full spectral range 1.0–2.3  $\mu\text{m}$  excluding the 1.28  $\mu\text{m}$  peak to quantitatively determine surface emissivity. In addition, this approach enables a retrieval error analysis (the present paper). But the required computational resources are considerably larger, and only a selection of spectra can be studied that covers targets that were beforehand identified to be of special geological interest, for instance by using the results by Müller et al. (2008); Smrekar et al. (2010), or geologic maps (Ivanov and Head, 2006) derived from Magellan radar data (Pettengill et al., 1991; Ford and Pettengill, 1992), or the radiance ratio based quick-look method by Arnold et al. (2008). Note that always the full ranges of the respective spectral windows are utilized here and in the following. Exemplary tests show that noise impact on retrieved emissivities is 10–20% higher, when for each window only the single spectral band with the maximum radiance value is utilized, as should be done to achieve a direct comparison to Müller et al. (2008).

The emissivities  $e_i$  ( $e_{1.02}$ ,  $e_{1.10}$ , and  $e_{1.18}$ ) are now retrieved for the default reference model (scenario 1, Section 3.8) and for 100 noise repetitions. The global error measures  $\overline{RG}_i$ , color dispersions  $\overline{CD}$ , relative error measures  $\overline{RL}_i$  (noisy versions of Eqs. (1), (2), (6)), and noise measures  $N_i$  (Eq. (8)) can then be computed. The summarized error measures follow immediately. Results are listed in Table 3, where the values are rounded to the nearest whole percent. Note that  $t_{\text{HCl}}$ ,  $t_{c1.74}$ , and  $t_{c2.3}$  do not affect the relevant radiance peaks here. Also, emissivities need not be considered as external parameters. Note that the relative error measures for continuum parameters are zero by construction (Section 3.6, indicated by ‘/’ in the table).

The error measures are quite large and much worse at 1.10 and 1.18  $\mu\text{m}$  than at 1.02  $\mu\text{m}$ . Global and  $\overline{CD}$  *Atm* errors are dominated by  $t_{\text{H}_2\text{O}}$  at 1.10 and 1.18  $\mu\text{m}$ , a consequence of the direct influence of H<sub>2</sub>O in these peaks (Fig. 2).  $t_{\Delta T0}$  and  $t_{\Delta T25}$  errors dominate the global and  $\overline{CD}$  *Spa* error measures, but not the relative ones that are dominated by surface elevation uncertainties. Note that relative errors due to  $t_{\Delta h}$  uncertainties are very similar to the corresponding global errors. This is in accordance with the expectation that over a distance (100 km for definition of relative errors) that exceeds the correlation length (50 km for  $t_{\Delta h}$ ), relative spatial errors should not provide other information than the global errors already do. This also justifies the factor 1/2 in definition of the relative error measures (Eq. (6)). *Cnt* errors are worst at 1.10  $\mu\text{m}$ . Noise impact is also quite large, and at 1.10 and 1.18  $\mu\text{m}$ , it is three times as large as at 1.02  $\mu\text{m}$ . This is mainly a consequence of the lower surface information content of those peaks, not just of the signal-to-noise ratio, which is for the 1.18  $\mu\text{m}$  peak twice as large as for 1.10  $\mu\text{m}$  (Fig. 4). The summarized error and noise measures show that single-spectrum retrieval of emissivities is not possible for this retrieval pipeline (but note that Müller et al. (2008) average over many spectra). Still, apart from impacts due to surface elevation uncertainties, spatial fine structure can be determined quite reliably as can be seen from the  $\overline{RL}_i$ -columns.

While the utilization of a wider spectral range than here allows to roughly estimate continuum parameters by applying MSR (shown by Kappel (2014)), this is not possible when only including the 1.31  $\mu\text{m}$  radiance

$t$	$\overline{RG}_i$			$\overline{CD}$	$\overline{RL}_i$		
$m_1$	4	<b>12</b>	8	8	1	<b>2</b>	1
$m_2$	4	<b>14</b>	10	8	1	<b>2</b>	2
$m_{2'}$	3	<b>4</b>	2	2	0	<b>1</b>	0
$m_3$	8	<b>25</b>	16	21	1	<b>5</b>	3
$\Delta t_1$	1	<b>2</b>	<b>2</b>	2	0	<b>1</b>	<b>1</b>
$\Delta t_{2'}$	1	<b>3</b>	3	2	0	<b>1</b>	1
$\Delta t_3$	0	<b>1</b>	1	1	0	<b>1</b>	1
$\Delta B$	7	20	<b>22</b>	17	0	<b>1</b>	<b>1</b>
$\Delta t_2$	1	<b>3</b>	2	2	0	<b>2</b>	2
$\text{H}_2\text{SO}_4$	1	<b>7</b>	5	9	0	<b>1</b>	1
$\text{H}_2\text{O}$	5	66	<b>72</b>	115	0	7	<b>8</b>
$\Delta h$	12	<b>18</b>	16	7	12	<b>20</b>	18
$\Delta T_0$	27	<b>51</b>	46	48	0	<b>2</b>	2
$\Delta T_{25}$	18	36	<b>43</b>	32	0	1	<b>1</b>
$c_{1.02}$	<b>18</b>	1	1	17		/	
$c_{1.10}$	0	<b>86</b>	1	150		/	
$c_{1.18}$	1	2	<b>61</b>	88		/	
$c_{1.31}$	26	66	<b>70</b>	92		/	
$\langle \text{Atm} \rangle$	14	76	<b>78</b>	119	2	9	<b>9</b>
$\langle \text{Spa} \rangle$	35	65	<b>65</b>	58	12	<b>20</b>	18
$\langle \text{Cnt} \rangle$	32	<b>108</b>	93	198		/	
$N_i$	7	<b>24</b>	23				

Table 3: Surface emissivity retrieval error and noise measures in percent, when the total cloud factor is retrieved as auxiliary parameter from the 1.31  $\mu\text{m}$  peak. Scenario 1, 100 noise repetitions. The three kinds of parameters are grouped by the horizontal lines. The first group comprises *Atm* parameters, the second *Spa*, and the third *Cnt* parameters. The corresponding summarized error measures are given below the double horizontal line. The noise measures  $N_{1.02}$ ,  $N_{1.10}$ , and  $N_{1.18}$  are given in the last row in the respective  $\overline{RG}_i$  columns. The largest of the values at the different wavelengths is always typed in bold.

peak in addition to the surface window peaks. Also, for real spectra, the FWHM of the VIRTIS-M-IR instrumental response function varies with detector temperature and location on detector and is currently not sufficiently well predictable in the calibration procedure, as is the spectral band-to-wavelength mapping (Cardesin-Moinelo et al., 2010; Kappel et al., 2012c). By utilizing the spectral range 1.0–2.3  $\mu\text{m}$ , it is possible to retrieve these instrumental parameters from the spectra themselves (shown for FWHM by Kappel (2014)), which is not well possible with the present pipeline.

#### 4.3. Three cloud modes from 1.295–2.300 $\mu\text{m}$

A wider spectral range shall now be utilized. This also increases the available spectral information content and allows to retrieve more auxiliary parameters. When using the full spectral range considered in this work, information can be retrieved from six radiance peaks at most, three of which are carrying surface information. It is thus natural to assume that at most three auxiliary parameters can be retrieved. These comprise three of the four cloud mode factors  $m_i$ , and the remaining one has to be treated as external parameter.

When three cloud mode factors are retrieved from the 1.31, 1.74, and 2.3  $\mu\text{m}$  peaks and then the emissivities, it turns out that the lowest overall error measures for emissivity retrieval are achieved when  $m_1$  or  $m_2$  are external parameters, whereas the lowest noise measures result for the external parameters  $m_{2'}$  or  $m_3$ . However, the error and noise measures are in all these cases not better and for some external parameters even slightly worse than for the pipeline in Section 4.2. The reason is that three auxiliary parameters with Jacobians that are very similar longward of 1.295  $\mu\text{m}$  (Fig. 2) are retrieved from three radiance peaks, where one of the peaks is small (1.31  $\mu\text{m}$  peak radiance  $\approx 10^{-2}$  W/(m<sup>2</sup> sr  $\mu\text{m}$ ), Fig. 2) and has a small signal-to-noise ratio, compare also Kappel (2014, Section 5). In the best pipeline (external  $m_2$ ), the retrieved cloud mode factors  $m_1$ ,  $m_{2'}$ , and  $m_3$  typically scatter with double standard deviations of respectively 1.6 (!), 0.4, and 0.1 for the 100 noise repetitions, indicating a bad disentanglement and leading to the high emissivity errors.

#### 4.4. Cloud modes 2' and 3 as auxiliary parameters

To overcome this problem, only two cloud mode factors are now retrieved as auxiliary parameters from the range 1.295–2.300  $\mu\text{m}$  and then the emissivities. The lowest error and noise measures result, when both

$m_1$  and  $m_2$  are treated as external, and  $m_{2'}$  and  $m_3$  are retrieved as auxiliary parameters. However, when for the same parameter setup only the radiance peaks at 1.74 and 2.3  $\mu\text{m}$  are utilized, results are virtually identical: nearly all error and noise measures differ by 1 percentage point at most, and differences are evenly distributed in both directions, indicating a negligible information loss when discarding the 1.31  $\mu\text{m}$  peak.

The most important difference to the results in Sections 4.2 and 4.3 is that the resulting noise measures are now much smaller (by 80%–90%). This can be explained by the much smaller scatter of the retrieved cloud mode factors  $m_{2'}$  (double standard deviation of scatter 0.01) and  $m_3$  (0.005) for the 100 noise repetitions. Compared to Section 4.2, summarized global and  $\overline{CD}$  *Atm* error measures are slightly improved, whereas relative *Atm* errors are worse. Summarized *Spa* measures are slightly improved. Summarized *Cnt* measures are larger, because the radiance peaks at 1.74 and 2.3  $\mu\text{m}$  are now included, resulting in  $t_{c1.74}$  and  $t_{c2.3}$  errors that naturally do not occur in Section 4.2. Compared to Section 4.3 (best case), summarized global *Atm* error measures are smaller, but color dispersion as well as most relative errors are slightly worse. Summarized *Spa* errors are slightly better except for  $\overline{RG}_{1.10}$  and  $\overline{CD}$ . Relative error measures are almost the same, and summarized *Cnt* measures are slightly larger. The retrieval of  $m_{2'}$  and  $m_3$  as auxiliary parameters from the 1.74 and 2.3  $\mu\text{m}$  peaks is preferable to the pipelines in Sections 4.2 and 4.3, mainly because of the now vastly better noise behavior that leads to much better defined retrieved parameters.

#### 4.5. Cloud modes 2, 2', and 3 from 1.00–2.30 $\mu\text{m}$

Finally, the case is studied, where, along with emissivities, cloud mode factors  $m_2$ ,  $m_{2'}$ , and  $m_3$  are retrieved from the range 1.0–2.3  $\mu\text{m}$  excluding the 1.28  $\mu\text{m}$  peak. This retrieval pipeline leads to the best overall results and is now discussed in more detail. Note that the emissivities do not affect the range longward of 1.295  $\mu\text{m}$ , while the cloud modes affect the whole range. Emissivities directly interfere with the cloud modes, therefore, but the cloud modes are mainly determined from the spectral information content in the range longward of 1.295  $\mu\text{m}$  with the added information that they have to be compatible with a realistic emissivity. This avoids cloud mode factors that allow a good fit longward of 1.295  $\mu\text{m}$  but not in the surface windows. Note that, when  $m_2$  instead of  $m_1$  is treated as external parameter, most error and noise measures slightly worsen. When the 1.31  $\mu\text{m}$  radiance peak is excluded, results do not significantly change, except that the  $c1.31$  measures naturally vanish and that the global  $\text{H}_2\text{O}$  error measure gets worse at 1.02  $\mu\text{m}$  and improves at 1.10  $\mu\text{m}$ , both by about ten percentage points.

Figure 5 displays retrieved emissivities for a number of external parameters that lead to large error measures. For each depicted parameter, results are shown at least for the wavelength that leads to the worst case global error measure. The emergence of the error measures can be qualitatively comprehended from such graphs. The deviation of retrieved emissivities from their true values 0.65 depends on wavelength and typically increases with increasing deviation of the external parameter from its true value. Since some of the graphs are not monotonic, the error measures in Section 2 were not defined by using the deviations of retrieved from true emissivities at the *CR* boundaries, but by computing the maxima over the entire respective *CR* intervals (Eqs. (1), (2), and (6)). Large color dispersions occur where the retrieved emissivities at the three surface windows strongly disagree (in the figure for  $\text{H}_2\text{O}$  and to a lesser degree for  $\Delta T0$ ). The relative error measure for a parameter is closely related to the product of  $I_t$  and the maximal slope in the graph (Eq. (6)). The top left display exemplarily demonstrates that the double standard deviations of the retrieved emissivities for the 100 noise repetitions vary only slightly with the external parameter, see discussion for Eq. (8). It also shows a case where the entire parameter interval  $CR = [72.5\%, 87.5\%]$  over which the respective external parameter is expected to vary (Table 1) yields valid retrievals ( $QF \leq 0.8\% \leq 3\%$ ).

Table 4 lists the error and noise measures. The rows  $m_2$ ,  $m_{2'}$ ,  $m_3$ , and  $ei$  do not represent the error measures for external parameter variations, but for their *a priori* mean value variations over the range *CR*. This is indicated by a '+' in the first column. These entries are not taken into account when computing the summarized error measures, but they are negligibly small anyway.  $t_{\text{H}_2\text{O}}$  and  $t_{c1.74}$  variations (for both  $CR = [0.75, 1.25]$ ) outside of the intervals  $[0.79, 1.25]$  and  $[0.84, 1.24]$ , respectively, lead to qualities of fits  $QF$  that on average over the noise repetitions exceed the threshold of 3% above which a retrieval is not regarded as valid and no error and noise measures are calculated (Section 2.2). Row  $e1.02$  shows the 1.10 and 1.18  $\mu\text{m}$  emissivity errors, when the 1.02  $\mu\text{m}$  emissivity is treated as external parameter. Since a parameter that is treated as external parameter cannot be retrieved, the entries for  $\overline{RG}_{1.02}$  and  $\overline{RL}_{1.02}$  are flagged by '/'. The

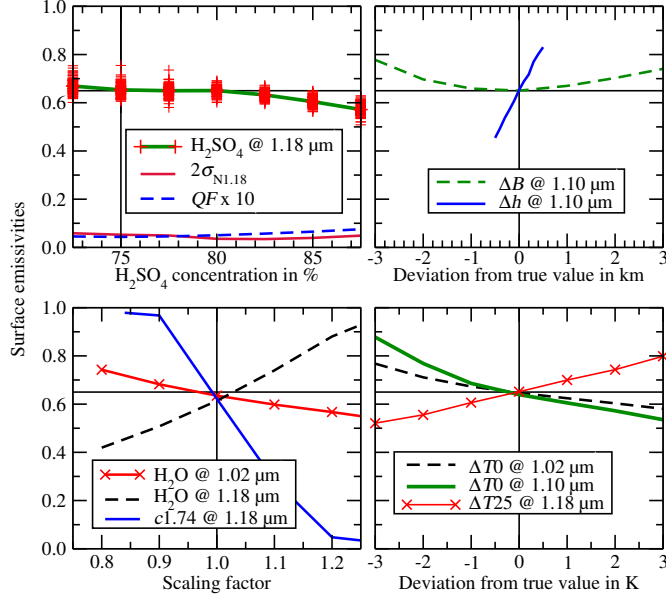


Figure 5: Retrieved surface emissivities for variation of selected external parameters (compare Tables 1 and 4) for the pipeline in Section 4.5. Parameter names and wavelengths of retrieved emissivities are given in the legends. All values are mean values over the 100 noise repetitions, except for top left display where the scattered symbols also show retrieved 1.18  $\mu\text{m}$  emissivities for all noise repetitions. Top left display also shows quality of fit  $QF$  (averaged over noise repetitions and scaled by factor of 10 for better representation) and double standard deviation  $2\sigma_{N1.18}$  of retrieved 1.18  $\mu\text{m}$  emissivities. Only results for the valid intervals ( $QF \leq 3\%$ ) are shown. True emissivities (0.65) and parameter values are indicated by continuous black lines.

rows  $e1.10$  and  $e1.18$  show analogous results for external 1.10 and 1.18  $\mu\text{m}$  emissivity variations, respectively. The entries in the rows  $e1.02$ ,  $e1.10$ , and  $e1.18$  contribute to the summarized  $Spa$  errors.

Compared to Section 4.4, summarized  $Atm$  and  $Cnt$  errors have reduced. In particular, the relative  $Atm$  errors are much smaller. They show that the retrieval errors in the spatial fine structure of a retrieved emissivity map are here essentially not due to  $Atm$  parameters. The summarized  $Spa$  errors are smaller than in Section 4.4, except for the relative error measures, where external emissivity variations now contribute. Noise measures have become worse (by a factor of 3) but are still much lower than for any of the other pipelines. The retrieved cloud mode factors  $m_2$ ,  $m_{2'}$ , and  $m_3$  typically scatter with double standard deviations of respectively 0.2, 0.1, and 0.01 for the 100 noise repetitions. The error measures are still very large even for this pipeline. It thus seems to be impossible to obtain reliable emissivity data from single spectra.

However, if a surface bin is covered by several measurements, some error and the noise measures can be improved. This can be done by applying MSR, and by not retrieving absolute emissivities but emissivity maps relative to an initial emissivity and, when appropriate, a subsequent renormalization for targets with limited spatial extension. This was discussed in Section 2.3, where also the corresponding scaling properties for the error and noise measures were given. Table 4 illustrates this by showing the scaled summarized error and noise measures for  $N_r = 25$  measurement repetitions, excluding the  $Cnt$  impacts that are assumed to be negligible. The spatial fine structure part of the  $Spa$  errors,  ${}^s\langle Spa \rangle_F$ , includes the contributions from  $t_{\Delta h}$ ,  $t_{e1.02}$ ,  $t_{e1.10}$ , and  $t_{e1.18}$ . The spatially slowly varying part  ${}^s\langle Spa \rangle_S$  comprises the impacts of  $t_{\Delta T0}$  and  $t_{\Delta T25}$ . While noise measures are negligible now, the other errors are still quite large. Scaled global and  $\overline{CD}$   $Atm$  errors are dominated by  $t_{H_2O}$  and  $t_{\Delta B}$  uncertainties. Scaled global and  $\overline{CD}$   $Spa$  errors are very large and are dominated by the temperature uncertainties. The large  $\overline{CD}$   ${}^s\langle Spa \rangle_S$  error indicates that the value by which a retrieved emissivity map is corrected in a renormalization may differ for the three surface windows. Spatial fine structure errors are completely dominated by  $t_{\Delta h}$  uncertainties, as can be seen from the relative errors.

Spatial  $t_{\Delta T0}$  and  $t_{\Delta T25}$  variations should not depend on longitude (or local time) but may quite weakly depend only on latitude. The additive renormalization of the retrieved relative emissivity map can remove

$t$		$\overline{RG}_i$			$\overline{CD}$	$\overline{RL}_i$		
$m1$		1	<b>2</b>	1	2	0	<b>0</b>	0
$m2$	+	0	<b>1</b>	0	1	0	<b>0</b>	0
$m2'$	+	0	0	<b>0</b>	0	0	0	<b>0</b>
$m3$	+	0	<b>1</b>	1	1	0	<b>0</b>	0
$\Delta t1$		0	<b>1</b>	1	1	0	0	<b>0</b>
$\Delta t2'$		1	<b>2</b>	1	1	0	<b>1</b>	0
$\Delta t3$		1	<b>2</b>	1	1	0	<b>1</b>	0
$\Delta B$		5	<b>20</b>	14	13	0	<b>1</b>	0
$\Delta t2$		0	<b>1</b>	1	1	0	0	<b>0</b>
$H_2SO_4$		2	3	<b>12</b>	14	1	1	<b>2</b>
$H_2O$		15	26	<b>43</b>	58	1	1	<b>2</b>
$HCl$		2	3	<b>8</b>	9	0	1	<b>1</b>
$\Delta h$		9	<b>12</b>	12	3	10	<b>13</b>	12
$ei$	+	0	<b>1</b>	1	0	0	<b>1</b>	1
$e1.02$		/	<b>26</b>	22	5	/	<b>30</b>	24
$e1.10$		6	/	<b>15</b>	10	7	/	<b>18</b>
$e1.18$		7	<b>22</b>	/	18	7	<b>27</b>	/
$\Delta T0$		18	<b>35</b>	35	13	0	<b>1</b>	1
$\Delta T25$		12	21	<b>23</b>	12	0	0	<b>1</b>
$c1.02$		<b>18</b>	1	1	17	/	/	/
$c1.10$		7	<b>85</b>	18	150	/	/	/
$c1.18$		14	42	<b>49</b>	75	/	/	/
$c1.31$		3	<b>9</b>	8	6	/	/	/
$c1.74$		24	83	<b>94</b>	171	/	/	/
$c2.3$		17	71	<b>73</b>	101	/	/	/
$\langle Atm \rangle$		17	33	<b>48</b>	62	1	2	<b>3</b>
$\langle Spa \rangle$		25	<b>54</b>	51	28	14	<b>43</b>	32
$\langle Cnt \rangle$		38	<b>145</b>	130	260	/	/	/
$N_i$		4	<b>11</b>	8				
${}^s\langle Atm \rangle$		3	7	<b>10</b>	12	1	2	<b>3</b>
${}^s\langle Spa \rangle_F$		9	<b>14</b>	13	5	10	<b>15</b>	13
${}^s\langle Spa \rangle_S$		22	41	<b>42</b>	18	0	1	<b>1</b>
${}^sN_i$		1	<b>2</b>	2				

Table 4: Surface emissivity error and noise measures in percent for the pipeline in Section 4.5, also compare Fig. 5. Scenario 1, 100 noise repetitions. Values below the last double horizontal line are exemplarily scaled (flagged by superscript ‘ $s$ ’ in ‘ $t$ ’ column) to emulate MSR retrieval of emissivity relative to the initial emissivity 0.65 for  $N_r = 25$  measurement repetitions. The largest of the values at the different wavelengths is always typed in bold.

their error contributions for surface targets with limited latitudinal extension. Then, the retrieval errors due to  $Spa$  parameters effectively comprise only the spatial fine structure part  ${}^s\langle Spa \rangle_F$ . Not much can be done with respect to  $t_{\Delta h}$  interferences that dominate  ${}^s\langle Spa \rangle_F$ , except that areas with more rugged mountains should be avoided to keep the elevation uncertainties small. Note that for areas with a wide latitudinal extension, consistent unreasonable large-scale latitudinal trends of retrieved emissivities may hint at a corresponding deviation of the deep atmospheric temperature from VIRA. To be able to recognize the significance of such a deviation and its order of magnitude,  ${}^s\langle Spa \rangle_S$  is kept in the table.

Finally, note that for any two different sets of spectra that cover the same surface bin,  $Spa$  parameters and their uncertainty impacts are identical (at least up to first order in perturbations by the other external parameters). When the repeatability of MSR retrievals with different sets of spectra is tested, differences in the results should thus be mainly explainable with  $Atm$  errors and noise. The large  $t_{H_2O}$  dominated  $\overline{CD} {}^s\langle Atm \rangle$  error indicates that errors of the retrieved emissivity maps can spatially differ for the three surface windows but not on the fine structure scale, see relative  ${}^s\langle Atm \rangle$  errors and the correlation length for  $t_{H_2O}$  (Section 3.4). Note that there may be unsystematic calibration or preprocessing errors. For instance, an imperfection in the removal of straylight (Kappel et al., 2012c) can introduce unsystematic additional emissivity errors, and repeatability of MSR retrievals with different sets of spectra may not be given anymore. Furthermore, multi-spectrum retrieval errors are here only roughly estimated by scaling single-spectrum error and noise measures according to plausible rules. This may not sufficiently well describe the true

multi-spectrum retrieval errors. Running into subsidiary multi-spectrum retrieval solutions, for example, can increase retrieval errors, and inclusion of information from adjacent measurements can strongly decrease retrieval errors. But the presented scaling should at least provide a rough guideline.

#### 4.6. Other reference spectra

Until now, only the default reference spectrum (scenario 1) has been studied. This section presents for the pipeline from Section 4.5 the error and noise measures for the other reference spectra that were introduced in Section 3.8. Results are shown in Table 5. However, considerable computer resources are required to

$t$	$RG_{1.02}$						$RG_{1.10}$						$RG_{1.18}$						$CD$								
	S0	S1	S2	S3	S4	S5	S0	S1	S2	S3	S4	S5	S0	S1	S2	S3	S4	S5	S0	S1	S2	S3	S4	S5			
$m1$	1	1	1	<b>1</b>	1	<i>1</i>	2	<b>1</b>	1	1	1	<i>1</i>	1	1	<b>1</b>	1	0	1	0	0	0	2	<b>2</b>	2	<i>1</i>	2	2
$\Delta t2'$	1	<i>1</i>	1	<b>1</b>	1	1	2	1	<b>1</b>	1	<i>1</i>	1	1	1	0	<b>1</b>	0	0	0	0	1	1	<b>2</b>	<i>1</i>	1	1	1
$\Delta t3$	1	0	<b>2</b>	2	0	1	2	2	<b>4</b>	3	<i>2</i>	2	2	1	2	<b>4</b>	2	1	1	1	1	2	<b>5</b>	<i>1</i>	1	1	1
$\Delta B$	5	5	6	<b>16</b>	6	5	20	20	15	<b>25</b>	<i>11</i>	20	14	14	17	<b>20</b>	<i>11</i>	13	13	13	13	<b>13</b>	7	5	13	13	
$H_2SO_4$	2	<i>1</i>	<b>7</b>	2	1	2	3	2	<b>19</b>	4	2	<i>2</i>	12	12	6	10	10	<b>13</b>	14	14	<b>24</b>	8	11	15	15	15	
$H_2O$	15	31	22	<b>41</b>	<i>16</i>	28	26	17	21	14	<i>10</i>	<b>22</b>	43	42	6	11	20	<b>44</b>	58	51	<i>19</i>	28	32	<b>52</b>	52	52	
HCl	2	3	<i>2</i>	3	4	<b>6</b>	3	5	4	3	4	<b>14</b>	8	9	6	5	6	<b>10</b>	9	10	7	6	6	6	<b>11</b>	11	
$\Delta h$	9	10	10	<b>11</b>	10	<i>9</i>	12	12	<b>12</b>	9	8	12	12	11	11	9	8	<b>11</b>	3	<b>3</b>	3	2	<i>2</i>	3	3	3	
$e1.02$	/	/	/	/	/	/	26	22	<b>43</b>	8	14	22	22	19	<b>40</b>	7	12	18	5	4	<b>5</b>	1	2	4	4	4	
$e1.10$	6	<b>6</b>	5	6	4	5	/	/	/	/	/	/	15	<b>15</b>	14	7	8	11	10	<b>10</b>	9	1	4	7	7	7	
$e1.18$	7	7	7	<b>9</b>	6	7	22	22	<b>24</b>	<i>14</i>	14	21	/	/	/	/	/	/	18	17	<b>20</b>	6	9	17	17	17	
$\Delta T0$	18	17	16	12	<i>11</i>	<b>18</b>	35	34	30	<i>12</i>	14	<b>40</b>	35	33	28	<i>12</i>	14	<b>40</b>	13	14	11	1	4	<b>17</b>	17	17	
$\Delta T25$	12	12	<b>13</b>	7	11	12	21	21	<b>22</b>	7	15	21	23	23	<b>23</b>	8	17	23	12	12	12	1	7	<b>12</b>	12	12	
$c1.02$	18	18	<b>19</b>	<i>11</i>	16	18	1	1	1	0	0	1	1	1	1	0	0	1	17	17	<b>18</b>	<i>11</i>	18	17	17	17	
$c1.10$	7	3	3	2	<b>3</b>	3	85	84	<b>84</b>	<i>43</i>	70	84	18	8	<b>8</b>	2	6	8	150	148	<b>149</b>	<i>38</i>	112	148	148	148	
$c1.18$	14	16	17	9	11	<b>25</b>	42	46	46	<i>12</i>	21	<b>54</b>	49	50	48	<i>18</i>	28	<b>52</b>	75	<b>72</b>	71	<i>26</i>	43	65	65	65	
$c1.31$	3	3	<i>3</i>	7	3	3	9	9	9	<b>10</b>	7	9	8	8	8	<b>9</b>	6	7	6	<b>6</b>	6	3	4	6	6	6	
$c1.74$	24	27	<i>22</i>	<b>39</b>	27	27	83	<b>94</b>	72	<i>54</i>	83	94	94	96	76	<i>54</i>	85	<b>96</b>	171	180	108	<i>47</i>	130	<b>181</b>	181	181	
$c2.3$	17	16	9	17	16	<b>18</b>	71	62	39	<i>33</i>	51	<b>70</b>	73	65	41	<i>36</i>	52	<b>72</b>	101	84	33	<i>24</i>	55	<b>99</b>	99	99	
$\langle Atm \rangle$	17	32	24	<b>44</b>	18	29	33	27	<b>33</b>	29	16	33	48	47	21	25	25	<b>49</b>	62	56	34	<i>31</i>	35	<b>57</b>	57	57	
$\langle Spa \rangle$	25	25	<b>25</b>	21	20	25	54	52	<b>63</b>	<i>23</i>	30	55	51	48	<b>57</b>	20	27	52	28	28	28	7	13	<b>29</b>	29	29	
$\langle Cnt \rangle$	38	40	<i>35</i>	<b>45</b>	37	45	145	148	126	78	122	<b>154</b>	130	127	99	<i>68</i>	104	<b>131</b>	260	259	201	71	186	<b>262</b>	262	262	
$N_i$	4	3	<b>7</b>	6	4	5	11	11	<b>18</b>	8	8	16	8	8	<b>15</b>	6	6	12	8	8	6	6	12	12	12	12	
$^s \langle Atm \rangle$	3	6	5	<b>9</b>	4	6	7	5	<b>7</b>	6	3	7	10	9	4	5	5	<b>10</b>	12	11	7	6	7	<b>11</b>	11	11	
$^s \langle Spa \rangle_F$	9	10	10	<b>11</b>	10	9	14	14	<b>16</b>	10	9	13	13	12	<b>14</b>	9	9	12	5	5	<b>5</b>	2	3	5	5	5	
$^s \langle Spa \rangle_S$	22	21	21	<i>14</i>	16	<b>22</b>	41	40	37	<i>14</i>	21	<b>45</b>	42	40	36	<i>14</i>	22	<b>46</b>	18	18	16	1	8	<b>21</b>	21	21	
$^s N_i$	1	<i>1</i>	1	1	1	1	2	2	<b>4</b>	2	<i>2</i>	3	2	2	<b>3</b>	1	1	2	2	2	2	2	2	2	2	2	

Table 5: Emissivity error and noise measures in percent for the pipeline in Section 4.5. S1–S5: noiseless error measures and exemplary noise measures for scenarios 1–5. S0: noisy results for scenario 1, for comparison to S1 carried over from Table 4. Largest value out of S1–S5: bold, smallest: italic. Rows with just a few entries 1 and otherwise zero are omitted ( $\Delta t1$ ,  $\Delta t2$ , and  $a priori$  data modification results for  $m2$ ,  $m2'$ ,  $m3$ ,  $ei$ ). Other notation as in Table 4, but initial emissivity for S4 is 0.9.

generate a single table like Table 4. Therefore, no noise repetitions are performed now, and only the noiseless cases are analyzed. To still get an exemplary estimate on noise measures, one single parameter ( $m2'$ ), while its  $a priori$  mean value is varied to cover its  $CR$ , is additionally studied for all 100 noise repetitions as before. Both the reduced parameter intervals ensuring valid  $QF$ 's and the  $RL_i$  values do not differ much from the corresponding values for noisy scenario 1 and are not shown here. The  $RL_i$  values for rows  $e1.02$ ,  $e1.10$ , and  $e1.18$  are not much different from the respective  $RG_i$  values.

The noisy results from Table 4 (S0 columns in Table 5) agree well with the noiseless error measures, respectively the 'exemplary' noise measures for scenario 1 (S1). Only  $t_{H_2O}$  ( $RG_{1.02}$ ,  $RG_{1.10}$ ),  $t_{c1.10}$  ( $RG_{1.02}$ ,  $RG_{1.18}$ ),  $t_{c1.74}$  ( $RG_{1.10}$ ,  $CD$ ), and  $t_{c2.3}$  ( $RG_{1.10}$ ,  $RG_{1.18}$ ,  $CD$ ) have larger differences in their impacts. These can be explained, for instance, by the convergence to different subsidiary minima of the retrieval cost function. But in general, the noiseless results seem to be usable as guidelines for the noisy error measures (see also the summarized error measures), and the noise measures can be reasonably approximated by the exemplary noise measures. The standard deviations of the retrieved emissivities over the 100 noise repetitions are very similar to the averages over the  $a posteriori$  standard deviations of the individual retrieval runs.

While the larger cloud opacity in scenario 2 leads to a significant degradation of the retrieved emissivities for the pipeline in Section 4.4, it does not consistently lead to worse error measures for the present pipeline. For instance,  $RG_{1.10}$  improves even for  $t_{\Delta B}$ , and while  $RG_{1.02}$  and  $RG_{1.10}$  significantly increase for  $t_{\text{H}_2\text{SO}_4}$ ,  $RG_{1.18}$  improves significantly. Also note that for  $t_{\text{H}_2\text{O}}$ , S2 leads to a much smaller  $RG_{1.18}$ .  $Spa$  errors are predominantly unaffected by the larger cloud opacity, except for external  $t_{e1.02}$  variations where retrieved emissivities at 1.10 and 1.18  $\mu\text{m}$  significantly worsen.  $Cnt$  error measures improve for  $t_{c1.74}$  and  $t_{c2.3}$  and are unaffected elsewhere. The generally smaller radiances for S2 (Fig. 4) lead to a decreased signal-to-noise ratio, which results in accordingly increased noise measures. Except for a few parameters that lead to higher error measures at 1.02  $\mu\text{m}$ , the error measures for S3 mostly decrease compared to S1, partially strongly. Apparently, in spite of the smaller radiances in the surface windows (Fig. 4), the increased surface elevation of 5 km leads for S3 to a better sounding of surface properties. This is because at this high elevation, the surface contributes a higher fraction to the observable surface window radiances, as can be checked by simulations, see also Haus and Arnold (2010, Fig. 9). For S4, most error measures are better than for S1. This is likely a result of the higher surface window radiances that follow from the higher true emissivities of 0.9, and of the correspondingly increased fraction of surface information in those radiance peaks. The noise measures for S3 and S4 are mostly slightly improved compared to S1, probably due to the increased surface information content, except for  $N_{1.02}$ , where for S3 the decreased signal-to-noise ratio following from the smaller surface window radiances (Fig. 4) apparently outweighs this effect. S5 exhibits larger  $t_{\text{HCl}}$ ,  $t_{\Delta T0}$ ,  $t_{c1.18}$ , and  $t_{c2.3}$  impacts than S1, whereas most other error measures are unaffected. Most summarized error measures slightly increase. The smaller radiances for S5 that are due to the larger observation angle (Fig. 4) lead to a decreased signal-to-noise ratio and increased noise measures.

The scaled summarized error and noise measures for the renormalized MSR retrieval of relative emissivity with  $N_r = 25$  repetitions (Section 2.3) exhibit higher noise measures compared to those for Section 4.4, but they are still small. While  ${}^s\langle Atm \rangle$  error measures have become slightly worse at 1.02  $\mu\text{m}$ , they have become significantly smaller in the other two surface windows.  ${}^s\langle Spa \rangle_F$  and  ${}^s\langle Spa \rangle_S$  errors have reduced.

The renormalized global errors are dominated by  ${}^s\langle Spa \rangle_F$ . Mostly, the error and noise measures are larger at 1.10 and 1.18 than at 1.02  $\mu\text{m}$ . There is an overall tendency for better results at low cloud opacity, low observation angle, higher surface elevation, or true emissivity, but differences are not large. This can be seen from the quadratic means over the three wavelengths of the quadratically added global error and noise measures, which read 13%, 14%, 12%, 10%, 14% for S1–S5. The analogous results for the pipeline in Section 4.4 read 18%, 19%, 13%, 13%, 18% and show slightly more pronounced differences between the five scenarios. They are also slightly larger, which underlines the preference of the present pipeline. The  ${}^s\langle Atm \rangle$  error measures alone show no clear tendency except that S4 leads to best results. For a significantly higher measurement repetition number than  $N_r = 25$ , essentially only  $t_{\Delta h}$  errors remain. The just discussed overall tendencies are still valid in this case, except at 1.02  $\mu\text{m}$  where the influence of the scenario is negligible.

Without the renormalization for targets with limited latitudinal extension,  ${}^s\langle Spa \rangle_S$  errors are unacceptably large, especially at 1.10 and 1.18  $\mu\text{m}$ . Systematic latitudinal deep atmospheric temperature deviations from VIRA in the order of 3 K should be easily recognizable.

Since the FWHM of the VIRTIS-M-IR instrumental response function is currently not sufficiently well predictable in the calibration procedure (Kappel et al., 2012c), it shall be noted that spectra with a smaller FWHM (for instance 10 instead of 17 nm) tend to yield somewhat smaller (typically by a few percent) error and noise measures. Note that a tenfold increase of the spectral resolution (1.7 nm FWHM, 0.949 nm spectral sampling), does not lead, for instance in case of the retrieval pipeline from Section 4.2, to large improvements compared to VIRTIS-M-IR resolution. However, at this high spectral resolution it is possible, for instance for the retrieval of  $e_{1.02}$ , to largely eliminate the interferences of the  $Atm$  parameters (unscaled summarized global 1.02  $\mu\text{m}$   $Atm$  error measure  $\leq 2\%$ ) by retrieving only a total cloud factor along with  $e_{1.02}$  from the 1.02  $\mu\text{m}$  peak alone. The spectral range has to be restricted to 1.01–1.05  $\mu\text{m}$  to exclude  $\text{H}_2\text{O}$  interferences at shorter wavelengths. This improvement underlines the importance of the interfering parameters' color dispersion effects that are still present in the first case but largely vanish over the narrow spectral range utilized in the second case. But even for the second case, a number of parameters ( $t_{\Delta h}$ ,  $t_{\Delta T0}$ ,  $t_{\Delta T25}$ ,  $t_{c1.02}$ ) with Jacobians that are very similar to the emissivity Jacobian in the 1.02  $\mu\text{m}$  window lead to retrieval errors that are still large (noiseless  $RG_{1.02}$  10%, 12%, 10%, 24%). They could be removed by a renormalized

relative emissivity retrieval, but errors due to topography uncertainties cannot be improved this way.

## 5. Summary and conclusion

The impacts were investigated that a number of interfering parameters and measurement noise can have on 1.02, 1.10, and 1.18  $\mu\text{m}$  surface emissivities that are retrieved from VIRTIS-M-IR spectra of Venus' nightside emissions in the range 1.0–2.3  $\mu\text{m}$ . Spectral information content is limited, and a number of parameters required for the numerical simulation of spectra cannot be uniquely derived from the measured spectra. These parameters are usually set to certain reasonable values that are compatible with current knowledge on atmospheric and surface conditions and allow to generate realistic synthetic spectra. When the true physical values deviate from these assumptions, this may cause surface emissivity retrieval errors.

Synthetic spectra were used as reference spectra where all underlying parameters are known by construction, and Gaussian noise was added to emulate measurement noise. Along with the emissivities, auxiliary parameters had to be retrieved from these synthetic measurements to allow for useful fits at all. The retrieved emissivities were compared to their true values. This was repeated for several different realizations of the Gaussian noise and with each studied interfering parameter set to a series of reasonable values. A number of emissivity retrieval error and noise measures were monitored to allow a categorization of the success of the emissivity retrieval. Global error measures represent for each studied interfering parameter the maximum relative impact on the retrieved emissivities when the interfering parameter is varied over the range it is expected to typically cover. A further characteristic error measure is the color dispersion between the emissivities at the three accessible wavelengths. The third error measure estimates relative errors between two simultaneously measured surface bins that are separated by 100 km, the order of the expected achievable spatial resolution of surface data. The noise measure represents typical retrieval errors due to measurement noise.

Several retrieval pipelines were tested to find out, which auxiliary parameters should be retrieved along with the emissivities from which spectral range in order to achieve the lowest retrieval errors. The best pipeline retrieves cloud mode factors  $m_2$ ,  $m_{2'}$ , and  $m_3$  as auxiliary parameters from the spectral range 1.0–2.3  $\mu\text{m}$  excluding the 1.28  $\mu\text{m}$  peak that is in practice contaminated by non-LTE  $\text{O}_2$  nightglow. It was found that even for this pipeline, error and noise measures are so large that emissivity cannot be derived within reasonable error margins from a single VIRTIS-M-IR spectrum, even though the fits well match the synthetic measurements in each case.

This situation can be improved by retrieving the emissivity of a surface bin as a parameter that is common to several measurements that cover this bin by applying the MSR multi-spectrum retrieval algorithm (Kappel, 2014). This assumes negligible geologic activity. MSR can also incorporate correlation lengths and times of the retrieval parameters and thus leads to more reliable results. But even with MSR, absolute emissivity values are difficult to obtain. To minimize errors that are due to  $\text{CO}_2$  opacity uncertainties, it can be useful to set an initial emissivity value (e.g. 0.65) and to pre-determine wavelength dependent  $\text{CO}_2$  opacity corrections as parameters that are compatible with as many and diverse as possible observed spectra by applying MSR. Using these pre-determined parameters, a surface emissivity map relative to the initial emissivity can be retrieved with MSR for the actual surface target. Moreover, spatially slowly varying systematic interferences from other parameters than the opacity corrections can lead to spatially large scale deviations of the retrieved from the true emissivity map. When only the spatial emissivity fine structure (100 km scale) of a target with limited spatial extensions (a few thousand km) is of interest, the retrieved relative emissivity map can be additively renormalized such that its new median value corresponds to the initial emissivity. Then the spatially slowly varying interferences can be neglected as spatially almost constant emissivity offsets that have been removed by the renormalization. On the other hand, for areas with a wide latitudinal extension, consistent unreasonable large-scale latitudinal trends of retrieved emissivities may hint at a corresponding deep atmospheric temperature deviation from VIRA (Seiff et al., 1985) in latitude direction.

A full multi-spectrum retrieval error analysis is very complex and requires considerable computer resources. But the retrieval error and noise measures can be roughly estimated by plausible scaling rules that emulate the statistical impact when each surface bin is covered by  $N_r$  repeated measurements. This was exemplarily done for  $N_r = 25$  and five different reference spectra that represent various typical environmental and observational conditions (Table 5). Renormalized relative emissivity retrieval errors due to uncertainties



of statistically independent temporally varying interfering parameters like cloud properties and minor gas abundances can add up to 3%–10% of the true emissivity value with color dispersions between 6% and 12%, depending on surface window and spectrum. Temporally constant interfering parameters with a spatial fine structure on a scale of 100 km (surface elevation, interfering emissivities from other surface windows) contribute another 9%–16% with color dispersions between 2% and 5%. Measurement noise with a standard deviation of  $10^{-4}$  W/(m<sup>2</sup> sr  $\mu$ m) leads to additional 1%–4%. Reasonable *a priori* data modifications were checked to have negligible impacts. Spatial fine structure errors can add up to 10%, 15%, and 13% for the three wavelengths and are completely dominated by the impacts of assumed surface elevation uncertainties of 200 m (Table 4). Useful results are thus possible, but the errors are still not small. Retrieved emissivities are most reliable at 1.02  $\mu$ m. The most suitable spectra for emissivity retrieval tend to be those with a small cloud opacity, high surface elevation, high emissivity, and small observation angle, but this depends on the emissivity window, retrieval pipeline, and measurement repetition number.

It may be possible that utilization of a wider spectral range of VIRTIS-M-IR could help to reduce errors (but not those due to surface elevation uncertainties). In particular, the cloud mode abundances can probably be better disentangled by additionally using the range around the 4.3  $\mu$ m CO<sub>2</sub> absorption band, compare Haus et al. (2013). But then, the strong spatial and temporal variations of the mesospheric temperature altitude profiles must be considered by treating them as additional auxiliary parameter vector. A further approach could be to choose more or different auxiliary parameters or to pre-estimate some parameters by other means (ratio based estimation of cloud bottom altitude or H<sub>2</sub>SO<sub>4</sub> concentration of cloud droplets (Barstow et al., 2012)), but note that this does not concern the parameters that are the most problematic for emissivity retrieval. Also note that even a tenfold spectral resolution does not help to avoid errors due to topography uncertainties. Geologic input and the ranges of the retrieved renormalized relative emissivities and the quality of the fits may help to constrain realistic initial emissivities.

Haus and Arnold (2010) showed 1.02  $\mu$ m emissivities retrieved from many VIRTIS-M-IR spectra. The authors noted that these emissivities increased for increasing topography. It was mentioned that the results were preliminary and had to be interpreted with much care due to interferences from other parameters. In fact, they applied a single-spectrum retrieval. Even though the fits had been very good, it follows from the present work that these results were probably not reliable due to precisely such interferences.

This work has studied retrieval errors only. Calibration, preprocessing, and forward model errors can lead to additional emissivity errors. But calibration and preprocessing were improved by Kappel et al. (2012c), 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 and preprocessing and functional descriptions of the simulated spectra are not expected to be severely wrong, therefore. Even if they were wrong, renormalized relative emissivities would not be strongly affected by that. Interfering parameters that were not studied here 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), etc. But impacts of quasi-random time variations of these parameters should disappear with the inverse square root of the repetition number, and impacts of systematically wrong assumptions should be greatly diminished through the renormalized relative emissivity retrieval. Overall, the potential additional errors mentioned in this paragraph are expected to be not large compared to the already determined MSR retrieval errors, and the results of this paper should give good advice on the typical errors that may occur.

A subsequent paper is planned by the present authors, where, based on the current results, a renormalized relative emissivity map will be retrieved with MSR, and where retrieval errors and reproducibility of results for modified external parameters and initial emissivity will be discussed. It will also be checked, whether maps retrieved from different selections of measurements reasonably agree. For this kind of test, 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 that are too large to be explained by the presented results may be due to calibration and preprocessing errors or may indicate that there are temporally varying interfering parameters that have not been considered here.

## Acknowledgements

This work is funded by the German Aerospace Center (DLR) e.V. and the German Research Foundation (DFG). The authors would like to thank the reviewers for their valuable suggestions to improve this paper.

## References

- Abdrakhimov, A. M., Basilevsky, A. T., 2002. Geology of the Venera and Vega landing-site regions. *Solar Syst. Res.* 36 (2), 136–159.
- Anderson, T. W., 1958. An introduction to multivariate statistical analysis. Wiley publications in statistics. Wiley, New York, NY, USA.
- Arnold, G., Haus, R., Kappel, D., Drossart, P., Piccioni, G., 2008. Venus surface data extraction from VIRTIS/Venus Express measurements: Estimation of a quantitative approach. *J. Geophys. Res.* 113 (E5), E00B10.
- Arnold, G. E., Haus, R., Kappel, D., Piccioni, G., Drossart, P., 2012. VIRTIS/VEX observations of Venus: overview of selected scientific results. *J. Appl. Remote Sens.* 6 (1), 063580–1–063580–20.
- Balgovind, R., Dalcher, A., Ghil, M., Kalnay, E., 1983. A stochastic-dynamic model for the spectral structure of forecast errors. *Mon. Weather Rev.* 111 (4), 701–721, doi:10.1175/1520-0493(1983)111<0701:ASDMFT>2.0.CO;2.
- Barstow, J. K., Tsang, C. C. C., Wilson, C. F., Irwin, P. G. J., Taylor, F. W., McGouldrick, K., Drossart, P., Piccioni, G., Tellmann, S., 2012. Models of the global cloud structure on Venus derived from Venus Express observations. *Icarus* 217 (2), 542–560, in: *Advances in Venus Science*.
- Basilevsky, A. T., Head, J. W., Schaber, G. G., Strom, R. G., 1997. The resurfacing history of Venus. In: Bougher, S. W., Hunten, D. M., Phillips, R. J. (Eds.), *Venus II: Geology, Geophysics, Atmosphere, and Solar Wind Environment*. Space science series. University of Arizona Press, Tucson, pp. 1047–1086.
- Bézar, B., Fedorova, A., Bertaux, J.-L., Rodin, A., Korablev, O., 2011. The 1.10- and 1.18- $\mu\text{m}$  nightside windows of Venus observed by SPICAV-IR aboard Venus Express. *Icarus* 216 (1), 173–183.
- Bjoraker, G. L., Larson, H. P., Mumma, M. J., Timmermann, R., Montani, J. L., 6 1992. Airborne observations of the gas composition of Venus above the cloud tops: measurements of  $\text{H}_2\text{O}$ , HDO, HF, and the D/H and  $^{18}\text{O}/^{16}\text{O}$  isotopic ratios. In: *AAS/Division for Planetary Sciences Meeting Abstracts*. Vol. 24 of *Bulletin of the American Astronomical Society*. p. 995.
- Cardesin-Moinelo, A., Piccioni, G., Ammannito, E., Filacchione, G., Drossart, P., 9 2010. Calibration of hyperspectral imaging data: VIRTIS-M onboard Venus Express. *IEEE Trans. Geosci. Remote* 48, 3941–3950.
- Carlson, R., Anderson, M., 2011. Absorption properties of sulfuric acid in Venus’s infrared spectral windows region. In: *EPSC-DPS Joint Meeting 2011*. p. 1171.
- Daley, R., 1991. *Atmospheric Data Analysis (Cambridge Atmospheric and Space Science Series)*. Cambridge University Press, Cambridge, UK.
- de Bergh, C., Bézar, B., Owen, T., Crisp, D., Maillard, J.-P., Lutz, B. L., 1991. Deuterium on Venus: observations from Earth. *Science* 251 (4993), 547–549.
- 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., Afanasenko, T. Z., 2007. Scientific goals for the observation of Venus by VIRTIS on ESA/Venus Express mission. *Planet. Space Sci.* 55 (12), 1653–1672.
- Fedorova, A., Korablev, O., Vandaele, A.-C., Bertaux, J.-L., Belyaev, D., Mahieux, A., Neefs, E., Wilquet, W. V., Drummond, R., Montmessin, F., Villard, E., 2008. HDO and  $\text{H}_2\text{O}$  vertical distributions and isotopic ratio in the Venus mesosphere by solar occultation at infrared spectrometer on board Venus Express. *J. Geophys. Res.: Planets* 113 (E5).
- Ford, P. G., Pettengill, G. H., 8 1992. Venus topography and kilometer-scale slopes. *J. Geophys. Res.: Planets* 97 (E8), 13103–13114.
- Gaspari, G., Cohn, S. E., 1999. Construction of correlation functions in two and three dimensions. *Q. J. Roy. Meteor. Soc.* 125 (554), 723–757.
- Gelb, A., 1974. *Applied Optimal Estimation*. MIT Press, Cambridge, MA., USA.
- Grieger, B., Ignatiev, N. I., Hoekzema, N. M., Keller, H. U., 2 2004. Indication of a near surface cloud layer on Venus from reanalysis of Venera 13/14 spectrophotometer data. In: Wilson, A. (Ed.), *Planetary Probe Atmospheric Entry and Descent Trajectory Analysis and Science*. Vol. 544 of *ESA Special Publication*. pp. 63–70.
- Hansen, J. E., Travis, L. D., 1974. Light scattering in planetary atmospheres. *Space Sci. Rev.* 16 (4), 527–610.
- Haus, R., Arnold, G., 2010. Radiative transfer in the atmosphere of Venus and application to surface emissivity retrieval from VIRTIS/VEX measurements. *Planet. Space Sci.* 58 (12), 1578–1598.
- Haus, R., Kappel, D., Arnold, G., 2013. 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.
- Haus, R., Kappel, D., Arnold, G., 2014. Atmospheric thermal structure and cloud features in the southern hemisphere of Venus as retrieved from VIRTIS/VEX radiation measurements. *Icarus* 232, 232–248.
- Haus, R., Kappel, D., Arnold, G., 2015. Lower atmosphere minor gas abundances as retrieved from Venus Express VIRTIS-M-IR data at 2.3  $\mu\text{m}$ . *Planet. Space Sci.* 105, 159–174.
- Ignatiev, N. I., Titov, D. V., Piccioni, G., Drossart, P., Markiewicz, W. J., Cottini, V., Roatsch, T., Almeida, M., Manoel, N., 2009. Altimetry of the Venus cloud tops from the Venus Express observations. *J. Geophys. Res.* 114 (E9), E00B43.

- Imamura, T., Hashimoto, G. L., 2001. Microphysics of Venusian clouds in rising tropical air. *J. Atmos. Sci.* 58 (23), 3597–3612, doi:10.1175/1520-0469(2001)058<3597:MOVCIR>2.0.CO;2.
- Ivanov, M. A., Head, J. W., 2006. Geologic map of the Mylitta Fluctus Quadrangle (V-61), Venus. US Geological Survey Scientific Investigations Map 2920, <http://pubs.usgs.gov/sim/2006/2920/>.
- Ivanov, M. A., Head, J. W., 2011. Global geological map of Venus. *Planet. Space Sci.* 59 (13), 1559–1600.
- Kappel, D., 2014. MSR, a multi-spectrum retrieval technique for spatially-temporally correlated or common Venus surface and atmosphere parameters. *J. Quant. Spectrosc. Rad.* 133, 153–176.
- Kappel, D., Arnold, G., Haus, R., 2012a. Retrieval of surface emissivity in a Venus coordinate patch as parameter common to repeated measurements by VIRTIS/VEX. In: EGU General Assembly 2012. Vol. 14. p. 9708.
- Kappel, D., Arnold, G., Haus, R., 2012b. Sensitivity of Venus surface emissivity retrieval to model variations of CO<sub>2</sub> opacity, cloud features, and deep atmosphere temperature field. In: 39th COSPAR Scientific Assembly. B0.8-0008-12.
- Kappel, D., Arnold, G., Haus, R., Piccioni, G., Drossart, P., 2012c. Refinements in the data analysis of VIRTIS-M-IR Venus nightside spectra. *Adv. Space Res.* 50 (2), 228–255.
- Knollenberg, R. G., Hunten, D. M., 1980. 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.
- Krasnopolsky, V. A., Belyaev, D. A., Gordon, I. E., Li, G., Rothman, L. S., 2013. Observations of D/H ratios in H<sub>2</sub>O, HCl, and HF on Venus and new DCI and DF line strengths. *Icarus* 224, 57–65.
- Lebonnois, S., Hourdin, F., Forget, F., Eymet, V., Fournier, R., 2010. The LMD Venus general circulation model: Improvements and questions. In: International Venus Conference, Aussois, France 20–26 June 2010. [http://lesia.obspm.fr/venus2010/sites/venus2010/IMG/pdf/04-04\\_Lebonnois.pdf](http://lesia.obspm.fr/venus2010/sites/venus2010/IMG/pdf/04-04_Lebonnois.pdf).
- Marov, M. Y., Lystsev, V. E., Lebedev, V. N., Lukashevich, N. L., Shari, V. P., 1980. The structure and microphysical properties of the Venus clouds: Venera 9, 10, and 11 data. *Icarus* 44 (3), 608–639.
- Meadows, V. S., Crisp, D., 1996. Ground-based near-infrared observations of the Venus nightside: The thermal structure and water abundance near the surface. *J. Geophys. Res.* 101 (E2), 4595–4622.
- More, J., 1978. The Levenberg-Marquardt algorithm: Implementation and theory. In: Watson, G. A. (Ed.), *Numerical Analysis*. Vol. 630 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, Ch. 10, pp. 105–116.
- Moroz, V. I., 2002. Estimates of visibility of the surface of Venus from descent probes and balloons. *Planet. Space Sci.* 50 (3), 287–297.
- Müller, N., Helbert, J., Hashimoto, G. L., Tsang, C. C. C., Erard, S., Piccioni, G., Drossart, P., 2008. Venus surface thermal emission at 1  $\mu$ m in VIRTIS imaging observations: Evidence for variation of crust and mantle differentiation conditions. *J. Geophys. Res.* 113 (E5), E00B17.
- Müller, N., Helbert, J., Stofan, E. R., Smrekar, S. E., Piccioni, G., Drossart, P., 2012. Search for active lava flows with VIRTIS on Venus Express. In: American Geophysical Union, 45th Fall Meeting. pp. P24B–02.
- Palmer, K. F., Williams, D., 1975. Optical constants of sulfuric acid; application to the clouds of Venus? *Appl. Opt.* 14 (1), 208–219.
- Pettengill, G. H., Ford, P. G., Johnson, W. T. K., Raney, R. K., Soderblom, L. A., 1991. Magellan: Radar performance and data products. *Science* 252 (5003), 260–265.
- 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., Zambelli, M., 2007. South-polar features on Venus similar to those near the north pole. *Nature* 450 (7170), 637–640.
- Piccioni, G., Zasova, L., Migliorini, A., Drossart, P., Shakun, A., Garcia Munoz, A., Mills, F. P., Cardesin-Moinelo, A., 2009. Near-IR oxygen nightglow observed by VIRTIS in the Venus upper atmosphere. *J. Geophys. Res.: Planets* 114 (E5).
- Pollack, J. B., Dalton, J. B., Grinspoon, D., Wattson, R. B., Freedman, R., Crisp, D., Allen, D. A., Bézard, B., DeBergh, C., Giver, L. P., Ma, Q., Tipping, R., 1993. Near-infrared light from Venus' nightside: A spectroscopic analysis. *Icarus* 103 (1), 1–42.
- Rappaport, N. J., Konopliv, A. S., Kucinkas, A. B., Ford, P. G., 1999. An improved 360 degree and order model of Venus topography. *Icarus* 139 (1), 19–31.
- Rodgers, C. D., 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice* (Series on Atmospheric, Oceanic and Planetary Physics). World Scientific Publishing Company, Singapore.
- 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., Miller, 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., Auwera, J. V., 2009. The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Rad.* 110 (9–10), 533–572.
- Rothman, L. S., Gordon, I. E., Barber, R. J., Dothe, H., Gamache, R. R., Goldman, A., Perevalov, V. I., Tashkun, S. A., Tennyson, J., 2010. HITRAN, the high-temperature molecular spectroscopic database. *J. Quant. Spectrosc. Rad.* 111 (15), 2139–2150.
- Rothman, L. S., Wattson, R. B., Gamache, R., Schroeder, J. W., McCann, A., 1995. HITRAN HAWKS and HITEMP: high-temperature molecular database. In: *Proc. SPIE 2471, Atmospheric Propagation and Remote Sensing IV*, 105 (June 15, 1995). pp. 105–111.
- 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., Marov, M. Y., 1985. Models of the structure of the atmosphere of Venus from the surface to 100 kilometers altitude. *Adv. Space Res.* 5 (11), 3–58.
- Smrekar, S. E., Stofan, E. R., Müller, N., Treiman, A., Elkins-Tanton, L., Helbert, J., Piccioni, G., Drossart, P., 2010. Recent hotspot volcanism on Venus from VIRTIS emissivity data. *Science* 328 (5978), 605–608.
- Tashkun, S. A., Perevalov, V. I., Teffo, J.-L., Bykov, A. D., Lavrentieva, N. N., 2003. CDS-1000, the high-temperature carbon dioxide spectroscopic databank. *J. Quant. Spectrosc. Rad.* 82 (1-4), 165–196.
- Tellmann, S., Pätzold, M., Häusler, B., Bird, M. K., Tyler, G. L., 2009. Structure of the Venus neutral atmosphere as observed by the radio science experiment VeRa on Venus Express. *J. Geophys. Res.* 114 (E9), E00B36.
- 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., Moissl, R., 2012. Morphology of the cloud tops as observed by the Venus Express Monitoring Camera. *Icarus* 217 (2), 682–701, *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., Calcutt, S. B., 2009. Variability of CO concentrations in the Venus troposphere from Venus Express/VIRTIS using a band ratio technique. *Icarus* 201 (2), 432–443.
- Wiscombe, W. J., 1980. Improved Mie scattering algorithms. *Appl. Opt.* 19 (9), 1505–1509.