CIE, VITAMIN D AND DNA DAMAGE: A SYNERGETIC STUDY IN THESSALONIKI, GREECE

Melina Maria Zempila [(1*,2)], Michael Taylor [(2)], Ilias Fountoulakis [(2)], Maria Elissavet Koukouli [(2)], Alkiviadis Bais[(2)], Antii Arola [(3)], Jos van Geffen[(4)], Michiel van Weele[(4)], Ronald van der A[(4)], Natalia Kouremeti [(5)], Stelios Kazadzis [(5)], Chariklia Meleti [(2)], Dimitrios Balis [(2)]

[(1)] *Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523, U.S.A.*
Email: melina.zempila@colostate.edu
[(2)] *Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, PO Box 149, 54124, Thessaloniki, Greece.*
[(3)] *Finnish Meteorological Institute, FI-70211 Kuopio, Finland.*
[(4)] *Koninklijk Nederlands Meteorologisch Instituut (KNMI), De Bilt, the Netherlands.*
[(5)] *Physikalisch-Meteorologisches Observatorium Davos, World Radiation Center, Dorfstrasse 33, 7260 Davos Dorf, Switzerland.*

ABSTRACT

The present study aims to validate different approaches for the estimation of three photobiological effective doses: the erythemal UV, the vitamin D and that for DNA damage, using high temporal resolution surface-based measurements of solar UV from 2005-2015. Data from a UV spectrophotometer, a multi-filter radiometer, and a UV radiation pyranometer that are located in Thessaloniki, Greece are used together with empirical relations, algorithms and models in order to calculate the desired quantities. In addition to the surface-based dose retrievals, OMI/Aura and the combined SCIAMACHY/Envisat and GOME/MetopA satellite products are also used in order to assess the accuracy of each method for deriving the photobiological doses.

1. INTRODUCTION

During the last few decades, the danger of overexposure to UV sunlight has been well analysed and a causal link has been established to skin diseases and cancer, and that mutation of DNA can be triggered by extreme UV-B doses [1]. Furthermore, the cutaneous production of vitamin D is also activated by spectral UV radiation, hence accurate knowledge of the safe UV doses for humans is paramount [2]. Of particular relevance is the *Commission Internationale de l'Éclairage* (CIE) action spectrum as a model for the susceptibility of skin to sunburn (erythema).

As a result of advances in the fields of photobiology and ground-based measurements of UV using different types of instrumentation, a variety of methods now exist to obtain erythemal, vitamin D and DNA damage dose rates. In parallel, space technology has been making huge steps forward to monitor the Earth's surface and atmosphere at higher spatial and temporal resolution and erythemal, vitamin D and DNA damage dose rates and doses can now be retrieved globally from surface backscatter radiation observations from different satellite sensors.

In this study, photobiological doses retrieved from ground-based measurements using empirical models and satellite estimates are cross-validated to assess their accuracy and potential utility.

2. DATASETS AND INSTRUMENTATION

2.1. Ground-based measurements

The calculation of the photobiological doses over Thessaloniki (40.63^{o}E, 22.96^{o}N), are based on measurements taken by three different types of instruments in continuous operation at the Laboratory of Atmospheric Physics of the Aristotle University of Thessaloniki (http://lap.physics.auth.gr).

Firstly, a Brewer MKIII spectrophotometer with serial number #086 (B086) is equipped with a double monochromator and measures the UV solar spectrum (286.5 - 363 nm) with a wavelength step of 0.5 nm using a triangular slit that has a full width at half maximum (FWHM) of 0.55 nm, within an interval of 7 minutes. The more simple, single monochromator Brewer with serial number #005 (B005) has been operational in Thessaloniki since 1982 and has been providing continuous, well-calibrated and documented total ozone column measurements [3, 4, 5]. The spectra used in this study have recently been subjected to quality control and re-evaluation [6] after which the remaining 1-sigma uncertainty is estimated to be ~5% [7] for wavelengths higher than 305 nm and solar zenith angles (SZA) smaller than 80°. For lower wavelengths and higher SZA the uncertainty is greater as a consequence of the photon noise that dominates due to the low recorded signal [8]. The B086 spectra have been extended to 400 nm using the SHICrivm algorithm [9] and weighted with the action spectra for: i) the erythemal dose [10], ii) the formation of vitamin D in the human skin [11], and iii) DNA damage [12]. For the DNA damage, the effective spectrum is normalized at 300 nm. The corresponding effective doses have been calculated by integrating the weighted spectra over the nominal wavelength range.

Secondly, a NILU-UV multi-filter radiometer has been fully operational in Thessaloniki since 2005 and forms part of the UVnet network of NILU-UV

radiometers (http://www.uvnet.gr). The NILU-UV with serial number 04103 (NILU103) provides one-minute measurements in 5 UV channels with nominal central wavelength at 302, 312, 320, 340 and 380 nm and a FWHM of 10 nm; while its sixth channel that measures the Photosynthetically Active Radiation (PAR) is used to determine cloud free cases based on the cloud detection algorithm proposed by [13]. Although the Brewer #086 measures the UV spectrum with high spectral resolution, the time frequency of the scans can vary from 30 to 60 minutes. Nevertheless, Brewer spectrophotometers are a very powerful means for calibrating other UV measuring instruments that provide higher temporal resolution measurements. By calibrating the NILU103 measurements with the B086 coincident irradiances, we estimate that the uncertainties of the NILU103 irradiance measurements used in this study are less than 5.5% [14].

Finally, a Yankee UVB-1 pyranometer (YES) operating also in Thessaloniki, provides one minute erythemal dose measurements with a spectral response very similar to the official erythemal action spectrum [10]. Using libRadtran radiative transfer model simulations [15] proper weighting factors are calculated with respect to SZA and the total ozone column (TOC) which are used to transform the YES measurements into erythemal irradiance [16]. In addition, the Brewer measurements have been used to correct the pyranometer observations for the degradation of its absolute spectral response and for sudden changes in the behaviour of the instrument. Thus, the datasets from the pyranometer and the NILU-UV radiometer are not completely independent since the Brewer instrument has had to be used for the calibration of both instruments.

2.2. Satellite measurements

Satellite estimates of erythemal UV, vitamin D and DNA damage effective dose time series from the OMI/Aura, SCIAMACHY/Envisat and GOME2/MetopA instruments are also used to provide a comparison for the surface-based estimates. The OMI/Aura surface UV irradiance data from 2005 to 2015 include the erythemally-weighted daily dose and erythemal dose rate both at the overpass time and also at local solar noon. Surface UV overpass data for Thessaloniki have been extracted from NASA's Aura Data Validation Centre, http://avdc.gsfc.nasa.gov/. The OMIUVB algorithm is described in the Algorithm Theoretical Basis Document [17] and examples of its validation may be viewed in [18, 14].

The SCIAMACHY/Envisat and GOME2/MetopA joint UV product, that includes the same sub-products as the OMI/Aura dataset, were downloaded from ESA's Tropospheric Emission Monitoring Internet Service (TEMIS), http://www.temis.nl. The data are described in detail in [19, 20, 21]. The effective spectrum suggested by [10] is used to derive the erythemal dose from the OMI/Aura, SCIAMACHY/ Envisat and GOME2/MetopA instruments. Using the effective spectra suggested by [11] and [12] the effective doses for the formation of the vitamin D and the DNA damage are also calculated from measurements from the GOME2/MetopA instrument.

3. METHODOLOGY

3.1. Effective UV doses from the Brewer spectrophotometer

The B086 spectra were processed by the SCHICrivm algorithm and extended to 400 nm and have been weighted with the action spectra for erythema [10], the formation of vitamin D in the human skin [11] and the DNA damage [12] and the corresponding effective doses have been calculated. The 1-sigma uncertainty of the derived effective doses for the erythema and the vitamin D is estimated to be ~5% since the contribution of photons with wavelengths lower than 305 nm (where the signal may be very low) is small. However, the uncertainty in the calculated effective dose for the DNA damage is higher because of the important contribution of lower wavelengths (very low signal levels) and may exceed 20% for SZAs near 80° or during overcast conditions. The effective doses from B086 are used for training a neural network (NN) to derive effective doses from NILU103 radiances. The B086 effective doses are also used to derive empirical relationships and to calculate the vitamin D and DNA damage effective doses from the YES UVB-1 pyranometer.

3.2. Effective UV doses from NILU-UV radiances using a neural network model

A feed-forward function-approximating NN model [22] was coded using MATLAB's object-oriented scripting language in conjunction with its Neural Network Toolbox [23]. As inputs, the NN has NILU103 irradiance measurements at 305, 312, 320, 340 and 380nm and temporal variables including the SZA, the day of the week (DOW) and the day of the year (DOY) and its sinusoidal components cos(DOY) and sin(DOY). As outputs, the NN calculates the biological UV products resulting from B086 response weighted spectra: i.e. the erythemal UV (CIE), vitamin D and DNA damage effective dose. The rationale behind including temporal variables in the inputs is that geophysical variables very often exhibit periodicity associated with an annual or diurnal cycle and are now commonly incorporated into atmospheric chemistry models [24]. A key feature of our modeling approach is signal to noise separation. The NN model is constructed using denoised time series of the NILU-UV irradiances and denoised time series of the photobiological products. Once constructed, the real (noisy) data is input to the model to calculate the real (noisy) photobiological outputs. In order to achieve this, we applied singular spectrum analysis (SSA) to separate the total trend-cycle plus periodicity from the total noise component

for each time series (see [25] for a review of this approach).

From the NILU103 data, a matrix of $n = 47,908$ co-located input-output vectors was extracted to train and validate the model. A subset of t-vectors was chosen randomly as the training set, with the remainder being used as a validation set containing ($v = n - t$) vectors. All output variables were found to correlate strongly and positively on all 5 of the irradiances ($0.922 \leq r \leq 0.995$), strongly anti-correlate with SZA ($-0.891 \leq r \leq -0.909$), and to weakly anti-correlate with the temporal variables. The input and output vectors used in our study were connected via 2 network layers – the first containing hidden neurons with hyperbolic tangent (*tanh*) activation functions and the second containing linear activation functions as depicted in *Fig. 1*.
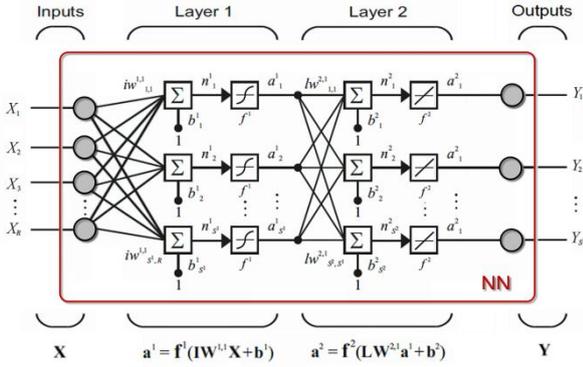
Figure 1. Schematic showing the neural connectivity between input and output parameters in the NILU-UV NN model.

The exact mathematical equation relating the outputs to the inputs for this type of NN model has been shown to have the matrix equation [26]:

$$\mathbf{Y} = f^2(\mathbf{LW}^{2,1}f^1(\mathbf{IW}^{1,1}\mathbf{X} + \boldsymbol{b^1}) + \boldsymbol{b^2}) \qquad (1)$$

where in this study: the matrix of input vectors is $\mathbf{X} =$ [Ir(305), Ir(312), Ir(320), Ir(340), Ir(380), SZA, DOY, Sin(DOY), Cos(DOY), DOW]^T and the concurrent matrix of output vectors is $\mathbf{Y}$ = [CIE, vitamin D, DNA]^T. The optimal NN architecture was then found by minimizing the mean squared error (MSE) between the NN estimates and Brewer reference output data for each NN in a grid of 100 NN architectures where the number of hidden neurons was varied from 5 to 15 and the proportion of training data (t/n) was varied from 50% 95% in steps of 5%. During each of 100 iterations of the learning process, the weights and biases of each NN are tuned with the Bayesian regularization back-propagation optimization algorithm [27] to minimize the MSE cost function over the set of input-output vectors. We have used the Bayesian regularization scheme using a Laplace prior discussed in [28]. As a result of this robustness analysis, the optimal NN was found to require 13 hidden neurons and a training to validation ratio of 90%:10%. The optimal NN is valid for the

range of parameters determined by the training data shown in Tab. 1 (temporal variables are not listed but have the following expected ranges: DOY=[0,366], Cos(DOY)=[-1,1], Sin(DOY)=[-1,1] and DOW=[1,7]).

Table 1. Range of validity of the trained optimal NN as determined by its input parameters (upper list) and output parameters (lower list).

Parameter	Min	Max	Mean	St. Dev.
Ir(305)	0	0.017	0.003	0.004
Ir(312)	0	0.229	0.064	0.055
Ir(320)	0	0.333	0.108	0.079
Ir(340)	0	0.678	0.252	0.159
Ir(380)	0	0.871	0.327	0.208
SZA	15.63	81.162	54.373	16.120
CIE	0	0.234	0.056	0.054
vitamin D	0	0.460	0.103	0.107
DNA	0	0.011	0.002	0.002

The optimal trained NN was then fed with the remaining ("unseen") input vectors from the 10% of the training data and its estimates are then compared against the target measurements of the output vector to evaluate its performance. The correlation of NILU103 NN estimates with target outputs was high (r=0.988 to 0.990) and with very low bias (0.000 to 0.011 absolute units).
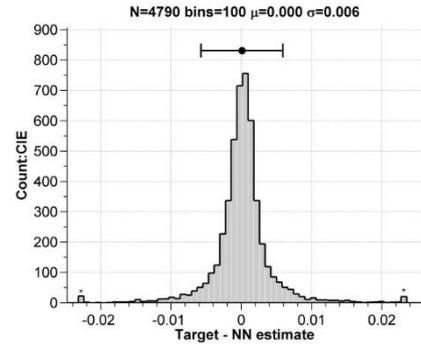
Figure 2. Histogram of differences for the validation of the NILU103 NN case of the erythemal UV dose (CIE). The mean (μ) and standard deviation (σ) are indicated.

Having trained and validated the NILU103 NN, we then proceeded to remove the constraint on having to co-locate input and output vectors and ran it in *unsupervised mode* (i.e. simulation mode) on the full record of nearly 2.5 million (2,471,529 records) input vectors of coincident NILU103 irradiances together with the associated temporal variables. On each day of the simulation record, the input-output vector closest to local noon and satellite overpass was stored. This was also used to extract all vectors within ± 30 minutes for estimation of output values closest to local noon time and satellite overpass time. In order to estimate daily doses rather than dose rates, the full record of input-

output vectors on each day was integrated with the Trapezium rule for days that did not contain gaps greater than 15-minutes in their record.

3.3 Effective UV doses from the YES UVB-1 pyranometer

The Yankee UVB-1 pyranometer provides measurements of the erythemal dose rates whereas the dataset's validity is monitored by coincident measurements from the B086. Using the effective doses derived from the B086, we adopted the empirical relationship suggested by [29] to convert erythemal irradiance to effective dose for the formation of vitamin D which, if we note that UVI = CIE/40, can be written:

$$vitamin\ D = f(TOC, UVI, \cos\theta) \qquad (2)$$

where TOC is the total ozone column, UVI is the UV index [10] and θ is the SZA. It was found that for UV index values below 2, the vitamin D is overestimated significantly and should be divided by the following correction factor (cf) obtained empirically by a least squares fit to the data:

$$cf = -0.086\ UVI^3 + 0.379\ UVI^2 - 0.575\ UVI + 1.317 \quad (3)$$

This correction, applied to the calculated vitamin D when UVI < 2, allowed us to derive vitamin D from the YES data in a way that avoids the limitation to clear sky conditions. In a similar way, the DNA damage effective spectrum was estimated from an empirical relationship that also consists of TOC, UVI and cosθ, but also involves the ratio between the UVI and the climatological value of UVI on each day (UVI_{clim}):

$$DNA = g(TOC, UVI, \cos\theta, UVI_{clim}) \qquad (4)$$

The empirical rule given by Eq. (4) was found to be valid for UVIs greater than 0.5. The daily mean TOC from the B005 was used in Eq. (2) and Eq. (4). In cases of missing data, daily climatological means derived from the 30-year record of B005 were used. Using the effective doses from the B086, we estimated that the 1-sigma uncertainty in the determination of vitamin D from Eq. (2) and Eq. (3) is smaller than 3% for UVI greater than 2 and raises to 10% for UVI near 0. The 1-sigma uncertainty in the calculation of the effective dose for the DNA damage are smaller than 7% for the range of used UVIs (i.e. greater than 0.5). The mean ratio between semi-simultaneous measurements of the clear sky erythemal irradiance from the Brewer B086 and the pyranometer (± 1 minute differences between the mean time of the spectral scan and the YES measurements) for SZAs below 80° for the period 2004 – 2014 is 1.00 ± 0.04, indicating that the uncertainty in the erythemal irradiance from the pyranometer is similar to that of the Brewer B086.

4. RESULTS & DISCUSSION

In the following section comparisons among the different datasets are performed.

4.1. NILU-UV and OMI /Aura CIE products

The OMI/Aura algorithm provides erythemal doses and dose rates at overpass time as well as at local noon. Both cases were investigated, while at the same time discriminating cloud free cases in two different ways: i) a cloud screening algorithm based on NILU-PAR measurements was used to define the NILU103 clear sky cases, ii) the limitation of cloud optical depth less than 20 was applied to satellite estimates in order to derive the satellite cloudless cases. In *Table 2* we present a statistical summary of the comparisons performed for the overpass and noon local time.

Table 2. Statistical analysis of the differences between erythemal dose rates provided by NILU103 and OMI/Aura for overpass and local noon time.

	Overpass			Local Noon		
	All Skies	OMI Clear	NILU Clear	All Skies	OMI Clear	NILU Clear
N counts	2152	1361	610	2426	2208	663
R^2	0.92	0.91	0.93	0.89	0.88	0.91
Mean (%)	14.07	9.97	13.28	18.47	14.38	17.68
STD (%)	49.22	42.77	49.63	68.03	58.87	62.85
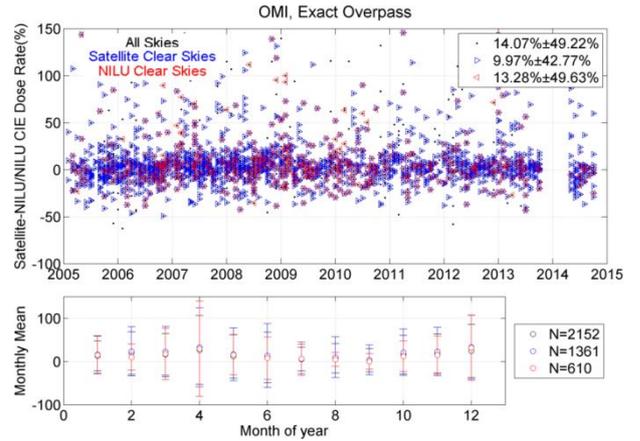
Figure 3. Percentage differences of the NILU103 and OMI erythemal dose rates at the overpass time (upper level). Monthly mean percentage differences are provided as well (lower level). NILU103 and OMI clear days are provided in blue and red triangles respectively.

For all skies cases the agreement between the ground-based overpass erythemal dose rates is 14.07% while the satellite overestimates by 18.47% at local noon. Limiting the dataset to cloud free cases based on OMI observations leads to smaller relative percentage differences, 9.97% and 14.38% for the overpass and the local noon time respectively. In all cases, the coefficient of determination is between 0.88 and 0.93.

The OMI dataset also includes the daily erythemal doses, i.e. daily integrals were calculated for days that did not have more than 15 minute gaps in their

record. The NILU103 daily values were characterized as clear when more than 70% of the measurements were identified as cloud free by the PAR cloud screening algorithm.
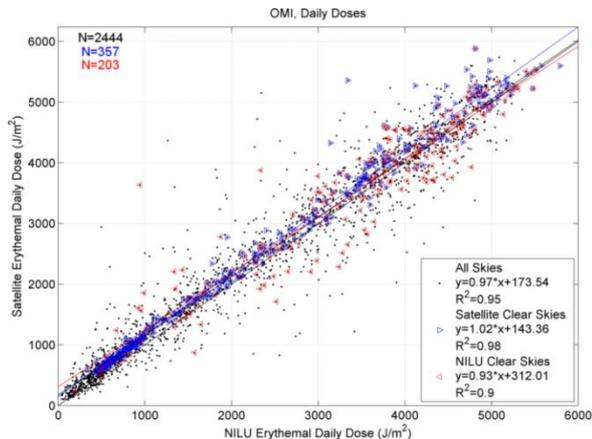
Figure 4. Scatter plot of daily erythemal values provided by OMI (y axis) and NILU103 (x axis) in J/m². Colour scale is used to denote the all skies from cloud free cases based on the satellite and ground-based criteria used in Figure 3.

The R² values are above 0.9 for all three cases (all skies, OMI- and NILU103-defined cloud free cases) with slopes close to unity. The OMI estimates tend to provide higher values when the doses are relatively low however in general they appear to underestimate the highest erythemal daily doses.

4.2. NILU-UV, SCIAMACHY/Envisat and GOME2/MetopA products

The SCIAMACHY/Envisat and GOME2/MetopA joint UV products are discussed in this section. Time series analysis and correlation statistics were performed on the daily erythemal, vitamin D and DNA damage reported doses while local noon UVI values were also included to the dataset.

As seen in *Figure 5* the daily all skies erythemal doses agree within ~15% while a seasonal pattern is observable. Although all daily datasets achieve high correlations of 0.92, 091 and 0.92 for CIE, vitamin D and DNA damage respectively, the observed differences for the latter two are not consistent. For vitamin D, the action spectrum used in satellite retrievals is shifted 3nm towards shorter wavelengths than that used in the NILU103 retrievals. This discrepancy leads to a significant underestimation of the satellite vitamin D doses by almost 50% for both all skies (-48.16% ± 25.17%) and clear sky (-46.57% ± 25.44%) cases. For DNA damage, the reverse is observed. Even though the applied action spectra are identical, the satellite overestimates by 48.84% for all skies and 55.4% for clear skies. These inconsistencies can be attributed to high uncertainty levels on B086 spectra at wavelengths lower than 305 nm that can be as high as 20% for SZAs of 80°. In addition, the aerosol

climatology used in the satellite retrieval algorithm and the uncertainties introduced in the algorithm retrieval at low UV-B wavelengths (e.g. ozone retrieval errors) can affect the accuracy of its performance. Both these limiting factors are under further investigation.
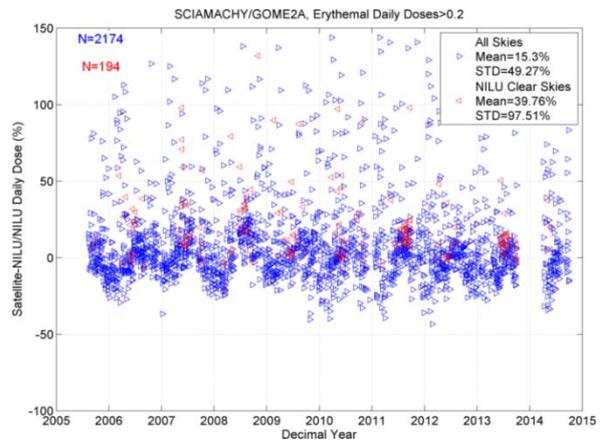
Figure 5 Time series of the relative percentage differences between the SCIA/GOME2A and NILU erythemal daily doses. Blue triangles depict the all skies daily data while red ones provide the comparisons for the characterized clear days.

Besides a mismatch in the absolute values of vitamin D and DNA, both NILU103 and SCIAMACHY/GOME2A daily datasets present high coefficients of determination and low bias (small y-intercepts). As seen in **Error! Reference source not found.** the common dataset of 2250 days results in a coefficient of determination of 0.92 for all skies and 0.85 for cloudless days.
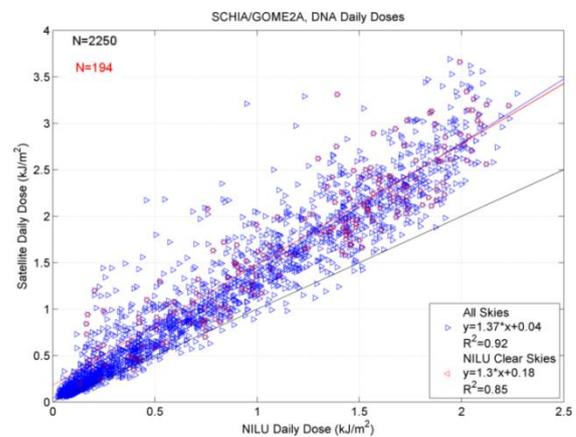
Figure 6. Scatter plot of daily erythemal values provided by the joint SCIA/GOME2A UV products (y-axis) and NILU103 (x-axis) in kJ/m². Colour scale is used to separate the all skies from cloud free cases based on the ground-based criteria used in Figure 5.

At this point we would like to briefly elaborate on the aspect of the determination of cloud free days. The optical geometry of the two monitoring systems is different and the point measurements of the NILU compared to the 0.5°x0.5° spatial analysis of the joined SCHIAMACHY/GOME2A product may be a source of

discrepancy. The algorithm's performance and the lower cut off filter of 70% of clear sky-characterized minute measurements, may also lead to higher differences. Furthermore, since satellite estimates are based on one daily observation, we expect that this could also augment the uncertainty on the daily doses estimated by the satellite algorithm.

4.3. NILU-UV and YES products

Following the appropriate methodologies already discussed in Sect. 3.3, vitamin D and DNA damage dose rates can be obtained from an erythemal-like measuring instrument, in this case a YES radiometer. Even though the YES data were corrected for the degradation of its absolute response and random incidences with B086 data, the validity of its measurements as absolute values can be used to evaluate the performance of the NN used to derive all of the biological dose products.
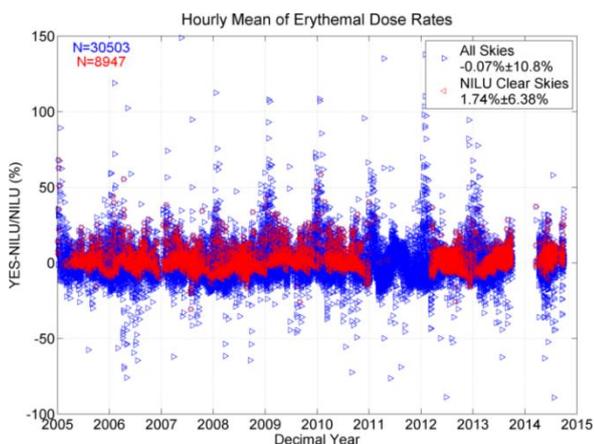
Figure 7 Time series of the relative percentage differences of hourly mean values for the NILU103 and YES erythemal dose rates. Red triangles indicate the cases were more than 70% of the data were identified as cloud free based on NILU103 PAR measurements.

The agreement between the two ground-based datasets is very good with close to zero mean differences and low standard deviation (< 11%). Similarly, it was found that NILU-UV calculated vitamin D hourly means also have a low (1.28%) overestimation in comparison with the YES data. While the YES data are implicitly corrected for any cosine errors through libRadtran model assimilations [15], it is possible that the correction of the cosine response of the NILU-UV may be slightly different and could result in a SZA dependency in the differences that dominates for SZAs of 70° and above for cloud free skies, as seen in *Figure 8*.

DNA damage presents higher scatter than that of the CIE and vitamin D dose rates due to higher sensitivity at lower UV-B wavelengths. The mean differences are of the order of ~5% ± 50% with the YES dataset providing lower values under all sky conditions than those from NILU103. On the contrary, under cloud free skies, YES data overestimates this dose by only 2.23%. In both

cases the correlation of the two datasets is close to unity ($R^2 \approx 0.99$) suggesting close agreement over the full range.
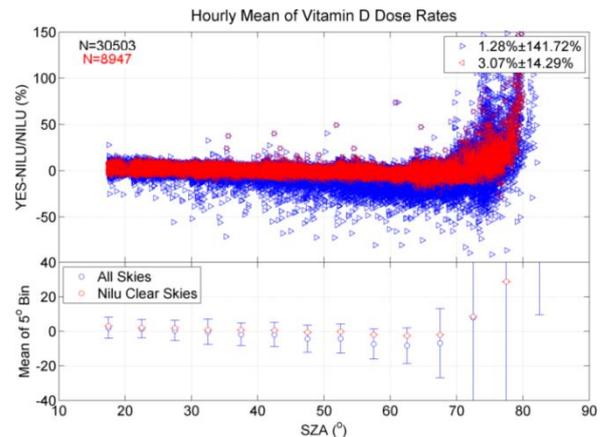
Figure 8 Hourly mean relative percentage differences of vitamin D estimates form the YES and NILU103 radiometers against SZA (upper panel) and the same datasets averaged on 5° SZA bins (lower panel).
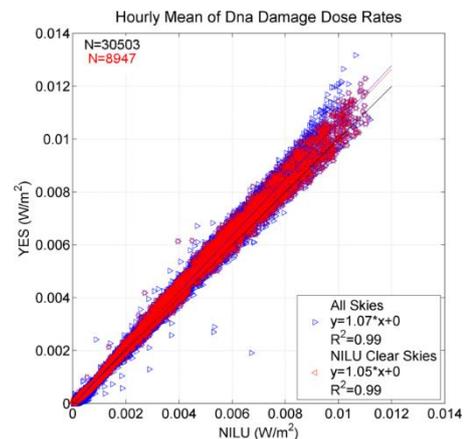
Figure 9 Scatter plot of hourly mean DNA damage doses as estimated from the YES (y-axis) and the NILU103 (x-axis) radiometers. Cloud free cases are distinguished by red triangles as in Figure 7.

5. CONCLUSIONS

In this work ground-based measurements, model estimates, and satellite-retrievals of the important photobiological UV products - CIE, vitamin D and DNA damage effective dose have been produced, compiled and compared so as to thoroughly discuss their accuracy and limitations at the mid-latitude UV and Ozone monitoring station in the Laboratory of Atmospheric Physics of the Aristotle University of Thessaloniki, Greece.

We show how a NN can be trained on NILU-UV multi-filter radiometer irradiances at 5 different wavelengths together with weighted action-spectra from a Brewer MKIII spectrophotometer to produce 1-minute time series of erythemal CIE, vitamin D and DNA

damage dose rates. Furthermore, while the NN estimated erythemal UV dose rates can be directly compared with YES calibrated measurements, we show how appropriate methodologies can be applied to the original YES data to also produce vitamin D and DNA damage dose rates at the same temporal resolution as the NILU-UV instrument to enable a ground-based verification and evaluation of NN model estimates resulting from NILU103 measurements. The comparisons with the full YES dataset revealed very good agreement. In particular:

- The hourly means of erythemal UV (CIE) dose rates from temporally aligned NILU-UV NN and the YES ground-based datasets (30,503 coincident 'all skies' records) did not show percentage fractional differences of more than 0.07% and had a moderately low standard deviation of 10.8%.
- For vitamin D, the disagreement was approximately 1.3% for 'all skies' data but showed high variance (141.7%) associated with an SZA dependence at high zenith angles. For cloud free cases this effect diminished (14.3%).
- The DNA dose rates, the most demanding of the three doses discussed in this study, agree to within about 5%, dropping to 2.23% for the cloud free cases.

In the context of space-retrievals of the photobiological doses, satellite estimates from OMI/Aura, SCIAMACHY/Envisat and GOME2/MetopA instruments were used, and which are provided as dose rates around the overpass time or the local noon time, as well as daily doses. The comparisons of the NILU-UV derived products with the satellite retrievals revealed the following major points:

- OMI exact overpass time comparisons show an agreement of 14.07%, 9.97% and 13.28% for all skies, and satellite- and NILU-derived cloud free cases when compared with the erythemal UV dose rates obtained by the NILU-UV NN model.
- Local noon time values reported by OMI have higher differences with respect to NILU-UV NN model estimates erythemal UV than those at overpass time but R^2 values are consistently high and in the range 0.88 to 0.93 for both all skies and clear sky cases at both times.
- OMI daily erythemal UV doses result in R^2 values of 0.90-0.98 when compared those calculated from integrating nearly complete daily records of NILU103 data.
- SCIAMACHY/GOME2A erythemal UV daily product revealed, on average, a 15% overestimation with respect to NILU103 daily doses. Despite the presence of a visually apparent seasonal pattern, the R^2 value was again found to be robustly high and equal to 0.92.
- For the vitamin D from SCIAMACHY/GOME2A daily doses, the differences between the two action

spectra of the satellite and ground estimations resulted in higher differences of 50%, with the satellite significantly underestimating the dose - possibly due to a shift of the action spectrum towards lower UV wavelengths.

- The SCIAMACHY/GOME2A DNA damage dose was also found to be overestimated by the satellite algorithm by almost 50%. This behaviour may be attributed to the higher uncertainty of the B086 spectrophotometer for wavelengths less than 305nm. Furthermore, it is possible that the *a priori* aerosol information that the satellite retrieval algorithm uses as input is contributing to high discrepancies in the shorter wavelength part of the UV-B spectral region where errors in measuring the total ozone column can have a high impact in accurate retrieving the DNA damage dose, as well as the vitamin D like dose.

In conclusion, this comprehensive study has revealed the merits, limitations and accuracy of ground-based and satellite-based estimates of erythemal UV, vitamin D and DNA damage doses and dose rates. Although calibration procedures, *a priori* information and constraints of the methods applied in the original datasets can limit the accuracy of the calculated photobiological products, increasing awareness of the harmful effects of overexposure to UV radiation means that such comparisons are of high importance.

6. REFERENCES

1. Lucas R., McMichael T., Smith W. and Armstrong, (2006), Solar Ultraviolet Radiation: Global burden of disease from solar ultraviolet radiation, World Health Organization, Public Health and the Environment.
2. McKenzie R.L. and Liley J.B., (2010), *UV Radiation in Global Climate Change: Measurements, Modeling and Effects on Ecosystems*. Chapter 2: Balancing the risks and benefits of Ultraviolet Radiation. Springer-Verlag Berlin Heidelberg and Tsinghua University Press, Beijing, China.
3. Bais, A. F., Zerefos, C. S., Ziomas, I. C., Zoumakis, N., Mantis, H. T., Hofmann, D. J., and Fiocco, G., (1985): Decreases in the Ozone and the S02 Columns Following the Appearence of the El Chichon Aerosol Cloud at Midlatitude. In: *Atmospheric Ozone*, Zerefos, C. S. and Ghazi, A. (Eds.), Springer Netherlands.
4. Meleti, C., Fragkos, K., Bais, A. F., Tourpali, K., Balis, D., and Zerefos, C. S., (2012): Thirty years of total ozone measurements at Thessaloniki with a MKII Brewer spectrophotometer, *Quadrennial Ozone Symposium 2012*, Toronto.
5. Zerefos, C., (1984): Evidence of the El Chichón stratospheric volcanic cloud in Northern Greece, *Geofísica Internacional*, 23.
6. Fountoulakis I., Bais A. F., Fragkos K., et al. (2016(a)) Short- and long-term variability of spectral solar UV irradiance at Thessaloniki,

Greece: effects of changes in aerosols, total ozone and clouds, *Atmos. Chem. Phys.*, 16, 2493-2505.

7. Garane K., Bais A. F., Kazadzis S., et al, (2006), Monitoring of UV spectral irradiance at Thessaloniki (1990-2005): data re-evaluation and quality control, *Ann. Geophys.*, 24, 3215-3228.

8. Fountoulakis I., Redondas A., Bais A. F., et al., (2016(b)), Dead time effect on the Brewer measurements: correction and estimated uncertainties, *Atmos. Meas. Tech.*, 9, 1799-1816.

9. Slaper H., Reinen H. A. J. M., Blumthaler M., et al. (1995), Comparing ground-level spectrally resolved solar UV measurements using various instruments: A technique resolving effects of wavelength shift and slit width, *Geophysical Research Letters*, 22, 2721-2724.

10. McKinlay A. F. and Diffey B. L., (1987), A reference action spectrum for ultraviolet induced erythema in human skin, *CIE J*, 6, 17-22.

11. Bouillon R., Eisman J., Garabedian M., et al. (2006), Action Spectrum for the Production of Previtamin D3 in Human Skin, *CIE Report No 174*, Vienna.

12. Setlow R. B., (1974), The Wavelengths in Sunlight Effective in Producing Skin Cancer: A Theoretical Analysis, *Proceedings of the National Academy of Sciences of the United States of America*, 71, 3363-3366.

13. Zempila M. M., Giannaros T. M., Bais A., et al. (2016(a)), Evaluation of WRF shortwave radiation parameterizations in predicting Global Horizontal Irradiance in Greece, *Renewable Energy*, Vol. 86 (pp 831-840), doi:10.1016/j.renene.2015.08.057.

14. Zempila M. M., Koukouli M. E., Bais, A., et al. (2016(b)), OMI/Aura UV product validation using NILU-UV ground-based measurements in Thessaloniki, Greece, *Atmospheric Environment*, under review.

15. Emde C., Buras-Schnell R., Kylling A., Mayer B., Gasteiger J., Hamann U., Kylling J., Richter B., Pause C., Dowling T. and Bugliaro L., (2015), The libRadtran software package for radiative transfer calculations (Version 2.0), *Geoscientific Model Development Discussions*, VOL. 8, doi: 10.5194 /gmdd-8-10237-2015

16. Lantz O.K. and Disterhoft P., (1998), Methodology for deriving Clear-sky erythemal calibration factors for UV broadband radiometers of the U.S. Central UV calibration facility, *Journal of Atmospheric and Oceanic Technology*, Vol. 16 (pp 1736-1752).

17. Krotkov, N. A., J. Herman, P. K. Bhartia, et al., (2002), OMI Surface UV Irradiance Algorithm, *Algorithm Theoretical Baseline Document: Clouds, Aerosols, and Surface UV Irradiance, P. Stammes (ed.)*, vol. III, ATBD-OMI-03, version 2.0, http://eospso.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-03.pdf, last accessed : May 2nd, 2016.

18. Tanskanen, A., A. Lindfors, A. Maatta ,et al. (2007), Validation of daily erythemal doses from Ozone Monitoring Instrument with ground-based UV measurement data, *J. Geophys. Res.*, 112, D24S44, doi:10.1029/2007JD008830.

19. Allaart, M., M. van Weele, P. Fortuin and H. Kelder, (2004), An Empirical model to predict the UV-index based on Solar Zenith Angle and Total Ozone, *Meteorological Applications (Royal Met. Society)*, doi:10.1017/S1350482703001130.

20. van Geffen, R. van der A, M. van Weele, et al., (2005), Surface UV radiation monitoring based on GOME and SCIAMACHY in: *Proceedings of the ENVISAT & ERS Symposium*, 6-10 September 2004, Salzburg, Austria, ESA publication SP-572.

21. de Laat, A. T. J., R. J. van der A, M. A. F. Allaart, et al., (2010), Extreme sunbathing: Three weeks of small total O3 columns and high UV radiation over the southern tip of South America during the 2009 Antarctic O3 hole season, *Geophys. Res. Lett.*, 37, L14805, doi:10.1029/2010GL043699.

22. Hornik, K., Stinchcombe M., White, H. (1989). Multilayer Feedforward Networks are Universal Approximators. *Neural Networks 2*, 359-366. doi:10.1016/0893-6080(89)90020-8.

23. Beale, M. H., Hagan, M. T., & Demuth, H. B. (2012). Neural network toolbox™ user's guide. In R2012a, *The MathWorks, Inc.,* 3 Apple Hill Drive Natick, MA 01760-2098,, www. mathworks. com.

24. Kolehmainen, M., Martikainen, H., Ruuskanen, J. (2001). Neural networks and periodic components used in air quality forecasting, *Atmos. Env.*, 35(5): 815-825, doi:10.1016/S1352-2310(00)00385-X.

25. Ghil M, Allen MR, Dettinger MD et al (2002) Advanced spectral methods for climatic time series, Rev Geophys 40(1:3):1-41.

26. Taylor, M., Kazadzis, S., Tsekeri, A., et al. (2014). Satellite retrieval of aerosol microphysical and optical parameters using neural networks: a new methodology applied to the Sahara desert dust peak, *Atmos. Meas. Tech*. 7, 3151-3175. doi:10.5194/amt-7-3151-2014.

27. Rumelhart, D.E., Hinton, G.E., Williams, R. J. (1986). Learning representations by backpropagating errors. *Nature* 323, 533–536.

28. Foxall, R.J., Cawley, G.C., Dorling, S.R. and Mandic, D.P. (2002). *Error functions for prediction of episodes of poor air quality. In Artificial Neural Networks-ICANN 2002* (pp. 1031-1036). Springer Berlin, Heidelberg.

29. Fioletov V. E., McArthur L. J. B., Mathews T. W. and Marrett L., (2009), On the relationship between erythemal and vitamin D action spectrum weighted ultraviolet radiation, *Journal of Photochemistry and Photobiology B: Biology*, 95, 9-16.