

Answers to reviewer 1

October 31, 2025

This paper describes an intercomparison of the properties of cirrus measured by in-situ probes vs those retrieved from a ground-based infrared spectrometer that spans from the mid-infrared to the far-infrared (i.e., from 1500 cm^{-1} to 400 cm^{-1}). The authors use a ceilometer to ascertain the base and top of the cirrus clouds, and to estimate their optical depth. Lastly, their infrared retrieval framework also retrieves thermodynamic profiles, which they compare against ERA5 model output and retrieved profiles from a collocated ground-based microwave radiometer. The paper achieves what it sent out to do, namely, to perform an intercomparison of derived cirrus properties, which includes an evaluation of the assumption of the habit of the ice particles, and thermodynamic profiles. It is well written, and the references included are sufficient.

We sincerely thank the reviewer for providing a precise and insightful summary of our manuscript. The careful evaluation not only highlighted aspects requiring clearer explanation but also recognized the validity and potential impact of our work. We greatly appreciate the thoughtful comments, which have helped us improve the clarity and strength of the manuscript.

My main concern is that this intercomparison evaluate 11 samples that were collected within a 30-minute period on a single day at a single location. Thus, I must ask: how representative are these results? Would these same results hold for different atmospheric state conditions (i.e., temperature and humidity profiles) or cirrus conditions (i.e., optically thicker clouds, cirrus generated by other means, etc). This is further complicated by the sky image shown in Figure 2, which suggests that their in-situ aircraft is leaving a contrail (based upon the flight pattern on the right and the clear spiral signature in the sky image). Can we assume the ice habit in the contrail is consistent with the habit of the background cirrus? How much is the contrail impacting either the radiometric observations made by the FINESSE or the in-situ obs (when aircraft starts its next circle)? Additionally, the ice optical depths sampled here are small; less than 0.2 – so again, do the conclusions (e.g., the importance of the far-infrared channels) hold when in cases when the optical depth is larger? Before this paper should be accepted, I would hope that the authors can address these concerns well.

We understand the reviewer’s concern and we agree that measurements limited to a time interval of less than one hour are not representative of a wide range of atmospheric and cloud scenarios. However, the purpose of this work is to assess as much as possible the consistency of spectral radiance simulations using currently available databases of the optical properties of ice clouds, in combination with both in situ and remote sensing measurements of high-altitude ice clouds. Unfortunately, the great difficulty of performing such measurements — especially in situ observations with aircraft — means that they are scarce and always time-limited. The availability of in situ aircraft measurements of the microphysics of ice particle size distributions, together with an estimate of crystal habits within the cloud, represents a unique asset that must be exploited as much as possible, especially when simultaneous remote sensing measurements across different spectral bands are available, even though the amount of data is not as large as desired.

In addition, the capability to correctly retrieve cirrus properties under a wide range of both atmospheric conditions (temperature and water vapor) and cloud conditions (particle size and optical depth) has already been clearly demonstrated through results that are partly published and partly in preparation, based on the analysis of measurements from the REFIR-PAD spectrometer (a Fourier spectrometer similar to FINESSE covering the FIR spectrum) at Dome-C (Antarctica) in the presence of a backscattering and depolarization lidar. Within such studies, test of consistency have already performed by using various ground-based measurements including also radar, covering different cloud scenarios, but never available unique and precious in situ measurements from aircraft or balloon sounding inside the clouds. In previously published studies on the analysis of REFIR-PAD data, (e.g. Di Natale et al. (2020); Maestri et al. (2014)), a non-perfect matching of the residuals in the FIR has already been observed — although it is less evident due to the higher instrumental noise of REFIR-PAD compared to FINESSE in the spectral range above 400

cm^{-1} . This same result is also found in the present work, even though the residuals are averaged over a much more limited set of measurements than those collected at Dome-C.

From REFIR-PAD measurements, we were able to produce a Level-2 dataset of atmospheric profiles and properties of ice and mixed-phase clouds over 11 years of observations, from 2012 to 2022, covering all seasons with a measurement frequency of 12 minutes. The retrieval products have been thoroughly validated using local radiosonde data, and consistency in the microphysical properties of precipitating clouds has been confirmed through measurements from an ICE-CAMERA — capable of collecting precipitating ice crystals — and a K-band radar, both also installed at Dome C.

We state that we aim at achieving consistency for cirrus clouds. Cirrus are generally optically thin clouds, ranging from subvisible up to 0.3-0.5 in optical depth; for larger values we are already talking about thick ice clouds. Our analysis shows the occurrence of two clearly distinguishable cloud scenarios in which the optical depth of ice clouds decreases, in fact halving from 0.17 to about 0.07/0.08. Thus, the cloud evolves from thin to very thin, covering two rather different values. While this is certainly not representative of a broader range of cases, it nevertheless spans a range of optical depths that well characterize this type of cloud, which permanently envelops the planet.

The scarcity of this type of measurement, due also to their complexity and operational difficulty, highlights the need to fund new campaigns aimed at confirming the results obtained and, if possible, expanding the range of case studies.

Regarding the reviewer’s concern about the impact of the aircraft contrail on the retrieval of the measurements, we tried to trace the flight path from the beginning of the acquisition, as shown in Fig. 1 (which has been added to the manuscript), and verified that the contrail did, in fact, intersect FINESSE’s field of view at 08:25 UTC — when the thicker cirrus clouds were passing — while in the second measurement sequence, where the cirrus was thinner, there was no overlap.

For this reason, we first decided to test the impact on the simulated FINESSE radiance when introducing a droxtal component for small crystal lengths below $50 \mu\text{m}$. We verified that a resonance occurs for $\text{OD} = 0.2$ and $D_{ei} = 20 \mu\text{m}$, where the differences between the radiances exceed the instrumental noise, as shown in Fig. 3.

So we regenerated the database of mean optical properties by adding, for smaller particles (below $50 \mu\text{m}$) a droxtal component to account for the fact that aircraft engines produce small droplets or microparticles that can act as condensation nuclei, promoting the formation of embryonic ice crystals. The most plausible particle shape in this case is therefore that of a hexagonal-faced spheroid, i.e., a droxtal. The new habit distribution is then represented by:

$$f_h(L) = \begin{cases} \text{BR} & L > 50 \mu\text{m} \\ \frac{1}{2}\text{DX} + \frac{1}{2}\text{PL} & \text{@ 8:25 UTC, PL @ 8:37 UTC } L \leq 50 \mu\text{m} \end{cases} \quad (1)$$

and similarly for the other $f_{2,3,4}$ as reported in the manuscript.

We then repeated the retrieval. The results obtained by assuming the two habit composition for smaller crystal lengths are shown in Fig. 2. For the first sequence when the INCAS contrails intersected FINESSE’s FOV we see that there is not much effect of introducing the droxtal component, however in the second sequence, when there was no contrail overlap, we obtain much better agreement with the INCAS retrieved D_{ei} with the droxtal component included. We updated Figs. 17 and 19 in the manuscript showing only the retrievals with droxtal+plate for smaller lengths for the first sequence and only plates for the second one. We also added the calculated standard deviation ($\approx 10 \mu\text{m}$) of the INCAS D_{ei} in Fig. 2 along with the corresponding error, showing that all retrieved diameters in when assuming HBRs are consistent with the internal variability of the INCAS data. As a consequence of the new retrieval we also updated Figs. 14 and 15 of the manuscript. For convenience and simplify, and to make the procedure more streamlined and faster, we use the sum of square root of nesr and calibration error in the VCM of measurements. This has been explained in the manuscript in section 3.1.3.

Minor concerns:

Fig 15 (question 1): why do the INCAS points change mean values and have different error bar ranges in the 4 panels? This is most easily seen looking at the bottom-most panel against any of the other three?

The mean value must change because it is calculated using different habits or mixtures of habits. Different habits have different geometric properties. As a result, for a given maximum crystal length, the corresponding values of volume and projected area differ. Consequently, the effective diameter, calculated as reported in Eq. (1), changes.

Regarding the associated error, the reviewer is correct, and we thank them for pointing this out, as it was indeed a mistake. The error associated with the mean was uniquely estimated in Section 3.4, and therefore it does not change

with the variation of the calculated diameter value. The error only affects the first value in the second panel (case $\frac{1}{2}\text{HBR} + \frac{1}{2}\text{SBR}$). We have corrected this accordingly.

Fig 15 (question 2): Why are the FINESSE error bars so much larger for cases 8-11 when HBR is assumed relative to when you assume $\frac{1}{2}\text{HBR} + \frac{1}{2}\text{SBR}$? This does not make sense to me.

This is not intuitive, but in the simulations $\frac{1}{2}\text{HBR} + \frac{1}{2}\text{SBR}$ the error bars are larger due to the calibration error. The error on the diameter is strongly correlated with the error on the slope of the spectrum in the atmospheric window. This is mainly controlled by the calibration error which has a more pronounced slope than the NESR in this region. For lower OD values this manifests as an enhanced error in the retrievals.

You are using a Bayesian retrieval framework, which is able to provide degrees of freedom for signal. For these very small optical depths, what is the DFS for the retrieved effective radius? Does it change when you assume different habits? Does it change when you only use the midinfrared vs when you use both the midinfrared and far infrared? (This would be a pretty convincing point to make, if the far infrared is indeed important).

The reviewer’s point is correct; therefore, we show the distribution of the information content (IC), the correlations between D_{ei} -OD, and the number of degrees of freedom for the different habits. We find that there are essentially no differences in IC and DOFs — up to 0.3/0.4 and 10^{-4} , respectively — and the differences in correlations are smaller than 0.25. The plots are shown in Fig. 4. This indicates that, a priori, there is no greater or lower sensitivity in distinguishing cloud parameters when using different habit distributions.

In contrast, when we consider whether or not to include the FIR component in the retrieval procedure, Fig. 5 shows that IC changes significantly, decreasing from 9 to 6; the correlation between D_{ei} and OD parameters increases notably to about 0.8/0.9, and the difference in DOFs in this case is on the order of 10^{-2} — that is, one hundred times larger than in the case of considering different habits.

We performed several retrieval tests simulating FINESSE observations, both for the values obtained from the analysis in this work and for other cases with larger optical depths and particle effective diameters. Specifically, we tested the capability of the retrieval algorithm to recover the “true” values with and without including the FIR spectral region. To this end, we generated synthetic observations using the FINESSE noise and two different atmospheric profiles, together with varying cloud parameters — optical depths ranging from 0.07 and 0.18 (representative of the values obtained from the FINESSE measurements) up to 0.8 and 1.5 — while fixing the D_{ei} value at $30\ \mu\text{m}$.

Our tests demonstrated that the retrieval algorithm is capable of reproducing the true state from synthetic observations whether or not we include the FIR if we fit the observations with habit distribution equal to those used to generate the “truth”, as shown in Fig. 6 and 7. However, when we invert the real measurement, we find that we can only match the values measured by INCAS when the FIR is included. This is likely due to the fact that the actual microphysics of the cloud is more complex than what our models can reproduce, and by adding spectral information in the FIR we can approach more closely the absolute minimum of the χ^2 distribution.

In conclusion, we can state that using a more comprehensive set of spectral information, including the FIR, allows for more stable solutions that better represent reality. However, a much more detailed investigation would be required to fully explore this, which is beyond the scope of the present work, although it remains an interesting topic for future research. We added the following sentence in the text in section 5: “We would like to emphasize that, based on the FINESSE measurements presented and analyzed, performing the inversion with limited information that excludes the FIR leads to an underestimation of the effective diameter values, but this result cannot, at this time, be generalized to every possible scenario. This clearly calls for a more in-depth investigation, which will be addressed in future work”. In addition, we have modified the abstract as follows: “Furthermore, we show that the radiance information contained within the far-infrared (wavenumbers $> 650\ \text{cm}^{-1}$) spectrum is critical to achieving this level of agreement with the in-situ aircraft observations. The results emphasize why it is vital to expand the current limited database of measurements encompassing the far-infrared spectrum, particularly in the presence of cirrus, to explore whether this finding holds over a wider range of conditions.”

Fig 17: caption errors: INCAS measurements are green dots, and only using the mid-IR is the violet diamonds

We thank the reviewer, we have remade the Fig. 17, now Fig. 19 in the manuscript.

You spend very little time talking about the thermodynamic retrieval and its accuracy. Personally, I do not feel it adds anything to the paper; indeed, it is more distracting. If you elect to keep that

portion in the paper for the next iteration, I would want to see much more discussion about it (recognizing that you don't have strong sources of truth). But, in particular, your statement on line 468 about "excellent agreement...inside the cirrus" needs to be tempered. Does the FINESSE retrieval actually have any information content (i.e., degrees of freedom for signal) at that altitude? I would be very surprised if it did. And if the DFS is very small, then the agreement in temperature within the cirrus is more of a happy coincidence (provided by the statistics of the prior dataset used in the constraint, which you did not discuss at all) than real skill by the retrieval.

The retrieval of atmospheric profiles is crucial to achieving consistency among all the available measurements. The reviewer is correct in noting that we lack sensitivity to cloud altitude; however, that is not the goal of our analysis. The reason for using the inversion technique is to "adjust", as also suggested by reviewer 2, the parameters around the minimum, since we cannot know, within the associated uncertainties, how much the measurements should vary in order to achieve full consistency among them. This is because local factors and variables come into play that are difficult to account for when setting up the simulation framework, even when attempting to incorporate all available information.

For this reason, when we say "very good agreement", we mean that there is consistency between the temperature measurements inside the cirrus and the scaled temperature profile above the last fitted level. We understand, however, that this phrase may be misunderstood, and we have therefore sought to soften and clarify it. We have rephrased it as follows in the text: "The retrieved temperature profile above the last fitted level is merely scaled but it turned out consistent with the temperatures measured by INCAS probe inside the cirrus, and the vertical profile provided by ERA5."

The a priori/initial guess of the atmospheric profiles are now reported in Table 2.

We have introduced a sensitivity study based on the vertical distribution of the averaging kernels (AK) for temperature and water vapor, as shown in Fig. 8, where it can be seen that the sensitivity for temperature and water vapor is mainly limited to about 1 km and 4 km above the surface, respectively. So, as the reviewer pointed out, there is not much sensitivity to the internal temperature of the cloud. The average degrees of freedom for signal (i.e., the trace of the AK) were found to be 3 and 7, respectively. For this reason, the vertical levels were limited to about 1 km for temperature (3 levels) and to about 9 levels for water vapor below 4 km as stated in the manuscript.

This choice allows us to adjust the profiles around a minimum that is assumed to be close to, but not exactly coincident with, the ERA5 profiles — as indeed turns out to be the case. In this context, there is strong motivation for this approach: ERA5 profiles are not radiosonde measurements but are derived from models and data assimilation. Nevertheless, even radiosondes may not be exactly co-located with the measurements and can be affected by small biases and local variability in the lower atmospheric layers.

For these reasons, we believe that presenting the retrieval of the atmospheric profiles is an integral part of the study and can help colleagues when considering how to approach similar studies in future, so should not be overlooked.

We have added a discussion on this aspect and included the corresponding figure in section 3.1.3 of the manuscript.

References

- Di Natale, G., Bianchini, G., Del Guasta, M., Ridolfi, M., Maestri, T., Cossich, W., Magurno, D., and Palchetti, L.: Characterization of the Far Infrared Properties and Radiative Forcing of Antarctic Ice and Water Clouds Exploiting the Spectrometer-LiDAR Synergy, *Remote Sensing*, 12, <https://doi.org/10.3390/rs12213574>, 2020.
- Maestri, T., Rizzi, R., Tosi, E., Veglio, P., Palchetti, L., Bianchini, G., Girolamo, P. D., Masiello, G., Serio, C., and Summa, D.: Analysis of cirrus cloud spectral signatures in the far infrared, *Journal of Geophysical Research*, 141, 49–64, 2014.

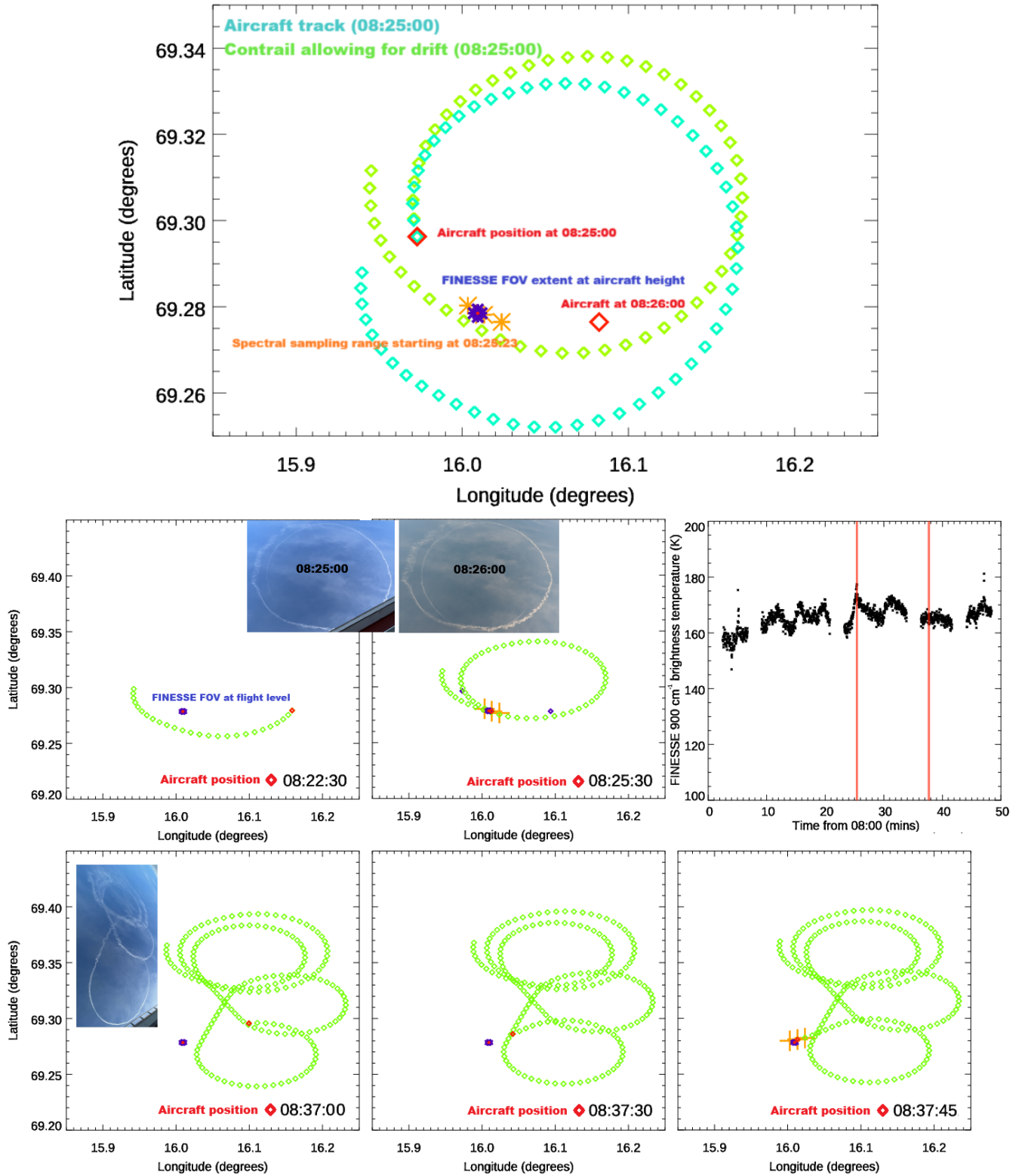


Figure 1: Upper panel: the turquoise diamonds map the aircraft track over the Alomar site up to 08:25:00 UTC, 30 seconds before the first overpass. We use photographs taken at this time and at 08:27:00, 30 seconds after the overpass to estimate the drift of the aircraft contrail, shown by the green diamonds. Extrapolating this to 08:26:30 UTC indicates that at the time of the overpass the contrail released earlier in the flight, to the South-West, is directly over Alomar at the time of the overpass, this time also coincides with a distinct rise in radiance at 900 cm⁻¹, observed by FINESSE. Middle and lower panels: reconstructed INCAS track with the corresponding photos of the contrail in the sky for the two sequences measure by FINESSE. In particular in the middle right panel is reported the increasing radiance at 900 cm⁻¹ detected by FINESSE at the first sequence, when the contrail intersects the FINESSE's FOV.

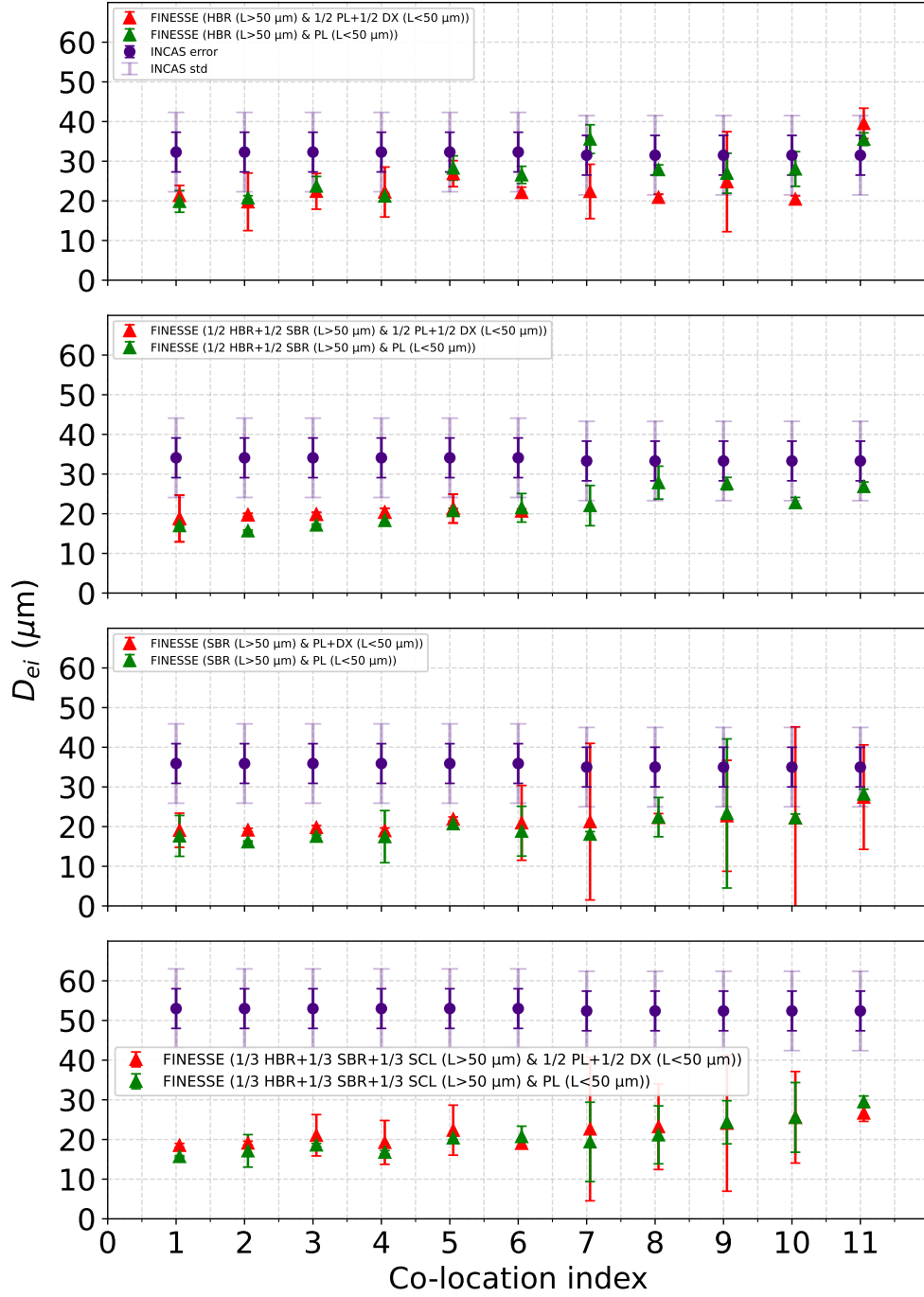


Figure 2: Top panel: comparison of the effective diameters calculated from INCAS KA data (violet dots) with the corresponding error and standard deviation between 08:19–09:00 UTC on 17 February 2023 and those retrieved from FINESSE (red triangles) between 08:25–08:38 UTC of the same day by assuming the habit distribution f_{1h} in Eq. 1. Second panel: same comparison but assuming f_{3h} . Third panel: same comparison but assuming f_{2h} . Bottom panel: same comparison but assuming f_{4h} . For brevity, the caption labels do not explicitly state the fraction of plates considered for crystal lengths below $50 \mu\text{m}$.

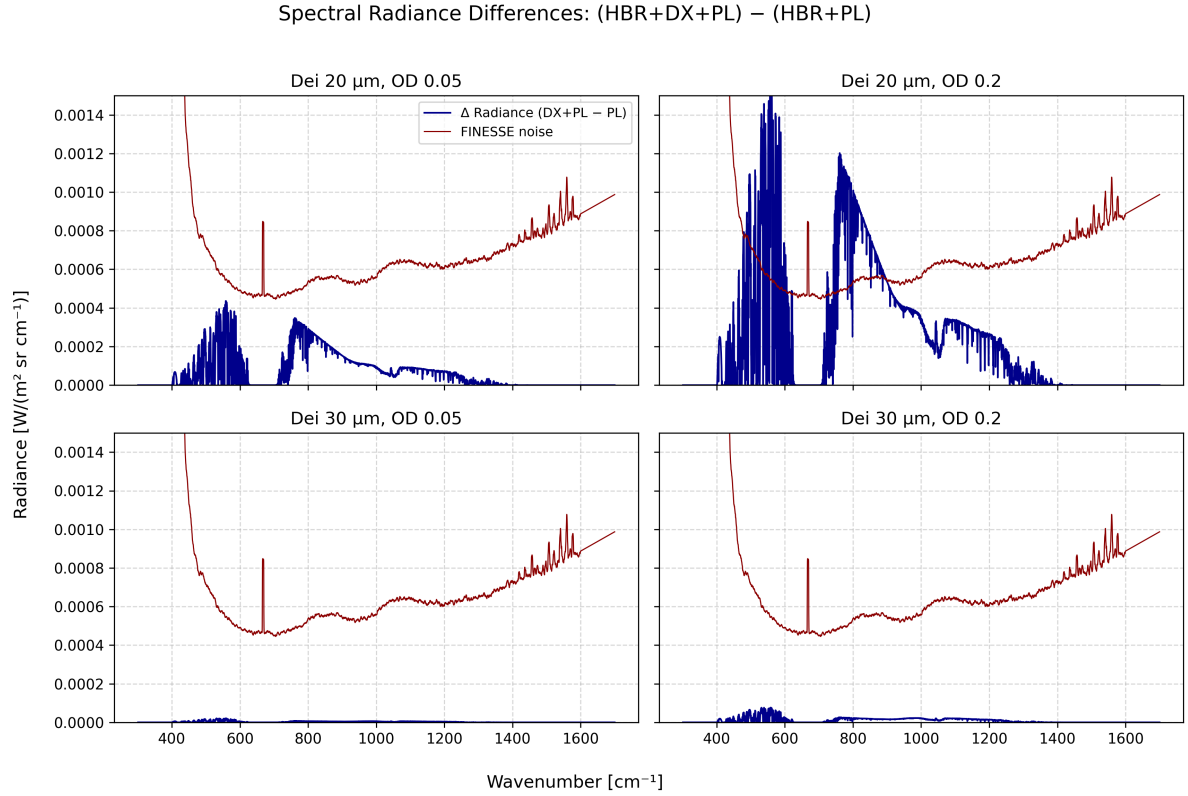


Figure 3: Differences between radiances simulated by considering and non considering the droxtal contribution introduced in Eq. (1) in comparison with the FINESSE noise. Top left panel: considering $OD = 0.05$ and $D_{ei} = 20 \mu\text{m}$. Right top: considering $OD = 0.2$ and $D_{ei} = 20 \mu\text{m}$. Left bottom panel: $OD = 0.05$ and $D_{ei} = 30 \mu\text{m}$. Right bottom panel: considering $OD = 0.2$ and $D_{ei} = 30 \mu\text{m}$.

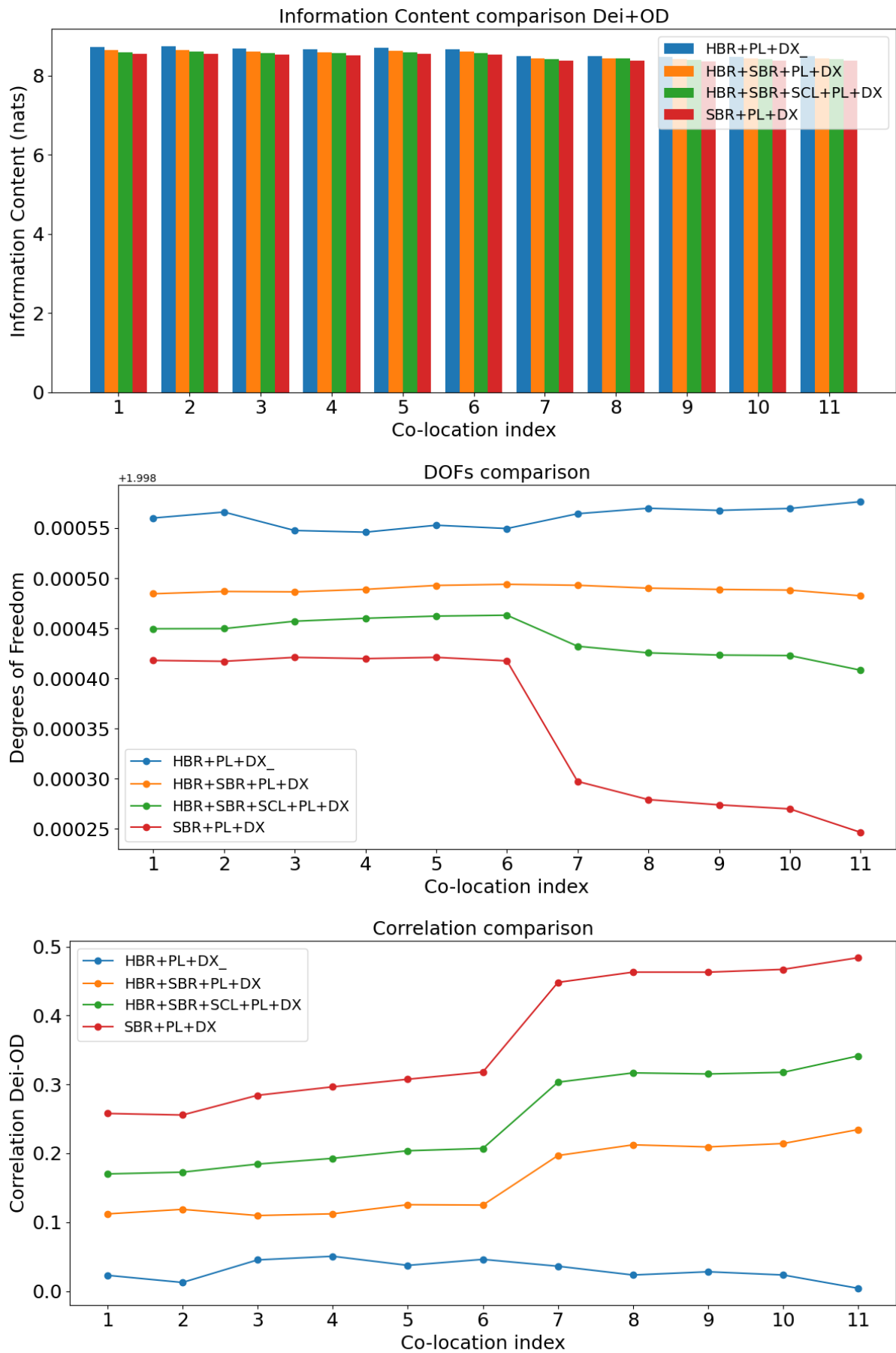


Figure 4: Upper panel: Information content of retrieval for the 4 habit distributions, corresponding to the different colors, considered in the analysis. Middle panel: degrees of freedom. Bottom panel: correlations OD- D_{ei} .

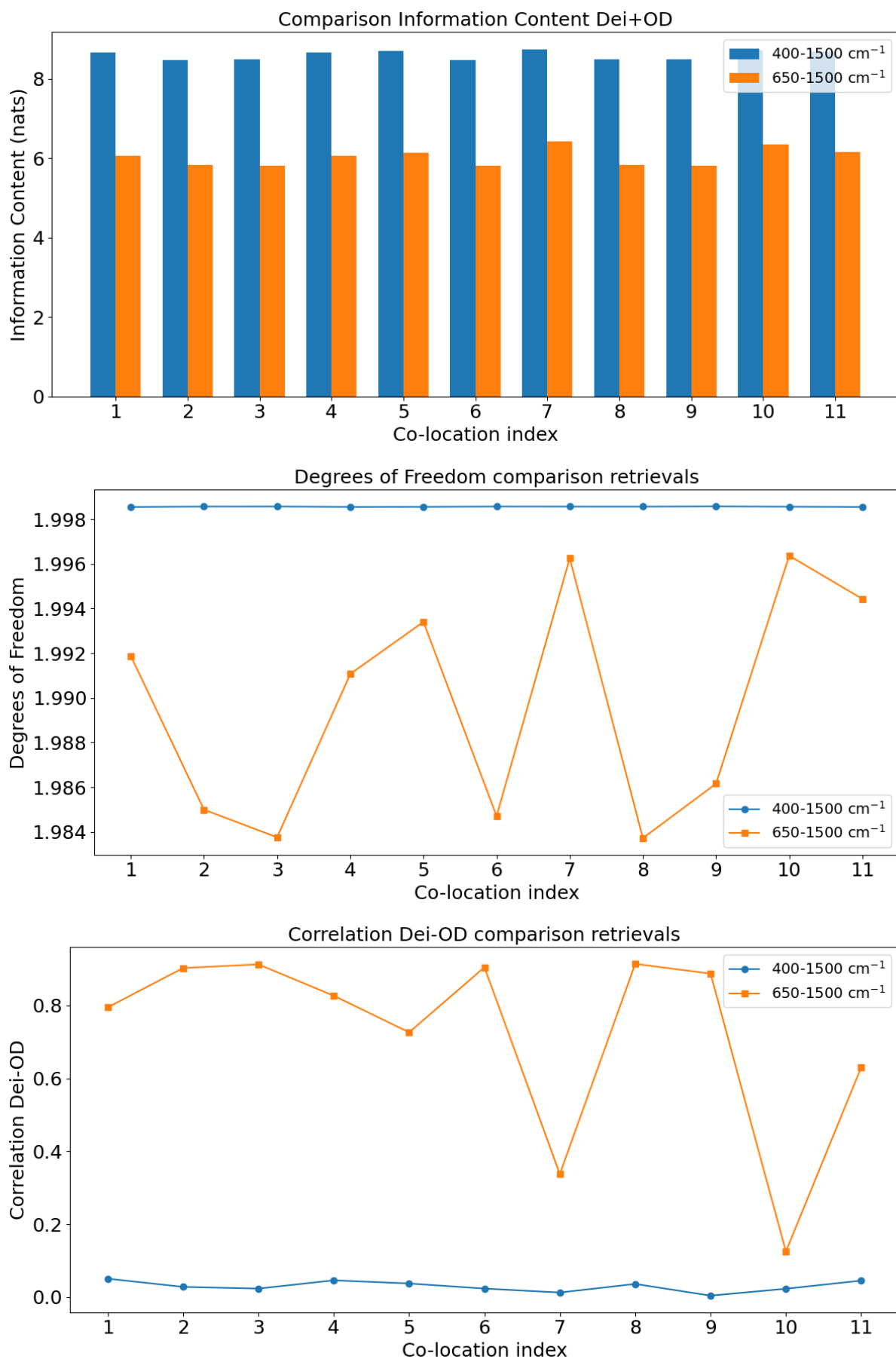


Figure 5: Upper panel: Information content of retrieval for the case FIR was considered (blue) and neglected (orange). Middle panel: degrees of freedom. Bottom pane: correlations OD- D_{ei} .

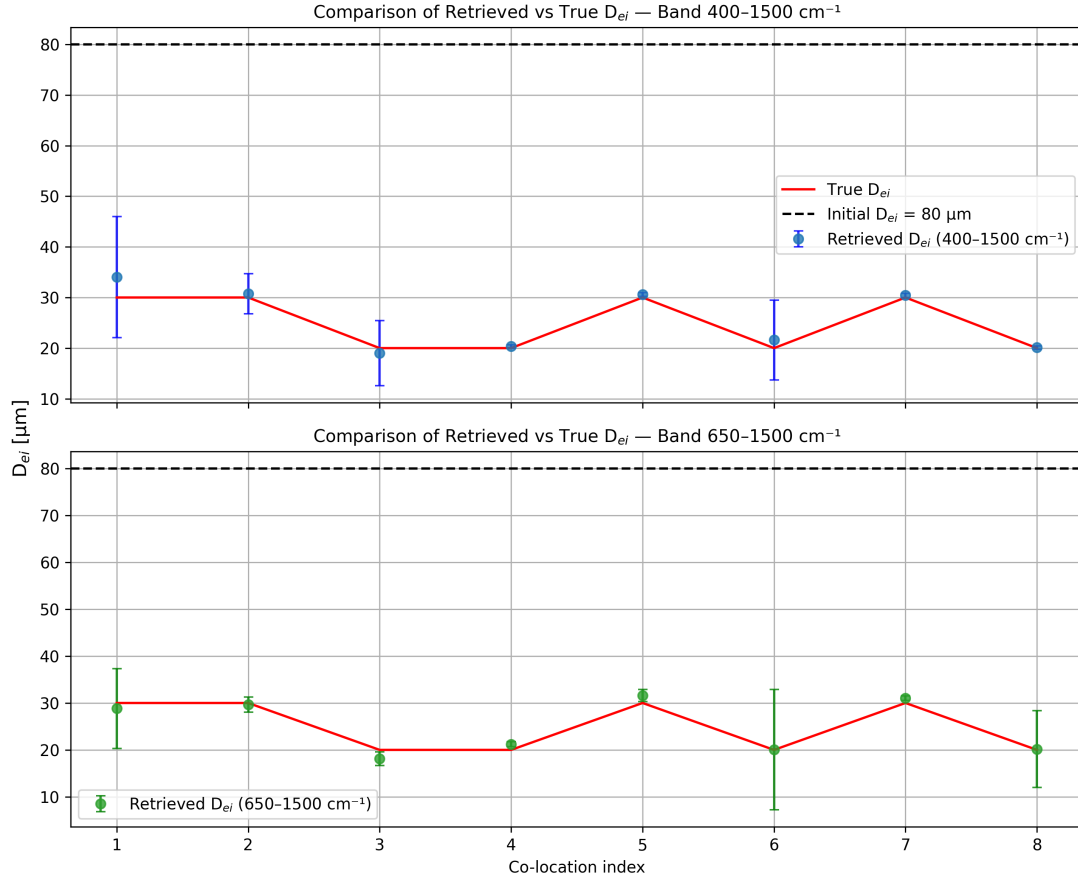


Figure 6: Top panel: comparison of the retrieved D_{ei} by including the FIR (blue dots) with respect to the "truth" (red curve). Bottom panel: same but neglecting the FIR (green dots). In both cases the initial guess and a priori is 80 μm as done with the analysis.

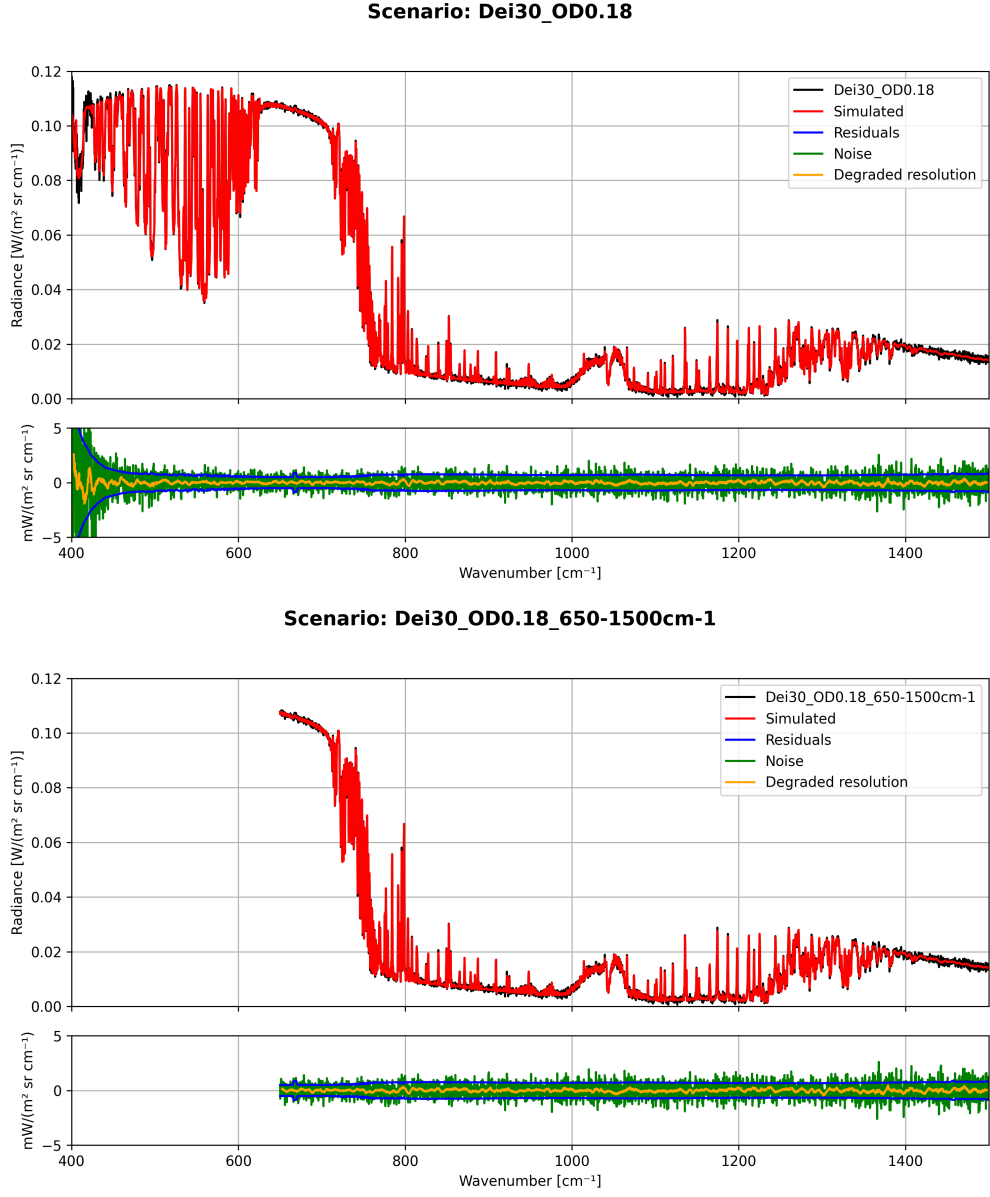


Figure 7: Top panel: fit and residuals by using the FIR for the case $OD = 0.18$ and $D_{ei} = 30 \mu\text{m}$. Bottom panel: same by neglecting the FIR.

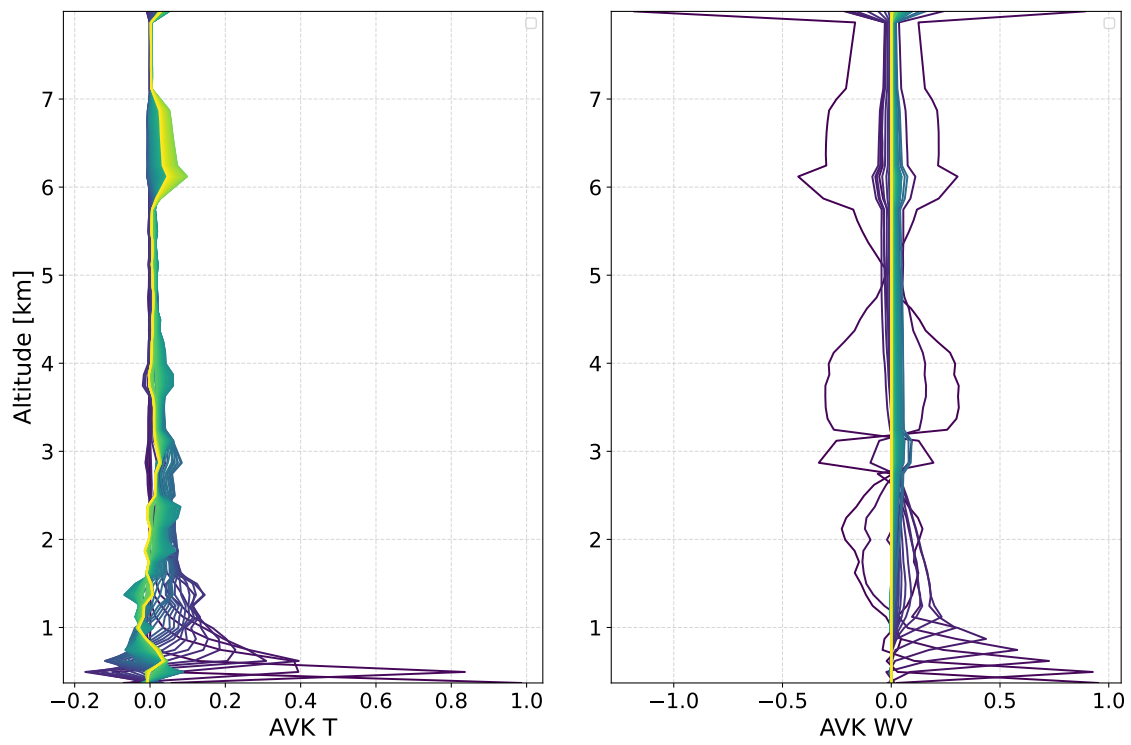


Figure 8: Left panel: vertical distribution of averaging kernels of temperature. Right panel: same for water vapour.