CAMS Service Evolution

D4.3 Single value radiation uncertainty patterns

Due date of deliverable	29/02/2024
Submission date	29/02/2024
File Name	D4.3 Single value radiation uncertainty patterns
Work Package /Task	Tasks T4.3.3 and T4.3.4
Organisation Responsible of Deliverable	FMI
Author name(s)	Antti Lipponen, Antti Arola
Revision number	1
Status	Issued
Dissemination Level	PUBLIC

Funded by the European Union

The CAMEO project (grant agreement No 101082125) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

In this project, we targeted the enhancement of computational models used for simulating and forecasting global and ultraviolet radiation at the surface level. Recognising the limitations of current models, our primary objective was to identify the factors contributing to inaccuracies in radiation data, leveraging advanced data-driven methodologies.

Central to our approach was the development and implementation of a post-process correction method, employing machine learning techniques, notably the XGBoost algorithm, to correct for the biases within the global horizontal irradiation (GHI) and ultraviolet index (UVI) data from the CAMS service. The integration of SHapley Additive exPlanations (SHAP) analysis with this model provided deep insights into the influence of individual variables on the predictions, enhancing the interpretability and accuracy of the model.

A significant achievement of our work was the notable improvement in the precision of radiation data, as demonstrated by the reduced error standard deviation in post-process corrected data across various radiation intensities. Furthermore, we introduced a data-driven approach for uncertainty quantification, employing an ensemble technique that quantified the radiation data's reliability and revealed patterns in uncertainty distribution.

The outcomes of this project mark an advancement in the field of radiation simulation and forecasting. The refined models and methodologies developed promise improved accuracy and reliability in radiation data and provide a robust framework for future research and development in this domain.

Table of Contents

2 Introduction

2.1 Background

Monitoring the composition of the atmosphere is a key objective of the European Union’s flagship Space programme Copernicus, with the Copernicus Atmosphere Monitoring Service (CAMS) providing free and continuous data and information on atmospheric composition.
The CAMS Service Evolution (CAMEO) project will enhance the quality and efficiency of the CAMS service and help CAMS to better respond to policy needs such as air pollution and greenhouse gases monitoring, the fulfilment of sustainable development goals, and sustainable and clean energy.

CAMEO will help prepare CAMS for the uptake of forthcoming satellite data, including Sentinel-4, -5 and 3MI, and advance the aerosol and trace gas data assimilation methods and inversion capacity of the global and regional CAMS production systems.
CAMEO will develop methods to provide uncertainty information about CAMS products, in particular for emissions, policy, solar radiation and deposition products in response to prominent requests from current CAMS users.

CAMEO will contribute to the medium- to long-term evolution of the CAMS production systems and products.

The transfer of developments from CAMEO into subsequent improvements of CAMS operational service elements is a main driver for the project and is the main pathway to impact for CAMEO.

The CAMEO consortium, led by ECMWF, the entity entrusted to operate CAMS, includes several CAMS partners thus allowing CAMEO developments to be carried out directly within the CAMS production systems and facilitating the transition of CAMEO results to future upgrades of the CAMS service.

This will maximise the impact and outcomes of CAMEO as it can make full use of the existing CAMS infrastructure for data sharing, data delivery and communication, thus supporting policymakers, business and citizens with enhanced atmospheric environmental information.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

This document provides the results of the analysis related to global horizontal irradiation (GHI) and UV index (UVI) errors and uncertainties in CAMS radiation services and forecast data products aiming at revealing patterns in CAMS radiation product biases and uncertainties.

2.2.2 Work performed in this deliverable

In this deliverable the work as planned in the Description of Action (DoA, WP4 T4.3.3 and T4.3.4) was performed.

2.2.3 Deviations and counter measures

No deviations have been encountered.

2.2.4 CAMEO Project Partners:

ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
Met Norway	METEOROLOGISK INSTITUTT
BSC	BARCELONA SUPERCOMPUTING CENTER-CENTRO NACIONAL DE SUPERCOMPUTACION
KNMI	KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT-KNMi
SMHI	SVERIGES METEOROLOGISKA OCH HYDROLOGISKA INSTITUT
BIRA-IASB	INSTITUT ROYAL D'AERONOMIE SPATIALEDE BELGIQUE
HYGEOS	HYGEOS SARL
FMI	ILMATIETEEN LAITOS
DLR	DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV
ARMINES	ASSOCIATION POUR LA RECHERCHE ET LE DEVELOPPEMENT DES METHODES ET PROCESSUS INDUSTRIELS
CNRS	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS
GRASP-SAS	GENERALIZED RETRIEVAL OF ATMOSPHERE AND SURFACE PROPERTIES EN ABREGE GRASP
CU	UNIVERZITA KARLOVA
CEA	COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
MF	METEO-FRANCE
TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
INERIS	INSTITUT NATIONAL DE L ENVIRONNEMENT INDUSTRIEL ET DES RISQUES - INERIS
IOS-PIB	INSTYTUT OCHRONY SRODOWISKA - PANSTWOWY INSTYTUT BADAWCZY
FZJ	FORSCHUNGSZENTRUM JULICH GMBH
AU	AARHUS UNIVERSITET
ENEA	AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE

3 Radiation datasets

3.1 Global radiation

3.1.1 Observed data

Ground-based global radiation measurements are obtained from 21 stations. The stations and data are from the BSRN, SAURAN, and ENERMA networks. The data has undergone a thorough quality control and we have only selected to use data from stations that have been marked as good or very good quality flag by the quality control procedures to ensure good enough quality data for our analyses. Figure 1 displays the locations of the stations on a map. Table 1 reports the stations, their location coordinates and altitude. Data from full year 2019 is used.

Table 1. List of global radiation stations with latitude, longitude, and elevation information.

Station	Latitude (°)	Longitude (°)	Elevation (m)
BSRN_Bud	47.43	19.18	139
BSRN_Cab	51.97	4.93	0
BSRN_Cnr	42.82	-1.6	471
BSRN_Flo	-27.6	-48.52	11
BSRN_Gob	-23.56	15.04	407
BSRN_Ino	44.34	26.01	110
BSRN_Iza	28.31	-16.5	2373
BSRN_Lin	52.21	14.12	125
BSRN_Pal	48.71	2.21	156
BSRN_Pay	46.81	6.94	491
BSRN_Run	-20.9	55.48	116
BSRN_Son	47.05	12.96	3109
SAURAN_Csir	-25.75	28.28	1400
SAURAN_Cut	-29.12	26.22	1397
SAURAN_Nmu	-34.01	25.67	35
SAURAN_Sun	-33.93	18.87	119
SAURAN_Upr	-25.75	28.23	1410
ENERMENA_Erf	31.49	-4.22	859
ENERMENA_Jor	30.17	35.82	1012
ENERMENA_Tn	32.97	10.48	210
ENERMENA_Zag	30.27	-5.85	783

Table 2. List of CAMS radiation service expert mode quantities used as model inputs in global radiation study.

Variable	Description
albedo	Ground albedo
AODAM	Partial aerosol optical depth at 550 nm for ammonium
AODBC	Partial aerosol optical depth at 550 nm for black carbon
AODDU	Partial aerosol optical depth at 550 nm for dust
AODNI	Partial aerosol optical depth at 550 nm for nitrate
AODOR	Partial aerosol optical depth at 550 nm for organic matter
AODSO	Partial aerosol optical depth at 550 nm for secondary organics
AODSS	Partial aerosol optical depth at 550 nm for sea salt
AODSU	Partial aerosol optical depth at 550 nm for sulphate
Cloud Modification Factor	Global horizontal irradiation / Clear sky global horizontal irradiation
CloudCoverage	Cloud coverage of the pixel
CloudOpticalDepth	Cloud optical depth
CloudType	Cloud type 0=no clouds 5=low-level cloud 6=medium-level cloud 7=high-level cloud 8=thin cloud
fgeo	MODIS-like surface reflectance BRDF parameter fgeo
fiso	MODIS-like surface reflectance BRDF parameter fiso
fvol	MODIS-like surface reflectance BRDF parameter fvol
GHI	Global irradiation on horizontal plane at ground level (Wh/m2)
Snow_probability	Snow probability of the pixel
SummerWinterSplit	Summer/winter split. 1.0 means summer, 0.0 means winter
SZA	Solar zenithal angle for the middle of the summarization (deg)
TCO3	Total column content of ozone (Dobson unit)
TCWV	Total column content of water vapour (kg/m2)
TOA	Irradiation on horizontal plane at the top of atmosphere (Wh/m2)

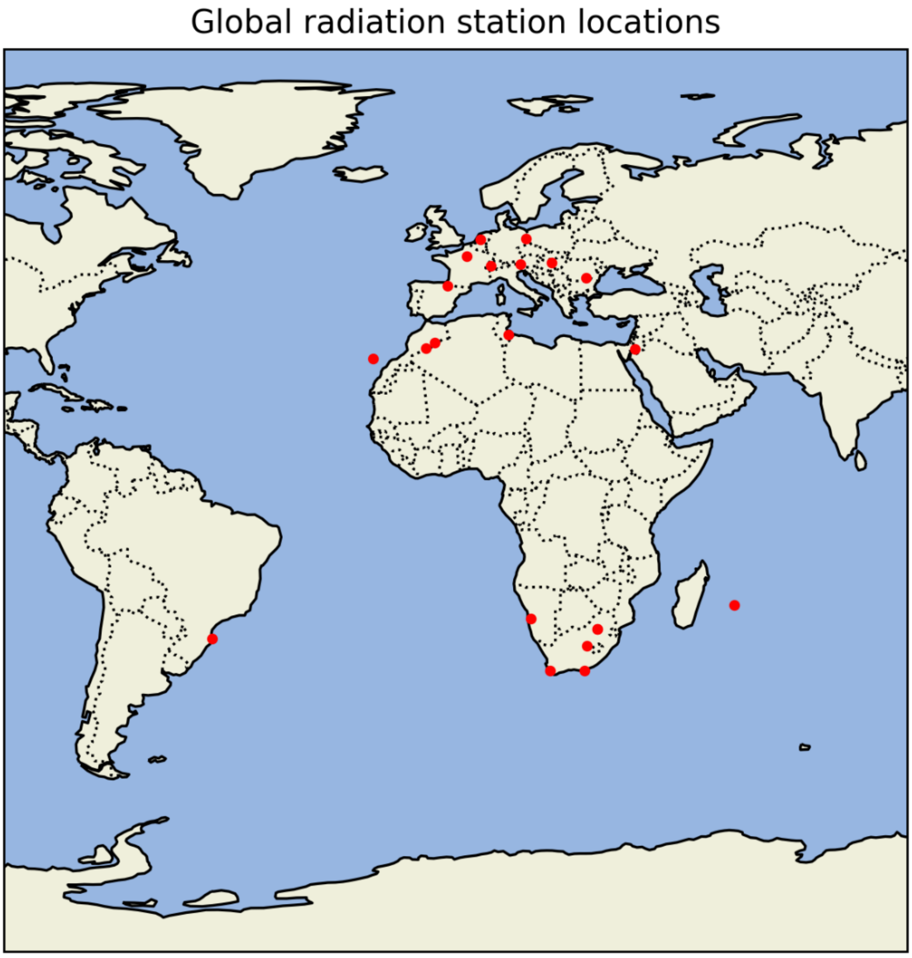

Figure 1. Locations of global radiation stations.

3.1.2 Model data

As the model data, we use the CAMS radiation service expert mode data (Schroedter-Homscheidt et al., 2022, Qu et al., 2017, Lefèvre et al., 2013, Gschwind et al., 2019). The data provides radiation quantities as well as the input parameters used in the radiative transfer modelling. We collocate the model data with the ground-based observations and form our dataset for the analysis. The analysis dataset consists of 285052 data points. The temporal resolution of the data is 1 minute. For the list of parameters used in our analyses, see Table 2.

3.2 Ultraviolet (UV) index

3.2.1 Observed data

Ground-based measurements are obtained from 33 stations. The stations and data are the same as used for the CAMS radiation service UV index validation. The data has undergone quality checks to ensure good quality data. Figure 2 displays the locations of the stations on a map. Table 2 report the stations, their location coordinates and altitude.

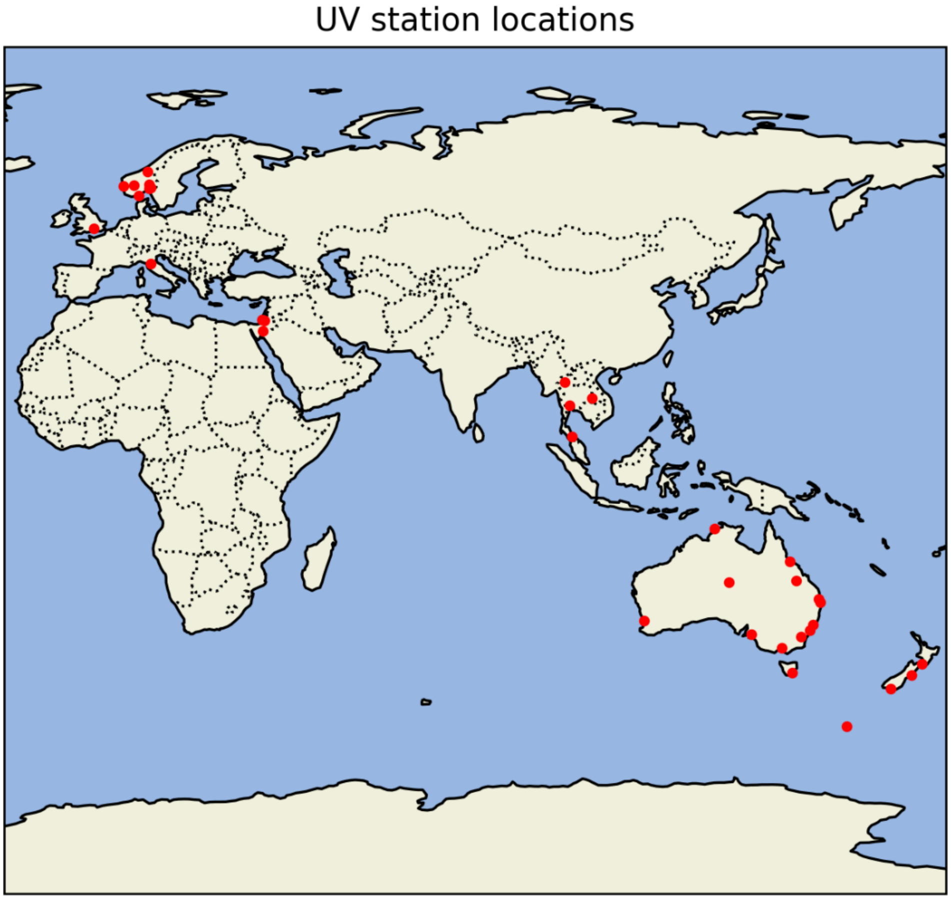

Figure 2. Map of UV station locations.

3.2.1 Model data

As the model data, we use the CAMS global atmospheric composition forecasts UV biologically effective dose. The dose is multiplied by 40 to obtain UV index (McKinlay, 1987). We use the time 00:00 for each day and lead time hours 0-23 to construct our hourly model dataset. We use data from years 2019-2022. The temporal resolution of the data we use is 1 minute. We use linear interpolation to collocate the model data with observations both spatially and temporally. Due to relatively coarse model resolution, both in time and space, we have not carried out any smoothing or averaging of the data in the collocation. As input data for our correction model we use the The same CAMS radiation service expert mode input data as for the global radiation, with the only exception of the global horizontal irradiation was replaced with the CAMS model UV index. The total amount of collocated data points in our dataset is 901796.

Table 3. List of UV stations with latitude, longitude, and elevation information.

Station	Latitude (°)	Longitude (°)	Elevation (m)
Adelaide	-34.92	138.62	10
Alicesprings	-23.8	133.9	550
Bergen	60.38	5.33	40
Betdagan	32	34.81	25
Brisbane	-27.45	153.03	20
Canberra	-35.31	149.2	580
Chiangmai	18.78	98.98	312
Christchurch	-43.53	172.61	6
Darwin	-12.43	130.89	30
Eilat	29.55	34.96	10
Emerald	-23.53	148.16	190
Finse	60.58	7.57	1200
Florence	43.82	11.2	50
Goldcoast	-28	153.37	10
Invercargill	-46.42	168.33	0
Jerusalem	31.78	35.21	700
Kingston	-42.99	147.29	50
Kise	60.78	10.82	140
Kjeller	59.98	11.05	143
Landvik	58.33	8.52	0
Macquarieisland	-54.5	158.94	0
Melbourne	-37.73	145.1	60
Nakhonpathom	13.82	100.04	72
Newcastle	-32.9	151.72	20
Osteras	59.92	10.75	150
Perth	-31.92	115.96	15
Reading	51.44	-0.94	66
Songkhla	7.2	100.6	15
Sydney	-34.04	151.1	20
Townsville	-19.33	146.76	10
Trondheim	63.42	10.38	70
Ubonratchathani	15.25	104.87	120
Wellington	-41.24	174.92	14

4 Methods

In this work, we have applied the post-process correction approach to CAMS surface global radiation and ultraviolet index data to minimize the biases in these data. The post-process correction models were trained using datasets consisting of various atmosphere and radiation variables and co-located model and observed radiation data. Following the post-process bias-correction, we quantified the uncertainty in the radiation data using an ensemble technique and carefully analysed the factors affecting the biases and uncertainties in the CAMS radiation data. The methods used in our work are presented below.

4.1 Post-process correction

In machine-learning-based post-process correction approach, we correct for the simulation model error using a post-processing step. In the conventional, fully learned machine learning approach, a machine learning model is usually trained to directly predict the simulation model output. In our post-process correction, we take different approach: we train a machine learning algorithm (e.g deep neural network or tree-based regression algorithm) to predict the error in the simulation model output and use the error estimate to correct for the simulation results. This way we can combine the information both from the physics-based simulation model and data-driven machine learning. The post-process correction approach has been found to perform more stable and produce more accurate results than the conventional fully learned machine learning approach, for example, in generation of surrogate simulation models (Lipponen et al., 2018) and in medical imaging, see for example Hamilton et al. (2019). Intuitively, the reason why our approach can be expected to improve over the conventional machine learning approach is that the model error is often less complicated function for machine learning regression than the physical process corresponding to the simulation model outputs. The key advantages of our approach over the conventional approaches are 1) improved accuracy, 2) possibility to post-correct existing (past) simulation-based datasets with no need to run the full simulations again, and possibly 3) need of less training data than with the conventional machine learning approaches. It should be noted that in case of updates to the original simulation model, the post-process correction model needs to be re-trained to correspond to the actual simulation model.

For more technical description of the post-process correction approach and more use cases see, for example, Lipponen et al. (2021), Taskinen et al. (2022) or Lipponen et al. (2022).

In this work, we will use an ensemble technique for the post-process correction by training an ensemble of post-process correction models consisting of 10 models trained with different random initializations. This provides us means to estimate and evaluate the uncertainties in the corrected simulation model outputs using the spread of the outputs of the ensemble members. For more technical description of the ensemble uncertainty modelling of the machine learning models see, for example, Lakshminarayanan et al. (2017) or Fort et al. (2019).

4.2 SHapley Additive exPlanations (SHAP) analysis

SHapley Additive exPlanations (SHAP) analysis is a method derived from cooperative game theory and is primarily used to interpret complex machine learning models. It assigns each feature or input variable an importance value for a particular prediction, offering a detailed insight into how each feature contributes, positively or negatively, to the target outcome.

Imagine a team of factors working together to predict the amount of radiation that reaches the Earth surface. Each factor, such as cloud cover, aerosol levels, or surface reflectance, plays a role like a team member in a cooperative game, where the prediction of radiation level is the total payout. SHAP values explain the payout (prediction) by fairly distributing the credit among all the contributing factors (features).

For a more technical description of the SHAP analysis see, for example, Lundberg and Lee (2017).

Example Let's consider a practical example in the context of our radiation simulation models:
1. **Cloud Cover**: Suppose the model predicts a decrease in surface radiation levels. SHAP analysis could reveal that dense cloud cover is the most influential factor in this prediction, assigning it a high negative SHAP value. This indicates that the presence of clouds is significantly blocking the sunlight, leading to a lower simulation of radiation at the surface.
2. **Aerosol Information**: Aerosols can either absorb or reflect sunlight. SHAP analysis might attribute a positive value to aerosol levels if the model predicts higher radiation at the surface, suggesting that in this specific case, aerosols are reflecting more sunlight back into space and thus reducing the amount of radiation reaching the ground.
3. **Surface Reflectance**: Different surfaces (water, forest, urban areas) reflect and absorb sunlight differently. SHAP can provide insights into how much the surface type is contributing to the prediction. For instance, a high positive SHAP value for snow-covered areas would indicate that the high reflectivity of snow is contributing significantly to the prediction of lower radiation absorption at the surface level.

In these examples, SHAP analysis not only quantifies the impact of each factor on the radiation levels but also offers a nuanced understanding of how different conditions interact to affect the model's predictions. This insight is invaluable for identifying and addressing the sources of error in radiation simulations, ultimately leading to more accurate and reliable forecasts.

In our work, we applied the SHAP analysis to the post-process error correction model. Therefore, in our case, the SHAP analysis did not reveal the factors affecting the solar radiation quantities, that are relatively well known, but the factors that affect the errors in CAMS solar radiation quantities. This way our analysis will help revealing the strengths and weaknesses in the CAMS radiation models and in the future help improving the models.

4.3 Machine-learning-based regression model eXtreme Gradient Boosting (XGBoost)

In our work, we chose to use the XGBoost regression model, a decision driven by its robustness, efficiency, and accuracy in handling complex prediction tasks (Chen and Guestrin, 2016). XGBoost stands for eXtreme Gradient Boosting, an advanced implementation of gradient boosted decision trees designed for speed and performance. This model operates by constructing an ensemble of trees, where each tree is built in a sequential manner, learning and improving from the mistakes of its predecessors. Essentially, each new tree attempts to correct the errors made by the ensemble of previously built trees. This iterative approach enables the model to focus on and learn from the most challenging cases in the dataset.

The XGBoost model has capacity to handle a variety of data types, missing data, and various non-linear feature relationships, making it a versatile tool for predictive modeling. Its inherent regularization feature helps in reducing overfitting, making the model generalizable to new,

unseen data. Moreover, XGBoost's ability to perform parallel processing significantly cuts down on computational time, allowing for quicker model development and iteration.
Pairing XGBoost with SHAP analysis leverages the strengths of both methodologies, resulting in a powerful synergy for our work. XGBoost's capability in modeling complex, non-linear relationships and interactions between variables lays a solid foundation for predictive accuracy. When SHAP analysis is applied to this model, it unpacks the 'black box', providing a clear, detailed, and quantifiable explanation of how each input feature influences the model's predictions.

The use of tree-based models, particularly XGBoost, in SHAP analysis is highly regarded due to several intrinsic properties that make them suitable for explaining model predictions. SHAP values are based on game theory and provide a measure of the contribution of each feature to the prediction of a particular instance, taking into account all possible combinations of features. The tree structure of models like XGBoost allows for an efficient calculation of these values through the Tree SHAP algorithms that exploit the model's hierarchical decision paths to compute exact SHAP values in polynomial time. Computational efficiency and algorithms that produce exact SHAP values are significant advantages over other model types, such as neural networks, where the computation of SHAP values can be more computationally intensive and less precise.

There are many advantages in this method combination. First, it enhances the interpretability of complex models, making the results more transparent and understandable. This is particularly beneficial in a research setting where explaining the 'why' behind predictions is just as crucial as the predictions themselves. Second, it aids in feature selection, helping to identify the most impactful factors that drive the model's outputs. This insight is invaluable for refining the model and focusing on the most relevant inputs. Finally, the use of XGBoost with SHAP analysis fosters trust and confidence in the model's predictions, as stakeholders can see and understand the rationale behind the model's decisions, making it a potent tool in our work on improving radiation simulation models.

In our XGBoost model training, we use the Python XGBoost library. We tested over and extensive set of parameter combinations and found the following training parameters produced the best results and decided to use them in the training:

- max_depth: 15
- colsample_bytree: 0.3
- learning_rate: 0.9

The objective function to be minimized in the training was the mean squared error.

5 Results

5.1 Global radiation

To advance our understanding and predictive capabilities regarding global and ultraviolet radiation at the surface level, we conducted a comprehensive investigation, ultimately aiming to refine the computational models currently in use in CAMS. We adopted a data-driven approach to analyze and interpret the factors that predominantly contribute to the inaccuracies and errors observed in existing radiation simulations and forecasts.

Our methodology involved applying the post-process correction approach to the CAMS global horizontal irradiation (GHI) data at the collocated station locations. The primary objective of this correction was to correct for the inherent biases within the GHI data, thereby aligning the computational outputs more closely with empirical observations and thus improving the reliability and precision of radiation forecasting and simulations.

Figure 3 shows a comparative analysis of the GHI data — both with and without the implementation of post-process correction — against ground-based observations. In the comparison, we treat the ground-based observations as the accurate ground truth against which we compare the simulation model data.

Figure 4 presents histograms that show the errors associated with the original CAMS GHI data and the data subjected to post-process correction. This visual representation highlights the discrepancies between the two datasets and provides an intuitive understanding of the error distribution, a crucial factor in comprehending the overall impact of our correction methodology.

Figure 5 presents both the CAMS GHI and the post-process corrected GHI across various binned GHI ranges in a more granulated examination of the error distribution. This detailed analysis highlights the consistency and reliability of the post-process correction across a spectrum of conditions, reaffirming the robustness of our approach.

Our findings reveal an improvement in the post-process corrected data. Notably, the post-process correction method slightly reduced the overall GHI bias even though in some size bins the correction resulted in slightly increased bias. Moreover, a marked reduction in the error standard deviation was observed, underscoring our approach's effectiveness in enhancing the radiation data's precision. Interestingly, our findings also highlighted that the improvements by the post-process correction were not confined to specific GHI ranges; rather, the benefits were uniformly distributed, affirming the correction method's applicability across different radiation intensities.

From a quantitative standpoint, the GHI error standard deviation — a simplified yet insightful uncertainty estimate — exhibited a noteworthy decrease in the post-corrected data. Specifically, in the post-process corrected GHI data, the standard deviation ranged between 65 and 103 W/m^2, significantly improving from the 84 to 145 W/m^2 range observed in the uncorrected GHI data. This quantifiable reduction in standard deviation signifies a substantial enhancement in data accuracy. It reinforces our correction methodology's potential to serve as a cornerstone in future radiation modelling and forecasting.

Tables 4-9 show the accuracy metrics station-by-station for both the 1-minute and 1-hour averaged data, both in terms absolute and relative values and both with and without the post-process correction.

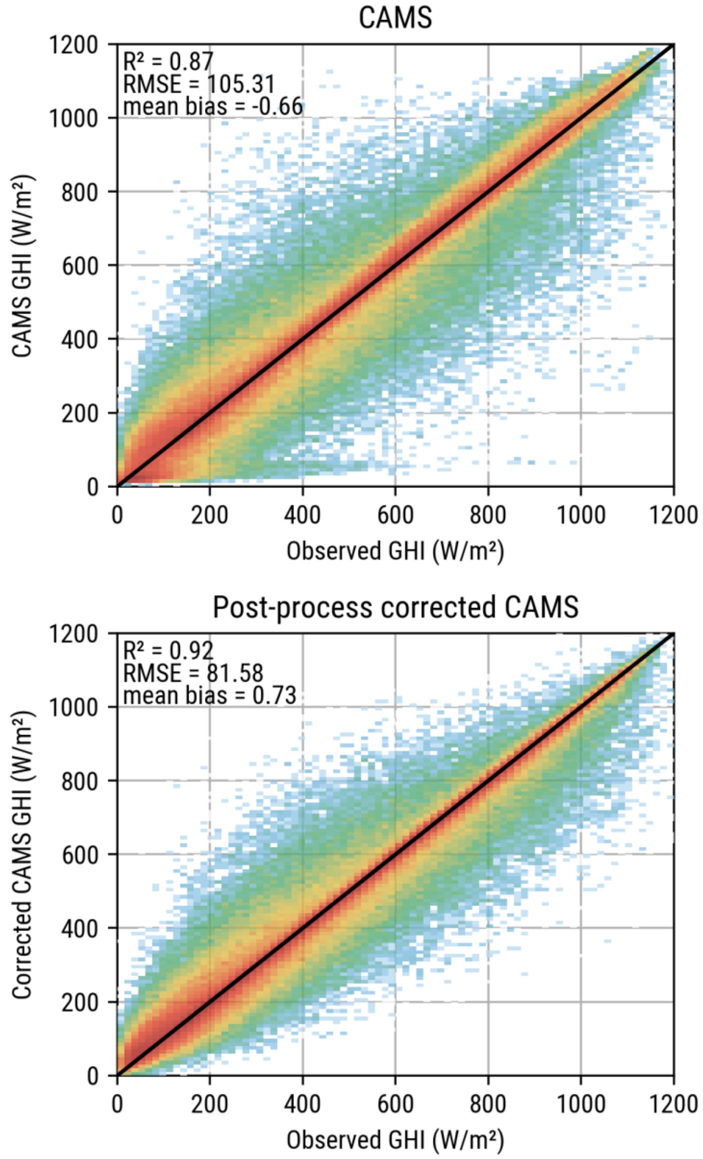

Figure 3. Global horizontal irradiation (GHI). Top: CAMS model GHI as function of observed GHI. Bottom: Post-process corrected CAMS GHI as function of observed GHI.

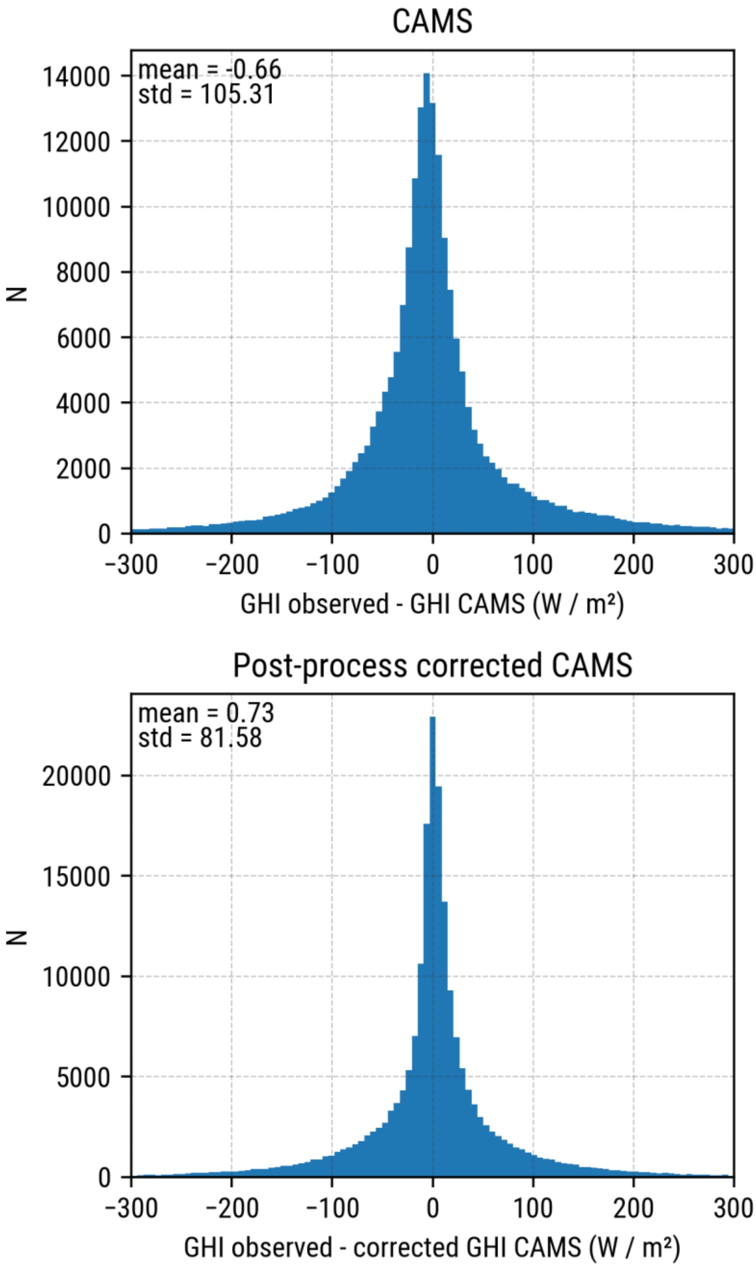

Figure 4. Error histograms for the CAMS (top) and post-process corrected CAMS (bottom) global horizontal irradiation (GHI) at surface level.

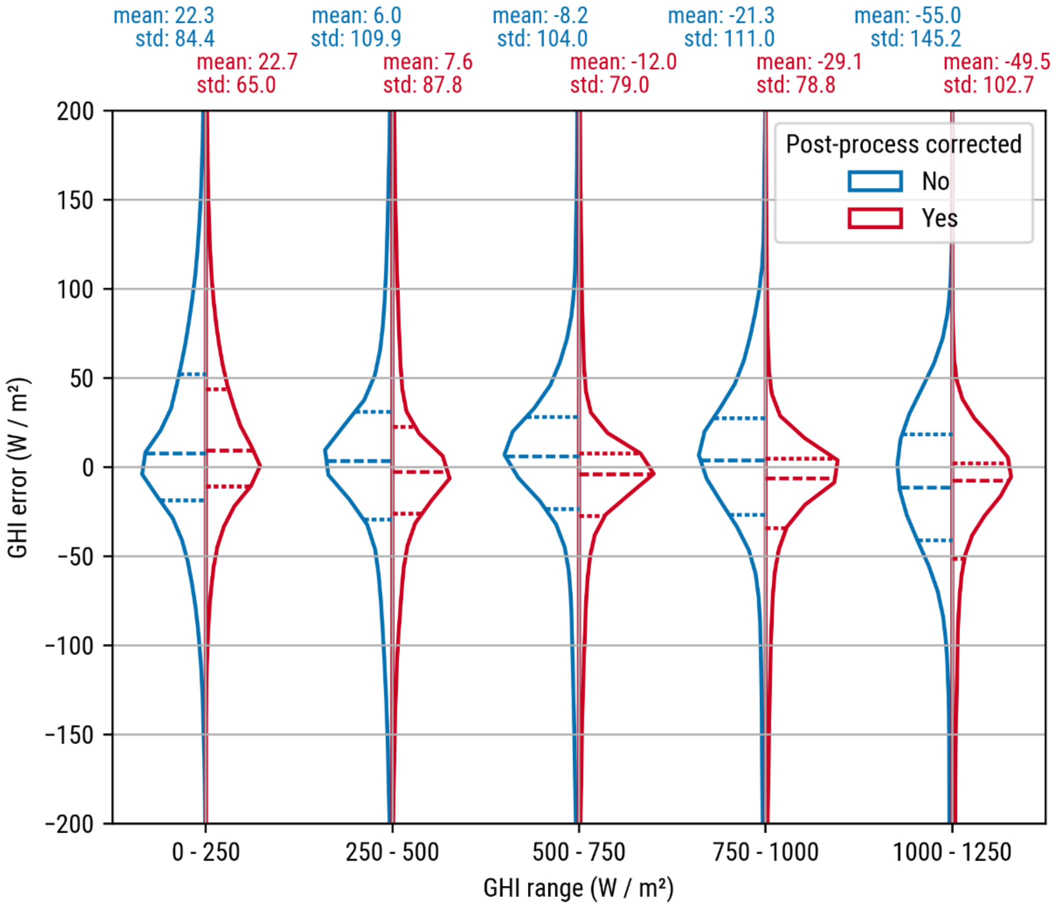

Figure 5. Global horizontal irradiation (GHI) error distributions for the uncorrected (blue) and post-process corrected (red) CAMS GHI for different GHI ranges. The metrics on the top of the figure show the mean and standard deviation (std) of the error.

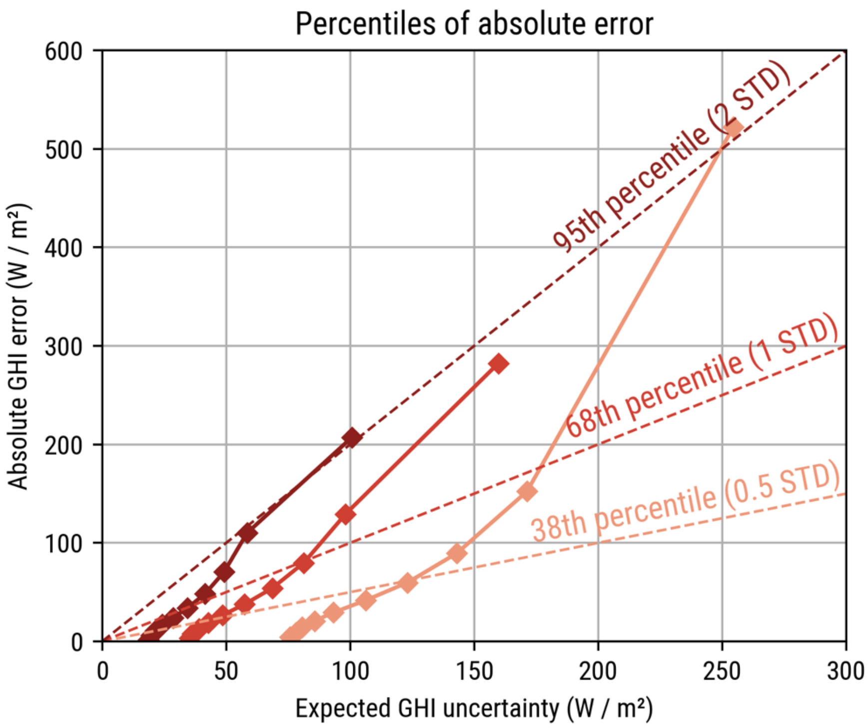

Figure 6. Absolute global horizontal irradiation (GHI) error as function of expected GHI uncertainty computed as the ensemble member spread. The GHI data was binned to 10 equally populated bins and the diamonds indicate different GHI bins. The lightest tone corresponds to 38th percentile (0.5 standard deviations), the middle tone corresponds to 68th percentile (1 standard deviation), and the darkest to the 95th percentile (2 standard deviations). The dashed lines show the theoretical ideal values for the uncertainties. The lines below the corresponding dashed lines indicate overestimated uncertainty estimates and the lines above the corresponding dashed lines indicate underestimated uncertainty estimates.

Building on the foundation of our post-process correction work, we adopted an ensemble approach and generated uncertainty estimates for the post-process corrected GHI. This step enhanced and facilitated a more comprehensive understanding of the GHI data reliability and precision.

Figure 6 shows the relationship between the absolute GHI error and the expected GHI uncertainty in a similar manner of the approach utilized by Sayer et al. (2020) for evaluating uncertainty estimates in satellite-based aerosol data. An observation from Figure 6 reveals a pattern in the behaviour of the ensemble-based uncertainty estimates. At lower GHI levels, a trend of overestimation in uncertainty estimates became apparent. Conversely, at higher GHI levels, the trend shifted towards underestimating uncertainty. Underestimating uncertainties carry the risk of unwarranted confidence in the data, potentially leading users to place excessive trust in the values, especially in scenarios where precision is essential. However, when viewed in entirety, the ensemble-based uncertainty estimates are balanced, aligning with acceptable standards.

Figures 7-10 show the ensemble-based uncertainty estimates in relation to all model input parameters. This comprehensive examination was designed to reveal potential correlations, or the lack thereof, between the levels of uncertainty and the post-process correction model's input parameters. The investigation did not reveal any significant correlations between the model inputs and estimated uncertainty levels.

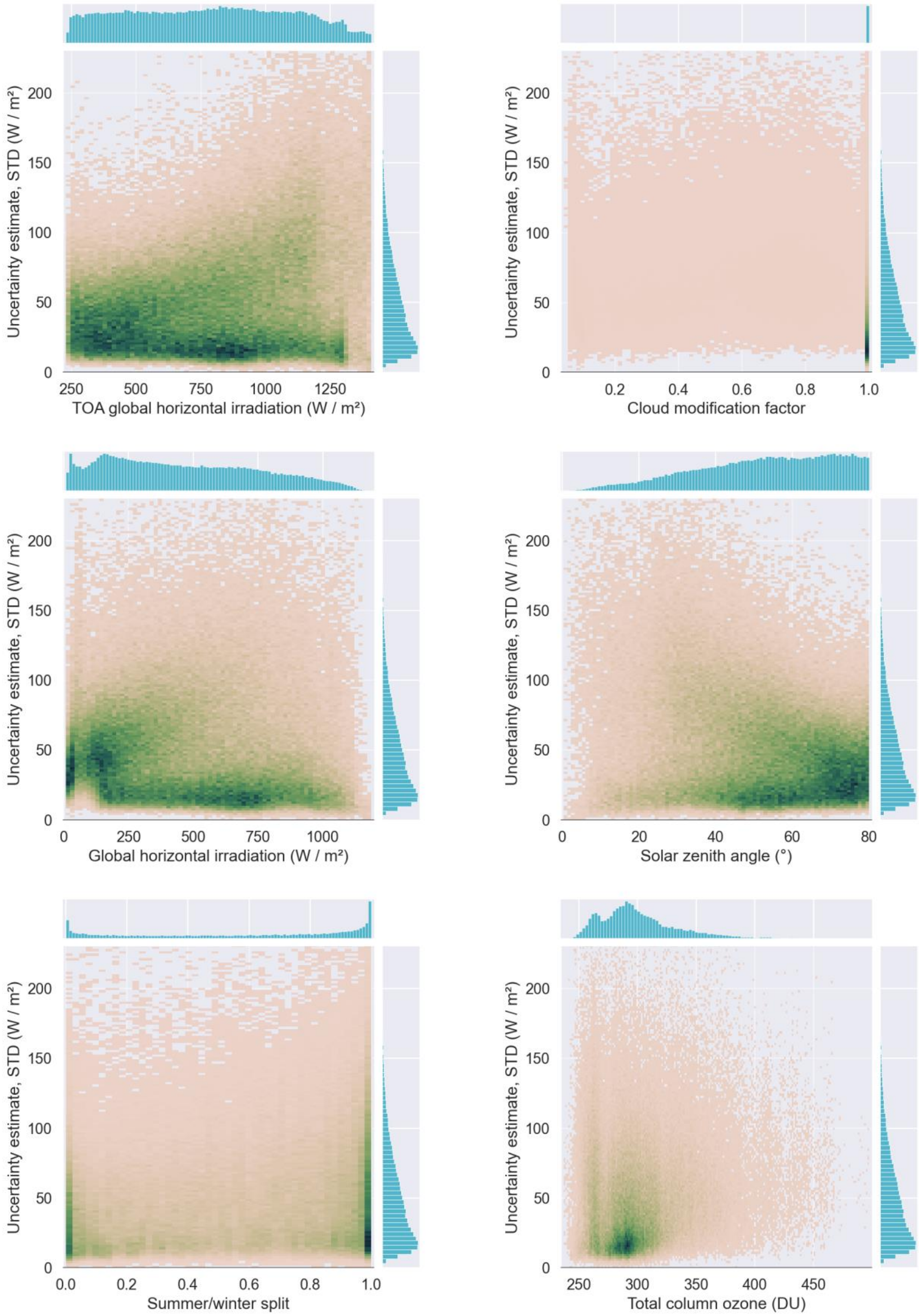

Figure 7. Uncertainty estimates for global horizontal irradiation (GHI) as function model input parameters. Top left: TOA global horizontal irradiation. Top right: Cloud modification factor. Middle left: GHI. Middle right: Solar zenith angle. Bottom left: Summer/winter split. Bottom right: Total column ozone.

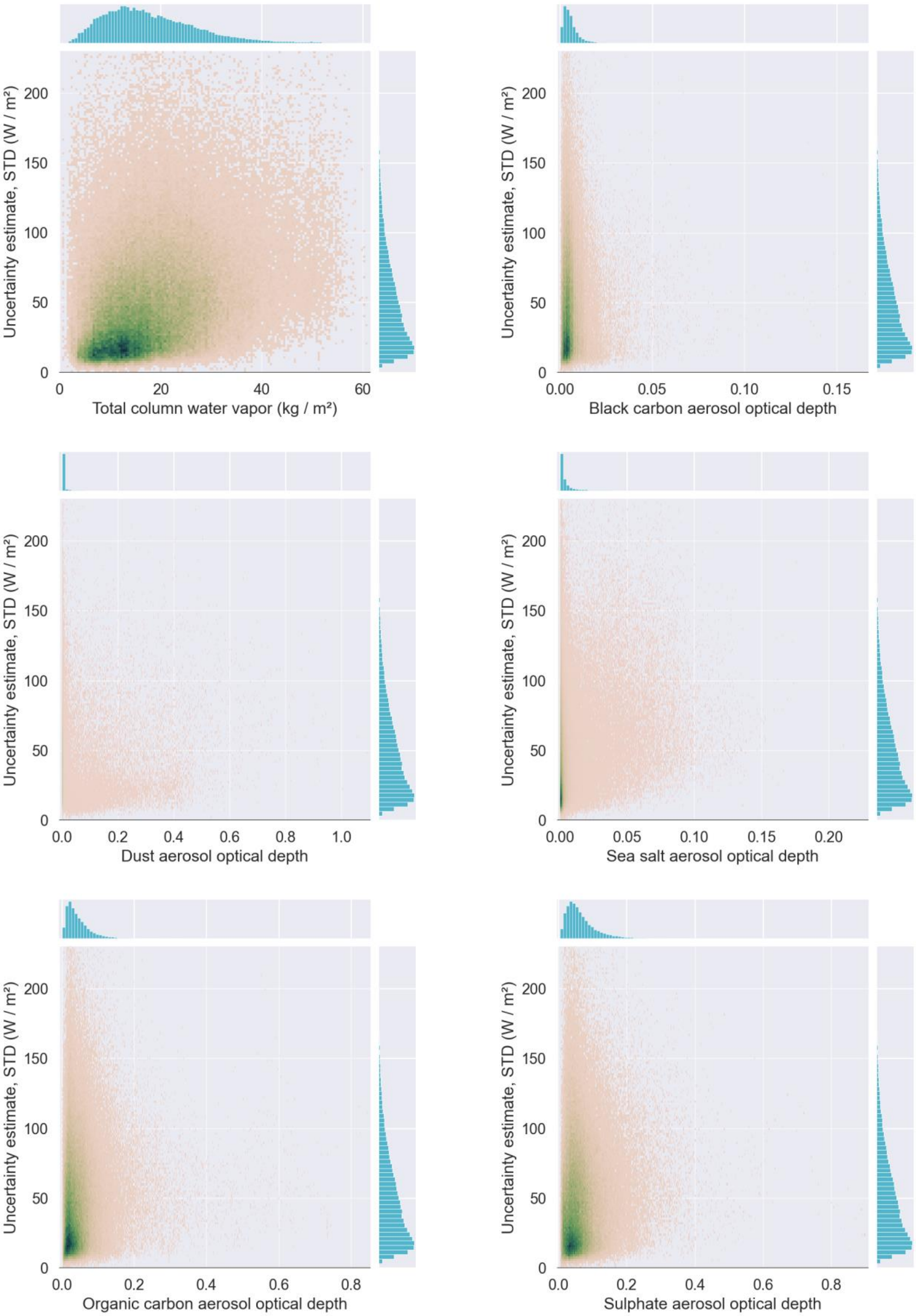

Figure 8. Uncertainty estimates for global horizontal irradiation (GHI) as function model input parameters. Top left: Total column water vapor. Top right: Black carbon aerosol optical depth. Middle left: Dust aerosol optical depth. Middle right: Sea salt aerosol optical depth. Bottom left: Organic carbon aerosol optical depth. Bottom right: Sulphate aerosol optical depth.

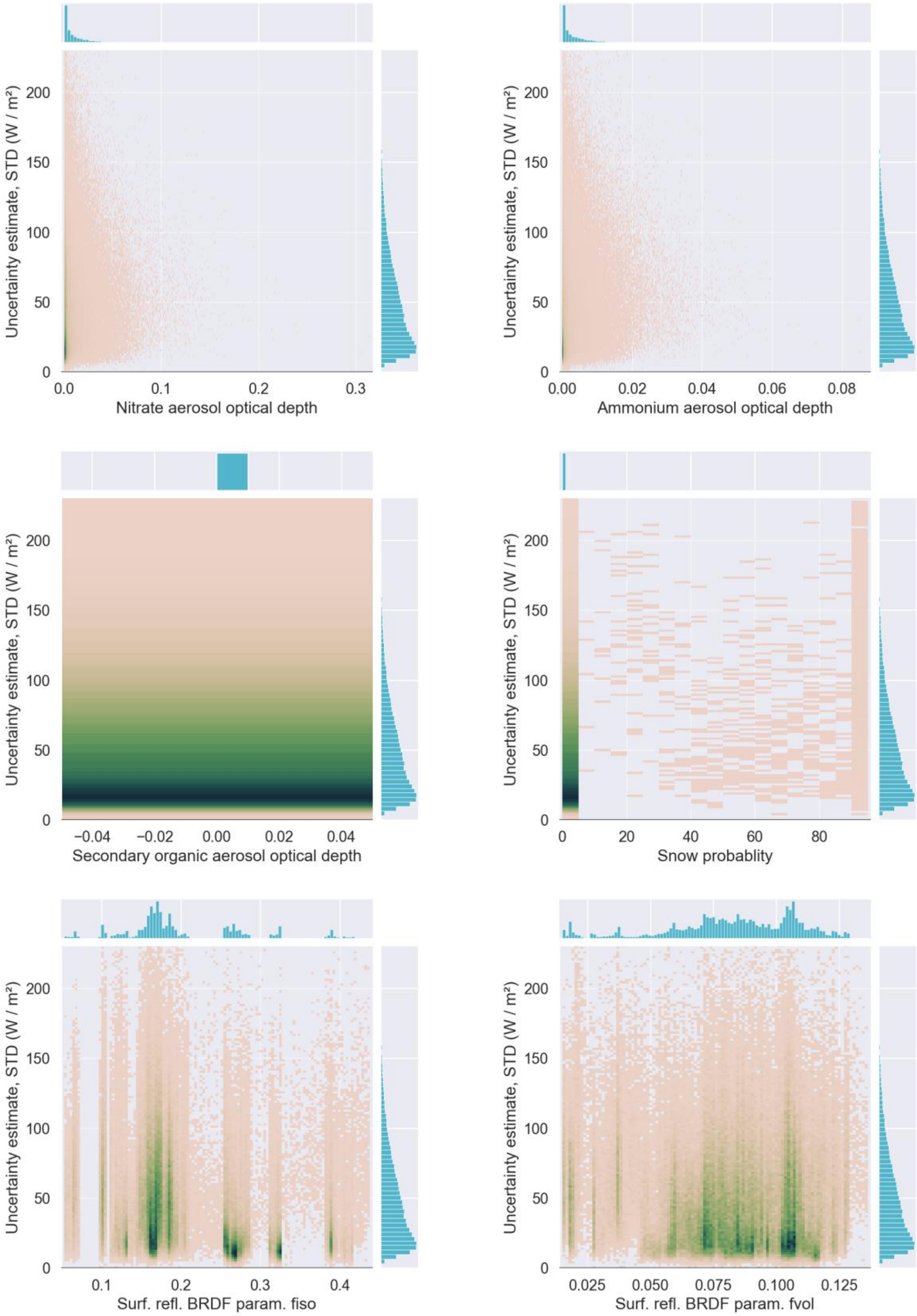

Figure 9. Uncertainty estimates for global horizontal irradiation (GHI) as function model input parameters. Top left: Nitrate aerosol optical depth. Top right: Ammonium aerosol optical depth. Middle left: Secondary organic aerosol optical depth. Middle right: Snow probability. Bottom left: Surface reflectance BRDF parameter fiso. Bottom right: Surface reflectance BRDF parameter fvol.

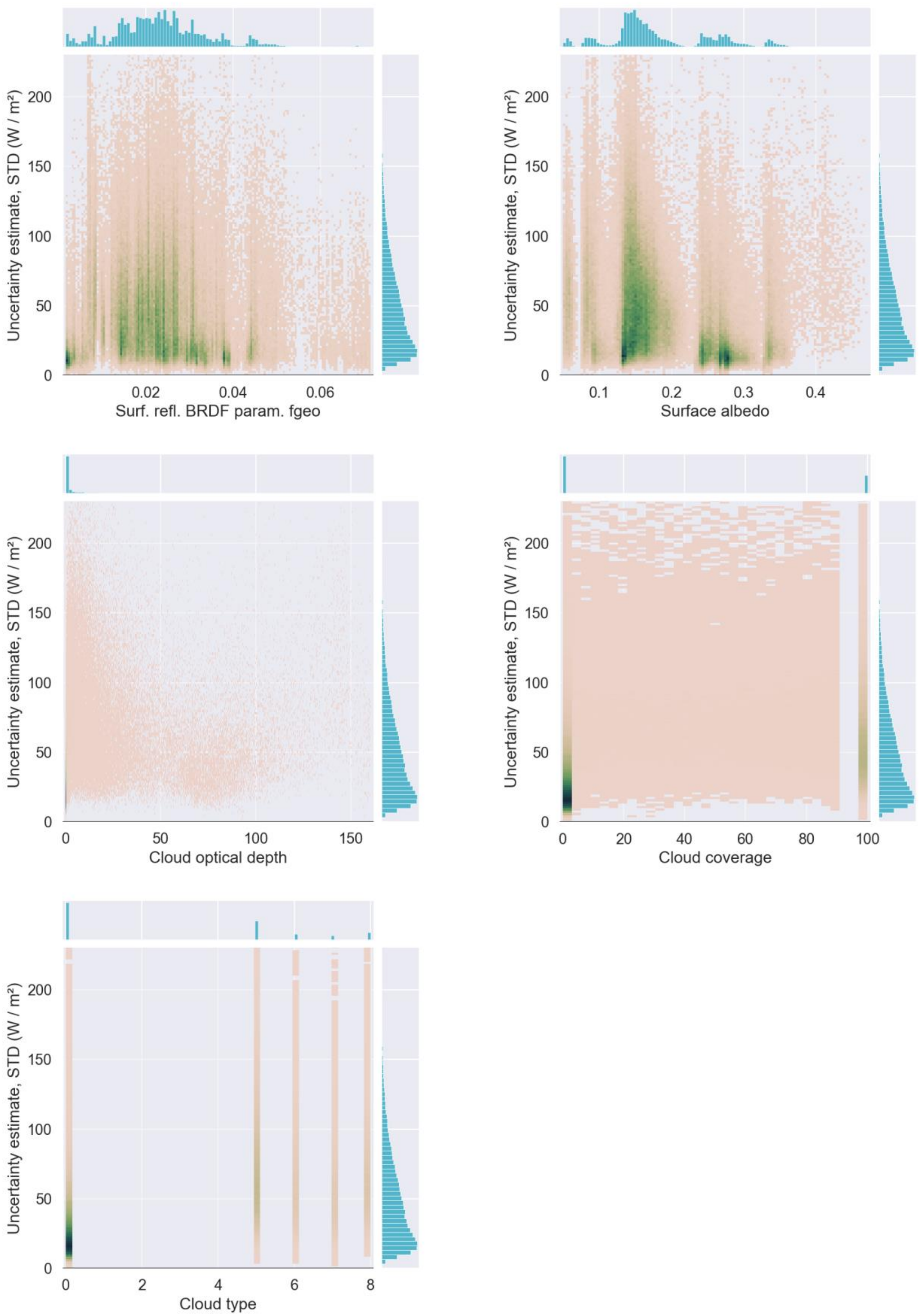

Figure 10. Uncertainty estimates for global horizontal irradiation (GHI) as function model input parameters. Top left: Surface reflectance BRDF parameter fgeo. Top right: Surface albedo. Middle left: Cloud optical depth. Middle right: Cloud coverage. Bottom left: Cloud type.

In our analysis of the factors influencing the accuracy and uncertainties of the CAMS GHI data, we embarked on a comprehensive SHAP (SHapley Additive exPlanations) analysis of the post-process correction model. This analytical approach was employed to uncover the most important features that play the most significant roles in the bias observed in the CAMS GHI data.

Figure 11 shows a result of our detailed SHAP analysis, presenting a SHAP diagram for a single ensemble member correction model. This diagram provides a comprehensive view of the SHAP values corresponding to the entire GHI dataset, offering a broad overview of the impact of various features on the post-process correction model's predictions.

In this model, cloud coverage emerged as the most significant feature influencing the CAMS model error. Our SHAP analysis showed that larger-than-average cloud coverage almost always coincides with positive SHAP values. This correlation suggests a tendency towards underestimating CAMS radiation data under such conditions. Conversely, smaller-than-average cloud coverage frequently yields negative SHAP values, indicating a probable overestimation of CAMS radiation data. These findings underscore the critical impact of cloud coverage on the model's accuracy in predicting GHI, highlighting the need for precise cloud representation in atmospheric models. The study also uncovered a similar pattern with another cloud-related variable: cloud optical depth. Large cloud optical depths were consistently associated with positive SHAP values, while small optical depths correlated with negative SHAP values. Most SHAP values were relatively close to zero when examining variables unrelated to clouds. This observation indicates their relatively minor importance in adjusting the CAMS model's predictions or at least stresses the more considerable importance of the cloud-related variables in contributing to the CAMS model error.

Figure 12 shows the SHAP values for a single model and single prediction. This detailed breakdown explains how each input value contributes explicitly to the prediction in this particular instance, offering a detailed perspective on the model's decision-making process.

In Figures 13-16, the mean SHAP values, representative of the entire ensemble of models, are catalogued in the order of their significance. These figures highlight the model's input features and their effect on the model's predictions.

The cloud modification factor is the most significant input variable affecting the post-process correction. The relationship between the cloud modification factor and SHAP value is obvious. The SHAP values are positive for smaller cloud modification factor values, showing a positive impact on the GHI predictions. However, as the cloud modification factor grows to values close to 1, the SHAP values have mostly negative values, mainly between 0 and -80 W/m2. This trend explains a critical insight: larger cloud modification factors are likely to be linked to the underestimation of GHI.

The GHI value itself was identified as the second most critical input variable. A pattern emerged here: smaller GHI values (below 150 W/m^2) are primarily associated with positive SHAP values (10-40 W/m^2), indicating a tendency towards GHI underestimation. In contrast, larger GHI values (above 750 W/m^2) are commonly linked with negative SHAP values (-20 to 0 W/m2), signalling GHI overestimation.

The third significant factor was the cloud optical thickness. Large cloud optical thickness values (>50) were clearly associated with large SHAP values mostly between 0-70 indicating GHI underestimation.

Further down the list of significance, the cloud type emerged as the fourth most significant variable, with cloud type 5 — low-level clouds — distinguished by its prominently positive SHAP values ranging from 0 to 50 W/m^2. This positive skew indicates that low-level clouds are another substantial factor contributing to the overestimation of GHI values in the CAMS dataset.

The SHAP analysis did not reveal any other significant input factors. Almost all SHAP values associated with the other input variables remained modestly confined within the -20 to 20 W/m^2 range. This finding underscores a lack of significant correlation between these additional input variables and the GHI error SHAP values. This refined understanding by the SHAP analysis enables more targeted and effective enhancements to the post-process correction and CAMS radiation models, leading to more accurate and reliable GHI predictions.

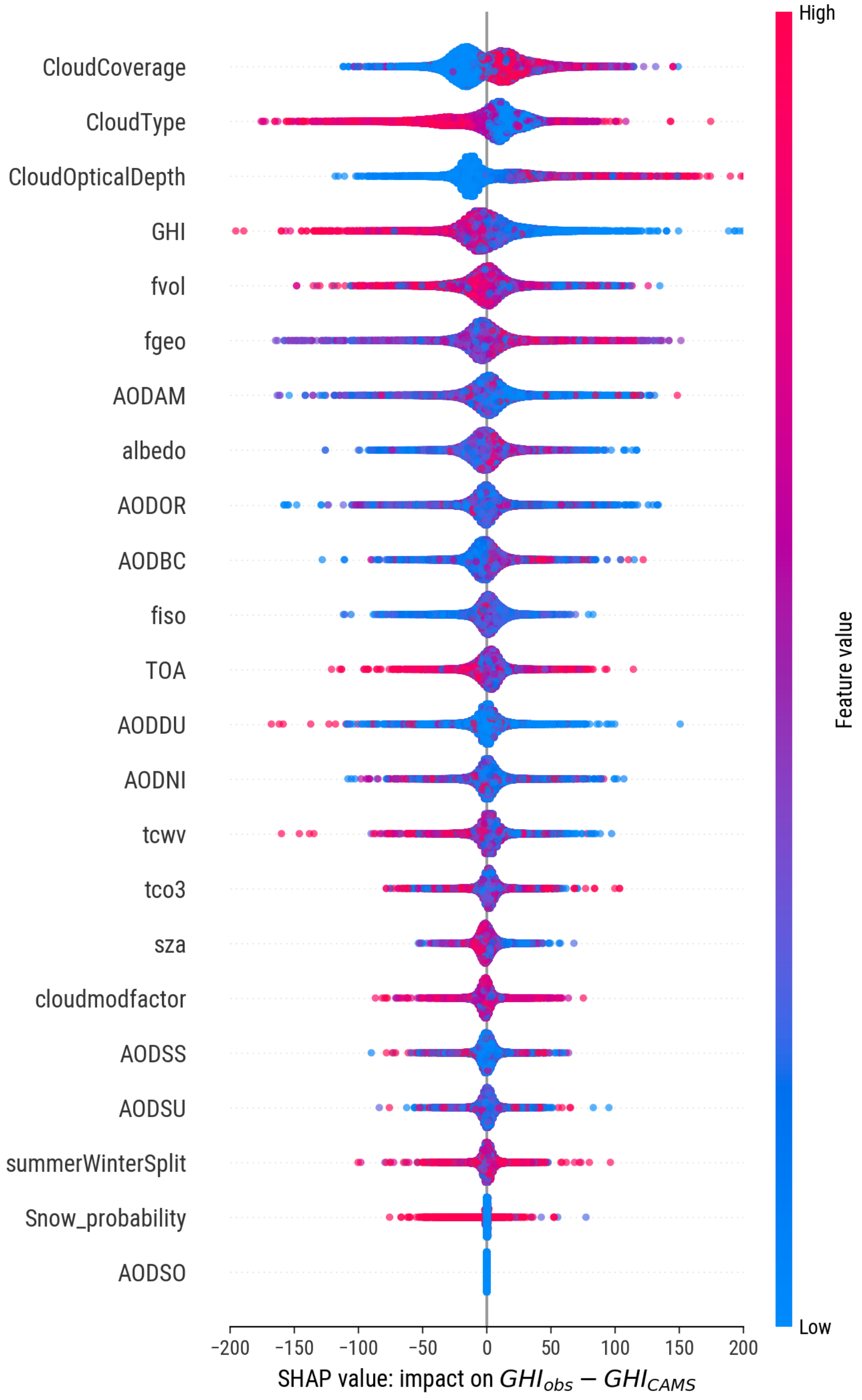

Figure 11. SHAP diagram for the global horizontal irradiation (GHI) error (GHI_{obs} - GHI_{CAMS}).

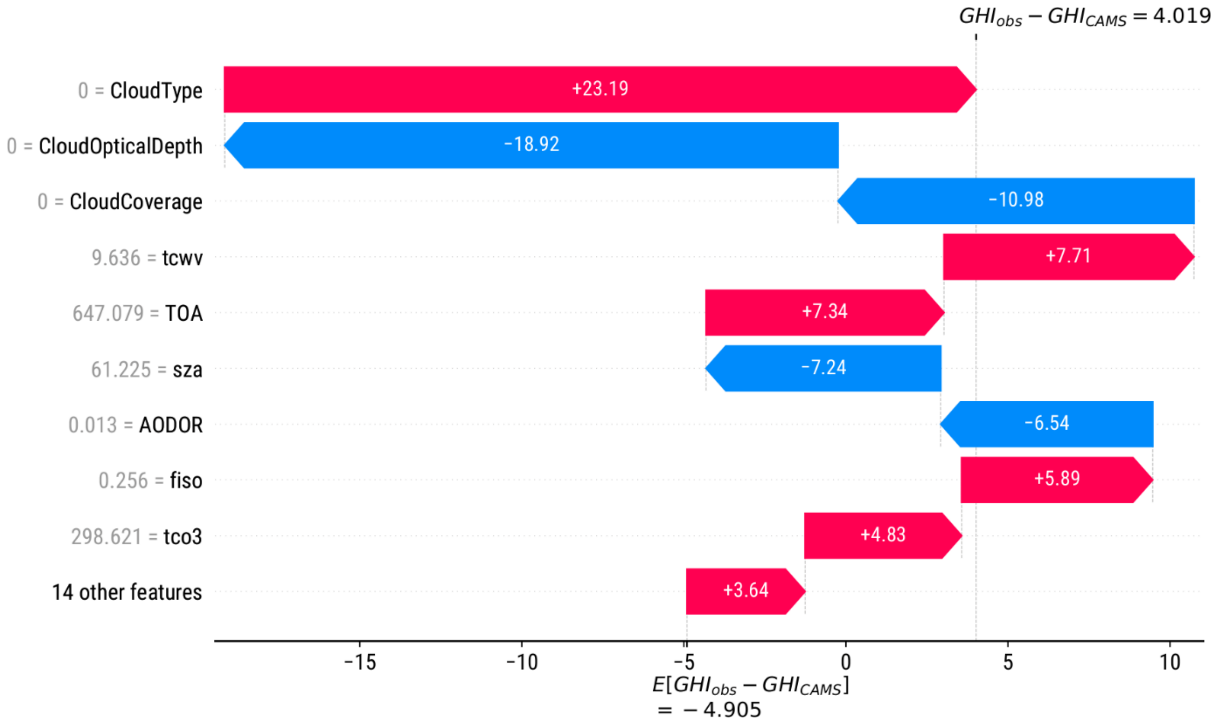

Figure 12. SHAP values corresponding to a single prediction by the post-process GHI error correction model.

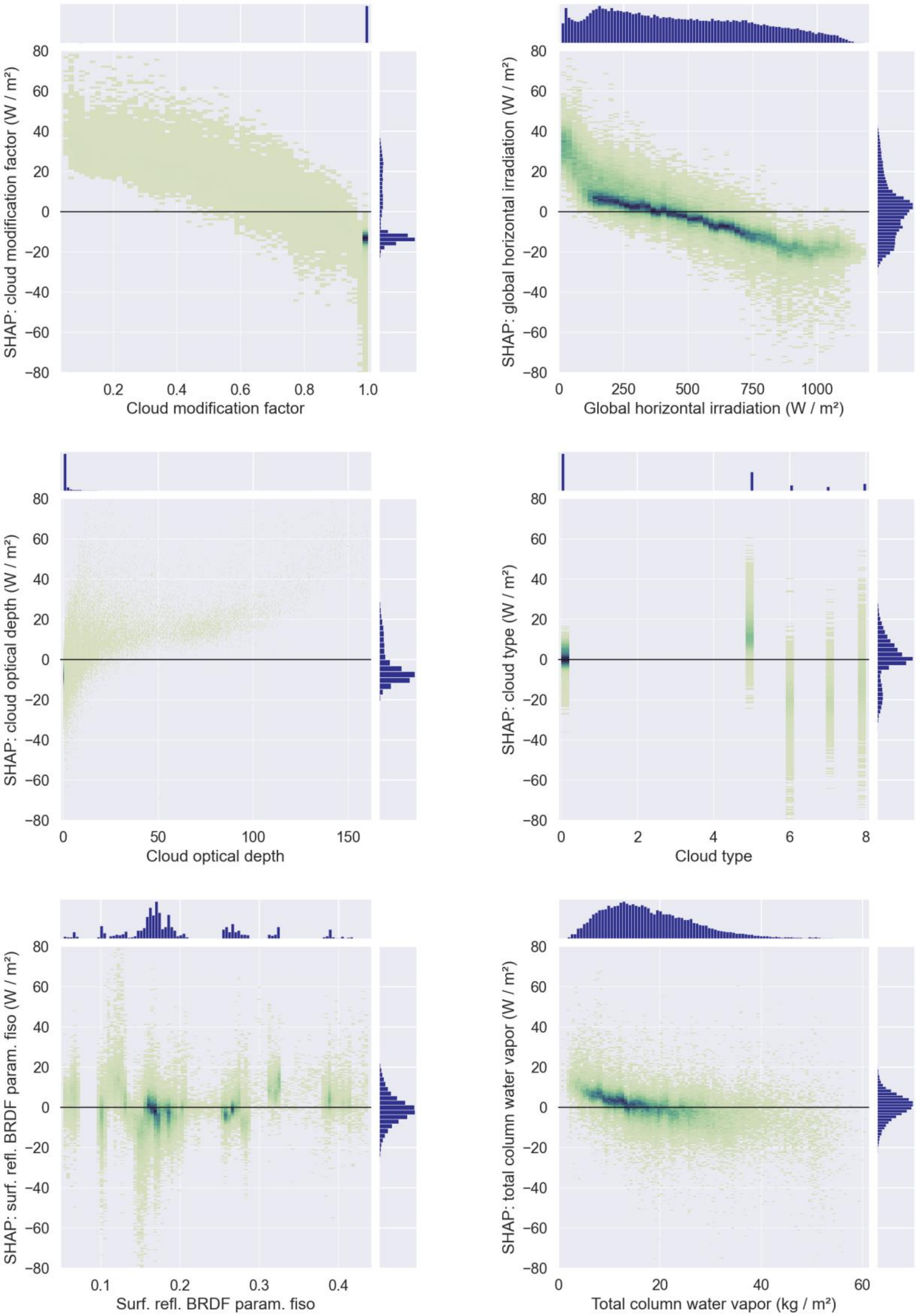

Figure 13. SHAP value for global horizontal irradiation (GHI) as function model input parameters. Top left: Cloud modification factor. Top right: GHI. Middle left: Cloud optical depth. Middle right: Cloud type. Bottom left: Surface reflectance BRDF parameter fiso. Bottom right: Total column water vapor.

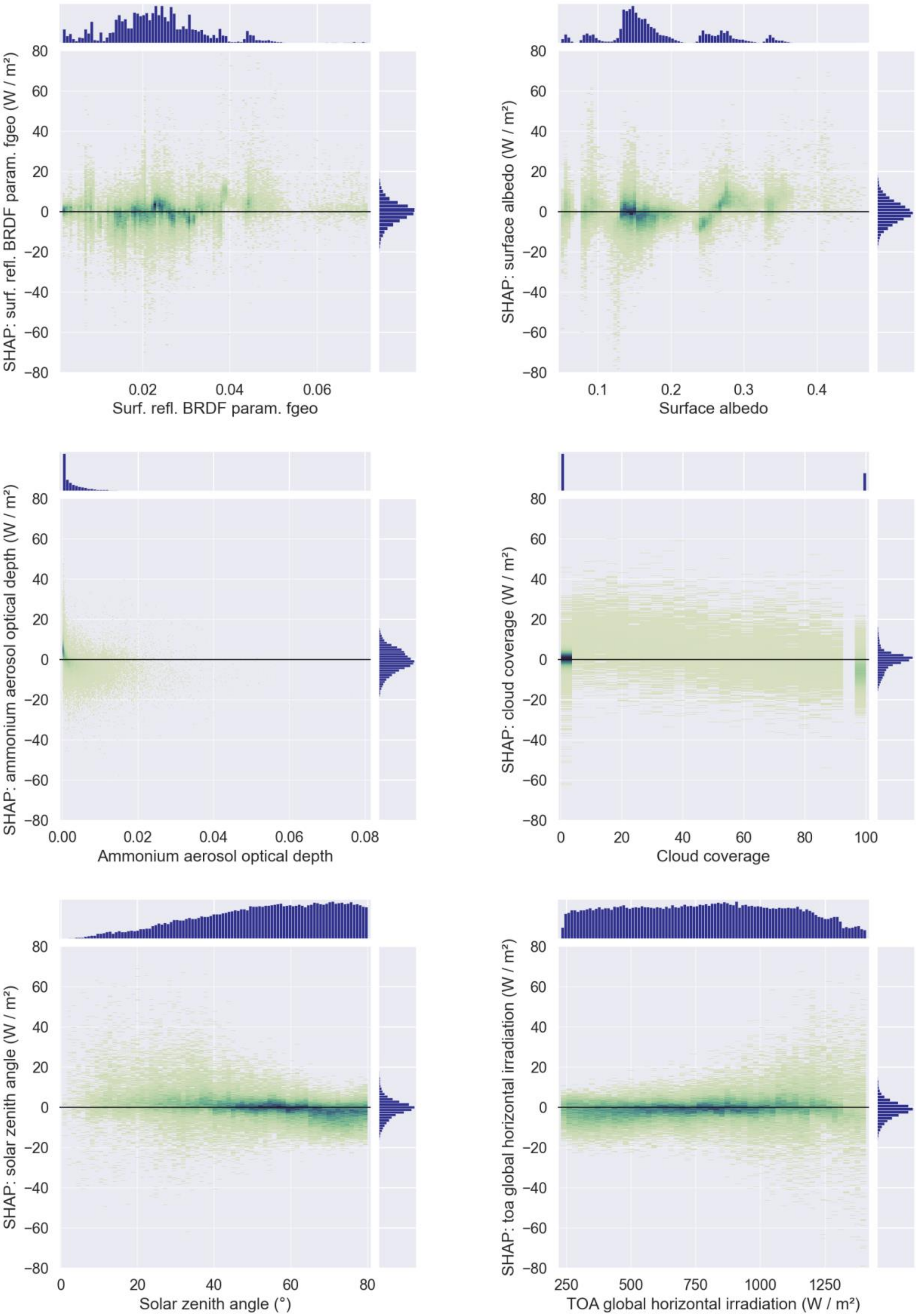

Figure 14. SHAP value for global horizontal irradiation (GHI) as function model input parameters. Top left: Surface reflectance BRDF parameter fgeo. Top right:Surface albedo. Middle left: Ammonium aerosol optical depth. Middle right: Cloud coverage. Bottom left: Solar zenith angle. Bottom right: TOA GHI.

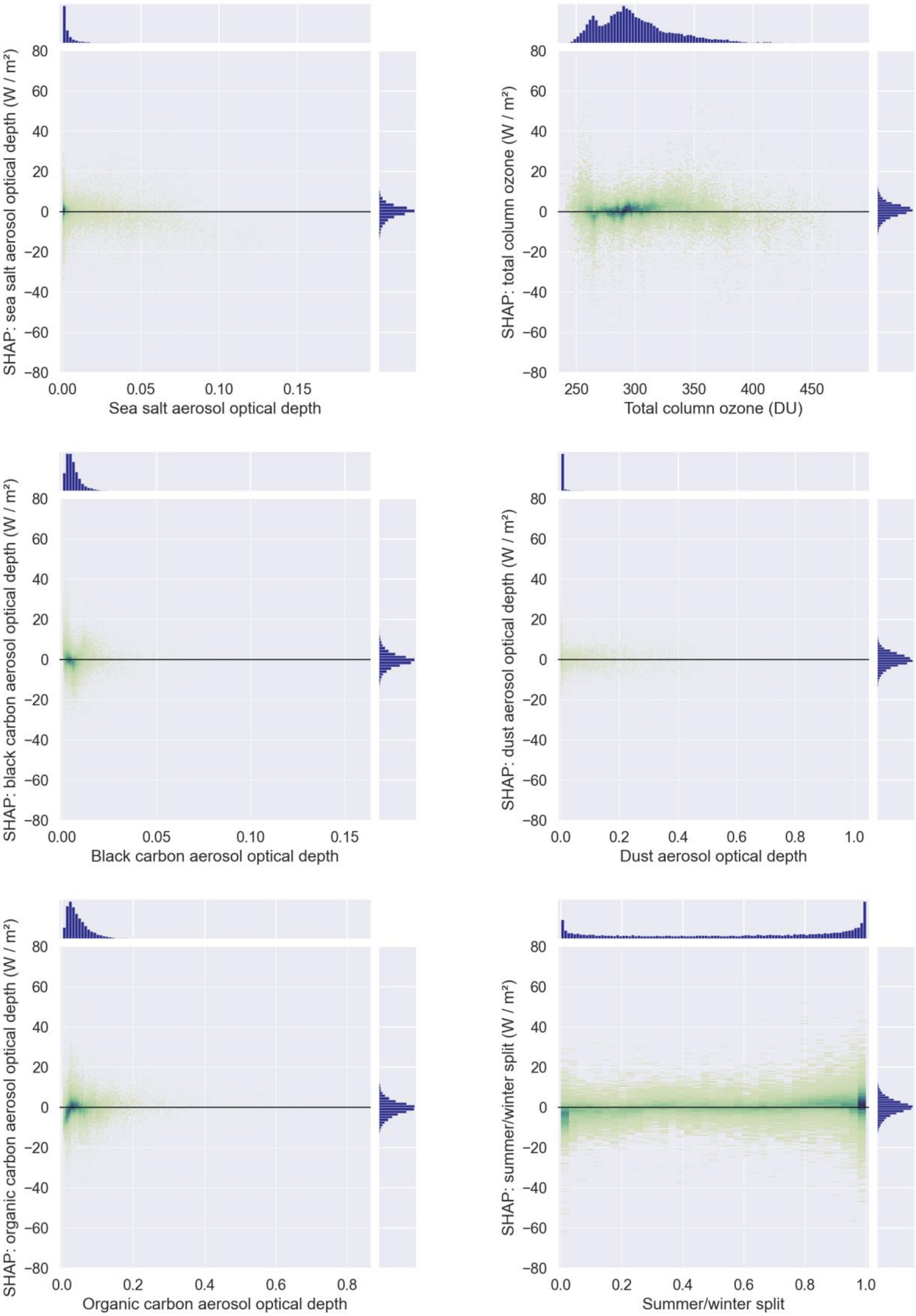

Figure 15. SHAP value for global horizontal irradiation (GHI) as function model input parameters. Top left: Sea salt aerosol optical depth. Top right: Total column ozone. Middle left: Black carbon aerosol optical depth. Middle right: Dust aerosol optical depth. Bottom left: Organic carbon aerosol optical depth. Bottom right: Summer/winter split.

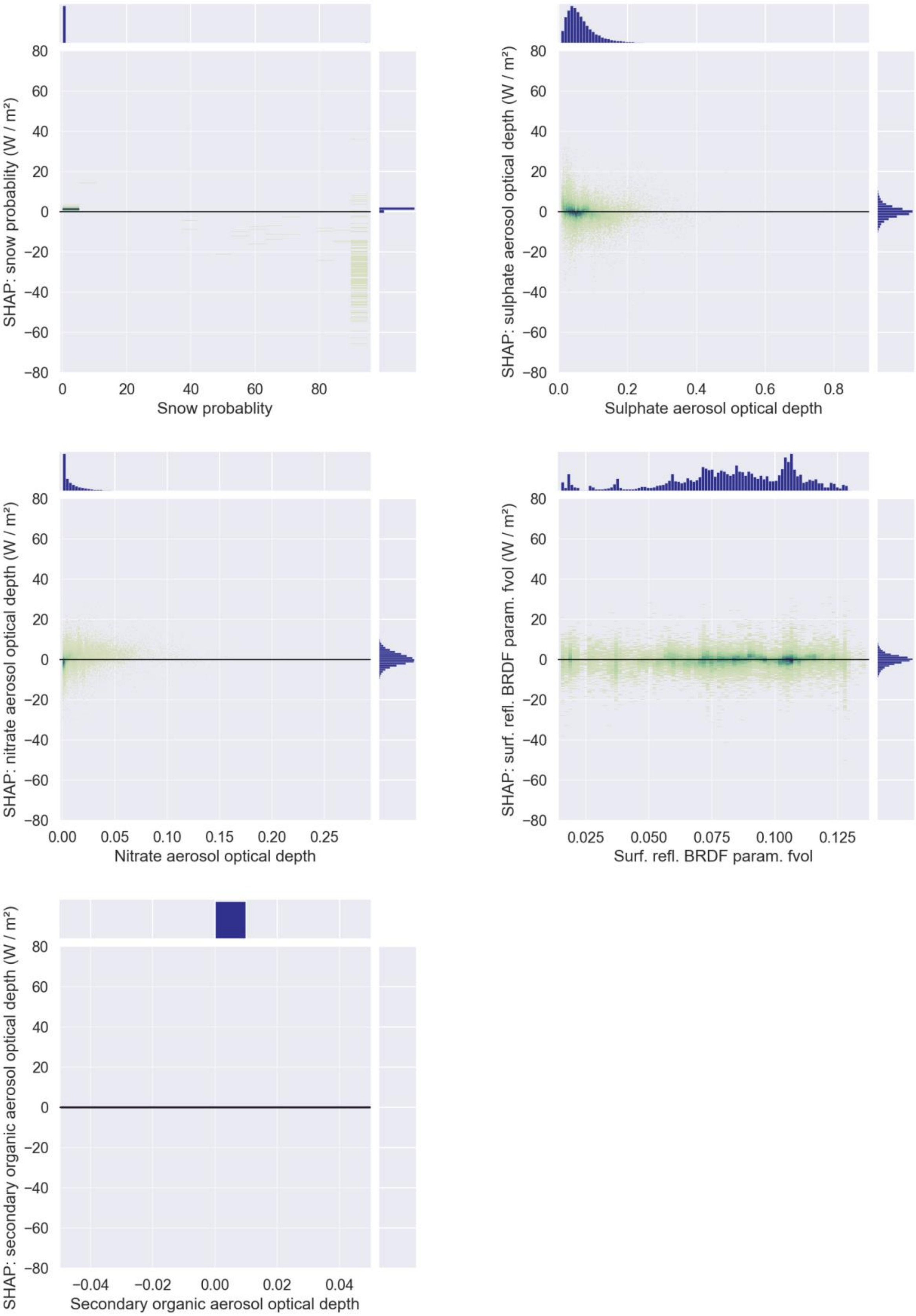

Figure 16. SHAP value for global horizontal irradiation (GHI) as function model input parameters. Top left: Snow probability. Top right: Sulphate aerosol optical depth. Middle left: Nitrate aerosol optical depth. Middle right: Surface reflectance BRDF parameter fvol. Bottom left: Secondary organic aerosol optical depth.

5.2 UV index

Next, we extended the methodology successfully employed in the analysis of GHI to improve the CAMS ultraviolet index (UVI) forecast data. We adopted the post-process correction approach, previously applied to GHI data, to UVI data to correct biases and match the forecast data more closely with empirical observations.

Figure 17 compares the UVI data before and after applying post-process correction against the ground-based observations. These observations, recognized as the accurate ground truth, provide a robust benchmark against which the efficacy of our correction measures can be evaluated.

Figure 18 presents histograms detailing the error distributions associated with the original CAMS UVI data and the data subjected to post-process correction. This graphical representation provides an intuitive understanding of the error distribution, highlighting the differences between the two datasets and the impact of our post-process correction methodology.

Figure 19 illustrates the error distributions for both the original and the post-process corrected UVI data across a spectrum of binned UVI ranges. This granular analysis underscores the consistency of the post-process correction across different levels of UVI and shows the method's robustness and applicability.

The overall finding in all the results is the enhanced precision and reliability of the UVI data. The post-process correction method reduced the bias in UVI data, as evidenced by the average reduction across various measurements. Furthermore, a notable decrease in the error standard deviation was observed, highlighting the method's effectiveness in enhancing the precision of the UVI data. The improvement in standard deviations was not confined to specific UVI ranges, and the benefits of post-process correction were uniformly distributed.

The standard deviation showed a significant improvement in the post-process corrected data over the original CAMS UVI forecast. The post-process corrected UVI data exhibited a standard deviation ranging between 0.4 and 1.4, marking an enhancement compared to the 0.6 to 2.0 range observed in the uncorrected UVI data.

Tables 10-12 show both the relative and absolute accuracy metrics station-by-station data both with and without the post-process correction.

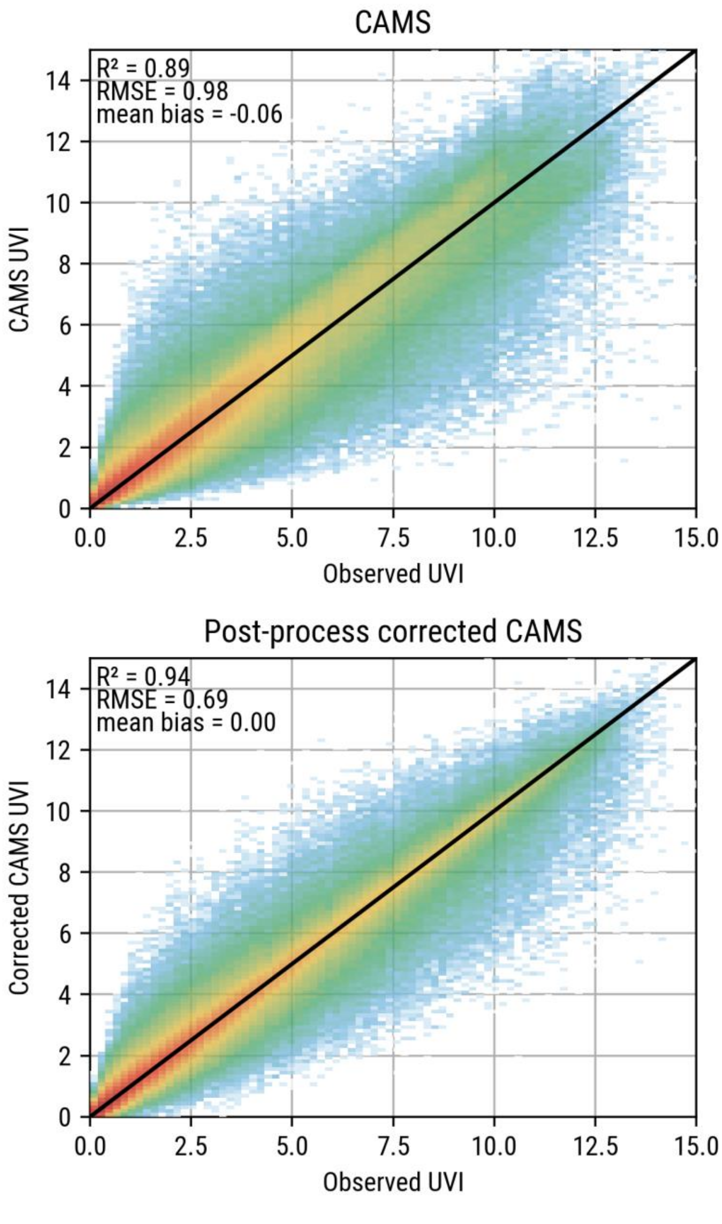

Figure 17. UV index (UVI). Top: CAMS forecast UVI as function of observed UVI. Bottom: Post-process corrected CAMS forecast UVI as function of observed UVI.

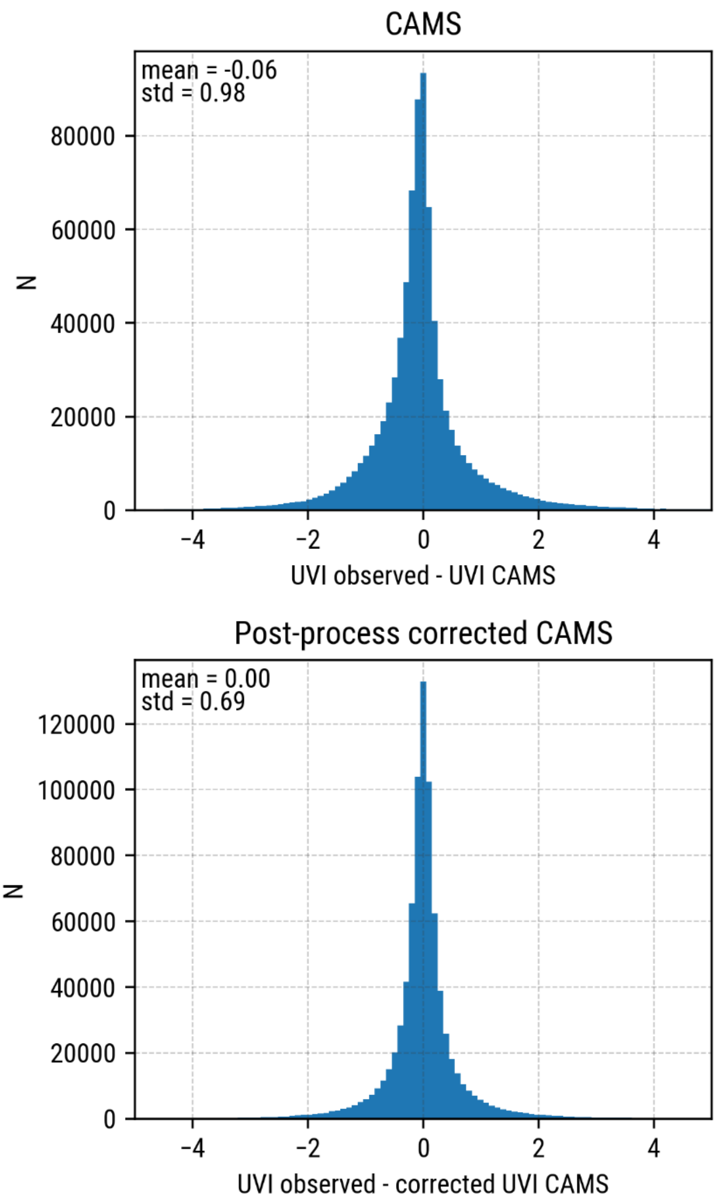

Figure 18. Error histograms for the CAMS (top) and post-process corrected CAMS (bottom) ultraviolet index (UVI) at surface level.

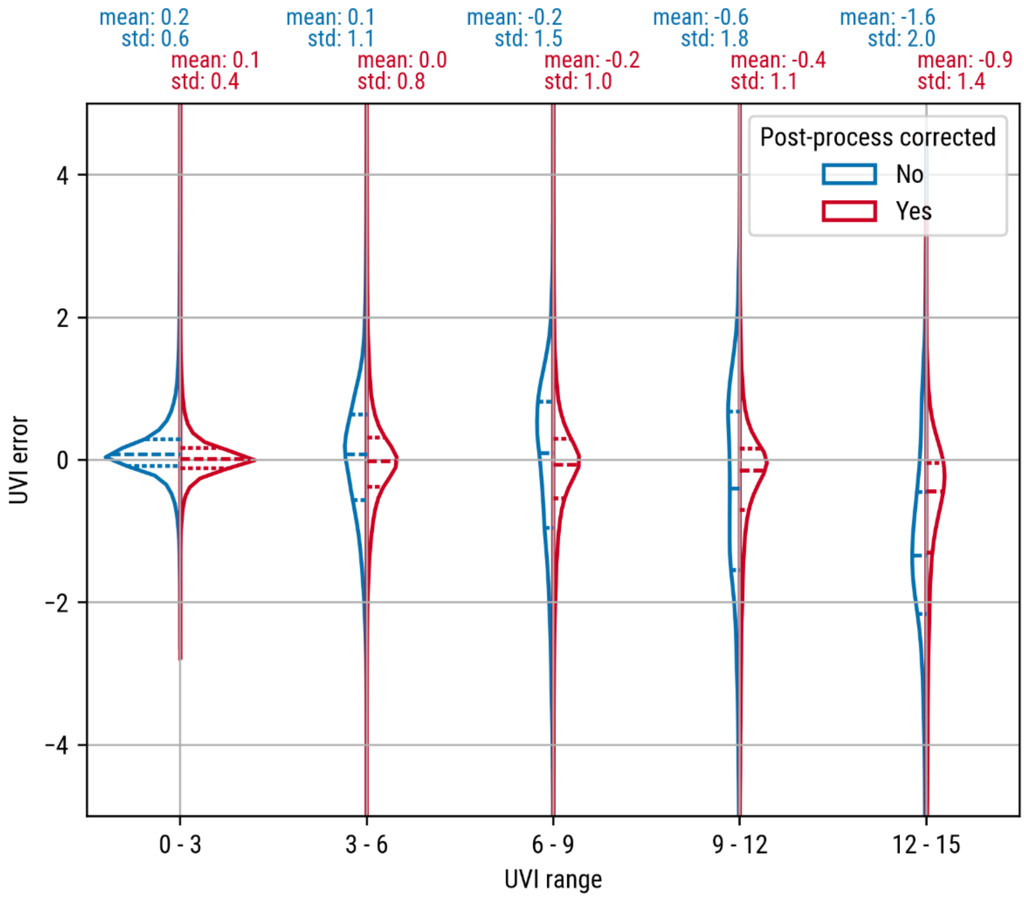

Figure 19. Ultraviolet index (UVI) error distributions for the uncorrected (blue) and post-process corrected (red) CAMS UVI for different UVI ranges. The metrics on the top of the figure show the mean and standard deviation (std) of the error.

Building upon the methodologies applied to GHI, our work further evolved to encompass the refinement of UVI forecast data through an ensemble-based post-process correction approach. This strategy was needed to generate the uncertainty estimates for the post-process corrected UVI, following the approach previously employed with GHI data.

Figure 20 shows the relationship between the absolute UVI error and the expected UVI uncertainty. The analysis revealed a pattern reminiscent of our findings with GHI. The ensemble-based uncertainty estimates for UVI tended to be overestimated at lower UVI levels, while a trend of underestimation was observed at higher UVI levels. Despite these nuances at the extremes, on average, the ensemble-based uncertainty estimates were within acceptable bounds, suggesting a robust overall performance of our correction and uncertainty estimation methodology.

We further extended our analysis to investigate how the uncertainty estimates behave in relation to all the model input parameters. Figures 21-24 describe the UVI uncertainty estimates as a function of the post-process correction model's input parameters. Our analysis showed some correlations between the input variables and the uncertainty estimates, revealing patterns that were not as pronounced in our GHI analysis. The top-of-atmosphere (TOA) global horizontal irradiation positively correlated with the ensemble-based UVI uncertainty estimate, indicating a direct relationship between increased irradiation and elevated uncertainty levels.

We made an observation regarding the cloud optical depth. Specifically, cloud optical depth values exceeding 100 were consistently linked with larger-than-average uncertainty estimates. This finding highlights the significant influence of dense cloud cover on the uncertainty of UVI forecasts, marking it as a critical factor in the predictability of ultraviolet radiation levels. Also there was a positive correlation between the UVI uncertainty and total column water wapor.

These insights, derived from our detailed analysis, enhanced our understanding of the uncertainties in UVI forecasting and gave us ways to refine our post-process correction model further. By identifying and quantifying the relationships between input variables and forecast uncertainty, we are better positioned to develop more accurate and reliable predictive models for ultraviolet radiation.

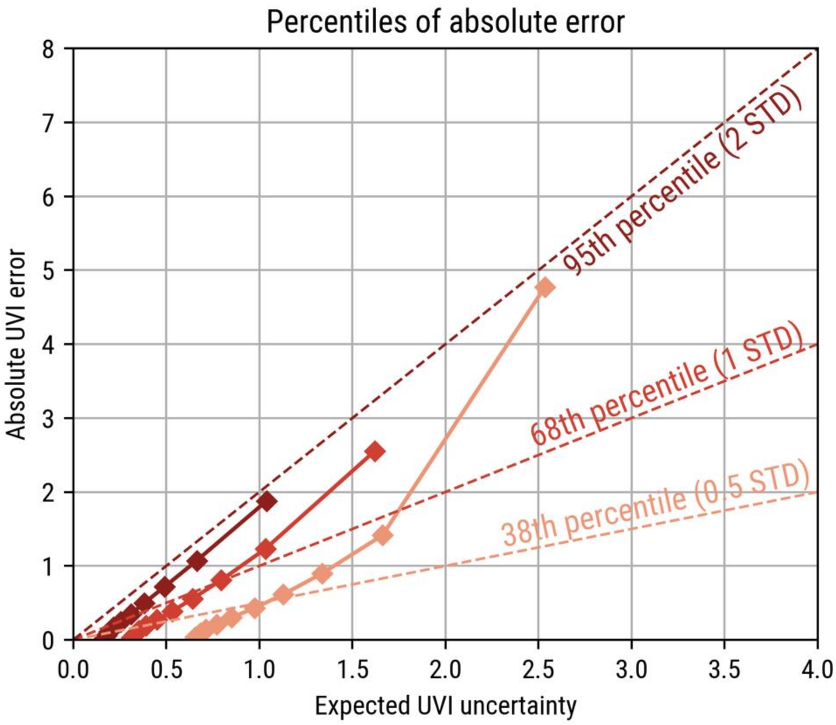

Figure 20. Absolute ultraviolet index (UVI) error as function of expected UVI uncertainty computed as the ensemble member spread. The UVI data was binned to 10 equally populated bins and the diamonds indicate different UVI bins. The lightest tone corresponds to 38th percentile (0.5 standard deviations), the middle tone corresponds to 68th percentile (1 standard deviation), and the darkest to the 95th percentile (2 standard deviations). The dashed lines show the theoretical ideal values for the uncertainties. The lines below the corresponding dashed lines indicate overestimated uncertainty estimates and the lines above the corresponding dashed lines indicate underestimated uncertainty estimates.

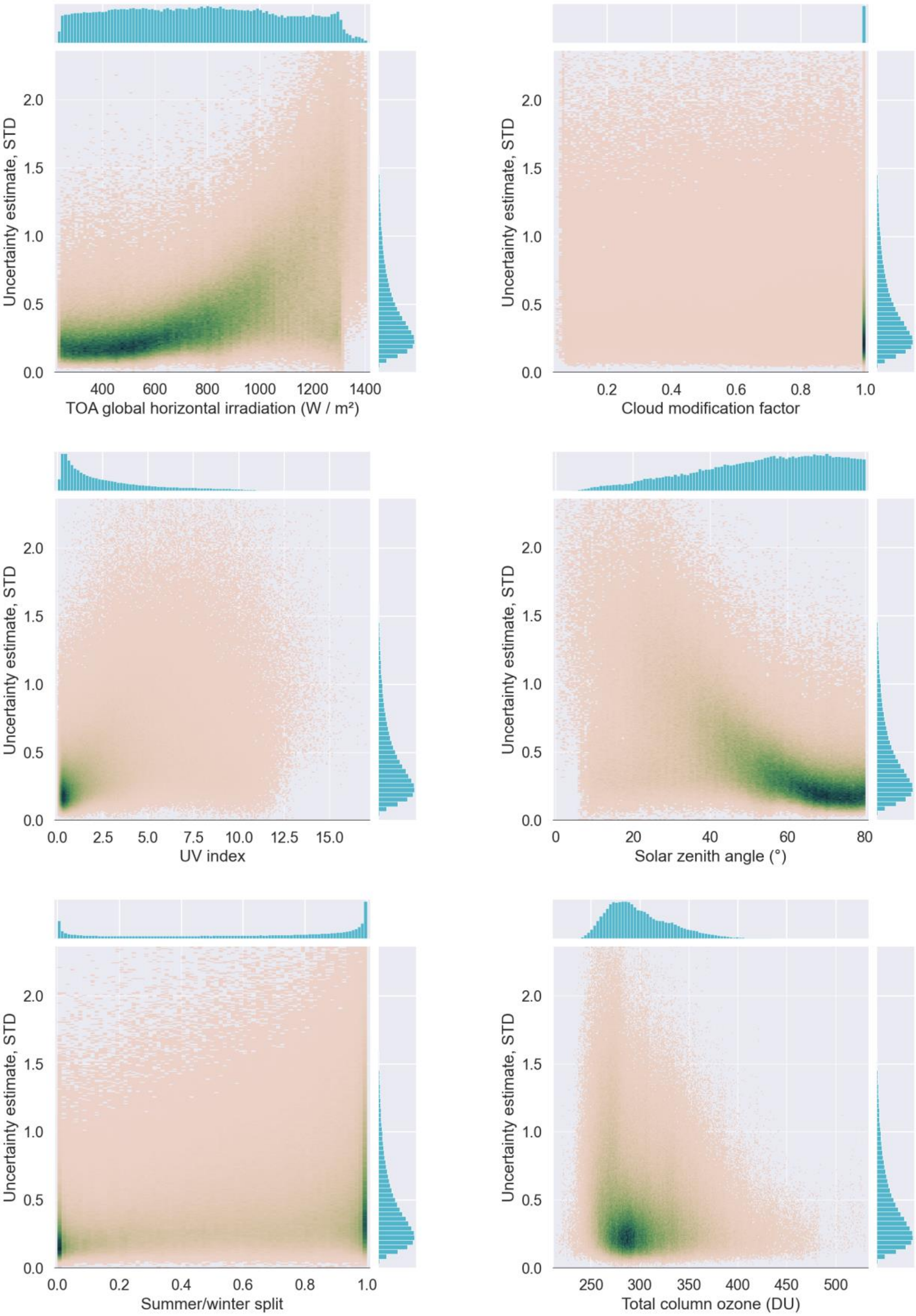

Figure 21. Uncertainty estimates for ultraviolet index (UVI) as function model input parameters. Top left: TOA global horizontal irradiation. Top right: Cloud modification factor. Middle left: UVI. Middle right: Solar zenith angle. Bottom left: Summer/winter split. Bottom right: Total column ozone.

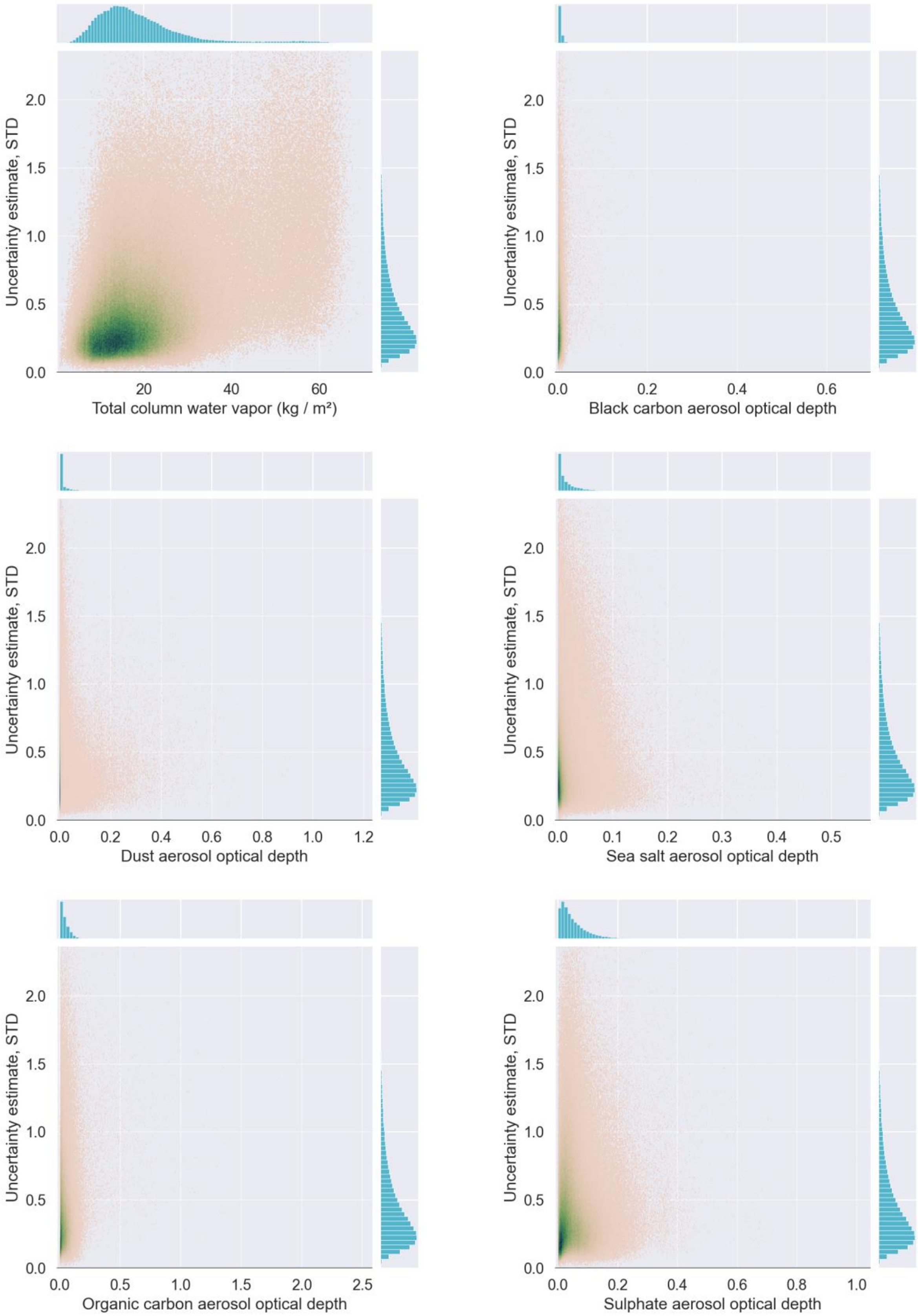

Figure 22. Uncertainty estimates for ultraviolet index (UVI) as function model input parameters. Top left: Total column water vapor. Top right: Black carbon aerosol optical depth. Middle left: Dust aerosol optical depth. Middle right: Sea salt aerosol optical depth. Bottom left: Organic carbon aerosol optical depth. Bottom right: Sulphate aerosol optical depth.

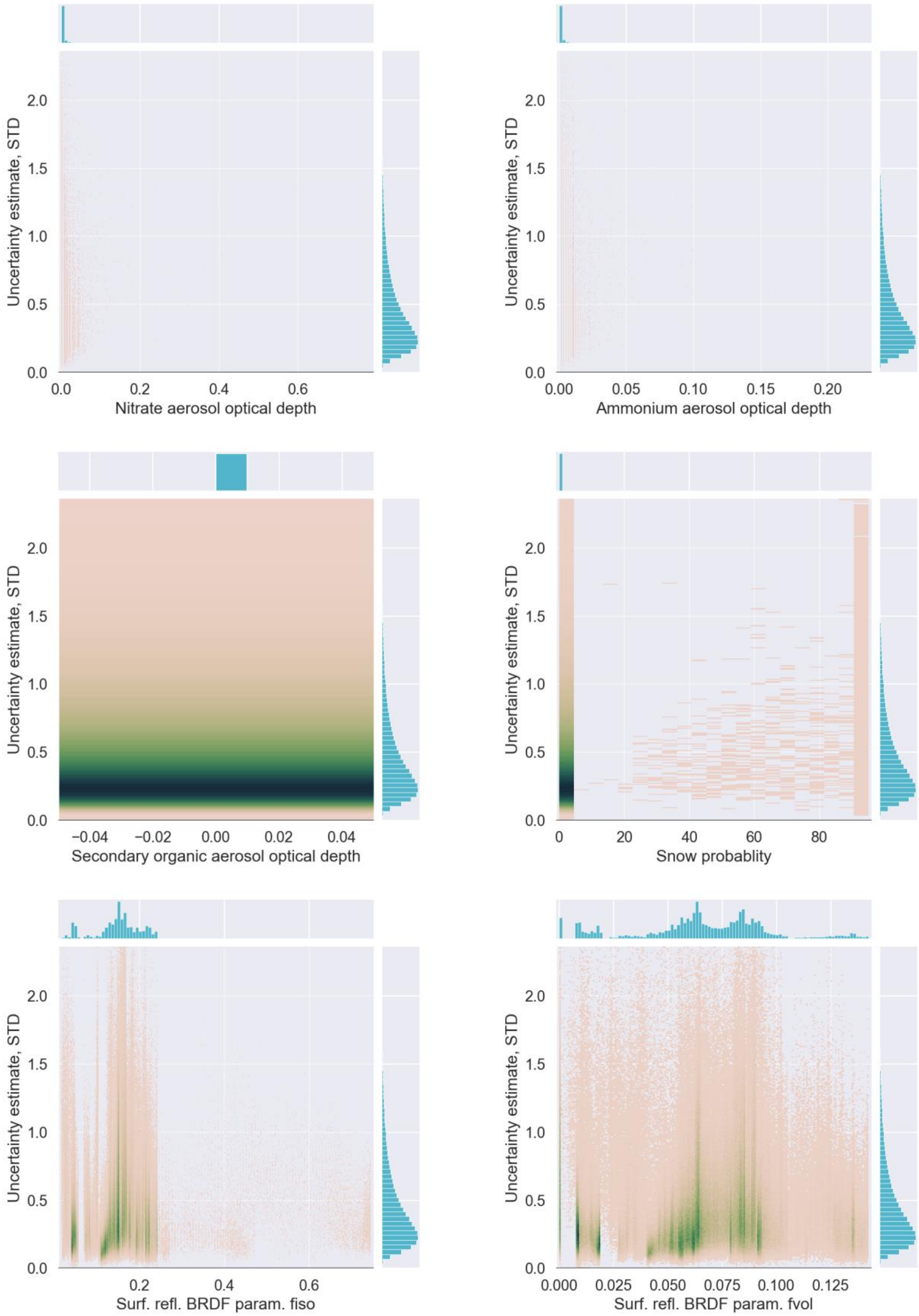

Figure 23. Uncertainty estimates for ultraviolet index (UVI) as function model input parameters. Top left: Nitrate aerosol optical depth. Top right: Ammonium aerosol optical depth. Middle left: Secondary organic aerosol optical depth. Middle right: Snow probability. Bottom left: Surface reflectance BRDF parameter fiso. Bottom right: Surface reflectance BRDF parameter fvol.

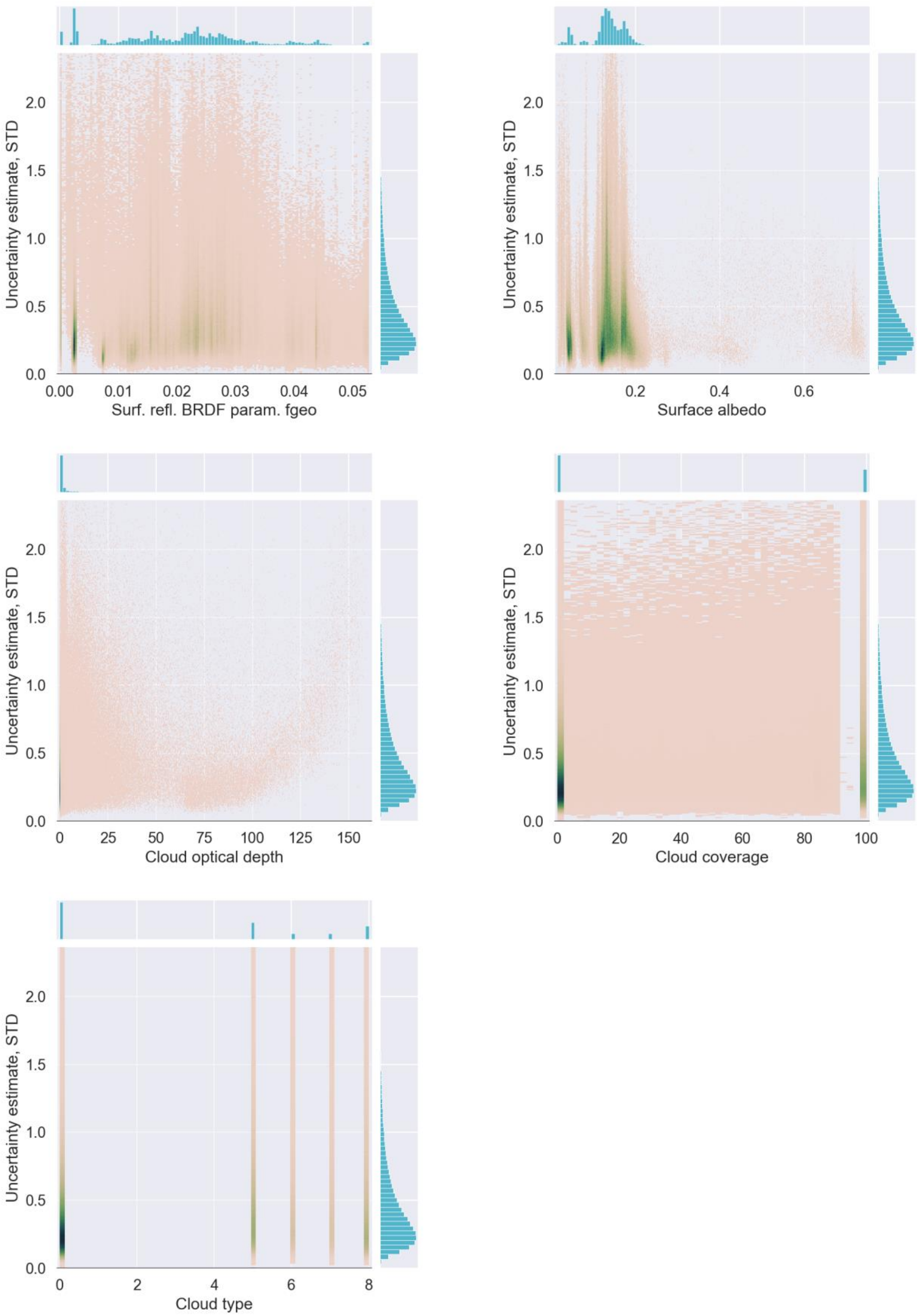

Figure 24. Uncertainty estimates for ultraviolet index (UVI) as function model input parameters. Top left: Surface reflectance BRDF parameter fgeo. Top right: Surface albedo. Middle left: Cloud optical depth. Middle right: Cloud coverage. Bottom left: Cloud type.

Similarly, as for the GHI, we employed the SHAP (SHapley Additive exPlanations) analysis to the UVI post-process correction model. This approach was harnessed to unveil the key features that predominantly influence the bias observed in the CAMS UVI data.

Figure 25 presents the SHAP diagram for a single ensemble member UVI post-process correction model. This diagram offers a comprehensive encapsulation of the SHAP values across the entire UVI dataset, providing a view of the impact of various features on the UVI correction model's predictions.

Figures 26-29 show the mean SHAP values corresponding to the full ensemble of models, arranged in the order of their significance. This ordered presentation highlights the hierarchy of influence among the input variables. Also, it facilitates a clear and structured understanding of the input parameter's individual and collective impact on the model's output.

The UVI itself emerged as the most influential factor affecting the post-process correction. A nuanced pattern was observed in the SHAP values corresponding to UVI: for lower UVI values (below 2.5), the SHAP values were predominantly positive, indicating a tendency for these values to increase the CAMS UVI output from its original estimate, pointing towards an underestimation by the model. Conversely, for higher UVI values (above 4), the SHAP values tended to be negative (mostly between -0.25 and -0.75), suggesting that the CAMS UVI is overestimated in these cases.

Following UVI, the top-of-atmosphere (TOA) GHI was the second most significant input variable. The relationship between the TOA GHI and SHAP values revealed that small TOA GHI values were associated with negative SHAP values (ranging between -0.25 and 0).

The cloud optical depth emerged as the third most critical input variable. Cloud optical depth values larger than 25 correspond to negative SHAP values ranging between -0.1 and -1.0 indicating correction to the negative direction and meaning possible overestimation of UVI.

We identified the solar zenith angle as the fourth most significant input variable. The SHAP analysis unveiled a negative correlation with the solar zenith angle, delineating a clear trend: smaller solar zenith angle values generally corresponded to positive SHAP values. In comparison, larger solar zenith angle values were typically associated with negative SHAP values (ranging between -0.2 and 0). This finding underscores the critical influence of the sun's elevation on the UVI predictions.

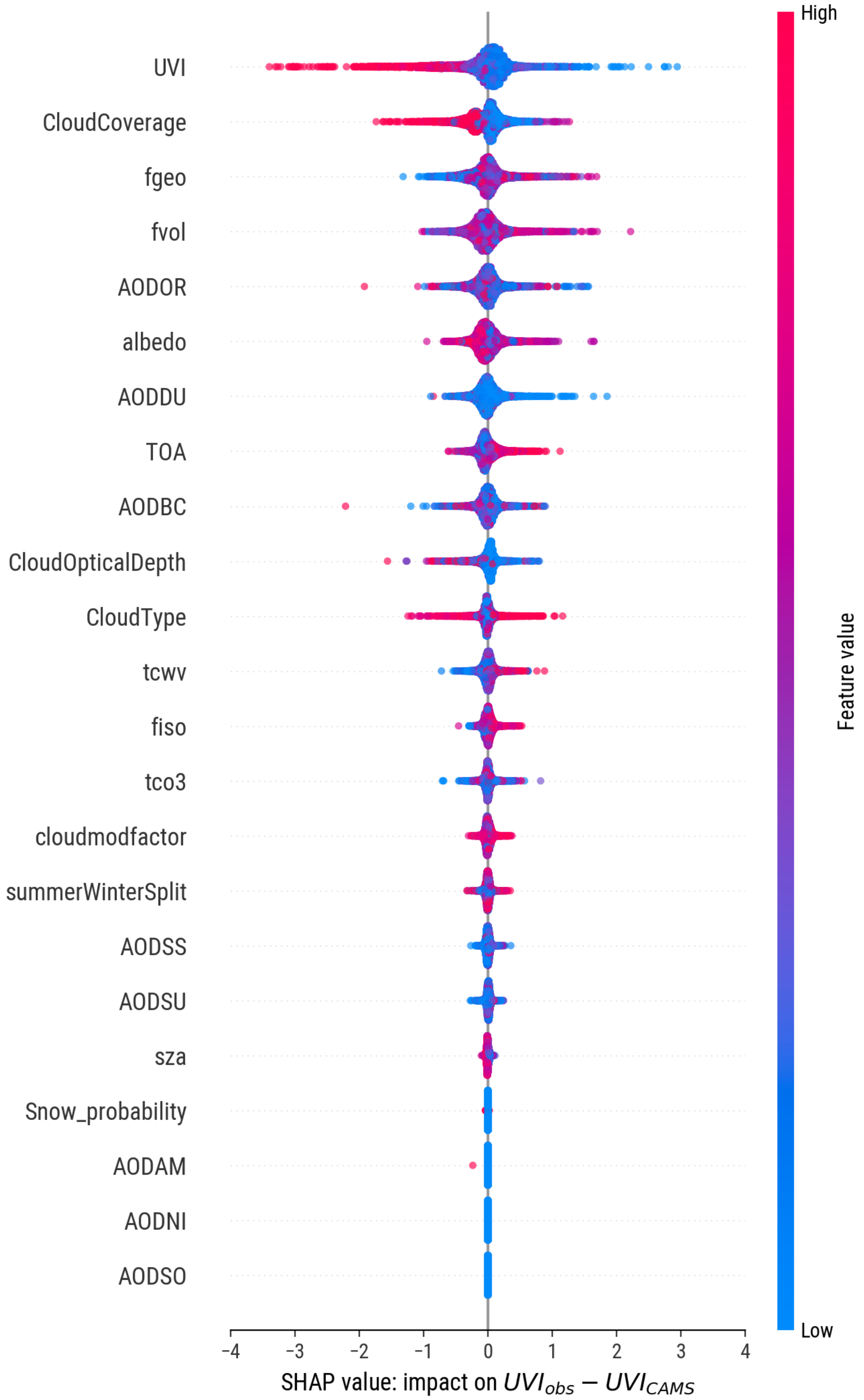

Figure 25. SHAP diagram for the ultraviolet index (UVI) error (UVI_{obs} - UVI_{CAMS}).

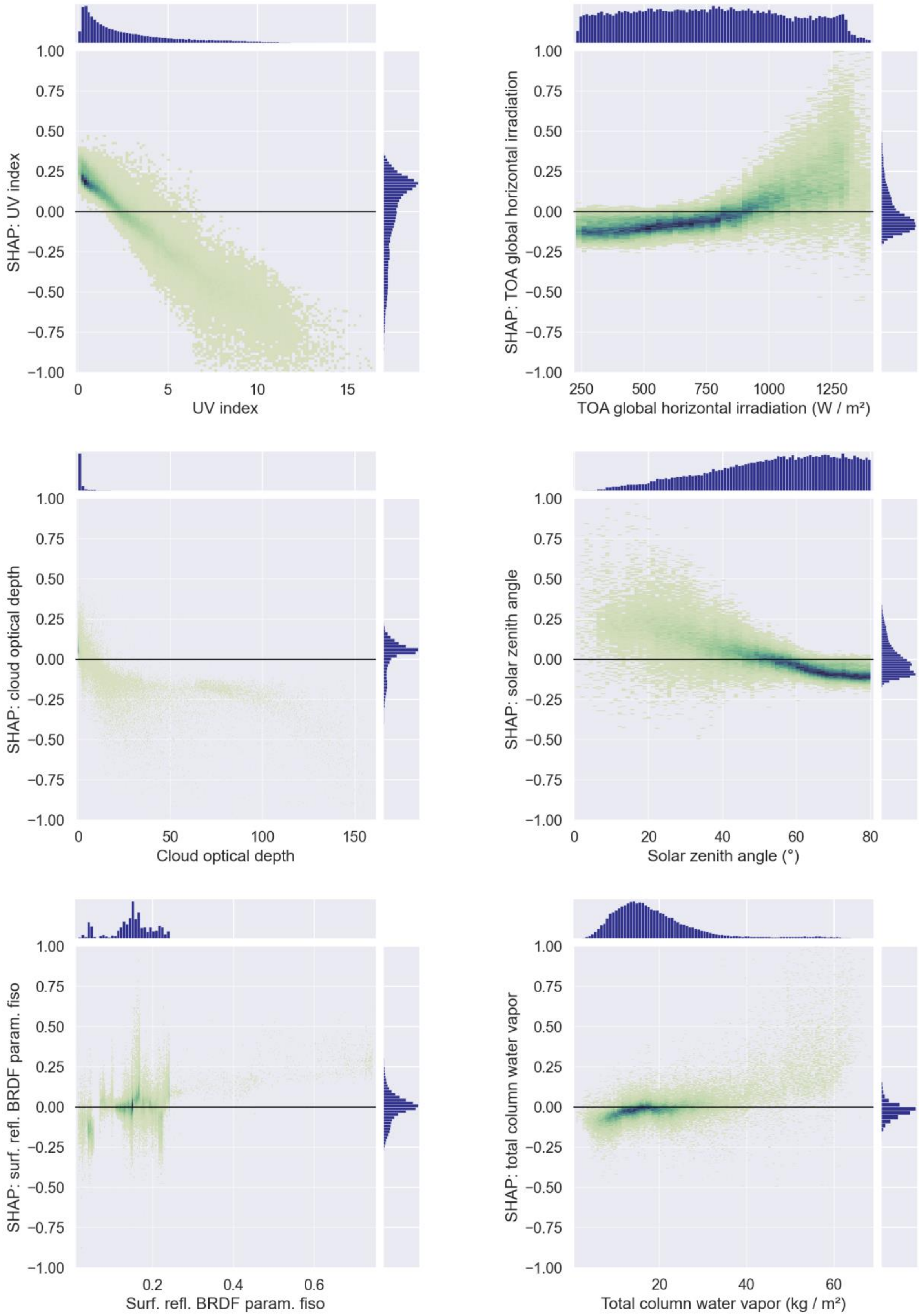

Figure 26. SHAP value for ultraviolet index (UVI) as function model input parameters. Top left: UVI. Top right: TOA global horizontal irradiation. Middle left: Cloud optical depth. Middle right: Solar zenith angle. Bottom left: Surface reflectance BRDF parameter fiso. Bottom right: Total column water vapor.

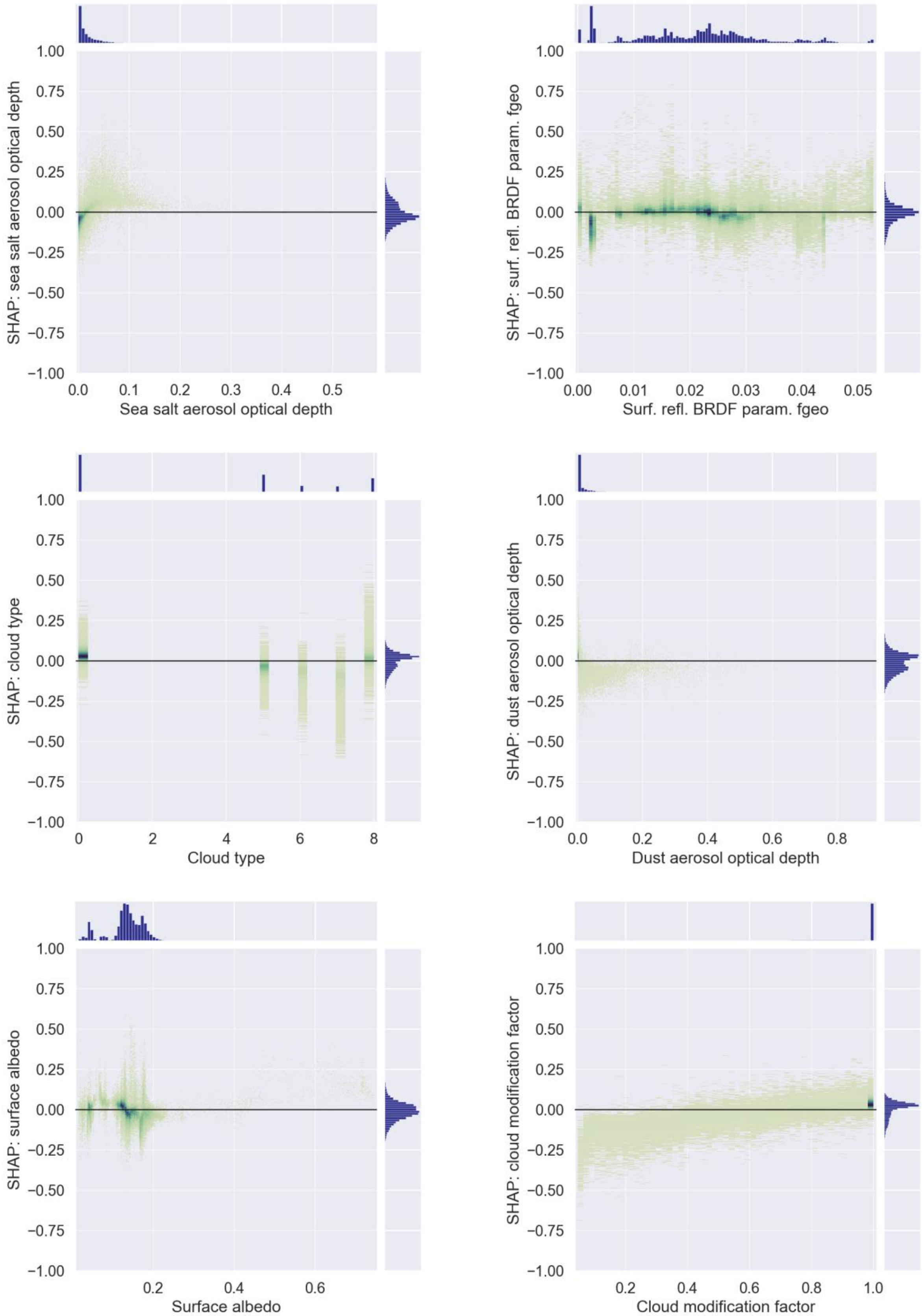

Figure 27. SHAP value for ultraviolet index (UVI) as function model input parameters. Top left: Sea salt aerosol optical depth. Top right: Surface reflectance BRDF parameter fgeo. Middle left: Cloud type. Middle right: Dust aerosol optical depth. Bottom left: Surface albedo. Bottom right: Cloud modification factor.

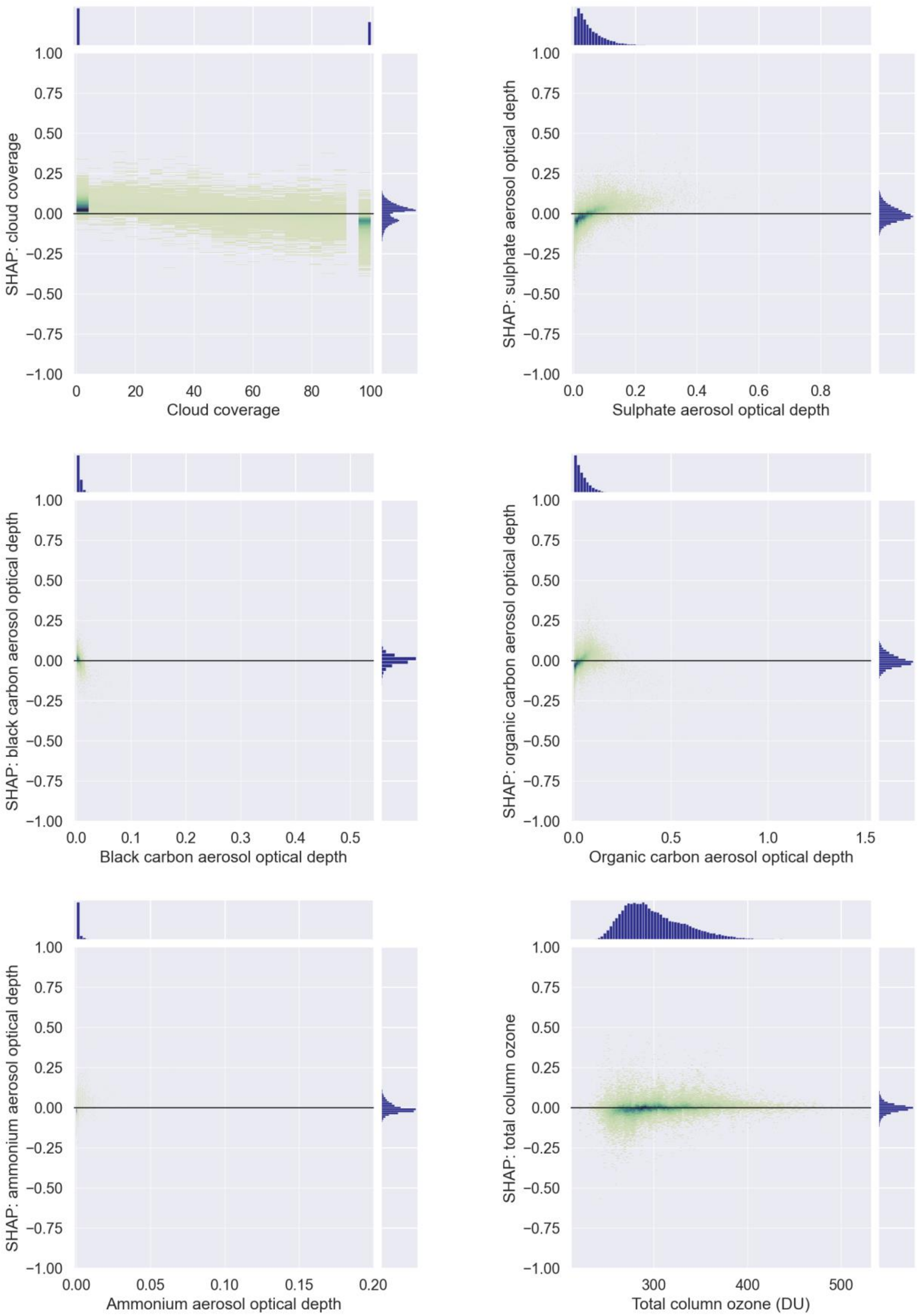

Figure 28. SHAP value for ultraviolet index (UVI) as function model input parameters. Top left: Cloud coverage.. Top right: Sulphate aerosol optical depth. Middle left: Black carbon aerosol optical depth. Middle right: Organic carbon aerosol optical depth. Bottom left: Ammonium aerosol optical depth. Bottom right: Total column ozone.

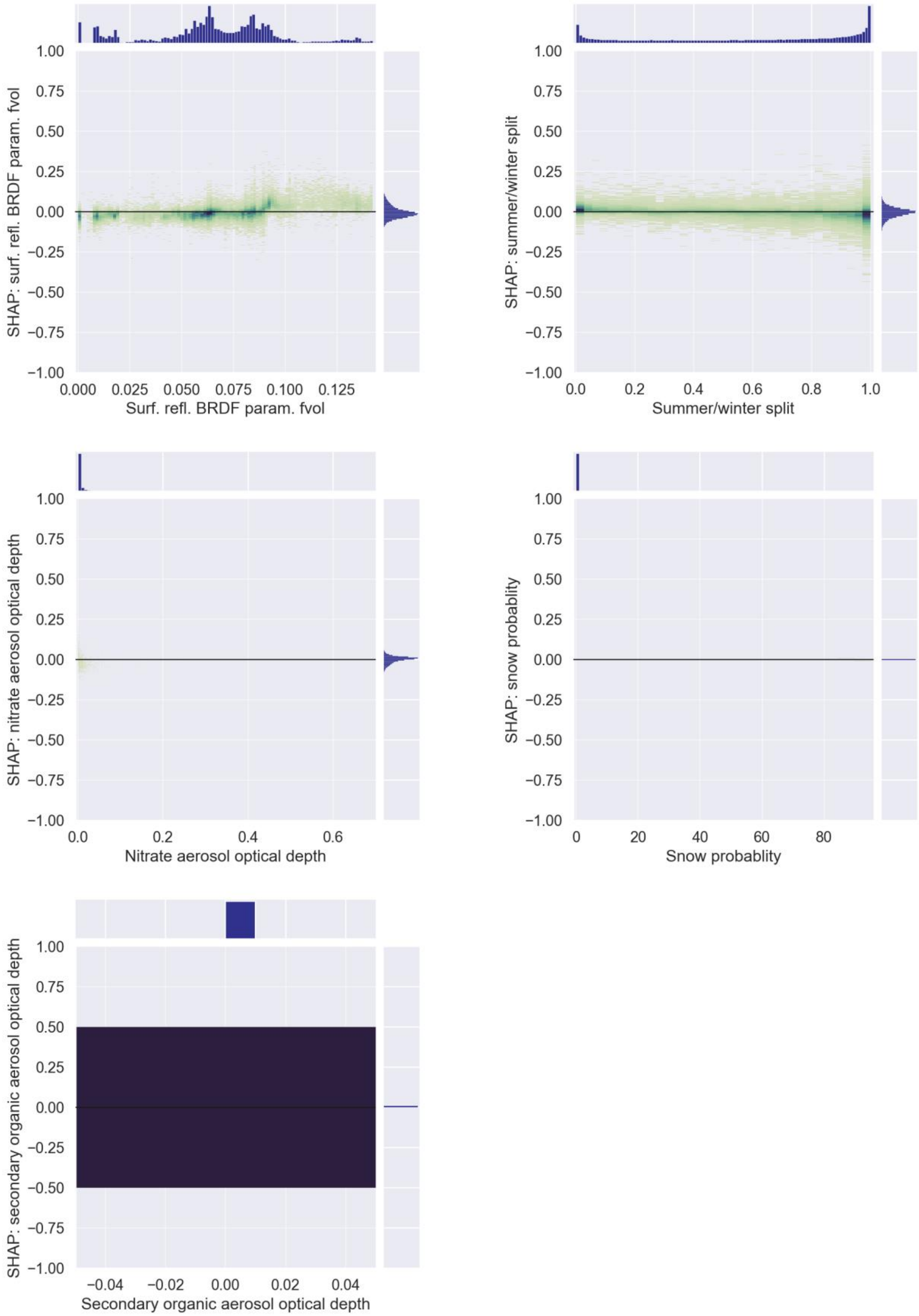

Figure 29. SHAP value for ultraviolet index (UVI) as function model input parameters. Top left: Surface reflectance BRDF parameter fvol. Top right: Summer/winter split. Middle left: Nitrate aerosol optical depth. Middle right: Snow probability. Bottom left: Secondary organic aerosol optical depth.

6 Conclusion

Our work with global and ultraviolet radiation simulation and forecast models has resulted in findings and methodological advancements that promise to enhance the accuracy, reliability and uncertainty quantification of radiation simulation and forecasting. To address the current inaccuracies in CAMS radiation products, we took the data-driven approach to analyse the factors contributing to the discrepancies between CAMS model data and ground-based observations. We took the post-process correction approach to correct the simulation outputs to match actual observations more closely.

Our methodology revolved around the post-process correction of global horizontal irradiation (GHI) and ultraviolet index (UVI) data. By applying machine learning techniques, specifically the XGBoost algorithm, we managed to identify and rectify the biases embedded within the CAMS simulation data. The synergy of XGBoost with SHAP (SHapley Additive exPlanations) analysis provided a transparent and quantifiable insight into how each input variable influenced the post-process correction model's predictions. Our approach enhanced the interpretability of complex models and facilitated a feature selection process.

A notable aspect of our work was the detailed analysis of error distribution in the post-process corrected data. The histograms and comparative analyses we conducted revealed an improvement in precision, as evidenced by the reduction in error standard deviation across various ranges of GHI and UVI. The results showed the efficacy of our post-process correction approach, which proved its merit not only in specific conditions but across a broad spectrum of radiation intensities.

Recognising the significance of uncertainty quantification in computational models, we explored a data-driven approach to assess the reliability of computed radiation data. We generated and evaluated uncertainty estimates using an ensemble technique, providing an understanding of the data's precision. This method also illuminated patterns in uncertainty distribution, offering valuable insights into the model's behaviour across different radiation intensity levels.

We leveraged the SHAP analysis to analyse the factors influencing the accuracy of the CAMS GHI and UVI data. This technique gave us information about the most significant features impacting the model's predictions. For instance, we uncovered that the cloud optical depth and cloud type were paramount in influencing the GHI predictions. At the same time, the UVI and top-of-atmosphere GHI emerged as critical factors in UVI post-process correction. These revelations deepened our understanding of the model's inner workings and guided us in refining our approach for even more accurate and reliable predictions.

Our work led to an increased understanding of the computational models for global and ultraviolet radiation and the information on the most important variables steering the model's accuracy. The methodologies developed, and insights obtained from this project will serve as a cornerstone for further advancements in radiation model development, ultimately benefiting many stakeholders, from policymakers to the general public seeking reliable environmental information.

7 References

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Fort, S., Hu, H., & Lakshminarayanan, B. (2019). Deep ensembles: A loss landscape perspective. *arXiv preprint arXiv:1912.02757*.

Gschwind, B., Wald, L., Blanc, P., Lefèvre, M., Schroedter-Homscheidt, M., & Arola, A. (2019). Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions–McClear-v3. *Meteorologische Zeitschrift*, *28*(2), 147-163.

Hamilton, S. J., Hänninen, A., Hauptmann, A., & Kolehmainen, V. (2019). Beltrami-net: domain-independent deep D-bar learning for absolute imaging with electrical impedance tomography (a-EIT). *Physiological measurement*, *40*(7), 074002.

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in neural information processing systems*, *30*.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., ... & Morcrette, J. J. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmospheric Measurement Techniques*, *6*(9), 2403-2418.

Lipponen, A., Huttunen, J. M., Romakkaniemi, S., Kokkola, H., & Kolehmainen, V. (2018). Correction of model reduction errors in simulations. *SIAM Journal on Scientific Computing*, *40*(1), B305-B327.

Lipponen, A., Kolehmainen, V., Kolmonen, P., Kukkurainen, A., Mielonen, T., Sabater, N., ... & Arola, A. (2021). Model-enforced post-process correction of satellite aerosol retrievals. *Atmospheric Measurement Techniques*, *14*(4), 2981-2992.

Lipponen, A., Reinvall, J., Väisänen, A., Taskinen, H., Lähivaara, T., Sogacheva, L., ... & Kolehmainen, V. (2022). Deep-learning-based post-process correction of the aerosol parameters in the high-resolution Sentinel-3 Level-2 Synergy product. *Atmospheric Measurement Techniques*, *15*(4), 895-914.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

McKinlay, A. F. (1987). A reference action spectrum for ultraviolet induced erythema in human skin. *CIE j*, *6*, 17-22.

Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., ... & Witek, M. L. (2020). A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing. *Atmospheric Measurement Techniques*, *13*(2), 373-404.

Schroedter-Homscheidt, M., Azam, F., Betcke, J., Hanrieder, N., Lefevre, M., Saboret, L., & Saint-Drenan, Y. M. (2022). Surface solar irradiation retrieval from MSG/SEVIRI based on APOLLO Next Generation and HELIOSAT-4 methods. *Meteorologische Zeitschrift*, *31*(6), 455-476.

Taskinen, H., Väisänen, A., Hatakka, L., Virtanen, T. H., Lähivaara, T., Arola, A., ... & Lipponen, A. (2022). High-Resolution Post-Process Corrected Satellite AOD. *Geophysical Research Letters*, *49*(18), e2022GL099733.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., ... & Wald, L. (2017). Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorologische Zeitschrift*, *26*(1), 33-57.

8 Appendix: Global horizontal irradiation and UV-index accuracy metrics by station

Table 4. Global horizontal irradiation 1-minute data accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Mean bias (W / m2)		**RMSE (W / m2)**		**MAE (W / m2)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
BSRN_BUD	5.0	**-1.5**	52.3	**47.8**	87.0	**78.6**
BSRN_CAB	-2.9	**-3.0**	53.8	**47.7**	81.8	**73.7**
BSRN_CNR	10.8	**3.2**	56.7	**49.0**	88.4	**77.5**
BSRN_FLO	4.1	**0.8**	72.1	**64.6**	119.2	**101.3**
BSRN_GOB	9.8	**0.3**	29.3	**18.2**	54.8	**42.3**
BSRN_INO	2.0	**-0.5**	49.2	**45.6**	84.3	**77.7**
BSRN_IZA	-23.8	**-7.1**	65.1	**36.9**	131.9	**70.9**
BSRN_LIN	-4.5	**0.7**	60.5	**53.9**	92.5	**83.3**
BSRN_PAL	-4.6	**-3.8**	57.4	**51.7**	91.4	**82.5**
BSRN_PAY	7.1	**2.7**	59.4	**48.6**	95.6	**78.6**
BSRN_RUN	-26.6	**-2.9**	107.9	**81.0**	172.6	**122.7**
BSRN_SON	-24.5	**-4.2**	129.5	**74.2**	189.0	**111.6**
SAURAN_CSIR	15.4	**2.9**	68.4	**59.4**	116.9	**101.9**
SAURAN_CUT	-4.1	**-2.6**	62.5	**56.0**	114.6	**100.8**
SAURAN_NMU	-10.8	**-1.2**	64.2	**55.3**	97.6	**84.8**
SAURAN_SUN	8.2	**-1.0**	52.5	**43.4**	93.6	**76.9**
SAURAN_UPR	19.2	**0.6**	61.0	**50.3**	99.4	**88.5**
enerMENA_ERF	28.2	**1.7**	59.3	**31.7**	84.0	**58.5**
enerMENA_JOR	-19.5	**-1.2**	33.0	**20.4**	63.5	**44.3**
enerMENA_TN	-3.1	**-0.7**	43.2	**31.5**	74.6	**57.9**
enerMENA_ZAG	23.1	**0.7**	53.3	**31.1**	81.8	**60.1**

Table 5. Global horizontal irradiation 1-minute data relative (pointwise) accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Relative bias (%)		**Relative RMSE (%)**		**Relative MAE (%)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
BSRN_BUD	13.4	**11.2**	28.7	**24.5**	72.0	**60.5**
BSRN_CAB	**9.6**	10.3	32.6	**26.9**	63.8	**53.7**
BSRN_CNR	17.7	**15.9**	34.0	**28.7**	**157.4**	214.1
BSRN_FLO	**16.8**	20.2	38.6	**36.7**	**109.4**	109.9
BSRN_GOB	2.1	**1.9**	7.5	**5.6**	**22.7**	25.7
BSRN_INO	**12.1**	12.3	28.4	**24.5**	77.3	**69.0**
BSRN_IZA	**-0.6**	1.6	14.2	**9.3**	33.2	**25.1**
BSRN_LIN	**9.4**	14.7	35.1	**30.7**	69.5	**63.4**
BSRN_PAL	**11.7**	12.7	32.0	**28.2**	86.9	**85.3**
BSRN_PAY	18.1	**14.9**	36.7	**28.3**	85.4	**69.7**
BSRN_RUN	**13.1**	13.9	37.2	**29.2**	87.9	**69.4**
BSRN_SON	**12.0**	13.2	51.0	**30.7**	120.5	**90.9**
SAURAN_CSIR	19.3	**13.5**	32.5	**25.0**	166.2	**84.9**
SAURAN_CUT	**13.0**	15.8	26.7	**26.4**	**119.8**	157.4
SAURAN_NMU	**3.1**	9.4	24.5	**23.0**	**46.9**	53.8
SAURAN_SUN	9.8	**7.6**	21.1	**17.9**	71.6	**54.8**
SAURAN_UPR	16.3	**11.7**	27.7	**22.5**	128.0	**103.1**
enerMENA_ERF	**11.5**	11.8	19.4	**18.4**	**290.0**	680.1
enerMENA_JOR	-2.7	**2.0**	8.2	**6.4**	**21.3**	34.6
enerMENA_TN	**-0.6**	3.3	12.6	**10.3**	**26.7**	29.0
enerMENA_ZAG	6.7	**2.9**	15.3	**9.7**	58.2	**30.7**

Table 6. Global horizontal irradiation 1-minute data relative (average, avg) accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Relative biasavg (%)		**Relative RMSEavg (%)**		**Relative MAEavg (%)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
BSRN_BUD	1.3	**-0.4**	23.1	**20.8**	13.9	**12.7**
BSRN_CAB	**-1.0**	**-1.0**	27.1	**24.5**	17.8	**15.8**
BSRN_CNR	2.8	**0.8**	22.8	**20.0**	14.6	**12.6**
BSRN_FLO	1.0	**0.2**	30.5	**25.9**	18.4	**16.5**
BSRN_GOB	1.5	**0.0**	8.7	**6.7**	4.6	**2.9**

BSRN_INO	0.5	**-0.1**	20.3	**18.7**	11.9	**11.0**
BSRN_IZA	-3.8	**-1.2**	21.3	**11.4**	10.5	**6.0**
BSRN_LIN	-1.5	**0.2**	30.5	**27.5**	19.9	**17.8**
BSRN_PAL	-1.4	**-1.1**	26.9	**24.3**	16.9	**15.2**
BSRN_PAY	2.0	**0.8**	26.7	**21.9**	16.6	**13.6**
BSRN_RUN	-5.2	**-0.6**	34.0	**24.2**	21.3	**16.0**
BSRN_SON	-6.3	**-1.1**	48.7	**28.8**	33.4	**19.1**
SAURAN_CSIR	3.0	**0.6**	23.0	**20.1**	13.5	**11.7**
SAURAN_CUT	-0.7	**-0.4**	20.0	**17.6**	10.9	**9.8**
SAURAN_NMU	-2.4	**-0.3**	21.7	**18.9**	14.3	**12.3**
SAURAN_SUN	1.6	**-0.2**	18.4	**15.1**	10.4	**8.5**
SAURAN_UPR	3.9	**0.1**	20.1	**17.9**	12.4	**10.2**
enerMENA_ERF	5.4	**0.3**	16.0	**11.1**	11.3	**6.0**
enerMENA_JOR	-3.2	**-0.2**	10.2	**7.1**	5.3	**3.3**
enerMENA_TN	-0.6	**-0.1**	14.1	**11.0**	8.2	**6.0**
enerMENA_ZAG	4.2	**0.1**	15.0	**11.1**	9.8	**5.7**

Table 7. Global horizontal irradiation 1-hour data accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Mean bias (W / m2)		**RMSE (W / m2)**		**MAE (W / m2)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
BSRN_BUD	5.1	**-1.2**	41.7	**35.4**	66.9	**57.2**
BSRN_CAB	**-2.6**	-3.0	40.9	**34.4**	61.1	**52.1**
BSRN_CNR	11.6	**3.6**	45.8	**36.4**	70.2	**57.3**
BSRN_FLO	4.3	**1.4**	59.0	**48.8**	96.0	**76.0**
BSRN_GOB	9.1	**0.1**	26.3	**14.2**	45.2	**31.4**
BSRN_INO	2.4	**-0.4**	38.4	**33.4**	61.9	**54.7**
BSRN_IZA	-22.0	**-6.0**	58.1	**28.7**	118.1	**53.4**
BSRN_LIN	-4.4	**0.6**	46.7	**39.4**	69.9	**59.7**
BSRN_PAL	-3.5	**-3.4**	43.6	**37.7**	67.7	**59.0**
BSRN_PAY	6.9	**2.5**	47.9	**36.3**	75.9	**57.9**
BSRN_RUN	-23.4	**-1.4**	92.7	**62.0**	146.5	**92.5**
BSRN_SON	-23.6	**-2.8**	117.8	**59.7**	171.5	**89.0**
SAURAN_CSIR	17.2	**4.4**	55.5	**44.3**	96.7	**77.5**
SAURAN_CUT	-1.7	**-1.6**	50.4	**42.3**	93.7	**75.8**
SAURAN_NMU	-7.6	**0.1**	52.4	**43.5**	78.4	**65.8**
SAURAN_SUN	8.2	**-1.3**	44.4	**33.4**	77.1	**58.3**
SAURAN_UPR	19.4	**1.0**	47.7	**35.3**	73.9	**60.6**
enerMENA_ERF	26.4	**1.5**	49.7	**23.4**	67.2	**41.5**
enerMENA_JOR	-18.8	**-0.9**	29.3	**16.0**	52.4	**32.6**
enerMENA_TN	-3.5	**-0.7**	36.0	**23.1**	57.2	**40.5**

enerMENA_ZAG	21.0	**0.5**	45.1	**23.4**	65.3	**43.7**

Table 8. Global horizontal irradiation 1-hour data relative (pointwise) accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Relative bias (%)		**Relative RMSE (%)**		**Relative MAE (%)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
BSRN_BUD	11.7	**9.1**	24.8	**19.9**	65.3	**55.3**
BSRN_CAB	7.2	**7.0**	26.4	**20.6**	49.6	**39.4**
BSRN_CNR	16.2	**11.8**	29.9	**22.0**	126.2	**74.6**
BSRN_FLO	**13.5**	16.5	33.0	**29.9**	**81.6**	85.9
BSRN_GOB	2.0	**1.4**	7.0	**4.6**	21.9	**16.1**
BSRN_INO	11.0	**9.5**	25.1	**19.6**	73.9	**49.1**
BSRN_IZA	**-0.6**	1.2	13.2	**7.7**	29.7	**20.9**
BSRN_LIN	**6.2**	10.8	28.8	**23.8**	54.6	**46.0**
BSRN_PAL	**9.6**	**9.6**	26.2	**22.2**	**70.7**	75.6
BSRN_PAY	16.9	**12.0**	32.2	**22.8**	80.9	**53.6**
BSRN_RUN	12.4	**11.8**	34.2	**24.3**	73.1	**56.0**
BSRN_SON	**9.2**	10.5	45.8	**24.7**	85.2	**55.6**
SAURAN_CSIR	15.1	**10.3**	25.9	**19.2**	97.8	**58.6**
SAURAN_CUT	**10.2**	11.1	21.6	**20.0**	**81.0**	87.4
SAURAN_NMU	**2.9**	7.6	20.8	**18.8**	**38.8**	41.0
SAURAN_SUN	8.4	**5.7**	17.9	**14.2**	55.1	**43.6**
SAURAN_UPR	12.4	**7.9**	21.4	**16.3**	84.0	**80.6**
enerMENA_ERF	**15.3**	22.5	**22.0**	27.9	**473.4**	1115.8
enerMENA_JOR	-3.2	**1.3**	7.4	**5.1**	**17.2**	17.5
enerMENA_TN	**-1.3**	2.7	11.1	**8.4**	**23.9**	30.9
enerMENA_ZAG	5.3	**1.8**	13.2	**7.7**	47.8	**22.8**

Table 9. Global horizontal irradiation 1-hour data relative (average, avg) accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Relative biasavg (%)		**Relative RMSEavg (%)**		**Relative MAEavg (%)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
BSRN_BUD	1.4	**-0.3**	18.5	**15.8**	11.5	**9.8**
BSRN_CAB	**-0.9**	-1.0	21.0	**17.9**	14.1	**11.8**
BSRN_CNR	3.1	**1.0**	18.8	**15.4**	12.3	**9.8**
BSRN_FLO	1.1	**0.4**	25.6	**20.2**	15.7	**13.0**
BSRN_GOB	1.5	**0.0**	7.4	**5.1**	4.3	**2.3**
BSRN_INO	0.6	**-0.1**	15.5	**13.7**	9.6	**8.4**
BSRN_IZA	-3.7	**-1.0**	19.7	**8.9**	9.7	**4.8**

BSRN_LIN	-1.5	**0.2**	24.0	**20.5**	16.0	**13.5**
BSRN_PAL	-1.1	**-1.0**	20.7	**18.1**	13.4	**11.6**
BSRN_PAY	2.0	**0.7**	22.1	**16.9**	13.9	**10.6**
BSRN_RUN	-4.8	**-0.3**	30.3	**19.1**	19.2	**12.8**
BSRN_SON	-6.3	**-0.8**	45.6	**23.7**	31.3	**15.9**
SAURAN_CSIR	3.5	**0.9**	19.8	**15.9**	11.4	**9.1**
SAURAN_CUT	**-0.3**	**-0.3**	17.0	**13.7**	9.1	**7.7**
SAURAN_NMU	-1.8	**0.0**	18.3	**15.4**	12.2	**10.1**
SAURAN_SUN	1.7	**-0.3**	15.6	**11.8**	9.0	**6.8**
SAURAN_UPR	4.1	**0.2**	15.6	**12.8**	10.1	**7.4**
enerMENA_ERF	5.2	**0.3**	13.3	**8.2**	9.8	**4.6**
enerMENA_JOR	-3.2	**-0.1**	8.8	**5.5**	4.9	**2.7**
enerMENA_TN	-0.7	**-0.1**	11.2	**8.0**	7.1	**4.5**
enerMENA_ZAG	4.0	**0.1**	12.6	**8.4**	8.7	**4.5**

Table 10. UV index 1-minute data accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Bias		**RMSE**		**MAE**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
Adelaide	0.2	**0.0**	0.6	**0.4**	1.0	**0.7**
Alicesprings	0.6	**0.0**	0.8	**0.5**	1.3	**0.9**
Bergen	**0.0**	**0.0**	0.3	**0.2**	0.5	**0.4**
Betdagan	0.3	**0.0**	0.5	**0.3**	0.7	**0.4**
Brisbane	0.3	**0.1**	0.7	**0.5**	1.0	**0.8**
Canberra	0.3	**0.1**	0.7	**0.5**	1.1	**0.8**
Chiangmai	-1.2	**-0.3**	1.6	**1.0**	2.4	**1.5**
Christchurch	0.1	**0.0**	0.6	**0.4**	1.0	**0.7**
Darwin	**0.0**	**0.0**	1.0	**0.8**	1.7	**1.3**
Eilat	0.5	**0.0**	0.5	**0.2**	0.7	**0.3**
Emerald	0.2	**0.0**	0.7	**0.5**	1.1	**0.9**
Finse	-0.2	**-0.1**	0.5	**0.3**	0.7	**0.5**
Florence	0.2	**0.0**	0.4	**0.3**	0.7	**0.5**
Goldcoast	**0.0**	**0.0**	0.7	**0.6**	1.2	**0.9**
Invercargill	**0.0**	**0.0**	0.6	**0.4**	0.9	**0.7**
Jerusalem	-0.3	**-0.1**	0.5	**0.3**	0.8	**0.5**
Kingston	**0.0**	**0.0**	0.5	**0.4**	0.9	**0.7**
Kise	**0.0**	**0.0**	**0.3**	**0.3**	0.6	**0.4**
Kjeller	**0.0**	**0.0**	0.4	**0.3**	**0.5**	**0.5**
Landvik	0.1	**0.0**	0.3	**0.2**	0.5	**0.4**
Macquarieisland	0.2	**0.0**	0.4	**0.3**	0.6	**0.5**
Melbourne	**0.0**	**0.0**	0.6	**0.4**	0.9	**0.7**

Nakhonpathom	-0.2	**0.1**	0.9	**0.8**	1.4	**1.1**
Newcastle	0.1	**0.0**	0.6	**0.5**	1.0	**0.7**
Osteras	**0.0**	**0.0**	**0.3**	**0.3**	0.5	**0.4**
Perth	0.3	**0.0**	0.6	**0.4**	0.9	**0.6**
Songkhla	**-0.1**	**0.1**	1.0	**0.8**	1.5	**1.2**
Sydney	-0.1	**0.0**	0.7	**0.5**	1.1	**0.9**
Townsville	0.1	**0.0**	0.8	**0.7**	1.3	**1.0**
Trondheim	-0.1	**0.0**	0.3	**0.2**	0.5	**0.4**
Ubonratchathani	-0.5	**-0.1**	1.1	**0.8**	1.7	**1.3**
Wellington	-0.1	**0.0**	0.7	**0.5**	1.2	**0.8**

Table 11. UV index 1-minute data relative (pointwise) accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Relative bias (%)		**Relative RMSE (%)**		**Relative MAE (%)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
Adelaide	8.5	**-0.1**	25.1	**20.8**	39.2	**32.6**
Alicesprings	13.2	**0.0**	22.4	**16.3**	34.4	**28.9**
Bergen	15.3	**4.4**	**34.7**	35.1	**57.4**	60.4
Betdagan	15.3	**3.8**	23.0	**14.8**	43.9	**30.4**
Brisbane	20.6	**9.4**	29.4	**23.5**	49.8	**42.3**
Canberra	24.9	**5.7**	36.8	**25.8**	62.4	**43.4**
Chiangmai	-11.1	**5.4**	34.9	**31.3**	**47.9**	63.1
Christchurch	20.7	**7.0**	40.2	**26.3**	69.6	**46.8**
Darwin	9.5	**7.5**	28.5	**24.2**	56.5	**45.8**
Eilat	19.8	**-0.1**	22.2	**10.3**	32.9	**19.0**
Emerald	10.7	**1.1**	21.9	**17.5**	35.1	**28.2**
Finse	**-1.6**	5.3	29.8	**26.3**	46.3	**44.8**
Florence	30.3	**8.9**	39.8	**28.5**	73.1	**52.2**
Goldcoast	12.1	**6.2**	27.9	**23.0**	47.5	**39.8**
Invercargill	25.4	**9.7**	46.0	**32.6**	79.2	**58.2**
Jerusalem	4.1	**3.9**	21.5	**15.6**	45.9	**39.4**
Kingston	13.1	**6.5**	30.6	**25.0**	51.1	**41.1**
Kise	14.2	**8.2**	33.4	**30.5**	57.6	**54.2**
Kjeller	15.3	**6.2**	34.5	**33.3**	**53.9**	54.4
Landvik	20.4	**8.5**	32.3	**30.0**	61.7	**60.1**
Macquarieisland	22.4	**2.5**	36.9	**35.5**	**54.5**	54.6
Melbourne	6.8	**-0.8**	28.7	**23.1**	47.3	**36.4**
Nakhonpathom	**5.3**	11.1	**29.3**	29.7	**44.5**	54.1

Newcastle	16.3	**7.8**	30.1	**23.7**	54.2	**44.0**
Osteras	20.2	**5.7**	35.7	**33.3**	65.2	**63.0**
Perth	12.4	**3.6**	21.9	**16.8**	33.3	**26.8**
Songkhla	15.3	**11.6**	35.5	**32.0**	74.1	**71.0**
Sydney	11.6	**8.9**	30.6	**25.8**	54.2	**46.3**
Townsville	7.6	**4.7**	24.2	**21.5**	**38.1**	38.7
Trondheim	**4.4**	5.8	29.3	**28.5**	**45.6**	47.4
Ubonratchathani	-1.4	**9.3**	29.5	**28.8**	**49.0**	88.6
Wellington	17.6	**8.4**	40.0	**29.4**	69.5	**54.6**

Table 12. UV index 1-minute data relative (average, avg) accuracy metrics mean bias, root-mean-squared-error (RMSE) and mean absolute error (MAE) for each station. The bold font indicates better performance. See the Appendix for the error metrics definitions.

	Relative biasavg (%)		**Relative RMSEavg (%)**		**Relative MAEavg (%)**	
Station	CAMS	corrected	CAMS	corrected	CAMS	corrected
Adelaide	4.8	**-0.4**	29.8	**21.7**	18.7	**13.4**
Alicesprings	12.0	**0.1**	27.7	**19.3**	18.0	**10.8**
Bergen	1.5	**-0.2**	39.1	**31.9**	23.8	**20.3**
Betdagan	8.9	**1.4**	22.5	**13.6**	15.1	**8.2**
Brisbane	8.2	**3.5**	30.3	**23.6**	19.0	**14.8**
Canberra	11.9	**1.8**	40.0	**28.7**	25.3	**17.5**
Chiangmai	-25.0	**-7.3**	51.8	**32.0**	33.7	**21.0**
Christchurch	3.0	**-0.3**	43.5	**29.0**	25.7	**16.6**
Darwin	0.9	**0.6**	33.4	**26.4**	21.2	**16.5**
Eilat	12.3	**0.1**	19.7	**9.1**	14.4	**5.7**
Emerald	4.5	**-0.7**	24.6	**19.4**	15.7	**12.0**
Finse	-12.1	**-4.3**	38.8	**28.6**	24.7	**17.8**
Florence	8.3	**2.2**	31.6	**23.8**	20.5	**14.8**
Goldcoast	0.7	**0.2**	32.4	**24.7**	19.9	**15.4**
Invercargill	1.2	**0.8**	49.4	**34.3**	29.1	**19.9**
Jerusalem	-7.2	**-1.3**	20.6	**13.6**	13.4	**7.8**
Kingston	**-0.2**	-0.3	37.4	**30.2**	22.3	**18.0**
Kise	-0.7	**0.6**	40.8	**30.8**	24.7	**19.5**
Kjeller	**0.3**	1.2	36.5	**30.5**	23.4	**20.3**
Landvik	3.7	**1.8**	31.7	**24.2**	18.6	**15.0**
Macquarieisland	15.7	**2.1**	46.9	**37.0**	28.8	**23.6**
Melbourne	1.2	**-1.1**	33.3	**25.0**	20.7	**15.8**
Nakhonpathom	-5.1	**1.7**	34.5	**27.2**	22.7	**18.5**
Newcastle	1.7	**0.3**	31.3	**23.6**	19.2	**14.5**

Osteras	1.6	**0.0**	36.7	**28.4**	22.1	**17.9**
Perth	7.1	**1.1**	24.0	**17.1**	15.6	**10.9**
Songkhla	-1.5	**1.2**	31.8	**25.9**	21.5	**17.9**
Sydney	-2.2	**-0.1**	36.4	**28.2**	21.6	**17.0**
Townsville	1.3	**0.6**	28.7	**23.0**	18.6	**14.6**
Trondheim	-5.7	**-0.8**	37.9	**29.5**	22.7	**18.7**
Ubonratchathani	-11.8	**-1.9**	37.2	**27.5**	24.0	**18.3**
Wellington	-2.6	**-0.5**	42.3	**29.7**	25.7	**17.3**

9 Appendix: Definitions of accuracy metrics

In the following formulas, j indicates the j^{th} value of the quantity of interest, n is the total number of values, Y is the quantity of interest (GHI or UVI), and $\overline{Y}$ denotes the average of Y.

Absolute metrics:

$$Bias = \frac{1}{n}\sum_{j=1}^{n} Y_{estimated(j)} - Y_{measured(j)}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(Y_{estimated(j)} - Y_{measured(j)}\right)^2}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|Y_{estimated(j)} - Y_{measured(j)}\right|$$

Relative metrics:

$$rBias_{avg} = \frac{100\%}{\overline{Y_{measured}}} Bias$$

$$rRMSE_{avg} = \frac{100\%}{\overline{Y_{measured}}} RMSE$$

$$rMAE_{avg} = \frac{100\%}{\overline{Y_{measured}}} MAE$$

$$rBias = 100\% \cdot \frac{1}{n}\sum_{j=1}^{n}\left(Y_{estimated(j)} - Y_{measured(j)}\right)/Y_{measured(j)}$$

$$rRMSE = 100\% \cdot \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left[\left(Y_{estimated(j)} - Y_{measured(j)}\right)/Y_{measured(j)}\right]^2}$$

$$rMAE = 100\% \cdot \frac{1}{n}\sum_{j=1}^{n}\left|\left(Y_{estimated(j)} - Y_{measured(j)}\right)/Y_{measured(j)}\right|$$

Document History

Version	Author(s)	Date	Changes
1.0	Antti Lipponen, Antti Arola	21/02/2024	Initial version

Internal Review History

Internal Reviewers	Date	Comments
Sam Remy (HYGEOS) and Vincent Guidard (MF)	Feb 2024	Internal review