
Response to Reviewer #2

Manuscript ID: egusphere-2026-2510

Title: A blue-band surface residual correction method for land aerosol optical depth retrieval from DQ-1 WSI

Journal: Atmospheric Measurement Techniques

Dear Reviewer,

We thank the reviewer for the careful and constructive assessment. In response to the concerns about validation scope, aerosol-model representativeness, and methodological reproducibility, we substantially revised the manuscript and Supplement.

Most importantly, the independent AERONET evaluation now covers three periods and 104 station–date matchups in total: 27 in spring 2025, 49 in autumn–winter 2025, and 28 in spring 2026. The autumn–winter analysis used a season-specific static prior together with a seasonally adjusted LUT and RF model, whereas the spring-2026 analysis used the spring-2025 configuration without retraining or recalibration. We also evaluated the prescribed aerosol spectral dependence using 678 AERONET records from four sites spanning 34 observation dates and added methodological details on the 6S/LUT implementation, gas-transmittance treatment, angular interpolation, matching

cost, regional comparison, and data availability. The revised manuscript also states the scope and remaining limitations more explicitly.

Our detailed responses are provided below. For clarity, the reviewer's comments are shown in **black** font, our responses in "***ITALIC BOLD BLUE***" font, and quoted revisions to the manuscript in "***Red***" font.

General comment

This manuscript describes a method to obtain an a priori surface reflectance for aerosol optical depth (AOD) retrieval from the WSI satellite instrument over the Beijing–Tianjin–Hebei (BTH) region in China. It is a two-step approach, first creating a static reference data base (beforehand) and then applying a correction based on mismatch at 443 nm (on a scene-by-scene basis). The static data base is obtained using atmospheric correction of WSI data from low-AOD days using MODIS MAIAC AOD retrievals as an input. The correction step improves performance of the AOD retrieval, based on comparisons of results using the static vs. corrected surface reflectance against ground-based AERONET observations at one location. The idea seems promising and to my knowledge is novel. Generally a prior surface constraint is necessary because surface reflectance modeling errors are a large contributor to AOD retrieval uncertainty, particularly when AOD is low, and a single-view multispectral imager generally does not have sufficient information content to constrain both spectral aerosol and surface information from a single overpass. As such the manuscript falls within AMT's scope. I appreciate the authors having incorporated suggestions made during the Quick Review process as this has made the manuscript easier to follow.

My main concern is about the extent of the data processed and analysed.

It is very restricted: to the BTH region only and for three months only (March to May 2025). It is not clear how generalisable to other parts of the world or times the results would be (despite the authors' claims that it is, it is not supported by the analysis). This short scope also makes it hard to judge the quantitative comparisons with AERONET. The data volume is not enough to make robust conclusions, and it is common for new algorithm papers to describe higher data volumes than this. Since the satellite launched in April 2022 it should be easy to process more data, even if only in the same study region. If nothing else this will also shed light on issues such as whether the initial 3-month data base is valid for other years, or times of year. Additionally, the aerosol optical models seem to have been selected for convenience and are not well justified by the analysis shown in the manuscript. These concerns are touched on by the authors in section 5.2 but I think more should be done on them in the current work. As a result, I recommend major revisions and encourage the authors to expand the analysis to make the results more robust. My specific comments are below:

Response: *We thank the reviewer for the positive assessment of the motivation and novelty of the study and for identifying the limited validation scope and aerosol-model justification as the main concerns. We agree that the original spring-2025 analysis was too limited to support broad conclusions regarding temporal robustness or general applicability. We have therefore substantially expanded the analysis in the revised manuscript.*

The independent AERONET evaluation now covers three periods and 104 station–

date matchups in total: 27 in spring 2025, 49 in autumn–winter 2025, and 28 in spring 2026. For autumn–winter 2025, a season-specific static prior was constructed, the LUT was adjusted for the seasonal viewing and atmospheric conditions, and the RF model was retrained using representative scenes from that period. In contrast, the spring-2026 evaluation used the spring-2025 retrieval configuration without retraining or recalibration, providing an interannual evaluation under a similar seasonal regime. We also broadened the regional spatial analysis by replacing the earlier municipal cutouts with a BTH-scale comparison. To address the concern regarding the predefined aerosol models, we added an observational comparison using 678 AERONET records from four BTH sites spanning 34 observation dates to examine whether the prescribed spectral AOD dependence is representative of the study region. The revised manuscript also states more explicitly that the present results do not establish year-round or cross-regional transferability, and that dense summer vegetation, snow-dominated surfaces, broader aerosol variability, and regions outside BTH remain to be evaluated. These revisions directly address the two main issues raised in the general comment—the limited data volume and the justification of the aerosol assumptions. The corresponding changes and additional analyses are described in detail in our responses to Comments 6, 7, and 12 below.

Specific comments

1. Line 122: What is the satellite overpass time window? This should be stated here.

Response: *We have added the DQ-1 WSI overpass time in both the sensor description and the AERONET matchup section. The nominal overpass time over the study region is approximately 13:30 Beijing time, corresponding to approximately 05:30 UTC. We have also clarified the temporal matching windows used for the different validation periods.*

Changes in the manuscript:

Section 2.2 now states:

“The DQ-1 atmospheric environment monitoring satellite was launched on 16 April 2022 and operates in a Sun-synchronous orbit at an altitude of approximately 700 km, with an overpass time of about 13:30 Beijing time over the study region (Wang et al., 2024b).”

Section 2.3 further states:

“The nominal DQ-1 WSI overpass time over the study region is approximately 13:30 Beijing time (approximately 05:30 UTC).”

For spring 2025 and spring 2026, AERONET observations between 02:00 and 06:00 UTC were used, while the autumn–winter 2025 evaluation used observations within ± 60 min of the nominal overpass time.

2. Line 126: A citation is needed for the MAIAC aerosol data set here.

Probably: Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC algorithm, Atmos. Meas. Tech., 11, 5741–5765, <https://doi.org/10.5194/amt-11-5741-2018>, 2018.

Response: *Thank you for pointing this out. We have added the suggested MAIAC reference when introducing the MCD19A2 aerosol product.*

Changes in the manuscript:

Section 2.3 now states:

“The MODIS MAIAC MCD19A2 AOD product (Lyapustin et al., 2018) was also used to support AOD-screened surface-prior construction and regional cross-product comparison.”

3. Line 136-140: I think everything from “To ensure” to “navigation data” should be deleted because it is to be expected that the algorithm is working in calibrated physical units and that geometry is needed. If this text is kept in some form then a link or reference to the CRESDA calibration data should be given (if this calibration is not something

which is included by standard in the WSI level 1 data).

Response: *We agree that the original wording was unnecessarily general. We have therefore revised this part to describe only the preprocessing steps actually applied to the WSI data. Because the visible-to-shortwave-infrared channels were converted from digital numbers to TOA reflectance using the official CRESDA calibration coefficients, we retained this information for reproducibility and added the corresponding CRESDA calibration-data source.*

We have also added the CRESDA data portal for the DQ-1 WSI observations in the Code and data availability section.

Changes in the manuscript:

Section 2.4 now states:

“The WSI visible-to-shortwave-infrared channels were converted from digital numbers to top-of-atmosphere (TOA) reflectance using the official radiometric calibration coefficients released by the China Centre for Resources Satellite Data and Application (CRESDA). The 10.8 μm thermal infrared channel was converted to brightness temperature, while solar zenith angle, viewing zenith angle, and relative azimuth angle were obtained from the accompanying navigation data.”

The Code and data availability section now provides both the DQ-1 WSI data portal and the official calibration-data webpage:

“DQ-1 WSI data used in this study were obtained from the [China Centre for Resources Satellite Data and Application \(CRESDA\) data portal](#) and remain subject to its data-access policy. The official DQ-1 radiometric calibration coefficients and spectral response information are available from the [CRESDA calibration-data webpage](#). The processed matchup data, validation statistics, and figure data generated in this study are available from the corresponding author upon reasonable request.”

4. Lines 144, 150: Acronyms such as MNDWI, NDVI, NDSI should be defined at first use.

Response: *We have defined these indices at their first use in the revised manuscript.*

Changes in the manuscript:

Section 2.4 now states:

“Water pixels were identified mainly using the modified normalized difference water index (MNDWI; Xu, 2006), the normalized difference vegetation index (NDVI), and related spectral constraints, while snow pixels were detected using the normalized difference snow index (NDSI).”

5. Sections 3.1, 3.1 and in general: the manuscript is not always clear about what it means by “reflectance”, e.g. line 170 talking about “surface reflectance terms” (what quantities?). I recommend adopting nomenclature from Schaepman-Strub et al (2006) where possible:

<https://www.sciencedirect.com/science/article/pii/S0034425706001167>

Reading sections 3.1 and 3.2 talking about the use of MAIAC AODs and WSI reflectance, what exactly is being calculated for the static data base?

The paper mentions Lambertian assumption using 6S for atmospheric correction but then talks about geometry being used for the random forest correction later. What I infer is being done is that the static data base is created using Lambertian atmospheric correction of WSI images using MAIAC AOD (at which wavelength, this needs to be stated?) but that aerosol optical properties (i.e. spectral and directional dependence of scattering and absorption) come from OPAC which is built into 6S (not from MAIAC – which creates an inconsistency). The multiple scenes used to generate the data base are combined to give the final data base value using rules based on brightness and NDVI (are these the “matching criterion” used to pick aerosol model mentioned in line 208? This is also

unclear). Thus there is no direct accounting for BRDF effects in here. Then the random forest (RF) method used in the correction approach takes in geometry so I guess is making the first-order correction for BRDF effects (since if I read correctly it is trained on 3 of the individual scenes going into the data base), but there is no inherent physical BRDF model. Is all this interpretation correct? This all feels very empirical but I suppose if it works it works, at least within this limited domain. Although the authors state that the retrieval is not circular because the MAIAC AOD is not input into the process here, I am not fully convinced, because the training target of this RF method is a discrepancy between the data base and the surface reflectance obtained using MAIAC AOD so it is still indirectly in there. However, since the training was done based on low-AOD conditions it seems likely that the bulk of the RF correction targets e.g. BRDF effects and not indirectly consistency with MAIAC aerosols.

Response: *We agree that the original description did not sufficiently distinguish the different reflectance quantities and the roles of MCD19A2, the 6S aerosol models, the reference-date selection, and the RF correction. We have therefore reorganized and expanded Sects. 3.1–3.4.*

In the revised manuscript, “surface reflectance” in the LUT framework is explicitly defined as a Lambertian-equivalent quantity under the Lambertian surface assumption. During offline construction of the static surface prior, MCD19A2 AOD at 550 nm is used only to constrain aerosol loading in the LUT inversion. The spectral and angular aerosol optical properties are specified by the predefined 6S aerosol model rather than taken from MAIAC. The brightness–NDVI score is used only to select the reference date for each pixel and is not the criterion used for

aerosol-model selection.

We have also clarified the role of viewing geometry in the RF model. Geometry-related variables allow the RF to represent geometry-dependent variability in the surface residual empirically, but the current framework does not contain an explicit physical BRDF model. We therefore do not interpret the RF correction as a physical BRDF correction.

Finally, we agree with the reviewer that MCD19A2 is not completely absent from the residual-learning process. Because MCD19A2 AOD is used in the offline LUT inversion that generates the reference surface term, it contributes indirectly to the RF training target. We now state this explicitly. MCD19A2 AOD is not an RF input feature and is not required in the online retrieval.

Changes in the manuscript:

Section 3.1 now clearly separates the offline and online stages:

“The proposed framework consists of an offline preparation stage and an online AOD retrieval stage. The offline stage is used to construct the static surface prior and train the blue-band residual model, whereas the online stage performs AOD retrieval without requiring any external AOD product.”

“For these scenes, collocated MCD19A2 AOD is used only as an external atmospheric constraint in the LUT inversion, together with WSI TOA reflectance, observation geometry, and elevation, to derive LUT-based reference surface terms.”

Section 3.2 now states:

“Under the Lambertian surface assumption, the 6S radiative transfer model (Kotchenova et al., 2006; Vermote et al., 1997) was used to simulate top-of-atmosphere (TOA) reflectance, with surface reflectance represented as a Lambertian-equivalent quantity in the LUT retrieval (Schaeppman-Strub et al., 2006).”

Section 3.3 further clarifies how the static prior is constructed:

“After invalid pixels affected by clouds, water, and snow were removed, the collocated MCD19A2 AOD at 550 nm was used to constrain aerosol loading in the LUT inversion, while the spectral and angular aerosol optical properties were specified by the predefined 6S continental aerosol model. Together with DQ-1 WSI

TOA reflectance, observation geometry, and elevation, these inputs were used to interpolate the LUT and inversely estimate the Lambertian-equivalent surface reflectance for the blue, green, red, and near-infrared bands on a pixel-by-pixel basis.”

Section 3.4 now states:

“The inclusion of solar and viewing geometry allows the RF to represent geometry-dependent variability in the surface residual empirically. No explicit physical BRDF model is included in the current framework. MCD19A2 AOD is used in the offline derivation of the LUT-derived reference surface term and therefore contributes indirectly to the RF training target. It is not included among the RF input features and is not required during online retrieval.”

6. This does lead in to a general concern (touched on in section 5.2 but not enough in my view): the study presents results only for a small geographic area and for three months (March-May 2025). It is not clear how generalisable this will be. For example, I would expect changing vegetation phenology and changing solar angles throughout the year to influence apparent surface reflectivity which would likely mean that additional data bases are needed for other times of year (this is what e.g. the NASA Deep Blue approach does). It’s also not clear that a data base created for one year would be suitable for use in other years.

Additionally, while this is a proof of concept for the BTH region, it’s not clear how the availability of suitable scenes to construct the static data base would be elsewhere in the world (due to e.g. availability of low-AOD, cloud-free, snow-free days and sufficient stability of surface characteristics). Presumably the WSI satellite is intended to eventually lead to AOD retrievals outside the BTH region (indeed, outside the cities

themselves – the only maps shown are described as urban areas, we don’t actually see any broader maps) so I would have appreciated some more analysis (or at least discussion) of that. It is not unusual for new algorithm papers to have a somewhat broader spatial/temporal scope.

Response: *We agree that the original spring-2025 analysis was too limited to support broader claims of temporal generalizability. We have therefore substantially expanded the temporal evaluation.*

First, we added an autumn–winter 2025 evaluation. For this period, a season-specific static surface prior was constructed from four autumn–winter reference scenes, the LUT was extended to larger solar zenith angles using the Midlatitude Winter atmospheric profile, and the RF model was retrained using the corresponding seasonal scenes. The resulting AERONET evaluation contains 49 station–date matchups from two sites spanning 26 independent observation dates. Second, we added a spring-2026 interannual evaluation. In this case, the spring-2025 residual model and retrieval configuration were applied without retraining or recalibration. The evaluation contains 28 station–date matchups from three AERONET sites spanning 17 independent observation dates. This evaluates the spring-2025 configuration in a different year under a similar seasonal regime. We also replaced the earlier municipal cutouts with a broader BTH regional comparison and added a true-color image to provide spatial context. At the same time, we have revised the manuscript to avoid claiming broader transferability than is supported by the data. The current study still does not systematically evaluate dense summer vegetation, snow-dominated conditions, or regions outside BTH. We now state these limitations explicitly and identify extension to additional seasons and regions as future work.

Changes in the manuscript:

Section 4.1 now reports:

“Figure 5 shows the AERONET validation results for autumn–winter 2025. After

excluding scenes used for RF training and held-out internal validation, 49 station–date matchups from two AERONET sites were retained, covering 26 independent observation dates.”

“To examine interannual performance, the spring-2025 retrieval configuration was applied to spring 2026 without using 2026 observations for retraining or recalibration. The evaluation included 28 station–date matchups from three AERONET sites spanning 17 independent observation dates.”

Section 5.2 now states:

“Across the three evaluation periods, the AERONET analysis includes 104 station–date matchups: 27 in spring 2025, 49 from two sites spanning 26 independent observation dates in autumn–winter 2025, and 28 from three sites spanning 17 independent dates in spring 2026. The autumn–winter evaluation was performed using a season-specific static prior together with a seasonally adjusted LUT and RF model, whereas the spring-2026 evaluation used the spring-2025 configuration without retraining or recalibration. These additional evaluations substantially extend the temporal coverage of the original analysis, although validation remains confined to the BTH region and does not constitute a systematic year-round or cross-regional assessment.”

“Dense summer vegetation, snow-dominated surfaces, and regions outside the BTH domain were not systematically evaluated.”

Section 5.3 further states:

“Broader aerosol-type constraints and more systematic validation over additional regions, seasons, and independent ground-based sites will also be needed before the framework can be evaluated for broader operational application.”

7. Line 195: the continental and urban OPAC models in 6S seem to have been picked because of their names and because they are convenient to use. There is no scientific justification provided. There should be some discussion of why this is a reasonable choice, or if it was because it was

convenient, at least say that. These OPAC models are very old and I'm not sure how reasonable the urban one in particular is. Looking at the Supplement they have fairly similar spectral dependence (and there's no evidence provided that this spectral dependence is common for the BTH region) and are quite absorbing (especially the urban model, with single scattering albedo, SSA, below 0.7 at all wavelengths). As one simple step the authors could look at e.g. AERONET retrievals in this region and see how closely these models fall to observed spectral dependence of AOD and SSA. If they are similar, this is evidence that it is a reasonable choice. If not, then more justification and/or a better model is needed.

Response: *We agree that the original manuscript did not provide sufficient justification for the predefined aerosol models. We have therefore added both their representative optical properties and an observational comparison with AERONET. The revised Supplement now reports the single-scattering albedo and normalized spectral AOD of the predefined continental and urban 6S models. Their regional spectral-AOD representativeness was then examined using 678 AERONET records from four BTH sites spanning 34 observation dates. Both predefined models fall within the observed 5th–95th percentile range of normalized spectral AOD. The continental model is closer to the AERONET median spectral dependence, whereas the urban model has a steeper spectral dependence.*

The continental model is therefore used for offline construction of the static surface prior and LUT-derived reference surface terms, while both continental and urban models are retained as candidates during online retrieval and are selected at the scene scale according to the matching cost.

The added AERONET analysis evaluates the normalized spectral dependence of AOD only; it does not validate the prescribed SSA. This limitation is now stated

explicitly in Sect. 3.2.

Changes in the manuscript:

Section 3.2 now states:

“The spectral AOD dependence prescribed by the two models was compared with AERONET observations from the BTH region. The comparison included 678 records from four sites spanning 34 observation dates and was used to characterize the regional spectral AOD range. Both models fell within the observed 5th–95th percentile range, with the continental model closer to the median spectral dependence and the urban model showing a steeper spectral dependence (Fig. S1). The AERONET comparison evaluates the prescribed spectral AOD dependence only; the predefined SSA values were not independently evaluated in this study.”

Table S5 now provides the SSA and normalized spectral AOD values for the two predefined 6S aerosol models, and Fig. S1 compares their normalized spectral AOD with the AERONET median, interquartile range, and 5th–95th percentile range.

8. Lines 201-203: More information is needed on the gas absorption cross sections and how the correction was done. Spectral responses are not shown, absorption cross sections are not shown, source of gas concentrations is not given. I’m assuming either the assumption of scattering-absorption decoupling was made or the cross-sections and spectral responses were built into 6S and it was handled there. But none of this is stated.

Response: *We have expanded the description of the spectral and gaseous-absorption treatment. The 6S simulations were performed at the nominal WSI central wavelengths and did not explicitly convolve the measured WSI spectral response functions. Gaseous absorption was handled within 6S using the gaseous contents specified by the selected atmospheric profile; no additional external gas-concentration data or absorption cross sections were introduced.*

We have also clarified how the directional global-gas transmittances were combined with the total-scattering transmittances used in the LUT calculation. In addition, we added a sensitivity test in the Supplement to quantify the influence of this gas-transmittance treatment on both simulated TOA reflectance and retrieved AOD.

Changes in the manuscript:

Section 3.2 now states:

“The 6S simulations were performed at the nominal WSI central wavelengths listed in Table 2 without explicit convolution with the measured spectral response functions.”

“Gaseous absorption was treated within 6S using the gaseous contents specified by this profile, without additional external gas-concentration data or absorption cross sections. The downward and upward scattering transmittances stored in the LUT were obtained from the total-scattering transmittances provided by 6S, and the corresponding directional global-gas transmittances were applied to obtain the transmittances used in the TOA reflectance calculation.”

The sensitivity test in Table S7 includes 240 cases. The median absolute AOD change was 0.002, and 97.1 % of the cases showed $|ΔAOD| ≤ 0.02$.

9. Table 2: the angular steps here feel really coarse, is this truly sufficient? Most other approaches use finer angular grids. Do the authors have an estimate of uncertainty introduced by this angular sampling? I would expect some uncertainty from the interpolation step because top of atmosphere (TOA) reflectance often does not vary linearly with angular changes.

Response: *We agree that the relatively coarse angular grid requires an explicit interpolation-error assessment. We therefore added a dedicated sensitivity test for the baseline spring 0.443 μm LUT.*

A total of 160 geometries located between the LUT angular nodes were evaluated by comparing direct 6S simulations with LUT-interpolated TOA reflectance. The

overall interpolation bias was 0.0052, the RMSE was 0.0078, and the 95th percentile absolute error was 0.0185. The error increased toward large viewing zenith angles. These results and the angular dependence of the interpolation error are now reported in Fig. S2 and Table S6.

Changes in the manuscript:

Section 3.2 now states:

“For the baseline spring LUT, the effect of angular discretization was evaluated by comparing direct 6S simulations with LUT-interpolated 0.443 μm TOA reflectance for 160 geometries located between the LUT nodes. The resulting bias and RMSE were 0.0052 and 0.0078, respectively, and the 95th percentile of the absolute difference was 0.0185. The interpolation error increased toward large viewing zenith angles (Fig. S2 and Table S6).”

10. Section 3.4: Just commenting to note that while I understand the intent of the correction approach I am not an expert on the RF method so would defer to other reviewers on whether the chosen method and level of analysis are sufficient for this part.

Response: *Thank you for noting this. Although no specific methodological change was requested, we expanded Sect. 3.4 and Table S2 so that the RF component can be assessed more directly. The revision now reports the training/validation split, hyperparameters, input variables, feature-importance calculation, residual constraints, and the role of MCD19A2 in constructing the training target.*

Changes in the manuscript:

Section 3.4 now provides the RF inputs and clarifies the role of MCD19A2 as follows:

“A random forest (RF) model was used to learn the nonlinear mapping between the blue-band surface reflectance residual and the input features (Breiman, 2001; Wei et al., 2019). The input features included blue-band surface reflectance from the static library, blue-band TOA reflectance, NDVI derived from red and near-infrared TOA reflectance, solar zenith angle, viewing zenith angle, relative azimuth angle, and

elevation. The inclusion of solar and viewing geometry allows the RF to represent geometry-dependent variability in the surface residual empirically. No explicit physical BRDF model is included in the current framework. MCD19A2 AOD is used in the offline derivation of the LUT-derived reference surface term and therefore contributes indirectly to the RF training target. It is not included among the RF input features and is not required during online retrieval.”

The training and validation configuration and the main RF settings are also stated explicitly:

“For the spring configuration, the three reference scenes used for static-prior construction in Sect. 3.3 were also used to generate RF residual-training samples. The scene acquired on 28 April 2025 was reserved for held-out internal validation and was not used for static-prior construction, RF training, feature selection, or hyperparameter tuning. Up to 80,000 pixels were randomly sampled from each training scene, and up to 60,000 pixels were sampled from the validation scene. Following previous studies on random-forest hyperparameters and tuning strategies (Probst et al., 2019), the main RF settings were as follows: 500 trees, maximum depth of 22, minimum leaf size of 40, minimum split size of 80, and maximum feature fraction of 0.7, while other parameters were kept at their default values.”

Table S2 in the Supplement further summarizes the residual-learning configuration, including the target definition, training and validation scenes, input features, sample limits, RF hyperparameters, residual outlier filter, predicted-residual constraint, and corrected-surface-term constraint.

11. Line 302: the cost function used needs to be defined. I am guessing it is either absolute residual TOA reflectance, square residual TOA reflectance, or one of these normalised by some assumed measurement uncertainty? Since the thresholds mentioned later in this paragraph are small numbers my guess is probably not normalised?

Response: *We have now defined the matching cost explicitly. It is the squared*

residual between observed and LUT-simulated 0.443 μm TOA reflectance, minimized over the discrete AOD nodes. The cost is not normalized by measurement uncertainty. We have also clarified how this cost is used for the scene-level aerosol-model selection and added a sensitivity check for the matching-cost threshold.

Changes in the manuscript:

Section 3.4 now defines:

$$(5) J_{i,m} = \min_{\tau \in A} [\rho_{i,0.443}^{\text{obs}} - \rho_{i,0.443,m}^{\text{sim}}(\tau)]^2,$$

and states:

“The matching cost was not normalized by measurement uncertainty. Retrieval was performed separately under the continental and urban aerosol models, and the model with the lower median $J_{i,m}$ over all valid pixels was selected at the scene scale.”

12. Section 4.1: the data volume is really not adequate to make a meaningful assessment of performance. It is probably safe to say that the residual correction approach works better than the than the static LUT, but not much else. This is only 27 points from a single location, and no telling how representative that is of the broader region. The authors acknowledge this is limited but in my view agreeing it is limited is not enough, more data should be processed to get towards a more robust data volume. I would also suggest the authors do the AOD comparison at multiple wavelengths; since the spectral dependence of AOD is fixed in their retrieval approach to one of two models (continental or urban) it would be interesting to see if any one wavelength seems more stable than the others. Scatter plots could also be coloured by the Ångström exponent to see whether that seems to covary with the error of individual points. That would provide a useful indication of whether the

remaining main error source is aerosol spectral dependence as opposed to something else (e.g. surface reflectance, SSA).

Response: *We agree that the original 27 spring-2025 matchups were too few for a robust assessment of retrieval performance. We therefore expanded the independent AERONET evaluation to three periods, giving 104 station–date matchups in total: 27 in spring 2025, 49 in autumn–winter 2025, and 28 in spring 2026.*

For autumn–winter 2025, after excluding RF training and internal-validation scenes, 49 matchups from two AERONET sites spanning 26 independent observation dates were retained. After residual correction, the RMSE decreased from 0.086 to 0.071, while the fraction within the expected-error envelope increased from 67.3 % to 85.7 %.

For spring 2026, the spring-2025 retrieval configuration was applied without retraining or recalibration. The evaluation included 28 matchups from three AERONET sites spanning 17 independent observation dates. The RMSE decreased from 0.330 to 0.153 and the MBE from 0.269 to 0.115 after residual correction.

The reviewer also suggested examining the spectral dependence of the retrieval. In the present framework, AOD is parameterized at 550 nm, while the final matching objective uses the 0.443 μm TOA reflectance with spectral dependence prescribed by the selected 6S aerosol model. Separate wavelength-by-wavelength AOD scatterplots would therefore not constitute independent retrieval tests. Rather than further subdividing the original small matchup set by Ångström exponent, we evaluated the prescribed spectral dependence directly using 678 AERONET records from four BTH sites spanning 34 observation dates. This provides a more direct observational assessment of whether the assumed aerosol spectral dependence is representative of the study region.

We therefore focused the revision on the reviewer’s primary concern—the limited validation volume—while also directly examining the prescribed aerosol spectral dependence. A fully multi-band retrieval analysis would require band-dependent aerosol sensitivity and retrieval errors to be treated independently and is left for

future work.

Changes in the manuscript:

Section 4.1 now reports the autumn–winter 2025 and spring-2026 validation results:

“After excluding scenes used for RF training and held-out internal validation, 49 station–date matchups from two AERONET sites were retained, covering 26 independent observation dates.”

“To examine interannual performance, the spring-2025 retrieval configuration was applied to spring 2026 without using 2026 observations for retraining or recalibration. The evaluation included 28 station–date matchups from three AERONET sites spanning 17 independent observation dates.”

The aerosol spectral-dependence analysis is now provided in Sect. 3.2 and Fig. S1, using 678 AERONET records from four sites spanning 34 observation dates.

Section 5.3 also states:

“The residual-correction concept could also be extended from the 0.443 μm band to additional visible bands if band-dependent aerosol sensitivity and cross-band error characteristics are explicitly considered (Litvinov et al., 2025).”

13. Section 4.4: For me, presenting Figure 9 as a trio of cutouts of municipal boundaries is not so useful without the broader regional context. My suggestion is just to show the whole broader BTH area these three locations all sit in (i.e. including the spaces between them). and add a panel showing a true colour image from this day so the reader had broader context. So this would be 5 larger panels instead of 12. About all I glean from the current image is that the residual-corrected method matches MAIAC better than the static one, but it’s hard to gauge much more.

Response: *We agree that the earlier municipal cutouts did not provide sufficient regional context. We have therefore replaced this presentation with a broader BTH-*

scale comparison and added a DQ-1 WSI true-color panel.

The revised figure now presents the true-color image, static blue-band surface prior, static-prior LUT AOD, residual-corrected AOD, and MCD19A2 AOD over the same broader BTH extent. This allows the spatial structure of the retrievals and the locations of cloud-affected or masked areas to be interpreted within a common regional context.

Changes in the manuscript:

Figure 11 has been revised as:

“Figure 11. Spatial comparison over the broader Beijing–Tianjin–Hebei (BTH) region on 28 May 2025. (a) DQ-1 WSI true-color image; (b) static blue-band surface reflectance used as the surface prior; (c) static-prior LUT AOD; (d) residual-corrected AOD; and (e) MCD19A2 AOD.”

The corresponding text now states:

“Figure 11 presents the DQ-1 WSI true-color image, the static blue-band surface prior, the static-prior LUT AOD, the residual-corrected AOD, and MCD19A2 AOD over the same spatial extent.”

14. Line 620: links and/or data DOIs are needed for this section.

Response: *We have revised the Code and data availability section to provide direct access information for the principal datasets used in this study. The revised text now includes the NASA AERONET data portal, the DOI for the MODIS MAIAC MCD19A2 Version 6.1 product, the CRESDA data portal for DQ-1 WSI observations, and the official CRESDA webpage for DQ-1 radiometric calibration coefficients and spectral response information.*

Changes in the manuscript:

The Code and data availability section now states:

“The code used for QA-mask generation, blue-band surface-residual learning, and LUT-based AOD retrieval is available from the corresponding author upon reasonable request. AERONET Version 3 observations are publicly available from the [NASA AERONET data portal](#). The MODIS MAIAC MCD19A2 Version 6.1 aerosol product

is publicly available from NASA Earthdata under DOI 10.5067/MODIS/MCD19A2.061. DQ-1 WSI data used in this study were obtained from the [China Centre for Resources Satellite Data and Application \(CRESDA\) data portal](#) and remain subject to its data-access policy. The official DQ-1 radiometric calibration coefficients and spectral response information are available from the [CRESDA calibration-data webpage](#). The processed matchup data, validation statistics, and figure data generated in this study are available from the corresponding author upon reasonable request.”