Response to Reviewer #1

Review for “The JUICE 2024 close flyby of the Moon: Thermal assessment from MAJIS”:

Overall, the manuscript is well written considering all aspects of the study. While comparing the results, the authors have also highlighted the limitations of each algorithm. However, I have a few concerns that need to be addressed before its acceptance for publication in ANGEO.

We thank the Reviewer for the careful and constructive review of our manuscript and for the detailed technical suggestions. We appreciate the Reviewer’s emphasis on improving methodological clarity and on strengthening the quantitative assessment of the retrieval robustness. Following these recommendations, we have revised the manuscript to (i) better document key elements of the Bayesian inversion framework (including the role of the a priori and the definition of covariance matrices), (ii) clarify the treatment of topography and the limits of the ellipsoid-based geometry, (iii) justify the comparison between MAJIS-derived physical temperatures and Diviner brightness-temperature products, and (iv) improve the presentation of several figures for enhanced readability. Overall, we believe that these revisions have significantly strengthened the manuscript and improved its transparency and reproducibility.

Technical comments:

1) The retrievals from the Bayesian approach are reported to be dependent on the a priori. It would be more informative if the authors also reported the degree of dependence. Sensitivity analysis of the retrievals on a priori will add more value to the retrievals. Such a kind of diagnostics is available from the algorithm itself and straightaway tell how much observations (and a priori) are contributing to the retrievals.

Response: We thank the Reviewer for this valuable suggestion. We agree that quantifying the dependence of the Bayesian retrieval on the a priori improves the interpretation of the results and provides an objective measure of retrieval robustness. To address this point, we performed a sensitivity analysis by repeating the inversion using different emissivity priors (ε_0 = 0.70, 0.80, 0.90, and 0.95; Table A1). We quantify the degree of dependence by evaluating the spread of the retrieved mean temperatures across these emissivity priors for each MAJIS observation. For the 4.5–5.56 µm retrieval window, the mean temperature varies by $\Delta T\_mean \approx$ 8.6 K (C1), 2.2 K (C2), 9.2 K (C3), and 12.9 K (C4) across the tested ε_0 values. These results confirm that the sensitivity to the emissivity prior is generally modest for intermediate illumination conditions (C2), while it becomes more significant for extreme cases, particularly for high-temperature conditions where the radiance–temperature relationship is highly nonlinear. We have added a dedicated paragraph in Sect. 2.1, immediately after the paragraph introducing the dependence on the emissivity prior, to explicitly report this quantitative sensitivity.

2) Bayesian inversion is all about giving proper weights to the observations and *a priori*. There is no mention of the covariance matrices used in the retrieval. Although the authors have mentioned the impact of the covariance matrix at many places in the text. Please mention how the covariance matrices (both observation and *a priori*) are prepared. Whether full covariance matrices are used for both observations and *a priori*? If only diagonal matrices are used, then what is done to compensate for the missing covariances? Hyperspectral observations are usually correlated. In practice, only the diagonal observation error

covariance matrix is used, and errors are inflated. For *a priori* covariance matrix, the emissivity may be assumed to be uncorrelated, but temperature and emissivity are highly correlated, so their correlations must be included. Ignoring covariances will result in suboptimal retrievals. Physically inconsistent results mentioned on page no. 13, line no. 272 can be better explained with the help of this information.

Response: We thank the Reviewer for this important and technically well-founded comment. We agree that the construction of the observation error covariance matrix (S_e) and the a priori covariance matrix (S_a) is central to the Bayesian (optimal estimation) inversion and should be described explicitly. We have therefore expanded Sect. 2.1 to detail how both covariance matrices are defined in our implementation.

For the observation error covariance matrix (S_e), we assume uncorrelated errors between spectral channels and adopt a diagonal matrix whose elements are given by the in-flight Noise Equivalent Spectral Radiance (NESR) of MAJIS, evaluated at each wavelength. This choice is motivated by the lack of a validated spectral correlation model for MAJIS radiance noise at the time of writing, and it follows common practice in hyperspectral retrievals when channel-to-channel correlations are not available. We note that the diagonal S_e provides conservative uncertainties because the NESR already includes the dominant instrumental noise sources and increases toward longer wavelengths due to thermal background.

For the a priori covariance matrix (S_a), we apply (i) a temperature prior variance corresponding to ±30 K (soft constraint), and (ii) a wavelength-dependent emissivity prior with physically motivated bounds ($\varepsilon \leq 1$) and smoothness enforced through spectral correlations between neighbouring channels. In practice, S_a includes off-diagonal terms for emissivity, modelled with a Gaussian correlation function whose correlation length is comparable to the spectral sampling/resolution. This regularizes the emissivity solution and mitigates unrealistic channel-to-channel oscillations. The coupling between temperature and emissivity is naturally represented through the Jacobian and the posterior covariance, which captures their correlation in the retrieved state.

We clarify in the revised manuscript that the remaining physically inconsistent emissivity behaviour near the crossover region is primarily linked to the intrinsic temperature–emissivity degeneracy and to local observational conditions (e.g. low thermal signal and/or high incidence angle), rather than to a failure of the Bayesian formulation. We have added this information in Sect. 2.1 and referenced Rodgers (2000) for the formalism.

3) It is not clear whether topographic corrections are implemented in the Bayesian approach for retrieval. The sentence on page no. 6, line 184, “Topographic effects may be mitigated…” is creating doubt. It’s not certain. Further, the sentence on page no. 7., line no. 185, “However, while such corrections are critical… a few hundred meters.” is creating more suspicion. On page no. 13, line no. 298, “The Moon’s global shape can be approximated by a smooth ellipsoid… rather than geometric scattering.” emphasize that topographic effects are not important at MAJIS observations’ spatial resolution, but on page no. 16, line no. 370, “This discrepancy likely arises from unresolved topographic effects… can dominate the measured radiance and produce local deviations from the expected thermal behavior.” and Page no. 23, line no. 491, “At higher… in temperature retrievals”, the authors attribute the retrieval errors to local topography.

Response: We thank the Reviewer for highlighting this ambiguity. We agree that the manuscript could be interpreted as contradictory regarding the role of topography. We clarify here that no explicit topographic correction is implemented in the Bayesian retrieval (Tosi et al., 2014, Sect. 2.1

of the paper) and in the empirical thermal correction (Li and Milliken, 2016, Sect. 2.2 of the paper), i.e. we do not use a DEM to compute local slopes, local incidence/emission angles, or sub-pixel shadowing corrections. The MAJIS data are georeferenced using SPICE-based geometry and an ellipsoidal lunar reference surface. However, a LOLA-derived digital roughness map is used in Sect. 2.3 to properly inform the thermal model, following the framework of Wohlfarth et al. (2023).

The statement that an ellipsoidal approximation is acceptable refers to the global mapping scale and to the fact that the MAJIS pixel size during the flyby is of the order of a few hundred metres, so that large-scale trends in temperature and emissivity can be robustly interpreted without DEM-based corrections. However, local discrepancies between MAJIS retrievals and Diviner-derived products can still arise from unresolved sub-pixel topographic effects (e.g. mixed slopes and shadowing within a pixel), which are not represented in the ellipsoid-based geometry and can locally bias the measured radiance. We have revised the relevant sentences in Sect. 2.1 to explicitly state that topographic corrections are not applied, and to distinguish between large-scale mapping robustness and local residual effects.

4) The retrievals from the three algorithms mentioned in the study provide the physical lunar surface temperature, but LRO Diviner provides bolometric temperature. Authors have compared the surface temperature retrievals against the Diviner peak brightness temperature (determined by parabolic fitting of channels 3, 4, and 5). It is still brightness temperature not physical surface temperature. Authors should add a justification for making such a comparison and should also mention this while discussing discrepencies in the results.

Response: We thank the Reviewer for this important clarification. We agree that MAJIS retrievals aim at estimating a physical surface temperature (together with wavelength-dependent emissivity), whereas Diviner products provide brightness temperatures (per channel) and a peak brightness temperature derived from a parabolic fit of channels 3–5 around the Christiansen Feature (CF). These quantities are not strictly equivalent and should not be interpreted as identical physical temperatures.

Our comparison is therefore intended as a consistency check using Diviner peak brightness temperature as a practical proxy for the surface thermal state at the time of observation, rather than a one-to-one validation of absolute thermodynamic temperature. The Diviner peak brightness temperature is particularly suitable for this purpose because it is derived near the CF, where the effective emissivity is close to unity and the dependence on emissivity is reduced compared to single-channel brightness temperatures. We have revised the manuscript to explicitly state this limitation and to clarify that residual discrepancies between MAJIS and Diviner can partly arise from the intrinsic difference between physical temperature estimates and brightness-temperature products, in addition to spatial resolution and unresolved topographic effects.

5) In the “Roughness-informed thermal model” used in this study, it is not clear from the text whether whole model (Reflectance + Emission) given by Wohlfarth et al. (2023) is implemented or only the emission model is used. Wohlfarth et al. (2023) used Hapke model for reflectance which involves various parameters. Moreover, single scattering albedo is used as a free parameter in reflectance model. If whole model is used then the results are self-explanatory.

Response: We thank the Reviewer for this useful remark. In the original version of the manuscript, we implemented only the thermal formulation presented in Wohlfarth et al. (2023). To improve the robustness of the analysis, we have revised this implementation by developing and applying a combined reflectance and thermal emission model following the full framework described in

Wohlfarth et al. (2023). For the reflected component of the signal, we adopt the definitions provided in Hapke (2005). In addition, for consistency with previous studies (e.g., Bandfield et al., 2015; Müller et al., 2021; Wohlfarth et al., 2023), we derived an average estimate of the cm–mm scale roughness for selected areas within C4, which is characterized by high signal-to-noise ratio (SNR) and stronger thermal emission compared to C1–C3. This approach provides a more representative description of the actual sub-pixel roughness than the meter-scale roughness derived from LRO/LOLA data. In light of these changes, Section 2.3 of the manuscript has been updated accordingly, including revised figures and tables (also in the Appendix), and the analysis has been repeated using this new approach. The Hapke parameters adopted for the implementation of the full model are taken from Li and Li (2011).

6) Figure A1. in Appendix A, (a) where is the noise plot? The plot in figure (b) is skeptical, how can the posterior uncentainty be so low ($\sigma \approx 0.01$ K)? It is given as (see Rodgers 2000):

$$\boldsymbol{S} = (\boldsymbol{S}_a^{-1} + \boldsymbol{K}^T \boldsymbol{S}_\varepsilon^{-1} \boldsymbol{K})^{-1}$$

Is it verified? It can only happen in the case of almost perfect observations. It is difficult to say anything without having the knowledge about the observation and *a priori* errors used in the retrieval.

Response: We thank the Reviewer for this useful technical remark. In Fig. A1(a), the noise curve (NESR) is plotted in light blue. However, for the selected best-case pixel its amplitude is much smaller than the measured/modelled radiance in the 4.5–5.5 µm interval, and therefore it appears nearly overlapped by the other curves at this plotting scale.

Regarding the very low posterior uncertainty shown in Fig. A1(b) ($\sigma \approx 0.01$ K), we agree that such a narrow posterior can only occur in particularly favourable conditions. In the original version, Fig. A1 was intentionally constructed using a "best-case" pixel from MAJIS observation C3, selected to provide a clear illustration of the Bayesian retrieval behaviour under high signal-to-noise conditions. In this regime, the posterior covariance matrix can yield very small formal uncertainties when the radiance sensitivity to temperature is strong and the observational noise is low.

To address the Reviewer's concern and improve transparency, we have now revised Appendix A by updating Fig. A1 to include both (i) a best-case example and (ii) a representative suboptimal pixel from the same MAJIS cube, characterized by a significantly larger formal uncertainty. The revised figure therefore illustrates how the posterior temperature distribution broadens under less favourable observational conditions, as expected from the Rodgers (2000) formulation.

Minor comments:

1) Page no. 5, line no. 138, What does "Kirchhoff-derived estimate" mean? Kirchoff's law is $r = 1 - \varepsilon$. Both r and ε are unknown. Is estimate of r available from any other source?

Response: We thank the Reviewer for pointing out this ambiguity. In our retrieval, "Kirchhoff-derived estimate" does not refer to an external measurement of reflectance. Rather, it refers to an internal consistency check based on Kirchhoff's law applied in a spectral, directional sense under the standard assumptions of thermal equilibrium and an opaque surface (negligible transmittance). In this framework, spectral directional emissivity equals the corresponding spectral directional absorptivity; for an opaque surface this can be written as

$$\varepsilon_d(\lambda) = 1 - r_{hd}(\lambda),$$

where r_{hd} is the hemispherical–directional spectral reflectance (consistent with the directional emissivity definition). Practically, we compute a Kirchhoff-consistent reflectance proxy $r_{hd}(\lambda) = 1 - \varepsilon_d(\lambda)$ from the retrieved emissivity and use it only as a diagnostic to flag spectral segments near the crossover region where the inversion becomes ill-conditioned and may produce non-physical emissivity behaviour. We have revised the manuscript to clarify the emissivity/reflectance definitions and to avoid implying that reflectance is independently known or that a bolometric Kirchhoff relation is used.

2) Figure 13(b), there are three plots in this figure, all shown in shades of gray. It is difficult to identify which is what. Authors should consider revising this figure for better presentation.

Response: We thank the Reviewer for this suggestion. We agree that the three MAJIS-derived datasets in Fig. 13 were difficult to distinguish in grayscale. We have revised Fig. 13 to improve readability by using colors other than grey. The updated figure now allows a better identification of MAJIS-derived temperature values trends for each of the three approaches.

3) Figure 14(d), I think line plots showing the mean along with error bars would be better. It requires binning in latitude and calculate the values (mean and standard deviation, or median and mean absolute deviation) in each bin.

Response: We thank the Reviewer for this helpful recommendation. We have revised Fig. 14(d) accordingly by binning the data in longitude and plotting the binned central tendency together with uncertainty bars. Specifically, for each latitude bin we compute the mean value and its dispersion (standard deviation), which is shown as error bars. This revised representation improves clarity and makes the latitude trends easier to interpret compared to the original scatter plot.

4) I think it should be “Kelvin” not “kelvin”, throughout the text.

Response: We thank the Reviewer for the remark. According to SI conventions, the unit name is written in lowercase (“kelvin”), while the unit symbol is uppercase (“K”). We have ensured consistency throughout the manuscript, using “K” for temperature values and “kelvin” for text/captions.

Response to Reviewer #2

Review for “The JUICE 2024 close flyby of the Moon: Thermal assessment from MAJIS”:

Reviewer comment: Overall evaluation.
The authors present an analysis of a novel dataset of mid-infrared measurements of the lunar surface. These data were acquired by the MAJIS instrument during a gravity assist maneuver of the JUICE spacecraft in August 2024 and represent a rare opportunity to observe the Moon in the mid-infrared wavelength range. The paper integrates recent developments in thermal modeling and lunar science and serves as a preparatory step for the upcoming MAJIS observations at Europa, Ganymede, and Callisto.
While the dataset is valuable and provides a challenging test case for thermal models, there are several significant issues regarding the line of argumentation, technical implementation, and scientific evaluation. In its current form, both the study and the manuscript fall short of the standards expected for publication. Substantial revision and additional work will be required to address these issues before the manuscript is ready for publication.

Response:
We thank the Reviewer for the thorough and constructive review. We appreciate the recognition of the novelty and community value of the MAJIS lunar mid-infrared dataset, as well as the careful assessment of the strengths and shortcomings of our initial submission. We have addressed the Reviewer’s main concerns by revising the scope and terminology of the manuscript, improving the clarity and internal coherence of the line of argumentation, expanding and sharpening the methodological descriptions where needed, and substantially rewriting the conclusions to ensure that all statements are supported by the analyses presented. We have also incorporated additional relevant literature in the 3–5 μm domain and revised the emissivity discussion to clarify definitions, limitations, and cross-instrument comparisons.

Reviewer comment: Revise the scope of the work.
Several claims in the abstract regarding validation and preparation are not supported by the analysis and are not consistently addressed in the conclusion. I recommend revising the stated scope and improving the overall coherence of the manuscript to ensure alignment between abstract, analysis, and conclusions. Please consider the following points:
Validation of MAJIS capabilities: The abstract states that the study aims to validate the capabilities of MAJIS (line 15), and the term “validation” is used repeatedly throughout the manuscript. However, it is unclear what specific type of validation is actually performed. The paper does not explicitly address, for example, absolute calibration, pointing performance, thermal stability, or instrument-level performance metrics. Instrument calibration and validation appear to be primarily covered by Langevin et al. (this issue). In contrast, the present manuscript does not read as a systematic validation of MAJIS performance, more like a plausibility check. Please revise.

Response: We thank the Reviewer for this important and constructive remark. We agree that the term “validation” was used too broadly in the previous version of the manuscript and may have suggested an instrument-level performance assessment (e.g., radiometric calibration accuracy, geometric calibration, and thermal stability), which is indeed addressed in separate papers of the same special issue.
The present work does not constitute a formal instrument validation. Instead, it investigates the internal consistency and physical plausibility of thermal retrieval approaches applied to MAJIS mid-infrared lunar observations. Our objective is to assess whether physically motivated thermal modelling frameworks can coherently reproduce the observed radiances and yield temperature and emissivity estimates consistent with established lunar properties.

Following the Reviewer's suggestion, we have revised the abstract, introduction, and conclusion to clarify this point. The terminology has been adjusted throughout the manuscript, replacing "validation" with expressions such as "assess the instrumental performance", "assessing the physical consistency and methodological robustness", and "methodological and scientific consistency assessment". We now explicitly state that this study represents an exploratory analysis of a new dataset and a methodological preparation for future applications, rather than a systematic validation of MAJIS instrument performance.
We believe this revision better reflects both the scope and the scientific contribution of the work.

Reviewer comment:
Application of methods to JUICE's main mission: At several points (e.g. line 93–95), the study claims to prepare for the main mission. "can help validate techniques for interpreting thermal data from Ganymede, Callisto and Europa, where similar sunlit observations are planned. This is a relevant and interesting activity. However, I do not explicitly see how the results from this paper will be transferred to the Jovian moons.
Are similar thermal models/thermal retrievals required? Can they be used right away, or must they be adapted? If so, how?

Response: We thank the Reviewer for highlighting the need to clarify how the present analysis relates to the upcoming MAJIS observations of the Galilean moons.
Our intention was not to imply that the lunar thermal results can be directly extrapolated to Ganymede, Callisto, or Europa in a quantitative sense. Rather, the lunar dataset provides a controlled and well-characterised test case to evaluate the behaviour of thermal retrieval approaches in the spectral region beyond 5 μm.
In particular, the analysis allows us to:

1. Assess the numerical stability of the temperature–emissivity separation in a regime where the thermal contribution is moderate but non-negligible;
2. Evaluate the sensitivity of the retrieval to assumptions on emissivity smoothness and prior constraints;
3. Quantify degeneracies inherent to the spectral crossover region.

This regime is directly relevant for Callisto, where the subsolar temperature (~165 K) implies that thermal emission becomes detectable toward the long-wavelength edge of the MAJIS range (5.0–5.56 μm). For Ganymede, the thermal contribution is weaker but still potentially measurable under favourable illumination. Europa, being colder, represents a more reflection-dominated scenario, which can be treated within the same inversion framework by appropriately adjusting prior constraints and noise assumptions.
The mathematical structure of the Bayesian inversion and the thermophysical modelling framework are body-independent. However, their application to the Jovian moons will require adaptation of physical parameters, including albedo, emissivity spectra, thermal inertia, and surface roughness characteristics specific to icy regoliths.
We have revised the manuscript to clarify this distinction and now describe the lunar analysis as a methodological test case for the interpretation of future MAJIS observations, rather than a direct validation of performance at the Jovian system.

Reviewer comment:
Comparison of different thermal models/retrieval approaches: Using three different models/approaches makes the study more robust and also provides interesting insights on how the models compare. The investigations have close ties to lunar hydration analysis, as the thermal correction approaches of Li and Milliken 2017 and Wohlfarth et al. 2023 have been primarily

developed for this purpose. If not already covered in the companion paper Langevin et al. 2026 (this issue), the manuscript could become more interesting to a wider audience by discussing the implications of the three different thermal approaches/methods for lunar hydration analysis (See also Schörghofer et al 2021).

Response:
We thank the Reviewer for this constructive suggestion. We agree that the empirical correction approach (Li and Milliken, 2016) and the roughness-informed framework (Wohlfarth et al., 2023) have strong heritage in lunar hydration studies, where an accurate thermal correction is essential to interpret the 3-µm band. In the ANGEO special issue, the MAJIS companion papers do not present a full hydration analysis, but they highlight why this connection matters: Langevin et al. report a possible weak OH absorption near ~2.9 µm at mid-to-high incidence and emphasize that band-strength estimates in this wavelength region are highly sensitive to the model-dependent subtraction of the thermal contribution because the feature lies close to the filter boundary and near the crossover between reflected and emitted regimes.
Zambon et al. similarly provide a brief overview of the 2.6–5.4 µm range after thermal removal and caution that the 2.85–2.95 µm interval is affected by an LVF-interface gap, complicating a robust interpretation of the 3-µm feature.
Accordingly, in the revised manuscript we add a short discussion clarifying how differences among the three thermal approaches tested here may propagate into reflectance products around 3 µm (e.g., residual thermal contamination and geometry-dependent biases), while keeping the scope focused on thermal retrieval performance rather than a dedicated hydration investigation.

Reviewer comment:
Reframe the scope of the work (recommendation): In my view, the main contribution of this study lies in the exploration of a novel dataset in particularly challenging wavelength region, as well as in the development and testing of thermal modeling and retrieval approaches in preparation for future studies of the Jovian moons. In my view, the analysis therefore appears to serve as a scientific plausibility check, based on comparisons with well-established lunar properties, rather than as a formal instrument validation.

I therefore recommend revising the stated scope of the work and adjusting the terminology accordingly, for example by:

- Framing the study as an exploratory investigation of the challenging lunar mid-infrared wavelength region using a new dataset.
- Describing the analysis as a scientific plausibility check for the MAJIS instrument (useful for assessing instrument performance), rather than a formal validation.
- Emphasizing the role of this work in preparing thermal modeling and retrieval methods for application to the Jovian moons.

Response:
We thank the Reviewer for this thoughtful recommendation and fully agree with the proposed reframing. In the revised manuscript we have clarified that the primary contribution of this study is (i) the exploratory investigation of a new MAJIS lunar dataset acquired in a particularly challenging spectral region (4.5–5.5 µm), and (ii) the development and intercomparison of three thermal modelling/retrieval strategies as a methodological test case relevant to future MAJIS observations of airless bodies, including the Galilean moons.
Accordingly, we have revised the stated scope and terminology throughout the manuscript. In particular, we removed or rephrased statements implying a formal instrument-level validation and replaced the term “validation” with more accurate expressions such as “physical consistency” or “independent consistency assessment”. We now explicitly state that instrument calibration and

performance validation (e.g., radiometric calibration, geometric calibration, stability) are addressed in dedicated companion papers, whereas our study focuses on the physical consistency of temperature/emissivity retrievals when applied to MAJIS mid-infrared lunar observations. These changes are reflected in the revised abstract, in the closing paragraph of the introduction, and in the revised conclusion.
We believe this reframing improves internal coherence between abstract, introduction, and conclusions, and more accurately represents the scientific value of the work.

Reviewer comment: Revise the Bayesian approach.
The authors state Kirchhoff’s law as simply “emissivity = 1 – reflectance,” which is not accurate. Kirchhoff’s law is derived from energy conservation and properly states:

$$e_{d} = 1 - r_{hd},$$

where is the directional spectral emissivity and is the hemispherical–directional spectral reflectance. Using an incorrect formulation can introduce significant errors in the crossover region, as demonstrated by Myhrvold et al. (2018) in response to thermal modeling in the NeoWISE mission, which erroneously simplified Kirchhoff’s law to .
It is important to note that Kirchhoff’s law applies to directional spectral emissivity, not bolometric emissivity. This distinction is particularly critical when making graybody assumptions. The authors should ensure that the correct definition is applied, energy conservation is properly satisfied, and the models are re-run, if any changes had to be made.

Response: We thank the Reviewer for this important clarification. We agree that the wording in the previous version (“$\varepsilon = 1 - r$”) was overly simplified and could be misleading. In the Bayesian retrieval, we work with spectral directional emissivity (i.e., wavelength-dependent emissivity relevant to the observation geometry), not with bolometric emissivity. The Bayesian algorithm was adapted from previous works by Tosi et al. (e.g., 2014, 2018, 2019), where surface temperature and spectral emissivity were retrieved pixel by pixel based on the geometric information and wavelength-dependent formulations.
We have revised the text, including the caption of Figure 15, to explicitly state these definitions and avoid the ambiguous shorthand. No re-running of the models was required.

Reviewer comment: Empirical thermal correction:
No comments on this section.

Reviewer comment: Revise the thermal roughness model approach.
Quantifying surface roughness on planetary surfaces is inherently challenging due to its scale dependence and statistical complexity. For example, Figure 1(a) in Rubanenko et al. (2020) shows lunar surface roughness at multiple scales, indicating that the RMS slope generally increases as the spatial scale decreases. As discussed by Bandfield et al. (2015, Figure 11), the spatial scales most relevant for thermal modeling are those at which thermal isolation becomes significant, typically in the millimeter-to-centimeter range. At these scales, common lunar roughness estimates are approximately 20° near nadir (Bandfield et al. 2015; Wohlfarth et al. 2023) and 30–35° in the TIR (Bandfield et al. 2015; Rubanenko et al. 2020; Müller et al. 2021; Wohlfarth et al. 2023).
In contrast, the authors derive surface roughness from LOLA measurements, which sample much larger spatial scales (meters) and therefore likely underestimate roughness at the scales controlling thermal emission. Additionally, the manuscript does not provide specific roughness estimates in terms of RMS slope or Hapke θ.

To improve the study, the authors should:

1. Include relevant literature on lunar surface roughness at thermal-relevant scales.
2. Revise the roughness implementation to reflect the appropriate spatial scales.
3. Investigate the effects of surface roughness on their results and discuss the implications accordingly.

Response: We thank the Reviewer for the insightful comment, which helped us to strengthen the robustness of our analysis. First of all, we added and discussed the relevant literature on lunar roughness at thermally relevant scales (Bandfield et al., 2015; Rubanenko et al., 2020; Müller et al., 2021; Wohlfarth et al., 2023). As suggested, we implemented a refined thermophysical inversion to better constrain the cm-mm scale roughness of the analyzed regions. A full pixel-by-pixel inversion based on synthetic fractal surfaces would be computationally demanding and not warranted by the information content of all hyperspectral cubes. Therefore, we retrieved representative sub-pixel roughness values only for selected ROIs within cube C4, which provides the highest thermal signal and most favorable illumination geometry. These terrain-representative values were then used to generate the synthetic fractal surfaces required for the temperature and emissivity retrievals across the MAJIS observations. The corresponding results are presented in the revised manuscript, including updated figures (Fig.10 and Fig. 11) and a new table (Table 2). The retrieved roughness values (Table 2) are reported in degrees and represent the RMS surface slope for both cm–mm retrieved sub-pixel roughness and LRO/LOLA macroscopic values (meter-scale).

Furthermore, we compare the new results obtained using cm–mm scale roughness with those derived with the revised model assuming the sub-pixel scale roughness to be equal to the macroscopic roughness inferred from LRO/LOLA measurements. For this purpose, the previous Figure 11 has been replaced with a new version that directly compares the updated results. An additional figure to map these differences has been added in the Appendix. This comparison was introduced to better quantify the differences between the two assumptions and to assess the potential biases introduced by assuming that LRO/LOLA macroscopic roughness is representative of the sub-pixel roughness.

Moreover, the development of the combined reflectance and thermal emission model allowed us to revisit and refine the codes implemented for the modeling. By adopting average roughness values representative of different terrain units to describe the sub-pixel roughness, we were able to generate a reduced number of simulated fractal surfaces. In the previous implementation, the number of synthetic surfaces was significantly larger due to the non-negligible variability of the macroscopic roughness derived from LRO/LOLA measurements, which required sampling a grid of incidence angle, emission angle, and roughness parameters as was described in Section 2.3 of the original manuscript. In the updated implementation, the reflectance contribution is treated using the Hapke formalism consistently with Wohlfarth et al. (2023). We clarify, however, that under the very high phase-angle conditions of the JUICE lunar flyby (~90°), no photometric treatment can be regarded as strictly optimal, including Hapke-based formulations; in this context, the Hapke term should be interpreted as a physically motivated approximation of the reflectance contribution rather than as a unique or fully constrained solution.

The manuscript has therefore been revised accordingly, and the previous description of this discretization scheme has been removed, as it is no longer required in the updated modeling approach.

Reviewer comment: Line 494: The statement that the model “explicitly accounts for local topography through a fractal surface formulation” is misleading. Typically, “explicitly accounting for topography” implies the use of actual topographic information, such as a digital elevation model. Please clarify what is meant here.

Response: The sentence has been revised for clarity. Specifically, we intended to convey that the use of a synthetic fractal surface formulation would account for simulated sub-pixel roughness, thereby potentially capturing anisothermal effects. The manuscript has been updated to clarify this aspect. Furthermore, a reference to Rozitis et al. (2020) has been included to clarify the utilization of fractal surfaces to model sub-pixel roughness.

Reviewer comment: The thermal model uses albedo, but the terminology in the manuscript is inconsistent: the authors refer to it as directional-hemispherical albedo (Lines 435, 443) and bolometric albedo (Line 395). This should be made consistent, and the derivation of this quantity should be explained. Since albedo varies between mare and highland regions, it likely has a significant effect on the energy balance. How does this parameter influence the results?

Response: We thank the Reviewer for this comment. In the revised manuscript, we clarified the terminology and ensured consistency throughout the text. The model derives two related quantities: the hemispherical–directional reflectance (r_{hd}) and the corresponding directional–hemispherical albedo for reflected sunlight (A_{dh}). The directional–hemispherical albedo for self-heating ($A_{\mathrm{dh,th}}$) is fixed at 0.05 following Wohlfarth et al. (2023).
We also revised the manuscript to include a reference to Wöhler et al. (2017), which presents the method we adopted to compute A_{dh} directly from the JUICE/MAJIS data.

Reviewer comment: The model assumes an effective emissivity of 0.95. Is this based on a graybody assumption? I checked the original publication and Wohlfarth et al. (2023) use directional spectral emissivity derived from Apollo data. Please compare the graybody assumption using a single effective emissivity with an approach that combines a bolometric emissivity of 0.95 and spectral emissivity in the relevant wavelength range. Please revise the manuscript accordingly.

Response: The use of a coupled reflectance–emission model enables the retrieval of the directional spectral emissivity in the 5.0-5.5 µm spectral range without imposing prior assumptions, starting from the retrieved single-scattering albedo (Hapke 2005; Wohlfarth et al., 2023). This formulation allows the emissivity to vary spectrally, rather than assuming a constant value.

Reviewer comment: The authors also derive maps of macroscopic surface roughness. It is unclear whether this parameter is meaningful, given that thermal modelling typically requires roughness at millimeter-to-centimeter scales. Furthermore, there is a concern that the roughness map may partially absorb variations in bolometric albedo. Please analyze this interaction and discuss its implications.

Response: To address this concern, we have implemented the reflectance + thermal emission model by Wohlfarth et al. (2023), which explicitly accounts for sub-pixel roughness at mm–cm scales. In this framework, the retrieved roughness parameter represents the unresolved surface roughness within each pixel and is therefore consistent with the roughness scale typically required by thermophysical modelling.
We also evaluated the potential degeneracy between surface roughness and single scattering albedo. To investigate this effect, we performed 20 independent runs on cube C4, varying the initial guesses of the parameters. The results show a good convergence for the roughness values, indicating that this parameter is relatively well constrained. In contrast, the single-scattering albedo exhibits a larger dispersion, suggesting that it is more affected by degeneracy effects even in the case of cube C4.

Our tests therefore indicate that, within the spectral range considered, the model is only partially sensitive to the roughness–albedo degeneracy under optimal illumination conditions and high signal-to-noise ratio (SNR), as in the case of cube C4. For the other cubes (e.g., C1–C2), the degeneracy is significantly stronger due to their lower SNR and weaker signal, resulting from both lower surface temperatures and a reduced reflected component, which makes the separation between roughness and albedo more difficult to constrain.

Reviewer comment: Lines 493–494: The phrase "and thus thermal beaming" is misleading. Thermal beaming occurs primarily when the vectors pointing to the observer and the Sun are roughly aligned. Is this truly the effect being referred to?
Figure 11: The numbers in the legend are unclear. Roughness is typically expressed in terms of RMS slope or Hapke θ, so the manuscript should clarify what these values represent.
Line 496: The points listed here (sub-pixel topography, surface composition, and wavelength-dependent emissivity) are in fact the critical factors that drive differences in thermal emission. Their importance should be highlighted and discussed.

Response: We thank the Reviewer for these comments. We revised the sentences highlighted by the reviewer to improve clarity. In the revised version macroscopic roughness is expressed in terms of RMS slope (in degrees) for improved readability (see previous comments/responses). Furthermore, we revised the analysis presented in the last part of Section 2.3. In the previous version of the manuscript, the macroscopic roughness derived from LRO/LOLA measurements was used to constrain the sub-pixel roughness, which may introduce a formulation that is not fully representative of the physical processes occurring at sub-pixel scales (Bandfield et al. (2015)). In agreement with the reviewer's comments and the related revisions throughout this section, we replaced the previous Figure 11 with a new figure that compares the results obtained using the average sub-pixel roughness values modeled following Wohlfarth et al. (2023) with those derived under the previous assumption that macroscopic and cm-mm sub-pixel roughness are equivalent. The differences between these approaches are now presented and discussed in Section 2.3.

Reviewer comment: Comparison with Diviner:
No comments on this section.

Reviewer comment: Temperature analysis.
I do not agree with the conclusion of this section. I do not see complementary strengths in the approaches themselves. It is more due to implementation choices in this particular study, that some approaches perform better than others. Please revise.

Response:
We thank the Reviewer for this important remark and agree that the relative performance of the three approaches is influenced not only by their conceptual differences, but also by the specific implementation choices adopted in this study (e.g., spectral windows, priors/regularization strength, assumptions on roughness, and the treatment of geometry at high phase angle). We have revised the concluding statements of Sect. 2.5.1 to avoid implying that the approaches are inherently complementary in a general sense. Instead, we now describe the results as an intercomparison under the particular configurations implemented here, highlighting the corresponding trade-offs and sensitivities. We also clarify that the main value of applying multiple approaches is to provide a consistency check and to identify regimes where assumptions or implementation choices drive systematic biases.

Reviewer comment: Revise the emissivity analysis.

This section requires substantial revision.

Emissivity definitions: It is unclear whether the emissivity quantities are directly comparable. Several different emissivity definitions exist, with subtle but important differences. The manuscript does not specify which emissivity the Bayesian approach uses; if Kirchhoff's law was applied correctly, it should be the spectral directional emissivity. What type of emissivity does the empirical approach use? The roughness model assumes an effective emissivity, while the original publication used a spectral emissivity. Please ensure that all emissivities compared in Figure 14 follow the same definition.

Comparison with other lunar mid-infrared measurements: Wu et al. (2018) present disk-resolved infrared measurements of the lunar nearside at 3.8 µm, acquired by the Gaofen-4 weather satellite. These observations cover most regions targeted by MAJIS and fall within its spectral range, making the Gaofen-4 dataset well suited for comparison. Including such a comparison would allow a clear examination of:

- Emissivity differences between mare and highland terrains,
- Maturity-related effects, and
- The consistency of emissivity estimates across instruments.

Comparison with laboratory samples: Reflectance and emissivity spectra of Apollo return samples are available and could provide a direct spectral comparison. I suggest adding these spectra to Figure 15. Notably, the highland emissivity spectrum in Figure 15 appears generally higher than the mare spectrum around 4 µm, which contradicts the spectral shapes of Apollo samples (e.g., Müller et al., 2021, Figure 6). Please discuss this discrepancy.

Diviner comparison: The comparison with Diviner appears incomplete and potentially misleading. Diviner channel 3 measures emissivity at 7.8 µm, near the Christiansen feature (CF) of highland materials, whereas the CF of mare materials occurs at longer wavelengths (Greenhagen et al., 2010). This naturally produces higher emissivity for highlands at the specific wavelength of channel 3, but does not imply that highlands are generally more emissive than mare regions across the mid-infrared. Including additional channels around 8 µm would provide a more complete picture. Please clarify the meaning of this comparison.

Comparison with IIRS data: A comparison with data from the IIRS instrument onboard Chandrayaan-3 would, in general, be insightful. However, I think there is not yet a definitive dataset suitable for comparison.

Lunar hydration: Lunar hydration has been extensively studied, focusing on the 3 µm absorption band in reflectance spectra. How does the bump around 3 µm in Figure 15 relate to this hydration feature? Even though the companion paper (Langevin et al., this issue) addresses lunar hydration in MAJIS data, a brief discussion here would be valuable.

Retrieved reflectance spectrum: Figure 15 currently shows only retrieved emissivities. For evaluating the Bayesian retrieval algorithm, it would be helpful to also plot the retrieved reflectance spectrum, which could be directly compared with existing laboratory samples for consistency. How do the other methods treat reflectance spectra? Please add this analysis.

Lines 634–636: The statement that "The spectral crossover, governed by vibrational properties of silicate minerals and by particle-scale roughness, reconciles the MAJIS and Diviner observations and highlights their complementarity" is confusing and scientifically inaccurate.

- For lunar silicates, vibrational effects are primarily observed in the thermal infrared (TIR). To my knowledge, the spectral crossover region of lunar silicates is not significantly influenced by vibrational effects. While CO_3 or organic absorption bands exist in this range, they are not relevant for lunar materials. I recommend examining Apollo sample reflectance spectra available from the RELAB database for guidance.

- The term “particle-scale roughness” is unclear. For thermal emission, the relevant effect is macroscopic roughness at the thermal isolation scale (millimeter to centimeter scale), not individual particle properties.
- It is unclear what is meant by “reconciles”. Was there a specific discrepancy between MAJIS and Diviner observations? This should be clearly stated and justified.

The entire paragraph requires revision for scientific clarity and accuracy.

Lines 636–643: This paragraph currently reads more like an advertisement and does not add scientific value. Also, it is quite vague. Please revise to focus on evidence, analysis, and interpretation rather than general statements.

Spectral feature around 3.5 μm: This feature is unusual and requires further investigation:

- How does the radiance spectrum appear? Is the feature present in the raw radiance, or only in the retrieved emissivity spectrum?
    - If it is present in the radiance, it could indicate an instrumental effect.
    - If only in the retrieved emissivity, it may point to a retrieval artifact.
- How does the retrieved reflectance spectrum compare?
- What emissivity spectra are obtained from the roughness model and the empirical model in this region?

Addressing these questions will clarify the origin of this feature and improve confidence in the retrievals.

Response:

We thank the Reviewer for this detailed and helpful set of comments. We agree that the emissivity section required clarification and strengthening. In the revised manuscript we have:

(i) defined the emissivity quantity produced by each method and clarified that the Fig. 14 intercomparison is restricted to a common observable, namely an effective directional emissivity at 5.5 μm at the MAJIS observing geometry (Method 1: spectral directional emissivity retrieved by maximum-a-posteriori estimate (Rodgers, 2000); Method 2: emissivity inferred from $I_{5.5} \approx \varepsilon_{5.5} B_{\lambda}(\mathrm{T})$ assuming negligible reflected contribution; Method 3: grey effective emissivity retrieved over 5.0–5.5 μm within the roughness framework);

(ii) revised the Diviner discussion to avoid implying that highlands are generally more emissive than maria in the mid-IR, clarifying that the Diviner channel-3 behaviour at 7.8 μm reflects the position of the Christiansen Feature and is a narrow-band effect;

(iii) removed/rewrote statements that were overly general or physically inaccurate (including the sentence attributing the crossover to “vibrational properties” and “particle-scale roughness”);

(iv) added a diagnostic discussion of the ~3.5 μm structure, explicitly stating whether it is present in the original radiance data or emerges only in retrieved emissivity. The revised text now clarifies that no scene-wide absorption is seen in the C4 median radiance spectrum.

(v) added brief context and references regarding complementary mid-infrared datasets (e.g., Gaofen-4 at 3.8 μm).

(vi) added laboratory emissivity spectra of Apollo soil samples to Fig. 15. The comparison highlights that the apparent ordering between the MAJIS terrain-averaged emissivity curves around ~4 μm should not be interpreted as a robust mineralogical trend. In our dataset, the 3.0–4.7 μm interval corresponds to the reflected–thermal crossover regime, where the emissivity retrieval is ill-conditioned and small residual calibration/reflectance-modelling errors can propagate into the inferred emissivity, as discussed in Sect. 2.5.2 (and supported by the radiance-space check around 3.5 μm). Moreover, the MAJIS products represent effective directional emissivities under the JUICE flyby geometry (phase angle ~90°), whereas Apollo laboratory measurements represent emissivity under controlled conditions and are not strictly equivalent. We therefore restrict the geophysical interpretation to the thermally dominated part of the MAJIS window, where the mare–highland contrast is consistent across methods and broadly compatible with published laboratory and remote-sensing constraints.

Reviewer comment: Line 679: No, it's the Christiansen feature, not the Reststrahlen bands.
Response: Ok.

Reviewer comment: Lines 679 – 683: Is this really what was shown? Please revise!
Response: We thank the Reviewer. We agree that the original wording was too general and could be interpreted as claiming more than directly demonstrated. We have revised the paragraph to focus on the specific observational evidence shown in Fig. 15, emphasizing the robustness of the relative emissivity separation among the selected terrains within the MAJIS retrieval window and explicitly acknowledging the limitations in the crossover region.

Reviewer comment: Revise the conclusion.
The conclusion requires substantial revision. It appears to have been written in haste and contains multiple incorrect statements, unsupported logical inferences, and vague or confusing formulations. Please carefully check consistency with the results presented in the manuscript and revise accordingly. In particular, the following points should be addressed:
Line 694: The phrase "where reflected and thermal contributions overlap" is misleading. Reflected and thermally emitted radiance are superimposed across the entire scene, independent of incidence angle. Please revise this statement.
Line 698: The term "anisothermal regions" is not defined. Please clarify what is meant or remove the term.
Lines 697–698: I do not agree with this conclusion, as it is not explicitly demonstrated in the manuscript.
Line 701: The phrase "key outcome" is vague and informal. Please restate more precisely.
Lines 702–706: This paragraph is incorrect on multiple levels and must be revised. The authors attribute emissivity differences between mare and highland regions primarily to porosity, which contradicts the established literature. In the visible and near-infrared, mare basalts appear darker than feldspathic highlands primarily due to composition, notably higher abundances of pyroxene, olivine, and ilmenite, with maturity effects also playing a significant role. The wavelength range up to appx. 6-8 µm is affected by nanophase-iron–related space-weathering effects. Porosity may contribute, but only as a secondary effect. Please revise.
Line 705: The statement "display higher thermal inertia" is misplaced. Thermal inertia cannot be determined in the present observational scenario, as it requires nighttime measurements.
Line 706: Similarly, enhanced thermal conductivity cannot be inferred without nighttime data and should not be claimed.
Line 707: The statement "The comparison with Diviner further refines this picture" is vague. There is no refinement shown; only an additional data point. Please revise.
Lines 707–711: The argument in this paragraph is unclear and potentially misleading. Diviner Channel 3 measures emissivity at 7.8 µm, near the Christiansen feature (CF) of lunar highland materials, while mare CFs occur at longer wavelengths (Greenhagen et al., 2010). Due to the characteristic shape of silicate emissivity spectra, this naturally results in higher emissivity values for highlands at this specific wavelength. This does not imply that highlands are generally more emissive than mare regions across the mid-infrared, where mare materials typically exhibit higher emissivity due to compositional differences. The wording "with highlands appearing more emissive" should therefore be revised to clarify that this is a narrow-band effect driven by CF position, not a general emissivity property.
Moreover, the purpose of this comparison remains unclear. As stated before, a more meaningful discussion would involve comparisons with Gaofen-4 observations in the same wavelength range and with mid-infrared laboratory measurements of Apollo samples and analog materials.

Line 727: The statement “display higher emissivity and warmer daytime temperatures, consistent with denser and less weathered regolith” is incorrect. As discussed above, composition is the dominant factor.
Line 734: The claim that the study captures “regolith properties that control emissivity and roughness” is not supported by the presented analysis.
Lines 733–736: This paragraph is vague and lacks a clear scientific argument. In particular, the phrase “inversion across the Christiansen feature” does not constitute a meaningful analysis in this context.
Line 739: The statement regarding “how mineralogy, grain size and roughness jointly shape the thermal environment” partly repeats earlier claims and introduces grain size, which has not been discussed elsewhere in the manuscript. This does not align with the presented work.
Lines 744–750: Terms such as “unique,” “critical benchmark,” and “pivotal” read more like promotional language than scientific conclusions. Several claims are also unclear:

- Why does this constitute a *robust framework* for MAJIS exploration?
- Why is Callisto only mentioned at the very end, despite being a major MAJIS target?

Please rework the entire conclusion, ensuring that all statements are supported by the analysis, that interpretations are physically justified, and that limitations and shortcomings of the emissivity retrieval are clearly acknowledged. The revised conclusion should also explicitly reflect and integrate the concerns raised throughout the earlier sections of the review.

Response:
We thank the Reviewer for this detailed assessment. We have thoroughly reworked the Conclusions to ensure full consistency with the results presented in the manuscript and to remove unsupported or ambiguous statements.
In particular, we (i) revised the wording related to the reflected/thermal “crossover” regime to avoid implying that overlap occurs only under specific geometries, and instead framed it in terms of geometry-dependent relative contributions; (ii) removed undefined terminology (e.g., “anisothermal regions”) and replaced it with an explicit description of unresolved sub-pixel slopes/shadowing and sub-pixel temperature variability; (iii) removed claims that cannot be supported by the present daytime flyby scenario (e.g., thermal inertia/thermal conductivity inferences); (iv) revised the Diviner comparison to clarify that Diviner channel-3 emissivity at 7.8 µm is a narrow-band effect driven by Christiansen Feature position and should not be interpreted as a general mid-infrared emissivity reversal; and (v) removed overly general or promotional phrasing (“key outcome”, “unique”, “pivotal”, etc.), replacing it with evidence-based statements and explicitly acknowledging limitations of emissivity retrievals in the 3–5 µm crossover interval.
We believe the revised Conclusions now accurately reflect the scope of the work (a methodological/scientific consistency assessment rather than instrument-level validation) while preserving the main scientific value of the MAJIS lunar dataset and its relevance for future MAJIS analyses.