https://doi.org/10.5194/egusphere-2025-6537
Preprint. Discussion started: 26 January 2026

Constructing a Two-Dimensional Ultraviolet Atmospheric Transmittance Field Using Gaussian Process Regression

Tao Jiang[1,2], Xiaohu Yang[1],Yu Huang[1],Zhanfeng Li[1], Yue Li[1], Guanrui Li[1,2],Qi Yu[1,2]

1Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun,130033, China
2University of Chinese Academy of Sciences, Beijing, 100049, China

Correspondence to: yangxiaohu@ciomp.ac.cn

Abstract. Accurate characterization of slant-path atmospheric transmittance in the 280–400 nm ultraviolet (UV) window is essential for radiometric calibration of ground-based UV observations and link-budget assessments of UV optoelectronic systems. Nighttime measurements, however, typically rely on stellar radiometers that provide only sparse and irregular samples along stellar tracks, which complicates construction of a continuous two-dimensional (zenith–azimuth) transmittance field. Here we retrieve slant-path transmittance from multi-star, multi-channel stellar UV radiometer observations using a stellar Langley-type calibration, and reconstruct a 2-D transmittance field by modeling optical depth with Gaussian process regression (GPR) as a function of zenith angle and azimuth. Predictive uncertainty is quantified and rescaled using cross-validated standardized residuals, and performance is benchmarked against conventional interpolation approaches. The reconstructed fields robustly capture the zenith-angle-dominated gradient, achieving cross-validated $R^2 = 0.965–0.984$ and RMSE = 0.004–0.007 across four UV channels. After calibration, 67%–81% of standardized residuals fall within ±1 and 93%–96% within ±2 standard deviations, indicating well-calibrated uncertainty; uncertainty increases in sparsely sampled regions and with local sampling spacing. The proposed framework enables practical construction of nighttime UV transmittance fields with well-characterized uncertainty for observation correction and quantitative assessment of site transparency.

1 Introduction

The "UV window" of the atmosphere in the 280–400 nm band is essential for ground-based UV observations, the design of optoelectronic systems, and radiative-transfer corrections. The total-column transmittance of this band directly governs the UV irradiance reaching the detector (Chubarova et al., 2020; Robert et al., 2018). Existing methods for obtaining atmospheric transmittance are broadly categorized into radiative-transfer model calculations and ground-based observational retrievals. Radiative-transfer models can offer a priori estimates of multi-band transmittance or irradiance based on atmospheric profiles and aerosol scenarios. However, their ability to represent spatiotemporal variability under specific observational conditions is often constrained by uncertainties in input parameters, aerosol scenarios, and profile assumptions (Campanelli et al., 2022; Mateos et al., 2014). Ground-based sun-photometer networks, such as AERONET, have established mature observing systems and aerosol retrieval methods in the visible–near-infrared range. However, their spectral channels primarily cover 340–1640 nm, with limited coverage in the UV window, making them unsuitable for nighttime transmittance variability measurements (Estellés et al., 2012; Giles et al., 2019; Perrone et al., 2021). Lunar photometers can measure AOD and transmittance at night, extending observations into the day–night continuum. However, the changing lunar phase necessitates frequent calibration, and reliance on a single light source limits angular-domain coverage (Román et al., 2020; Uchiyama et al., 2019). In contrast, Román and colleagues utilized all-sky cameras in the visible–near-infrared range to obtain dense angular distributions of sky radiance, facilitating multi-directional retrievals (Graßl et al., 2024; Román et al., 2022; Román et al., 2025).

In the UV range, ozone absorbs strongly in the 250–330 nm range, Rayleigh scattering increases rapidly at shorter wavelengths (Codrean et al., 2025), and aerosol absorption and scattering become more pronounced. These processes, in conjunction with enhanced ozone absorption and molecular scattering, severely limit the available UV radiative flux at the surface. Conventional area-array imaging systems, constrained by detector quantum efficiency and system signal-to-noise ratio, struggle to achieve high-density quantitative sampling over the UV angular domain. As a result, practical applications more frequently rely on point-measurement instruments, such as stellar radiometers, to obtain transmittance samples along discrete lines of sight (Barreto et al., 2019). This creates a key challenge: multi-star samples are typically clustered along star tracks and distributed non-uniformly. As a result, supporting observation correction and engineering applications requires converting these limited discrete samples into a continuous two-dimensional transmittance field on the zenith–azimuth plane through fitting and field-reconstruction methods, accompanied by credible uncertainty characterization.

In recent years, data-driven approaches for atmospheric retrieval and field modeling have advanced rapidly. Notably, Gaussian process regression (GPR), a nonparametric Bayesian method, offers distinct advantages for reconstructing spatial fields from irregular sampling, characterizing multiscale spatial dependence, and quantifying predictive uncertainty (Camps-Valls et al., 2016; Susiluoto et al., 2020). In Earth observation and atmospheric science, GPR has been applied to large-scale spatial interpolation of remote-sensing products, multi-source data fusion, and statistical emulation of complex radiative-transfer models, demonstrating stable mean predictions and reasonable confidence intervals even with high-dimensional inputs and limited samples (Gómez-Dans et al., 2016; Lamminpää et al., 2025; Lubbe et al., 2020). Within the covariance-kernel framework, GPR naturally incorporates anisotropic correlation length scales, enabling scale-dependent dependence to be represented along the zenith- and azimuthal-angle directions; it can therefore exploit sparse line-of-sight observations to capture the spatial continuity of the transmittance field, predict unobserved directions, and provide posterior variance as a measure of uncertainty (Liang et al., 2025; Perry et al., 2025; Yu et al., 2017). However, existing studies have largely focused on solar irradiance forecasting, retrieval of aerosol or pollutant concentration fields, and radiative-transfer modelling in the visible–near-infrared, and these approaches have not yet been widely applied to the two-dimensional angular-domain reconstruction of total-column transmittance in the 280–400 nm UV window. In particular, under limited sightlines and multi-star joint observations, how to use GPR to effectively reconstruct a UV transmittance field with relatively complete directional coverage and reliable uncertainty quantification remains an open gap in current research.

https://doi.org/10.5194/egusphere-2025-6537
Preprint. Discussion started: 26 January 2026

To address the above issues, this study proposes a workflow for constructing and validating a two-dimensional field of nighttime total-column UV transmittance based on multi-star observations, with the main contributions summarized as follows:

1.Using data from a multi-channel stellar UV radiometer, we retrieve total-column transmittance samples in four narrowband channels spanning 280–400 nm via the Langley-plot method, providing highly traceable inputs for two-dimensional angular-domain modelling;

2.We develop a GPR-based spatial dependence model in the optical-depth domain with zenith angle and azimuth as predictors, and we incorporate a multi-star track-envelope constraint to reconstruct a continuous field within the band-covered region while simultaneously producing the corresponding uncertainty field;

3.We scale-calibrate the predictive variance via cross-validation and, through comparisons with inverse-distance weighting and piecewise linear interpolation together with residual diagnostics, we systematically evaluate reconstruction accuracy, stability, and uncertainty consistency, thereby delineating the applicability and confidence of the results.

2 Methods

To construct a continuous ultraviolet transmittance field on the zenith–azimuth space from discrete line-of-sight samples, we introduce a Gaussian process regression model in the optical-depth τ domain. For each observation, the zenith angle and azimuth are denoted by z and A , respectively; after applying the Langley-plot calibration (Michalsky and Mcconville, 2024), the total-column slant-path transmittance $T(z,A)$ and the corresponding optical depth $\tau(z,A)$ are obtained. Compared with modelling directly in the transmittance T domain, modelling in the τ domain is more consistent with the physical processes of atmospheric radiative transfer. On the one hand, τ is obtained by integrating the extinction coefficients along the line of sight, and the total optical depth can be expressed as a linear superposition of components such as Rayleigh scattering, ozone absorption, and aerosol extinction; this representation is widely used in standard radiative-transfer models and optical-depth retrievals and is therefore more suitable as a “linear” state variable for modelling spatial correlations (Emde et al., 2016). Extensive statistical analyses based on observations such as AERONET indicate that the logarithm of aerosol optical depth and related optical-depth quantities is closer to a Gaussian distribution; therefore, applying a logarithmic transform in the transmittance domain and converting it to optical depth τ helps the errors behave as approximately additive Gaussian noise in the τ domain, which is consistent with the Gaussian prior and Gaussian-noise assumptions commonly adopted in Gaussian process regression (N. T. et al., 2000; Kauppi et al., 2024). Accordingly, we use optical depth as the target variable in GPR to build a spatial correlation model and then map the predictions back to the transmittance domain, enabling errors along different azimuth directions to be assessed on a unified optical-depth scale and yielding both the full-field transmittance and its uncertainty.

We denote the input vector of each observation as $\boldsymbol{x}_{\mathrm{i}} = (z_{\mathrm{i}}, A_{\mathrm{i}})^{\mathrm{T}}$, with the corresponding output being the optical depth $y_i = \tau_i, i = 1, \ldots, N$.

We assume that the target field $f(\boldsymbol{x})$ follows a zero-mean Gaussian process:

$$f(\boldsymbol{x}) \sim \mathcal{GP}\big(0, k(\boldsymbol{x}, \boldsymbol{x}')\big), \tag{1}$$

where $k(\boldsymbol{x}, \boldsymbol{x}')$ denotes the covariance kernel defined over the angular domain; the observation noise is modelled as additive independent Gaussian noise:

$$y_i = f(\boldsymbol{x}_i) + \varepsilon_i, \varepsilon_i \sim \mathcal{N}(0, \sigma_n^2), \tag{2}$$

To improve numerical stability and ensure consistency across different input units, we first standardize the input variables (z, A) by subtracting the sample mean and dividing by the sample standard deviation. For the kernel design, to capture both large-scale smooth structure and local variability while allowing different correlation length scales along the zenith and azimuth directions, we employ a sum of a Matérn kernel and a radial basis function (RBF) kernel, with an explicit white-noise kernel term:

$$k(\boldsymbol{x}, \boldsymbol{x}') = k_{Matern}(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\vartheta}_1, \nu) + \sigma_c^2 k_{RBF}(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\vartheta}_2) + \sigma_w^2 \delta_{\boldsymbol{x}\boldsymbol{x}'}, \tag{3}$$

Here, the Matérn smoothness parameter ν, the anisotropic length-scale vector $\boldsymbol{\vartheta}_1$, $\boldsymbol{\vartheta}_2$, as well as the amplitude coefficients σ_c^2 and the noise level σ_w^2 are treated as hyperparameters and are jointly estimated by maximizing the log marginal likelihood. To mitigate the influence of local optima, we apply multiple random restarts in the hyperparameter optimization.

Let the training sample set be $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]^{\mathrm{T}}$ with corresponding optical-depth observations $\boldsymbol{y} = [y_1, \ldots, y_N]^{\mathrm{T}}$ .According to Gaussian process regression theory, the covariance matrix of the training set can be written as

$$\boldsymbol{K} = \boldsymbol{K}_f + \sigma_n^2 \boldsymbol{I}, \tag{4}$$

where $\left[\boldsymbol{K}_f\right]_{ij} = k\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)$. Given the hyperparameters, the log marginal likelihood is

$$\log p(\boldsymbol{y}|\boldsymbol{X}) = -\tfrac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{K}^{-1}\boldsymbol{y} - \tfrac{1}{2}\log|\boldsymbol{K}| - \tfrac{N}{2}\log 2\pi, \tag{5}$$

Numerical optimization yields a set of hyperparameter estimates that maximize this likelihood. For an arbitrary prediction direction $\boldsymbol{x}_*$, we define the covariance vector with the training samples as

$$\boldsymbol{k}_* = [k(\boldsymbol{x}_*, \boldsymbol{x}_1), \ldots, k(\boldsymbol{x}_*, \boldsymbol{x}_N)]^{\mathrm{T}}, \tag{6}$$

and the self-covariance $k_{**} = k\left(\boldsymbol{x}_*, \boldsymbol{x}_*\right)$ .Then the posterior distribution of τ_* is a one-dimensional Gaussian, with predictive mean and variance given by

$$\mu_\tau(\boldsymbol{x}_*) = k_*^{\mathrm{T}} \boldsymbol{K}^{-1} \boldsymbol{y} \,, \tag{7}$$

$$\sigma_\tau^2(\boldsymbol{x}_*) = k_{**} - \boldsymbol{k}_*^{\mathrm{T}} \boldsymbol{K}^{-1} \boldsymbol{k}_* \,, \tag{8}$$

After obtaining μ_τ and σ_τ , we map them back to the transmittance domain as

$$T_*\left(x_*\right) = \exp\left(-m\left(z_*\right) \mu_\tau\left(x_*\right)\right), \tag{9}$$

$$\sigma_T\left(\boldsymbol{x}_*\right) \approx m\left(z_*\right) T_*\left(\boldsymbol{x}_*\right) \sigma_\tau\left(\boldsymbol{x}_*\right), \tag{10}$$

where the transmittance uncertainty is estimated using first-order error propagation. In the following sections, we take the μ_τ corresponding to $T_*\left(z, A\right)$ as the reconstructed two-dimensional UV transmittance field, and use to σ_T describe the angular-domain confidence distribution of the reconstructed field.

Moreover, to improve the statistical reliability of the predictive uncertainty, we further calibrate the GPR output variance using cross-validation optical-depth residuals (Acharki et al., 2023; Gneiting et al., 2007; Jimenez and Katzfuss, 2022). Specifically, we construct standardized residuals from the cross-validated residuals $\Delta\tau^{CV}$ and the corresponding predictive standard deviations σ_τ^{CV} as

$$r_\tau^{CV} = \Delta\tau^{CV} / \sigma_\tau^{CV} \quad , \tag{11}$$

We compute its empirical standard deviation $\hat{\sigma}_r$ and rescale all predictive standard deviations by an inflation factor $\hat{\sigma}_r$ .After calibration, the variance of the standardized residuals becomes closer to 1 and the coverage is more consistent with the theoretical values of the standard normal distribution, thereby increasing confidence in the transmittance uncertainty field.

3 Data Acquisition

The UV transmittance data used in this study were obtained with an in-house stellar UV radiometer (Zhao et al., 2025), whose optical layout is shown in Fig. 1. The instrument adopts a coaxial reflective telescope and a dual-channel architecture, in which the imaging channel is used for star acquisition imaging and closed-loop guiding control, while the UV detection channel performs photon-counting measurements of stellar UV radiation using a narrowband filter wheel and a high-sensitivity photomultiplier tube (PMT).

The filter wheel is equipped with four narrowband channels covering the 280–400 nm UV window: U1 (370–400 nm), U2 (340–370 nm), U3 (300–340 nm), and U4 (288–314 nm), with bandwidths of approximately 20–30 nm. From the longest to the shortest wavelengths, the four channels exhibit progressively higher sensitivity to ozone absorption and Rayleigh scattering, enabling characterization of the wavelength dependence of total-column atmospheric transmittance.

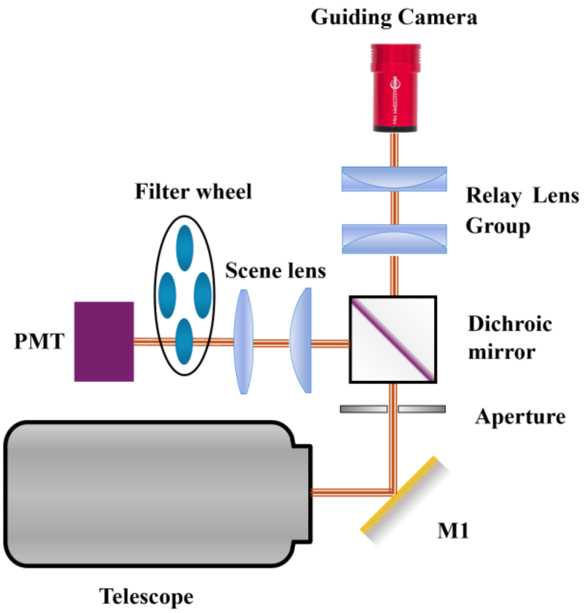

Figure 1: Instrument Optical Path Diagram.

Field observations were conducted in the Changchun region in October 2025 at a site located at 43°50′56″ N, 125°24′01″ E, with an elevation of approximately 200 m. Four moderately bright stars distributed across different right ascensions and declinations were selected as target sources, and their basic parameters are listed in Table 1. Using the site coordinates and observation times, the stars' right ascension and declination can be converted to zenith angle z and azimuth A , thereby obtaining the apparent motion tracks of each star during the observation period.

Table 1 Summary of the target stars used for UV transmittance measurements, including their right ascension, declination, magnitude, and spectral type.

Star name	Right ascension	Declination	Magnitude	Spectral type
HD12533	2h05m31.13s	42d27m19.7s	2.15	K3II+B9.5V+A0V
HD13161	2h11m06.72s	37d59m00.3s	3.00	A5III
HD11443	1h54m35.05s	29d42m24.4s	3.40	F5III
HD13041	2h10m04.73s	37d59m00.3s	4.75	A5IV-V

At each time step, once the imaging channel completes closed-loop tracking, the UV detection channel sequentially switches among the four narrowband filters (U1–U4) to sample photon counts from the target star. After removing samples affected by cloud occlusion, guiding failures, and obvious count outliers, a total of N=98 valid samples were retained, with identical sample sizes across channels U1–U4. Based on the calibration results, the total-column slant-path transmittance $T\left(z, A; \lambda_k\right)$ $\left(k=1,\ldots,4\right)$ for the four channels can be retrieved along each line of sight. The resulting observations can be represented as a set of discrete samples $\left\{z_i, A_i; \mathrm{T}_i^k\right\}$, $i=1,\ldots,N$, $(k=1,\ldots,4)$, where N denotes the total number of valid observations of all target stars over the entire campaign. In the subsequent two-dimensional field reconstruction, each channel is modelled and analysed independently.

4 Analysis and Discussion

In this section, using the GPR model and multi-star observations from the Changchun site, we analyze the reconstructed two-dimensional UV transmittance fields, the associated uncertainty distributions, comparisons with simple interpolation methods, and residual behavior.

4.1 Ultraviolet Transmittance Field Construction Results

To obtain a continuous description of UV transmittance over the angular domain, we project all samples from the four stars during the observation period onto the (z, A) space to form a discrete sample set. Following the GPR formulation in Section 2, we train the model using zenith angle and azimuth as inputs and optical depth as the target variable, learning the spatial dependence structure of the transmittance field under a covariance function formed by the sum of a Matérn kernel and an RBF kernel. To avoid nonphysical extrapolation outside the observed region, we clip the prediction grid using a bounded band constructed from the observation tracks and report the two-dimensional transmittance field only within the multi-star track envelope and its limited expansion, as shown in Fig. 2.

https://doi.org/10.5194/egusphere-2025-6537
Preprint. Discussion started: 26 January 2026

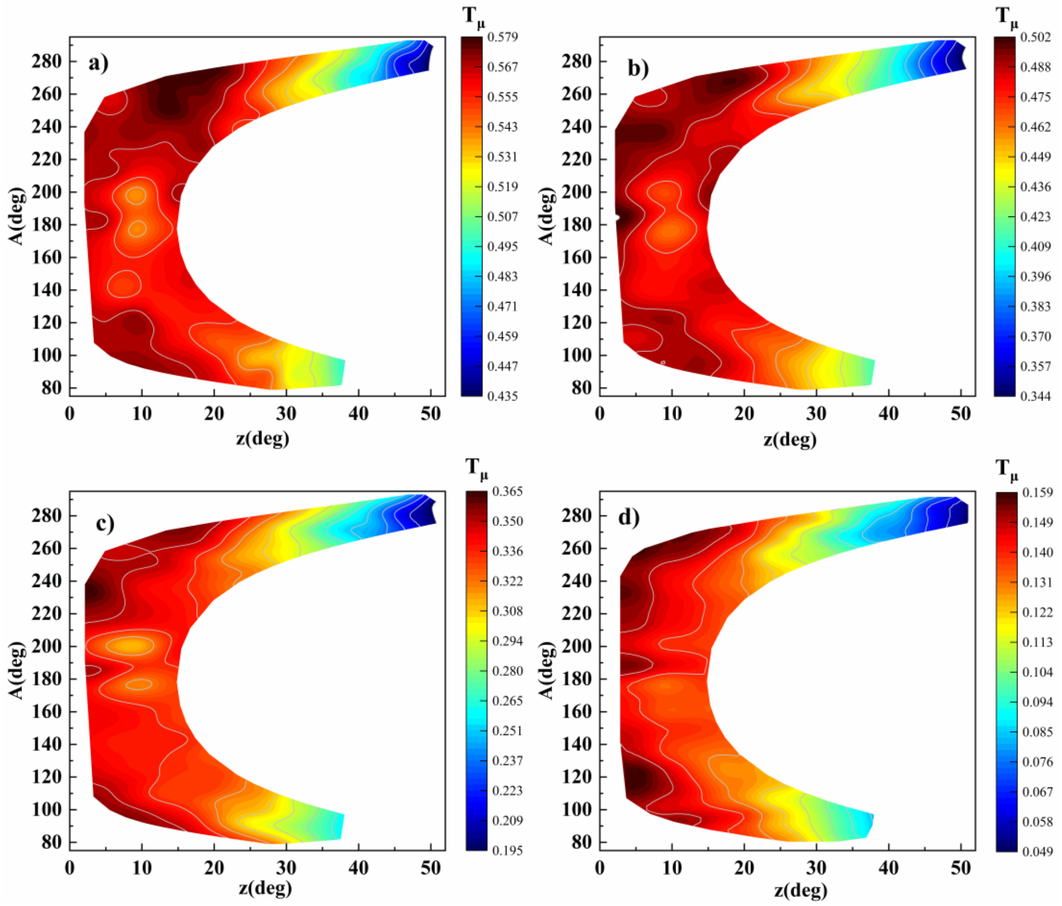

Figure 2: Construction of ultraviolet transmittance using Gaussian Process Regression (GPR). Panels (a)-(d) show the transmittance construction for channels U1–U4, respectively, with the corresponding zenith and azimuth angles.

As can be seen from the spatial patterns in Fig. 2, the transmittance fields in all four channels exhibit a systematic gradient dominated by zenith angle: transmittance is generally higher at small zenith angles ($z < 15°$), and decreases continuously as zenith angle increases (corresponding to longer line-of-sight paths). This trend is consistent with the physical expectation that total-column extinction accumulates along the path, indicating that the model can robustly recover the dominant structural features of the angular-domain field even when line-of-sight samples are sparse and unevenly distributed.

In terms of spectral differences, the shortwave channels (U3 and U4) exhibit a faster decrease in transmittance with increasing zenith angle than the longwave channels (U1 and U2) and respond more sensitively at large zenith angles, leading to stronger gradients and more pronounced morphological contrasts in the large-zenith-angle region. This behavior is consistent with enhanced ozone absorption, stronger Rayleigh scattering, and more pronounced aerosol UV extinction at shorter wavelengths, and it also highlights the necessity of multi-channel observations for characterizing wavelength-dependent extinction responses.

Beyond the systematic variation along zenith angle, Fig. 2 also shows localized fluctuations and heterogeneous structures in certain azimuth sectors. For example, at a similar small zenith angle (≈10°), an overall reduction in transmittance appears in the azimuth range of 170°–210°. This feature may be associated with localized atmospheric perturbations during the observations (e.g., thin high clouds aloft or locally enhanced aerosols); moreover, under the anisotropic length-scale setting and multi-kernel constraints of GPR, such local variations are not overly smoothed, suggesting that the atmospheric transmittance exhibited a degree of azimuthal heterogeneity over the observation period.

To further quantify the credibility of the reconstruction in different regions, the next section analyzes the results in conjunction with the spatial distribution of uncertainty.

4.2 Uncertainty Field and Reliability Analysis

Within the GPR framework, the model outputs both the predictive mean and the posterior standard deviation in the optical-depth domain. To facilitate engineering applications and observational correction, we propagate uncertainty from the τ domain back to the T domain and calibrate it using a variance inflation factor $\hat{\sigma}_r$ derived from cross-validated (CV) standardized residuals, thereby obtaining calibrated transmittance uncertainties within the band-covered region. Figure 3 presents the spatial distribution of σ_T for the four channels.

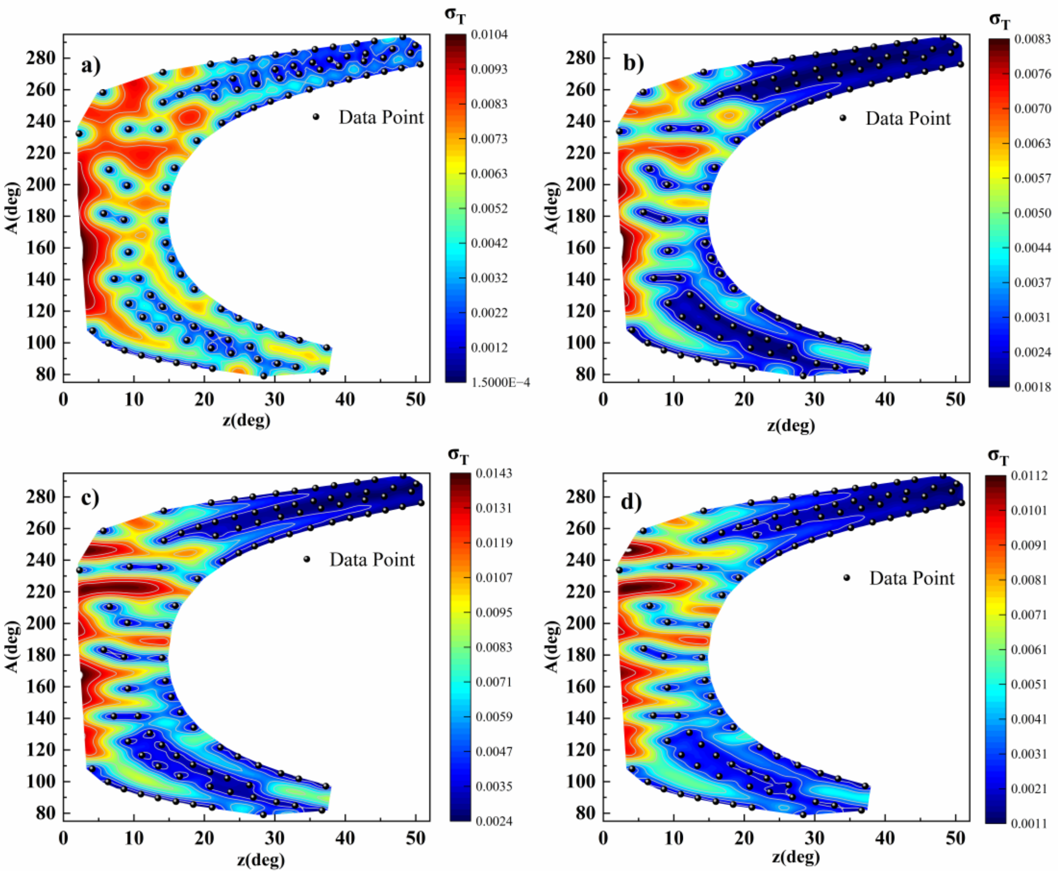

Figure 3: Prediction uncertainty and measured data distribution for the four UV channels. Panels (a)-(d) show the uncertainties for channels U1–U4, respectively, along with the distribution of measured data points.

Summary statistics over the entire reconstructed region show that the median σ_T ranges from 0.0033 to 0.0059, and the median relative uncertainty σ_T / T_μ is 0.69%–2.95%. The relative uncertainty is higher in the shortwave channels, consistent with the physical characteristics that shorter-wavelength UV signals are weaker and more sensitive to path length due to the combined effects of ozone absorption, Rayleigh scattering, and aerosol extinction.

Overall, uncertainty exhibits a structure consistent with sampling density: is small in densely sampled areas near the multi-star tracks, but increases markedly in locally sparse regions. To quantify the phenomenon that “sparser sampling leads to larger uncertainty,” we introduce the “local sampling spacing” $d_3(T)$ in the (z, A) space. For any prediction location , $d_3(T)$ is defined by computing and sorting its distances to all observation points $d_1(T) \le d_2(T) \le d_3(T) \le d_4(T)$..... $d_3(T)$ denotes the third-nearest-neighbour distance. We then analyze the relationship between $d_3(T)$ and σ_T , as shown in Fig. 4. The results indicate a robust positive association between $d_3(T)$ and σ_T: the Spearman correlation coefficient ρ ranges from 0.66 to 0.76 across the four channels (all significant with $p < 0.001$). As implied by Eq. (12), this increase in uncertainty under sparse sampling is consistent with the underlying mechanism of the model.

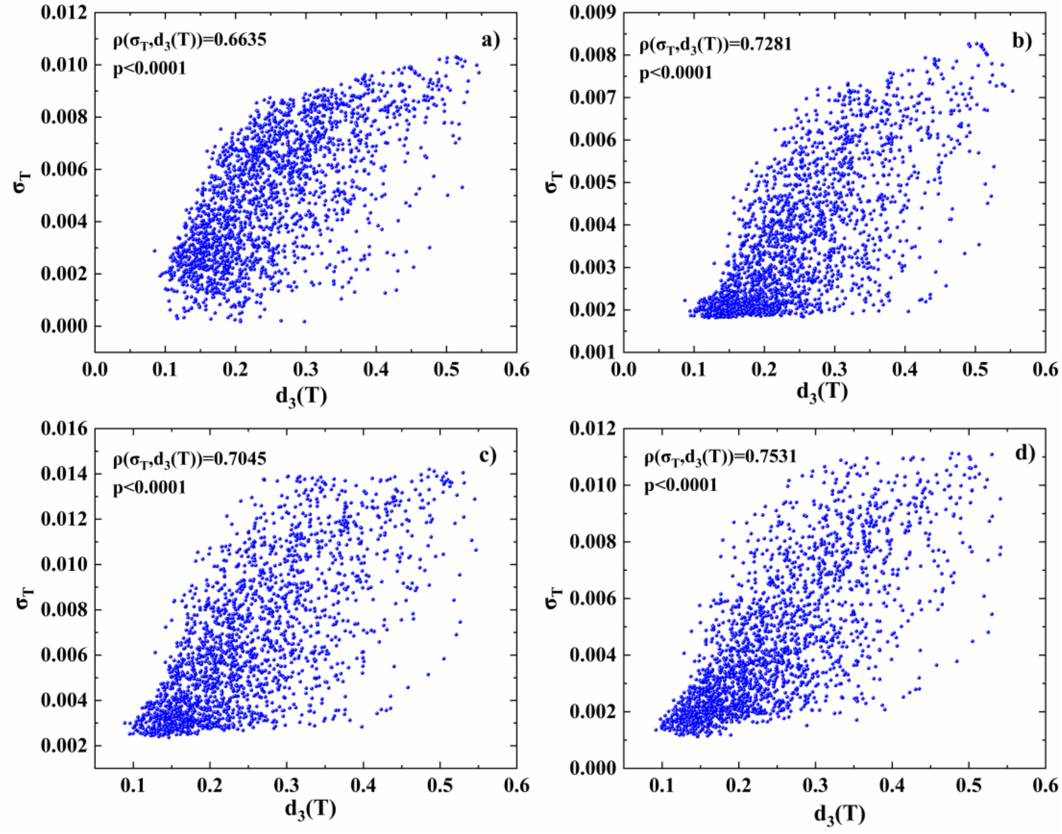

Figure 4. Correlation analysis between local sampling spacing and uncertainty. Panels (a)-(d) correspond to channels U1–U4, respectively, showing the relationship between the third-nearest-neighbor distance and the uncertainty in the transmittance predictions.

4.3 Accuracy Assessment and Comparison with Simple Interpolation Methods

To evaluate the accuracy of GPR for reconstructing two-dimensional UV transmittance fields and its advantages over commonly used interpolation methods, we adopt a cross-validation scheme and compare GPR with piecewise linear interpolation based on Delaunay triangulation (denoted as Linear ND) and inverse distance weighting (IDW, p=2). Table 2 reports the RMSE, MAE, and R^2 metrics for the four channels under the three methods, computed in the τ domain and the T domain, respectively.

Table 2 Comparison of reconstruction accuracy for GPR, piecewise linear interpolation (Linear ND), and inverse distance weighting (IDW) across the four UV channels. The table reports RMSE, MAE, and R2 metrics for both the optical-depth and transmittance domains.

Channel	Method	RMSE_τ	MAE_τ	R^2_τ	RMSE_T	MAE_T	R^2_T

https://doi.org/10.5194/egusphere-2025-6537
Preprint. Discussion started: 26 January 2026

U1	GPR	0.012	0.008	0.605	0.007	0.005	0.965
	Linear ND	0.012	0.008	0.587	0.007	0.005	0.962
	IDW(p=2)	0.013	0.009	0.520	0.007	0.005	0.958
U2	GPR	0.010	0.007	0.689	0.005	0.004	0.984
	Linear ND	0.010	0.007	0.665	0.006	0.004	0.983
	IDW(p=2)	0.014	0.009	0.411	0.007	0.005	0.970
U3	GPR	0.022	0.014	0.472	0.007	0.005	0.968
	Linear ND	0.023	0.014	0.427	0.008	0.005	0.966
	IDW(p=2)	0.023	0.016	0.411	0.008	0.005	0.964
U4	GPR	0.033	0.024	0.485	0.004	0.003	0.979
	Linear ND	0.031	0.021	0.532	0.004	0.003	0.982
	IDW(p=2)	0.033	0.025	0.466	0.004	0.003	0.976

In terms of overall error levels, all three methods achieve RMSE values on the order of 0.004–0.008 in the transmittance domain, with the corresponding R^2 generally exceeding 0.96, indicating that within the observation envelope all methods provide relatively accurate transmittance estimates. Against this background, GPR and Linear ND exhibit similar RMSE in the T domain; compared with IDW, GPR yields lower RMSE across all channels in the T domain, with an average reduction of about 20%–30%, and generally higher R^2, indicating that a statistical model that explicitly incorporates spatial correlation has a clear advantage over simple distance-weighted averaging under limited sightline sampling.

In the optical-depth domain, the RMSE of all three methods increases as the wavelength decreases (U1-U4); GPR attains values of R^2 about 0.4–0.7, generally higher than or comparable to Linear ND and clearly better than IDW. This partly reflects the fact that atmospheric processes are more complex in the short-wavelength UV channels and the dynamic range of τ is larger, which increases modelling difficulty; it also indicates that a GPR model with anisotropic length scales in the τ domain can effectively capture angular-domain correlations. Importantly, beyond achieving T domain errors comparable to or slightly better than Linear ND, GPR additionally provides statistically interpretable uncertainty estimates, which deterministic interpolation methods cannot offer.

4.4 Residual Analysis

Building on the cross-validation in Section 4.3, we assess the reliability of the GPR model for UV transmittance construction by conducting a statistical analysis of standardized residuals r_T^{CV} in the transmittance domain. Figure 5 shows probability histograms of the standardized residuals for each channel, together with their means and standard deviations. The standardized residual distributions for all four channels are unimodal and approximately symmetric, with means close to 0 and standard deviations close to 1, indicating no pronounced systematic bias in the model. Specifically, the empirical coverage for $\left|r_T^{CV}\right| < 1$ is 67.3%–80.6% (close to the nominal 68.3% for a standard normal), for $\left|r_T^{CV}\right| < 2$ is 92.9%–95.9%

(close to the nominal 95.4%), and the fraction of $\left|r_T^{CV}\right| > 3$ is 1.0%–3.1%. These few samples mainly occur in regions with larger uncertainty, consistent with the uncertainty patterns.

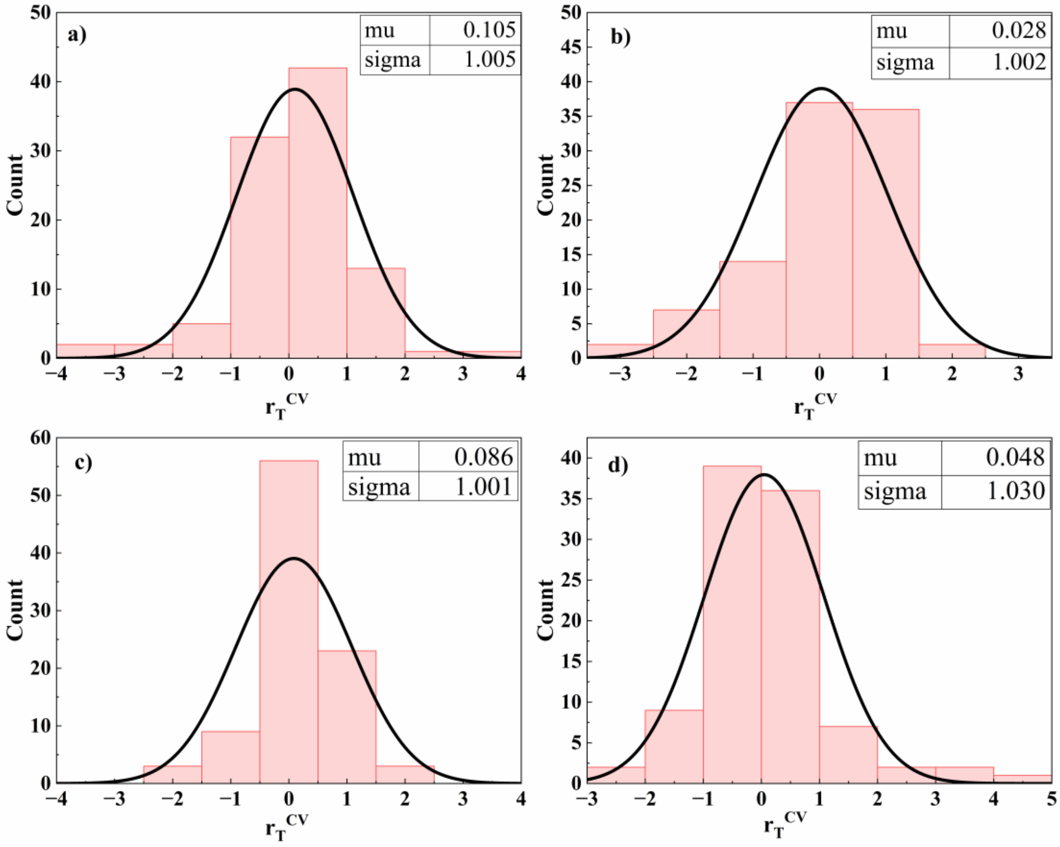

Figure 5: Probability histograms of standardized residuals, illustrating the mean and standard deviation of the distributions. Panels (a)-(d) represent the residual distributions for channels U1–U4, respectively.

Furthermore, Fig. 6 presents how the cross-validation residuals ΔT^{CV} vary with stellar target, zenith angle, and azimuth. The boxplots of residuals grouped by target star show that the absolute mean residual for each stellar direction is below 0.0042, with similar spread across targets, and no evidence of a systematic bias associated with any single star. Scatterplots of residuals against zenith angle and azimuth indicate that, within the main portion of the observed band, residuals form a trend-free cloud with no clear systematic dependence on either zenith angle or azimuth; only in the large-zenith-angle region where samples are sparser does the residual dispersion increase slightly, consistent with the larger uncertainties previously noted in sparsely sampled areas.

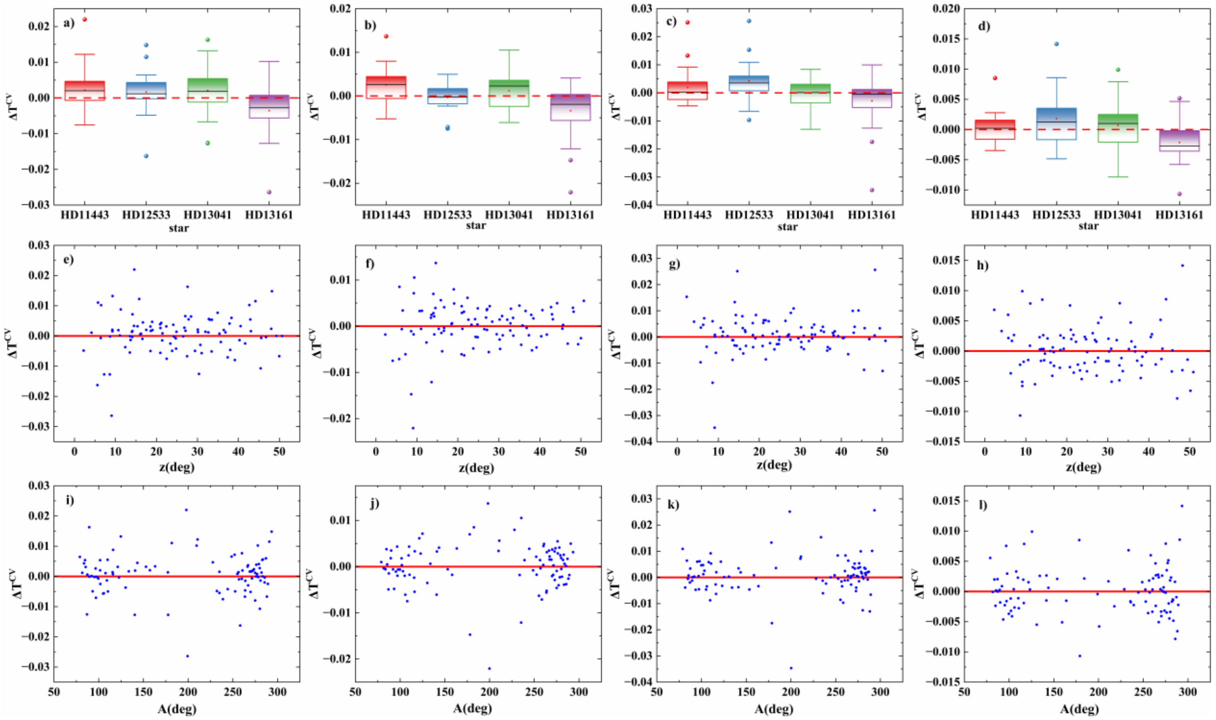

Figure 6: Residual distributions for different variables. Panels (a)-(d) show the relationship between residuals and target stars for channels U1–U4, respectively. Panels (e)-(h) illustrate the relationship between residuals and zenith angle, while panels (i)-(l) show the relationship between residuals and azimuth for each channel.

In summary, the cross-validation residual analysis indicates that the GPR model exhibits no obvious systematic bias within the observation band, and that the errors are dominated by local random fluctuations associated with sampling density and atmospheric complexity, corroborating the preceding error metrics and the spatial patterns of the uncertainty field.

5 Conclusion

This study addresses the two-dimensional construction of atmospheric UV transmittance on the zenith–azimuth plane by developing a Gaussian process regression model in the optical-depth domain and jointly estimating the transmittance field and its uncertainty through a multi-star track-envelope constraint and cross-validation–based variance inflation. The results show that, within the observed band, GPR achieves transmittance-domain reconstruction accuracy superior to or comparable with piecewise linear interpolation and inverse-distance weighting. Furthermore, GPR provides statistically calibrated confidence intervals, with spatial patterns consistent with the line-of-sight sampling density and the underlying atmospheric structure. Analysis of residual variations with stellar target, zenith angle, and azimuth indicates no significant systematic bias. Overall, the proposed framework enables the transformation from limited multi-star observations to a continuous UV transmittance field with reliable error characterization, providing a generalizable technical pathway for UV observation correction and quantitative assessment of site transparency.

Nevertheless, this study has several limitations:

(1) The current results are based on a single site and a single observing campaign, lacking robustness tests at seasonal and interannual scales. Future work will conduct multi-season, multi-year, and multi-site joint observations to evaluate the generalizability of the approach under different meteorological regimes and site conditions;

(2) Only four relatively broad narrowband channels are available, resulting in limited spectral resolution and restricting the separation of underlying physical contributions. We plan to increase the number of channels and narrow the bandwidths, and to incorporate cross-channel joint modelling to better disentangle the contributions from Rayleigh scattering, ozone absorption, and aerosols;

(3) The observing efficiency is currently limited, leading to a relatively small angular coverage of the reconstructed two-dimensional transmittance fields. Future improvements will optimize both hardware and sampling strategies (automation and parallelization, simultaneous multi-star observations, and enhanced pointing and tracking) with the goal of achieving stable fits with near–full-sky coverage.

Code, data, or code and data availability

The dataset analyzed is available at https://github.com/jiangtao185/jiangtao.git for the peer review process.

Author contributions

Tao Jiang led the study design, methodology development and data analysis, and drafted the manuscript. Xiaohu Yang supervised the project and contributed to study design and manuscript revision. Yu Huang supported method validation and result interpretation and revised the manuscript. Zhanfeng Li, Yue Li and Guanrui Li contributed to data acquisition and manuscript revision. Qi Yu provided overall supervision, resources/funding support and manuscript revision. All authors approved the final manuscript.

Competing interests

The authors declare no conflicts of interest.

Acknowledgements

We thank the field staff in Changchun for their technical support and for assisting with the October 2025 observation campaign.

Financial support

This work was supported in part by the National Key R&D Program funding under Grant 2022YFF0705600 and Grant 2022YFF0705603. We thank the field staff in Changchun for their technical support and for assisting with the October 2025 observation campaign.

References

Acharki, N., Bertoncello, A., and Garnier, J.: Robust prediction interval estimation for Gaussian processes by cross-validation method, Comput. Stat. Data Anal., 178, 24, https://doi.org/10.1016/j.csda.2022.107597, 2023.

Barreto, Á., Román, R., Cuevas, E., Pérez-Ramírez, D., Berjón, A., Kouremeti, N., Kazadzis, S., Gröbner, J., Mazzola, M., Toledano, C., Benavent-Oltra, J., Doppler, L., Jurysek, J., Almansa Rodríguez, A., Victori, S., Maupin, F., Guirado-Fuentes, C., González, R., Vitale, V., and Yela, M.: Evaluation of night-time aerosols measurements and lunar irradiance models in the frame of the first multi-instrument nocturnal intercomparison campaign, Atmospheric Environment, 202, https://doi.org/10.1016/j.atmosenv.2019.01.006, 2019.

Campanelli, M., Diémoz, H., Siani, A. M., di Sarra, A., Iannarelli, A. M., Kudo, R., Fasano, G., Casasanta, G., Tofful, L., Cacciani, M., Sanò, P., and Dietrich, S.: Aerosol optical characteristics in the urban area of Rome, Italy, and their impact on the UV index, Atmos. Meas. Tech., 15, 1171-1183, https://doi.org/10.5194/amt-15-1171-2022, 2022.

Camps-Valls, G., Verrelst, J., Munoz-Mari, J., Laparra, V., Mateo-Jimenez, F., and Gomez-Dans, J.: A Survey on Gaussian Processes for Earth-Observation Data Analysis: A Comprehensive Investigation, IEEE Geoscience and Remote Sensing Magazine, 4, 58-78, 10.1109/MGRS.2015.2510084, 2016.

Chubarova, N. E., Pastukhova, A. S., Zhdanova, E. Y., Volpert, E. V., Smyshlyaev, S. P., and Galin, V. Y.: Effects of Ozone and Clouds on Temporal Variability of Surface UV Radiation and UV Resources over Northern Eurasia Derived from Measurements and Modeling, Atmosphere, 11, 59, https://doi.org/10.3390/atmos11010059, 2020.

Codrean, A.-F., Bunoiu, O. M., and Paulescu, M.: Identifying the Signature of the Solar UV Radiation Spectrum, Atmosphere, 16, 427, https://doi.org/10.3390/atmos16040427, 2025.

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T., and Bugliaro, L.: The libRadtran software package for radiative transfer calculations (version 2.0.1), Geosci. Model Dev., 9, 1647-1672, https://doi.org/10.5194/gmd-9-1647-2016, 2016.

Estellés, V., Campanelli, M., Smyth, T. J., Utrillas, M. P., and Martínez-Lozano, J. A.: Evaluation of the new ESR network software for the retrieval of direct sun products from CIMEL CE318 and PREDE POM01 sun-sky radiometers, Atmos. Chem. Phys., 12, 11619-11630, https://doi.org/10.5194/acp-12-11619-2012, 2012.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, Atmos. Meas. Tech., 12, 169-209, 10.5194/amt-12-169-2019, 2019.

Gneiting, T., Balabdaoui, F., and Raftery, A. E.: Probabilistic Forecasts, Calibration and Sharpness, Journal of the Royal Statistical Society Series B: Statistical Methodology, 69, 243-268, https://doi.org/10.1111/j.1467-9868.2007.00587.x, 2007.

Gómez-Dans, J. L., Lewis, P. E., and Disney, M.: Efficient Emulation of Radiative Transfer Codes Using Gaussian Processes and Application to Land Surface Parameter Inferences, Remote Sensing, 8, 119, https://doi.org/10.3390/rs8020119, 2016.

Graßl, S., Ritter, C., Wilsch, J., Herrmann, R., Doppler, L., and Román, R.: From Polar Day to Polar Night: A Comprehensive Sun and Star Photometer Study of Trends in Arctic Aerosol Properties in Ny-Ålesund, Svalbard, Remote Sensing, 16, 3725, https://doi.org/10.3390/rs16193725, 2024.

Jimenez, F. and Katzfuss, M.: Scalable Bayesian Optimization Using Vecchia Approximations of Gaussian Processes, arXiv e-prints, https://doi.org/10.48550/arXiv.2203.01459, 2022.

Kauppi, A., Kukkurainen, A., Lipponen, A., Laine, M., Arola, A., Lindqvist, H., and Tamminen, J.: A Bayesian Framework to Quantify Uncertainty in Aerosol Optical Model Selection Applied to TROPOMI Measurements, Remote Sensing, 16, 1945, https://doi.org/10.3390/rs16111945, 2024.

Lamminpää, O., Susiluoto, J., Hobbs, J., McDuffie, J., Braverman, A., and Owhadi, H.: Forward model emulator for atmospheric radiative transfer using Gaussian processes and cross validation, Atmos. Meas. Tech., 18, 673-694, https://doi.org/10.5194/amt-18-673-2025, 2025.

Liang, W., Zhang, Y., Chong, A., Cochran Hameen, E., and Loftness, V.: Exploring Gaussian Process Regression for indoor environmental quality: Spatiotemporal thermal and air quality modeling with mobile sensing, Building and Environment, 281, 113143, https://doi.org/10.1016/j.buildenv.2025.113143, 2025.

Lubbe, F., Maritz, J., and Harms, T.: Evaluating the Potential of Gaussian Process Regression for Solar Radiation Forecasting: A Case Study, Energies, 13, 5509, https://doi.org/10.3390/en13205509, 2020.

Mateos, D., Antón, M., Toledano, C., Cachorro, V. E., Alados-Arboledas, L., Sorribas, M., Costa, M. J., and Baldasano, J. M.: Aerosol radiative effects in the ultraviolet, visible, and near-infrared spectral ranges using long-term aerosol data series over the Iberian Peninsula, Atmospheric Chemistry and Physics, 14, 13497-13514, https://doi.org/10.5194/acp-14-13497-2014, 2014.

Michalsky, J. and McConville, G.: Ozone and aerosol optical depth retrievals using the ultraviolet multi-filter rotating shadow-band radiometer, Atmos. Meas. Tech., 17, 1017-1022, https://doi.org/10.5194/amt-17-1017-2024, 2024.

N. T., O. N., Ignatov, A., Holben, B. N., and Eck, T. F.: The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET Sunphotometer data, Geophysical Research Letters, 27, 3333-3336, https://doi.org/10.1029/2000GL011581, 2000.

Perrone, M. R., Lorusso, A., and Romano, S.: Diurnal and nocturnal aerosol properties by AERONET sun-sky-lunar photometer measurements along four years - ScienceDirect, Atmospheric Research, 265, https://doi.org/10.1016/j.atmosres.2021.105889, 2021.

Perry, N., Pedersen, P. P., Christensen, C. N., Nussli, E., Heinonen, S., Dagallier, L. G., Jacquat, R., Horstmann, S., and Franck, C.: Detecting urban PM2.5 hotspots with mobile sensing and Gaussian process regression, npj Clean Air, 1, 38, https://doi.org/10.48550/arXiv.2509.17175, 2025.

Robert, F., Didier, R., Emmanuel, B., Dominique, J., Mathieu, C., Jing, T., Heather, B., Thomas, J., Bryan, F., and Trevor, P.: Satellite Radiation Products for Ocean Biology and Biogeochemistry: Needs, State-of-the-Art, Gaps, Development Priorities, and Opportunities, Frontiers in Marine ence, 5, 3-, https://doi.org/10.3389/fmars.2018.00003, 2018.

Román, R., González, R., Toledano, C., Barreto, Á., Pérez-Ramírez, D., Benavent-Oltra, J. A., Olmo, F. J., Cachorro, V. E., Alados-Arboledas, L., and de Frutos, Á. M.: Correction of a lunar-irradiance model for aerosol optical depth retrieval and comparison with a star photometer, Atmos. Meas. Tech., 13, 6293-6310, https://doi.org/10.5194/amt-13-6293-2020, 2020.

Román, R., Antuña-Sánchez, J. C., Cachorro, V. E., Toledano, C., Torres, B., Mateos, D., Fuertes, D., López, C., González, R., Lapionok, T., Herreras-Giralda, M., Dubovik, O., and de Frutos, Á. M.: Retrieval of aerosol properties using relative radiance measurements from an all-sky camera, Atmos. Meas. Tech., 15, 407-433, https://doi.org/10.5194/amt-15-407-2022, 2022.

Román, R., González-Fernández, D., Antuña-Sánchez, J. C., Herrero del Barrio, C., Herrero-Anta, S., Barreto, Á., Cachorro, V. E., Doppler, L., González, R., Ritter, C., Mateos, D., Kouremeti, N., Copes, G., Calle, A., Granados-Muñoz, M. J., Toledano, C., and de Frutos, Á. M.: Star photometry with all-sky cameras to retrieve aerosol optical depth at nighttime, Atmos. Meas. Tech., 18, 2847-2875, https://doi.org/10.5194/amt-18-2847-2025, 2025.

Susiluoto, J., Spantini, A., Haario, H., Härkönen, T., and Marzouk, Y.: Efficient multi-scale Gaussian process regression for massive remote sensing data with satGP v0.1.2, Geosci. Model Dev., 13, 3439-3463, https://doi.org/10.5194/gmd-13-3439-2020, 2020.

Uchiyama, A., Shiobara, M., Kobayashi, H., Matsunaga, T., Yamazaki, A., Inei, K., Kawai, K., and Watanabe, Y.: Nocturnal aerosol optical depth measurements with modified sky radiometer POM-02 using the moon as a light source, Atmos. Meas. Tech., 12, 6465-6488, https://doi.org/10.5194/amt-12-6465-2019, 2019.

Yu, W., Liu, Y., Ma, Z., and Bi, J.: Improving satellite-based PM2.5 estimates in China using Gaussian processes modeling in a Bayesian hierarchical setting, Scientific Reports, 7, 7048, https://doi.org/10.1038/s41598-017-07478-0, 2017.

Zhao, Z., Yang, X., Huang, Y., Zhang, Z., Li, Z., Wu, H., and Li, G.: Measurement Technology of Total Atmospheric Transmittance of Ultraviolet at Night, Acta Optica Sinica, 45, http://dx.doi.org/10.3788/AOS250571, 2025.