https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

Emissivity Retrievals with FORUM's End-to-end Simulator: Challenges and Recommendations

Maya Ben-Yami[1], Hilke Oetjen[1], Helen Brindley[4], William Cossich[3], Dulce Lajas[1], Tiziano Maestri[3], Davide Magurno[3], Piera Raspollini[5], Luca Sgheri[2], and Laura Warwick[4]

[1]ESA – ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[2]IAC – CNR, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy
[3]Università di Bologna – Dipartimento di Fisica e Astronomia, Viale Berti Pichat 6/2, 40126 Bologna, Italy
[4]Space and Atmospheric Physics Group, Department of Physics, Imperial College London, SW7 2AZ, United Kingdom
[5]IFAC – CNR, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy

Correspondence: Maya Ben-Yami (mayayami@pik-potsdam.de)

Abstract. Spectral emissivity is a key property of the Earth surface of which only very few measurements exist so far in the far-infrared (FIR) spectral region, even though recent work has shown its FIR contribution is important for accurate modelling of global climate. The European Space Agency's 9th Earth Explorer, FORUM (Far-infrared Outgoing Radiation Understanding and Monitoring) will provide the first global spectrally resolved measurements of the Earth's top-of-the-atmosphere (TOA) spectrum in the FIR. In clear-sky conditions with low water vapour content, these measurements will provide a unique opportunity to retrieve spectrally resolved FIR surface emissivity. In preparation for the FORUM mission with an expected launch in 2026, this study takes the first steps towards the development of an operational emissivity retrieval for FORUM by investigating the sensitivity of the emissivity product of a full spectrum optimal estimation retrieval method to different physical and operational parameters. The tool used for the sensitivity tests is the FORUM mission's end-to-end simulator. These tests show that spectral emissivity of most surface types can be retrieved for dry scenes in the 350-600 cm^{-1} region with an uncertainty ranging from 0.005 to 0.01. In addition, the quality of retrieval is quantified with respect to the precipitable water vapour content of the scene, and the uncertainty caused by the correlation of emissivity with surface temperature is investigated. Two main recommendations are made based on these investigations: (1) As the extent of TOA sensitivity to the surface in the FIR depends on the atmospheric state, the spectral region of the emissivity product should be decided using a so-called information quantifier, calculated from the ratio of the retrieval uncertainty to the a-priori uncertainty. (2) Depending on retrieval input parameters, the correlation of emissivity with surface temperature allows for retrieved emissivities within a small range around the true emissivity. Thus the impact of this correlation on the uncertainty estimates of the product should be quantified in detail during further development of the operational retrieval.

1 Introduction

The European Space Agency's 9th Earth Explorer, FORUM (Far-infrared Outgoing Radiation Understanding and Monitoring, (Palchetti et al., 2020)) is scheduled to launch in 2026. FORUM will provide spectrally resolved measurements of the Earth's

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

top-of-the-atmosphere (TOA) outgoing radiation from 100 cm^{-1} to 1600 cm^{-1}, with the goal of filling the observational gap in the far-infrared (FIR, defined here as below 667 cm^{-1}). Even though simulations suggest that around 50% of the outgoing longwave radiation (OLR) to space is in the FIR in the global mean, due to technical reasons it has never been observed, spectrally resolved, in its entirety. FORUM's novel measurements will be provided by the mission's core instrument, a nadir-viewing Fourier Transform Spectrometer (FTS), which will measure the Earth's upwelling spectral radiance. While the primary goal of FORUM is to provide these calibrated spectral radiances, its further aim is to exploit instantaneous radiance observations to retrieve atmospheric and surface properties (Level 2 products). This work focuses on the retrieval objectives of FORUM in the case of clear skies, in particular on the retrieval of FIR surface emissivity.

FORUM clear sky radiances will be used to retrieve temperature and water vapour profiles as well as surface emissivity and surface temperature. Surface emissivity is the material property determining how much thermal radiation a surface emits at a given temperature and for a surface (or skin) temperature T_s it is defined as the ratio of surface emission to the Blackbody emission at T_s. Emissivity is not constant across the spectrum, and the emissivities of different surfaces exhibit distinct spectral variation. The possibility to retrieve spectrally resolved FIR emissivity is particularly exciting given the absence of other instrumentation capable of providing global coverage across the FIR and its potential influence on the surface and top of atmosphere energy budget (Feldman et al., 2014; Kuo et al., 2018).

Surface emissivity across the globe is routinely retrieved in the mid-infrared (MIR) from satellite observations (Susskind et al., 2014; Capelle et al., 2012; Masiello and Serio, 2013; Wan, 2014; Wang et al., 2005). These are complemented by laboratory measurement datasets such as the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Spectral Library, which includes more than 2300 different spectral emissivities down to 650 cm^{-1} (Baldridge et al., 2009). Due to the absence of spectrally resolved TOA radiance observations in the FIR no global retrievals of surface emissivity are available in the FIR, and there is also a lack of laboratory measurements. Bellisario et al. (2017) and Murray et al. (2020) were the first to retrieve FIR snow emissivities from aircraft measurements (during the CIRCCREX/COSMICS projects over Greenland), and there are also studies under way to measure the emissivity of snow from ground-based measurements (see Palchetti et al. (2021)). But while the work of Murray et al. (2020) confirms the feasibility of retrieving FIR surface emissivity from OLR spectral measurements, agreement between their work and the earlier Bellisario et al. (2017) study implies that no theoretical snow/ice model could fit their retrieved values in the MIR and FIR simultaneously. This indicates that further testing of the theoretical models using global emissivity retrievals is vital to extending surface emissivity datasets into the FIR.

The potential value in knowing the spectral variation of surface emissivity in the FIR is significant. In recent years there has been increasing focus on the inadequate representation of surface emissivity in global climate models (GCMs), which almost all assume Black-body or Gray-body emissivity (Huang et al., 2018). To test the validity of this assumption in the FIR, Feldman et al. (2014) incorporated spectrally varying FIR emissivity into the Community Earth System Model I (CESM I) and showed significant changes in its predictions after 25 years: at high latitudes as much as a 2 K change in surface temperature and 10 Wm^{-2} in the outgoing longwave radiation. The authors also identified a possible feedback mechanism associated with FIR emissivity: in the FIR the emissivity of snow can be substantially higher than that of water (while in the MIR the difference is less significant, see Figure 1). This means that as sea ice melts in a warming climate it exposes a potentially less emissive

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

water surface, exacerbating the warming. Further work with the CESM has confirmed that this feedback is present, if small, and has shown that the inclusion of realistic surface emissivity in fact significantly reduces the persistent cold-pole bias of climate models (Kuo et al., 2018; Huang et al., 2018). Critically, by comparing the assumption of ice vs snow emissivity in the models it was shown that the size and sign of the feedback depends on the properties of the surface (Huang et al., 2018).

Much work has already been done to analyse the performance of the geophysical products (including emissivity) expected from FORUM clear-sky measurements (e.g. Ridolfi et al., 2020; Sgheri et al., 2021). In this study we focus on spectral surface emissivity in particular, and investigate its retrieval using the FORUM mission's end-to-end simulator (FEES) described in Sgheri et al. (2021). As this work is meant to provide the first steps towards the development of an operational emissivity retrieval for FORUM the focus is placed on investigating the effect of various factors on the retrieval of a range of typical scenes. The aim of an operational retrieval is to provide the users with a retrieved product that is transparent and accessible, and thus the focus of this work is not on extreme cases or on optimizing the retrieval for specific scenes, but rather on highlighting general features which need to be investigated.

The paper is structured as follows: Section 2 describes the FEES, and Section 3 the experimental set-up. In Section 4 the general FEES retrieval result is introduced together with the different quantifiers used for its analysis. The retrieval quality is compared against scene pwv content in Section 5. Section 6 introduces the correlation of surface emissivity with surface temperature, investigates its consequences and analyses its behaviour. Finally, Section 7 summarises the results, focusing on the main challenges and on the recommendations this study has for further development towards an operational emissivity retrieval for FORUM. Appendix A looks at the emissivity-T_s parameter space in more detail, and in Appendix B the choice of the emissivity a-priori uncertainty is investigated.

2 The FORUM End-to-end simulator and the Optimal Estimation method

The FORUM mission's end-to-end simulator (FEES) constitutes a chain of modules which simulate the elements relevant to the mission performance. A full description of the FEES can be found in Sgheri et al. (2021), together with a discussion of the geophysical products not shown in this work. Our study uses the first five modules of the simulator: The Geometry Module (GM), the Scene Generator Module (SGM), the FORUM Sounding Instrument (FSI) Module, the FORUM Embedded Imager (FEI) Module and the Level 2 Module (L2M). For the purpose of this work it is enough to note that when the first four modules are run in the default chain (see Sgheri et al. (2021)) they generate synthetic FORUM observations. This is what is done in this work for various geographic scenes in clear sky conditions. The L2M uses these synthetic observations to retrieve the geophysical properties of the scene, and in this work this retrieval algorithm is tested with a focus on the retrieved spectral surface emissivity.

The L2M retrieves the atmospheric state from the synthetic FORUM measurements using the Optimal Estimation (OE) method, which deals with the ill-posed nature of the inverse problem using an a-priori regularization (Rodgers, 1976, 2000). Starting from an initial guess of the n-dimensional atmospheric state vector $\mathbf{x}$, the algorithm arrives at a best estimate $\hat{\mathbf{x}}$ by

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

minimising the cost function ξ^2:

$$\xi^2(\mathbf{x}) = (\mathbf{y} - \mathbf{f}(\mathbf{x}))^T {\mathbf{S_y}}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{x})) + ((\mathbf{x_a} - \mathbf{x})^T {\mathbf{S_a}}^{-1} (\mathbf{x_a} - \mathbf{x})) \quad (1)$$

The first term on the right-hand side is the χ^2 of the forward model, in essence the difference between the m-dimensional observation vector $\mathbf{y}$ and the forward model $\mathbf{f}(\mathbf{x})$ calculated from the atmospheric state vector $\mathbf{x}$, where the covariance matrix $\mathbf{S_y}$ represents the uncertainty on the observations. The second term is the regularization term, which takes into account the difference of the state vector $\mathbf{x}$ from an a-priori (model) atmospheric state $\mathbf{x_a}$ with uncertainty $\mathbf{S_a}$. In this work the retrieved atmospheric state vector constitutes the atmospheric water vapour profile, the temperature profile, the spectral surface emissivity and the surface temperature. For more details on the forward model and minimization technique see Sgheri et al. (2021, 2020), and for the parameters and assumptions used in this work see Section 3.

To understand the parameters influencing the quality of the retrieved emissivity it is useful to keep in mind the role emissivity plays in the forward model, which is in the simulation of the atmospheric radiative transfer. For nadir-looking observations the clear-sky TOA spectral radiance $S_{\text{toa},\sigma}$ at wavenumber σ can be written as:

$$S_{\text{toa},\sigma} = S_{\text{surf},\sigma} \mathcal{T}_\sigma(z_1) + \int_{z_0}^{z_1} B_\sigma(T(z)) \frac{\partial \mathcal{T}_\sigma(z))}{\partial z} dz \quad (2)$$

where $B(T)$ is the Planck function, $T(z)$ is the atmospheric temperature profile, $\mathcal{T}(z)$ is the transmittance between the surface and height z, and the integral is over the height z from the surface z_0 to the TOA z_1. $\mathcal{T}(z_1)$ is the transmittance from the surface to the TOA. The emissivity contributes to the surface part of the radiance:

$$S_{\text{surf},\sigma} = L_{d,\sigma}(1 - \epsilon_\sigma) + \epsilon_\sigma B_\sigma(T_s) \quad (3)$$

Here $L_{d,\sigma}$ is the downwelling radiance at the surface, T_s is the surface (or skin) temperature, and ϵ_σ is the emissivity of the surface at wavenumber σ. In this work the emissivity is always assumed to have no directional dependence.

Equations 2 and 3 show that the value for the surface emissivity at one wavenumber, although in itself only influencing the OLR at that wavenumber, depends in different ways on the surface temperature and on the full water vapour and temperature profiles and through them depends on a large range of wavenumbers. This work attempts to disentangle some of these dependencies and investigate what effect they have on the uncertainties of the OE method used and whether the retrieval parameters can be optimised to reduce these uncertainties.

3 Experimental Set-up

In this work a range of parameters are modified in the FEES to investigate their effect on retrieval quality:

- the water vapour profile in the SGM (Sections 4 and 5)
- the instrumental random noise in the FSI module (Sections 4 and 6)

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

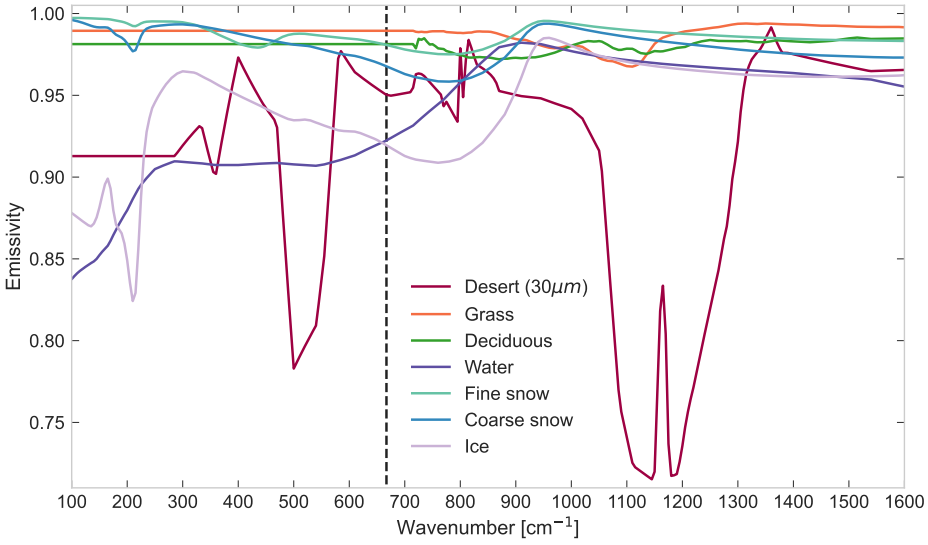


Figure 1. Spectral surface emissivity between 100 and 1600 cm^{-1} for seven out of the 11 surface types in Huang et al. (2016). Desert subtype is $r_e = 30\mu m$ as used in the FEES. The far-infrared is defined in this work to the left of the dashed line.

– the retrieval parameters in the L2M associated with surface temperature and emissivity: the value of the retrieval a-priori, the a-priori uncertainty and the initial guess (Section 6)

This section describes the experimental set-up of the FEES in more detail, and defines a baseline retrieval scenario as a basis for these modifications.

3.1 FEES modules

All the results presented in this work are the products of FEES runs. A complete description of this simulator and its modules can be found in Sgheri et al. (2021) and unless otherwise stated the same parameters and settings are used as described in that work for homogeneous clear sky cases.

Only the first five modules of the FEES are used in this work. No modification is made to the Geometry Module. The next module is the Scene Generator Module (SGM), which uses geographic coordinates to compute high-resolution TOA spectral radiances in clear sky conditions using the radiative transfer model LBLRTM version 12.8 (Clough et al., 2005) and auxiliary databases prepared for the FEES. For a detailed description of the auxiliary datasets see Sgheri et al. (2021), but for reference note that the water vapour and temperature profiles and the surface temperature are taken from ERA5 reanalysis data (Hersbach et al., 2020). In this work all scenes used are from 15 January 2018 12:00:00 UTC for consistency, and they are identified using their geographic coordinates (see Table 1). The emissivity dataset used by the SGM is based on the Huang et al. (2016) geolocated dataset of spectral emissivity and uses the 11 surface types defined by Huang et al. (2016) (out of the multiple Desert subtypes the $r_e = 30$ µm subtype is used). Each scene is generated using the surface type out of these 11 that

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

best matches the January value in the geolocated dataset for the given coordinates. Seven of these 11 surface types can be seen in Figure 1.

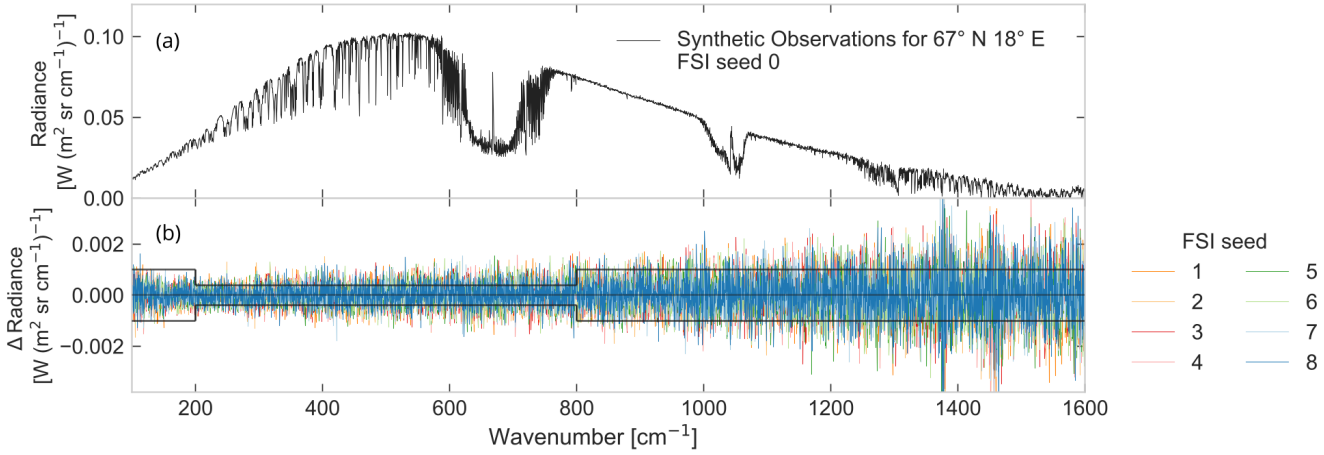


Figure 2. Figure (a) shows an example of synthetic observations generated by the first four modules of the FEES for the scene at 67° N 18° E on 15 January 2018 12:00. Figure (b) shows the difference in radiance between spectra generated with FORUM Sounding Instrument (FSI) random noise seeds from 1-8 and the spectrum generated with seed 0 (shown in Figure (a)). The FORUM goal noise equivalent spectral radiance (NESR) is shown as the black lines in Figure (b).

The third and fourth modules of the FEES simulate the observing system, and in this work are used without modification for the same instrument concept. The only change made is to vary a so-called seed used to generate the random noise associated with the FORUM Sounding Instrument (FSI). The synthetic observations thus generated and the variation of random noise is illustrated in Figure 2.

The final module used is the L2M, which has been described in more detail above in Section 2. This is the module used to test emissivity retrieval properties and in which the major modifications were made.

3.2 The baseline retrieval parameters

For the purpose of this study we define a baseline/default retrieval case, which is used as the basis for all modifications and tests. Unless otherwise stated, all parameters are the standard parameters for clear sky retrievals in Sgheri et al. (2021). Only two parameters differ between the baseline retrieval in this study and the standard of Sgheri et al. (2021): for the emissivity a-priori, this work uses a flat a-priori value instead of a perturbed climatological one, and a 0.1 uncertainty instead of 0.05 (see Sections 3.3 and Appendix B for a justification of these choices).

For comparison with later modifications some of these baseline parameters are listed:

- Emissivity initial guess: constant and equal to 1
- Emissivity a-priori: constant and equal to 1

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

Coordinates	Surface Temperature T_s [K]	$T_0 - T_s$ [K]	pwv [mm]	Surface Height [m]	Surface Type
21° N 15° E	307.1	12.7	7.31	516	Desert
21° N 18° E	313.5	17.9	3.05	1513	Grass
25° N 09° E	302.4	15.0	2.24	1415	45% desert and 55% grass
47° N 25° E	267.3	2.9	1.87	1022	Deciduous
55° N 20° E	271.9	4.6	4.09	8	Water
66° N 17° E	264.6	-0.4	4.14	572	Fine snow
67° N 18° E	262.4	-0.7	3.55	755	Fine snow
67° N 29° E	266.5	0.1	4.8	261	Coarse snow
71° N 29° E	278.5	5.8	4.07	38	Conifer

Table 1. Atmospheric and surface data for the various scenes used in this study. All data is for 15 January 2018 at 12:00:00. Surface temperatures, surface heights, water vapour profiles and temperature profiles are from ERA5 reanalysis data. Surface types are fitted to the Huang et al. (2016) dataset as detailed in the text. Pwv stands for precipitable water vapour, and $T_0 - T_s$ is the difference between the lowest point of the temperature profile and the surface temperature

- Emissivity a-priori uncertainty matrix: defined using uncertainty $\Delta\epsilon = 0.1$ and correlation length (CL) of 50 cm^{-1} (see Appendix B for an explanation of these terms and a justification of the choice of uncertainty matrix)
- Emissivity retrieval grid: evenly spaced 5 cm^{-1} grid for the full FORUM spectral range
- Surface temperature initial guess: climatological value from ERA5 monthly averages (different from the daily value used for the SGM)
- Surface temperature a-priori: a random perturbation of the true value with a 2 K standard deviation (the perturbation is the same for the same geographical scene)
- Surface temperature a-priori uncertainty: 2 K

In the baseline retrieval the same instrumental noise is used for all cases (i.e. the seed used to generate the instrumental random noise is kept the same at a value of 0, see Figure 2).

To test the retrieval of surface emissivity in the FIR, we choose to use geographic scenes with low precipitable water vapour (pwv), which is defined as the depth of water produced if all water in the atmospheric column precipitated as rain. For reference, the full list of scenes used in the tests shown in this work can be found in Table 1, together with some of their relevant atmospheric and surface properties.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

3.3 The emissivity a-priori

The probabilistic interpretation of the Optimal Estimation method from Rodgers (2000) is in the Bayesian frame, wherein the solution represents the estimate of maximum a-posteriori probability. In this interpretation, $\mathbf{x}_a$ represents our a-priori knowledge of the atmospheric state and $\mathbf{S}_a$ our uncertainty in that estimate (assuming Gaussian error distributions). However the formalism of the method can be used without giving a probabilistic interpretation to $\mathbf{x}_a$ and $\mathbf{S}_a$ and simply tuning them to best regularize the retrieval (von Clarmann et al., 2020). For example, the smaller the uncertainties in $\mathbf{S}_a$ are, the closer the solution will be on average to $\mathbf{x}_a$ - this can be thought of as giving the retrieval more or less freedom to converge to the true state.

In this work we choose not to use climatological datasets for the emissivity a-priori. As can be seen from Figure 1, if the underlying surface type is mis-categorised and for example is in fact grass instead of desert, the difference in emissivity can be on the order of 0.3. Therefore the emissivity part of $\mathbf{S}_a$ must be such that the retrieval has the freedom to converge to any value between 0.7 and 1, which covers almost all the possible theoretical emissivity model values in the spectral region considered. Introducing fine a-priori structure is of little use if the retrieval has such a range of freedom. Therefore for the purposes of this work a *flat* a-priori of a constant value is chosen, as it ensures consistency across cases and allows for easier comparison of different retrieval setups. In addition Appendix B shows an analysis of the effect of varying the parameters defining the emissivity submatrix of $\mathbf{S}_a$. An uncertainty $\Delta\epsilon = 0.1$ and correlation length (CL) of 50 cm^{-1} (see Appendix B for a definition) are chosen to calculate the emissivity submatrix of $\mathbf{S}_a$, as these provide the freedom needed and represent a compromise between accuracy and sensitivity.

4 The emissivity product, its quantifiers and water vapour

The retrieval process can only give information on a retrieved quantity where the forward model $\mathbf{f}(\hat{\mathbf{x}})$ is sensitive to this quantity, whereas it returns the a priori where there is no sensitivity. In this work the analysis of the quality of the retrieved surface emissivity is developed to only consider the retrieved values where they are in fact retrieved and not purely drawn from the a-priori. To this end four different quantifiers are defined in this section, and their behaviour for varying water vapour content is analysed. This analysis is used to introduce a criterion for deciding on the spectral region where emissivity is retrieved.

TOA measurements are only sensitive to surface emissivity where there is non-zero transmission from the surface to the TOA, and so the sensitivity of the retrieval to emissivity is dependent on the atmospheric transmission. The reason for the focus on water vapour in this Section and Section 5 is that in the FIR it is one of the most important factors influencing the transmission (Harries et al., 2008), as the water vapour rotational band dominates the atmospheric absorption in this region. In fact for high water vapour content the TOA is opaque to the surface in the FIR, but becomes more transparent as the atmosphere gets drier. The distinct characteristic of atmospheric transmission in the FIR is that as the pwv decreases, transmittance does not increase uniformly but in so-called *microwindows*, which become deeper as pwv decreases.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

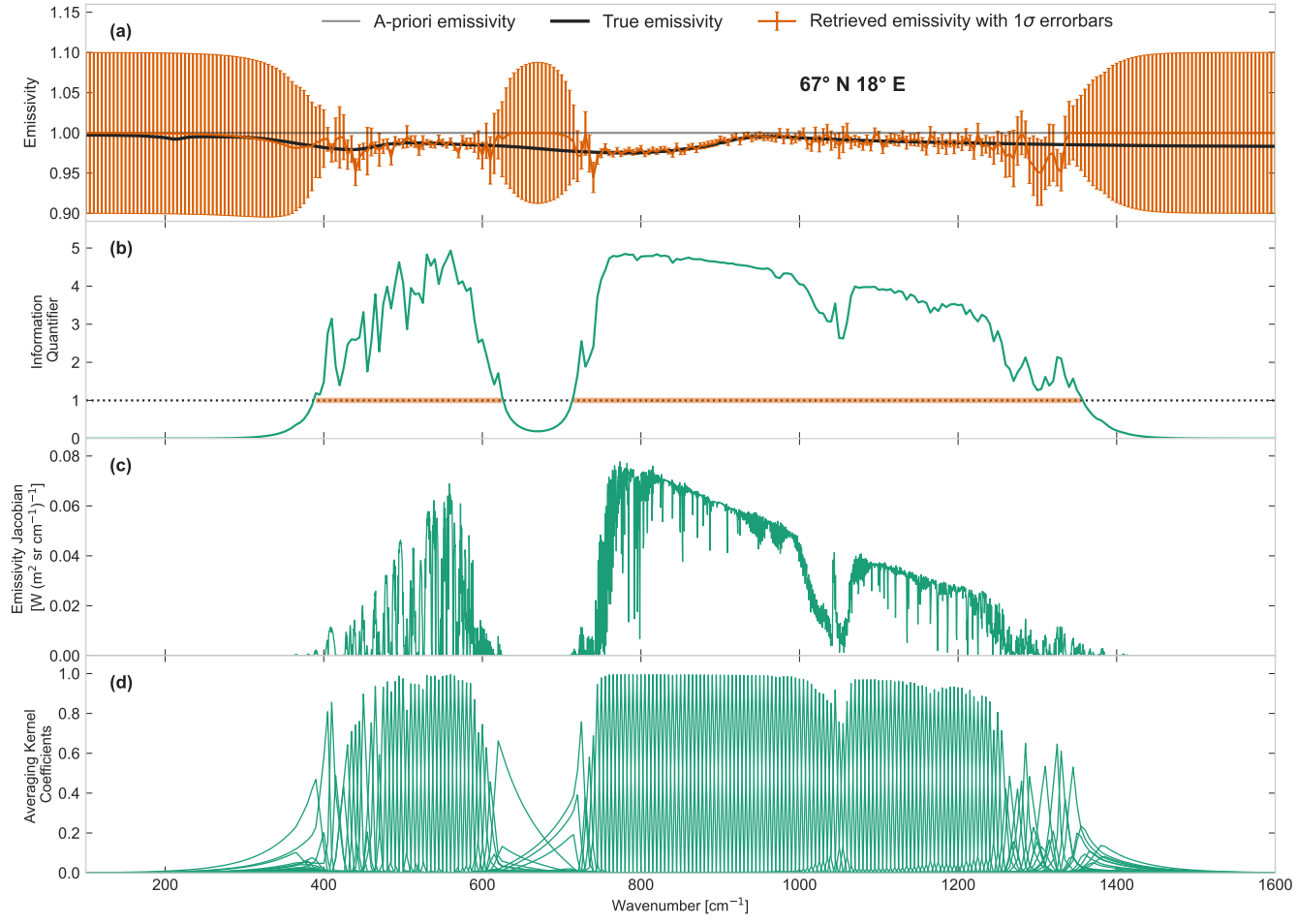


Figure 3. The retrieved emissivity and four quantifiers used to assess the retrieval quality. The retrieved scene is of fine snow emissivity at 67°N 18°E 15 January 2018 12:00, with default parameters as outlined in Section 3. Figure (a) shows the retrieved emissivity with 1σ retrieval uncertainty errorbars in orange, the true fine snow emissivity from Huang et al. (2016) in black and the a-priori emissivity in grey. Figure (b) shows the Information Quantifier (IQ) for this retrieval, defined in Equation 5. Figure (c) shows the diagonal elements of the spectral-resolution emissivity Jacobian of the radiative transfer part of the retrieval forward model at convergence, defined in Equation 6. In Figure (d) the rows of the emissivity submatrix of the averaging kernel (defined in Equation 7) are plotted, all in the same color. The orange bars in figure (b) show the spectral regions in which the IQ is larger than 1.

4.1 The quantifiers

Four different quantifiers are described in the following Section to illustrate retrieval quality. These are shown in Figure 3 for the baseline retrieval of the scene at 67° N 18° E.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

Figure 3(a) shows the baseline retrieved emissivity, which is the emissivity part of the best estimate atmospheric state vector $\hat{\mathbf{x}}$ introduced in Section 2. Note that the emissivity is retrieved on a $5\ \mathrm{cm}^{-1}$ spectral grid which is much coarser than the $\sim 0.4\ \mathrm{cm}^{-1}$ resolution of the synthetic observations, and thus the emissivity ϵ_σ used in the atmospheric radiative transfer calculations of the forward model (Equation 3) is in fact a linear interpolation of the emissivity elements of the retrieval vector $\hat{\mathbf{x}}$.

The first quantifier is the retrieval uncertainty, shown as the error-bars in Figure 3(a). These are derived from the retrieval uncertainty covariance matrix $\mathbf{S_x}$, defined as in Rodgers (2000):

$$\mathbf{S_x} = (\mathbf{K}^T \mathbf{S_y}^{-1} \mathbf{K} + \mathbf{S_a}^{-1})^{-1} \tag{4}$$

where $\mathbf{S_y}$ and $\mathbf{S_a}$ are as in Equation 1, and $\mathbf{K}$ is the jacobian of the full forward model at convergence with respect to the retrieval vector. The retrieval standard deviation σ_x is the square root of the diagonal of $\mathbf{S_x}$. σ_x is called the retrieval uncertainty in this work, whilst the systematic uncertainty is defined as the true value minus the retrieved value.

The second quantifier shown in Figure 3(b) makes further use of the information contained in σ_x, in particular that in regions where there is no sensitivity the retrieval vector will equal the a-priori and the σ_x will equal the a-priori uncertainty σ_a. Recognising this, Dinelli et al. (2009) defined the information quantifier (IQ) as:

$$\mathrm{IQ} = -\frac{1}{2}\log_2\left(\frac{\sigma_x}{\sigma_a}\right) \tag{5}$$

where σ_x is as above and σ_a is the square root of the diagonal of the retrieval a-priori covariance matrix $\mathbf{S_a}$. The IQ thus tends to 0 in regions with low sensitivity as σ_x approaches σ_a. Note that while the IQ can be defined for the full retrieval vector, in this work it is only used for the retrieved emissivity.

The third quantifier is the Jacobian of the TOA radiances with respect to emissivity. Whilst Equation 4 uses the Jacobian $\mathbf{K}$ with respect to the full forward model, to directly quantify the emissivity retrieval quality a different Jacobian $\mathbf{J}$ is used, which is calculated with respect to the radiative transfer simulation at convergence:

$$J_{ij} = \frac{\partial F_{\mathrm{toa},\sigma_i}}{\partial \epsilon_{\sigma_j}} \tag{6}$$

where ϵ_{σ_j} are the emissivity values used in the radiative transfer calculations of LBLRTM and $F_{\mathrm{toa},\sigma_i}$ are the resulting TOA radiances at wavenumbers σ_i and σ_j (see Equations 2 and 3 for their physical definitions). From Equation 3 we can see that at the measurement spectral resolution J_{ij} is diagonal in the emissivity, and so the diagonal J_{ii} values are plotted in Figure 3(c).

The final quantifier is the averaging kernel $\mathbf{A}$, which is frequently used to evaluate OE retrievals (see Rodgers (2000); von Clarmann et al. (2020)) and gives more information on the retrieval process itself. $\mathbf{A}$ is defined as the derivative of the retrieved atmospheric state vector $\hat{\mathbf{x}}$ with respect to the true state vector $\mathbf{x}$ (where $\mathbf{x}$ is the interpolation of the true atmospheric components onto their respective retrieval grids):

$$A_{ij} = \frac{\partial \hat{x_i}}{\partial x_j} \tag{7}$$

Considering the diagonal submatrix of $\mathbf{A}$ that corresponds to emissivity in the retrieval vector, the rows of that submatrix represent the sensitivity of the retrieved emissivity at a particular wavenumber to the true emissivity at all wavenumbers. These

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

emissivity submatrix rows are plotted in Figure 3(d). **A** approaches the identity matrix **I** when the contribution of the a-priori is negligible with respect to the measurements.

The scene shown in Figure 3 has a pwv content of 3.55 mm and so as discussed above its retrieval is sensitivite to the surface in the FIR. The quantifiers in Figures 3(b)-(d) thus show the distinct pattern of the TOA's sensitivity to the surface in such dry atmospheric scenes:

- Significant transmission in the FIR below the CO_2. This is the so called *dirty window* of the water vapour rotational band where emission is still strong but the transmission is in microwindows. The microwindow structure can clearly be seen in the jacobian and is also reflected in the varying strength of the averaging kernel
- Low sensitivity below 400 cm^{-1} as the absorption of the water vapour rotational band increases
- Uniform transmittance in the MIR atmospheric window, resulting in an averaging kernel close to 1
- A small decrease in sensitivity in the ozone band around 1000 cm^{-1}
- Decreasing sensitivity at MIR wavenumbers higher than 1200 cm^{-1} because of a combined increase of noise in the measurements and absorption by water vapour
- No sensitivity in the CO_2 band between roughly 600 and 750 cm^{-1}

4.2 Spectral quantifiers and water vapour content

These quantifiers can also be used to investigate how the retrieval quality changes across the spectral range as the pwv content of the atmosphere is varied. To this end the pwv content of the scene at 67° N 18° E was modified by multiplying its climatological water vapour profile by a range of constant factors and generating synthetic observations from these modified scenes (resulting in pwv content ranging from 0.4 to 17.8 mm). The baseline retrieval is run for six such modified scenes, and the retrievals and their quantifiers are shown in Figure 4. The only difference from Figure 3 is that for clarity in subplot (a), instead of using error bars the 1σ retrieval uncertainty in emissivity is shown as a shaded region.

Figure 4 shows that the basic spectral characteristics of the quantifiers described above do not change when the pwv is varied. However while changing the water vapour content does not significantly affect the retrieval in the MIR, it is an important factor determining the sensitivity to emissivity in the FIR.

The Jacobians in Figure 4(c) show that while the transmission maintains its microwindow structure in the FIR, these windows gradually weaken and disappear as the pwv content is raised. This is reflected in the averaging kernels in Figure 4(d), where at low pwv the retrieved emissivity in the FIR has high sensitivity to the true value, but this sensitivity decreases to almost 0 for the highest pwv content. The consequence of this change for the retrieval result itself can clearly be seen in Figure 4(a). As noted above, where there is no sensitivity to the true value the retrieval uncertainty will approach the a-priori uncertainty (here 0.1), and Figure 4(a) shows this: for dry scenes there is small retrieval uncertainty as low as 300 cm^{-1}, whilst at high pwv the retrieval uncertainty is equal to the large a-priori uncertainty value through most of the FIR. Thus the spectral region where the emissivity values *are in fact retrieved* changes depending on the pwv.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

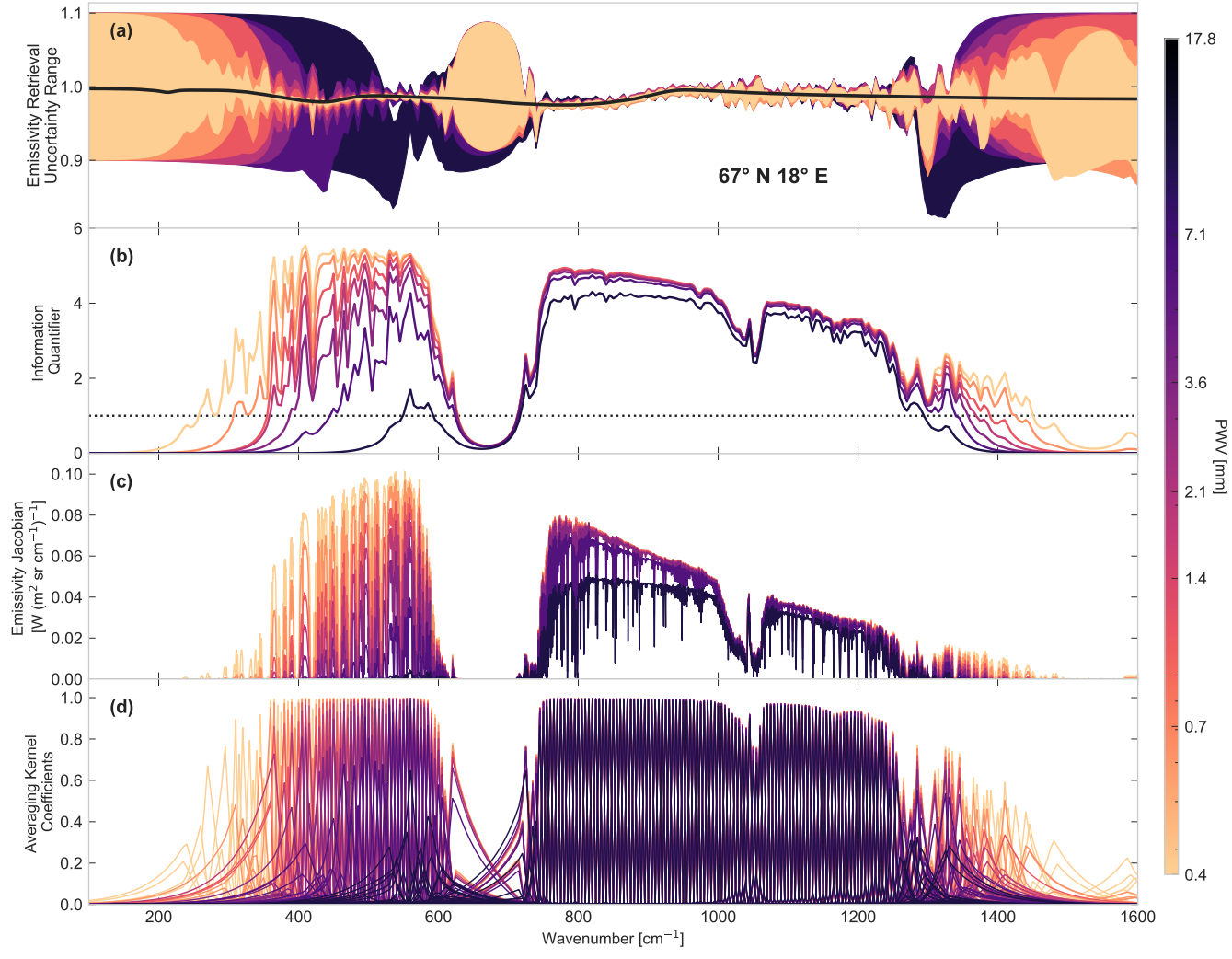


Figure 4. The same retrieval as in Figure 3 is run for scene 67°N 18°E with modified precipitable water vapour (pwv) content, and the same quantifiers are shown as in Figure 3. Plot (a) shows the $\pm 1\sigma$ emissivity retrieval uncertainty range as a shaded colored region as well as the true emissivity as a black line, and plots (b)-(d) are the same as in Figure 3. The colors from dark to light indicate the true pwv content of the retrieved scene from high to low, and the exact pwv values are marked on the color scale on the right of the Figure.

4.3 Retrieved emissivity criterion

In this work a criterion is defined to produce an *emissivity product* from the retrieval vector that represents only the retrieved emissivity values that contain information on the true emissivity. The IQ is well suited for use as part of a sensitivity criterion as it is defined to show which retrieved values contain information on the true state and which do not. Figure 4(b) shows that it reflects the variations in sensitivity caused by variations in pwv content. And unlike the jacobian or the averaging kernel, the

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

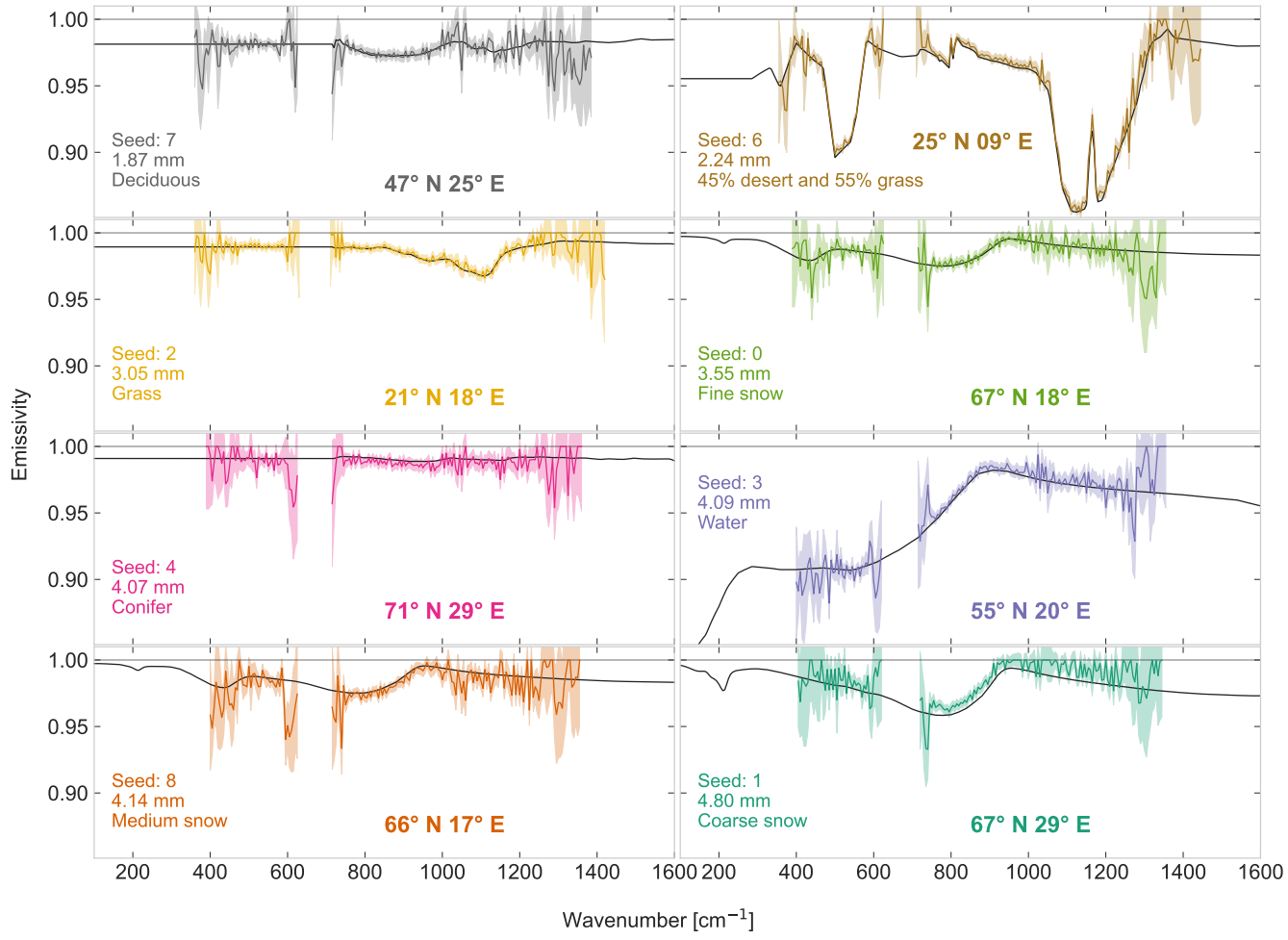


Figure 5. Retrieved emissivity for eight scenes with various surface types. The retrieved emissivity is shown as a dark colored line, with the 1σ retrieval uncertainty range shown as a shaded region of the same color. The true emissivity is shown as a solid black line in each Figure, and the a-priori emissivity is in grey. All retrievals use the baseline parameters from Section 3. Each scene uses atmospheric and surface data from the coordinates indicated on the Figure for 15 January 2018 at 12:00. More information for each scene is listed in Table 1 - the true precipitable water vapour content and surface type is indicated on the figures. Each retrieval uses synthetic observations generated with a different random instrumental noise seed to mirror true retrieval conditions, and the seed is indicated on the corresponding figure. The retrieved emissivity is only shown in the spectral regions in which the Information Quantifier (see Equation 5) is larger than 1.

IQ is smooth enough to lend itself to a threshold criterion. In this work, a value of 1 is chosen as such a threshold for the IQ following Dinelli et al. (2009). Thus for all further sections of this work, the emissivity range considered as retrieved (i.e. in regions of sensitivity) is that for which:

$$\mathrm{IQ} > 1 \tag{8}$$

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

The possibility to produce an emissivity product for a range of cases using this criterion is illustrated in Figure 5, which shows such a retrieved emissivity product for eight dry geographic scenes with different surface emissivities. Only scenes with pwv below 5 mm are shown here to demonstrate the viability of FORUM FIR emissivity retrievals. For FEES emissivity retrievals of scenes with pwv higher than 5 mm see Sgheri et al. (2021). As already seen in Figure 3, the emissivity in dry scenes is retrieved in two sections above and below the CO_2 band, with the uncertainty in the retrieval highest in the edge regions of these sections. Figure 5 thus illustrates the potential of FORUM to retrieve FIR emissivity for a range of surface types and locations on the globe. Finally, note that the criterion defined in Equation 8 is flexible: if there is a requirement on the uncertainty level of the retrieval (for example if only the *cleanest* data is needed), the choice of IQ boundary for the criterion can be tuned to shorten or extend these boundary regions.

5 Impact on retrieval quality by precipitable water vapour

pwv [mm]	0.4	0.7	1.4	1.8	2.1	2.5	2.8	3.2	3.6	3.9	4.3	4.6	5.0	7.1	10.7	17.8
extent [cm^{-1}]	260	310	360	360	360	360	380	385	390	395	400	400	400	445	480	550

Table 2. Data points from Figure 6 for case 67° N 18° E. Pwv is the true precipitable water vapour content of the retrieved scene and extent is the minimum wavenumber at which the retrieved emissivity satisfies the criterion in Equation 8.

In this section the analysis of the variation in the retrieval quality with water vapour content shown in Figure 4 for the scene at 67° N 18° E is extended and compared for multiple geographic scenes. The procedure for modifying the pwv content is identical: leaving all other atmospheric and surface properties untouched, the climatological water vapour profile of the scene was multiplied by a constant value (ranging from 0.05 to 120). The four scenes were 25°N 09°E, 21°N 15°E, 67°N 18°E and 67°N 29°E, and the corresponding maximum and minimum water vapour profiles used can be seen in Figure 6(a). The synthetic observations generated from these modified scenes were then used to run the baseline retrieval (see Section 3). Although these modified scenes included some un-physical water vapour profiles, there was no significant change in the retrieval quality of the atmospheric profiles.

In Section 4 it was seen that as pwv decreased the retrieval quality at a given FIR wavelength improved as microwindows deepened and the retrieval sensitivity extended farther into the FIR as new microwindows opened up. To complement the spectral analysis of Figure 4 and compare the variation in quality for multiple scenes in this section three single-value quantifiers are analysed for the retrievals. All three are shown in Figure 6 plotted against the true pwv content of the scene.

The first quantifier in Figure 6(b) shows the extent of the retrieval sensitivity into the FIR by plotting the minimum wavenumber which satisfies the criterion for retrieval (see Equation 8). The data for 67° N 18° E is also listed in Table 2. This wavenumber value decreases as the scene becomes drier and the weaker microwindows become transparent enough for the emissivity to be retrieved at lower wavenumbers. The second quantifier in Figure 6(c) shows the root-mean-square (RMS) error of the retrieved emissivity in the 500-600 cm^{-1} region for the cases that are fully sensitive in that region. Whilst the region in which

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

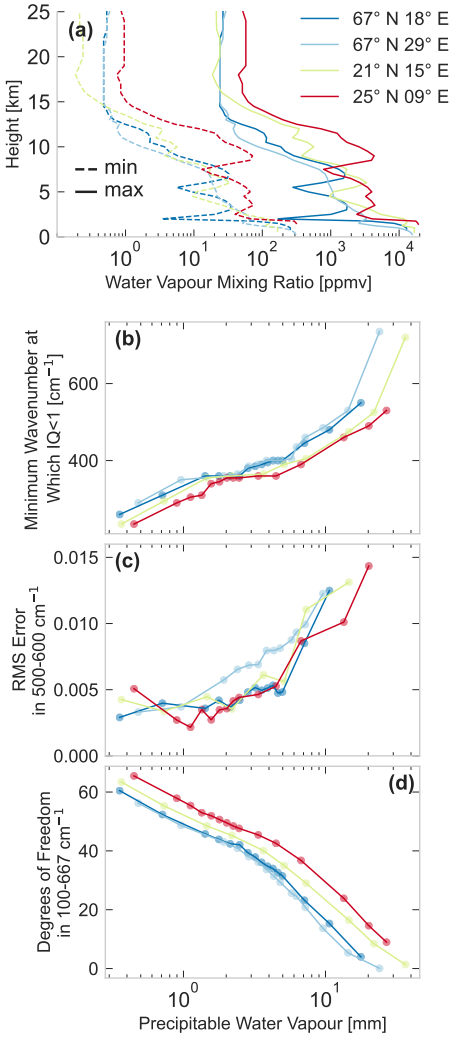


Figure 6. Retrieval quantifiers as a function of scene precipitable water vapour. For each scene the climatological water vapour profile was multiplied by a constant factor when generating the synthetic observations so as to keep everything constant except for the water vapour content. Figure (a) shows the minimum (dashed line) and maximum (full line) modified water vapour profile. Figures (b), (c) and (d) have a shared x-axis showing the precipitable water vapour of each scene, with the color-coding of the lines and markers the same as in Figure (a). Figure (b) shows the extent of the retrieval into the far-infrared using the minimum wavenumber at which the retrieved emissivity satisfies the criterion in Equation 8. Figure (c) shows the root-mean-square (RMS) error of the retrieved-true emissivity in the 500-600 cm^{-1} range for the cases with full sensitivity in that range (thus even for this conservative range the RMS is not calculated for the highest pwv values). Figure (d) shows the degrees of freedom in the 100-667 cm^{-1} range (the sum of the emissivity averaging kernel submatrix rows corresponding to that range). All three figures show the quantifiers improve as the water vapour content decreases.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

the emissivity is being retrieved in the FIR can be larger than 500-600 cm^{-1} for many of these cases, the RMS is calculated for a constant region to avoid influence from the fluctuations at the edge of the sensitive regions. Figure 6(c) shows that not only the extent of the retrieval but its quality also increases as the scene becomes drier. The final quantifier in Figure 6(d) shows the degrees of freedom of the emissivity retrieval in the full 100-667 cm^{-1} FIR region, calculated from the averaging kernel matrix. It is noteworthy that unlike the other qualifiers which have occasional plateaus in their trends, the information content in the FIR increases monotonically as the pwv decreases.

All cases individually show the same improvement in quality with pwv discussed in detail for Figure 4, and the results are only weakly dependent on the scene. However there is a small difference in the scene specific behaviour in all three plots, of which Figure 6(d) gives the clearest view. In general for the same value of pwv 25°N 09°E has the best retrieval quality, with 21°N 15°E next in quality and 67°N 18°E and 67°N 29°E lowest and about equal in quality. Although the many parameters of the atmospheric state and the small number of scenes investigated make attribution of this difference difficult, a plausible explanation can still be identified: the difference in surface temperature and surface-atmosphere contrast between these scenes. 25°N and 21°N are hot scenes (T_s > 300K, see Table 1), and their higher surface temperatures lead to a larger sensitivity to emissivity through the stronger T_s-emissivity correlation (see Section 6 and Appendix A). And though the 21°N scene surface temperature is in fact 4 K warmer than the 25°N surface temperature, the temperature contrast with respect to the atmosphere is 12.7 K and 15 K in the scenes, respectively. A larger difference between the air and surface could mean that the surface emission is easier to separate from the atmospheric emission, and would also reduce the reflected downwelling radiation. Further work should extend the analysis to a larger number of geographic scenes to better quantify this effect.

Overall the analysis of Figure 6 shows that FORUM measurements will provide significant information on emissivity in the FIR in a range of scenes. Finally, the gradual improvement in quality with pwv decrease seen across the figures supports the use of a retrieval product threshold such as the one introduced in Equation 8.

6 The correlation of surface temperature and emissivity and its consequences

The difficulty in surface emissivity retrieval caused by the connection of emissivity to surface temperature is widely recognized in the field of remote sensing (Li et al., 2013). In many cases one is only interested in either emissivity or surface temperature, but Equation 3 shows that from radiance measurements these cannot be determined independently. Even if one is only interested in the surface properties, the difficulty in Equation 3 arises from two sources: imperfect knowledge of $\mathcal{T}(z)$, the atmospheric transmittance between the surface and the instrument, and at measurement resolution the degeneracy of the surface emission itself with regards to the parameters of interest. The FEES retrieves the surface temperature and the atmospheric state that defines $\mathcal{T}(z)$ at the same time as the spectral emissivity. The main effect of $\mathcal{T}(z)$ is through the pwv content of the atmosphere which was discussed in Sections 4 and 5. This section focuses on the surface temperature on the T_s-emissivity correlation that arises from Equation 3 and investigates its impact on the retrieved emissivity.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

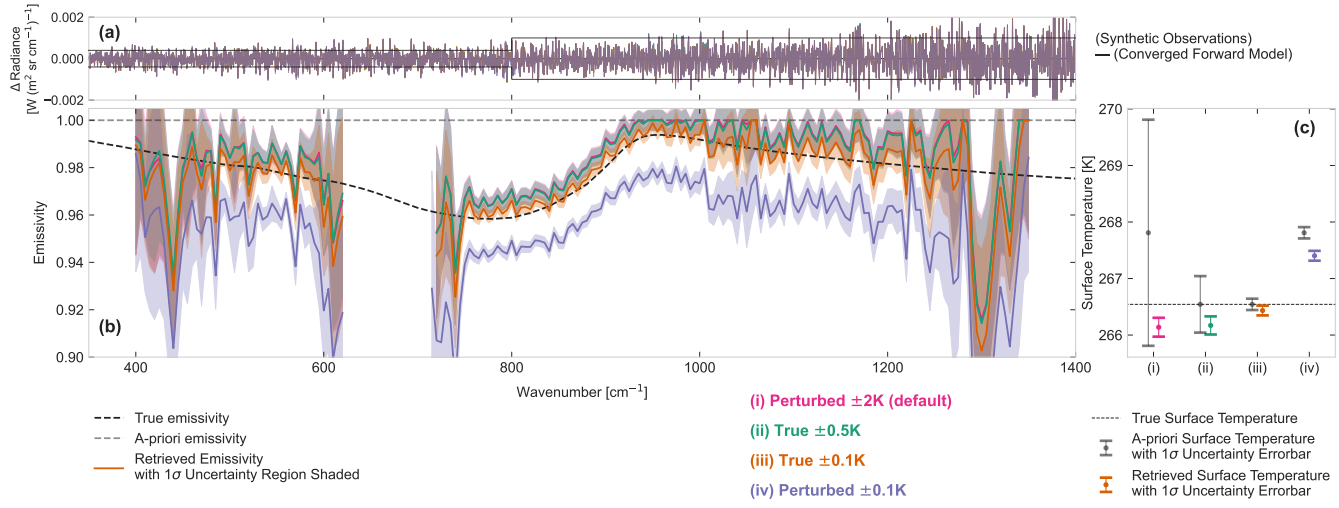


Figure 7. Different constraints on the surface temperature retrieval. All four retrievals, labeled (i)-(iv), are of the coarse snow emissivity scene at 67° N 29° E on 15 January 2018 12:00, with default parameters as outlined in Section 3. This scene was chosen as an illustrative example of the emissivity-surface temperature correlation due to the small positive offset in the default retrieved emissivity (associated with the specific instrumental random noise used in this case). In Panel (b) the colored lines are the retrieved emissivity with the $\pm 1\sigma$ retrieval uncertainty range shaded in the same color. The a-priori emissivity is shown in dashed light grey, and the true coarse snow emissivity is shown as a dashed black line. Panel (c) shows the retrieved surface temperatures with $\pm 1\sigma$ retrieval uncertainty for the same retrieval runs as Panel (b) as well as the surface temperature a-priori $\pm 1\sigma$ uncerainty for the four cases (labeled (i)-(iv) on the x-axis). The true surface temperature is shown in Panel (c) as a dashed black line. Panel (a) shows the difference between the synthetic observations, which are the same in all four cases, and the converged forward model. In all figures the retrieved quantities are color-coded in the same order: case (i) in pink is the default retrieval, with a perturbed surface temperature a-priori with a 2 K uncertainty and an initial guess at 261.5 K (not shown in the figure). Case (ii) in green has both the temperature a-priori and initial guess at the true value with a 0.5 K a-priori uncertainty. Case (iii) in orange is the same as (ii) with a 0.1 K a-priori uncertainty. Case (iv) in purple is the "misconstrained" case: it has a 0.1 K a-priori uncertainty, but the a-priori is the same as in (i) and the initial guess is the same as the a-priori. Note that the scale in Figure (b) only shows the emissivity from 0.9-1.0, and that even the miscontrained case only results in an uncertainty on a scale of ∼0.02.

6.1 Surface temperature and emissivity in the surface emission equation

The surface emission equation (Equation 3) as written is degenerate. Even if the atmospheric state is known and so $L_{d,\sigma}$ is given, measurements of $S_{\text{surf},\sigma}$ at N wavenumbers still leave N+1 unknowns to solve for: N spectral emissivity values and the surface temperature T_s. So in theory at spectral resolution for any value of the surface temperature T_s it is possible to find a corresponding surface spectral emissivity ϵ_σ that produces the correct surface radiance.

Different methods have been developed to deal with this degeneracy in the MIR when it occurs (see Li et al. (2013) for a review). While most methods make assumptions or use empirical relations which cannot be extended into the FIR, as Murray et al. (2020) and Bellisario et al. (2017) have shown, MIR measurements can be used to retrieve a T_s which can then be used

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

for the FIR emissivity retrieval, and future work could investigate incorporating such methods in tandem with the full-spectrum simultaneous OE retrieval used in this work.

In the FEES OE retrieval the assumption that breaks the degeneracy of Equation 3 is the retrieval of emissivity on a coarser grid than the measurements. As discussed in Section 4 the $\sim$0.4 cm^{-1} spaced ϵ_σ used to calculate $S_{\mathrm{toa},\sigma}$ is computed by linearly interpolating between the emissivity values retrieved on a coarser 5 cm^{-1} grid. Thus the retrieval vector $\hat{\mathbf{x}}$ has less elements than the observations vector $\mathbf{y}$ and the retrieval is not ill-posed, only ill-conditioned. This interpolation uses the assumption that the emissivity is smooth, and so breaks the degeneracy in a similar way to the retrieval method seen in Murray et al. (2020) and Knuteson et al. (2004). If in the FEES OE forward model the emissivity and T_s move away from the true value, to keep $S_{\mathrm{surf},\sigma}$ the same in Equation 3 the spectral emissivity would have to take up a shape with sharp high-resolution spectral features corresponding to the spectral pattern of $L_{d,\sigma}$. These cannot be reproduced by the interpolated coarser grid, and so ξ^2 is larger farther away from the correct emissivity. Thus the smoothing means that an incorrect emissivity introduces errors in the forward model, and this penalization leads the algorithm to nudge the retrieval vector towards the true value.

However for small shifts away from the true emissivity and true T_s the errors introduced in $S_{\mathrm{surf},\sigma}$ can be within the FORUM instrumental uncertainty. Thus to a limited extent the functional form of the emissivity and surface temperature still allows the retrieval to converge to a range of different emissivities. Such a parameter combination is sometimes called *sloppy*: moving along a sloppy direction in the parameter space has little effect on the behaviour of the model (see Transtrum et al. (2011)). The combination of T_s and emissivity form a *sloppy valley* in the model parameter space.

Figure 7 is shown both as an illustration of how surface temperature and emissivity compensate for each other and as a comparison of different a-priori constraint scenarios. The retrieval of scene 67° N 29° E with instrumental noise seed 0 was specifically chosen for this figure due to the $\sim$0.01 shift seen in the default retrieval, and is not necessarily a representative case.

As mentioned above, it is likely that for operational FORUM retrievals an estimate of T_s will be available either from independent observations, from synergy with IASI-NG, or from a different analysis of the FORUM observations. Thus the retrieval is run for four different scenarios of surface temperature a-priori information:

- (i) The default FEES retrieval, where a perturbation of the the true T_s is used as a-priori with a 2 K a-priori uncertainty that is characteristic of surface temperature measurements
- (ii) To model the ideal scenario of correct and accurate independent measurements the true T_s is used as both a-priori and initial guess with a smaller 0.5 K a-priori uncertainty
- (iii) A similar but less realistic scenario in which a high confidence in the independent measurement of the true T_s means that the true value is set as both a-priori and initial guess as in (ii), but in this case with a 0.1 K a-priori uncertainty
- (iv) To test whether using a tight a-priori constraint is advisable, the final retrieval uses the perturbed T_s of (i) as the a-priori and initial guess, with the 0.1 K a-priori uncertainty of (iii)

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

The first thing to note from the figure is the expected anti-correlation of surface temperature and emissivity systematic uncertainties in the retrieved values. Out of the four cases only retrieval (iii) has a retrieved surface temperature centred on the true value, with (i) and (ii) having lower and (iv) higher retrieved surface temperatures. These shifts in T_s cause upward/downward shifts of the whole spectral emissivity, with sign and size anti-correlated with the systematic uncertainty in surface temperature. It is interesting to note that even though the emissivity retrieval is shifted for the different cases, the emissivity retrieval uncertainty is the same for all of them, and when examined none of the standard quantifiers (see Figure 3) show which retrieval is better than the other. The reason can be seen in Figure 7(a) - all of these solutions are in the same *sloppy valley* of the parameter space, and so reproduce the observations to the same accuracy within the FORUM goal noise. This illustrates the effect that the functional form ($\epsilon B(T_s)$) of T_s and emissivity in the forward model can have on the retrieval.

Are the imposed constraints on T_s useful for mitigating such compensating shifts and reducing the systematic uncertainty on emissivity? There are two points to be made from the cases in Figure 7:

- Even a constraint of ± 0.5 K around the true value of T_s does not correct the shift seen in the default retrieval and can still result in an emissivity retrieval in which the true emissivity is outside the $\pm 1\sigma$ retrieval uncertainty range (but it should be noted that it is within both $\pm 2\sigma$ and the goal FORUM emissivity uncertainty of ± 0.01).

- Scenario (iii) shows that a constraint of ± 0.1 K is sufficiently small to result in the correct retrieved emissivity. However scenario (iv) shows that this is too tight of a constraint - if the a-priori T_s value is inaccurate even by ± 1.5 K, this already causes a much larger shift in the retrieved emissivity than is seen in the default scenario with more freedom for T_s. It is therefore not recommended to use such a tight a-priori constraint.

6.2 Impact on the retrieval by the a-priori and initial-guess choices

The retrievals shown in the previous section investigated possible T_s a-priori constraints. This Section investigates the impact allowed by the correlation of surface temperature and emissivity when varying the value of the emissivity initial guess and a-priori without changing the a-priori uncertainty constraints.

To explore the individual effects of the emissivity a-priori and initial guess on the retrieved emissivity their values are varied independently. The baseline retrieval was run for a combination of different constant a-prioris and initial guesses for four different geographical scenes, and the results are shown in Figure 8. The impact of the different combinations is shown by shading in the range between the maximum and minimum of systematic uncertainties in the retrieved emissivities for three color-coded scenarios, as well as shading in the maximum retrieval uncertainty range in grey. These scenarios are:

- The initial guess is kept constant at 0.9 and the a-priori is varied in steps of 0.1 from 0.7 to 1.0

- The a-priori is kept constant at 0.9 and the initial guess is varied from 0.7 to 1.0

- The initial guess and a-priori take on the same value and are jointly varied from 0.7 to 1.0

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

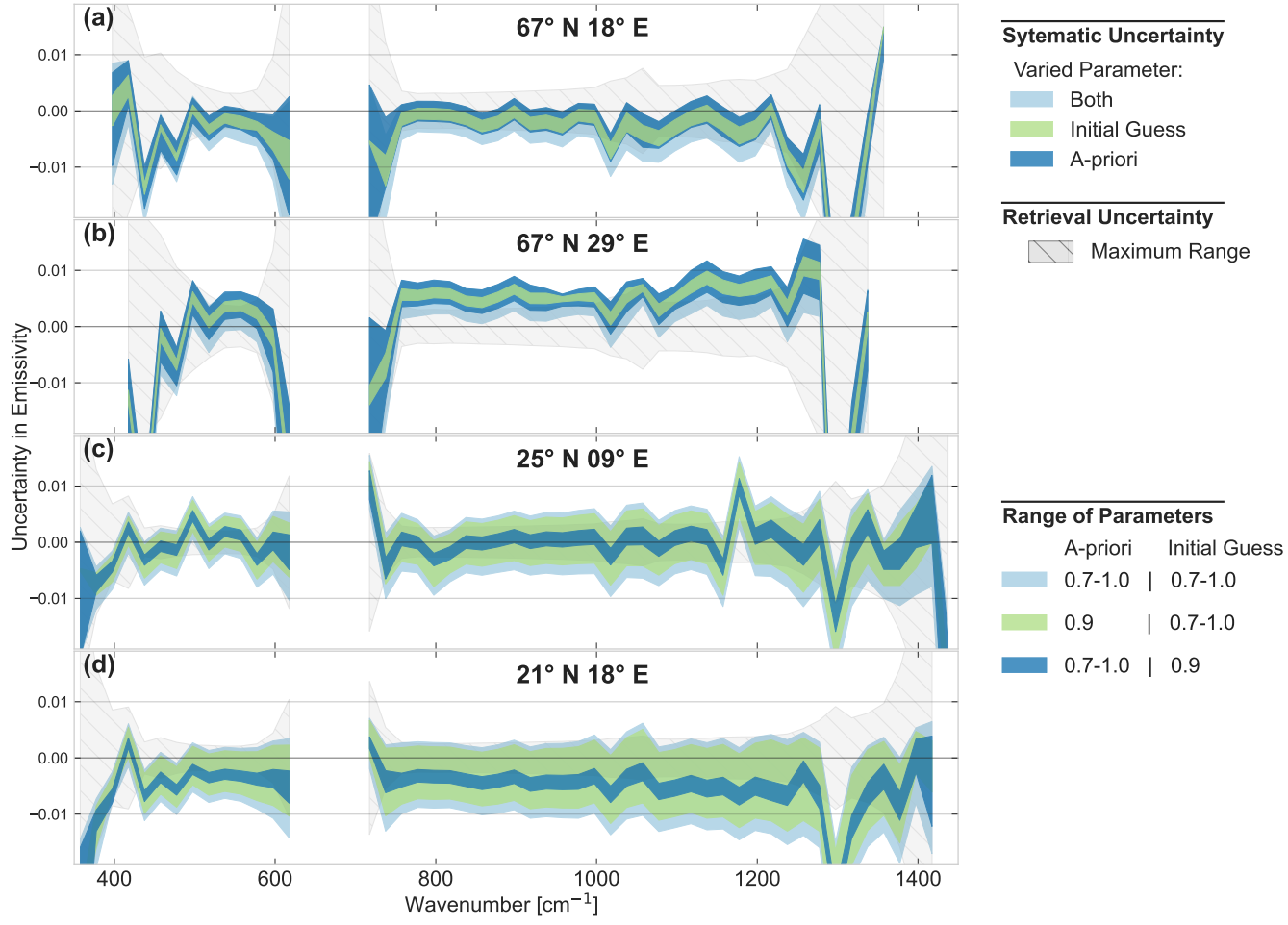


Figure 8. Range of systematic uncertainty in emissivity retrievals caused by different choice of emissivity a-priori or initial guess. Figures (a), (b), (c) and (d) show this range for four different geographical scenes: 67°N 18°E, 67°N 29°E, 25°N 09°E and 21°N 18°E, respectively. Except for the choice of emissivity initial guess and a-priori, all details can be found in Section 3. The colored regions are shaded between the maximum and minimum value of systematic uncertainty in the retrieved emissivity for different variations of the parameters. In dark blue the initial guess is kept constant at 0.9, and the a-priori takes the values of 0.7,0.8,0.9 and 1.0. In light green the a-priori is kept constant at 0.9 and the initial guess takes the values of 0.7,0.8,0.9 and 1.0. In light blue the a-priori and initial guess take the same values: 0.7, 0.8, 0.9 and 1.0. For reference the maximum value of the 1σ retrieval uncertainty from all 12 parameter combinations is shown as a dashed grey region in the background (the variation in this quantity between the retrievals is negligible). The uncertainties have been binned to a 20 cm^{-1} grid for clarity, and the sensitivity criterion (Equation 8) has been applied using the average IQ of the four cases for each shaded region. Note that the ranges in all four Figures are very small in extent, and the scale of the y-axis is ± 0.02 to highlight the differences.

While this is not an exhaustive list of the possible a-priori/initial guess combinations in the 0.7-1.0 range, the maximal impact that combinations in this range can have are represented by the difference between the case where both the a-priori and initial guess are 0.7 and that when they are both 1.0.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

Note however that all these retrievals are run for the same default instrumental noise seed, and so the specific higher/lower value of the retrieved emissivity is not necessarily characteristic. An in depth analysis would average retrievals run for at least 100 different versions of random instrumental noise as well as varying the L2M random seed, but this is outside the scope of the slow line-by-line forward model used by the L2M (which prioritises accuracy). On the other hand the choice of instrumental random noise should not affect the magnitude of the resulting emissivity ranges or their relation to each other, which is what is examined in this section (to confirm this the above analysis was in fact repeated for a small number of seeds and showed similar results, with the ranges shifted up or down by a small amount). A full analysis would also consider different a-priori uncertainties (see Appendix B) and T_s retrieval parameters.

Figure 8 shows the same full-spectrum upward/downward shifts in emissivity that were seen in Figure 7. In all of the scenes the impact of the a-priori/initial guess variation is not large overall, and the full range of variation amounts to at most a 0.015 relative difference in emissivity. The full range also appears to be additive in the impact of the two parameter choices (i.e. the range of the joint variation is the sum of varying each parameter individually).

However the relative and total size of the ranges show a different behaviour in scenes 8(a) and 8(b) than in scenes 8(c) and 8(d). While in the first two the variation of initial guess has slightly less of an influence than the a-priori, in the third and forth the sensitivity to the initial guess is stronger. This would not in general be expected from an OE retrieval, where usually the initial guess has little influence. However, the effect of the initial guess choice seen in Figures (c) and (d) is not due to a false convergence of the retrieval - the final forward model of all the retrievals for a given scene is almost identical. Thus they have the same final χ^2 (see Equation 1) and reach convergence in the same way. This is the same process that was seen in Figure 7(a), where the shifts in T_s and emissivity compensate for each other in a way that results in the same forward model within the FORUM noise. We can conclude that the *sloppy valley* of emissivity and T_s allows for a small range of solutions around the true value, and the choice of initial guess gives the retrieval a small nudge within this range.

The different behavior in the four scenes is likely due to their geophysical characteristics: while 67° N 18° E and 67° N 29° E both have low surface temperatures and low surface-to-air temperature contrast, 25° N 09° E and 21° N 18° E are hot scenes with high surface temperatures and high surface-to-air temperature contrast (see Table 1). This means the latter two have a stronger correlation of surface temperature with emissivity and so the retrieval vector can take larger steps in the parameter space. This effect of the path on the solution is discussed and analysed in more detail in Appendix A. For the purpose of this section it is sufficient to note that although the range is at least twice as large for the hotter scenes, even in the worst-case scenario the choice of initial guess and a-priori only change the emissivity by about 0.015, still close to the FORUM goal accuracy of 0.01.

However these shifts in the retrieval allowed by the *sloppy valley* in the T_s-emissivity parameter space are not well represented in a standard uncertainty analysis that only uses the diagonal elements of $\mathbf{S_x}$ and the emissivity submatrix of $\mathbf{A}$. This can already be seen from the fact that the retrieval uncertainty ranges shown in the background of Figure 8 are sometimes smaller than the parameter-variation induced ranges. The importance of the off-diagonal elements can be explained when comparing Figure 8(a) to the diagonal submatrix averaging-kernel-derived quantifiers of this scene shown in Figure 3. For example, if only $\mathbf{A}$ from Figure 3(d) is considered, one would conclude that the emissivity at 800 cm^{-1} is independent of the choice of

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

Coordinates	(a) FIR		(b) MIR	
	Slope	R	Slope	R
67° N 18° E	-0.017	-0.93	-0.019	-0.99
67° N 29° E	-0.021	-0.96	-0.018	-0.99
21° N 18° E	-0.011	-1.0	-0.013	-1.0
25° N 09° E	-0.011	-0.99	-0.014	-1.0

Table 3. Complementary table to Figures 9(a) and 9(b). A linear slope is fitted to the values for each scene using a least squares minimization, and as the intersect of all the fits is 0 only the slopes are quoted here. The sample Pearson correlation coefficient (R, see Equation 9) is also calculated for each set of points. The corresponding p-value (hypothesis test) for all 8 cases is smaller than 10^{-12}.

emissivity a-priori as the averaging kernel coefficient is equal to 1. But Figure 8(a) shows that the retrieval at 800 cm^{-1} for this scene does in fact vary when the a-priori is changed. In fact the effect is indirect: Figure 3(d) shows the choice of a-priori does directly influence the retrieved emissivity in other less-transparent regions, and some of these are correlated with the surface temperature. For example when the less sensitive retrieved emissivity at 400cm^{-1} is higher/lower depending on the a-priori choice the anti-correlated surface temperature is lower/higher. And thus through T_s the emissivity in the atmospheric window is also affected by the choice of a-priori, albeit indirectly. Therefore the uncertainty the correlation allows for should be evaluated separately from the standard OE quantifiers.

6.3 Spectral dependence of the emissivity-surface temperature correlation

The final step in understanding the variations allowed by the T_s-emissivity *sloppy valley* is to analyse the correlation strength in different spectral regions. Equation 3 shows that there are two main factors that could cause differences in correlation in the FIR and MIR. The first originates from $B(T_s)$ having a different shape in different spectral regions. The second is that even if the downwelling radiation L_d is known, its value still differs significantly between the FIR and MIR. This is for the same reasons as discussed in section 4: in the MIR the atmospheric window is transparent and so L_d is negligible, whilst in the FIR L_d is higher or lower depending on the amount of water vapour and on the microwindow structure (see e.g. Palchetti et al. (2016); Palchetti et al. (2020) for ground measurements of FIR downwelling radiation).

To investigate these effects Figure 9 shows an analysis of both the empirical correlation of 28 retrieved values for four scenes each and an analytic correlation calculated from the standard OE equations. The same geographic scenes are used as in Figure 8. For the empirical correlation the baseline retrieval is run for each scene using instrumental spectra with seven different versions of random instrumental noise (generated with FSI seeds of 0,1,2,3,4,5,6) and then each is retrieved with equal flat a-prioris and initial guesses set to 0.7,0.8,0.9 and 1.0, resulting in 28 cases for each scene. The variation of the instrumental noise and the a-priori and initial guess results in a range of different systematic uncertainties (as discussed in Section 6.2). These uncertainties are shown in Figures 9(a) and 9(b), which plot the average systematic uncertainty of emissivity in a specific spectral range against the systematic uncertainty in T_s. Constant and relatively small spectral ranges are chosen so that the variation of the

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

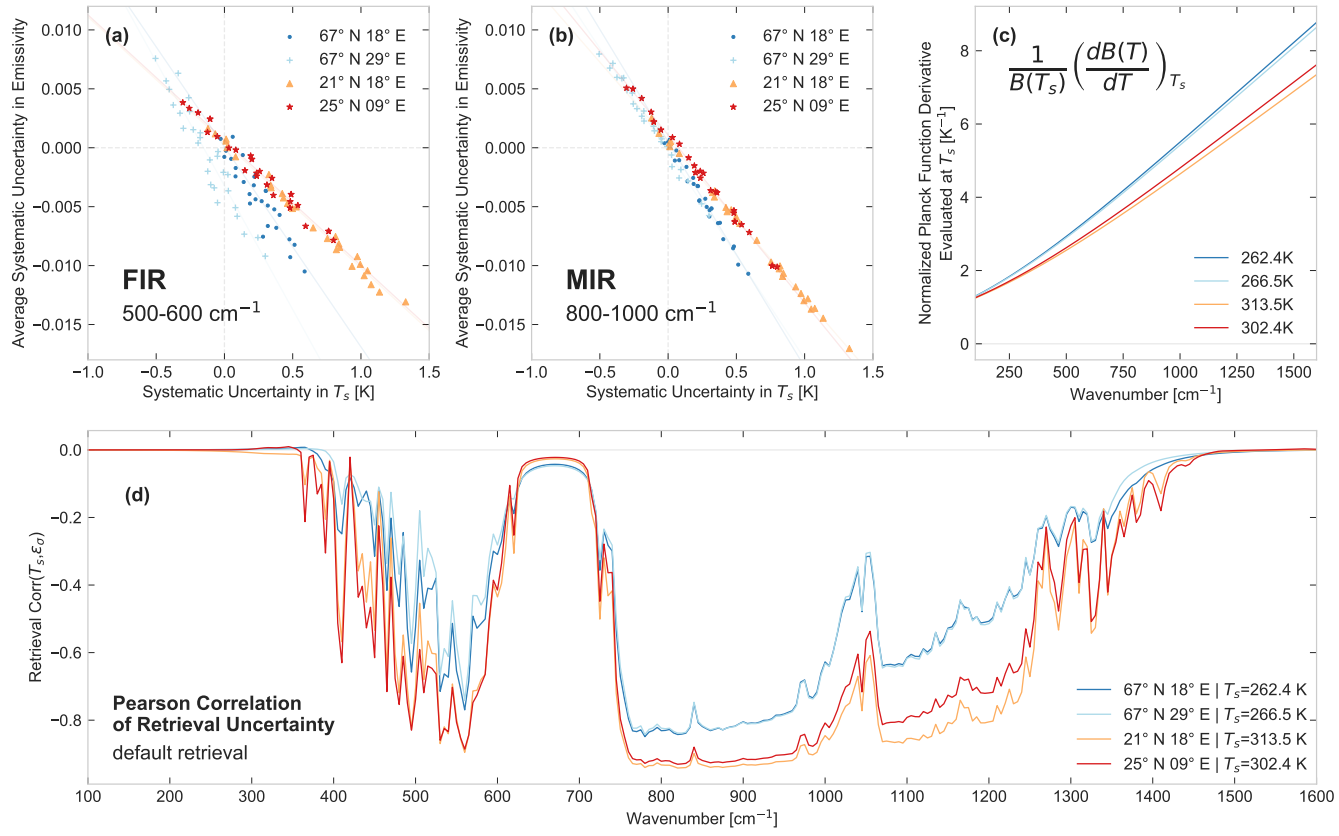


Figure 9. Correlation between emissivity and surface temperature in the atmospheric state retrieval. The four color-coded scenes are 67°N 18°E, 67°N 29°E, 21°N 18°E and 25°N 9°E in dark blue, light blue, orange and red, respectively, with details as outlined in Section 3. Figures (a) and (b) show the retrieval systematic uncertainties for 28 retrievals of each scene. The baseline retrieval is run on spectra generated with six versions of random instrumental noise (seeds of 0 to 6) and for equal flat a-priori and initial guess set to 0.7,0.8,0.9,1.0. In Figure (a) the average systematic uncertainty in emissivity in the 500-600cm^{-1} range is plotted against the systematic uncertainty in the surface temperature (T_s). Light grey dashed lines show the true emissivity and T_s. Table 3 details the slope of the linear trend fitted to the points (grouped by scene), as well as the sample Pearson's correlation coefficient (see Equation 9) for this data. The fitted trend is also plotted as a light line of the same color as the corresponding data. Figure (b) is similar, but with the 800-1000cm^{-1} systematic emissivity uncertainty on the y-axis. Its slope and correlation values are also detailed in Table 3. Figure (c) shows an analytic calculation of the value of the normalised Planck Function derivative $[dB(T)/dT]_{T_s}/B(T_s)$ at the four different true surface temperatures of the scenes, plotted over the full FORUM spectral range. Figure (d) shows the analytic Pearson correlation coefficient (see Equation 12) of the emissivity and T_s retrieval uncertainty over the full FORUM spectral range for all four cases from the retrieval run using the default setting outlined in Section 3. The correlation is calculated as shown in Equation 12 from the retrieval uncertainty covariance matrix at convergence (see Equation 4).

correlation slope and strength in the averaged range is small enough to allow a meaningful analysis. The spectral ranges of 500-600 and 800-1000 cm^{-1} are chosen to represent the FIR and MIR, respectively, as these are the spectral intervals with the

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

highest sensitivity in those regions. These are not representative of the variation in the full FIR/MIR, but only indicative of the difference between the regions.

As expected, there is a strong anti-correlation between the systematic uncertainties both in the FIR and the MIR. Table 3 lists the slopes of the linear trends fitted to the data in these figures (grouped by scene and spectral region), as well as the corresponding sample Pearson correlation coefficient R using the standard formula:

$$\mathrm{R} = \frac{\sum(\Delta T_s - m_{\Delta T_s})(\Delta\epsilon - m_{\Delta\epsilon})}{\sqrt{\sum \Delta T_s - m_{\Delta T_s})^2 \sum(\Delta\epsilon - m_{\Delta\epsilon})^2}} \tag{9}$$

where ΔT_s and $\Delta\epsilon$ are the data vectors of systematic uncertainties in surface temperature and emissivity, and $m_{\Delta T_s}$ and $m_{\Delta\epsilon}$ are the means of these vectors.

Three points can be highlighted from these results:

- With the exception of 67° N 29° E, the slope of the linear fit is steeper in the MIR than in the FIR
- In both spectral regions the slope of 67° N 18° E and 67° N 29° E is steeper than that of the other two scenes
- For scenes 67° N 18° E, 67° N 29° E and 25° N 09° E the scatter of values is larger in the FIR than the MIR (lower $\mathbf{R}$ in Table 3).

A possible cause for the variation in slopes can be found in the form of $B(T_s)$. To see this let $L_d = 0$, which primarily simplifies the analysis but is also a valid assumption for the MIR and for the centres of the FIR microwindows. Equation 3 then becomes:

$$S_{\text{surf}} = \epsilon B(T_s) \tag{10}$$

where the σ underscore has been dropped for convenience. Keeping S_{surf} constant, the equation is rearranged to get an expression for ϵ and then the derivative is taken with respect to T_s:

$$\frac{d\epsilon}{dT_s} = S_{\text{surf}} \frac{1}{-B^2(T_s)} \frac{dB(T)}{dT}\Big|_{T_s} = -\epsilon \left[\frac{1}{B(T_s)} \frac{dB(T)}{dT}\Big|_{T_s} \right] \tag{11}$$

The dominating factor in determining $d\epsilon/dT_s$ (the slope in Figures 9(a) and 9(b)) for a given scene and wavenumber is the expression in brackets on the right hand side of Equation 11, as although ϵ also varies spectrally and geographically its average variations are an order of magnitude smaller (20% as opposed to 800%). The plot in Figure 9(c) shows this expression for the surface temperatures of the four different scenes. This plot shows that the value of this expression increases with wavenumber and is lower for higher T_s. This behaviour could explain the difference in slopes observed in Figures 9(a) and 9(b). The larger value of this term and thus $d\epsilon/dT_s$ in the MIR would result in the steeper slope observed for most scenes in the MIR. And as 67° N 18° E and 67° N 29° E are much colder scenes, it is expected that the slope of their linear relation will be steeper than that of 21° N 18° E and 25° N 09° E.

Finally, an analytic correlation analysis can shed light on the scatter about the linear trend of the values in Figures (a) and (b) (also represented by the size of their R correlation value in Table 3). In Figure 9(d) the analytic Pearson correlation coefficient

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

of the retrieval uncertainties of T_s and emissivity is shown. The uncertainties are given in the retrieval uncertainty covariance matrix $\mathbf{S_x}$ defined in Equation 4. Using the standard formula for the analytic (population) Pearson correlation coefficient:

$$\mathrm{Corr}(T_s, \epsilon_i) = \frac{\mathrm{Cov}(T_s, \epsilon_i)}{\sigma_{T_s}\sigma_{\epsilon_i}} = \frac{\mathbf{S}_{T_s,\epsilon_i}}{\sqrt{\mathbf{S}_{T_s,T_s}}\sqrt{\mathbf{S}_{\epsilon_i,\epsilon_i}}} \tag{12}$$

where $\mathbf{S}_{T_s,T_s} = \sigma_{T_s}$ is the retrieval uncertainty standard deviation of T_s (dropping the $\mathbf{x}$ in $\mathbf{S}$ for visibility), and similarly for ϵ_i, the i-th value in the emissivity retrieval vector. Note that Figure 9(d) only shows this value for the baseline retrieval of the four scenes, and thus is meant as an illustration of the spectral structure of the correlation and not as a quantitative reference.

Figure 9(d) shows that as expected T_s and emissivity are not correlated to the same extent in different spectral regions. The correlation mirrors the spectral structure seen in the emissivity Jacobian (see Figure 3) - unsurprisingly, as it is calculated from $\mathbf{S_x}$ which in turn is calculated from the Jacobian (as well as from $\mathbf{S_a}$). There is a strong uniform correlation of the MIR emissivity points with T_s, while the correlation of the FIR values depends on the microwindow structure and with that on the dryness of the atmosphere. The $\sim$750-1250 cm^{-1} region of the MIR is called the atmospheric window as it is almost fully transparent to the surface, and thus in most of that region the strength of the correlation is determined solely by the value of the surface temperature. In the FIR the difference in correlation strength is harder to attribute precisely, as it is due to a combination of the pwv and the surface temperature. However its value for the four scenes analysed here can still be used to compare the correlation to the scatter seen in Figures 9(a) and 9(b). These show good agreement, as the scenes with a lower retrieval uncertainty correlation coefficient also have a smaller systematic uncertainty correlation and larger scatter around the linear trend.

In summary, the correlation of surface temperature and emissivity behaves as would be expected from the physics of the forward model. The range of systematic uncertainties in Figures 9(a) and 9(b) confirm what was already shown in Sections 6.2 and 6.1: that this correlation allows for a range of retrieved emissivities depending on the retrieval parameters. The predictability of the behaviour of the correlation is important for the evaluation of this effect, which should be thoroughly quantified during the development of the operational retrieval.

7 Conclusions and recommendations

This study follows from previous work on FORUM geophysical retrievals (e.g. Ridolfi et al., 2020; Sgheri et al., 2021) showing that FORUM measurements will be able to provide retrieved surface emissivity in a significant region of the FIR. Using the FEES, factors that influence OE retrievals of FIR emissivity were investigated with an emphasis on the development of operational retrievals for FORUM. More information could be gained from the retrieval by analysing individual scenes in detail and combining the OE retrieval with different methods and other observational products, and this should be addressed in future work. Additionally, we have only considered the use of FORUM measurements by themselves - see Ridolfi et al. (2020) for a discussion of how synergetic retrievals with IASI-NG observations can improve the FORUM geophysical products.

In Section 4 the retrieved emissivity was introduced together with the quantifiers used to analyse it. The spectral structure of these quantifiers and their variation with precipitable water vapour (pwv) content led to the definition of a criterion for

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

presenting an emissivity retrieval product. In Section 5 the variation in quality of the retrieval with pwv content was compared for multiple geographic scenes. Section 6 then investigated the consequences and characteristics of the surface temperature - emissivity correlation that arises from the functional form of the surface emission equation.

The main points to be highlighted from these results are:

- Emissivity retrieval quality, degrees of freedom, and extent of retrieval sensitivity towards shorter wavenumbers increases as the pwv of the scene decreases. This means that when producing the FORUM geophysical emissivity product, the spectral extent of emissivity in the product should be decided on a scene-by-scene basis (and not, for example, by applying a latitude cut-off). We recommend using the **information quantifier** of the scene as a basis for evaluation.

- The functional form of the surface emission equation leads to a strong anti-correlation of surface temperature and emissivity in the retrieval. Thus the retrieval can converge to a small range of solutions around the true value. Attempting to correct this by constraining the surface temperature retrieval (i.e. introducing more a-priori information) could lead to larger shifts away from the true emissivity when the a-priori for the surface temperature T_s is wrong. Thus the default FEES T_s **a-priori uncertainty of** $\pm$**2** K **is recommended**, as even in the worst cases investigated here it only results in an emissivity offset of an acceptable value around 0.01.

- For the cases investigated here, varying the value of the emissivity a-priori and initial guess between 0.7 and 1.0 results in relative differences in the FIR retrieved emissivity of up to 0.015 in the extreme.

- The correlation of emissivity with T_s leads to offsets in both retrieval parameters that are not accurately reflected in the standard quantifiers. **It is recommended that the systematic uncertainty originating from the** T_s**-emissivity correlation is evaluated** in detail during the development of the operational retrieval. Further work could also look into the possibility of using external constraints on T_s as well as other methods for T_s retrieval (such as that used in Murray et al. (2020)) to complement the OE.

- The information content of the emissivity retrieval with respect to further scene parameters such as the absolute surface temperature and its contrast to the atmospheric temperature should be investigated in the future.

In conclusion, the FORUM mission will be able to provide a unique contribution to our knowledge of surface emissivity in the FIR for many locations on the globe and potentially most types of surfaces. In this work we have taken the first steps towards the development of an operational emissivity geophysical retrieval for the FORUM mission by highlighting possibilities for optimization of the retrieval and the systematic uncertainties that still need to be quantified. As well as the development of the operational retrieval, complementary future work would include laboratory and aircraft measurements of emissivity, analysis of additional methods for surface temperature retrieval and an algorithmic optimisation of the emissivity retrieval grid.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

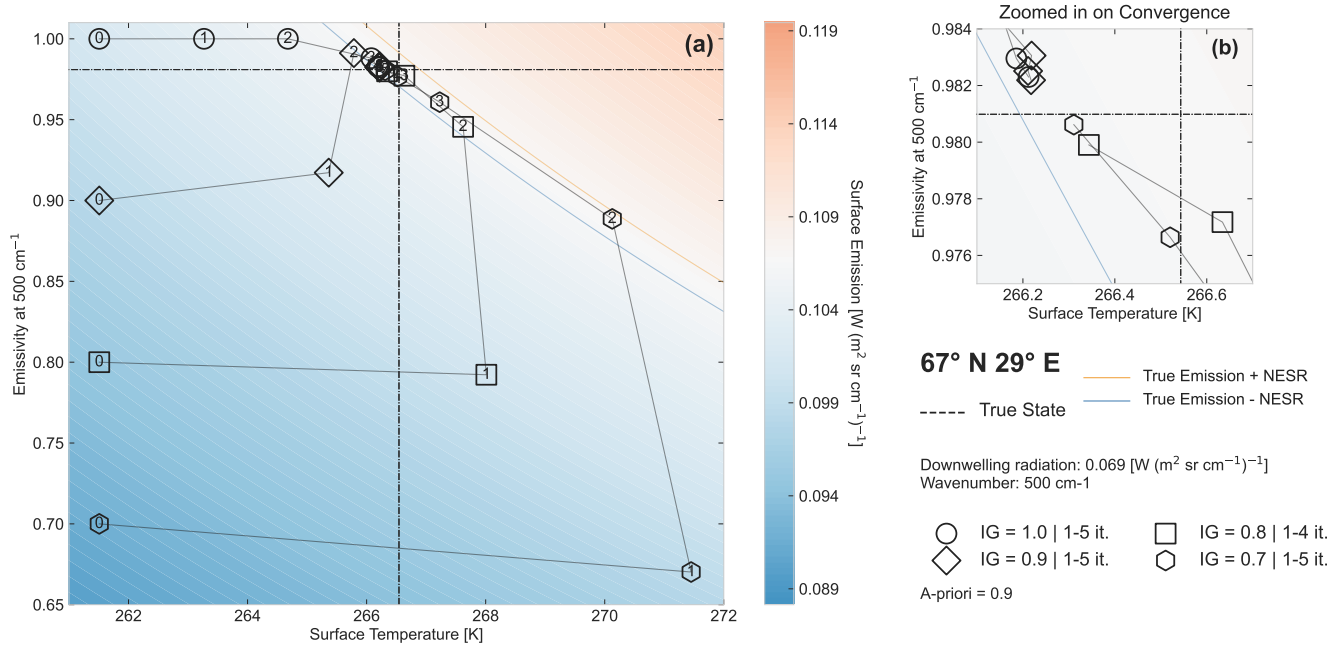


Figure A1. Value of two elements of the retrieval vector for all retrieval iterations in the emissivity - surface temperature (T_s) parameter space for the 67° N 29° E coarse snow scene. Figure (a) shows the parameter space defined by T_s and the emissivity at 500 cm^{-1}, with colored contours corresponding to the surface emission at 500 cm^{-1} (according to Equation 3) with the downwelling radiation set to the true value of 0.069 $\mathrm{W(m^2\ sr\ cm^{-1})^{-1}}$. The center of the diverging color scale is the true value of the surface emission. The true values of emissivity and surface temperatures are shown in dashed black lines. An orange (blue) line shows the contours of the true emission plus (minus) the FORUM noise equivalent spectral radiance (NESR) in the FIR. The different geometric marker shapes show the retrieval vector values for the iterations of four different retrievals, starting from the initial guess until the converged solution. In the four cases shown the emissivity a-priori is 0.9 and the initial guess takes the value of 1, 0.9, 0.8 and 0.7, shown using circles, diamonds, squares and hexagons, respectively. All other retrieval parameters are the same for the four cases and are the default values outlined in Section 3. Figure (b) shows an enlarged portion of Figure (a) centered on the true state for a clearer view of the later iterations (the color-scale and all other plotted values are the same as in (a)). The full spectral range of the final retrieved emissivity values for the four runs is shown in Figure 8 in Section 4 of the paper.

Appendix A: The Retrieval Path in the Emissivity-T_s Parameter Space

In Section 6 of the paper the concept of the T_s-emissivity parameter space and its *sloppy* nature was introduced in the context of the surface emission equation (Equation 3). The OE retrieval algorithm in the FEES minimises the cost function (Equation 1) using the Levenberg–Marquardt approach which interpolates between the Gauss–Newton algorithm and the method of gradient descent. The retrievals in this work converge after 4-6 iterations, and convergence is reached when the normalised change from one iteration to the next in χ^2 (the first term in Equation 1) is less than 0.01. The path the retrieval takes to convergence is hard to visualise, as the retrieval vector is stepping in a 300+ dimensional parameter space. However due to the linear contribution

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

of emissivity to the TOA radiance (see Equation 3) insight can be gained by plotting the steps in the surface temperature - emissivity slice of this parameter space, and two such plots are shown in this appendix. While the full forward model is far too complex and its computation too time-consuming to lend itself to contour plots or manifold visualisations, some insight can also be gained by showing these steps together with the contours of constant surface emission in Equation 3.

The issue addressed in this paper which benefits most from such parameter space path plots is that of the sensitivity of the retrieval to the initial guess discussed in Section 6.2. Figure A1 shows the convergence of the retrievals shown as the light green range in Figure 8(b), in which the emissivity of scene 67° N 29° E is retrieved with an a-priori of 0.9 and different initial guesses of 0.7,0.8,0.9 and 1.0. For each iteration, Figure A1(a) plots the value of two retrieval vector components, with the surface temperature on the x-axis and the emissivity at $500\mathrm{cm}^{-1}$ on the y-axis. The four different retrieval runs are represented by different geometric shapes. To put the convergence into context Equation 3 is used to plot the contours of the *true* surface emission value for each T_s-emissivity combination. For simplicity the surface emission is calculated at the surface (without the atmospheric transmission term). L_d is taken to be the true value at 500 cm^{-1}, calculated using a separate run of LBLRTM (version 12.10) using the true atmospheric state of the scene. Note that as the atmospheric profiles are also being retrieved, both the transmission and L_d in the forward model will not necessarily equal the true values at the early iterations, and so the background contours do not represent the surface emission used in the forward model at that iteration, but instead are there to give context to the later iterations (where the retrieval vector is close to the true). In addition these contours are not directly representative of the forward model, as $\mathbf{f}(\mathbf{x})$ includes many additional effects (for example those associated with the FORUM instrument). Finally, in Figure A1(b) a small region of the space has been zoomed in on so as to better show the behavior around convergence.

The behavior of the different retrievals in Figure A1 is typical of the emissivity retrieval. The retrieval vector starts from an initial guess which corresponds to a surface emission very different from the true value, and so takes large steps in the parameter space towards the correct surface emission contour (this is the gradient descent part of the Levenberg–Marquardt minimization). The existence of this true surface emission contour is the cause of the *sloppy valley* in the T_s-emissivity space discussed in Section 6. While in the MIR the contour is usually reached in one step in the FIR the retrieval usually takes 2-3 steps to reach the true emission value, as change in the water vapour part of the retrieval vector also change L_d. Once the true surface emission contour is reached, the retrieval proceeds along it, driven mainly by the T_s a-priori constraint and by the small forward model discrepancies caused by the emissivity smoothness assumption (discussed in Section 6 and difficult to visualise when plotting only the retrieval vector emissivity components). The main point seen in Figure A1 is that the direction of the shift of the final values from the true ones depends on whether the correct surface emission contour is first reached at a higher or lower value than the true emissivity (so that even though a retrieval might start from an emissivity at 0.9 that is lower than the true, due to the structure of the parameter space it will reach a final value that is higher than the true).

In Section 6.2, Figure 8 the comparison for the initial guess sensitivity for different geographical scenes revealed a different behavior of the colder 67° N 18° E and 67° N 29° E scenes and the warmer 25° N 09° E and 21° N 18° E scenes. The warmer scenes showed more sensitivity to the initial guess, in that there was a larger difference between the retrieved emissivity for different initial guesses. Figure A2 shows the convergence path of these four scenes in the FIR and MIR for the case when the

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

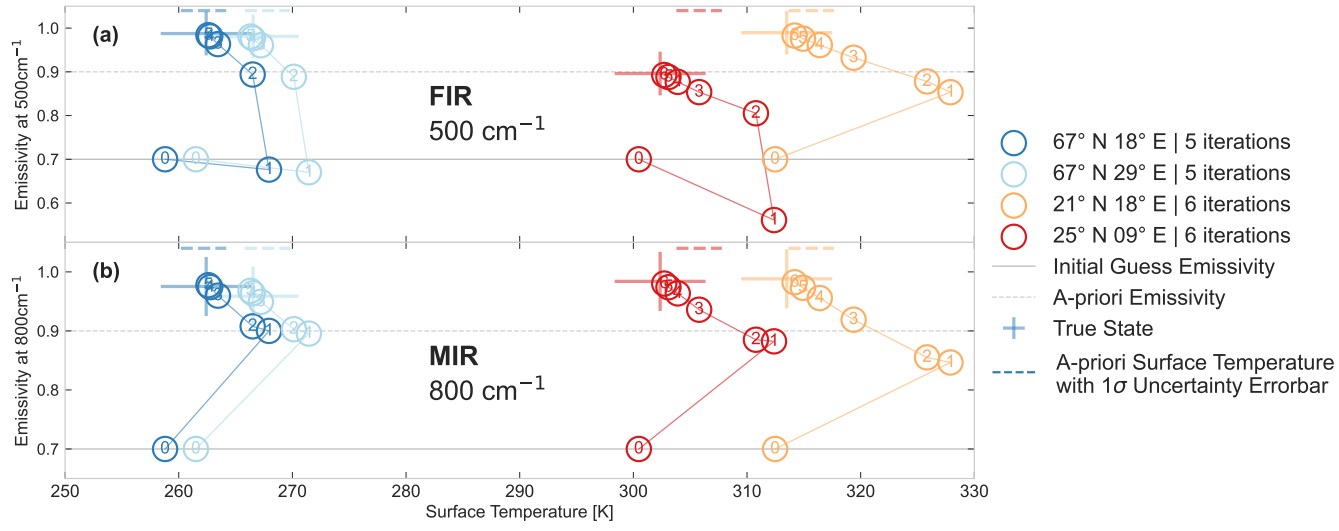
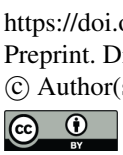


Figure A2. Value of two elements of the retrieval vector for all retrieval iterations in the emissivity - surface temperature (T_s) parameter space for four different scenes. The four color-coded scenes are 67°N 18°E, 67°N 29°E, 21°N 18°E and 25°N 09°E in dark blue, light blue, orange and red, respectively, with details as outlined in Section 3. The structure of the figure is similar to that of Figure A1. For the four retrievals the scattered colored circular markers show the values of the emissivity and T_s at each numbered iteration of the retrieval starting from the initial guess until the converged solution. The retrieval parameters are the same for all scenes, with an emissivity a-priori of 0.9 and an initial-guess of 0.7 (all others have the default values outlined in Section 3). The y axis of figures (a) and (b) show the value of the emissivity at 500 cm^{-1} and 800 cm^{-1}, respectively, with a shared x axis showing the surface temperature. Both figures also show the a-priori emissivity and initial guess emissivity as dashed and solid grey lines, respectively. The different $\pm 1\sigma$ a-priori uncertainty range for T_s is shown as a dashed line at the top of the plot (the y-axis location has no significance other than clarity), with the same color-coding as the circular markers. This color-coding is also used for the solid cross centered at the true values of emissivity and T_s of the scenes.

initial guess emissivity is set to 0.7 (and the a-priori is 0.9 as before). In this figure no surface emission contours are shown, as their different L_d values mean that the contours would differ for the four different scenes. Figure A2 shows that as discussed in Section 6.2, for the warmer scenes the retrieval vector takes larger initial steps in the parameter space in both spectral ranges. This is what we would expect from the stronger correlation associated with the warmer scenes, analysed in detail in Figure 9. Once the true surface emission contour is reached the steps are of similar magnitude for the four scenes. Figure A2 illustrates how such larger steps in the first iteration could potentially explain a higher sensitivity to the initial guess: by taking a larger initial step the retrieval approaches the true value from a lower emissivity value, and so also converges to a slightly lower emissivity value.

In summary, plotting the retrieval's path to convergence in the emissivity-T_s parameter space is a useful visualization tool. By comparing the paths of different cases it can provide further insight into the reasons underlying the sensitivity of the final retrieved product to different parameters.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021



Atmospheric Measurement Techniques Discussions

Appendix B: Choice of a-priori uncertainty

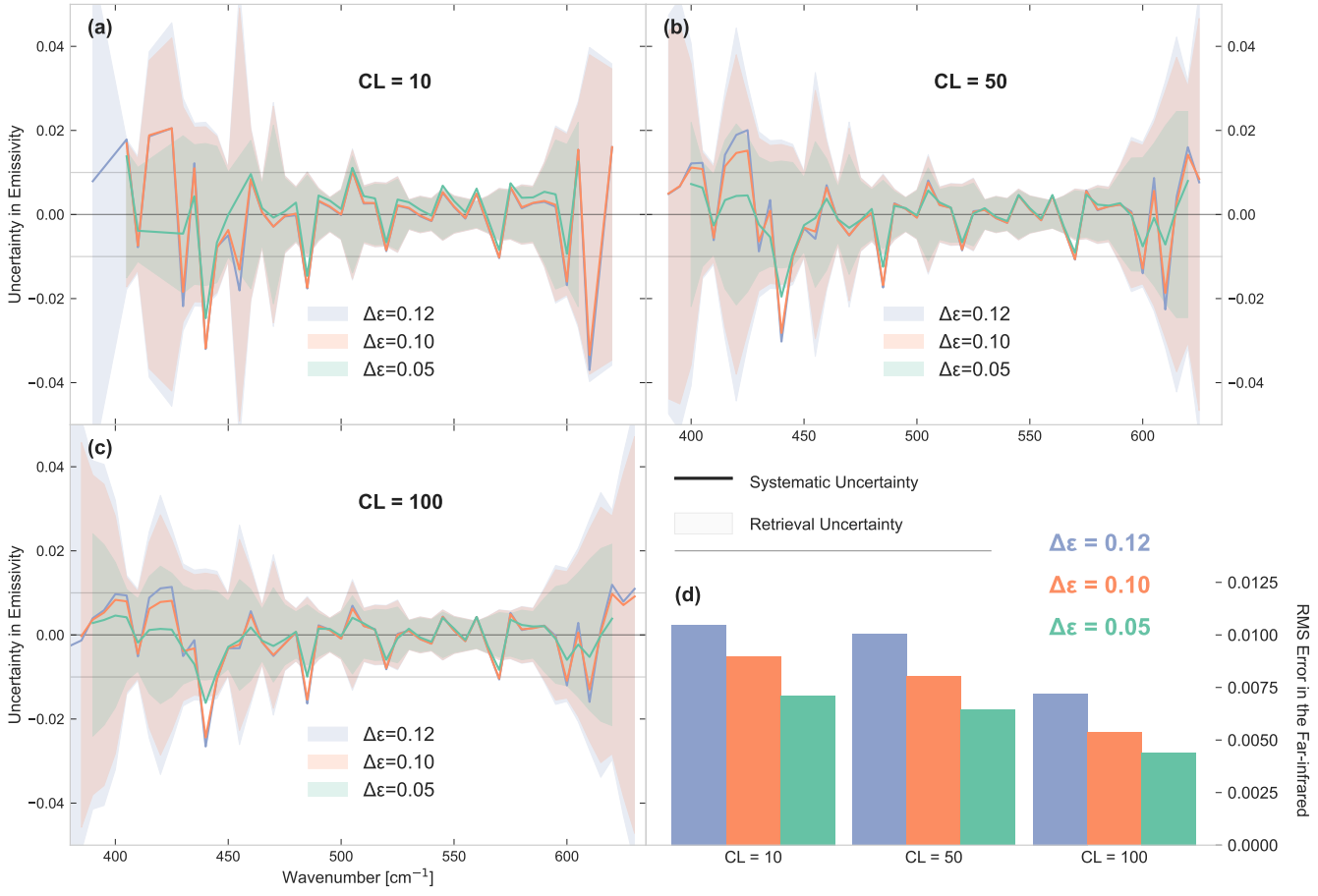


Figure B1. Systematic uncertainty in emissivity for different emissivity a-priori uncertainty parameters in the far-infrared. All nine retrievals are for the same fine snow emissivity at 67° N 18° E 15 January 2018 12:00:00, with default parameters as outlined in Section 3. Figures (a),(b) and (c) show the uncertainties for retrievals with a correlation length (CL) of 10, 50 and 100, respectively. Each figure shows the systematic uncertainty (true-retrieved emissivity) as a solid line for three values of the emissivity a-priori uncertainty $\Delta\epsilon$. The colors are green, orange and blue for $\Delta\epsilon$ values of 0.05, 0.1 and 0.12, respectively. The retrieval uncertainty 1σ range for the respective retrievals is shown as a shaded region of the same color. The uncertainties are shown only in the regions where the Information Quantifier for the respective retrievals is larger than 1 (not the same for each retrieval) between 350 and 650 cm^{-1}. Figure (d) shows the root-mean-square (RMS) error of the systematic uncertainty in the plotted spectral range for the nine cases.

Throughout this work the same emissivity a-priori uncertainty matrix $\mathbf{S_a}$ was used for all retrievals. Its value was chosen as a baseline case following the sensitivity tests shown in this Appendix.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

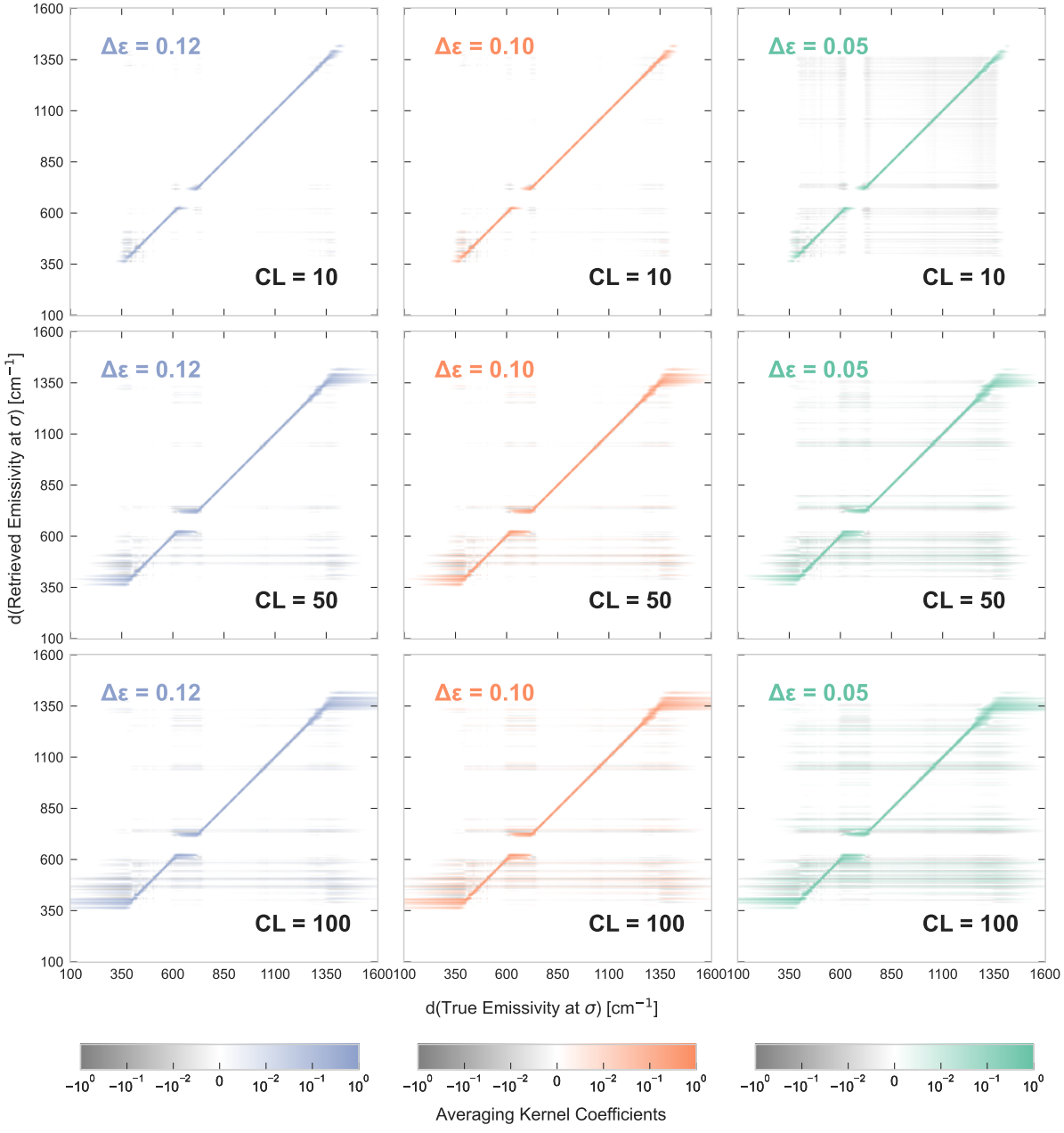


Figure B2. Averaging Kernel coefficients (see Equation 7) for the same nine cases shown in Figure B1. The color-coding of the emissivity a-priori uncertainty $\Delta\epsilon$ is kept the same as Figure B1, and its value is 0.12, 0.1 and 0.05 for columns 1, 2, and 3, respectively. Rows 1, 2, and 3 show cases with correlation length (CL) of 10, 50 and 100, respectively. The color scales are shown at the bottom of each column, and are the same for each figure. The values are plotted with a lower threshold of $\pm 10^{-3}$ for visibility.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

In the FEES, $\mathbf{S_a}$ is calculated using two parameters, the uncertainty and the correlation length. For the profiles the existence of reliable a-priori datasets justifies a nuanced calculation of the uncertainty matrix using uncertainty and correlation lengths that change with height (see Sgheri et al. (2021)). As there are no such datasets for FIR emissivity, as a starting point for optimising the uncertainty the same parameters are used for the full spectral range. Therefore a constant uncertainty $\Delta\epsilon$ and correlation length CL can be defined, and the $\mathbf{S_a}$ matrix elements for emissivity are then:

$$S_{\mathbf{a},ij} = \Delta\epsilon^2 \exp\left(-\frac{\Delta\sigma_{ij}}{\mathrm{CL}}\right) \tag{B1}$$

where $\Delta\sigma_{ij}$ is the wavenumber difference between the location of retrieved emissivity values ϵ_i and ϵ_j. In practice $\Delta\epsilon$ defines the *freedom* of the retrieval discussed in the previous Section 3, as a larger value will allow the profile to take larger steps at each iteration and reduces the penalization from $\mathbf{x} - \mathbf{x}_a$. The CL controls the off-diagonal elements in $\mathbf{S_a}$ - its presence means that the regularization term for the retrieved emissivity points is not minimized individually, but that the emissivity step at a given wavenumber is also affected by the difference of its neighbouring points from the a-priori. In practice this results in a smoother solution where the retrieval is sensitive to the a-priori.

These two parameters were varied in the retrieval setup and the results are shown in Figures B1 and B2. While the sensitivity tests were run for many different scenes, the analysis shown here is of the scene at 67° N 18° E, chosen as it is representative of the snow emissivity scenes that are the primary goal of FORUM's emissivity retrievals. The range of the uncertainty parameters shown is $\Delta\epsilon = 0.05, 0.1, 0.12$ and $\mathrm{CL} = 10, 50, 100\ \mathrm{cm}^{-1}$. Smaller values of $\Delta\epsilon$ were also considered, but are not shown as they did not give the retrieval the necessary freedom to converge to the right solution and caused a large systematic uncertainty. Figure B1 shows the systematic uncertainty and the retrieval uncertainty in the FIR for all nine cases, as well as the root mean square (RMS) error for the systematic uncertainty values shown. Figure B2 shows the averaging kernels of the nine cases for the full spectral range. The following points can be seen in these figures:

- The differences in uncertainties for a given correlation length are only present in the edge regions of the retrieval where the sensitivity is lower. This is because the a-priori uncertainty only matters where information is drawn from the a-priori and for a dry scene such as this (as discussed for Figure 3) in the centres of the FIR dirty window and of the MIR atmospheric window the retrieval is fully sensitive to the true state and thus the choice of a-priori uncertainty has no influence.

- Examining the averaging kernels shown in Figure B2 we can again see that the influence of the parameters is strongest in the edge regions of sensitivity. Here it can be seen that increasing CL leads to more information being drawn from regions to which the TOA is not in reality sensitive. Analysing the rows in the figure shows that decreasing $\Delta\epsilon$ decreases the diagonal averaging kernel values and increases its off-diagonal values.

These averaging kernels show that the lower RMS error of the high CL and low $\Delta\epsilon$ cases comes at the price of sensitivity to the true emissivity. In an ideal case the averaging kernel is a straight diagonal line. The more spread out the edges of this line are the more a-priori information is being used.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021


Atmospheric Measurement Techniques Discussions

The main conclusion of this analysis is that there is no abrupt transition in the explored a-priori uncertainty parameter space. All the parameter choices produced similar retrieval results, with differences only in less-sensitive regions. Therefore a choice in either direction will either give slightly more sensitivity or accuracy, and can be tuned to match the specific need of the user.

An additional option is to use a-posteriori regularization. Using a larger error and a smaller correlation length would be desirable to give the retrieval more precision and freedom. As seen in this Appendix, due to the ill conditioning of the retrieval the weaker regularization would cause the solution to oscillate more. An a-posteriori regularization method such as the IVS (Iterative Variable Strength) method introduced in Ridolfi and Sgheri (2011) and applied to FORUM atmospheric profile retrievals in Sgheri et al. (2020) could be used to smooth out these unphysical oscillations.

For the purpose of this study $\Delta\epsilon = 0.1$ and CL$= 50$ was used as the baseline parameter combination that represents a compromise between the two extremes of sensitivity and accuracy.

Author contributions. All authors contributed to the paper through discussions and comments

- Maya Ben-Yami wrote the paper
- Maya Ben-Yami and Hilke Oetjen devised the experiments
- Luca Sgheri, Piera Raspollini, Tiziano Maestri, William Cossich and Davide Magurno wrote the FEES code and helped in running the simulator
- Dulce Lajas was the ESA FEES code quality manager
- Laura Warwick and Helen Brindley provided the downwelling radiation values for Figure A1

Competing interests. No competing interests are present.

Acknowledgements. The work presented in this paper benefited significantly from discussions with the ESA FORUM Mission Advisory Group. The FEES was developed under ESA contract no. 4000123810. MBY acknowledges funding through the ESA Young Graduate Trainee program. LW acknowledges funding through the CASE partnership between EPSRC and the National Physical Laboratory.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Atmospheric Measurement Techniques Discussions

References

Baldridge, A., Hook, S., Grove, C., and Rivera, G.: The ASTER spectral library version 2.0, Remote Sensing of Environment, 113, 711–715, https://doi.org/https://doi.org/10.1016/j.rse.2008.11.007, 2009.

Bellisario, C., Brindley, H. E., Murray, J. E., Last, A., Pickering, J., Harlow, R. C., Fox, S., Fox, C., Newman, S. M., Smith, M., Anderson, D., Huang, X., and Chen, X.: Retrievals of the Far Infrared Surface Emissivity Over the Greenland Plateau Using the Tropospheric Airborne Fourier Transform Spectrometer (TAFTS), Journal of Geophysical Research (Atmospheres), 122, 12,152–12,166, https://doi.org/10.1002/2017JD027328, 2017.

Capelle, V., Chédin, A., Péquignot, E., Schlüssel, P., Newman, S. M., and Scott, N. A.: Infrared Continental Surface Emissivity Spectra and Skin Temperature Retrieved from IASI Observations over the Tropics, Journal of Applied Meteorology and Climatology, 51, 1164–1179, https://doi.org/10.1175/JAMC-D-11-0145.1, 2012.

Clough, S., Shephard, M., Mlawer, E., Delamere, J., Iacono, M., Cady-Pereira, K., Boukabara, S., and Brown, P.: Atmospheric radiative transfer modeling: a summary of the AER codes, Journal of Quantitative Spectroscopy and Radiative Transfer, 91, 233–244, https://doi.org/https://doi.org/10.1016/j.jqsrt.2004.05.058, 2005.

Dinelli, B. M., Castelli, E., Carli, B., Del Bianco, S., Gai, M., Santurri, L., Moyna, B. P., Oldfield, M., Siddans, R., Gerber, D., Reburn, W. J., Kerridge, B. J., and Keim, C.: Technical Note: Measurement of the tropical UTLS composition in presence of clouds using millimetre-wave heterodyne spectroscopy, Atmospheric Chemistry & Physics, 9, 1191–1207, https://doi.org/10.5194/acp-9-1191-2009, 2009.

Feldman, D. R., Collins, W. D., Pincus, R., Huang, X., and Chen, X.: Far-infrared surface emissivity and climate, Proceedings of the National Academy of Science, 111, 16 297–16 302, https://doi.org/10.1073/pnas.1413640111, 2014.

Harries, J., Carli, B., Rizzi, R., Serio, C., Mlynczak, M., Palchetti, L., Maestri, T., Brindley, H., and Masiello, G.: The Far-infrared Earth, Reviews of Geophysics, 46, https://doi.org/https://doi.org/10.1029/2007RG000233, 2008.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Huang, X., Chen, X., Zhou, D. K., and Liu, X.: An Observationally Based Global Band-by-Band Surface Emissivity Dataset for Climate and Weather Simulations, Journal of Atmospheric Sciences, 73, 3541–3555, https://doi.org/10.1175/JAS-D-15-0355.1, 2016.

Huang, X., Chen, X., Flanner, M., Yang, P., Feldman, D., and Kuo, C.: Improved Representation of Surface Spectral Emissivity in a Global Climate Model and Its Impact on Simulated Climate, Journal of Climate, 31, 3711–3727, https://doi.org/10.1175/JCLI-D-17-0125.1, 2018.

Knuteson, R., Best, F., DeSlover, D., Osborne, B., Revercomb, H., and Smith, W.: Infrared land surface remote sensing using high spectral resolution aircraft observations, Advances in Space Research, 33, 1114–1119, https://doi.org/https://doi.org/10.1016/S0273-1177(03)00752-X, climate Change Processes in the Stratosphere, Earth-Atmosphere-Ocean Systems, and Oceanographic Processes from Satellite Data, 2004.

Kuo, C., Feldman, D. R., Huang, X., Flanner, M., Yang, P., and Chen, X.: Time-Dependent Cryospheric Longwave Surface Emissivity Feedback in the Community Earth System Model, Journal of Geophysical Research: Atmospheres, 123, 789–813, https://doi.org/https://doi.org/10.1002/2017JD027595, 2018.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

Li, Z.-L., Wu, H., Wang, N., Qiu, S., Sobrino, J. A., Wan, Z., Tang, B.-H., and Yan, G.: Land surface emissivity retrieval from satellite data, International Journal of Remote Sensing, 34, 3084–3127, https://doi.org/10.1080/01431161.2012.716540, 2013.

Masiello, G. and Serio, C.: Simultaneous physical retrieval of surface emissivity spectrum and atmospheric parameters from infrared atmospheric sounder interferometer spectral radiances, Appl. Opt., 52, 2428–2446, https://doi.org/10.1364/AO.52.002428, 2013.

Murray, J. E., Brindley, H. E., Fox, S., Bellisario, C., Pickering, J. C., Fox, C., Harlow, C., Smith, M., Anderson, D., Huang, X., Chen, X., Last, A., and Bantges, R.: Retrievals of High-Latitude Surface Emissivity Across the Infrared From High-Altitude Aircraft Flights, Journal of Geophysical Research (Atmospheres), 125, e33672, https://doi.org/10.1029/2020JD033672, 2020.

Palchetti, L., Di Natale, G., and Bianchini, G.: Remote sensing of cirrus cloud microphysical properties using spectral measurements over the full range of their thermal emission, Journal of Geophysical Research (Atmospheres), 121, 10,804–10,819, https://doi.org/10.1002/2016JD025162, 2016.

Palchetti, L., Brindley, H., Bantges, R., Buehler, S. A., Camy-Peyret, C., Carli, B., Cortesi, U., Bianco, S. D., Natale, G. D., Dinelli, B. M., Feldman, D., Huang, X. L., C.-Labonnote, L., Libois, Q., Maestri, T., Mlynczak, M. G., Murray, J. E., Oetjen, H., Ridolfi, M., Riese, M., Russell, J., Saunders, R., and Serio, C.: FORUM: Unique Far-Infrared Satellite Observations to Better Understand How Earth Radiates Energy to Space, Bulletin of the American Meteorological Society, 101, E2030 – E2046, https://doi.org/10.1175/BAMS-D-19-0322.1, 2020.

Palchetti, L., Barucci, M., Belotti, C., Bianchini, G., Cluzet, B., D'Amato, F., Del Bianco, S., Di Natale, G., Gai, M., Khordakova, D., Montori, A., Oetjen, H., Rettinger, M., Rolf, C., Schuettemeyer, D., Sussmann, R., Viciani, S., Vogelmann, H., and Wienhold, F. G.: Observations of the downwelling far-infrared atmospheric emission at the Zugspitze observatory, Earth System Science Data Discussions, 2021, 1–18, https://doi.org/10.5194/essd-2020-377, 2021.

Ridolfi, M. and Sgheri, L.: Iterative approach to self-adapting and altitude-dependent regularization for atmospheric profile retrievals, Opt. Express, 19, 26 696–26 709, https://doi.org/10.1364/OE.19.026696, 2011.

Ridolfi, M., Del Bianco, S., Di Roma, A., Castelli, E., Belotti, C., Dandini, P., Di Natale, G., Dinelli, B. M., C.-Labonnote, L., and Palchetti, L.: FORUM Earth Explorer 9: Characteristics of Level 2 Products and Synergies with IASI-NG, Remote Sensing, 12, https://doi.org/10.3390/rs12091496, 2020.

Rodgers, C. D.: Retrieval of Atmospheric Temperature and Composition From Remote Measurements of Thermal Radiation, Reviews of Geophysics and Space Physics, 14, 609, https://doi.org/10.1029/RG014i004p00609, 1976.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, WORLD SCIENTIFIC, https://doi.org/10.1142/3171, 2000.

Sgheri, L., Raspollini, P., and Ridolfi, M.: Auto-adaptive Tikhonov regularization of water vapor profiles: application to FORUM measurements, Applicable Analysis, 0, 1–11, https://doi.org/10.1080/00036811.2020.1751825, 2020.

Sgheri, L., Belotti, C., Ben-Yami, M., Bianchini, G., Dominguez, B., Cortesi, U., Cossich, W., Del Bianco, S., Di Natale, G., Guardabrazo, T., Lajas, D., Maestri, T., Magurno, D., Oetjen, H., Raspollini, P., and Sgattoni, C.: The FORUM End-to-End Simulator project: architecture and results, submitted to Atmospheric Measurement Techniques, 2021.

Susskind, J., Blaisdell, J. M., and Iredell, L.: Improved methodology for surface and atmospheric soundings, error estimates, and quality control procedures: the atmospheric infrared sounder science team version-6 retrieval algorithm, Journal of Applied Remote Sensing, 8, 1 – 34, https://doi.org/10.1117/1.JRS.8.084994, 2014.

Transtrum, M. K., Machta, B. B., and Sethna, J. P.: Geometry of nonlinear least squares with applications to sloppy models and optimization, Phys. Rev. E, 83, 036 701, https://doi.org/10.1103/PhysRevE.83.036701, 2011.

https://doi.org/10.5194/amt-2021-232
Preprint. Discussion started: 7 September 2021

von Clarmann, T., Degenstein, D. A., Livesey, N. J., Bender, S., Braverman, A., Butz, A., Compernolle, S., Damadeo, R., Dueck, S., Eriksson, P., Funke, B., Johnson, M. C., Kasai, Y., Keppens, A., Kleinert, A., Kramarova, N. A., Laeng, A., Langerock, B., Payne, V. H., Rozanov, A., Sato, T. O., Schneider, M., Sheese, P., Sofieva, V., Stiller, G. P., von Savigny, C., and Zawada, D.: Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature, Atmospheric Measurement Techniques, 13, 4393–4436, https://doi.org/10.5194/amt-13-4393-2020, 2020.

Wan, Z.: New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product, Remote Sensing of Environment, 140, 36–45, https://doi.org/https://doi.org/10.1016/j.rse.2013.08.027, 2014.

Wang, K., Wan, Z., Wang, P., Sparrow, M., Liu, J., Zhou, X., and Haginoya, S.: Estimation of surface long wave radiation and broadband emissivity using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature/emissivity products, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/https://doi.org/10.1029/2004JD005566, 2005.