**Reviewers Comments and Authors Response**
**Manuscript No. egusphere-2026-1320**
**Title: Impact Comparison of Different Aerosol Types on Atmospheric Correction of Landsat 8 over Land**

Dear reviewers,

We sincerely appreciate the reviewer's careful evaluation and insightful comments on our manuscript. The suggestions have been very helpful in improving the clarity and quality of the study. We have carefully revised the manuscript accordingly and provide detailed point-by-point responses below. All corresponding revisions have been highlighted in the revised manuscript.

Thank you very much.

Shuning Zhang
On behalf of all authors
*****************************************************************************

Overall, the manuscript has clear research objectives, solid data, and standardized methods. It systematically compares the effects of three aerosol types on Landsat 8 atmospheric correction, demonstrating clear research value. However, improvements are still needed in the following aspects.

1.The introduction points out issues with LaSRC's Urban Clean assumption, but it does not clearly explain why it is necessary to compare the three aerosol types (MOD04/MOD09/Urban Clean), nor what dimensions of comparison are lacking in existing studies (global site coverage, full reflectance range, different land cover). The novelty and necessity of the study need to be further strengthened. The logical progression from fixed aerosol models to dynamic aerosol models is not sufficiently clear. It is recommended to restructure it as "commonly used models →limitations →three dynamic models selected in this study", highlighting the representativeness and rationality of the comparison.

Reply: Thank you for this constructive comment. We agree that the original Introduction did not present the motivation and logical progression clearly enough, particularly regarding the transition from conventional fixed aerosol models to dynamic aerosol models, the rationale for selecting the three dynamic aerosol models, and the specific comparison dimensions that are lacking in previous studies.

In the revised manuscript, we followed the reviewer's suggested logical sequence of commonly used models → limitations → three dynamic aerosol types selected in this study.

The first part, corresponding to commonly used models, was already discussed in the original manuscript and was therefore retained without major changes. In this part, we described the incorporation of dynamic aerosol models into satellite atmospheric correction algorithms, including the transition from the fixed continental aerosol type in LEDAPS to the Urban Clean dynamic aerosol model in LaSRC, as well as the dynamic aerosol modeling schemes used in MOD04/MYD04 and MOD09/MYD09 (Lines 89–105).

The second part, which we have revised to explicitly clarify the core research limitations, was also largely retained. We restructured this section to elaborate on two major deficiencies of current dynamic aerosol type research. First, the lack of comprehensive inter-comparison: the dynamic aerosol types adopted in MOD04, MOD09, and LaSRC have not been evaluated under a consistent AC framework, and differences in their aerosol subtype libraries and subtype-selection strategies may lead to variations in aerosol optical properties, inconsistencies in AOD retrieval, and subsequent SR uncertainties. Second, existing validations are insufficient, constrained by limited site coverage, narrow reflectance ranges, and inadequate land-cover-specific evaluations. We further retained the discussion on the urgent need for a unified evaluation framework that covers globally distributed sites, diverse land-cover types, multiple spectral bands, and the full 0–1 reflectance range. (Lines 106–129).

To better connect these two parts with the specific models investigated in this study, we added a new paragraph beginning with "To address these gaps" (Lines 130–139). This paragraph explicitly introduces the three representative dynamic aerosol types selected in this study: the MOD04-based type, the MOD09-based type, and the LaSRC Urban Clean type. We clarified that the MOD04-based and MOD09-based types are constructed according to the aerosol subtype definitions and parameterization strategies used in the MOD04 and MOD09 aerosol retrieval frameworks, respectively, while Urban Clean is one of the MOD09 aerosol subtypes adopted by LaSRC for Landsat 8 atmospheric correction. We also explained that all three types are dynamic because their aerosol size parameters and complex refractive index vary with AOD, but they differ in aerosol subtype libraries and subtype-selection strategies. This addition directly strengthens the rationale for comparing these three dynamic aerosol types in terms of their impacts on AOD and SR retrievals.

We hope these revisions have clarified the Introduction in terms of the necessity, representativeness, and rationality of comparing the three dynamic aerosol types in this study. The relevant text is provided in the Introduction section, Lines 89–139.

"

Lines 89–105: Meanwhile, dynamic aerosol models have been increasingly incorporated into satellite AC algorithms. A representative example is the evolution of Landsat AC processing, which transitioned from fixed aerosol assumptions to dynamic aerosol modeling. The LEDAPS algorithm (Vermote et al., 2007) applied to Landsat TM/ETM+ sensors adopts a fixed continental aerosol type, whereas the Land Surface Reflectance Code (LaSRC) algorithm for Landsat 8 integrates the Urban Clean dynamic aerosol model (Dubovik et al., 2002; Maciel et al., 2023; Vermote et al., 2016). Similarly, MODIS AC algorithms employ dynamic combinations of coarse-mode dust and fine-mode aerosol components (Kaufman et al., 1997; Remer et al., 2005), and later incorporated highly absorbing aerosol types to better represent global atmospheric variability (Ichoku et al., 2003; Remer et al., 2005). With the accumulation of global AERONET observations, generalized dynamic aerosol models have been developed to further improve MODIS aerosol products. Currently, aerosol information is available from two major MODIS aerosol retrieval algorithms: the MODIS Level-2 aerosol optical depth product (MOD04/MYD04) (Levy et al., 2007a; Levy et al., 2007b) and the MODIS atmospherically corrected SR product (MOD09/MYD09) (Dubovik et al., 2002; Lyapustin et al., 2021; Vermote et al., 2006b; Vermote et al., 2008), which employ distinct dynamic aerosol modeling schemes proposed by Levy et al. and Dubovik et al., respectively. These aerosol products have been widely utilized in operational applications, including NASA's Web-enabled Landsat Data project for improving Landsat 7 SR data and the AC of land observations acquired by the VIIRS sensor onboard the NPP satellite (Roy et al., 2010; Superczynski et al., 2017). Overall, while fixed aerosol models remain widely adopted due to their stability and ease of implementation, dynamic aerosol models provide improved adaptability under complex atmospheric conditions.

Lines 106–129: However, despite the widespread operational use of dynamic aerosol models, their performance has not been comprehensively evaluated under a consistent AC framework. Existing studies indicate that the use of different aerosol representations may introduce inconsistencies in AOD retrieval (Grey et al., 2006), which can subsequently propagate into SR uncertainties. In practice, although major aerosol and AC products, such as MOD04, MOD09, and the LaSRC algorithm, all adopt dynamic aerosol types, they employ different subtype libraries and subtype-selection strategies. This discrepancy may lead to variations in aerosol optical properties used in AC, whereas systematic inter-comparisons among these types remain limited. Therefore, a comparative assessment of widely used dynamic aerosol types is essential for improving retrieval accuracy and optimizing aerosol type selection (Zhang et al., 2022).

Additionally, existing validations of dynamic aerosol-related products are still insufficient. Among these products, the Landsat 8 SR product generated by the LaSRC algorithm—one of the most widely used dynamic aerosol-related

SR datasets—is a key dataset for operational applications. However, previous validation studies have reported systematic deviations over high-reflectance surfaces such as desert playas, where SR tends to be underestimated, particularly in short-wavelength visible bands including coastal aerosol and blue bands (Mann et al., 2024; Meghraj et al., 2023). Moreover, product documentation and independent validation studies indicate that SR retrieval uncertainties increase under challenging atmospheric or illumination conditions, with larger errors frequently observed in shorter visible wavelengths (Roy et al., 2014; Vermote et al., 2016). Despite these potential uncertainties, current evaluations of dynamic aerosol-related products remain limited in scope, with existing validation efforts constrained by narrow site coverage, restricted reflectance ranges, and insufficient land-cover-specific assessments. For example, Vermote et al. (2006b) validated the LaSRC algorithm using only 33 sites, with SR evaluation in visible bands primarily confined to low-reflectance conditions, which hampers a comprehensive understanding and effective use of the product.

The above limitations highlight that current evaluations of dynamic aerosol types are still incomplete. In particular, systematic inter-comparisons between different types remain lacking, and existing studies are restricted in terms of site coverage, reflectance range, spectral bands, and land-cover-specific validation. These limitations underscore the need for a systematic comparison of dynamic aerosol types to comprehensively assess their performance under diverse conditions.

Lines 130–139: To address these gaps, this study focuses on three representative dynamic aerosol types: the MOD04-based type, the MOD09-based type, and the LaSRC Urban Clean type. The MOD04-based and MOD09-based types denote dynamic aerosol types constructed according to the aerosol subtype definitions and parameterization strategies used in the MOD04 and MOD09 aerosol retrieval frameworks, respectively. Urban Clean is one of the MOD09 aerosol subtypes and is used as the aerosol model in the operational LaSRC algorithm. All three types are dynamic because their aerosol size parameters (radius, standard deviation, volume distribution) and complex refractive index vary with AOD. However, they differ in their aerosol subtype libraries and subtype-selection strategies. The MOD04-based type includes four aerosol subtypes selected according to location and season, whereas the MOD09-based type includes five aerosol subtypes selected according to the minimum aerosol retrieval error. Therefore, comparing these three types provides a basis for assessing how different dynamic aerosol subtype libraries and subtype-selection strategies influence AOD and SR retrievals in Landsat 8 AC.

”

2.There is too lengthy information in the first paragraph of the introduction, the research significance and research status of AC algorithm are all presented, and the logic is unclear. Reorganize this section and present it in several shorter paragraphs. In addition, the hybrid strategy proposed in the conclusion has no foreshadowing in the introduction, resulting in poor coherence between the two sections.

Reply: Thank you for this constructive comment. We agree that the original first paragraph of the Introduction was too lengthy and combined several different aspects, including the research significance of atmospheric correction, the development of AOD retrieval algorithms, aerosol-type classification, and the application of fixed aerosol models in radiative transfer models.

In the revised manuscript, we reorganized the original long paragraph into several shorter paragraphs and made minor revisions to improve the logical transitions among AC significance, AOD retrieval methods, aerosol-type classification, and fixed aerosol models. The revised structure first introduces the importance of AC and aerosol parameter retrieval, then summarizes the development and limitations of physically based, multi-angle/multi-sensor, and data-driven AOD retrieval methods, and subsequently introduces aerosol-type classification and fixed aerosol models within physically based AC frameworks.

In addition, we agree that the flexible/hybrid aerosol-type selection strategy discussed in the Conclusion section was not sufficiently foreshadowed in the Introduction section. To improve the coherence between these two sections, we revised the final sentence of the Introduction to indicate that the results can support more flexible, task-specific aerosol-model selection strategies. This statement provides a clearer connection to the condition-specific strategy discussed in the Conclusion section, which was derived from the comparative results of AOD and SR retrieval performance.

The reorganized description in the first part of the Introduction section and the revised final sentence of the Introduction (Lines 35–72 and Lines 145–148) are as follows:

“

Lines 35–72: Atmospheric correction (AC), a critical component of remote sensing data processing, aims to mitigate atmospheric scattering and absorption effects on electromagnetic signals and convert sensor-received radiance into accurate surface reflectance (SR). Aerosols are key atmospheric constituents that significantly influence radiative transfer processes, cloud formation mechanisms, and atmospheric environmental dynamics (Dubovik et al., 2000; Levy et al., 2007a; Satheesh et al., 2005). Accurate aerosol parameter retrieval therefore plays a fundamental role in atmospheric remote sensing and directly governs AC performance (Zhang et al., 2009). Numerous aerosol optical depth (AOD) retrieval algorithms have been developed to support aerosol characterization and AC applications. Physically based approaches remain the most widely adopted, with representative methods including the Dense Dark Vegetation algorithm (Kaufman et al., 1997) and the Deep Blue algorithm (Hsu et al., 2013). In recent years, multi-angle and multi-sensor retrieval strategies have further improved AOD estimation by enhancing surface ‒ atmosphere separation and reducing uncertainties in SR assumptions. For example, the Multi-angle Implementation of Atmospheric Correction (MAIAC) algorithm developed for MODIS (Lyapustin et al., 2018) and MISR multi-angle retrievals combining nine viewing geometries have demonstrated improved performance, particularly over bright and heterogeneous surfaces (Chen et al., 2024).

Data-driven approaches, such as random forest, XGBoost, and deep neural networks, have also shown promising performance in capturing nonlinear relationships between atmospheric parameters and spectral observations (Radosavljevic et al., 2007). Recent studies have further incorporated temporal and spectral dependencies through advanced architectures, such as Transformer-based models applied to Himawari-8 time series data (She et al., 2024) and multilayer aerosol retrieval integrating geostationary SEVIRI data with CALIOP vertical aerosol profiles (Pashayi et al., 2025). Despite these advances, data-driven and multi-sensor methods generally require extensive training datasets and complex model structures, and they often lack physical interpretability. Consequently, physically based algorithms remain the dominant approach in operational AC systems and serve as a standard reference for remote sensing applications.

Within physically based AC frameworks, aerosol-type classification provides an efficient strategy to represent aerosol optical properties and improve retrieval adaptability. The development of aerosol classification has been greatly supported by globally distributed ground-based observation networks, including the Aerosol Robotic Network (AERONET) (Holben et al., 1998), the Sun ‒ sky radiometer Observation Network (SONET) (Li et al., 2018), and SKYNET (Takamura et al., 2004). Early aerosol studies introduced simplified physical models with fixed parameters describing particle size distribution, chemical composition, and optical characteristics such as extinction coefficient, single scattering albedo, and phase function. Representative developments include lower-layer aerosol classifications proposed by Shettle et al. (Shettle et al., 1979), naval aerosol characterization by Gathman (Gathman et al., 1983), and desert aerosol models summarized by Longtin et al. (Longtin et al., 1988). Additionally, the World Meteorological Organization (WMO) identified representative aerosol categories including continental, maritime, urban, desert, biomass burning, and stratospheric aerosols in its reports (World Meteorological Organization, 1983; World Meteorological Organization, 1986).

These fixed-parameter aerosol models remain essential components in radiative transfer models (RTMs), ensuring computational efficiency and stable performance. For instance, the MODTRAN model includes predefined aerosol types such as rural, urban, maritime, and desert (Berk et al., 2018), while the 6S and vector 6SV models incorporate continental, urban, maritime, and desert aerosol models (Vermote et al., 1997). These classical aerosol representations have been widely applied in AC processing across diverse environmental conditions.

Lines 145–148: By providing a systematic comparison of widely used dynamic aerosol types, this study offers guidance for selecting appropriate aerosol types in operational AC and informs the development of more flexible, task-specific strategies, thereby improving AC performance in operational remote sensing applications.

”

3.The manuscript adopts a large number of acronyms, which reduces its readability for readers not familiar with atmospheric radiation community. It is recommended to eliminate unnecessary acronyms and retain only the commonly used ones. Moreover, since the author has already used acronyms, the full forms still used later in the text (e.g., atmospheric correction, lookup table), Please unify them.

Reply: Thank you for this valuable comment. We agree that the excessive use of acronyms in the original manuscript reduced readability, especially for readers who are not familiar with the atmospheric radiation and remote sensing communities. In the revised manuscript, we have carefully checked the entire text and removed unnecessary acronyms, and in particular, many acronyms that appeared only once, including ATREM, NDBI, and EOS, have been removed or replaced by their full terms to further improve readability, while retaining commonly used ones such as AC, AOD, SR, MOD04, MOD09, LaSRC, and AERONET. We have also unified the use of acronyms and full terms throughout the manuscript. Specifically, terms that are defined and subsequently used as acronyms, such as atmospheric correction (AC), surface reflectance (SR), aerosol optical depth (AOD), and lookup table (LUT), are now consistently referred to by their acronyms after their first occurrence.

These revisions have been applied throughout the manuscript. Representative changes can be found in the Introduction, Methods, and Results sections, including Lines 114–116, 255–256, and 357-358.

“

Lines 114–116: Among these products, the Landsat 8 SR product generated by the LaSRC algorithm—one of the most widely used dynamic aerosol-related SR datasets—is a key dataset for operational applications. (surface reflectance has been changed to SR)

Lines 255–256: A further approximation allows unknown atmospheric parameters ($\rho_{atm}$, $Tr_{atm}$, $S_{atm}$) to be computed only at standard pressure, leading to an additional reduction in LUT size. (lookup table has been changed to LUT)

Lines 357–358: Ozone data were derived from the Ozone Monitoring Instrument onboard NASA’s Aura satellite, which provides daily global ozone column densities in Dobson Units (DU) at a resolution of 2.5° × 2.5°. (OMI removed)

”

4. What is "BRDF" in Line 766.

Reply: Thank you for this helpful comment. After careful review, we found that the occurrence of BRDF at Line 766 was not its first occurrence. Therefore, we have provided the full form at the actual first appearance in the revised manuscript to improve readability.

The revised description in Section 5.2 (1) (Line 719–721) is as follows:

“

These surfaces are more likely to be affected by subpixel heterogeneity, bidirectional reflectance distribution function (BRDF) effects, adjacency effects, and shadowing.

”

5.The citation of references in the main text should use a standard format. For example, Line 74-75 should be " Ichoku et al. (2013) proposed highly absorbing aerosol...". Other nonstandard references in Line 98, Line 293-294, Line 218....

Reply: Thank you. We agree that several references in the original manuscript were not cited in a standard format. In the revised manuscript, we have checked and unified the citation format throughout the text. Specifically, when the author name is used as part of the sentence, the citation has been revised to the narrative format, such as “Author et al. (Year)”, instead of repeating the author name in both the sentence and the parenthetical citation.

For example, “Ichoku et al. proposed ... (Ichoku et al., 2003)” has been revised to “Ichoku et al. (2003) proposed ...”. Similar corrections have also been made for the citations of Vermote et al. (2006b), Levy et al. (2003), Friedl et al. (2002), and other nonstandard references throughout the manuscript.

Representative revisions in the manuscript (Lines 76–79, 239, and 314–315) are as follows:

“

Lines 76–79: Kaufman and Remer introduced fine-mode-dominated aerosol models such as urban/industrial and biomass burning/developing world aerosol types (Kaufman et al., 1997; Remer et al., 1998), while Ichoku et al. (2003) proposed highly absorbing aerosol models derived from African field observations (ECK et al., 2003).

Lines 239: To simplify the computational processes in Eq. (4), we employ the approach used by Vermote et al. (2006b) for MODIS.

Lines 314–315: As illustrated in Fig. 4, the quarterly spatial distributions of the three fine-mode and one coarse-mode aerosol subtype were adapted from the methodologies of Levy et al. (2003) and Friedl et al. (2002).

”

6.Figure 1 shows the distribution of AERONET sites and global climate zones. Is the atmospheric correction, the topic of this work, highly sensitive to climate regionalization? It is recommended to further present the dominant type or sub-types of aerosol at each AERONET site. In addition, in the subsequent validations of dynamic aerosol types, consider incorporating the spatial performances of AOD and SR, and revealing the possible impacts of different aerosol sub-types.

Reply: Thank you for this insightful comment. We agree that the original description of Fig. 1 could more clearly explain the purpose of using global climate zones as background information. In the revised manuscript, we have refined the description to emphasize that the climate-zone overlay is used to show the broad spatial and climatic representativeness of the selected AERONET sites, thereby providing a diverse environmental basis for evaluating the performance of dynamic aerosol models.

Regarding the suggestion to present the dominant aerosol type or subtype at each AERONET site, we agree that aerosol subtype information is important for understanding the effects of different aerosol models. However, we did not assign a single dominant subtype to each site because the aerosol subtypes in the evaluated dynamic aerosol models are not fixed site-level attributes. In the MOD04-based model, aerosol subtypes vary with geographic region and season, whereas in the MOD09-based model, aerosol subtypes are selected according to retrieval-error minimization. Therefore, the subtype associated with a given AERONET site may vary among observation dates, and assigning one static subtype to each site would oversimplify the temporal variability and selection mechanisms of the dynamic aerosol models.

To address this comment, we have revised the description related to Fig. 1 and added a statement in the Conclusion to acknowledge that further spatially explicit analyses of aerosol subtype occurrence and its impacts on AOD and SR retrievals would be valuable in future work.

The revised description related to Fig. 1 and the added statement in the Conclusion section (Lines 159–160 and Lines 912–914) are as follows:

“

Lines 159–160: This distribution demonstrates that the selected sites provide broad spatial and climatic representativeness, providing a diverse environmental basis for evaluating the performance of dynamic aerosol types.

Lines 912–914: Further spatially explicit analyses of aerosol subtype occurrence and its impacts on AOD and SR retrievals would also be valuable for better understanding the mechanisms underlying the performance differences among dynamic aerosol types.

”

7.Line 265, a unit error, the general O3 concentration is 300 DU or 0.3 atm-cm.

Reply: Thank you for pointing out this unit error. The ozone value should be expressed as 300 DU, which is equivalent to 0.3 atm-cm, rather than 300 cm-atm. In the revised manuscript, we have corrected the unit and revised the wording to describe water vapor and ozone as column quantities more accurately.

The revised description in Section 3.1 (2) (Line 285–287) is as follows:

“

The column WVC and total column ozone amount were fixed at 1.0 $g \cdot cm^{-2}$ and 300 DU, respectively, as their spatial and temporal variations exert only minor influences on the TOA.

”

8.In Equations (25)– (27), what do εi, ρe, ρt, and n denote respectively?

Reply: Thank you for your helpful comment. We agree that the variables in Equations (25)–(27) should be clearly defined for better understanding. In the revised manuscript, we have clarified that $\rho_i^e$ refers to the measured or computed value, $\rho_i^t$ represents the true value, $n$ indicates the number of calculations, and $\varepsilon_i$ denotes the difference between the computed value and the true value.

The revised description in the manuscript (Lines 388–389) is as follows:

“

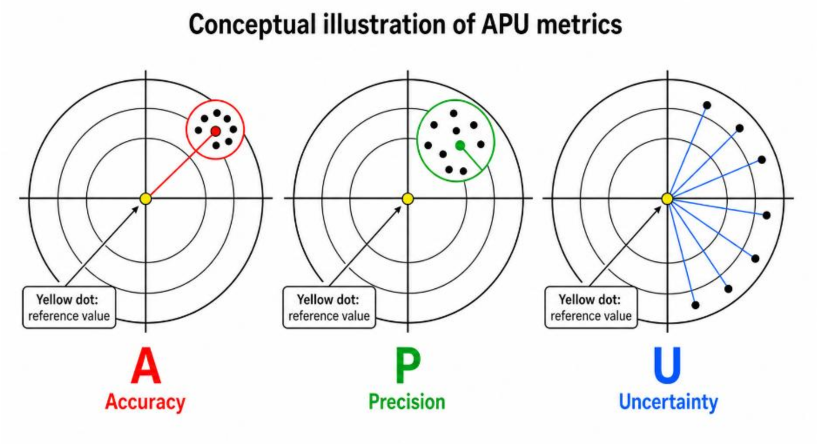
Where $\rho_i^e$ refers to the computed value, $\rho_i^t$ represents the true value, $n$ indicates the number of calculations, and $\varepsilon_i$ denotes the difference between the computed value and the true value.

”

9.Using the assessment metric "accuracy" may cause confusion with the word "accuracy" in the main text (e.g., Line 397, 420, 466, 483, 586...). The use of statistical terminology in the manuscript is confusing. The metric "precision" indicates "result consistency", how should we understand "result consistency"

Reply: Thank you for your constructive comment. We agree that the use of the terms "accuracy" and "precision" in the manuscript could potentially cause confusion, particularly given their common usage in other contexts.

In this study, A, P, and U are formula-based quantitative metrics in the APU framework, which characterize different aspects of retrieval deviations relative to the reference value. As illustrated in the schematic diagram, A describes the offset between the mean retrieval result and the reference value, P describes the spread of retrieval results around their mean retrieval result, and U describes the overall absolute deviation of retrieval results from the reference value. Specifically, P is represented by the radius of the green circle in the schematic diagram, indicating how dispersed the retrieval results are around their own mean. Therefore, in this study, “result consistency” refers to the degree to which retrieval results remain concentrated around their mean retrieval result. A smaller P indicates higher consistency among retrieval results, whereas a larger P indicates greater dispersion. Thus, P should be understood as a formula-based quantitative metric describing the consistency of retrieval results, rather than as “precision” in its broader statistical or colloquial sense.



To address this potential confusion, we have made several adjustments throughout the manuscript. In the revised manuscript, we have consistently used “A” (accuracy), “P” (precision), and “U” (uncertainty) throughout, replacing the full terms with these abbreviations, except in the definitions section. Additionally, in the abstract, we have clarified the terms by writing “accuracy (A), precision (P), and uncertainty (U)” to avoid any confusion. For example, the following revisions (Lines 466–467) were made:

“

The MOD04-based type had the lowest P (0.023184) and U (0.023366) values, indicating the most stable and least uncertain SR retrieval among the four products.

”

10.There are typos. Line 400: " MOD0-based"; Line 482: "domin·ant"

Reply: Thank you for your careful review. We agree that the typos pointed out in Line 400 ("MOD0-based") and Line 482 ("domin • ant") should be corrected. In the revised manuscript, we have made the necessary corrections.

Additionally, we have thoroughly reviewed the manuscript and made sure to address any similar errors, striving to prevent such issues from occurring elsewhere.

The specific revisions are as follows:

“

Lines 417–419: In terms of scatter distribution, most MOD04-based data points fell within the uncertainty range (blue solid lines) and were relatively concentrated.
Lines 500–501: Therefore, subsequent analysis focused on this range to reflect the dominant pixel distribution and ensure robust SR performance evaluation.

”

11.The texts and axis labels in many figures are too small to be legible; it is recommended to increase the font size. For the heatmap in Figure 6, only the numerical values can be placed within the plot, and the magnitude scale can be placed outside the plot (e.g., " ×10-2 " in plot title part or any blank space of the heatmap).
Reply: Thank you for this helpful comment. We agree that some figure elements in the original manuscript were too small, which reduced the readability of the figures. In the revised manuscript, we have enlarged the text elements and optimized the presentation of the relevant figures to improve readability.
Specifically, for Fig. 5, we enlarged and bolded the tick labels of both the x- and y-axes, as well as the in-figure text and legends. For Fig. 6, we followed the reviewer’ s suggestion and removed the scientific-notation parts, such as “e-02” or “e-03”, from the values inside the heatmap cells. To improve readability while maintaining a clean layout, the magnitude scale “$\times 10^{-3}$” was added after the colorbar title of each subplot. For Fig. 7, we simplified the x-axis tick labels from 21 values at 0.05 intervals to the more legible sequence of 0, 0.2, 0.4, 0.6, 0.8, and 1. Since the reflectance-bin definition has already been explained in the main text, the simplified tick labeling does not affect the interpretation of the figure and helps avoid overcrowding. The revised tick labels were also enlarged and bolded to further improve readability. In addition, the axis tick labels, axis titles, in-figure text, and legends in Fig. 7 were enlarged and bolded.
The revised versions of Fig. 5, Fig. 6, and Fig. 7 are as follows:

“

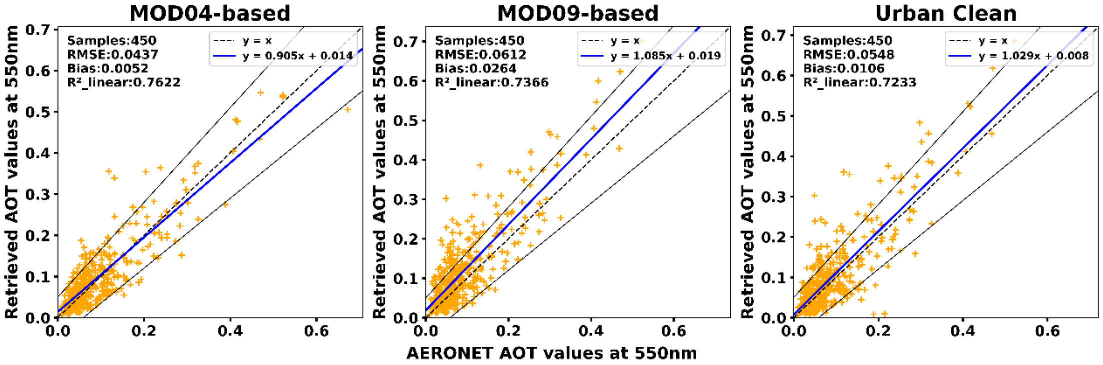
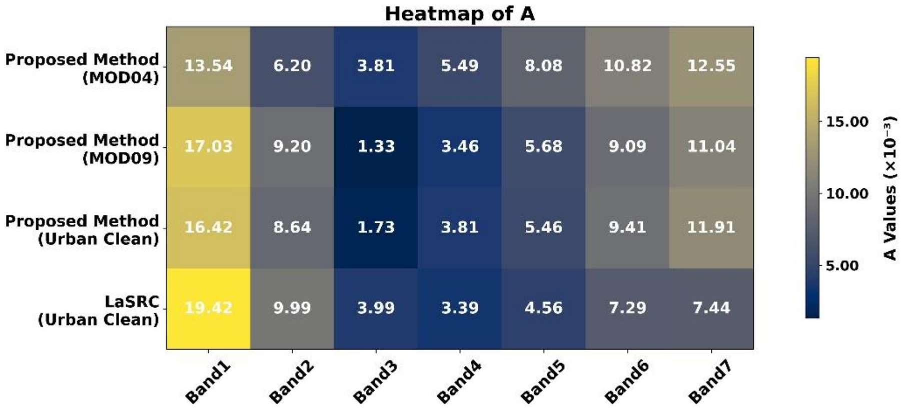
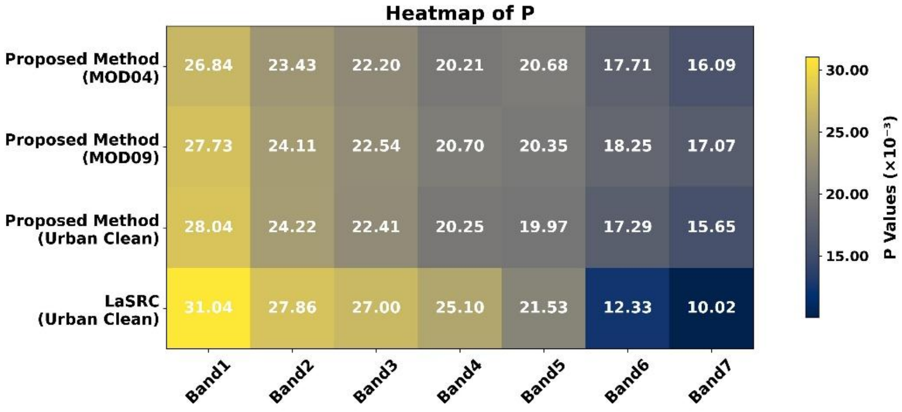


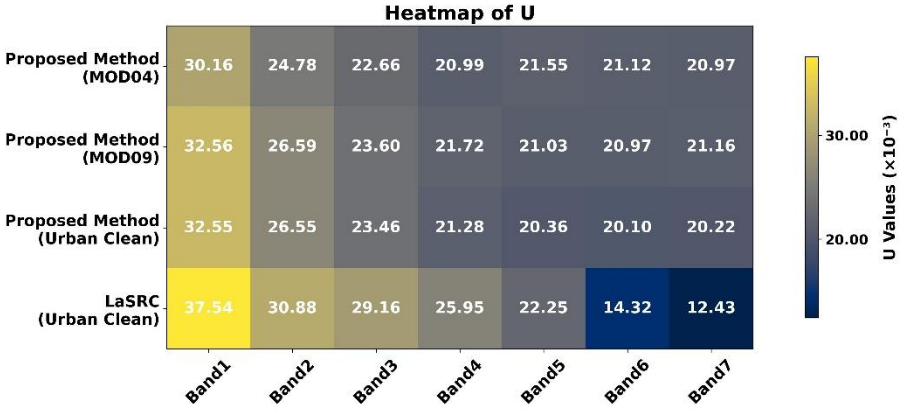
**Figure 5: Dispersion plot of 550 nm aerosol optical thickness(specs=0.15×AOD550+0.05).**



(a) A values of four SR products in single bands

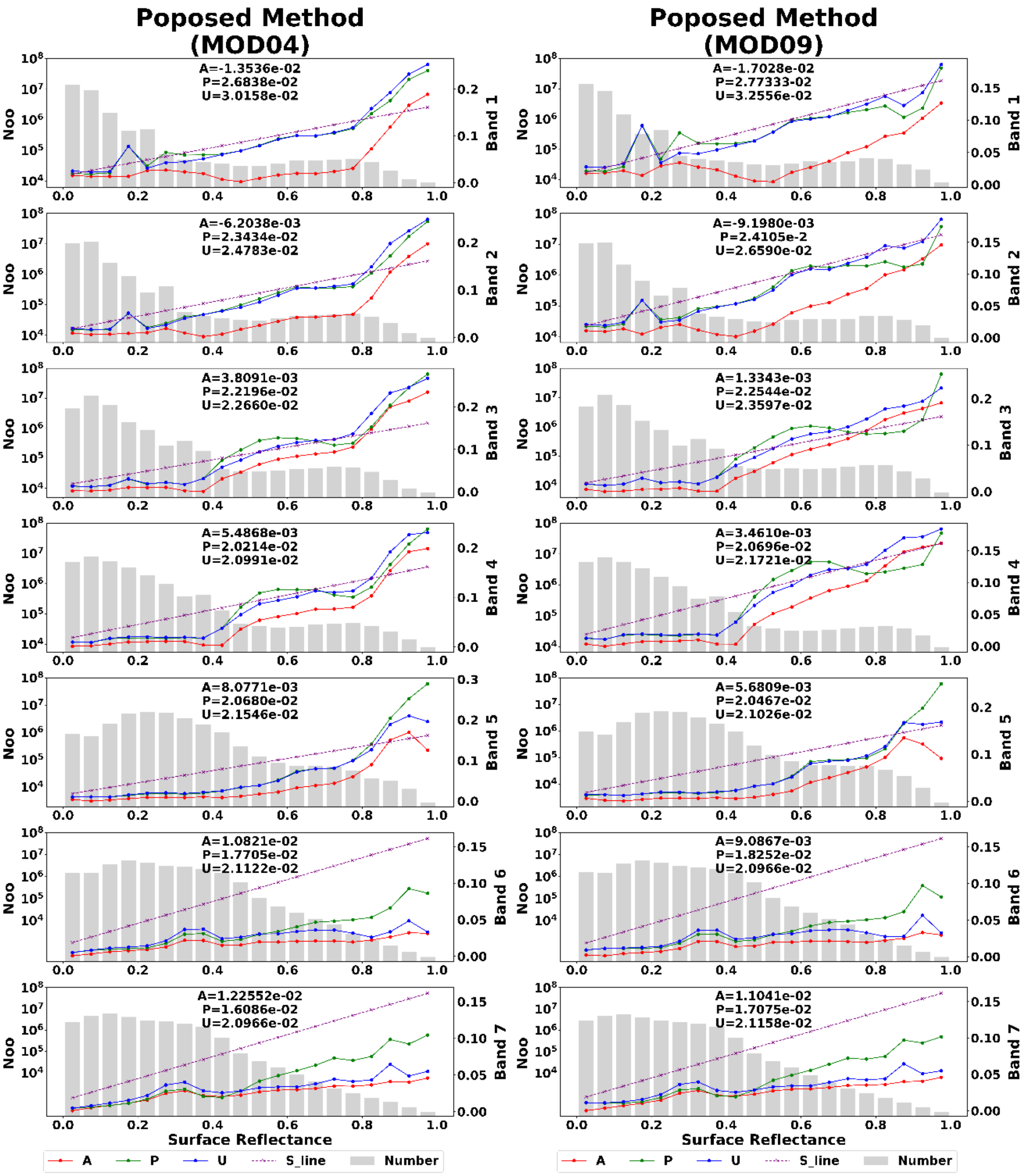


(b) P values of four SR products in single bands



(c) U values of four SR products in single bands

**Figure 6: Thermal map of the A, P, and U of four SR products in single bands.**

Poposed Method (MOD04)
Poposed Method (MOD09)
A=-1.3536e-02
P=2.6838e-02
U=3.0158e-02
A=-6.2038e-03
P=2.3434e-02
U=2.4783e-02
A=3.8091e-03
P=2.2196e-02
U=2.2660e-02
A=5.4868e-03
P=2.0214e-02
U=2.0991e-02
A=8.0771e-03
P=2.0680e-02
U=2.1546e-02
A=1.0821e-02
P=1.7705e-02
U=2.1122e-02
A=1.22552e-02
P=1.6086e-02
U=2.0966e-02
A=-1.7028e-02
P=2.77333e-02
U=3.2556e-02
A=-9.1980e-03
P=2.4105e-2
U=2.6590e-02
A=1.3343e-03
P=2.2544e-02
U=2.3597e-02
A=3.4610e-03
P=2.0696e-02
U=2.1721e-02
A=5.6809e-03
P=2.0467e-02
U=2.1026e-02
A=9.0867e-03
P=1.8252e-02
U=2.0966e-02
A=1.1041e-02
P=1.7075e-02
U=2.1158e-02
Noo
$10^4$ $10^5$ $10^6$ $10^7$ $10^8$
Band 1
Band 2
Band 3
Band 4
Band 5
Band 6
Band 7
0.0 0.2 0.4 0.6 0.8 1.0
0.00 0.05 0.10 0.15
0.0 0.1 0.2 0.3
Surface Reflectance
A
P
U
S_line
Number

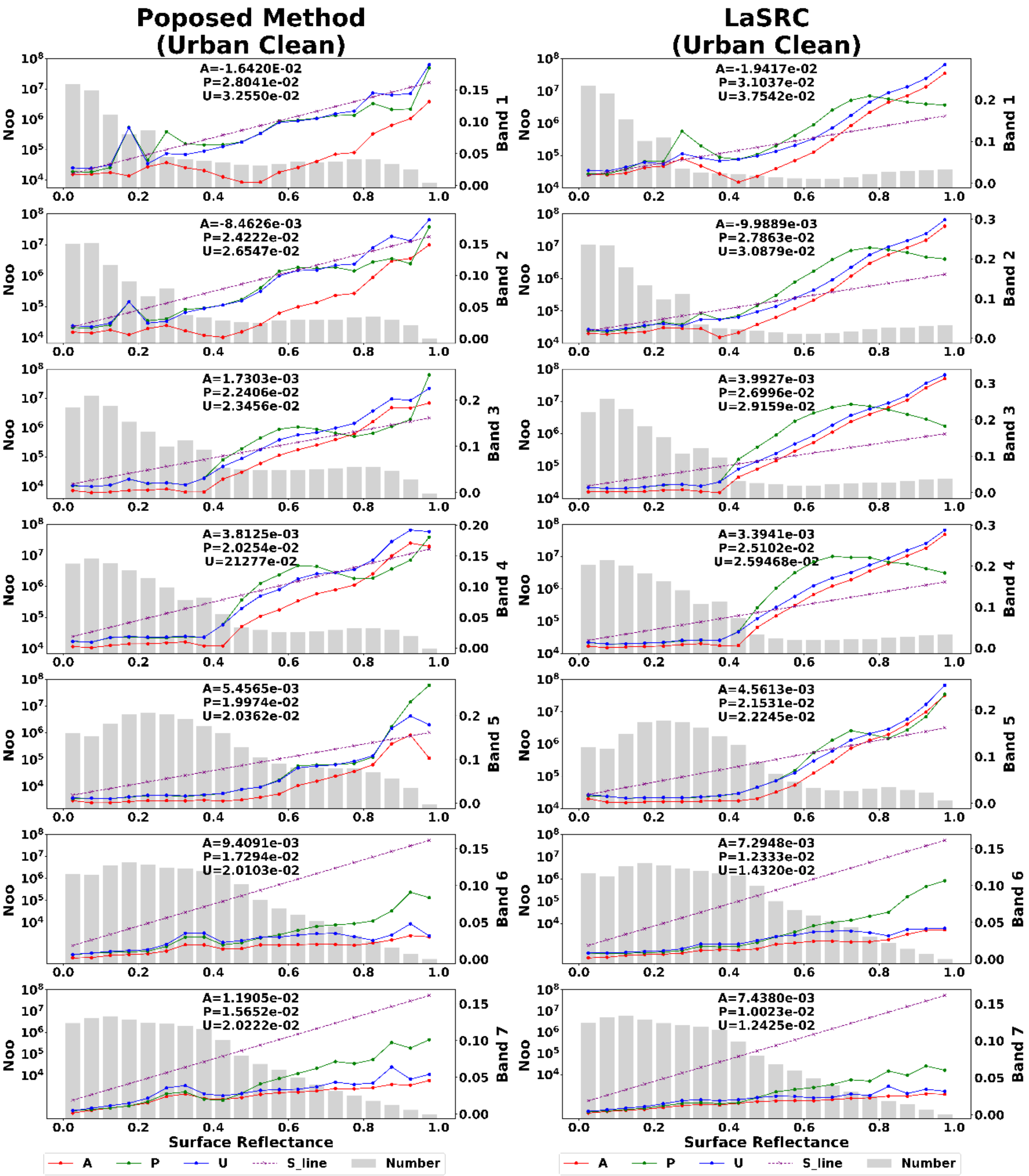


**Figure 7: Number of pixels (left y-axis) and accuracy metrics (A, P, and U; right y-axis) of the four SR products across different SR intervals (0–1 with a step of 0.1).**

”

12.Line 485-490: Provide some possible explanations for why these assessment metrics increase with the increase of SR. Are there any analogous studying results to confirm whether the current levels of accuracy, precision, and uncertainty are within acceptable ranges?

Reply: Thank you for this helpful comment. We agree that the original manuscript mainly described the increasing trend of A, P, and U with increasing SR, but did not sufficiently discuss the possible physical mechanisms behind this behavior.

In the revised manuscript, we have added possible explanations for the increase of A, P, and U with SR. Specifically, we explain that high-SR pixels are often associated with bright and heterogeneous surfaces, such as snow, bright bare soil, urban surfaces, and mixed pixels. These surfaces often exhibit strong spatial heterogeneity and anisotropic reflectance behavior, which can amplify BRDF effects, adjacency effects, shadow-related uncertainties, and mixed-pixel contamination. These effects can influence the estimation of path radiance and atmospheric transmittance and increase the discrepancy between retrieved and reference SR.

Regarding the interpretation of the metric levels, we further examined the APU ranges across different reflectance intervals. In the commonly observed SR range of 0 – 0.4, the APU values were generally low. For the visible bands, A, P, and U were mostly below 0.1 in this range. Specifically, for SR values of 0 – 0.2, A ranged from 0.0013 to 0.0374, P from 0.0096 to 0.0918, and U from 0.0097 to 0.0914; for SR values of 0.2 – 0.4, A ranged from 0.0044 to 0.0233, P from 0.0130 to 0.0634, and U from 0.0135 to 0.0539. The NIR and SWIR bands showed lower metric values in the same reflectance range. For the NIR band, A, P, and U ranged from 0.0036 to 0.0104, 0.0099 to 0.0197, and 0.0105 to 0.0217, respectively, for SR values of 0 – 0.4. For the SWIR bands, the maximum U value was approximately 0.037 in the SR range of 0 – 0.4.

Larger APU values mainly occurred in high-reflectance intervals. For example, when SR exceeded 0.6, the visible-band A, P, and U values increased to 0.0613 – 0.3127, 0.1009 – 0.2770, and 0.1005 – 0.2295, respectively. The NIR band also showed increased APU values under higher-reflectance conditions. In contrast, the SWIR bands remained relatively stable even in high-reflectance intervals, with maximum A, P, and U values of approximately 0.046, 0.106, and 0.055, respectively, for SR values greater than 0.6. These results indicate that the relatively large APU values were mainly associated with high-reflectance visible and NIR conditions, whereas the dominant low-to-medium reflectance ranges showed substantially lower errors.

We also compared these results with previous SR validation and AC intercomparison studies while considering differences in validation design, number of sites and scenes, spectral bands, reflectance ranges, and reference data. Vermote et al. (2016) evaluated the Landsat 8/OLI LaSRC SR product over AERONET sites using reference SR values generated by the 6SV radiative transfer model constrained with AERONET atmospheric measurements (e.g., AOD and water vapor). Based on the ranges shown in their validation results, the APU values were generally within approximately 0.01 in visible reflectance ranges below about 0.5 and within approximately 0.05 in SWIR reflectance ranges below about 0.6. Marujo et al. (2021) evaluated Landsat-8/OLI and Sentinel-2/MSI atmospheric correction products over five AERONET sites in Brazil using 154 Landsat-8/OLI scenes and 247 Sentinel-2/MSI scenes. Their reference SR values were generated using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software constrained with atmospheric measurements from AERONET sites. For the Landsat-8/OLI results, their validation showed band-dependent APU behavior, with uncertainties in the blue band generally remaining below approximately 0.07 for reflectance ranges of 0–0.25, below approximately 0.02 in the red band for ranges of 0–0.35, and below approximately 0.03 in SWIR Band 7 for ranges of 0–0.5. Doxani et al. (2023) assessed multiple Landsat 8 and Sentinel-2 atmospheric correction processors in the ACIX-II Land intercomparison using 79 AERONET sites worldwide. Their reference SR values were also generated using 6SV simulations constrained with atmospheric measurements from AERONET sites. Although their assessment did not separate reflectance intervals, the reported overall APU values across all spectral bands were generally below approximately 0.05, with most processors maintaining uncertainties around 0.01.

Compared with these studies, the SR uncertainties obtained in this study are generally within a comparable range for low-to-medium reflectance intervals. Slightly larger APU values mainly occurred in high-reflectance conditions, particularly for SR values above 0.6. This is reasonable because the present study evaluated a broader SR range (0–1), including highly reflective surfaces that are more susceptible to nonlinear reflectance effects, amplified path radiance, adjacency effects, and residual aerosol-

correction uncertainties.
These previous studies show that APU or SR uncertainty levels depend strongly on spectral band, reflectance range, surface condition, processor, number of sites/scenes, and reference dataset. Therefore, although the results are not directly comparable on a one-to-one basis, the APU levels observed in this study are considered reasonable for the dominant low-to-medium reflectance ranges, while the larger values in high-reflectance visible and NIR intervals reflect known challenging conditions for SR retrieval.

Vermote, E., Justice, C., Claverie, M., and Franch, B.: Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product, Remote Sens. Environ., 185, 46 – 56, https://doi.org/10.1016/j.rse.2016.04.008, 2016.
Marujo, R. F. B., Vermote, E., Roger, J.-C., Franch, B., and Skakun, S.: Evaluating the impact of LaSRC and Sen2Cor atmospheric correction algorithms on Landsat-8/OLI and Sentinel-2/MSI data over AERONET stations in Brazilian territory, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-3-2021, 271 – 278, https://doi.org/10.5194/isprs-annals-V-3-2021-271-2021, 2021.
Doxani, G., Vermote, E., Roger, J.-C., Gascon, F., Adriaensen, S., Frantz, D., Hagolle, O., Hollstein, A., Kirches, G., Li, F., Louis, J., Mangin, A., Pahlevan, N., Pflug, B., Vanhellemont, Q., and Vuolo, F.: Atmospheric Correction Inter-comparison eXercise, ACIX-II Land: an assessment of atmospheric correction processors for Landsat 8 and Sentinel-2 over land, Remote Sens. Environ., 285, 113412, https://doi.org/10.1016/j.rse.2022.113412, 2023.

The revised description in Section 5.2 (1) (Lines 708–723) is as follows:

"

The vertical analysis in Fig. 7 reveals consistent trends in the accuracy metrics (A, P, and U) across all four products. With increasing wavelength from the VIS to SWIR bands, all three metrics generally decreased, indicating an overall improvement in SR retrieval accuracy. In the VNIR bands, especially in the deep-blue and blue bands, the A, P, and U values were relatively large, and some high-reflectance intervals exceeded the expected error limits, indicating stronger retrieval uncertainties at shorter wavelengths. In contrast, the SWIR bands showed more stable performance, with all APU values remaining within the expected error limits. These trends can be attributed mainly to wavelength-dependent aerosol effects. Shorter wavelengths are more sensitive to aerosol scattering and AOD uncertainty, so residual errors in aerosol estimation can be amplified in the deep-blue and blue bands. Conversely, weaker aerosol scattering and reduced sensitivity to AOD uncertainty in the SWIR bands lead to more stable SR retrieval and lower uncertainty.
Horizontal analysis showed a systematic increase in A, P, and U with increasing SR across all products (Fig. 7), indicating reduced retrieval accuracy under high-reflectance conditions. This behavior may be related to the characteristics of bright and heterogeneous surfaces, such as snow, bright bare soil, urban surfaces, and mixed pixels. These surfaces are more likely to exhibit strong subpixel heterogeneity and anisotropic reflectance behavior, which can enhance bidirectional reflectance distribution function (BRDF) and adjacency effects and increase the influence of shadows and mixed pixels. These effects can influence the estimation of path radiance and atmospheric transmittance, thereby increasing the discrepancy between retrieved and reference SR.

"

13.Line 581: What is the data source for land-cover?
Reply: Thank you for this comment. To clarify the data source for the land-cover information used in the land-cover-specific SR accuracy assessment, we have revised Section 4.2(4). In this revision, we specified that the land-cover samples were derived from a subset of Landsat 8 scenes randomly selected from the

634 scenes used for SR validation. Spectral indices calculated from these Landsat 8 reflectance data—including NDBI, NDSI, NDVI, NDWI, and BSI—were used to identify representative pixels for building, snow, soil, vegetation, and water surfaces. For each land surface type, sample extraction continued until more than 5,000 pixels were obtained.

In addition, Section 3.2(2) was updated to include relevant citations supporting the use of these spectral indices, which documents the basis for using these indices in our analysis.

The corresponding revision in Section 3.2(2) (Lines 392–395) and Section 4.2(4) (Lines 601–604), is as follows:

“

Lines 392–395: For each land-cover category, the corresponding pixels were extracted using commonly-used remote sensing ecological index, including the Normalized Difference Building Index (Zha et al., 2003), the Normalized Difference Snow Index (Hall et al., 2011), the Bare Soil Index (Rikimaru et al., 2002), NDVI (Rouse et al., 1974) and the Normalized Difference Water Index (McFeeters et al., 1996).

Lines 601–604: To evaluate SR accuracy for different land surface types, land-cover samples were derived from a subset of Landsat 8 scenes randomly selected from the 634 scenes used for SR validation. Representative pixels for each surface type were extracted from these images using remote sensing spectral indices, including NDBI, NDSI, NDVI, NDWI, and BSI, ensuring more than 5,000 sample points for each type.

References:

Zha, Y., Gao, J., and Ni, S.: Use of normalized difference built-up index in automatically mapping urban areas from TM imagery, Int. J. Remote Sens., 24, 583 – 594, https://doi.org/10.1080/01431160304987, 2003.

Hall, D. K. and Riggs, G. A.: Normalized-difference snow index (NDSI), in: Encyclopedia of Snow, Ice and Glaciers, edited by: Singh, V. P., Singh, P., and Haritashya, U. K., Springer Netherlands, Dordrecht, 779–780, https://doi.org/10.1007/978-90-481-2642-2_376, 2011.

Rikimaru, A., Roy, P. S., and Miyatake, S.: Tropical forest cover density mapping, Trop. Ecol., 43, 39 – 47, 2002.

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W.: Monitoring vegetation systems in the Great Plains with ERTS, in: Third Earth Resources Technology Satellite-1 Symposium, NASA SP-351, Vol. 1, 309 – 317, 1974.

McFeeters, S. K.: The use of the normalized difference water index (NDWI) in the delineation of open water features, Int. J. Remote Sens., 17, 1425 – 1432, https://doi.org/10.1080/01431169608948714, 1996.

”

14.Regarding the captions of Figures 8 and 9, the results are not derived solely from building (Figure 8) and snow cover (Figure 9).

Reply: Thank you for pointing out this inconsistency. We confirm that the issue was caused by figure captions that were not properly updated during manuscript revision. The content description was correct, where Fig. 8 and Fig. 9 present RMSE and bias, respectively, across seven spectral bands.

The captions of Fig. 8 and Fig. 9 have now been revised to accurately reflect their contents. In addition, the in-text reference in Section 4.2(4) (Lines 584–586) has been slightly clarified to explicitly indicate that Fig. 8 corresponds to RMSE and Fig. 9 to bias, in order to improve clarity and avoid ambiguity.

The revised description in Section 4.2(4) (Lines 604–605) and the revised captions of Fig. 8 and Fig. 9 are as follows:

“

Lines 604–605: The performance of the four SR products was assessed using RMSE and bias across seven spectral bands, as shown in Fig. 8 (RMSE) and Fig. 9 (bias).

Figure 8: RMSE of the four surface reflectance products across seven spectral bands.

Figure 9: Bias of the four surface reflectance products across seven spectral bands.

”

15.Line 611-618: MOD04-based has the largest RMSE and bias over soil, this is opposite with those over other land covers. Can the author explain some of the reasons?

Reply: Thank you for this helpful comment. We agree that the original manuscript described the relatively larger RMSE and Bias of the MOD04-based model over soil surfaces, but did not sufficiently explain why this behavior differs from that over other land-cover types.

In the revised manuscript, we have added a possible explanation. Soil surfaces typically exhibit smooth spectral signatures and relatively homogeneous spatial distributions, which favor stable atmospheric correction performance. Under such conditions, aerosol models with stable single-type assumptions, such as Urban Clean and LaSRC, generally provide more consistent SR retrieval with lower RMSE and Bias. In contrast, the MOD04-based model employs dynamic aerosol subtype selection based on geographical location and seasonal information, while the MOD09-based model selects subtypes according to a residual-minimization criterion. Although aerosol types containing multiple subtypes can improve adaptability under complex or strongly aerosol-sensitive conditions, over soil surfaces they may introduce additional variability in aerosol optical properties due to the surface’s spectral uniformity. Such variability can perturb path-radiance and atmospheric-transmittance correction, especially in the short visible bands, leading to higher RMSE and stronger Bias compared with the more stable Urban Clean and LaSRC products. Therefore, the weaker performance of MOD04-based, and the limited advantage of MOD09-based, over soil surfaces does not necessarily contradict their advantages over other land-cover types. Instead, it suggests that the benefit of dynamic aerosol subtype selection is dependent on surface type and wavelength.

The revised description in Section 4.3 (3) (Lines 793 – 805) is as follows:

“

For soil surfaces, SR retrievals show relatively stable spectral behavior. Soil spectra typically vary smoothly with wavelength and often exhibit comparatively homogeneous spatial distributions. Although soil reflectance is influenced by mineral composition, organic matter, and moisture content, these factors usually introduce gradual spectral variations, which favor stable AC performance. Among the four products, Urban Clean and LaSRC generally demonstrate stable behavior over soil surfaces, with Urban Clean showing low RMSE in the VIS bands and LaSRC showing strong performance in several longer-wavelength bands. In contrast, the MOD04-based and MOD09-based models, both of which involve dynamic aerosol subtype selection, do not show consistent advantages over soil surfaces. This may be because, for surfaces with smooth spectral signatures and relatively stable reflectance behavior, additional aerosol subtype selection may introduce variability in aerosol optical properties without providing clear compensating benefits. Such variability can perturb path-radiance and atmospheric-transmittance correction, especially in the short visible bands, leading to larger RMSE or stronger Bias compared with the more stable Urban Clean and LaSRC products. Therefore, the weaker performance of MOD04-based and the limited advantage of MOD09-based over soil suggest that the benefit of dynamic aerosol subtype selection is surface- and wavelength-dependent.

”

16.It is recommended to strengthen the explanation of the physical mechanisms, as most of the analysis in the manuscript remains descriptive. The Discussions (Section 5) appears to repeat descriptions from the Results (Section 4).  It is advised to remove these redundant descriptions and instead add more mechanistic descriptions and physical interpretations. Add some discussion of research limitations, such as cloud contamination, snow/ice and urban heterogeneous underlying surfaces, and the limitations of AOD and reflectance range, and present them in the future directions.

Reply: Thank you for the constructive comments. We have carefully revised the Discussion section based on the reviewers' suggestions.

In the AOD Discussion, repeated descriptions as in the Results section have been removed, and physical mechanisms were strengthened, including explanations of the advantages of MOD04-based aerosol type in the VNIR bands and the influence of AOD discontinuities under the MOD09 residual-minimization strategy on SR retrieval. The last paragraph now includes the limitations in this study, such as the single-year dataset, 465 AERONET data points, and the predominance of AOD values below 0.2, along with corresponding future directions, including extending the analysis to multi-year datasets and incorporating additional high-quality AOD sources.

In the SR Discussion, Section 5.2(2) has been updated to enhance the mechanistic explanations for the MOD04 and MOD09 products, explicitly addressing the VNIR bands' sensitivity to AOD, the relative sensitivity of moderate-to-low reflectance pixels to AOD errors, and the effects of MOD09 residual constraints and smoothing on SR retrieval. Section 5.2(3), which focuses on SR retrieval across different land-cover types, has been updated to enhance mechanistic explanations for the performance of the four SR products across spectral bands and surface types. Limitations analyses have been added at the end of both Sections 5.2(2) and 5.2(3), highlighting factors such as the relatively limited number of Landsat 8 scenes (634), Landsat 8 spatial resolution, uncertainties in land-cover extraction, and the underrepresentation of complex surfaces, including dense forests, snow/ice, and heterogeneous urban areas. Future directions are integrated within the sections, including further refinement of MOD09 smoothing, analysis of LaSRC shortwave advantages, and the use of more precise land-cover classification products and expanded datasets to assess SR retrieval across diverse surface types, and the revised descriptions in Sections 5.1 (Lines 671–700) and 5.2 (2)(3) (Lines 724–852) are as follows:

"

### 5.1 AOD retrievals

The scatter plots and regression analyses presented in Fig. 5 show clear differences in AOD retrieval performance among the three aerosol types. MOD04-based retrievals generally align more closely with AERONET observations, with a larger proportion of points falling within the empirical uncertainty range. Urban Clean performs moderately, while MOD09-based exhibits greater dispersion and systematic deviations. Statistical analyses of RMSE and Bias indicate that these differences are significant, confirming that MOD04-based retrievals outperform the other two types under the evaluated conditions.

The superior performance of MOD04-based type can be attributed to its dynamic aerosol subtype-selection mechanism, which incorporates seasonal and geographic information. For a fixed location and season, MOD04 typically selects the same aerosol type for the majority of images, with only a very small fraction of images near type boundaries exhibiting two aerosol types. As a result, the continuity of AOD retrieval is generally closer to actual atmospheric conditions. In contrast, MOD09-based subtype selection, which is based on the residual-minimization criterion, is more likely to result in images in which multiple aerosol types are present. Examination of the AOD subtype results indicates that fine-scale, pixel-level mixtures of multiple aerosol types do not occur within individual images, while the presence of two aerosol types across the image is relatively common, with each subtype distributed over multiple contiguous pixels, forming extended continuous regions. To mitigate this, an overall smoothing process was applied to the AOD results, which may have contributed to the larger errors observed for the MOD09 type.

Urban Clean employs a single aerosol classification, limiting its adaptability under variable atmospheric conditions. While its retrievals are generally more consistent than MOD09-based ones in some cases, they are less accurate than MOD04-based when actual aerosol properties diverge from the assumed type. These patterns suggest that physically informed aerosol subtype selection with spatial and temporal variability, is crucial for achieving accurate AOD retrieval.

Several limitations of the current study should be acknowledged. First, the analysis is based on a single year and 465 AOD data points from AERONET, covering diverse climate types globally, limits both the spatial coverage of sites and the overall dataset size. In results, the robustness of AOD retrieval has not been assessed across a wider range of locations or with larger datasets, potentially restricting the generalizability of the findings. In addition, most AERONET observations correspond to AOD values below 0.2, indicating higher aerosol loading or partially cloud-obscured conditions less frequently recorded. Consequently, the performance of AOD retrieval under elevated aerosol loading, as well as under extreme surface conditions such as snow, ice, or heterogeneous urban areas, remains largely untested due to the limited representation of these scenarios in the current dataset.

Future research should extend these evaluations to multi-year datasets to better assess seasonal and interannual variations in AOD retrieval performance, and incorporating additional high-quality AOD datasets beyond AERONET could provide broader coverage of atmospheric conditions and improve the generalizability of findings.

## 5.2 SR retrievals

2) Different performance

The validation results revealed distinct performance variations among the four SR products.

MOD04-based: The MOD04-based type demonstrated superior retrieval accuracy in the VNIR bands across moderate-to-low reflectance ranges, outperforming the other products. This advantage is likely due to the incorporation of region- and season-specific aerosol parameters, which enhance aerosol subtype selection and optimize AOD retrieval. Because aerosol scattering effects are stronger at shorter wavelengths, inaccuracies in aerosol characterization can propagate more directly into SR retrievals in the VNIR region. This effect is particularly evident over moderate-to-low reflectance surfaces, where even small AOD retrieval errors may produce amplified SR deviations. Consequently, the MOD04-based SR product achieves higher accuracy in these bands and reflectance ranges, reflecting its ability to better represent spatiotemporal aerosol variability and reduce systematic biases under complex atmospheric conditions.

In summary, the MOD04-based type demonstrated exceptional inversion accuracy and generalizability in the VNIR bands under moderate-to-low reflectance scenarios, which can be attributed to its comprehensive consideration of the geographic and seasonal aerosol dynamics, making it a reliable reference for SR retrieval with broad potential for global-scale AC applications in the future.

MOD09-based: The validation results indicated that the SR product corresponding to the MOD09-based aerosol type performed exceptionally well in a relatively narrow high-reflectance range (>0.8). Across the first six bands examined in this study, although the errors for all four products exceeded the acceptable range in this interval, the MOD09-based type effectively controlled these errors and demonstrated a superior retrieval accuracy. This advantage can be attributed to the application of residual constraints during AOD retrieval, which enables effective error regulation, thereby improving the SR retrieval accuracy in high-reflectance ranges, particularly in remote sensing imagery of large-scale high-reflectance surfaces (e.g., snow). However, this residual-minimization strategy also carries certain risks. The introduction of residual constraints may cause abrupt changes in the subtypes of adjacent pixels, which may not reflect the reality for the spatial continuity of aerosol distributions. To mitigate the uncertainty resulting from these abrupt transitions, an overall smoothing process was applied to the AOD results. However, this approach has inherent limitations and may introduce additional small errors, which together contribute to the relatively moderate SR performance of the MOD09-based product outside the high-reflectance range. Consequently, future research should aim to enhance the spatial continuity of aerosol retrieval by investigating more appropriate smoothing strategies or alternative adjustment methods to minimize the potential adverse effects of these abrupt transitions on AOD and SR retrieval accuracies.

In conclusion, despite the exceptional performance of the MOD09-based aerosol type in high-reflectance ranges, making it a valuable tool for processing remote sensing images with extreme SR conditions, its subtype selection method under residual constraints is associated with certain risks.

Urban Clean: Among the four SR products, the one using the Urban Clean aerosol type, a subtype of the MOD09-based aerosol type integrated into the LaSRC algorithm, exhibited the weakest overall performance. It demonstrates limited advantages only at specific intervals, such as the moderate reflectance ranges in band 6, with generally inferior inversion accuracy compared with other products. Its error patterns partially align with the MOD09-based aerosol type but remain systematically less robust. This underperformance likely stems from the specialized design as a subtype of the MOD09-based design, which is optimized for urban areas with low air quality and stringent application constraints. Consequently, their narrow applicability limits their global utility, particularly in heterogeneous environments.

In summary, although Urban Clean shows localized efficacy, its restricted operational scope and demanding implementation criteria render it non-competitive with other types of global SR inversion tasks.

LaSRC product: The LaSRC product demonstrated marked superiority in the SWIR bands, especially in Band 7 (2.1 μm), where its SR estimates exhibited enhanced precision and reliability. This advantage is amplified with increasing reflectance levels, which is likely attributable to the refined handling of water vapor and other absorptive gases that critically influence the SWIR bands. The LaSRC algorithm effectively mitigates atmospheric interference, enabling high-fidelity retrieval in these spectral ranges. Notably, its performance surpassed that of the other products in high-reflectance SWIR scenarios.

However, LaSRC exhibits elevated error levels in the VNIR bands. The A-value (systematic bias) consistently exceeded that of the other products, suggesting pronounced deviations. This discrepancy may originate from two aspects: 1. the residual quality control method used in LaSRC, which flags pixels failing band 4/5/7 spectral tests as water bodies, potentially propagating AOD retrieval errors into VNIR reflectance estimates, and 2. the atmospheric parameters in the LaSRC method interpolated using AOD via cubic spline functions, which may introduce errors.

In conclusion, although the excellent performance of LaSRC in the SWIR band highlights its potential for high-precision reflectance retrieval, its systematic VNIR errors require further investigation to guide the refinement of algorithms.

These comparisons revealed the underlying factors contributing to the differences in product performance, such as the adaptability of MOD04-based aerosol parameterization. These insights offer practical guidance for optimizing the selection of dynamic aerosol types based on scene characteristics and highlight the potential advantage of the LaSRC algorithm in improving reflectance accuracy through a more effective correction of absorptive gas effects.

Several limitations of the current study should be acknowledged. The analysis is based on a relatively limited dataset of 634 Landsat 8 scenes, and the AERONET stations used for validation are generally located in open areas to ensure reliable measurements. As a result, the spatial distribution of validation sites is constrained, providing representativeness primarily at the level of climate zones but not capturing all land-cover types. Although some urban sites are included, complex surfaces such as dense forests, snow/ice-covered areas, and heterogeneous urban regions remain underrepresented. Expanding the dataset in future work could improve the generalizability of SR validation across a wider range of surface types.

### 3) SR retrievals with different landcover

Fig. 8 – 9 indicate that SR retrieval accuracy exhibits clear spectral dependence that varies with surface type. For snow and vegetation, retrieval errors are generally larger in shorter wavelengths and decrease toward longer wavelengths, reflecting the stronger influence of aerosol scattering in the VIS region. Soil shows relatively stable performance across spectral bands, while water exhibits a distinct spectral behavior, with retrieval accuracy

improving toward longer wavelengths. These variations reflect the combined effects of aerosol scattering, surface spectral characteristics, BRDF effects, subpixel heterogeneity, and the signal strength of different land-cover types. For soil surfaces, SR retrievals show relatively stable spectral behavior. Soil spectra typically vary smoothly with wavelength and often exhibit comparatively homogeneous spatial distributions. Although soil reflectance is influenced by mineral composition, organic matter, and moisture content, these factors usually introduce gradual spectral variations, which favor stable AC performance. Among the four products, Urban Clean and LaSRC generally demonstrate stable behavior over soil surfaces, with Urban Clean showing low RMSE in the VIS bands and LaSRC showing strong performance in several longer-wavelength bands. In contrast, the MOD04-based and MOD09-based types, both of which involve dynamic aerosol subtype selection, do not show consistent advantages over soil surfaces. This may be because, for surfaces with smooth spectral signatures and relatively stable reflectance behavior, additional aerosol subtype selection may introduce variability in aerosol optical properties without providing clear compensating benefits. Such variability can perturb path-radiance and atmospheric-transmittance correction, especially in the short visible bands, leading to larger RMSE or stronger Bias compared with the more stable Urban Clean and LaSRC products. Therefore, the weaker performance of MOD04-based and the limited advantage of MOD09-based over soil suggest that the benefit of dynamic aerosol subtype selection is surface- and wavelength-dependent.

For vegetation and building surfaces, retrieval performance is strongly influenced by surface complexity. Vegetation canopies consist of multilayer structures with strong anisotropic reflectance, while urban areas contain diverse materials and complex geometric configurations. These characteristics increase spatial heterogeneity and anisotropic reflectance behavior, which can amplify BRDF effects, adjacency effects, shadow-related uncertainties, and mixed-pixel contamination, thereby increasing AC uncertainty. MOD04-based shows advantages in the deep-blue and blue bands for these complex surfaces, especially for vegetation and building pixels in the short-wavelength region. This may be related to its aerosol subtype selection scheme that incorporates seasonal and geographic variability, which can better represent aerosol optical properties under conditions where short-wavelength SR retrieval is highly sensitive to aerosol scattering. In contrast, MOD09-based uses an error-minimization-based subtype selection strategy. Although this strategy can be effective in some retrieval conditions, it may not always provide additional advantages over spectrally complex surfaces because the selected aerosol subtype may be optimized for retrieval residuals rather than for land-cover-specific SR behavior. Urban Clean and LaSRC show relatively stable performance in the NIR and SWIR regions for several land-cover types. As aerosol scattering effects weaken with increasing wavelength, AC becomes less sensitive to aerosol-type selection, and stable aerosol assumptions can provide consistent retrieval behavior. This may explain why Urban Clean performs well in the NIR and SWIR regions for vegetation and water surfaces, and why LaSRC remains competitive in longer wavelengths. In contrast, multi-subtype selection strategies used in MOD04-based and MOD09-based may introduce additional variability when the sensitivity of SR retrieval to aerosol-type differences is reduced.

For water and snow surfaces, SR retrieval is strongly affected by their extreme reflectance characteristics. Water surfaces are characterized by very low reflectance, weak surface signals, and additional variability caused by surface roughness, adjacency effects, and specular reflection, all of which can increase retrieval uncertainty, especially in the VIS bands. Snow surfaces, on the other hand, exhibit extremely high reflectance in the VIS region and strong directional reflectance effects, which may increase retrieval sensitivity and introduce systematic biases. Among the four products, MOD04-based demonstrates stable performance over snow surfaces and in the VIS bands for water pixels, likely benefiting from its geographically and seasonally constrained aerosol parameterization in short-wavelength regions. Urban Clean shows advantages over water surfaces in longer wavelengths, reflecting its stability in spectral regions less sensitive to aerosol scattering. MOD09-based generally exhibits larger retrieval variability over snow and water surfaces, suggesting that its subtype selection based on retrieval-error minimization may be less stable under extreme reflectance conditions. LaSRC demonstrates stable performance for soil surfaces and maintains competitive accuracy in longer wavelengths, but shows relatively larger deviations over high-reflectance snow surfaces and low-reflectance water surfaces in the VIS region.

Overall, the analysis indicates that aerosol treatment strategies play a critical role in determining SR retrieval performance across different surface types. Physically constrained and seasonally/geographically adaptive aerosol schemes, such as MOD04-based types, demonstrate improved robustness in short-wavelength bands and under temporally varying conditions, while stable single-type assumptions provide more consistent performance over smoother surfaces or in spectral regions with reduced aerosol sensitivity. The performance of each aerosol type is not uniform across land-cover types, but depends on the combined effects of wavelength, SR magnitude, surface structural complexity, and aerosol sensitivity.

However, the current SR analysis over different landcover still has several inherent limitations. Both the 30-meter spatial resolution of Landsat 8 and the inherent uncertainties associated with extracting land-cover types using empirically derived RSEIs limit the accuracy of SR validation across different surface types. Although approximately 5,000 pixels were used in the analysis, this sample size remains insufficient to fully represent all land-cover types. Additionally, vegetation and snow surfaces exhibit pronounced seasonal and interannual variability, which was not explicitly stratified in the current study and may influence aerosol-type performance. The current results provide preliminary evidence of these effects, as MOD04-based aerosol types consistently achieve lower RMSE and Bias over snow surfaces and demonstrate advantages over vegetation in several short-wavelength bands. These patterns likely reflect the MOD04-based aerosol subtype selection scheme, which incorporates geographical and seasonal information, enabling better adaptation to temporally varying surface–atmosphere conditions. A more comprehensive assessment would require finer land-cover classification, explicit temporal stratification, and an expanded dataset. Future work could therefore involve validation using more refined land-cover products and larger pixel samples to further evaluate SR performance across surface types with pronounced temporal variability.

”