Scientific Reports

https://doi.org/10.1038/s41598-026-40878-9

Article in Press

UV Index from ERA5 reanalysis

Received: 2 July 2025

Accepted: 16 February 2026

Published online: 10 March 2026

Cite this article as: Teggi S., Costanzini S., Despini F. *et al.* UV Index from ERA5 reanalysis. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-40878-9

Sergio Teggi, Sofia Costanzini, Francesca Despini, Mazzoli Riccardo & Tommaso Filippini

Title: UV Index from ERA5 reanalysis

Authors: Sergio Teggi[1,*], Sofia Costanzini[1], Francesca Despini[1], Mazzoli Riccardo[2], Tommaso Filippini[2],

[1]Department of Engineering Enzo Ferrari, University of Modena and Reggio Emilia, Via P. Vivarelli 10, Modena 41125, Italy

[2]Environmental, Genetic and Nutritional Epidemiology Research Center (CREAGEN), Section of Public Health, Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy

[*] Corresponding author, Sergio.teggi@unimore.it

Abstract

Exposure to ultraviolet (UV) radiation significantly impacts human health. Consequently, comprehensive UV climatological databases are of great interest. UV exposure is evaluated by weighting UV spectra with spectral functions that describe physiological responses at each wavelength. The most widely used function is the erythemal weighting function, which is used to compute the UV Index (UVI) to assess the health risk associated with UV overexposure. The ERA5 datasets, produced by the Copernicus Climate Change Service (CDS), offer hourly ground-level UV radiation (UV_{E5}), but do not include UVI. This study proposes a model to compute hourly UVI using exclusively ERA5 data, enabling direct access through the CDS to derive UVI statistics for locations of interest and potentially supporting the integration of a dedicated UVI product into ERA5. The model was developed using UV spectra simulated under clear-sky conditions with the uvspec radiative transfer model, accounting for atmosphere type, solar zenith angle, visibility, altitude, albedo, total ozone, and aerosol type. For these parameters, representative values typical of tropical, mid-latitude, and subarctic regions were used, effectively excluding Arctic conditions, and considering UVI values ≤ 12. The resulting formulation expresses UVI as a function of UV_{E5}, sun elevation, and total ozone. The model was validated using ground-based UVI measurements from six stations (over 17000 cases) and, in addition, compared with UVI derived from Copernicus Atmospheric Monitoring Service (CAMS) products (over 6000 cases) taken as a reference. Performance was assessed through the statistics of the differences between measured/modelled values and CAMS data under three scenarios: clear-sky conditions, varying cloud cover, and all-sky conditions. Under clear-sky conditions, the model uncertainties showed a small positive bias (≤ 0.5), with the absolute difference (AD) < 1 in 73% of the cases for ground measurements and in 86% of the cases for CAMS. The root mean squared difference (RMS) and the mean absolute deviation (MAD) were 0.9 and 0.7, respectively, for ground measurements, and 0.7 and 0.6 for CAMS. Under cloudy-sky conditions, model performance worsens significantly for CC > 0.4, with RMS and MAD reaching values of about 1.5. However, when considering relative uncertainties (percentage ratio between RMS and reference values of UVI), up to CC < 0.7 the RMS% remains below 15% for Very High-to-Extreme WHO/ICNIRP exposure categories and below 20% for Moderate-to-Extreme categories. A comparison between CAMS UVI and ground measurements was also performed, yielding results consistent with those described above. As an example, Appendix A illustrates how the model can be applied to generate daily and monthly UVI

statistics over large geographical areas using only ERA5 data accessed through the CDS web portal.

1 Introduction

The overexposure to solar UV is a risk to public health. Common harmful effects are eye conditions, such as pterygiums, photokeratitis, and cataracts, immune suppression, photoaging, DNA damage, and skin cancer [1–6]. At the same time, UV exposure can have beneficial health effects, such as enabling the synthesis of vitamin D, essential for general well-being [7–10], and reducing the transmission of many aerodispersed viruses, including SARS-CoV-2 [11–15]. One of the actions to address the problem of overexposure to UV radiation was the definition of the global solar UV-Index (UVI), which is primarily aimed at informing people about the need to adopt protective measures when exposed to UV radiation. UVI was defined by the World Health Organization (WHO) in collaboration with the World Meteorological Organization (WMO), United Nations Environment Programme (UNEP), and International Commission on Non-Ionizing Radiation Protection (ICNIRP) [16–18], and is in operational use in more than 100 countries worldwide [19,20]. Due to the considerable interest, the number of ground station networks, of satellite platforms for UV[21] observation and global UV datasets and UV radiation level forecast models is continuously growing, as confirmed by the high number of scientific publications available in the literature.

The ERA5 dataset [22] generated by the Copernicus Climate Change Service, which is part of the Earth monitoring programme of the European Union and implemented by ECMWF, is among the most used climatological datasets in the world in many research fields, including UV solar radiation [14,23–28]. The “ERA5 hourly data on single levels from 1940 to present” dataset, hereinafter referred to as ERA5, is a reanalysis of hourly meteorological conditions from 1940 generated globally at 0.25° x 0.25° of spatial resolution ($\simeq 27$ km at the equator). The dataset is composed by many meteorological and geophysical quantities including the downward UV radiation in the 200 nm to 440 nm range at ground level (UV_{E5}; note that, following international standards, the UV region is strictly defined as 100–400 nm) but not including UVI. UV “biologically effective dose” (UV erythemal dose, conversion to UVI is trivial [29]) gridded maps are also provided by the Integrated Forecasting System of the Copernicus Atmospheric Monitoring Service (CAMS, atmosphere.copernicus.eu) using models implemented for clear-sky conditions and for all-sky conditions. Their grid step is 0.7° x 0.7° ($\simeq 76$ km at the equator) before June 2016 and 0.4° x 0.4° ($\simeq 43$ km at the equator) afterward and are forecast data and not reanalysis.

The definition of algorithms for the computation of UVI using ERA5 gridded quantities is therefore of great interest to produce UVI global climatology based on UVI gridded maps. The transformation of broad-band UV radiation, as UV_{E5}, to UVI is not straightforward since it is not linear and depends on erythemal action spectrum and on other quantities as cloud cover fraction and type, atmospheric ozone and aerosol content and solar elevation [20,23,30,31]. Moreover, there

are very few studies reported in literature focusing on this problem and this is the main objective of this work.

This work proposes a model for computing *UVI* using data exclusively from the ERA5 dataset. The model is formulated for Clear Sky Conditions (CSC) using UV ground radiation simulated by the *uvspec* (*LibRadtran*) radiative transfer model [38,39]. It is then compared with ground measurements and CAMS UVI, first under CSC and subsequently under cloudy-sky conditions, to evaluate its applicability under cloudy skies.

The manuscript is organized as follows: description of the quantities used; simulation of UVI and UV_{E5} at ground level under CSC using *uvspec*; formulation of the model; comparisons of the model with ground measurements and with CAMS data under both CSC and cluody-sky conditions; discussion of the results and the conclusions. Finally, Appendix A presents an example of a Python procedure that interfaces directly with the Copernicus Climate Center for the calculation of multitemporal UVI maps in NetCDF and CSV formats.

2 Materials and methods

2.1 Radiometric quantities

ERA5 reanalysis product includes the hourly (accumulated) downward UV radiation at the surface in $J\ m^{-2}$ in the electromagnetic spectrum range $200\ nm \div 440\ nm$. However, the UV radiation reaching the surface is in the range of $280\ nm \div 440\ nm$ due to the strong absorption of the atmospheric ozone layer. This quantity is transformed in hourly (mean) irradiance (UV_{E5}, $W\ m^{-2}$) by dividing it by $3600\ s$. The erythemally weighted radiation UV_{Ery} ($W\ m^{-2}$) and the UV Index UVI are defined as [16–18,20,33]:

$$UV_{Ery} = \int_0^{\infty} I(\lambda)\Phi_{Ery}(\lambda)\, d\lambda \quad (1)$$

$$UVI = UV_{Ery} \cdot k_0 \quad (2)$$

where λ is the wavelength in nm, $I(\lambda)$ is the solar spectral irradiance reaching the ground in $W\ m^{-2}$, and $\Phi_{er}(\lambda)$ is the erythemal weighting function (dimensionless), or sun burning action spectrum, that describes the effectiveness of different wavelengths of UV radiation in causing erythema. $k_0 = 40\ m^2\ W^{-1}$ is a constant[16]. The relation above can be used to define other quantities, such as doses, or indexes by changing the weighting function, such as those used to describe DNA damage and vitamin D synthesis (Figure 1) [29,34]. Considering that Φ_{Ery} is zero for $\lambda > 400\ nm$ and that the atmosphere absorbs completely the UV radiation for $\lambda < 280\ nm$, the integral in equation 1 can be limited in the range $280\ nm \div 400\ nm$.

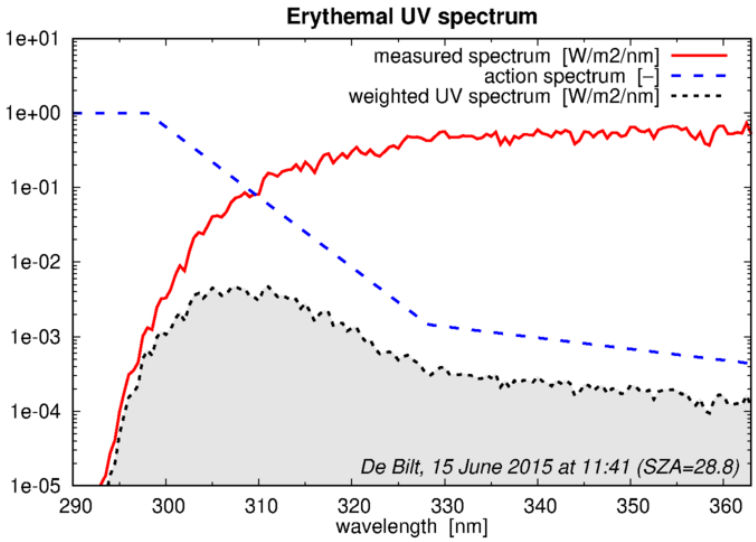


Figure 1. Erythemal weighting function [29]

2.2 Main assumption

UV radiation reaching the ground is influenced by many factors, such as atmospheric absorption and diffusion, ground altitude, surface albedo, sun elevation, day of the year, and clouds. The importance of these factors has been extensively reported in literature [20,23,30,35–37]. The model presented here assumes that most of these factors affect UVI and UV_{E5} in a very similar way. Then, the first step of this work was the study of the ratio $R_0 = UVI/UV_{E5}$ to individuate the relevant factors for the UVI vs UV_{E5} relationship. In the second step, this relationship was defined using linear regression analysis and simulations carried out using radiative transfer modelling.

2.3 Simulations and analysis

Simulated values of UV_{E5}, UVI, and UV_{Ery}, were generated using the radiative transfer code *uvspec* included in the *LibRadtran* library [38,39]. Given the difficulties in modelling radiative transfer under cloudy sky conditions [40–44] the simulations were limited to clear sky conditions. *uvspec* was run for many different cases defined by the set of input parameters[39,45] listed in Table 1. Some of these parameters have a small effect on the results of these simulations, so they were kept constant (fixed parameters) for all runs. The other parameters were changed for each run (variable parameters) according to the values reported in Table 1. The selection of fixed or variable parameters was based on preliminary runs and on the literature cited in 2.2. Assuming the reference case labelled “ref.” in Table 1, the variable parameters were changed one at a time. For example, for “Atmosphere type” six simulations were run by varying its value according to Table 1, while the other parameters were set to the reference values. Since the aim of this first set of simulations was the assessment of the importance of the variable parameters, they were varied considering extreme values (very infrequent in some cases). *uvspec* was used in the plan-parallel atmosphere assumption, so, the Sun zenith angle was limited to 60° to avoid relevant errors due to atmospheric refraction. Moreover, *UVI* is generally very low at a high sun zenith angle, reducing its concerns for population health and its interest in this study.

As an example, Figure 2 shows the spectral irradiance (black), the weighted spectral irradiance (red), and the integrated values UV_{E5}, UV_{Ery} and UVI obtained from a *uvspec* output.

Table 1. Scheme of *uvspec* input parameters used for the simulations (refer to[39,45] for the complete description of the parameters)

Fixed parameters fixed for all simulation
Wavelength range: 280 nm - 440 nm
Solar spectrum: *Atlas plus modtran spectrum* convolved with a triangular function (1 nm resolution)
Radiative transfer model: DISORT (discrete–ordinate–method radiative transfer) [46]
Molecular adsorption parameterization: *reptran*, based on HITRAN 2004 database [47] (resolution 15 cm^{-1})
Cloud cover fraction: 0 (clear sky conditions)
Aerosol profile above 2 km: set to moderate volcanic aerosol
Aerosol seasonality: set to spring-summer
Day of the year: 180
Variable parameters
Atmosphere type: midlatitude summer (MS), midlatitude winter (MW), subarctic summer (SS), subarctic winter (SW), tropics (TR), US-Standard (US) (ref. midlatitude summer)
Aerosol profile 0-2 km: rural (R), maritime (M), urban (U), tropospheric (T) (ref. rural)
Visibility: set to 5 km, 10 km, 15 km, 25 km, 50 km, 100 km (ref. 100 km)
Ground altitude: 0 to 4 km in 0.5 km increments (ref. 0 km)
Total columnar ozone content: set to 250 DU to 450 DU in 50 DU increments (ref. 250 DU)
Solar zenith angle: set to 0° to 60° in 10° increments (ref. 0°)
Surface albedo: 0 to 1 in 0.2 increments (ref. 0)

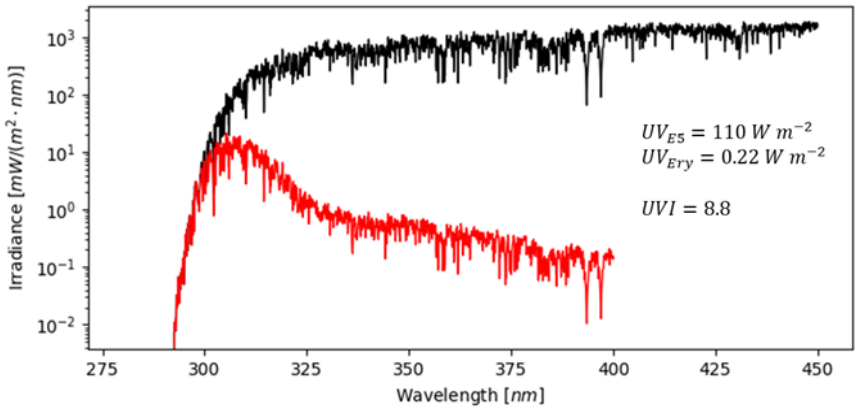


Figure 2. *uvspec* simulation: black line: $UV_{E5}(\lambda)$, red line: $UV_{Ery}(\lambda)$

To study the influence of the variable parameters, for each simulation the following ratios were computed:

$$R_0 = \frac{UVI}{UV_{E5}},\ R_1 = \frac{UVI}{UVI_{ref}},\ R_2 = \frac{UV_{E5}}{UV_{E5,ref}},\qquad R_3 = \frac{R_1}{R_2} = R_0\frac{UV_{E5,ref}}{UVI_{ref}} \tag{3}$$

where the subscript "*ref*" stands for the value obtained for the reference case. Figure 3 shows the result of this analysis. The reference case corresponds to the first value of the abscissa. It can be observed that:

- Atmosphere type: its influence on UV radiation is practically nil.
- Aerosol type: Urban aerosol is the only case where significant variations of R_1 and R_2 are observable. Nevertheless, R_3 is always very close to 1.
- Visibility: its influence is significant below 15 km. R_3 is weakly influenced and reaches 0.9 for very low visibility (5 km).
- Ozone: both R_3 and R_1 are strongly influenced by this factor.
- Sun zenith angle: all ratios are strongly influenced by this factor.
- Ground altitude: its influence is significant above 2 km. R_3 is poorly influenced and reaches 1.1 at 4 km.
- Albedo: both R_1 and R_2 are strongly influenced by this factor. Nevertheless, R_3 is weakly affected, for example, for albedo equal to 0.8 (typical value for snow cover) R_3 is close to 1.

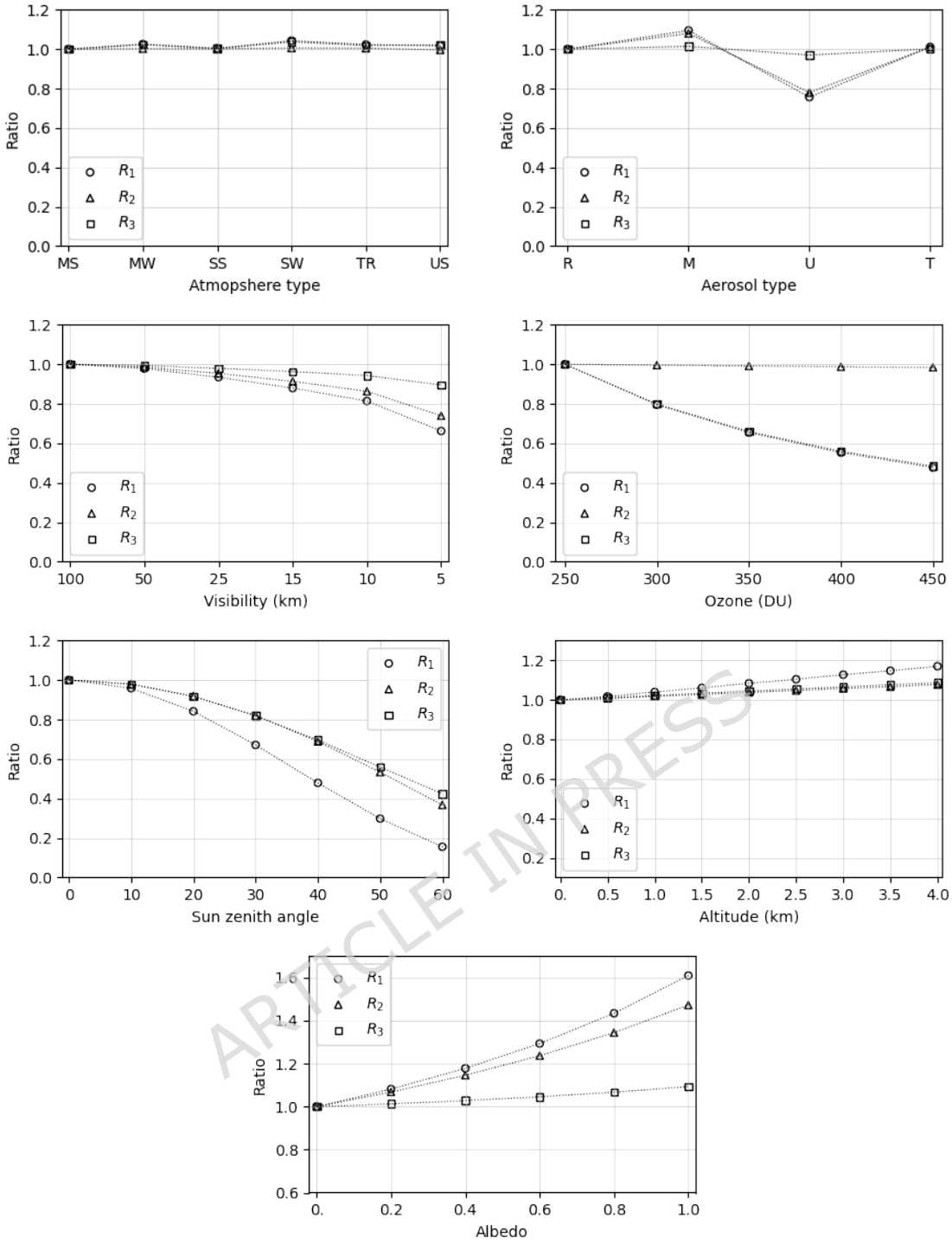


Figure 3. R_1, R_2 and R_3 ratios (Eq.(3))

Based on these results, a new set of simulations aimed at the definition of the UVI vs UV_{E5} relationship was run. Since the atmosphere type and the aerosol type do not have a significant effect on R_0, they were fixed to “US-standard” and to “Rural” respectively to avoid regression overfitting. Analogously, given that visibility, altitude, and albedo have a weak effect on R_3, only two values for each of them were used: 10 km and 50 km for visibility (very turbid and very clear atmosphere excluded), 0 km and 2 km for altitude (considering the ERA5 grid step, average altitude over 2 km are infrequent), 0.05 and 0.5 for albedo (values of 0 or greater than 0.5 are very infrequent [48,49]). To increase the number of cases used in the regression analysis, each

combination of the input parameters was considered (280 simulations). Some simulations returned very high UVI values, up to 20 units, which occur very rarely and in particular areas, such as in the case reported Cordero et al.[50]. The simulations dataset was limited to $UVI \leq 12$, which was assumed to be an extreme value for most inhabited mid-latitude regions[51–53], reducing its size to 241 simulations.

2.4 Model definition

The analysis of the second set of simulations indicates that the relevant factors driving R_0 are the sun zenith angle (θ_s) and the total ozone columnar content ($O3$), and that for both factors the dependence is not linear. To model the UV radiation at the ground, several studies reported in the literature[30,32,54] considered the ratio $R = \mu/O3$, where $\mu = \cos\theta_s$ describes the atmospheric path length of solar radiation. For example, Allaart et al.[32] used observations from a Brewer spectrophotometer to develop a predictive UVI model that depends solely on the Sun-Earth distance and the R ratio.

The correlation between R_0 and the quantities θ_s, μ, $O3$, and the simple ratio $R_* = \mu/O3$ was evaluated using Pearson's linear correlation coefficient (r), Spearman's rank correlation coefficient (ρ), and Kendall's Tau correlation coefficient (τ), which resulted (p-values $< 10^{-4}$ in all the cases):

$$r(R_0,\theta_s) = -0.63\ ,\quad r(R_0,\mu) = 0.68\ ,\quad r(R_0,O3) = -0.54\ ,\quad r(R_0,R_*) = 0.99$$
$$\rho(R_0,\theta_s) = -0.68\ ,\ \rho(R_0,\mu) = 0.68\ ,\quad \rho(R_0,O3) = -0.51\ ,\quad \rho(R_0,R_*) = 0.99$$
$$\tau(R_0,\theta_s) = -0.51\ ,\quad \tau(R_0,\mu) = 0.51\ ,\quad \tau(R_0,O3) = -0.39\ ,\quad \tau(R_0,R_*) = 0.92$$

These values indicate that R_* is the quantity that is best linearly correlated with R_0.

Figure 4 shows the results of the linear regression:

$$\widehat{R_0} = a_0 + a_1 \cdot R_* \qquad (4)$$

and of the relation resulting from the combination of Eq.(3) and Eq.(4):

$$\widehat{UVI} = UV_{E5}(a_0 + a_1 \cdot R_*) \qquad (5)$$

where the *hat* symbol ($\widehat{\ }$) indicates the value predicted by the regression (from here on also referred to as the *model*), $a_0 = -0.01\ m^2\ W^{-1}$ and $a_1 = 35.89\ m^2\ W^{-1}DU$ are the regression coefficients, and R^2 is the determination coefficient of the regression.

To evaluate the goodness of the model, the differences $\Delta R = \widehat{R_0} - R_0$ and $\Delta UVI = \widehat{UVI} - UVI$ were considered. Considering that the distribution of differences is slightly asymmetric and that in comparisons with measured data (next sections) some outliers may be present, the following metrics were computed:

- Mean of the differences: it is an estimate of the bias of the model;
- Root Mean Square (RMS) of the differences: is an estimate of the expected model error. RMS can provide overestimates of the error in the presence of outliers and is affected by the skewness of the differences.
- Median of the Absolute Differences;
- Scaled Median of the Absolute Deviation (MAD):

$$MAD(x) = 1.4826 \cdot Median(|x - Median(x)|)$$

were 1.4826 is the scaling factor. Like RMS, MAD is an estimator of the expected model error but it is less sensitive to outliers and to the skewness of the differences[55];

- 95th percentile of the absolute value of the differences (ABS95): is an estimate of the maximum differences by reducing the effect of outliers;
- Percentage of cases with absolute difference (AD) < 1 (P1, computed only for ΔUVI).

In this case the following values were obtained:

ΔR: Mean $< 10^{-6}$ m^2 W^{-1} , RMS = $2.9 \cdot 10^{-3}$ m^2 W^{-1} , MAD = $2.0 \cdot 10^{-3}$ m^2 W^{-1}, ABS95 = $5.0 \cdot 10^{-3}$ m^2 W^{-1}

ΔUVI: Mean < 0.1 , RMS = 0.3 , MAD = 0.2, ABS95 = 0.4, P1 = 100%

Both scatter plots and the statistics of the differences indicate that the goodness of the fit is high. In particular, ΔUVI is unbiased and ≤ 0.5 units in all cases.

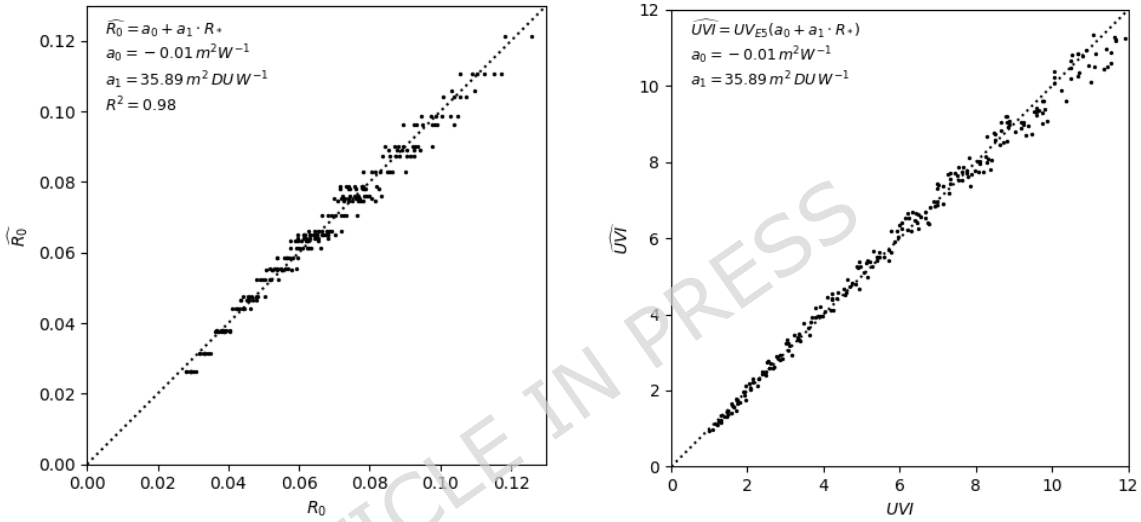


Figure 4. Scatter plots of R_0 (left) and UVI(right) vs the values $\widehat{R_0}$ and $\widehat{UVI}$ obtained from Eq.(4) and Eq.(5)

3 *UVI* measurements, ERA5 and CAMS data

The hourly UVI measurements (UVI_m) considered in this study for the application of the model were collected at the six ground stations described in Table 2. Data provider references are reported in Acknowledgments section. Seckmeyer et al. [56] suggest that the optimal uncertainty target for broadband instruments falls within the 10-15% range. However, since the specific measurement uncertainties for the UVI data from the individual stations are unavailable, we have adopted the approach of Pitkanen et al. [57], used for the validation of the CAMS Solar Radiation Products, to assign a standard uncertainty. Following this methodology, the uncertainty for our data is estimated to be slightly higher, at up to 20%. Further details on the measuring stations are given in the Discussion section.

ERA5 data are provided on a 0.25° × 0.25° grid cell. For stations located in mountainous areas, the measurement altitude may therefore differ significantly from the representative altitude of the corresponding grid cell. For the four mountain stations, the representative altitudes of the respective grid cells (ERA5 geopotential variable) were extracted and compared with the station

altitudes reported in Table 2. The absolute differences are DAVO 410 m, PALL 636 m, RENO 403 m, SEST 169 m. As a rough estimate, UVI can be assumed to vary by about 5% per 1000 m of altitude (TEMIS, https://www.temis.nl/uvradiation/product/height.html). Thus, even in the worst case, the altitude difference between ERA5 and ground-based measurements should result in errors below 4%.

Table 2. UVI measurements: ground stations

Acronym	Quantity	Location	Position (WGS84)	Altitude (m.a.s.l.)	Period (YYYY/MM/DD)	Time step
DAVO	SOLARLIGHT UV-Biometer	Davos (CH)	9.8444 E 46.8133 N	1590	2015/01/01 – 2016/31/12	Hourly
DIEK	SOLARLIGHT UV-Biometer	Diekirch (LU)	6.1700 E 49.8610 N	218	2015/02/07 – 2016/11/03	Central hours of the day
GOMO	Davis Vantage-Pro	Modena (IT)	10.9299 E 44.6481 N	64	2020/12/17 – 2022/10/07	Hourly
PALL	Davis Vantage-Pro	Pallanza (IT)	8.54820 E 45.9239 N	214	2019/01/01 – 2019/10/17	Hourly
RENO	SOLARLIGHT UV-Biometer	Renon (IT)	11.4331 E 46.5897 N	1770	2018/01/01 – 2019/12/31	Hourly
SEST	Davis Vantage-Pro	Sestriere (IT)	6.87470 E 44.9514 N	2020	2019/01/01 – 2021/12/31	Hourly

Data source/providers:
DAVO, DIEK: WOUDC (World Ozone and Ultraviolet Radiation Data Centre) (Canada)
GOMO: Geophysical Observatory of Modena (Italy)
SEST, PALL: ARPA Piemonte (Italy)
RENO: APPA Bolzano (Ialy)

In the periods listed in Table 2 the ERA5, gridded data (0.25° x 0.25°) of downward UV radiation at the surface converted to UV_{E5} [W/m^2], total cloud cover fraction (CC, in tenths), and total column ozone (O_3, in Dobson units, DU) on areas including the ground stations were downloaded from the Copernicus Climate Data Store service (https://cds.climate.copernicus.eu) and used for the comparisons. The ERA5 values at the ground station locations were estimated by spatial

bilinear interpolation. For each date and time of the measurements acquired at each ground station (longitude and latitude), the hourly value of μ was computed using the Python program *SunPos.py* [58]. Considering that low UV values are of minor interest (low impact on population health) and the difficulties in modelling cases with $\mu < 0.5$ (i.e. $\theta_s > 60°$) were excluded. In Table 3 some elements of the summary statistics of UVI, UV_{E5}, $O3$, μ, and CC of the dataset used for the comparison between the values calculated by the model and the ground measurements are reported.

Table 3. Summary statistics of the UVI_m, UV_{E5}, and $O3$ quantities for all the ground stations (joined).

	UVI_m [-]	UV_{E5} [W m^{-2}]	O3 [DU]	μ [-]	CC [-]
Count	17433	17433	17433	17433	17433
Mean	4.0	62	330	0.72	0.55
St. Dev.	2.0	22	33	0.12	0.35
Min.	1.0	2	242	0.50	0.00
25%	2.3	47	307	0.61	0.23
50%	3.6	62	326	0.73	0.57
75%	5.3	79	351	0.83	0.90
Max.	11.5	117	483	0.93	1.00

CAMS forecasts are provided on a 0.4° × 0.4° grid cell. For each CAMS variable, forecasts are available starting from 00:00 or 12:00 (t_0) for the following 120 hours (Δt). In this study, the values of "UV biologically effective dose" and "UV biologically effective dose, clear sky" (i.e. computed with a clear-sky algorithm), for $t_0 = 12{:}00$ and $\Delta t = 0$ (initialization data) were used, converted to UVI, and named $UVI_{C,ASC}$ and $UVI_{C,CSC}$ respectively (see documentation at https://ads.atmosphere.copernicus.eu/datasets/cams-global-atmospheric-composition-forecasts?tab=overview for details). To achieve greater similarity with ground-based measurements, both CAMS data and the corresponding ERA5 data were extracted at the same locations as the ground stations and for the period between 01/07/2016 (there is no data available before this data) and 31/12/2022, using the same method described above (μ computation and threshold, bilinear interpolation, etc.). Similarly to Table 3, Table 4 reports the summary statistics of this dataset.

Table 4. Summary statistics of the UVI_C, $UVI_{C,CSC}$, UV_{E5}, and $O3$ quantities of the dataset used for the comparison between the model and CAMS data.

	$UVI_{C,ASC}$ [-]	$UVI_{C,CSC}$ [-]	UV_{E5} [W m^{-2}]	O3 [DU]	μ [-]	CC [-]
Count	6537	6537	6537	6537	6537	6537
Mean	4.8	6.6	74	331	0.72	0.58
St. Dev.	2.0	1.8	21	35	0.10	0.34
Min.	1.0	1.9	5	243	0.50	0.35
25%	3.2	5.1	62	307	0.64	0.29
50%	4.7	6.7	77	326	0.75	0.62
75%	6.2	8.0	91	351	0.81	0.92
Max.	11.2	11.2	118	493	0.84	1.00

4 Application of the model

4.1 Comparison of the model with ground station measurements

Since the simulations were done in Clear Sky Conditions, the model defined by Eq.(5) was first applied to the experimental dataset imposing $CC \leq 0.1$. Figure 5 (upper plots), shows the scatter plot of the measured UVI *vs* modelled UVI ($\widehat{UVI}$) for all stations, and the box plots of $\Delta UVI = \widehat{UVI} - UVI$, for each individual station and for all stations. Table 5 reports the number of points and the metrics listed in section 2.4. Values reported in brackets refer to metrics calculated for $UVI \geq 3$, ("Moderate" to "Extreme" exposure category[16]). Subsequently, the model was applied to All Sky Conditions (Figure 5 , lower plots and Table 6).

Given that the effect of cloud cover is unlikely to cancel out perfectly in the R_0 ratio under cloudy skies, the mean and the RMS values were computed for each $CC_i \pm 0.05$ interval, where CC_i ranges from 0.05 to 0.95 with 0.1 steps, to analyze the impact of cloud fraction on the results. The results (all stations) are depicted in Figure 6.

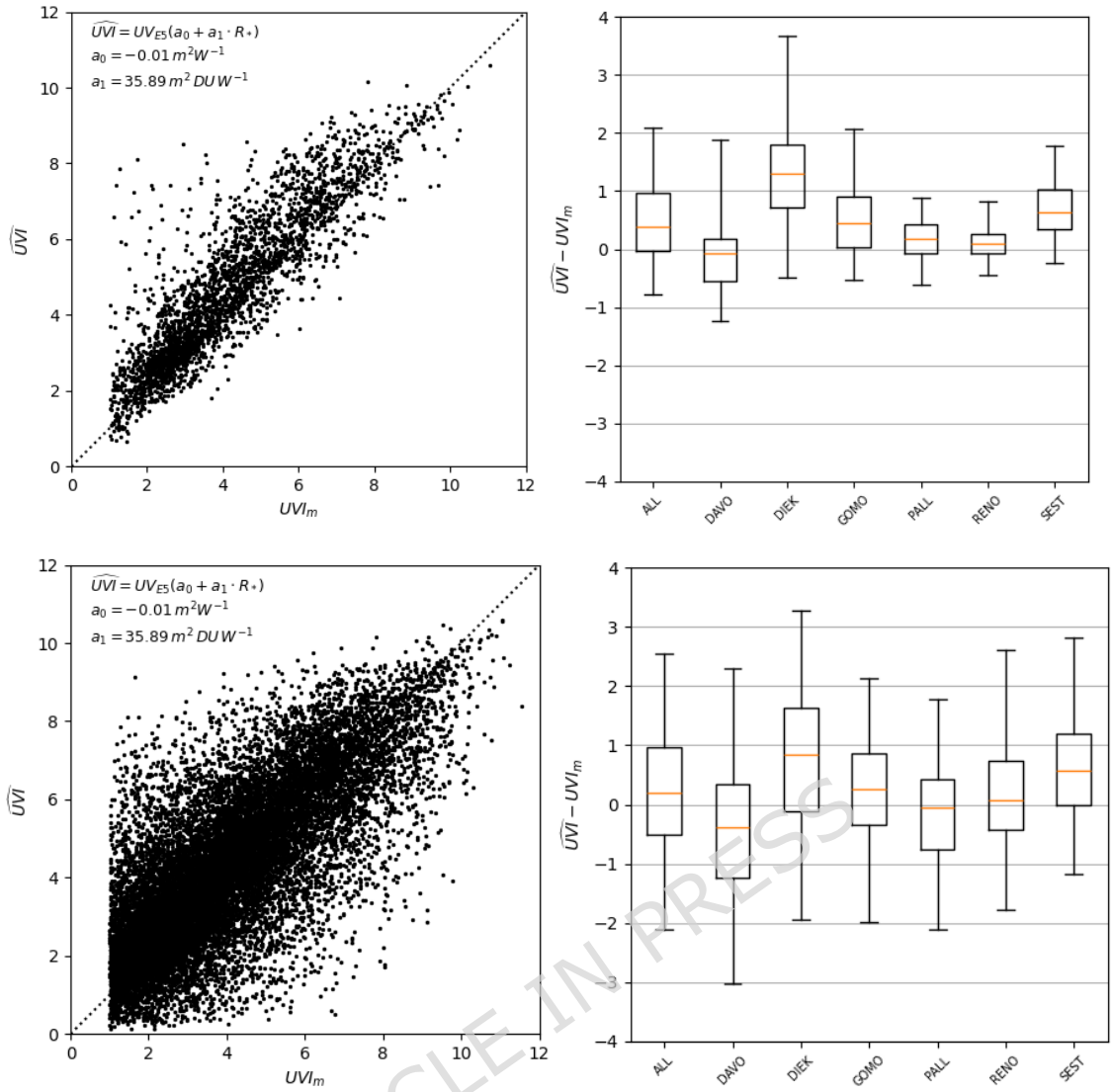


Figure 5. Measured UVI (UVI_m) and computed $\widehat{UVI}$ (Eq.(5)) comparison in CSC (upper plots) and in ASC (lower plots). Scatter plot (left, all stations) and box plots of the differences (right, all stations and single stations).

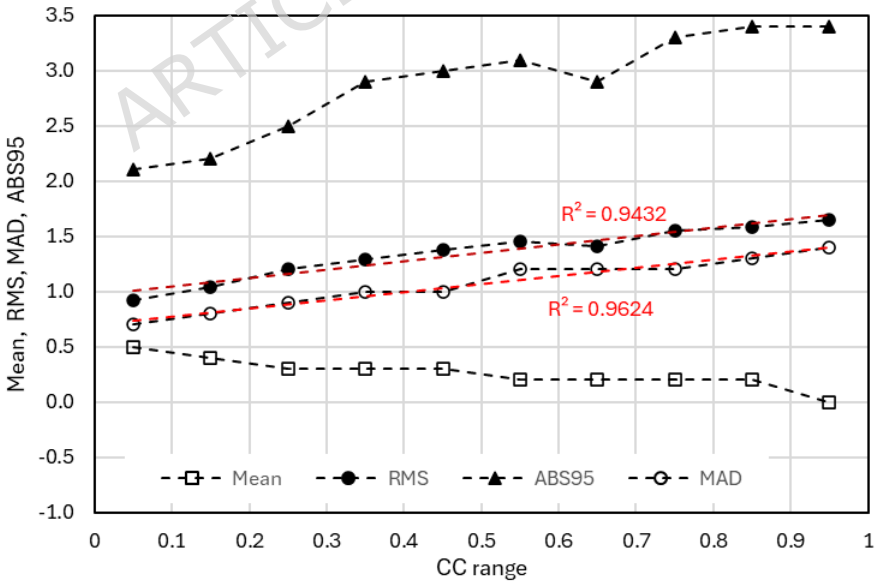
Table 5. Measured UVI and computed $\widehat{UVI}$ (Eq.(5)) in CSC: $\widehat{UVI}$ -UVI_m statistics for all stations and for single stations. Number of points (Size), Mean, Root Mean Squared value (RMS), scaled Median of Absolute Deviation (MAD), 95th percentile of the absolute values (ABS95), percentage of cases with absolute difference < 1 (P1). The metrics are calculated for UVI ≥ 1; values for UVI ≥ 3 are given in parentheses.

Station	Size	Mean	RMS	MAD	ABS95	P1
ALL	2863 (2259)	0.5 (0.5)	0.9 (0.9)	0.7 (0.7)	2.1 (2.9)	73% (72%)
DAVO	453 (403)	0.0 (-0.1)	0.9 (0.9)	0.5 (0.6)	1.9 (1.9)	78% (78%)
DIEK	420 (285)	1.4 (1.1)	1.3 (1.1)	0.8 (0.9)	3.7 (2.8)	38% (37%)
GOMO	645 (452)	0.5 (0.6)	0.8 (0.9)	0.7 (0.8)	2.1 (2.2)	76% (70%)
SEST	724 (562)	0.7 (0.7)	0.6 (0.6)	0.5 (0.6)	1.9 (1.8)	71% (66%)

RENO	354 (328)	0.1 (0.1)	0.5 (0.5)	0.2 (0.2)	0.8 (1.0)	94% (95%)
PALL	267 (229)	0.2 (0.2)	0.5 (0.5)	0.4 (0.3)	0.9 (1.0)	95% (94%)

Table 6. Measured UVI and computed $\widehat{UVI}$ (Eq.(5)) in ASC: $\widehat{UVI}$ -UVI_m statistics for all stations and for single stations. Number of points (Size), Mean, Root Mean Squared value (RMS), scaled Median of the Absolute Deviation (MAD), 95th percentile of the absolute values (ABS95), percentage of cases with absolute difference < 1 (P1). The metrics are calculated for UVI ≥ 1; values for UVI ≥ 3 are given in parentheses.

Station	Size	Mean	RMS	MAD	ABS95	P1
ALL	17443 (12414)	0.2 (0.0)	1.4 (1.4)	1.1 (1.2)	2.6 (3.0)	61% (59%)
DAVO	3318 (2388)	-0.4 (-0.7)	1.6 (1.5)	1.1 (1.2)	2.3 (3.5)	55% (51%)
DIEK	2036 (1178)	0.8 (0.4)	1.5 (1.5)	1.3 (1.5)	3.2 (2.9)	44% (43%)
GOMO	3214 (2227)	0.2 (0.1)	1.2 (1.4)	0.9 (1.1)	2.1 (2.8)	66% (61%)
SEST	4063 (3014)	0.6 (0.5)	1.2 (1.2)	0.9 (1.1)	2.8 (2.8)	61% (61%)
RENO	3264 (2327)	0.2 (0.0)	1.3 (1.2)	0.8 (0.8)	2.6 (2.7)	67% (68%)
PALL	1548 (1240)	-0.1 (-0.3)	1.2 (1.1)	0.6 (0.9)	1.8 (2.5)	70% (69%)



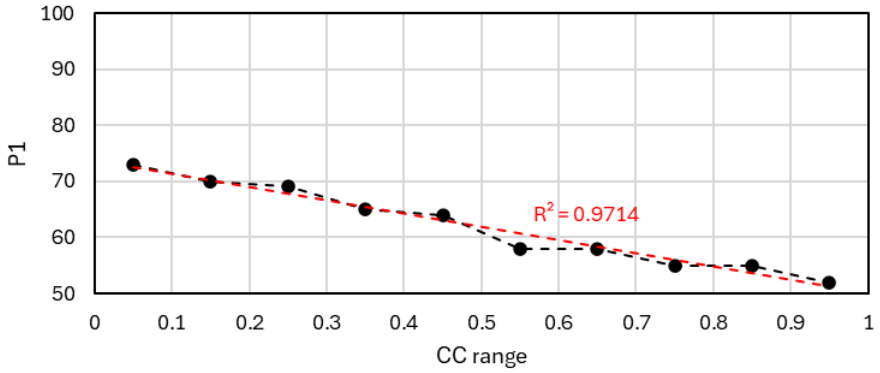
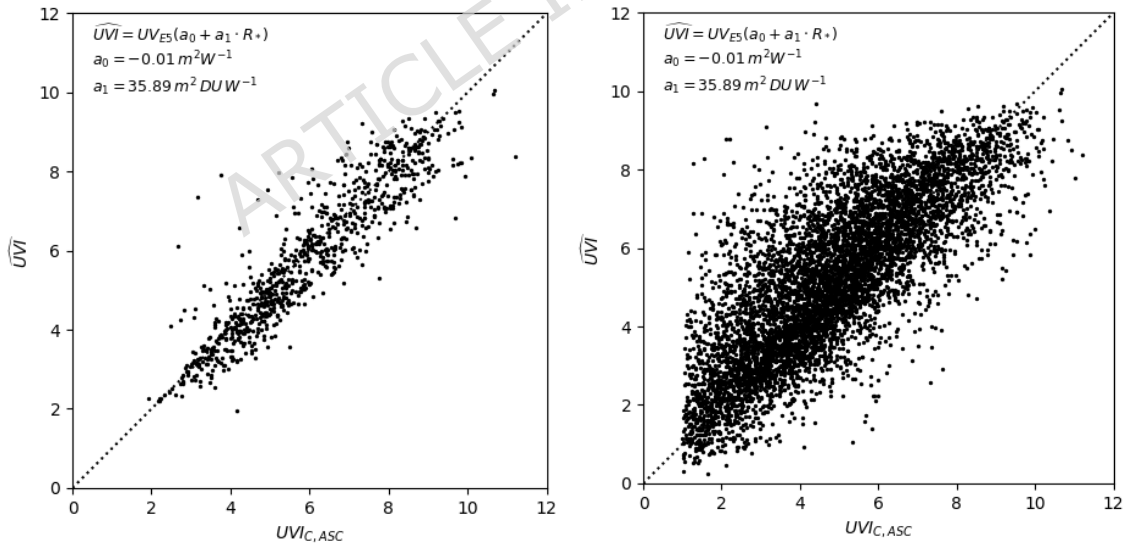


Figure 6. Mean (squares), RMS (filled circles), MAD (open circles), ABS95 (triangles), and P1 of the differences $\widehat{UVI}$ -UVI (all stations) for each $CC_i \pm 0.05$ interval, where CC_i ranges from 0.05 to 0.95 with 0.1. Red lines represent linear regressions.

4.2 Comparison of the model with CAMS

Using the same methodology described for the comparison between modelled UVI values ($\widehat{UVI}$) and those measured by ground stations, this section compares the modelled values with the UVI derived from CAMS dataset, specifically $UVI_{C,ASC}$ and $UVI_{C,CSC}$ datasets, and taking into account than in CSC $UVI_{C,ASC} = UVI_{C,CSC}$. The comparison between $UVI_{C,ASC}$ and $UVI_{C,CSC}$ in ASC has also been added. All comparisons were performed on the complete datasets, excluding any location-specific analysis. The results are reported in Figure 7 and in Table 7.



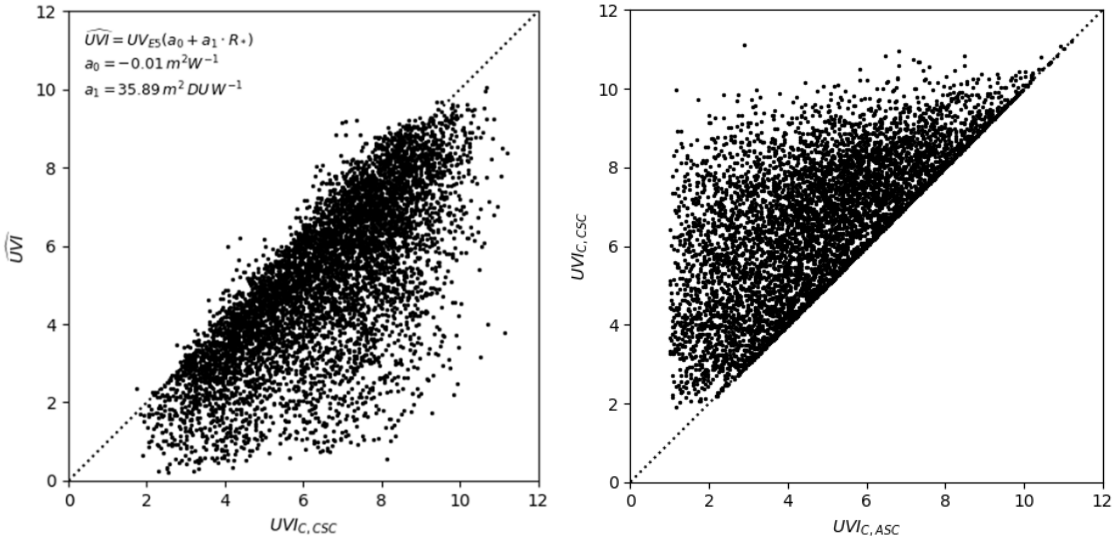


Figure 7. From top left, clockwise: $UVI_{C,ASC}$ *vs* $\widehat{UVI}$ in CSC, $UVI_{C,ASC}$ *vs* $\widehat{UVI}$ in ASC, $UVI_{C,CSC}$ *vs* $\widehat{UVI}$ in ASC, $UVI_{C,ASC}$ *vs* $UVI_{C,CSC}$ in ASC. $\widehat{UVI}$ is computed using Eq.(5).

Table 7. Comparisons between computed UVI ($\widehat{UVI}$, Eq.(5)), $UVI_{C,ASC}$ and $UVI_{C,CSC}$, in CSC and ASC: Number of points (Size), Mean, Root Mean Squared value (RMS), scaled Median of the Absolute Deviation (MAD), 95th percentile of the absolute values (ABS95), percentage of cases with absolute difference < 1 (P1). The metrics are calculated for $UVI_{C,ASC\ or\ CSC} \geq 1$; values for $UVI_{C,\ ASC\ or\ CSC} \geq 3$ are given in parentheses.

	CSC					
	Size	Mean	RMS	MAD	ABS95	P1
$\widehat{UVI}$ - $UVI_{C,ASC}$	835 (808)	0.1 (0.1)	0.7 (0.7)	0.6 (0.6)	1.5 (1.5)	86% (86%)
	ASC					
	Size	Mean	RMS	MAD	ABS95	P1
$\widehat{UVI}$ - $UVI_{C,ASC}$	6537 (5173)	-0.4 (-0.2)	1.3 (1.3)	1.1 (1.1)	2.8 (2.6)	61% (62%)
$\widehat{UVI}$ - $UVI_{C,CSC}$	6678 (6504)	1.4 (1.4)	1.4 (1.4)	1.2 (1.2)	4.3 (4.3)	48% (48%)
$UVI_{C,CSC}$ · UVI_{C}	6537 (5173)	1.7 (1.4)	1.6 (1.3)	1.7 (1.5)	4.8 (4.0)	40% (47%)

Similarly to what was done in the previous comparison, the RMS of the $\widehat{UVI}$ vs $UVI_{C,ASC}$ and of the $\widehat{UVI}$ vs $UVI_{C,CSC}$ comparisons were computed for each $CC_i \pm 0.05$ interval, where CC_i ranges from 0.05 to 0.95 with 0.1 steps, to analyze the impact of cloud fraction on the results. The results are shown in Figure 8 and in Figure 9.

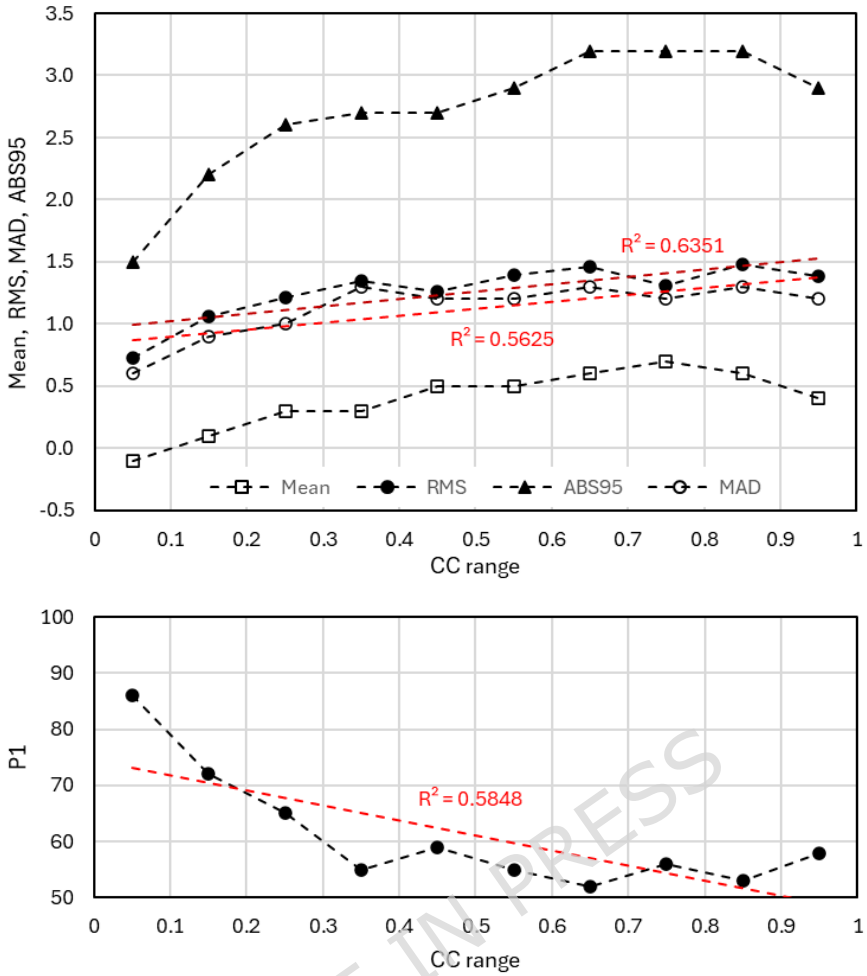
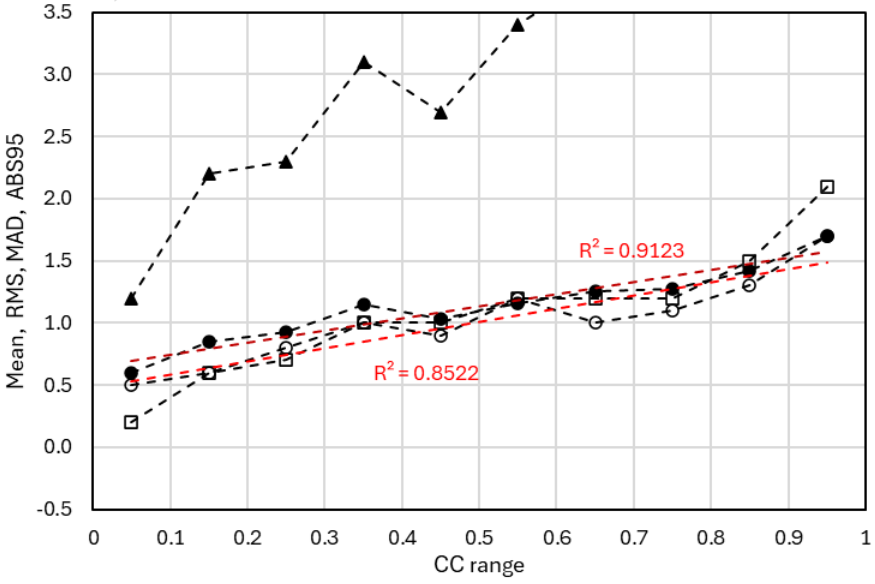


Figure 8. Mean (squares), RMS (filled circles), MAD (open circles), ABS95 (triangles), and P1 of the differences $\widehat{UVI}$ -$UVI_{C,ASC}$ (all stations) for each $CC_i \pm 0.05$ interval, where CC_i ranges from 0.05 to 0.95 with 0.1. Red lines represent linear regressions.



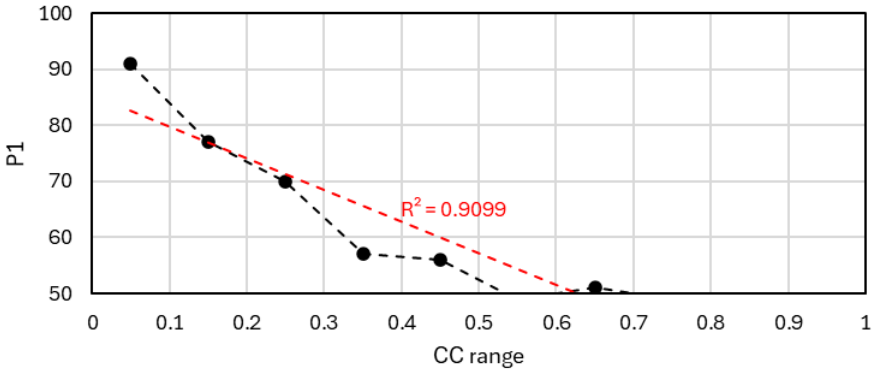
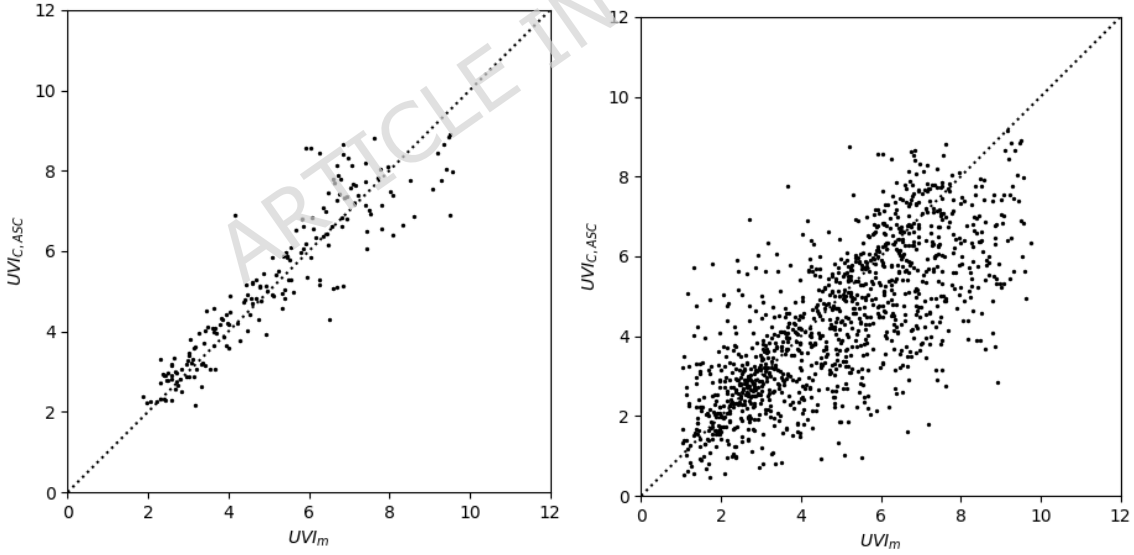


Figure 9. Mean (squares), RMS (filled circles), MAD (open circles), ABS95 (triangles), and P1 of the differences $\widehat{UVI}$ -$UVI_{C,CSC}$ (all stations) for each $CC_i \pm 0.05$ interval, where CC_i ranges from 0.05 to 0.95 with 0.1. Red lines represent linear regressions.

4.3 Comparison of CAMS with ground stations measurements

Although the analysis of CAMS products is not within the scope of this study, to complete the comparisons among the considered datasets, the UVI values measured by all ground stations were compared with the values extracted from the CAMS dataset, following the same methodology as in the previous section. The results are reported in Figure 10 and in Table 8. The analysis of the effect of cloud cover in these comparisons is shown in Figure 11 for the $UVI_{C,ASC}$ dataset and in Figure 12Figure 10 for the $UVI_{C,CSC}$ dataset.



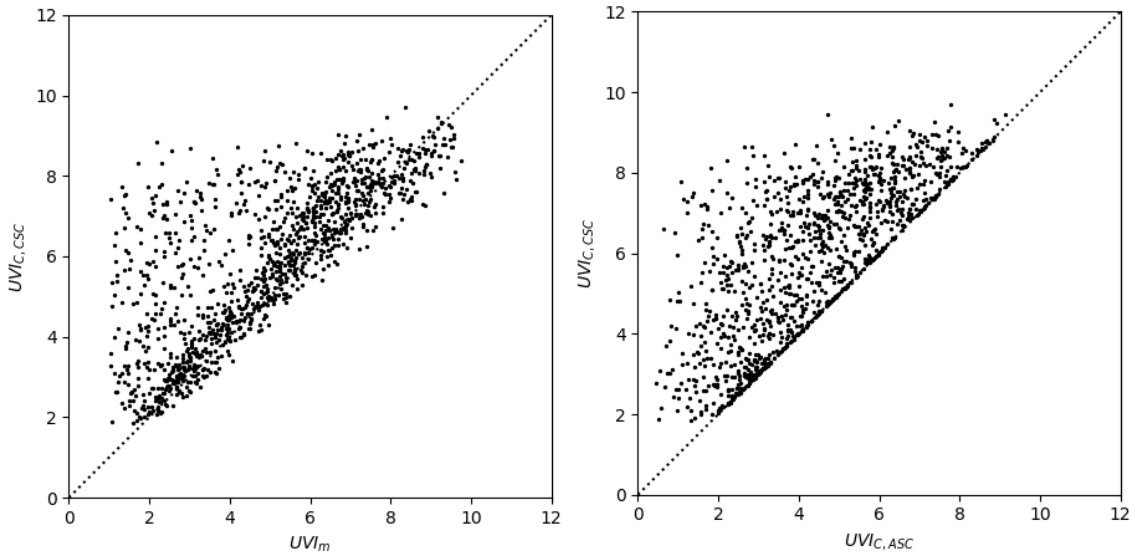


Figure 10. From top left, clockwise: UVI vs $UVI_{C,ASC}$ in CSC, UVI vs $UVI_{C,ASC}$ in ASC, UVI vs $UVI_{C,CSC}$ in ASC, $UVI_{C,ASC}$ *vs* $UVI_{C,CSC}$ in ASC. UVI indicates ground station measurements.

Table 8. Comparisons between measured UVI, $UVI_{C,ASC}$ and $UVI_{C,CSC}$, in CSC and ASC: Number of points (Size), Mean, Root Mean Squared value (RMS), scaled Median of the Absolute Deviation (MAD), 95th percentile of the absolute values (ABS95), percentage of cases with absolute difference < 1 (P1). The metrics are calculated for $UVI_{C,ASC\ or\ CSC} \geq 1$; values for $UVI_{C,\ ASC\ or\ CSC} \geq 3$ are given in parentheses.

	CSC					
	Size	Mean	RMS	MAD	ABS95	P1
$UVI_{C,ASC}$ - UVI_m	188 (156)	0.1 (0.1)	0.8 (0.9)	0.5 (0.6)	1.7 (1.7)	84% (81%)
	ASC					
	Size	Mean	RMS	MAD	ABS95	P1
$UVI_{C,ASC}$ - UVI_m	1184 (892)	-0.5 (-0.8)	1.5 (1.5)	1.3 (1.5)	3.2 (3.5)	56% (51%)
$UVI_{C,CSC}$ - UVI_m	1184 (892)	1.0 (0.7)	1.4 (1.1)	0.9 (0.9)	4.2 (3.1)	62% (68%)
$UVI_{C,CSC}$ · UVI_C	1184 (892)	1.6 (1.4)	1.4 (1.2)	1.5 (1.4)	4.1 (3.6)	43% (46%)

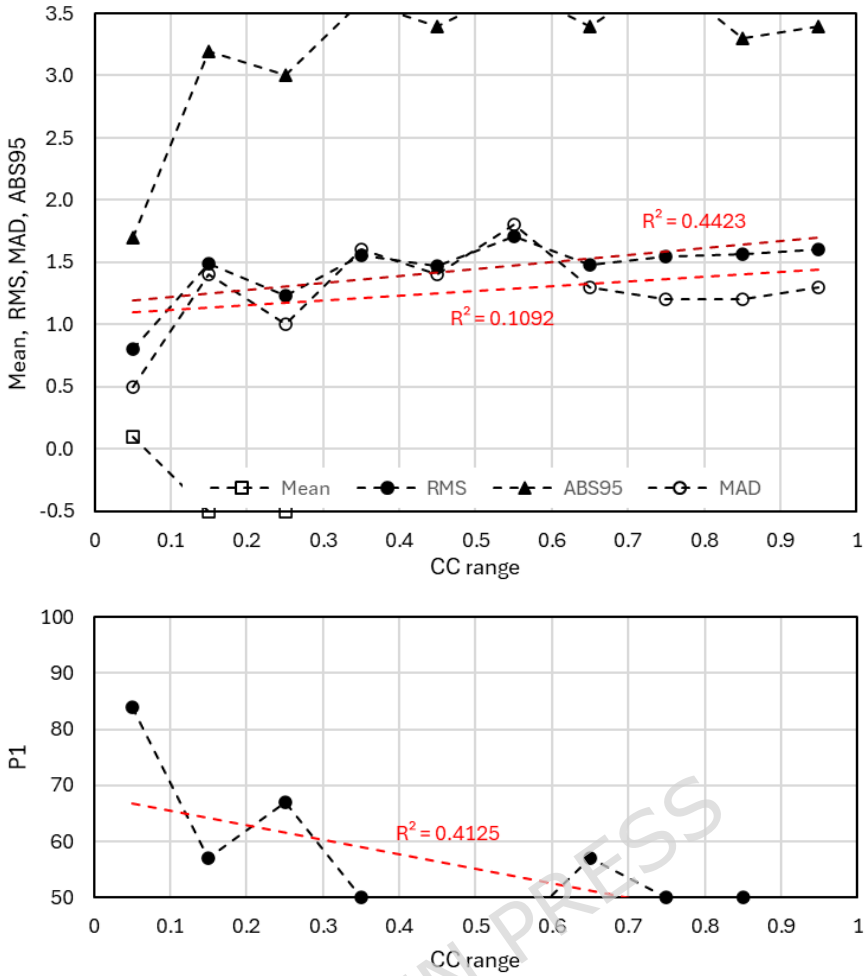
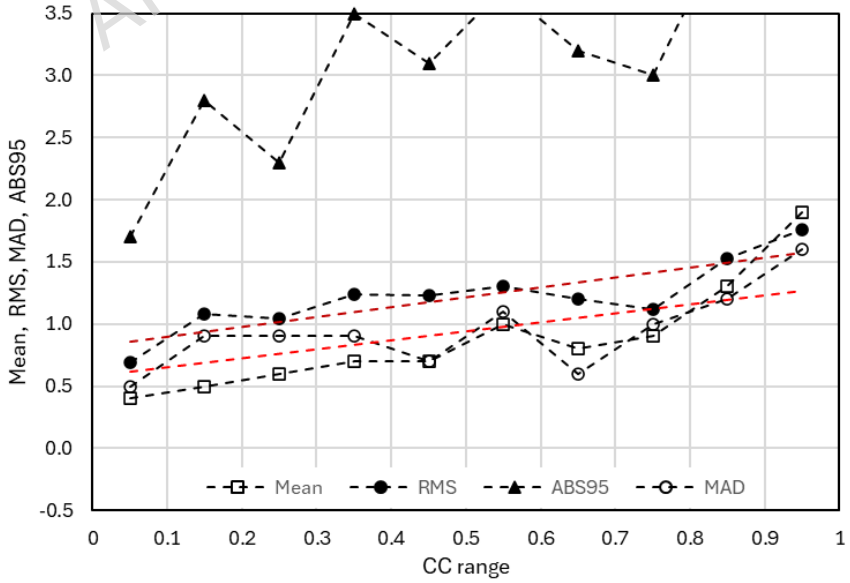


Figure 11. Mean (squares), RMS (filled circles), MAD (open circles), ABS95 (triangles), and P1 of the differences $UVI_{C,ASC}$ -UVI_m (all stations) for each CC_i ± 0.05 interval, where CC_i ranges from 0.05 to 0.95 with 0.1. Red lines represent linear regressions.



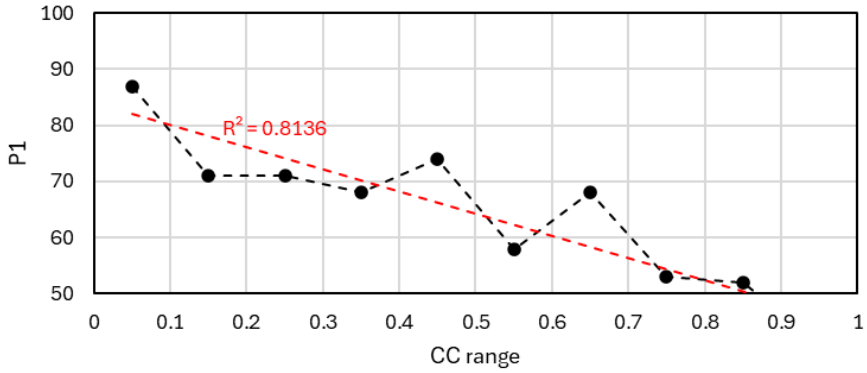


Figure 12. Mean (squares), RMS (filled circles), MAD (open circles), ABS95 (triangles), and P1 of the differences $UVI_{C,CSC}$ -UVI_m (all stations) for each $CC_i \pm 0.05$ interval, where CC_i ranges from 0.05 to 0.95 with 0.1. Red lines represent linear regressions.

5 Discussion

5.1 Comparison of the model with ground station measurements

The application of the model under CSC conditions provided good results (Table 6, $\Delta UVI = \widehat{UVI}$ -UVI_m statistics). Except for DIEK data, results showed mean difference ≤ 0.5, RMS < 1, MAD ≤ 0.9, and ABS95 ≤ 2.1 (i.e. $|\Delta UVI|$ is > 2.1 only in the 5% of the cases). The P1 percentage was between 70% and 95%. No substantial differences were observed in the statistics between $UVI \geq 1$ and $UVI \geq 3$ cases.

The application of the model under All Sky Conditions (ASC), i.e. with no limits to *CC*, resulted in greater uncertainty in UVI prediction compared to the CSC case. This was to be expected, given that the presence of clouds generates a large variability, very difficult to model, of UV radiation at the ground[44]. While the mean of ΔUVI is still low, other metrics worsened: except for DIEK data, results showed RMS ≤ 1.6, MAD ≤ 1.2, ABS95 ≤ 3.5, and P1 was between 51% and 70%. Also in this case, no substantial differences were observed in the statistics between $UVI \geq 1$ and $UVI \geq 3$ cases.

The DIEK data were thoroughly analysed to investigate the cause of the worsening of the performance of the model, but no plausible explanations were found. Therefore, for statistical integrity, these data were not excluded from the metric calculations.

The discrepancies in the statistics obtained for the different stations can be attributed to several factors affecting the representativeness of the ground-based measurement for the corresponding ERA5 grid cell. These factors include: the heterogeneity and type of land use/cover within the ERA5 grid cell (e.g., mountains, lakes, cities, rural areas); uncertainty of the sensors; the composition of the lower atmosphere (e.g., the specific conditions of the Po Valley versus the Alps); meteorological patterns (e.g., cloud formation and the persistence and frequency of snow cover in areas such as the Po Valley, the southern side of the Alps, and the northern side of the Alps).

The analysis of cloud cover (CC) influence on model performances (Figure 6) revealed that the mean ΔUVI was practically constant across all CC_i intervals, ranging between 0 and 0.5. This

indicates that CC does not introduce a bias effect in the model. All other metrics worsened as CC increased. RMS and MAD showed a nearly linear positive trend, while P1 showed a nearly linear negative trend, described by the regressions shown in Figure 6 (red lines).

A possible explanation of the worsening of the metrics is that the model was developed using only simulations computed in clear-sky conditions and that, as mentioned above, the effect of cloud cover does not perfectly cancel out in the ratio R_0. However, under this assumption, an effect should also be observed on the mean ΔUVI. To further investigate this aspect, a correlation analysis was performed using the Pearson (r), Spearman (ρ), and Kendall (τ) correlation coefficients for selected variables; the results are reported in Table 9. Under CSC conditions, the correlation between the measured UVI values and the UV_{E5} values extracted from ERA5 is relatively high and very similar (slightly lower) to that obtained for the corresponding simulated variables. Under ASC conditions, as expected, this correlation decreases substantially. At the same time, the correlation between cloud cover and the differences between measured and modelled UVI is also found to be low. By comparing the correlations between UVI_m and CC and between UV_{E5} and CC, it can be observed that the former is much weaker than the latter, suggesting that this difference may be responsible for the loss of correlation between the discrepancies of modelled and observed values and cloud cover. This can be explained considering that ERA5 *CC* is representative of 0.25° x 0.25° cells, while the *UVI* measurements are local (note that *local* does not mean *point-like* or *directional*, since the instrument observes a large portion of the sky within its field of view). This hypothesis is further supported by the observation that in 25% of the cases ($N = 1135$) where $CC > 0.9$, UVI measurements exceed 4, which is higher than expected in such conditions. It can therefore be concluded that the different spatial representativeness of UVI_m and CC introduces a quasi-random effect on UVI estimates, resulting in unbiased uncertainties that increase as CC increases.

Table 9 Correlation between several dataset used. Coefficients: Pearson (r), Spearman (ρ), and Kendall (τ).

Variables	Sky conditions	Correlation
(UVI_{uvspec}, $UV_{E5,uvspec}$)	CSC	$r = 0.89$, $\rho = 0.90$, $\tau = 0.72$
(UVI_m, UV_{E5})	CSC	$r = 0.80$, $\rho = 0.82$, $\tau = 0.63$
(UVI_m, UV_{E5})	ASC	$r = 0.62$, $\rho = 0.65$, $\tau = 0.47$
($\widehat{UVI}$ - UVI_m, CC)	ASC	$r = 0.12$, $\rho = 0.13$, $\tau = 0.09$
(UVI_m, CC)	ASC	$r = -0.21$, $\rho = -0.23$, $\tau = -0.15$
(UV_{E5}, CC)	ASC	$r = -0.40$, $\rho = -0.41$, $\tau = -0.28$
($UVI_{C,ASC}$, CC)	ASC	$r = -0.48$, $\rho = -0.41$, $\tau = 0. - 0.28$
($UVI_{C,ASC}$ - UVI_m, CC)	ASC	$r = -0.11$, $\rho = -0.13$, $\tau = 0.09$
($UVI_{C,CSC}$, CC)	ASC	$r = -0.47$, $\rho = -049$, $\tau = -0.34$
($UVI_{C,CSC}$ - UVI_m, CC)	ASC	$r = -0.50$, $\rho = -0.52$, $\tau = -0.36$

5.2 Comparison of the model with CAMS

The comparisons of the model-predicted UVI with the UVI obtained from the CAMS dataset is shown in Table 7. The statistics obtained under CSC, noting that in CSC $UVI_{C,ASC} = UVI_{C,CSC}$, are slightly better than those derived from the comparison with ground-based stations, particularly for ABS95 (1.5) and P1 (86%). Under ASC, the comparison between $\widehat{UVI}$ and $UVI_{C,ASC}$, is similar to that obtained with ground-based measurements, but it is noticeably better than the comparison between $UVI_{C,ASC}$ and $UVI_{C,CSC}$, highlighting the difference between the two dataset.
The analysis of the influence of cloud cover (Figure 8) revealed that the mean ΔUVI is close to zero in near clear-sky conditions, and increases to values between 0.5 and 0.7 per $CC > 0.4$. The RMS and MAD exhibit an almost linear positive trend, while P1 shows an almost linear negative trend for $CC < 0.4$; for $CC > 0.4$, these metrics remain nearly constant.

5.3 Comparison of CAMS with ground station measurements

The comparisons of the UVI extracted from CAMS dataset and those measured by the ground stations are shown in **Figure** 10 and in Table 8.
Compared to the model–ground measurement comparison, very similar RMS and MAD values are observed, along with a slight increase (in absolute value) in MAE and a deterioration of the extreme-value indicators ABS95 and P1, for both the $UVI_{C,ASC}$ and $UVI_{C,CSC}$ datasets. The variation of RMS and MAD parameters with increasing cloud cover, shown in Figure 11 and in Figure 12, is similar to that observed in the comparisons between the model and ground-based stations, with a slight improvement in the $UVI_{C,ASC}$ case (except for nearly complete cloud cover). For high CC values, a significant deterioration of the ABS95 and P1 parameters is observed in both cases. In the $UVI_{C,CSC}$ case, a general worsening of MAE is also noted. For $CC < 0.4$, both cases show behavior comparable to that of the model results.

5.4 Influence of cloud cover on the model

The comparisons of the model predictions with ground-based UVI measurements and with CAMS data show that Cloud Cover is not a suitable additional explanatory variable for UVI estimation under cloudy conditions.
ERA5 provides additional cloud cover types (high, medium, and low), which we tested as replacements for the total cloud cover (CC) used in the analyses above. However, these yielded worse results. Beyond cloud cover, we also evaluated other potential proxy variables for cloud effects (humidity, aerosol optical thickness, ...), though without success. The search for alternative proxies, such as the Cleaner Index used in Czerwińska and Krzyścin[55], will be part of future model developments.
With reference to the comparisons with ground stations, model uncertainty estimates can be derived from the linear regression equations depicted in red in Figure 6:

$$RMS = 0.97 + 0.76 \cdot CC\,, \quad MAD = 0.70 + 0.73 \cdot CC\,, \quad P1 = 73.7 - 23.7 \cdot CC \qquad (6)$$

For $CC \leq 0.4$, the metrics yield the following results: RMS < 1.2, MAD < 1.0, and P1 > 64%. RMS remains below 1 in more than 64% of the cases. Both RMS and MAD values stay below 1.5 for CC < 0.7. These estimates can also be used as numerical or categorical quality index (QI) for UVI computed with Eq. (5), an example is reported in Appendix A. When referring to the comparison between the model and CAMS, the uncertainties improve significantly for lightly cloudy skies (these estimates are made qualitatively considering the poor linearity of the metrics for low CC). Therefore, the uncertainties estimated using ground stations represent the worst-case. Moreover, although the RMS slightly exceeds 1, it should be noted that this metric tends to overestimate ΔUVI differences.

Table **10** summarizes the computed statistics for $CC \leq 0.4$.

The $CC < 0.4$ condition might appear very restrictive. However, from a human health perspective, in most cases (75th percentile, Q3) with a cloud cover > 0.4 and high solar elevation (> 45°), UVI is less than 6, precluding high exposure situations. This aspect is further highlighted in Figure 13, which shows the correlation between cloud cover and the model uncertainty (expressed as RMS percentage). The uncertainty values are derived from comparisons with both ground station measurements and CAMS. The uncertainty (RMS%) is calculated for reference UVI values of 5, 7, 9, and 11 (moderate, high, very high, and extreme exposure), and compared with the reference values (10% and 20%, black lines) indicated in the validation of the CAMS Solar Radiation Products report [57]. For example, for the moderate case the uncertainty is $RMS\% = 100 * RMS/UVI_{ref}$, where RMS is the mean RMS for UVI in the range [3.5,4.5] and $UVI_{ref} = 4$. These plots demonstrate that at high exposure levels, the model's relative uncertainty is comparable to the uncertainty of ground-based measurements. It also points to the need for more extensive validation using instruments with low, well-characterized uncertainty.

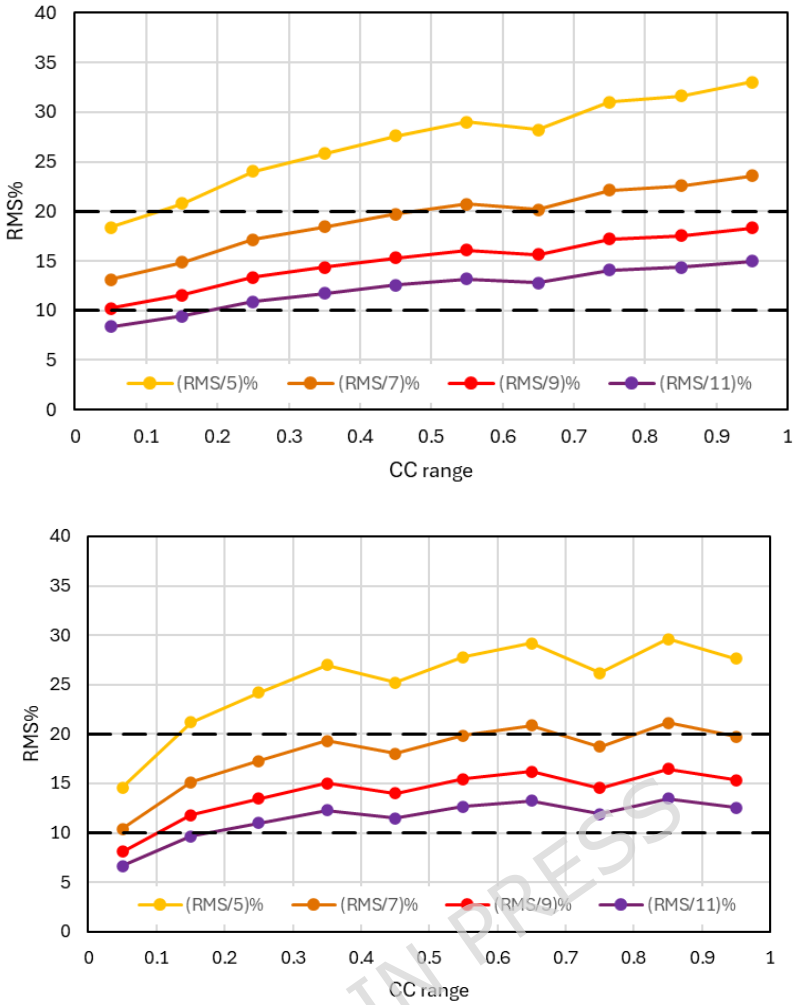


Figure 13. Ratio RMS% = 100 * RMS/UVI_{ref} , for reference UVI_{ref} values of 5, 7, 9, and 11 (moderate, high, very high, and extreme exposure) for each CC_i ± 0.05 interval, where CC_i ranges from 0.05 to 0.95 with 0.1. Top: RMS derived from ground stations measurements *vs* modelled UVI comparison. Bottom: RMS derived from CAMS UVI *vs* modelled UVI comparison. Black dashed lines indicates the uncertainty of ground measurements as reported in the validation of the CAMS Solar Radiation Products report [57].

Table 10. Statistics of the ΔUVI differences obtained for the various datasets for CC ≤ 0.4. Number of points (Size), Mean, Root Mean Squared value (RMS), scaled Median of the Absolute Deviation (MAD), 95th percentile of the absolute values (ABS95), percentage of cases with absolute difference < 1 (P1). The metrics are calculated for UVI ≥ 1; values for UVI ≥ 3 are given in parentheses.

ΔUVI	Size	Mean	RMS	MAD	ABS95	P1

$\widehat{UVI}$ - UVI	6592 (5281)	0.4 (0.3)	1.1 (1.0)	0.8 (0.8)	2.4 (2.4)	70% (70%)
$\widehat{UVI}$ -	2171 (2077)	0.1 (0.1)	1.1 (1.0)	0.9 (0.9)	2.3 (2.2)	72% (72%)
$\widehat{UVI}$ -	2171 (2077)	0.6 (0.6)	0.9 (0.9)	0.7 (0.7)	2.4 (2.4)	74% (73%)
$UVI_{C,ASC}$ · UVI	447 (377)	-0.3 (-0.4)	1.3 (1.3)	0.9 (1.1)	3.0 (3.0)	69% (65%)
$UVI_{C,CSC}$ · UVI	447 (377)	0.5 (0.5)	1.0 (0.9)	0.7 (0.8)	2.3 (2.3)	77% (77%)

6 Summary and conclusions

In this work, a model for the calculation of the UVI based exclusively on ERA5 data (Copernicus Climate Data Store) has been studied.

The model has been defined in clear-sky conditions starting from UVI and ERA5 hourly downward UV radiation (UV_{E5}) in the spectral range 280 nm – 440 nm simulated by the *uvspec* radiative transfer model in different atmospheric conditions, solar illumination and surface characteristics. The analysis of the simulations showed that the main parameters that influence the UVI/UV_{E5} are the total columnar content of ozone ($O3$) and the cosine of the solar zenith angle (μ). These quantities are included in the ERA5 dataset (the solar zenith angle is obtained from the geographical coordinates and time). The relation between UVI/UV_{E5} and the ratio $\mu/O3$ was defined using the simulation dataset and linear regression ($R^2 = 0.98$). The model uncertainties on the UVI estimates (ΔUVI, all stations, **Table 5**) are satisfactory: $Mean < 0.1$, $RMS = 0.3$, MAD = 0.2, $ABS95 = 0.4$, and P1 = 100%.

The model results were separately compared with two datasets: the UVI measured by 6 ground stations (UVI_m) and CAMS data (clear sky conditions model and all sky condition model, $UVI_{C,CSC}$ and $UVI_{C,ASC}$ respectively). In the first phase, the comparison was performed under clear-sky conditions by filtering the data using the hourly cloud cover (CC) data also included in the ERA5 dataset. These results can be considered satisfactory, especially for the comparison with CAMS data.

The comparison was then extended to all sky-conditions and, as expected, this resulted in a degradation of ΔUVI.

The comparisons with ground measurements (Table 6) yielded Mean = 0.2, RMS = 1.4, MAD = 1.1, ABS95 = 2.6 and P1 = 61%. For the comparison with $UVI_{C,ASC}$ values, virtually identical results were obtained (**Table 7**). This last comparison showed that these statistics are still significantly better than those obtained for the clear-sky CAMS model.

The analysis of the correlations between ERA5 CC and the independent variables, and between ERA5 CC and ΔUVI showed that the CC cannot be used as an additional independent variable to

improve the model in cloudy conditions. In particular, for ground stations measurements, this is partially due to the poor spatial representativeness of the ERA5 CC (0.25° x 0.25° grid) for local measurements. Other ERA5 cloud cover types (high, medium, and low), were tested as replacements for the total cloud cover (CC) but they yielded even less satisfactory results.

The statistical analysis of model uncertainties was repeated in three ways: using only cases with $UVI \geq 3$ (which are of greater interest for human health); for individual stations (only for the comparison of UVI_m *vs* $\widehat{UVI}$ comparison); and for different ranges of cloud cover. The statistics for cases with $UVI \geq 3$ do not show substantial differences compared to the complete dataset ($UVI \geq 1$).

Concerning ground stations measurements, the station-specific statistics yielded for the comparison with the model, are generally similar, except for the DIEK station, which shows a worsening of the metrics, an effect for which no explanation has been found so far.

The comparisons with ground measurements and CAMS data showed that, in general, model performance worsens significantly for $CC > 0.4$. However, considering that UVI is primarily used as a public health warning indicator, and thus the main interest lies in high UVI values, this limitation should not preclude the use of the model even under moderately cloudy conditions. The analysis showed that for CC < 0.7, the RMS% was below 15% for Very High-to-Extreme exposure categories and below than 20% for Moderate-to-Extreme categories.

To complete the analysis of the dataset comparisons, UVI values derived from the CAMS dataset were compared with ground-based measurements. The results obtained for the RMS and MAD metrics are generally consistent with those from the previous comparisons, particularly for cloud cover values below 0.4. However, in addition to a change in the mean bias, a significant deterioration of the ABS95 and P1 indicators, representative of extreme deviations, is observed, which may be attributed to the larger spatial extent of CAMS grid cells compared to those of ERA5. This latter aspect can be interpreted as further supporting evidence of the limitations related to the spatial representativeness of ground-based measurements when compared with modeled values, as discussed in the previous sections.

As practical application examples, the model could be used to generate maps of extreme UVI values over specific periods (month, season, year, etc.) using solely ERA5 data, and could be inserted into an automatic procedure that interfaces directly with the ERA5 dataset for the calculation of UVI maps, and consequently of UVB Dose and of UVB Erythemal radiation, with a spatial resolution of 0.25° x 0.25°. As an example, the procedure for calculating the daily maximum UVI and the 95th percentile of the daily maximum UVI, over a specific geographic area (Fig. S1 and Fig. S2) and for three cities (Paris, Milan, and Madrid, Fig. S3) is presented in Appendix A. Finally, the method may also serve as a reference for the calculation of other quantities related to UV solar radiation used in Health Sciences, such as the UVB for DNA damage and pre-vitamin D synthesis.

7 Data and scripts availability

The datasets generated and analyzed during this study, including the Python scripts used to generate the example presented in Appendix A, are available on Zenoodo, and are temporarily restricted during the peer-review process. Access can be granted upon request to the corresponding author (Sergio.teggi@unimore.it). The data will be made publicly available upon acceptance of the manuscript.

The original datasets of UVI measurements can be requested at the Institutions listed in the Acknowledgments Section.

The original ERA5 datasets are available at Copernicus Climate Data Store, “ERA5 hourly data on single levels from 1940 to present” (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels).

The original CAMS datasets are available at Copernicus Atmosphere Data Store, “CAMS global atmospheric composition forecasts” (https://ads.atmosphere.copernicus.eu/datasets/cams-global-atmospheric-composition-forecasts).

8 Acknowledgements

The UVI measurements at the Sestriere (SEST) and Pallanza (PALL) stations were kindly provided by the Dipartimento di Rischi Fisici e Tecnologici (Department of Physical and Technological Risks), ARPA Piemonte (Via Jervis, 30 – 10015 Ivrea (TO), Italy, dip.rischi.fisici.tecnologici@arpa.piemonte.it, https://www.arpa.piemonte.it/scheda-informativa/dipartimenti-tematici).

UVI data at the stations Davos (DAVO) and Diekirch (DIEK) was downloaded from the World Ozone and Ultraviolet Radiation Data Centre web portal (https://woudc.org/data/dataset_info.php?id=uv_index_hourly).

The UVI data of the Renon (RENO) station were kindly provided by the Labor für Luftanalysen und Strahlenschutz (Laboratory for Air Analysis and Radiation Protection) of the Autonome Provinz Bozen – Südtirol (Amba-Alagi-Straße 5, 39100 Bozen, Italy, labluftstrahlen@provinz.bz.it, https://umwelt.provinz.bz.it/de/home).

The UVI dataset of Modena was kindly provided by the Osservatorio Geofisico di Modena (Geophysical Observatory of Modena), Dipartimento di Ingegneria Enzo Ferrari, Università di Modena e Reggio Emilia (Via P.Vivarelli 10, 41125, Modena, Italy, ossgeo@unimore.it, https://www.ossgeo.unimore.it).

9 Author contributions

Conceptualization and Methodology S.T, T.F., and S.C.; Data curation, Formal analysis, F.D. and M.R All authors contributed to the writing of the article.

10 Competing interest

The authors declare no competing interests.

References

1. Blumthaler, M. UV monitoring for public health. *Int. J. Environ. Res. Public Health* **15**, (2018).
2. Godar, D. E. UV doses worldwide. *Photochem. Photobiol.* **81**, 736 – 749 (2005).
3. Turner, J. *et al.* A review on the ability of smartphones to detect ultraviolet (UV) radiation and their potential to be used in UV research and for public education purposes. *Sci. Total Environ.* **706**, (2020).
4. Lucas, R. M. *et al.* Human health in relation to exposure to solar ultraviolet radiation under changing stratospheric ozone and climate. *Photochem. Photobiol. Sci.* **18**, 641–680 (2019).
5. Newman, P. A. & McKenzie, R. UV impacts avoided by the Montreal Protocol. *Photochem. Photobiol. Sci.* **10**, 1152–1160 (2011).
6. Sánchez-Pérez, J. F., Vicente-Agullo, D., Barberá, M., Castro-Rodríguez, E. & Cánovas, M. Relationship between ultraviolet index (UVI) and first-, second- and third-degree sunburn using the Probit methodology. *Sci. Rep.* **9**, (2019).
7. Garland, C. F., French, C. B., Baggerly, L. L. & Heaney, R. P. Vitamin D supplement doses and serum 25-Hydroxyvitamin D in the range associated with cancer prevention. *Anticancer Res.* **31**, 607 – 612 (2011).
8. Matsui, M. S., Pelle, E., Dong, K. & Pernodet, N. Biological rhythms in the skin. *Int. J. Mol. Sci.* **17**, (2016).
9. Raksasat, R. *et al.* Accurate surface ultraviolet radiation forecasting for clinical applications with deep neural network. *Sci. Rep.* **11**, 1–12 (2021).
10. Buntoung, S. *et al.* Development of maps for monthly average hourly vitamin D–weighted solar ultraviolet radiation over Thailand using a semi-empirical model with ground-based and satellite-based inputs. *Theor. Appl. Climatol.* **155**, 2499 – 2507 (2024).
11. Tang, L. *et al.* Sunlight ultraviolet radiation dose is negatively correlated with the percent positive of SARS-CoV-2 and four other common human coronaviruses in the U.S. *Sci. Total Environ.* **751**, (2021).
12. Moozhipurath, R. K. & Kraft, L. Association of lockdowns with the protective role of ultraviolet-B (UVB) radiation in reducing COVID-19 deaths. *Sci. Rep.* **11**, 1–9 (2021).
13. Pérez-Gilaberte, J. B. *et al.* Correlation between UV Index, Temperature and Humidity with Respect to Incidence and Severity of COVID 19 in Spain. *Int. J. Environ. Res. Public Health* **20**, (2023).
14. Balboni, E. *et al.* The influence of meteorological factors on COVID-19 spread in Italy during the first and second wave. *Environ. Res.* **228**, (2023).
15. Smith, T. P. *et al.* Differential responses of SARS-CoV-2 variants to environmental drivers during their selective sweeps. *Sci. Rep.* **14**, (2024).
16. WHO. Global Solar UV Index: A Practical Guide. A Joint Recommendation of the World Health Organization, World Meteorological Organization, United Nations Environmental Programme, and the International Commission on Non-Ionizing Radiation Protection. *World Heal. Organ. Geneva, Switzerland, 2002; ISBN 92-4-159007 6* (2002).
17. CIE. ISO/CIE 17166:2019(en) Erythema reference action spectrum and standard erythema dose. *CIE (International Comm. Illum. Journnal* (2019).
18. WHO. Ultraviolet Radiation. Environmental Health Criteria (EHC) 160. *World Heal. Organ. Geneva, Switz.* (1994).
19. Schmalwieser, A. W. *et al.* UV Index monitoring in Europe. *Photochem. Photobiol. Sci.* **16**, 1349–1370 (2017).
20. Kerr, J. B. & Fioletov, V. E. Surface ultraviolet radiation. *Atmos. - Ocean* **46**, 159–184 (2008).
21. Kosmopoulos, P. G. *et al.* Real-time UV index retrieval in Europe using Earth observation-based techniques: system description and quality assessment. *Atmos. Meas. Tech.* **14**, 5657–5699 (2021).
22. Hersbach, H. *et al.* The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999 – 2049 (2020).
23. Xia, Y. *et al.* Concurrent hot extremes and high ultraviolet radiation in summer over the Yangtze Plain and their possible impact on surface ozone. *Environ. Res. Lett.* **17**, (2022).
24. Xia, Y., Hu, Y., Huang, Y., Bian, J. & Zhao, C. Stratospheric ozone loss-induced cloud effects lead to less surface ultraviolet radiation over the Siberian Arctic in spring. *Environ. Res. Lett.* **16**, (2021).
25. Kaaya, I., Mansour, D.-E., Gebhardt, P., Weiss, K.-A. & Philipp, D. Modelling and Validation of Photovoltaic Degradation under Ultraviolet-Damp-Heat Conditions. in *Conference Record*

of the IEEE Photovoltaic Specialists Conference 986 – 990 (2021). doi:10.1109/PVSC43889.2021.9519050
26. Kusumaningtyas, S. D. A. *et al.* Comprehensive analysis of long-term trends, meteorological influences, and ozone formation sensitivity in the Jakarta Greater Area. *Sci. Rep.* **14**, (2024).
27. Smith, T. P. *et al.* Differential responses of SARS-CoV-2 variants to environmental drivers during their selective sweeps. *Sci. Rep.* **14**, (2024).
28. Hoffmann, L. *et al.* Investigating the spatiotemporal associations between meteorological conditions and air pollution in the federal state Baden-Württemberg (Germany). *Sci. Rep.* **14**, (2024).
29. TEMIS. UV index and UV dose: details of the calculation data version 2.x. (2023). Available at: https://www.temis.nl/uvradiation/product/uvi-uvd.html. (Accessed: 30th July 2025)
30. Madronich, S. Analytic formula for the clear-sky UV index. *Photochem. Photobiol.* **83**, 1537–1538 (2007).
31. Lamy, K. *et al.* UV-Indien network: Ground-based measurements dedicated to the monitoring of UV radiation over the western Indian Ocean. *Earth Syst. Sci. Data* **13**, 4275 – 4301 (2021).
32. Allaart, M., van Weele, M., Fortuin, P. & Kelder, H. An empirical model to predict the UV-index based on solar zenith angles and total ozone. *Meteorol. Appl.* **11**, 59 – 65 (2004).
33. McKenzie, R., Smale, D. & Kotkamp, M. Relationship between UVB and erythemally weighted radiation. *Photochem. Photobiol. Sci.* **3**, 252 – 256 (2004).
34. Diffey, B. L., Whillock, M. J. & McKinlay, A. F. A preliminary study on photoaddition and erythema due to UVB radiation. *Phys. Med. Biol.* **29**, 419–425 (1984).
35. Meloni, D., Casale, G. R., Siani, A. M., Palmieri, S. & Cappellani, F. Solar UV Dose Patterns in Italy. *Photochem. Photobiol.* **71**, 681 (2000).
36. De Backer, H. *et al.* Comparison of measured and modelled uv indices for the assessment of health risks. *Meteorol. Appl.* **8**, 267–277 (2001).
37. Koepke, P. *et al.* Comparison of Models Used for UV Index Calculations. *Photochem. Photobiol.* **67**, 657 – 662 (1998).
38. Mayer, B. & Kylling, A. Technical note: The libRadtran software package for radiative transfer calculations - Description and examples of use. *Atmos. Chem. Phys.* **5**, 1855 – 1877 (2005).
39. Mayer, B. *et al.* libRadtran User ' s Guide. *Ed. Libr. version 2.0.2* 1–155 (2017).
40. Qin, W., Wang, L., Wei, J., Hu, B. & Liang, X. A novel efficient broadband model to derive daily surface solar Ultraviolet radiation (0.280–0.400 μm). *Sci. Total Environ.* **735**, (2020).
41. Feister, U., Cabrol, N. & Häder, D. UV irradiance enhancements by scattering of solar radiation from clouds. *Atmosphere (Basel).* **6**, 1211–1228 (2015).
42. McKenzie, R. L., Paulin, K. J., Bodeker, G. E., Liley, J. B. & Sturman, A. P. Cloud cover measured by satellite and from the ground: Relationship to UV radiation at the surface. *Int. J. Remote Sens.* **19**, 2969–2985 (1998).
43. Staiger, H. *et al.* Hourly resolved cloud modification factors in the ultraviolet. *Atmos. Chem. Phys.* **8**, (2008).
44. Lamy, K. *et al.* Clear-sky ultraviolet radiation modelling using output from the Chemistry Climate Model Initiative. *Atmos. Chem. Phys.* **19**, 10087 – 10110 (2019).
45. Mayer, B. & Kylling, A. Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use. *Atmos. Chem. Phys.* **5**, 1855–1877 (2005).
46. Stamnes, K., Tsay, S.-C., Wiscombe, W. & Jayaweera, K. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.* **27**, 2502–2509 (1988).
47. Rothman, L. S. *et al.* The HITRAN 2004 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **96**, 139–204 (2005).
48. Blumthaler, M. & Ambach, W. SOLAR UVB-ALBEDO OF VARIOUS SURFACES. *Photochem. Photobiol.* **48**, (1988).
49. Chadyšiene, R. & Girgždys, A. Ultraviolet radiation albedo of natural surfaces. *J. Environ. Eng. Landsc. Manag.* **16**, (2008).
50. Cordero, R. R. *et al.* Ultraviolet radiation in the Atacama Desert. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **111**, 1301 – 1313 (2018).
51. Casale, G. R. *et al.* Extreme UV index and solar exposures at plateau rosà (3500ma.s.l.) in valle d'aosta Region, Italy. *Sci. Total Environ.* **512–513**, 622 – 630 (2015).
52. Vitt, R. *et al.* UV-index climatology for europe based on satellite data. *Atmosphere (Basel).* **11**, (2020).

53. McKenzie, R., Bodeker, G., Scott, G., Slusser, J. & Lantz, K. Geographical differences in erythemally-weighted UV measured at mid-latitude USDA sites. *Photochem. Photobiol. Sci. Off. J. Eur. Photochem. Assoc. Eur. Soc. Photobiol.* **5**, 343–352 (2006).
54. Hall, E. S. Comparison of five modeling approaches to quantify and estimate the effect of clouds on the Radiation Amplification Factor (RAF) for solar ultraviolet radiation. *Atmosphere (Basel).* **8**, (2017).
55. Rousseeuw, P. J. & Croux, C. Alternatives to the median absolute deviation. *J. Am. Stat. Assoc.* **88**, (1993).
56. Seckmeyer, G. *et al.* Instruments to Measure Solar Ultraviolet Irradiance. Part 2: Broadband Instruments Measuring Erythemally Weighted Solar Irradiance. in 54 (2005).
57. Pitkanen, M. R. ., Wandji, W. & Arola, A. Validation Report of the CAMS UV processor Issue # 18 December 2020 / January-February 2020 (DJF) CAMS-72 : Solar radiation products. CAMS72_2018SC2_D72.2.1.1-2020Q2_UV_VAL_202006_v2. (2020).
58. Craig, J. C. Python Sun Position for Solar Energy and Research. (2021). Available at: https://levelup.gitconnected.com/python-sun-position-for-solar-energy-and-research-7a4ead801777.