Emissivity Retrievals with FORUM's End-to-end Simulator: Challenges and Recommendations | Author reply to referee comments

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)

Black text - referee comments

Blue text - author reply

GENERAL COMMENT

The reviewers have rightly commented on the fact that this paper does not strictly follow the "problem->solution" format of standard publications, but instead is an investigation which poses open questions. This is due to the circumstances in which it was produced - the lead author was only part of the FORUM team for one year during her time at ESA, and has since started a PhD in a completely different topic. Therefore we have chosen to include open questions in the paper as any investigation that is not included will most likely be lost. It is also worth noting that very little work has been done so far on FIR emissivity retrievals, and we hope that the detailed explanations of this paper's methods can be of use to a burgeoning field.

The reviewers' comments have been a great help in improving the structure and content of the paper, and we hope that the new version will allow readers to choose the section that is relevant to them, and that the details will now be easier to understand. We also hope that the modifications made to the conclusions will help clarify our points and have made the paper easier to read. We thank the reviewers for their help with this process.

REPLY TO REFEREES

Referee Comment 1

Thank you for sharing this work with the community. This paper represents good progress in the demonstration of the mathematical machinery which can be applied to the FORUM mission radiance dataset.

The main problem I see is that the term OE is used many times but the actual results do not include a climatology for surface emissvity so therefore this work is not Optimal Estimation nor is it the Maximum A Posteriori (MAP) method described in Rodger's paper. I believe it is more correct to refer to the method described as Minimum Information or Generalized Least Squares (GLS) using a diagonal matrix with a spectrally constant value. In other words this is the result you would get if you did not have any prior information about the land surface nor did you know anything about the spectral correlations of the land surface. For that reason it represents a kind of "worse case" scenario for retrieving information. As if you were flying FORUM on a different unknown planet somewhere outside the solar system. In reality the actual Earth's infrared surface emissivty is slowly varying both spectrally and temporally for any given latitude/longitude. The whole point of the FORUM mission is to provide the calibrated radiances globally for a long enough time (at least one year) that the radiance dataset can be used to derive a climatology of FarIR emissivity for a grid of latitudes/longitudes by month of the year. However the method described is far from optimal for doing that. For this reason I suggest that the abstract and conclusions be modified to indicate that the results are expected to greatly improve AFTER the FORUM mission is concluded and a climatology is derived (by some to-be-determined manner). Perhaps you can suggest some kind of timeline whereby currently you don't have a priori information so you can't use an OE (or MAP) method but by the end of the mission you hopefully can demostrate an OE method that uses a priori information. This is just the first step and demonstrates that the machinery of the mathematics has been coded correctly and the somewhat disappointing results are expected to improve by the end of the mission.

Thank you for this input - we were not clear enough in the text on the ways that we deviate from the OE method. We have rewritten Section 3.3 in a way that we hope addresses all the raised issues. However we do have a few points to make in reply to the reviewer's comment:

- The OE estimation method in the Bayesian framework does not specify the use of a climatology, only that a-priori information be used to constrain the retrieval. In taking the a-priori to be 1 and constraining the emissivity in the 0-1 region we are in fact using a-priori knowledge on what emissivity could be.
- The difference between this retrieval and one on an unknown planet is that we do use climatologies for the rest of the retrieval vector. The retrieval method doesn't distinguish between water vapour, temperature profile, surface temperature and emissivity in the retrieval vector - they are one quantity $\mathbf{x}$ that is being fitted. Our reasoning for not

using climatologies for surface emissivity is that with the large uncertainty that we use it doesn't make a difference. This has been tested by us and others, and using a climatology for emissivity did not change our results. However *the retrieval is strongly influenced by the climatologies used for the rest of the a-priori vector*. We have clarified in the text (section 3.3) that we are in a way "adding" a slightly differently constrained "emissivity bit" to a retrieval that without it would be exactly the classical OE from Rodger's.

- It is shown in the manuscript that the necessary emissivity a-priori uncertainty is so large that the main constraint on emissivity is the surface temperature a-priori, *which is climatological*. Finding a way to use climatological emissivity information in the retrieval would of course be a good step towards improving it, but the fact is that our results make extensive use of climatological information.

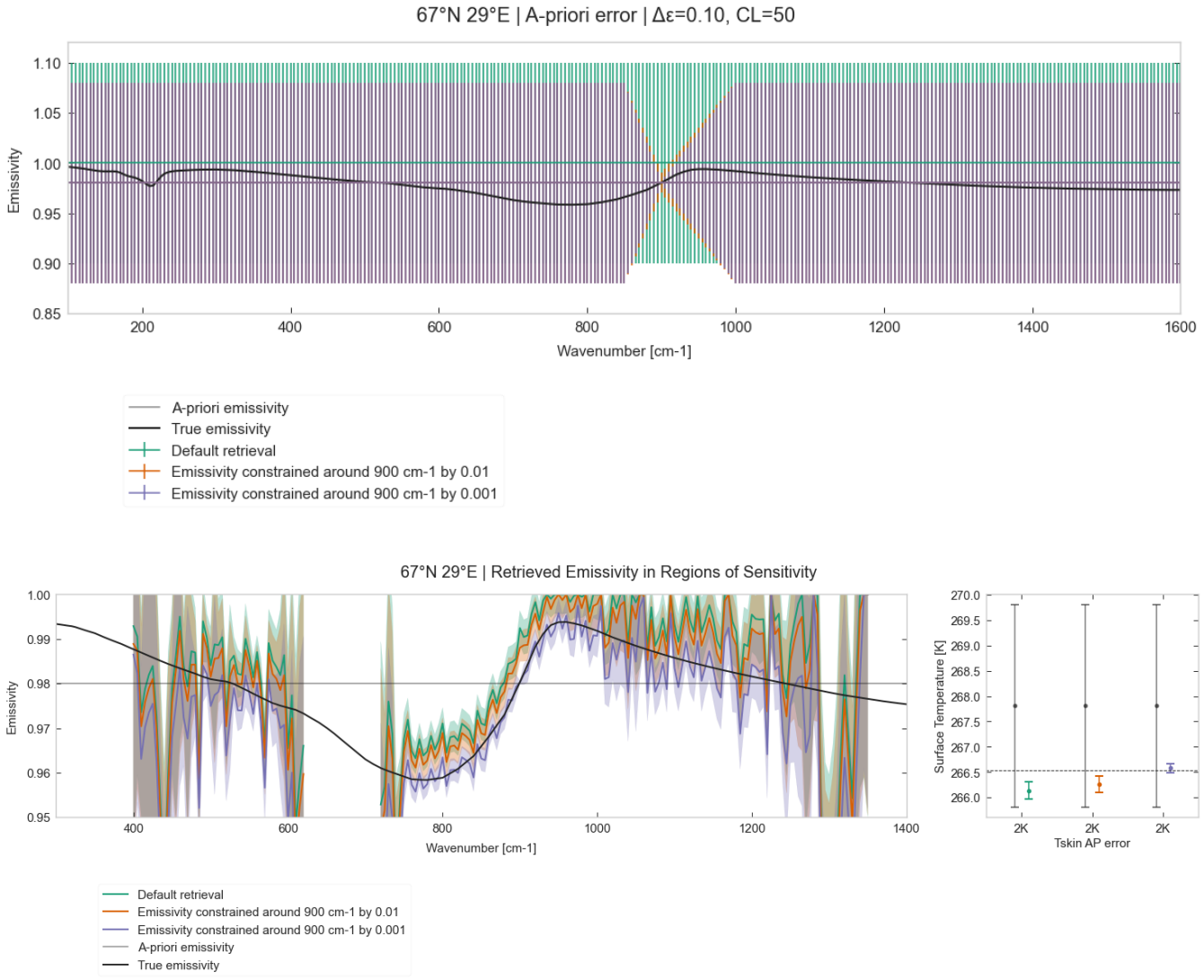
- We have clarified in the paper that integrating emissivity into the OE retrieval for the FIR has not been done before, and so this paper can be thought of as an experimental initial foray into an extension of the OE emissivity method, and an a-priori of 1 is just a first choice that can of course later be improved.

Comments:

1) In Section 3.3 I suggest rewriting and expanding the first paragraph to state a bit more clearly the role that the xa and Sa play in the OE method. I disagree that the Xa is the a priori knowledge and the Sa is the uncertainty. I don't think that is a proper interpretation. In fact the Sa represents the variance of the a priori climatology so for surface emissivity this is the natural variability of the actual surfaces included in the dataset. In particular, if you restrict the climatology to Snow/Ice scenes (by using some image classification for example) then the Sa matrix will "constrain" the solution so that it is consistent with the natural variability each spectra channel. This is critically important because as shown in Figure 7 the snow/ice emissivity is close to unity at 960cm-1 and the variablity is necessarly smallest there. In that case the Sa matrix will weight the spectral channels so that the fitted surface temperature is derived mainly from the 960 cm-1 region and that greatly reduces the error in the Ts. In that case the xa (guess) value at 960 cm-1 has a small variance and thus allows for the Ts to be derived in an true OE method with a physical constraint. This is how you narrow the "sloppy valley", kind of like a ford in the stream which provides the most obvious place to cross. At the same time the off-diagonal elements of Sa provide the spectral correlation for physical surface emissivities which is critical for estimating the surface emissivity in the FarIR using microwindows (mainly) that do not fully constrain the solution by themselves. So rather than just dismissing the entire purpose of the OE approach I think you could add some sentences that say why you want to use the OE method with a real climatology but you don't have one yet.

We have rewritten section 3 to better define Xa and Sa as suggested, and have clarified the difference between our approach and optimal estimation. We also appreciate the suggestion to constrain emissivity at certain wavelengths. We did try this at 900cm-1 and the results can be seen in the figures shown here below. However an unreasonable constraint of 0.001 was needed to improve the shift caused by the correlation with surface temperature. Such a large constraint

would result in a large error in emissivity if the scene is misclassified, and in this work we chose to prioritize the general flexibility of the method and so opted to not use such constraints.



The issue you point out that a single global Sa matrix can not distinguish between quartz and snow is true but the solution is not to throw out OE the better approach, much better is to use scene classification and develop climatologies for each scene type or mixture of scene types. The approach for doing this "divide and conquore" approach is described in the following references and is already implemented in RTTOV 12+ for the MidIR. The same approach should work for the FarIR.

Borbas, E.E., Hulley, G., Feltz, M., Knuteson, R. and Hook, S., 2018. The combined ASTER MODIS emissivity over land (CAMEL) part 1: Methodology and high spectral resolution application. Remote Sensing, 10(4), p.643.

Feltz, M., Borbas, E., Knuteson, R., Hulley, G. and Hook, S., 2018. The combined ASTER MODIS emissivity over land (CAMEL) part 2: Uncertainty and validation. Remote Sensing, 10(5), p.664.

Loveless, M., Borbas, E.E., Knuteson, R., Cawse-Nicholson, K., Hulley, G. and Hook, S., 2021. Climatology of the Combined ASTER MODIS Emissivity over Land (CAMEL) Version 2. Remote Sensing, 13(1), p.111.

We agree with the reviewer that a few years after launch of FORUM, the emissivity climatologies can be improved and a better emissivity product can be obtained, not just from FORUM, but also other FIR measurements. We have added this plan for future work at the end of section 3.3 with references to the MODIS retrievals.

2) In the conclusions I suggest that you change the wording slightly to emphasize that the current results are what would be possible without an a priori climatology of surface emissivity BUT at the end of the FORUM mission you expect that the results will improve based on the development of climatologies that are consistent with the FORUM radiances.

We agree, and have added this point to the conclusions.

Referee Comment 2

I have read the paper "Emissivity Retrievals with FORUM's End-to-end Simulator: Challenges and Recommendations" by Ben-Yami et al. with interest. It is clear that this paper is the result of very careful and extensive work, and I congratulate the authors for this. Most of the results are interesting (but see below) and I also liked the figures. Despite this, I have several general comments that could lead to an improved paper (see below). I leave it up to the authors to decide whether and to what extent these general comments are addressed. They are given merely to help increasing the impact of the paper. Even without addressing these, I think the paper is suitable for publication in AMT after minor revision.

General comments

(i) Explanations are not always expressed very clearly. Very frequently I found myself reading a paragraph twice or even three times to understand. The paper is also too long-winded, especially near the end. It is also repetitive at times. In all, I think the paper would really benefit from careful copy-editing, to increase its impact and make the paper more pleasurable to read (and to increase the chances that readers actually make it to the end).

We have rewritten parts of the text as well as the abstract and conclusions.

(ii) It is not always clear what the point is of the analysis or whether what is presented is really new. Some of the analysis is more appropriate for a thesis, or a technical report. A paper is not meant to show all results obtained in a study, but should present those results that are new and useful to the community. As an example, consider section 6.3 - did the reader really learn something? From the concluding paragraph (lines 509-515) one would say no. A ruthless shortening of the paper (both with respect to its wordiness and the material presented) would be highly beneficial in my opinion.

Please see the general comment above. Very little work has been done so far on FIR emissivity retrievals, so in fact all the results presented in this work are new. We agree though that section 6.3 was out of place in the main body of the paper as it did not contribute substantially to the conclusions, and so have moved it to Appendix C.

(iii) the study is geared towards FORUM and the specific tools that have been developed, but is this paper going to be useful to anyone else? (especially since it is mentioned several times that the paper is just a step in the development of a future product).

(iv) Point (2) in the abstract illustrates these three points. It is hard to read, difficult to understand what is meant, and not sure that it is useful to anyone but the developing team of FORUM.

(To both points above):

It takes many years to design, revise and implement a processor for an operational product of an Earth Explorer mission. This paper is an important step for the FORUM emissivity product. In addition, as FIR emissivity retrieval is a novel and expanding field, we hope it will prove useful for the development of other FIR mission products in the future.

We are aware that the results presented are not as helpful for the target user of the emissivity product. However as the FORUM spectra will be freely available there can always be scientific developments based on these independently from the provided emissivity product.

Specific Comments:

Line 2: its FIR contribution > that the FIR is..

Changed.

Line 10: absolute or relative uncertainty? is this

Added the clarification "absolute".

Line 13: not sure if I agree with the first recommendation (even after reading the paper). Why limit the range artificially with an ad hoc threshold? Why not simply provide the emissivity over the entire range together with the estimated uncertainty (or information content quantifier) and let the user decide what to use?

On further consideration we agree with the reviewer - the emissivity could be provided over the entire range for the user, and this is now explained in section 4.3. However for our analysis we still consider it important to use a criterion for retrieved emissivity, but now in section 4.3 this is only introduced as a part of the analysis and not as a major recommendation.

Line 15: Point (2) is very hard to understand, in particular the "Thus" - how does the second sentence follow from the first (which doesn't even mention uncertainty). An abstract should be understood by itself, without the need for going through the entire paper. If i understand it correctly, what you want to say is that because emissivity and temperature cannot be retrieved independently, a priori knowledge on at least one or both is needed - however errors on e.g. the a priori of the Tskin, might result in unrealistic uncertainty estimates; and a way has to be found to improve these estimates. If this is what is said, it is a rather weak way to close an abstract.

Abstract: overall, the abstract could be much stronger and would highly benefit from copy editing. It would also benefit from a few strong, easy to understand take-home messages.

We agree that the abstract was difficult to understand, and have edited it accordingly.

The sentence on line 15 was badly written. It was not stating that the estimates have to be improved, as we are in fact happy with this level of spread on the estimates. The problem is not in the spread but that the uncertainty estimation produced by the FEES does not include this spread.

As we agree with the reviewer that our two recommendations at the end of the abstract were too complicated for the abstract, we have removed them and changed the final sentence to "Based on these investigations a road-map is recommended for the development of the operational emissivity product". This road-map is now detailed in the conclusions in a way that we hope is easier to understand.

Introduction: I didn't see this paper cited https://doi.org/10.1175/BAMS-D-20-0155.1 which could be relevant

We have added a mention of the PREFIRE mission and this citation.

Line 25: "never been observed"; not exactly, as there have been aircraft measurements. Perhaps add "from satellite"

Added "from satellite".

Line 34-35: a bit strange to say that it is exciting to measure because only one instrument can measure it. There are many things unique that are very unexciting. The second reason is much more convincing.

Changed to "particularly exciting given its potential influence on the surface and top of atmosphere energy budget"

Line 95: I would move this to section 3.2, as it is there that the retrieval parameters are discussed.

Moved to the start of Section 3.2

Line 107: This deserves a reference or justification

We have added a citation to Bellisario et al. 2017, who discussed this question. The sentence now reads "Following the reasoning from Bellisario et al. 2017 in this work the emissivity is always assumed to have no directional dependence."

Line 115: I would move the start of Section 3 to the end of the introduction

Moved to the introduction (changed SGM/L2M to forward/retrieval for clarity).

Line 177 spectral region considered > considered spectral region

Done.

Line 187: to this end four > to this end, four

Added comma.

Line 192: "The reason of the focus on the water vapour" .. this sentence is very important, but should be mentioned much earlier in the manuscript (in the introduction), because water vapour is mentioned over and over, without saying why it is important.

Sentence moved to introduction.

Figure 3: This is a very nice figure, but here are a couple suggested improvements: (1) Panel a: plot 1 sigma uncertainties as a semi-transparent shaded region, because as-is, the error bars draw all the attention.

Done.

(2) I would add a 5th panel with the total transmittance (surface to TOA) as a function of wavenumber, since you discuss transmittance many times, without ever showing it. This property is probably available from the forward model without much effort. On this plot you could also indicate the main spectral ranges of absorption due to the different species (O3, CO2, H2O, CH4, ...)

We have added a 5th panel for the total transmittance, and have indicated the main absorbers in the Figure.

As a side comment, the paper emphasizes strongly on H2O but does not mention the other strong absorbers - I think it is a shame not too. This does not need to be long, but you could for instance raise the following question: With current trends in CO2 increases, can we expect that the retrieval range will become more narrow by the time FORUM is operational?

This is indeed an interesting point, but the effect of other absorbers is expected to be minimal as the emissivity is smoothly varying with wavenumber. As the paper is already very long and complex we have chosen not to add an analysis of the other absorbers.

(3) For the bottom panel, how about just displaying the diagonal of the averaging kernel? You mention that it is not smooth (line 267), but from the plot this does not seem to be the case for the diagonal. The diagonal will probably look similar to the IQ.

We have added the diagonal as a bold line. The reviewer is right to point out that it looks similar to the IQ, and in fact that choice of IQ instead of AK diagonal for the purpose of a "threshold" does not make a significant difference.

Line 236: "below the CO2" - might be unclear to some readers, but see previous comment on inclusion of extra panel.

We have now added an indication of the CO2 band in the figure and an approximate wavenumber in the text.

Figure 4: since a discrete number of of pwv are used, it would be good to use a discrete colorbar with the same number of colors.

Done.

Figure 5: From figure 5, and the visible "noise" in the retrieval, it can be argued that the retrieval is not smooth enough on the one hand, but gives too much importance to long range correlations. One way of tuning this is with the correlation length, as the paper does, but the results in Figure 5 suggest that the formula used for the covariance matrix (line 614) is far from optimal. Perhaps replacing Delta/CL with (Delta/CL)**n would already do it, with n>1 so that the short range correlations become relatively more important than the long). The value of CL would need to be reevaluated. You could optimize by minimizing the RMS for different choices of n and CL (like you do in appendix B).

This is an interesting point. However in the current retrieval method the covariance matrix does not have a very large influence on the results. This is not intuitive, but in fact the covariance matrix is only used where the algorithm relies on a-priori information, and it only does that in regions where there is less sensitivity to the surface. So in the regions of sensitivity, which are the regions of interest, it has little to no effect on the retrieval. In most regions of importance the noise is due to a combination of instrumental noise, microwindow transmittance and the retrieval grid. This can be seen in Figures B1 and B2 in the appendix.
However we would also like to improve the "noise" on the retrieval, and as we note in that appendix, there is an a-posteriori regularization method called IVS (Iterative Variable Strength) which can smooth the oscillations. This was introduced in Ridolfi and Sgheri (2011) and applied to synthetic FORUM atmospheric profile retrievals in Sgheri et al. (2020), and Luca Sgheri is implementing it for emissivity retrievals in the FEES as well (see Sgheri et al. 2021).

Figure 5: Please add the value of the RMS to each of the panels

Done.

Section 5: The word "extent" is a bit of misnomer, as it doesn't express the width of the interval. I would simply replace it with "lowest wavenumber" everywhere in figure, table and text.

Replaced everywhere.

Line 327: Only pwv is mentioned while obviously some other species are equally important (see comment above on the other species).

The focus on pwv in this work is due to its importance in the FIR and that it can also retrieved from FORUM observations. The sentence was misleading though, and we have changed it to: "The contribution of water vapour to $\mathcal{T}(z)$ was discussed in Sections ...".

Caption of Figure 7: please shorten. Only at the end of the caption the reader actually learns what the lines in the main plot represent. No need to give needless details (like the time of the observation) or repeat everything that is in the text, or what should be clear from the legend (like that the dashed line corresponds to the true emissivity).

The caption has been shortened.

"same in all four cases": do you mean exactly the same, this is very surprising!

This phrase was referring to the synthetic observations generated by the SGM used to run this retrieval - these were the same because we used the same run for all four retrievals, as the point was to test different retrieval parameters in the same scene. The text has been changed to "Panel (a) shows the difference between the same synthetic observations and the four different converged forward models" for clarity.

uncertainty > uncertainty.

Changed

Figure 7: the legend in the figure (i)-(iv) could be improved, e.g. true +/- 0.5K etc would be clearer if it is written as, True a priori, +/-0.5 K uncertainty.

Done.

Line 333-335: this is incorrect in the sense that we have a robust lower bound on the temperature (corresponding to an assumed emissivity of 1, whereas lower temperature would need an emissivity above 1). The statement is correct for temperatures higher than the actual temperature.

The text was in fact incorrect, and has now been changed to: "The constraint that the surface emissivity $\epsilon_\sigma \leq 1$ checks this degeneracy and provides a lower bound for the retrieved T_s. However for any higher values of T_s it is possible to find a corresponding surface spectral emissivity ϵ_σ that produces the correct surface radiance."

Line 337: unclear as you already use MIR? Please also define what you understand under MIR, as there are many definitions.

The text was in fact misleading, as the MIR measurements we had in mind were the independent IASI-NG observation. The text has been changed to : "Future work could investigate such methods by incorporating independent MIR measurements from synergy with IASI-NG in tandem with the full-spectrum simultaneous OE retrieval used in this work".

Section 6.1: I believe that temperature and emissivity will be easier to disentangle once you find a more appropriate covariance matrix per comment above.

See reply to comment above - constraint on the temperature comes from regions of sensitivity to the surface, where the a-priori uncertainty covariance matrix has little influence.

Line 362 "the the"

Removed second "the".

End of section 6.1 + section 6.2: I do not understand the choice of a priori value of 1 for the emissivity. Why not say a constant 0.98 or 0.99? This will in almost all cases be closer to true emissivity, and will therefore: (1) result in faster convergence and (2) result in retrieval results closer to true value (as your a priori value is closer to the truth). It is a free gain.

What we show in section 6.2 is that the AP choice doesn't make a big difference when the AP uncertainty is 0.1 - see also the Appendix. We agree 0.99 or 0.98 would also work, but we used 1 for consistency and to be compatible with previous tests. See attached figure that shows that 0.98 vs 1 would not make any noticeable difference in the retrieval.

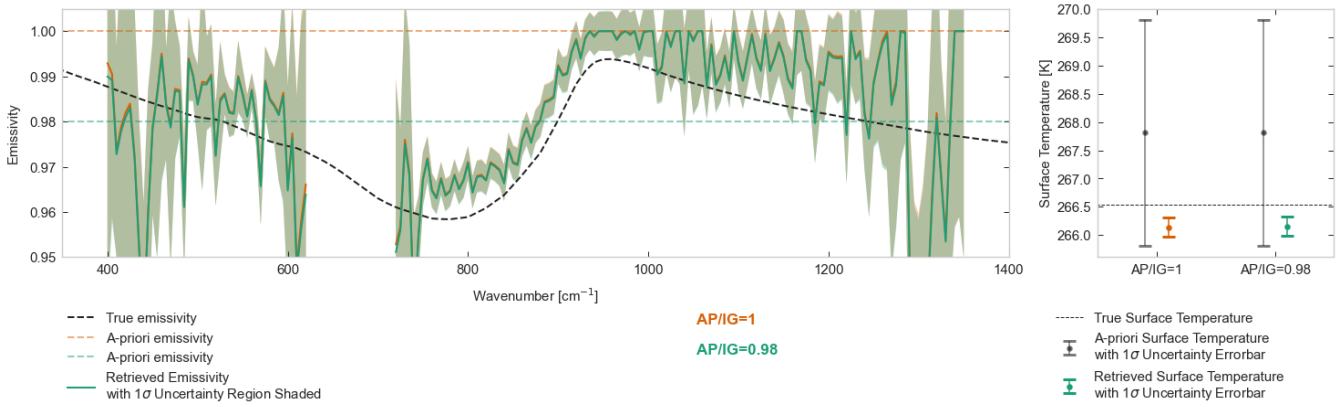


Figure 8: not a big fan of this figure, as it is very hard to read / interpret. First, the study of initial guess is not that interesting/relevant as there is little reason to take the initial guess different from the a priori. Easier, more comprehensive and more interesting would be a combined analysis of the a priori and constraint. How about a scatter plot with X-axis the a priori (say 0.8-1 in steps of 0.01), Y-axis the a priori uncertainties and then the datapoints colorcoded by RMS (averaged over the 4 cases)?

This is a very good point, and we have reconsidered the figure. However,

- In terms of the retrieval optimisation, the sensitivity to the initial guess is in fact interesting precisely because it is not expected. The sensitivity shown in this figure is a possible "red flag" that should be investigated.

- The figure suggested by the referee would be very interesting, but unfortunately requires significantly more runs than is computationally feasible with the FEES. We leave this study to research teams investigating synthetic FORUM retrievals with fast RTM codes.
- Averaging the 4 geographical cases would mean we lose interesting information on their differences.

We have therefore opted to simplify the existing figure by averaging the uncertainty over the FIR and the MIR, and plotting the uncertainty values for all the parameter combinations. We hope that this makes the figure easier to interpret.