Science of the Total Environment 590–591 (2017) 92–106

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

NILU-UV multi-filter radiometer total ozone columns: Comparison with satellite observations over Thessaloniki, Greece

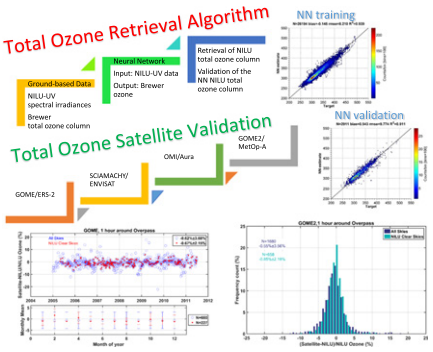
Melina Maria Zempila [a,b,*], Michael Taylor [b], Maria Elissavet Koukouli [b], Christophe Lerot [c], Konstantinos Fragkos [b], Ilias Fountoulakis [b], Alkiviadis Bais [b], Dimitrios Balis [b], Michel van Roozendael [c]

[a] USDA UV-B Monitoring and Research Program, Colorado State University, Fort Collins, CO 80523, USA
[b] Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, PO Box 149, 54124 Thessaloniki, Greece
[c] Belgian Institute for Space Aeronomy, BIRA-IASB, Brussels, Belgium

HIGHLIGHTS

- GOME, SCIAMACHY, OMI and GOME2 total ozone estimates in Thessaloniki Greece are evaluated
- NILU-UV irradiance measurements are used to produce high frequency total ozone estimates
- A neural network model converts NILU-UV data into total ozone. It can be applicable to any NILU-UV station within the range of the training data
- Both exact and 1 hour mean values around each satellite overpass are examined
- Ground-based and model comparisons reveal homogeneity while satellite sensors slightly underestimate ozone

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 November 2016
Received in revised form 1 February 2017
Accepted 21 February 2017
Available online 1 March 2017

Editor: D. Barcelo

Keywords:
Total ozone column
NILU-UV multi-filter radiometer
Brewer spectrophotometer
Neural network
GOME/ERS-2
SCIAMACHY/Envisat
OMI/Aura
GOME2/MetOpA

ABSTRACT

This study aims to construct and validate a neural network (NN) model for the production of high frequency (~1 min) ground-based estimates of total ozone column (TOC) at a mid-latitude UV and ozone monitoring station in the Laboratory of Atmospheric Physics of the Aristotle University of Thessaloniki (LAP/AUTh) for the years 2005–2014. In the first stage of model development, ~30,000 records of coincident solar UV spectral irradiance measurements from a Norsk Institutt for Luftforskning (NILU)-UV multi-filter radiometer and TOC measurements from a co-located Brewer spectroradiometer are used to train a NN to learn the nonlinear functional relation between the irradiances and TOC. The model is then subjected to sensitivity analysis and validation. Close agreement is obtained ($R^2 = 0.94$, RMSE = 8.21 DU and bias = −0.15 DU relative to the Brewer) for the training data in the correlation of NN estimates on Brewer derived TOC with 95% of the coincident data differing by less than 13 DU. In the second stage of development, a long time series (≥1 million records) of high frequency (~1 min) NILU-UV ground-based measurements are presented as inputs to the NN model to generate high frequency TOC estimates. The advantage of the NN model is that it is not site dependent and is applicable to any NILU input data lying within the range of the training data.
GOME/ERS-2, SCIAMACHY/Envisat, OMI/Aura and GOME2/MetOp-A TOC records are then used to perform a precise cross-validation analysis and comparison with the NILU TOC estimates over Thessaloniki. All 4 satellite TOC dataset are retrieved using the GOME Direct Fitting algorithm, version 3 (GODFIT_v3), for reasons of consistency. The NILU TOC estimates within ±30 min of the overpass times agree well with the satellite TOC retrievals with

* Corresponding author at: USDA UV-B Monitoring and Research Program, Colorado State University, Fort Collins, CO 80523, USA.
E-mail address: melina.zempila@colostate.edu (M.M. Zempila).

http://dx.doi.org/10.1016/j.scitotenv.2017.02.174

coefficient of determination in the range $0.88 \le R^2 \le 0.90$ for all sky conditions and $0.95 \le R^2 \le 0.96$ for clear sky conditions. The mean fractional differences are found to be $-0.67\% \pm 2.15\%$, $-1.44\% \pm 2.25\%$, $-2.09\% \pm 2.06\%$ and $-0.85\% \pm 2.19\%$ for GOME, SCIAMACHY, OMI and GOME2 respectively for the clear sky cases. The near constant standard deviation (~±2.2%) across the array of sensors testifies directly to the stability of both the GODFIT_v3 algorithm and the NN model for providing coherent and robust TOC records. Furthermore, the high Pearson product moment correlation coefficients ($0.94 < R < 0.98$) testify to the strength of the linear relationship between the satellite algorithm retrievals of TOC and ground-based estimates, while biases of less than 5 DU suggest that systematic errors are low. This novel methodology contributes to the ongoing assessment of the quality and consistency of ground and space-based measurements of total ozone columns.

1. Introduction

Since the mid-80s, after the discovery of ozone depletion in Antarctica (Farman et al., 1985), ozone has been one of the most studied trace gases due to its impact both in the stratosphere as well as in the troposphere. It is well known that ozone is responsible for the heating of the stratosphere by absorption of the harmful UV portion of the solar spectrum (de Forster et al., 1997; IPCC, 2013; Ramaswamy et al., 1996; Shine et al., 1995). Although stratospheric ozone protects the Earth ecosystem and humans from the harmful effects of exposure to solar UV radiation, its existence in the lower part of the atmosphere triggers photochemical reactions with detrimental impacts on plants and humans (Fowler et al., 2009; Sitch et al., 2007; WHO, 2003). Ozone is also an important tropospheric greenhouse gas and short-lived climate pollutant as is considered to be an essential climate variable (Shindell et al., 2012; Solomon et al., 2007).

Although the Montreal Protocol led to a reduction of the ozone depleting substances (Mäder et al., 2010; Shepherd and Jonsson, 2008), it is not yet clear that the stabilization, and the expected recovery predicted to be around 2050, is a result of a decrease in ozone depleting substances (WMO, 2015). Since the magnitude of the year to year variability of the total ozone column (TOC) is quite large (Fragkos et al., 2014) and the period since the stabilization of the TOC is not sufficient enough to arrive at a safe conclusion (Vyushin et al., 2010), the need for further investigation of the parameters affecting the total amount of this important trace gas is still an imperative. In this regard, high temporal frequency of point measurements of TOC provided by ground-based instruments in a global network of ozone monitoring stations are of paramount importance to help cross-validate satellite observations and TOC retrieval algorithms in general.

This study therefore aims to demonstrate how ground-based point measurements of TOC made in the global network of ozone monitoring stations, can be extended to high temporal resolution and cross-validated with satellite observations. To achieve this, we use ten years of coincident solar UV irradiance measurements from a Norsk Institutt for Luftforskning (NILU)-UV multi-filter radiometer and TOC retrievals from a co-located Brewer spectroradiometer to train a neural network (NN). In a similar vein, NN models have been developed to retrieve TOC from GOME/ERS2 spectral data (Müller et al., 2002). Rather than use TOC derived directly from ground-based instruments as we do here, Fan et al. (2014) used simulations from a radiative transfer model and a look-up table to train a NN with NILU-UV irradiances to produce combined estimates of TOC and the cloud optical depth; and found estimated TOC values to be comparable to those retrieved by OMI/Aura. The advantages of the proposed method described in this study are:

(i) no radiative model estimations are required; thus there is no need for other a priori input data such as aerosol optical depth or single scattering albedo,
(ii) aerosol effects are implicitly encoded in the spectral information contained at 5 different UV wavelengths; an effect that was not taken into account in the method of Fan et al. (2014) and which was deemed responsible for the under-estimation of NILU-retrieved TOC (OMI was also found to under-estimate TOC by almost 2.5% in other similar studies),
(iii) the results are directly compared with Brewer ozone retrievals and are accurate for a range of cloudiness conditions.

Although the method is developed using a long (decadal) time series of high frequency (~1-min) observations at one location (a UV monitoring station in Thessaloniki), this site samples a broad range of atmospheric conditions including widely varying aerosol loads due to dust incursions and biomass burning influxes, a range of distinct synoptic patterns and cloudiness conditions, and seasonal variation in both UV levels and solar zenith angles (SZA). Thus, the trained NN can be applied to data from any other NILU-UV station that is spanned by the min-max range of values used to train and validate the NN model (please see Table 3 *for details of the range of validity*). We would like to emphasize that here, our aim is not to produce a NN of global generality, rather it is to present and validate a new methodology that can easily be adopted locally.

The uncertainty in the NILU irradiances used in this study is assessed based on calibration of the NILU-UV channels using measured spectra provided by a second Brewer spectrophotometer also operating in Thessaloniki. We then feed the NN model with a long time series (≥1 million records) of high frequency (~1 min) NILU-UV ground-based measurements to generate high frequency, high accuracy TOC estimates which then provide coincidences with satellite retrieved TOC at overpass time of the OMI/Aura, GOME/ERS-2, GOME2/MetOp-A and SCIAMACHY/Envisat sensors. Although estimates of TOC in such high frequency are excessive, they provide the ability for variable time integrations while they provide larger statistical sample. Furthermore, comparisons of short time window averaging can reveal any possible discrepancies between the compared datasets regarding their angular behavior and/or their performance under various cloudiness conditions.

The rest of this paper is laid out as follows. In Section 2 we present the ground-based NILU-UV and Brewer instruments and datasets, as well as the GOME/ERS-2, SCIAMACHY/Envisat, OMI/AURA and GOME2/MetOp-A satellite-based TOC datasets used in this study. In the methodology Section 3, we describe the design and validation of the neural network (NN) model used to accurately estimate TOCs from coincident NILU-UV irradiances and Brewer derived TOC values. In Section 4 we present results of a detailed comparison of the NILU-UV derived TOC estimates with the overpass, satellite-based TOC retrievals. Finally, we sum up in Section 5 with our major findings.

2. Datasets and instrumentation

2.1. Ground-based measurements

Total ozone columns over Thessaloniki (40.63°N, 22.96°E) are being provided by a suite of instruments in continuous operation at the Laboratory of Atmospheric Physics of the Aristotle University of Thessaloniki (LAP/AUTh, http://lap.physics.auth.gr). The ones pertinent to this work are presented below.

2.1.1. The Brewer spectrophotometers

A SCI-TECH Inc. (http://www.sci-tec-inc.com/) Brewer spectrophotometer with serial number 005 (Brewer#005) has been operational in Thessaloniki since 1982 and has been providing continuous total ozone column measurements (Bais et al., 1985; Meleti et al., 2012). Brewer#005 has been calibrated in the past a number of times against the travel reference Brewer (Brewer#017; Meleti et al., 2012), while its latest calibration was held during the X intercomparison campaign of the Regional Brewer Calibration Center-Europe (RBCCE) at "El Arenosillo" Atmospheric Sounding Station, INTA (Huelva, Spain) during May 27–June 05, 2015. The agreement of Brewer#005 against the standard RBCCE traveling reference Brewer#185 during the blind days of the campaign was found to be quite satisfactory (within ±1%), while after the final calibration applied it was below ±0.5% (Redondas et al., 2016). In addition, the high quality TOC measurements from Brewer#005 can be depicted by the fact that the observed differences between Brewer#005 and satellite TOC retrievals are similar with those for other Brewers in the same latitude belt like Thessaloniki, as can be viewed in the official EUMETSAT Satellite Application Facility for Atmospheric Composition and UV Radiation Ozone Validation Web Services, http://lap3.physics.auth.gr/eumetsat/. In this study, we used TOC data from 2005 to 2014 which are available at the WOUDC database (www.woudc.org).

For the calibration of the NILU-UV measured irradiances (Section 2.1.2), comparisons with the double monocromator Brewer#086 were performed. Brewer#086 measures the spectrum of the global solar UV irradiance at wavelengths from 286.5 to 363 nm with a step of 0.5 nm using a triangular slit with a full width at half maximum (FWHM) of 0.55 nm. The spectra for the period 2005–2014, used in the present study, have been subjected to quality control and re-evaluation (Fountoulakis et al., 2016), after which the 1-sigma uncertainty in the final product is lower than ~5%, for wavelengths higher than 305 nm and solar zenith angles, smaller than 80° (Garane et al., 2006). The high quality of the used Brewer#086 spectra is also ensured by the good agreement between near-simultaneous measurements (Garane et al., 2006) of the two Brewers (Brewer#086 and #005). The mean value and the corresponding 1-sigma standard deviation of the monthly mean ratios between the 305–325 nm integrals, for the period 1994–2014, is 0.99 ± 0.02%.

In the past, the two Brewers have participated in a number of field campaigns where the good quality of their products has been verified (e.g. Bais et al., 2001, 2005; Balis et al., 2002). In addition, the high quality of both the TOC and UV datasets is outlined in a large number of older (Bais et al., 1996, 1993; Glandorf et al., 2005; Kazadzis et al., 2009; Meleti et al., 2009; Zerefos, 2002; Zerefos et al., 2012) and more recent articles (Fountoulakis et al., 2016; Fragkos et al., 2015; Zempila et al., 2016c). Detailed information regarding the calibration and re-evaluation of the UV spectra can be found in Garane et al. (2006) while methodological details of how TOC is retrieved with the Brewer#005 are presented in full in Section 3.1.

2.1.2. The NILU-UV multi-filter radiometer

A NILU-UV multi-filter radiometer with serial number 04103 ('NILU'), has been fully operational in Thessaloniki since 2005 and forms part of the network of NILU-UV radiometers (http://www.uvnet.gr). The NILU instrument provides one-minute measurements in 5 UV channels with nominal central wavelength at 302, 312, 320, 340 and 376 nm and a FWHM of 10 nm; while its sixth channel that measures the photosynthetically active radiation (PAR) (Zempila et al., 2016a) is used to determine cloud free cases based on the cloud detection algorithm proposed by Zempila et al. (2016b). The instrument has an internal temperature stabilizer, while the built in data logger has the ability to store 1-minute data for a period of 15 days.

In order to assure the quality of the measurements, a range of checks is performed every day. The internal temperature and the dark signal are recorded and monitored for their stability, while the diffuser of the instrument is cleaned on a regular basis. Furthermore, the raw data of all channels are plotted against coincidence data from a UVA radiometer manufactured by EKO instruments (http://eko-usa.com/) (for the NILU measurements at nominal wavelengths 340 and 376 nm) and a CM21 Kipp & Zonen pyranometer (for the PAR measurements) placed in the station. These comparisons aim to detect any discrepancies due to random incidences and to verify the synchronization of all instruments. The calibration and ensuing error analysis is described in detail in Zempila et al. (2016c).

The NILU dataset was subjected to intercomparisons with the Brewer#086 UV irradiance data for the whole period under investigation according to the methodology proposed by Zempila et al. (2016c). The uncertainties of the calibrated NILU irradiances based on Brewer#086 measurements are presented in Table 1.

For the estimation of the total uncertainty in the NILU irradiance data, the uncertainty in the measurements of the Brewer#086 has to also be quantified. Garane et al. (2006) stated that the estimated error in Brewer#086 spectra is less than 5% for wavelengths higher than 305 nm. According to error propagation methods ($\sqrt{[(5\%)^2 + (\text{NILU-UV channel error }\%)^2]}$, more details can be found in Zempila et al. (2016a, 2016c)) the uncertainty of the NILU calibrated data used in this study is 6.4% for the 305 nm irradiances and less than 5.4% for the greater wavelength measurements.

2.2. The GODFIT_v3 satellite algorithm and TOC record

Total ozone column records from GOME/ERS-2, SCIAMACHY/Envisat, OMI/Aura and GOME-2/MetOp-A have been reprocessed with the European Space Agency's Climate Change Initiative GODFIT (GOME-type Direct FITting) version 3 algorithm (Lerot et al., 2014). This algorithm is based on the direct-fitting of reflectances simulated in the Huggins bands to the observations, and is an evolved and upgraded algorithm from the version implemented in the operational GOME Data Processor v5 (van Roozendael et al., 2012). Inter-sensor comparisons and ground-based validation of this dataset (Koukouli et al., 2015) indicate that these ozone data sets are of unprecedented quality with stability better than 1% per decade, a precision of 1.7%, and systematic uncertainties less than 3.6% over a wide range of atmospheric states. In Table 2 the relevant information on the four instruments discussed further below is given.

Even though it has been already amply demonstrated that, on a global scale and at the monthly mean timescale, the TOC record of these four instruments has reached an unprecedented level of harmonisation, in the following, we compare the satellite TOC to the NILU TOC separately for each instrument in order to provide a thorough analysis of the specific characteristics of each one. Since we are able to constrain the temporal criteria of coincidence to within ±1 min should we wish to, the comparisons with the NILU NN model estimates, provide a unique opportunity to reveal possible causes of any remaining discrepancy between the four datasets due to the, admittedly, small difference in overpass time.

Table 1
Calibration uncertainties of the NILU irradiances based on Brewer#086 spectral measurements for the decade 2005–2014. The number of coincidences, the mean irradiances reported from Brewer#086 and the ratio (NILU/Brewer) of the mean values are also provided for each channel.

	302 nm	312 nm	320 nm	338 nm	376 nm
N	2105	3257	3575	4819	3364
NILU/Brewer	1	1	1	1	1
Brewer (W/m^2)	0.006	0.0759	0.12	0.29	0.37
STD	0.0734	0.039	0.03	0.03	0.03
RMSE	0.00	0.02	0.03	0.07	0.1
Error (%)	±4.05	±1.90	±1.73	±1.65	±1.57

Table 2
The characteristics of the four different sensors and platforms presented in this work.

Instrument platform	GOME ERS-2	SCIAMACHY ENVISAT	OMI Aura	GOME-2 MetOp-A
Spectral resolution	0.20 nm	0.26 nm	0.45 nm	0.26 nm
Spatial resolution	320 × 40 km^2	60 × 30 km^2	13 × 24 km^2	80 × 40 km^2
Swath width	960 km	960 km	2600 km	1920 km
Time period studied	Jan 1996 to June 2011	Jan 2003 to April 2012	Jan 2005 to Dec 2014	Jan 2007 to Dec 2014
Mean equatorial crossing time	10:30 L.T.	10:00 L.T.	13:45 L.T.	09:30 L.T.
Mean Thessaloniki visiting time	09:30 U.T.	09:00 U.T.	11:45 U.T.	08:30 U.T.
Main instrument reference	Burrows et al. (1999)	Bovensmann et al. (1999)	Levelt et al. (2006)	Hassinen et al. (2016)
Main GODFIT_v3 Algorithm Reference	Lerot et al. (2014)			
Main GODFIT_v3 Validation Reference	Koukouli et al. (2015)			

3. Methodology

In this section, a brief summary of the Brewer TOC retrieval is given, followed by a description of how the NN model was set up and validated to provide TOC based on the NILU-UV measurements and concurrent Brewer TOC values.

3.1. TOC values derived from Brewer spectrophotometer measurements

For TOC retrieval, the Brewer#005 spectrophotometer measures the direct solar irradiance at five distinct wavelengths, λ_i (nominally: 306.3, 310, 313.5, 316.8, and 320 nm). The TOC, is then derived using a relationship that compares the linear combinations of differential absorption and scattering coefficients with the direct spectral irradiances, at the four longest wavelengths (λ_2 to λ_5) (Kerr et al., 1981):

$$\mathrm{TOC} = \frac{MS9 - \mathrm{ETC}}{\mu * \Delta a} \quad (1)$$

where, ETC is the extraterrestrial constant which is defined during the calibration of the instrument, $\Delta\alpha$ is the differential absorption coefficient, μ is the relative optical air mass of the ozone layer, and *MS9* is defined from:

$$MS9 = \sum_{i=2}^{5} w_i \ln I_i \quad (2)$$

where I_i are the measured intensities at the respective wavelengths, corrected for the dark signal, dead time, and Rayleigh scattering; the weighting coefficients $w_i = (0, +1, -0.5, -2.2, +1.7)$ were selected to eliminate the effect of aerosol extinction and the absorption of SO_2. The differential ozone absorption coefficient $\Delta\alpha$ is derived from the Bass and Paur (1985) ozone absorption coefficients a_i for a fixed temperature of −45 °C:

$$\Delta a = \sum_{i=2}^{5} w_i a_i \quad (3)$$

3.2. TOC values derived from NILU-UV irradiances using a neural network model

In this section we describe the construction, training and validation of a feed-forward function-approximating NN model (Cybenko, 1989; Hornik et al., 1989) to retrieve total ozone from NILU-UV measured irradiances. As inputs, the NN model has NILU irradiance measurements at 302, 312, 320, 338 and 376 nm and temporal variables including the solar zenith angle (SZA), the day of the year (DOY) and its sinusoidal components cos(DOY) and sin(DOY), and the day of the week (DOW). The temporal variables, DOY, cos(DOY) and sin(DOY) were chosen to describe the annual cycle of ozone abundance, while it was proven that DOW does not have an impact on the retrieval (Fig. 1). The target variable for training the NN is the instantaneous Brewer#005 TOC. Once the NN is trained, it can be applied to time series of NILU irradiances in order to retrieve the NILU TOC estimates. The development of the NN model configuration is provided in Flow Chart 1.

The rationale behind including temporal variables in the model is that geophysical variables very often exhibit periodicity associated with an annual or diurnal cycle and are now commonly incorporated into atmospheric chemistry models (Kolehmainen et al., 2001; Taylor et al., 2016). While it is not expected that temporal variables like DOW will have a strong causal impact on TOC, we have adopted a pragmatic approach avoiding a priori assumptions - whereby they are first all included and then their redundancy is assessed through an analysis of neural weightings (see below). Another key step in preparing our data for model construction, was normalization to remove potentially undesirable variances that arise from parameters having very different min-max ranges. In the NN model we develop, all input and output matrices were pre-processed by converting each parameter into z-scores so that they have a mean equal to zero and a standard deviation of 1. In addition, a random number generator was used to ensure that identical initial weights were used in each run so that NN models with the same initial conditions could be compared. In particular, the twister algorithm (Matsumoto and Nishimura, 1998) based on the Marsenne prime (219937–1) was called with a constant integer seed value to return a single uniformly distributed pseudo-random number in the interval (0,1).

From the NILU data, a matrix of $n = 29{,}105$ co-located input-output vectors was extracted. 90% of this data (26,194 records) was used for NN training and 10% (2911 records) for NN validation. The Brewer derived TOC was not found to correlate positively or negatively on any of the 5 NILU-UV irradiances or the SZA ($-0.120 \le r \le 0.119$). It was found to strongly anti-correlate with the DOY ($r = -0.614$), and to strongly correlate with sin(DOY) with an r value of 0.717 as shown in Fig. 1.

There is also no evidence of a strong linear correlation between Brewer TOC and cos(DOY) or the DOW. At this point it should be mentioned that the pairwise linear Pearson correlation coefficients presented in Fig. 1, are only an indicator of the linear relationship between the input parameters and the expected output. While seasonal variation and cloud effects modulate ozone, it is not clear from the linear analysis above how combined effects manifest themselves in this framework. As such, we have retained them in the nonlinear NN model.

A key feature of our nonlinear 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 Brewer#005 TOC. Once constructed, the real (noisy) data is input to the model to calculate the real (noisy) NILU TOC output estimates. 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 irradiance time series (see Ghil et al. (2002) for a review of this approach). We set the size of the SSA window (i.e. the number of time delay versions of each time series) equal to 50. A clear break in the singular spectrum of ranked eigenvalues versus eigenvalue (or EOF) was observed for each of the irradiance input variables and also the TOC output respectively, separating the trend and periodic components (the signal) from a long tail of noise components. The separation was

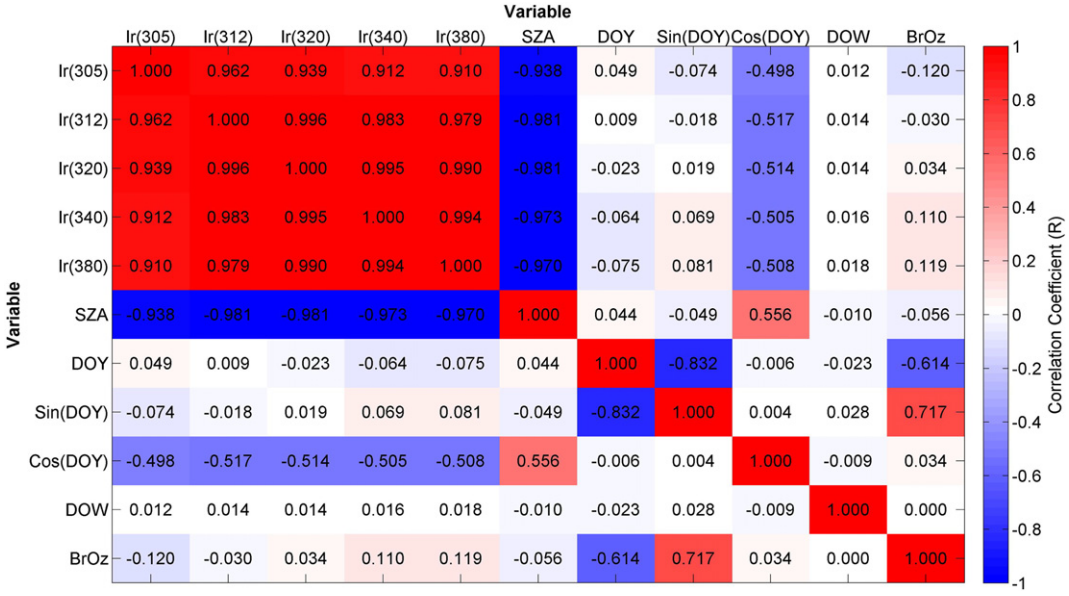


Fig. 1. The pairwise linear Pearson correlation coefficient of each input variable on Brewer#005 TOC ('BrOz'). Correlations are based on coincident NILU-UV and Brewer data which favours clear sky conditions due to the filtering of obscured sky cases based on the pointing error of the Brewer instrument during heavy cloud conditions.

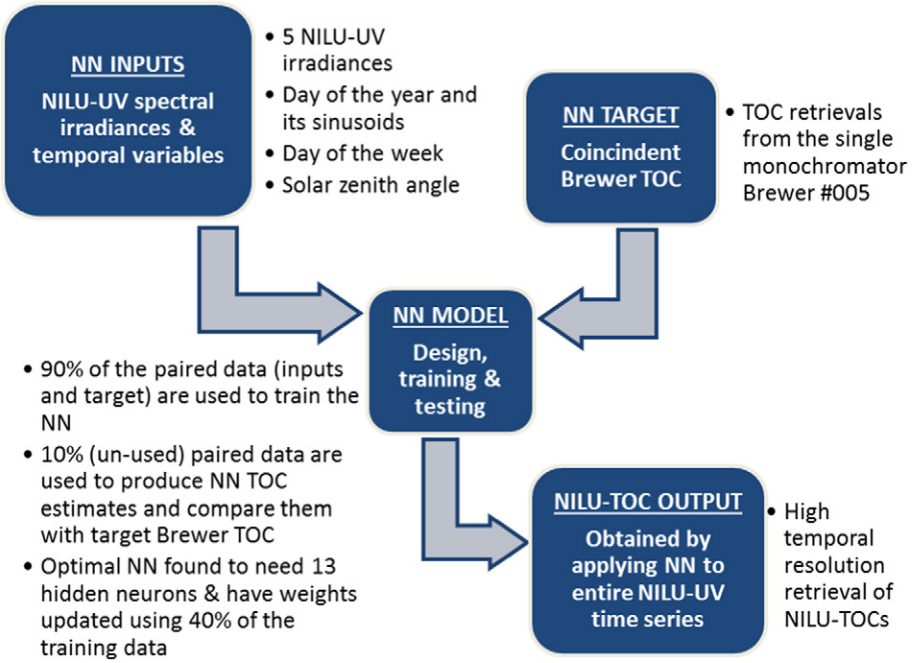
achieved at component s = 7 for Ir(302), Ir(312), Ir(320) and Ir(338), at component s = 6 for Ir(376) and at component s = 5 for TOC. The total noise obtained from summing components [s to 50] was then subtracted from each time series so that the NN model is developed using denoised signals in an attempt to capture as best as possible the underlying physics. We assess the sensitivity of this approach to noise in the data by feeding the trained network with measured (noisy) irradiances as inputs and comparing network NILU outputs against measured (noisy) Brewer#005 TOCs.

In the NN model, the input and output vectors were connected via 2 feed-forward network layers – the first containing hidden neurons with hyperbolic tangent (tanh) activation functions and the second containing a single neuron having a linear activation function connected to the output as depicted in Fig. 2.

The exact mathematical equation relating the outputs to the inputs for this type of NN model has been shown to have the matrix equation (Taylor et al., 2014):

$$\mathbf{Y} = f^2\left(\mathbf{LW}^{2,1} f^1\left(\mathbf{IW}^{1,1}\boldsymbol{X} + \boldsymbol{b}^1\right) + \boldsymbol{b}^2\right) \tag{4}$$

where in this study $\mathbf{Y} = [\mathrm{TOC}]^T$ and $\mathbf{X} = [\mathrm{Ir}(305), \mathrm{Ir}(312), \mathrm{Ir}(320), \mathrm{Ir}(340), \mathrm{Ir}(380), \mathrm{SZA}, \mathrm{DOY}, \mathrm{Sin}(\mathrm{DOY}), \mathrm{Cos}(\mathrm{DOY}), \mathrm{DOW}]^T$. The multiplication of the matrix $\mathbf{IW}^{1,1}$ and the vector $\mathbf{X}$ is a dot product equivalent to the summation of all input connections to each neuron in the hidden layer. Eq. (4) is the nonlinear relation between the output vector and the matrix of input vectors. The NN model was coded using MATLAB's object-oriented scripting language in conjunction with its Neural



Flow Chart 1. Schematic illustrating the construction of the NN model.

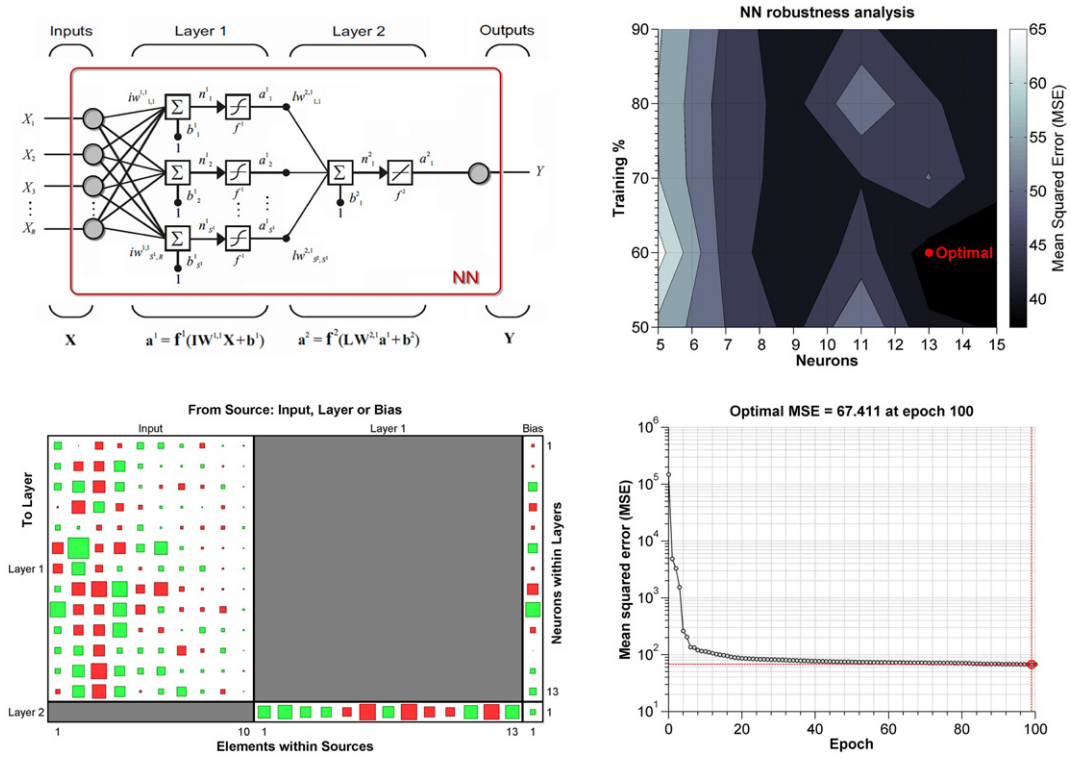


Fig. 2. Model selection. (Upper left) Schematic showing the neural connectivity between input and output parameters in the NILU-UV NN model. (Upper right) The robustness analysis on the grid of NN models using the minimum validation MSE as the criterion for selection of the optimal NN architecture which was found to have 13 hidden neurons and a training: validation data split of 60%:40%. (Bottom Left) The Hinton matrix of NN weights and biases with values represented by squares scaled to [−1,1]. Positive values are green and negative values are red. (Bottom Right) The progress of the NN training of the optimal architecture with backpropagation iteration.

Network Toolbox (Beale et al., 2015). Given enough hidden neurons and training data, such networks are capable of learning the exact mathematical relation between inputs and outputs.

Considering the lack of linear correlation between the Brewer#005 TOC and the input variables, this type of NN is ideally suited to nonlinear regression problems like the task at hand. The Hinton matrix of neuron weights and biases (see bottom left panel of Fig. 2) show that sizeable weights are apparent in the hidden neurons nonlinearly connected to the irradiances and the SZA, and the temporal variables have in general small non-zero neural weights. DOW, as expected, has the lowest weighting across all hidden neurons. We see then, that the NN implicitly takes into account dimensionality reduction in the data (in the classical sense of their impact) by assigning low/near-zero weights to neurons which are connected to "redundant" input variables.

The performance of NN models depends on their architecture (Bishop, 1995) and it is recommended that a sensitivity analysis is performed on the network parameters (Taylor et al., 2014). The optimal NN architecture was then found by minimizing the mean squared error (MSE) between the NN estimates and Brewer#005 TOC values used as reference output data for each NN in a grid of 100 NN architectures. From the training matrix of $n = 29.105$ co-located input-output vectors, a subset of t-vectors was chosen randomly as the training set. The number of hidden neurons was varied from 5 to 15 and the proportion of training data (t/n) was varied from 50% to 95% in steps of 5%. The number of learning iterations was set to 100 and the progression of training of the optimal NN architecture towards convergence at the horizontal asymptote is shown in the bottom-right panel of Fig. 2.

The training algorithm we used in this study is Bayesian regularization using a Laplace prior (Williams, 1995) which is based on the Bayesian framework described by Mackay (1992). Bayesian regularization converts a nonlinear regression into a "well-posed" statistical problem by minimizing a linear combination of squared errors and weights and modifies the linear combination so that at the end of training the resulting network has good generalization qualities (Beale et al., 2015). Bayesian regularization in MATLAB automates the determination of the optimal regularization parameters in its NN training function "trainbr" and takes place within the Levenberg-Marquardt algorithm where back-propagation is used to calculate the Jacobian of performance with respect to the weight and bias variables (Beale et al., 2015). During each of 100 iterations of the learning process, the weights and biases of each NN are tuned with the Bayesian regularization and the back-propagation optimization algorithm (Rumelhart et al., 1986) to minimize the mean squared error (MSE) between NN NILU outputs and Brewer measured TOC. As a result of this robustness analysis, the optimal NN was obtained with 13 hidden neurons and 60% of the training data as seen in the top-right panel of Fig. 2.

The optimal NN is valid for the range of parameters determined by the training data shown in Table 3 (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]).

The optimal trained NN was then fed with the remaining ("unseen") input vectors from the 10% of the training data (2988 records) and its estimates (NILU TOC) are then compared against the target measurements (Brewer#005 TOC) of the output vector to validate its performance.

Fig. 3 shows that a high level of correlation has been achieved between target Brewer#005 TOC values and the NILU TOC estimates with a high coefficient of determination for the training [left column] ($R^2 = 0.939$) and validation [right column] ($R^2 = 0.911$) data sets. Both present values of RMSE in the range 8.21–9.78 Dobson units (DU) and there is a close-to-zero bias in the range − 0.15–0.54 DU Robust regression was performed using the method of Theil (1950) and Sen (1968) and the peak of modelled TOC is, as expected at the mid-latitude station of Thessaloniki, in the range ≈280 to ≈350 DU in both cases. 95% of the differences between the NN estimates and the coincident target Brewer data are within ≈13 DU of the mean (which agree

Table 3
Range of validity of the trained optimal NN as determined by its input parameters and the Brewer TOC output parameter.

Parameter	Min	Max	Mean	St. dev.
Ir(302) (W/m^2)	0.001	0.019	0.004	0.004
Ir(312) (W/m^2)	0.006	0.228	0.090	0.052
Ir(320) (W/m^2)	0.030	0.334	0.150	0.071
Ir(338) (W/m^2)	0.102	0.735	0.342	0.135
Ir(376) (W/m^2)	0.134	0.929	0.447	0.171
SZA (°)	12.495	75.577	48.996	14.013
TOC (DU)	240.968	492.313	328.117	33.132

well between the training and validation samples: 328.18 DU and 326.67 DU respectively), suggesting that the NN is capturing the variability in the data satisfactorily. This is confirmed heuristically also by the identification of a clear annual cycle in both datasets as shown by a locally-weighted scatterplot smoothing (LOWESS) fit (Cleveland and Devlin, 1988) to the data using 10% of the time series as the span, in the bottom row of Fig. 3.

In order to further assure the validity of the NILU TOCs, an additional comparison was performed before they were used for alignment and comparison with satellite observations. Brewer#005 TOC data for the period 2012–2014 (i.e. a later period of Brewer data than that used to train and test the NN model) was compared against to generated NILU TOCs and showed an agreement of 0.41% ± 5.24% for this second test record of 7207 coincidences (Fig. 4). Since the training of the NN was performed for Brewer#005 TOC data up to the first semester of 2012, these may be considered as entirely independent measurements, also in the temporal sense. This self-standing evaluation of the validity of the NILU TOC is therefore ensured and is well within the uncertainty of the NILU irradiance measurements themselves (see Section 2.1).

It is important to note that the TOC measurements taken by the Brewer#005 are of much lower temporal resolution (optimally every 3–4 min) than the irradiance measurements taken by the NILU-UV (every minute). Furthermore, the uncertainty of the irradiances is assessed based on the calibration of the NILU-UV channels with the use of measured spectra by the second Brewer#086 spectrophotometer

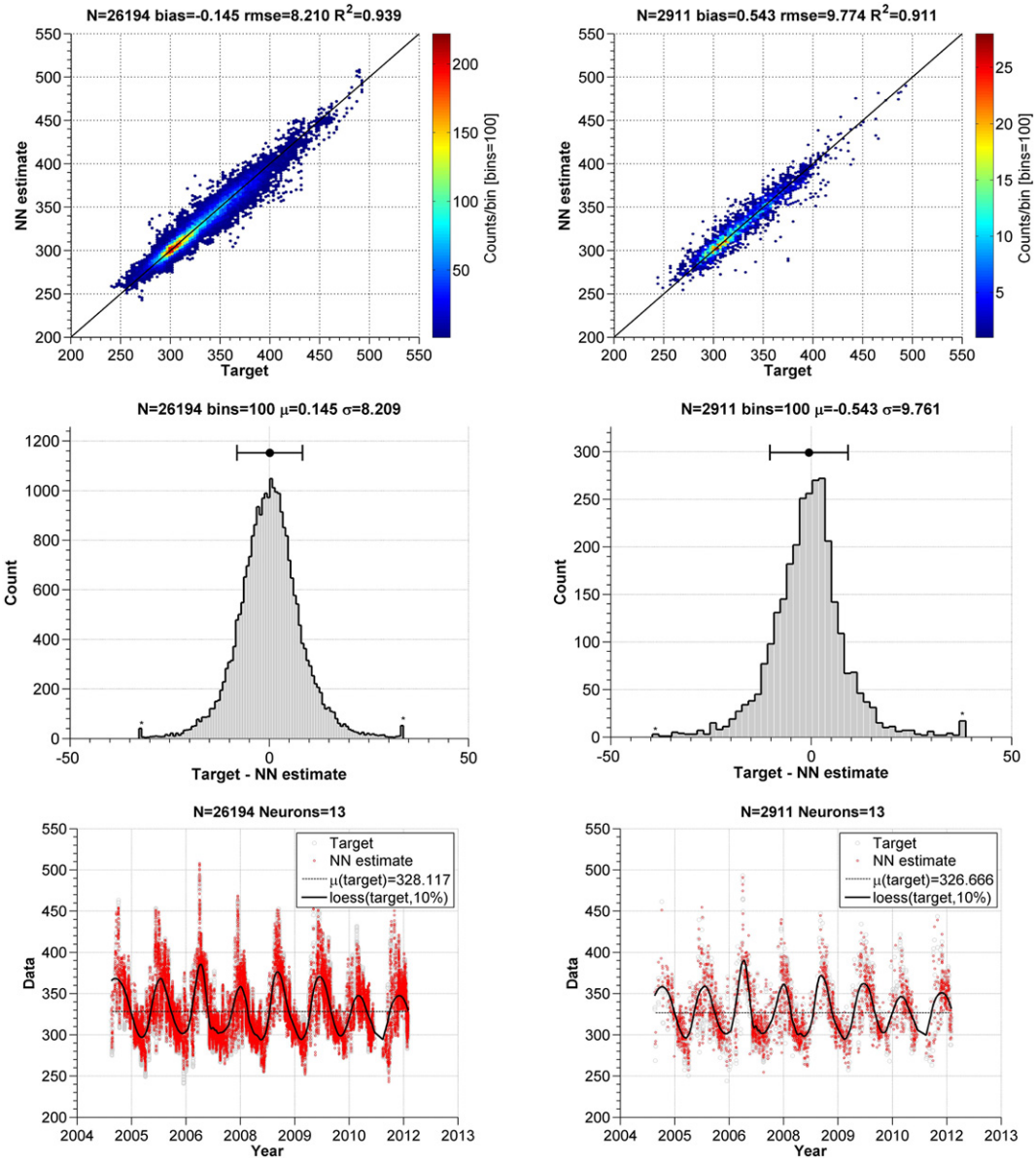


Fig. 3. NN training results (Left Panels) and NN validation results (Right Panels). (Top Row) Density plot of the regression of the NN estimates on the coincident Brewer derived TOC. (Middle Row) Histograms of the difference between NN estimates and coincident Brewer derived TOC. The mean (μ) and standard deviation (σ) are indicated. (Bottom Row) Time series of the NN estimates and coincident Brewer derived TOC. Also shown is the mean measured value as well as a LOWESS fit to the Brewer derived TOC data to indicate the annual variation. All quantities (with the exception of R^2 which is dimensionless) are measured in units of DU.

(detailed information in Table 1). The NN model developed here therefore, in effect, provides an approximation to Brewer TOC values but at the high resolution provided by NILU-UV.

The advantage of the proposed methodology in TOC retrievals from NILU-UV multi-filter radiometers is that the only requirements are the Brewer TOC values along with the NILU-UV irradiances (spectral or response weighted) without any other a-priori information like aerosol load and albedo as proposed by several other studies (Stamnes et al., 1991; Dahlback, 1996; Kazantzidis et al., 2009). Furthermore, in the case where spectral irradiances are used, the knowledge of the spectral response of each channel is not a prerequisite; while the 5 UV wavelengths used instead of only a pair can implicitly depict the aerosol and albedo wavelength dependence. Additionally, this method can provide ozone data under cloudy conditions if the sun disk is unobscured (with some limitations as discussed previously). In a similar basis methodology proposed by Zempila et al. (2016a) for retrieving PAR from Global Horizontal Irradiances (GHI), it was proven that the NNs trained with Thessaloniki's data had the same level of accuracy when applied to other NILU-UV and GHI stations of the Greek UV Network and Hellenic Network of Solar Energy (www.uvnet.gr and www.helionet.gr respectively). Thus, the NN originally trained in Thessaloniki's NILU-UV data is expected to perform likewise when is fed with any NILU-UV data within its training range limitations without any further adjustment.

4. Results and discussion

Having trained and validated the NN model, we then proceeded to remove the constraint of having to co-locate input and output vectors and ran it in unsupervised mode (i.e. simulation mode) hence calculating the NILU TOC dataset. Two cases are considered: one where all the NILU irradiances are used, irrespective of the cloudiness detected and one where only the clear-sky irradiances are considered. For the determination of the cloud presence measurements of the PAR from the sixth channel of the NILU were used. In particular, every 1 min measurement is characterized as cloud free or not by combining and comparing PAR measurements with estimations for clear skies resulting from the UVSPEC model of the library for radiative transfer (libRadtran) (Emde et al., 2016). A low cut-off filter is applied to PAR measurements based on model simulations where we assume a high aerosol optical depth of 0.4 at 500 nm, and an Angström exponent of 1.3. This corresponds to the smallest possible PAR under clear skies. Additionally, 1 min measurements with standard deviation larger than 10% of the mean were also considered to be cases influenced by clouds and were removed from the clear-sky dataset. In the final step, the relative percentage differences (RD) of the normalized (to the data point of interest) model estimations on five adjacent observations, were computed. For SZAs smaller than 75° a data point is characterized as cloud-free when the differences agree to within 2%. For larger SZAs, agreement within 5% is required. The evaluation of this cloud detection algorithm in Thessaloniki has been found to lead to an agreement of better than 80% compared with direct observations of cloudiness (Zempila et al., 2016b).

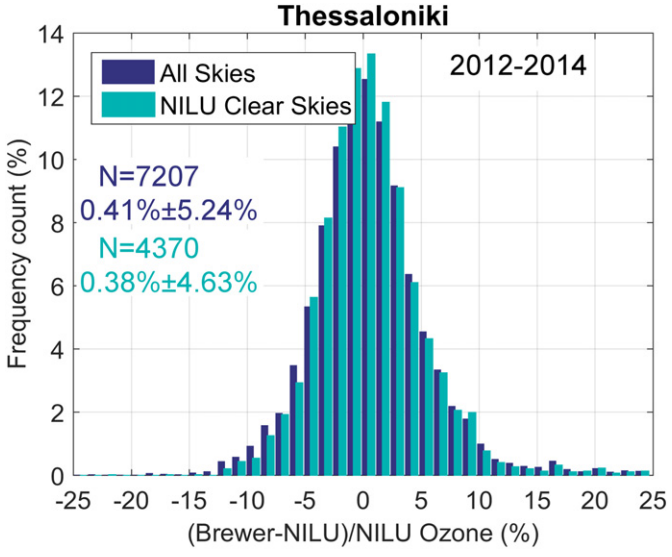


Fig. 4. Histogram of the percentage relative differences (RD) of the Brewer TOC and NILU TOC for the period 2012–2014 in Thessaloniki as an additional independent validation of the NN performance.

For cloud free cases in a 1 h interval (i.e. ±30 min), a filter was imposed on the ground-based data so that 1 h averages were characterized as clear if 70% of the available NILU TOC data were flagged as cloud-free by the cloud detection algorithm. The all skies and clear skies discrimination is made possible by the fact that Brewer#005 TOC retrievals are themselves based on direct solar radiation measurements. Thus the NILU derived TOC values are more reliable for cases where the sun disc is unobscured. So the NILU TOC data are expected to compare better under cloud free cases, while under cloudy conditions the NN will be struggling to perform if the sun disc is not clear (although the tests that were performed and described in Section 3.2 show that NN performs remarkably well for most cloud conditions).

In the case of all sky conditions 1,798,070 coincident NILU irradiance input records and associated temporal variables are available of which 475,467 are out of the range of validity of the NN training bounds given in Table 3. Furthermore, in order to take into account the sighting problem of the Brewer instrument during direct sun observations when clouds obscure the solar disc, we produced a daily climatology based on our long time series of Brewer TOC measurements from 04/01/2005–21/06/2012. For each DOY, Brewer TOC records were extracted and the mean and standard deviation was calculated to construct the daily climatology. NN estimates beyond 1 standard deviation of the daily climatology were excluded from the simulation (filtering out an additional 325,647 records). As a result of this filtering procedure ≈1 million NN TOC estimates (996,956 records) were produced for the entire decade of NILU measurements from 2005 to 2014 (inclusive). The same procedure was applied to clear sky cases flagged up by the NILU-UV algorithm. A total of 629,420 coincident input records and associated temporal variables are available of which 79,034 are out of the range of validity of the NN training bounds. The Brewer TOC daily climatology in this case leads to filtering of an additional 91,565 records. In total, 458,821 clear sky NN TOCs records were therefore also produced.

In order to compare NILU TOC with satellite TOC the same procedure for all four instruments was performed in the same manner. The percentage difference (relative to NILU TOC) between satellite and ground-based NILU TOC was calculated for each individual coincidence that was found to be co-located within a spatial scale of 150 km. Both the 'closest minute' coincidences and the 1 h mean of NILU TOC coincidences around satellite overpass co-location were examined. Since TOC values are not expected to have large variations at the 1 h timescale, no filter was imposed for the number of NILU TOC data within ±30 min of the satellite overpass when calculating the 1 h mean NILU TOC values. Temporal averaging is a common method when comparing ground-based measurements with satellite observations, and is well justified, as it can substantially increase the size of the coincident dataset (for e.g. Chubarova et al., 2002; Zempila et al., 2016c).

In studies that perform global validation of satellite TOC observations, daily mean ground-based measurements are used as ground truth (e.g. Fan et al., 2014; Zempila et al., 2016c). For validation of the GODFIT_v3 algorithm (Koukouli et al., 2015) daily mean Brewer and Dobson TOCs were used and resulted in global comparisons of the order of 0.75 ± 0.75% and 0.20 ± 0.50% respectively. This further points to the stability of the TOC within coincident timeframes containing satellite overpass retrievals and the ground-based observations. In this work, statistical analysis of the two different time-integration approaches (closest minute and ±30 min mean) showed that they result

in similar findings, almost identical coefficients of determination (R^2) and percentage fractional errors (both mean values and standard deviation). Furthermore, the use of 1 h mean NILU TOC values resulted in a larger number of coincidences (e.g. almost 4 times greater for GOME2 comparisons – see below). Therefore, in the following we will only focus on the comparisons based on the 1 h average ground-based NILU TOC values against the satellite TOCs.

For each of the satellite sensors, we present RDs between ground and satellite calculated from 100 ∗ (satellite − NILU) / NILU % using the ±30 min mean coincidences. In each case, a statistical summary is presented in a single composite figure. In the upper left panel of each figure a two panel composite graph is shown containing the time series of differences together with its seasonal time series. In the upper right panels, a SZA dependency plot is presented together with the mean of each 5° bin. In the lower left panel, a scatter plot shows the level of correlation between satellite TOC and NILU TOC values, and the lower right panel presents a histogram of the differences to help understand their distribution. Importantly, on all plots all sky cases are depicted by blue circles and clear sky cases by red dots.

4.1. GOME/ERS-2 comparisons

Comparisons between NILU TOC and with TOC retrievals from the GOME/ERS-2 satellite instrument were performed for 660 coincidences, while the cloud free days identified by the NILU PAR algorithm were found to be 227. The mean overpass time of ERS-2 over Thessaloniki for the entire time series that begins in 2006 and ends June 2011 was ~09:30 UT. The time series of RDs together with the seasonal mean values is shown in Fig. 5 (upper left panel). The agreement between the two datasets is remarkable with a mean value of −0.62 ± 3.68% for all skies and −0.67 ± 2.15% for clear sky cases. The fact that the standard deviation of the all sky cases is close to that of clear sky cases suggests that the comparison of NILU TOC and GOME TOC is quite robust also to cloudiness that might affect the ground-based observations. In Fig. 5, (upper right panel), there is no evidence of a SZA dependency for the cloud free cases. The coefficient of determination (R^2) as the measure of the goodness of fit of the GOME TOC on the NILU TOC is reasonably high (0.89) and the linear regression of GOME TOC on the NILU TOC in the scatter plot shown in Fig. 5 (lower left panel) present a strong correlation with the best fit line having a slope of 1.06 and a y-intersect of −23 DU. When limiting to cloud free cases, all statistical data were improved; the slope is almost unity (1.03), the coefficient of determination is 0.96 and the y-intersect is almost half (−12.7 DU). The histogram of the RDs in Fig. 5 (lower right) shows that the distribution of errors is fairly symmetrical and centered near zero (−0.62 and −0.67 for all sky and clear sky cases respectively). A small tail in the all sky differences suggests that NILU is slightly overestimating the TOC relative to GOME during partially cloudy conditions, most likely as a result of Brewer pointing error.

4.2. SCIAMACHY/Envisat comparisons

NILU TOC coincidences with SCIAMACHY on the Envisat satellite platform exist from January 2005 to April of 2012. Within this period, 821 cases of common data were found between the two TOC sets at overpass time ~09:00 UT. The time series of the RDs (Fig. 6, upper left panel) revealed a slight overestimation of the satellite data under clear skies during the month of June. However, there is no seasonal behavior detected for all sky cases; a fact that is reiterated to some extent also in the SZA dependency plot at the upper right panel of the same figure where the near-zero slopes (equal to 0.04) both for all-skies and clear-skies suggested small but statistically significant trends (*t*-test *p*-values <0.025 for both cases). In general, SCIAMACHY TOC was found to underestimate NILU TOC by 1.22% to 1.44% depending on the cloudiness. Standard deviations are very similar to those obtained for GOME (3.81% and 2.25% for all and clear skies respectively). In the lower left panel of Fig. 6, high coefficients of determination (0.88 and 0.95 respectively for all skies and clear skies) are shown and the least square linear fits have slopes of unity (1.01 and 1.00 respectively for all skies and clear skies) with low y-intersects (−7.37 and −3.81 D.U for all skies and clear skies respectively). The histogram of RDs (lower right panel of Fig. 6) presents a small positive skew towards positive differences (i.e. satellite TOC is higher than NILU TOC) and the distribution is centered

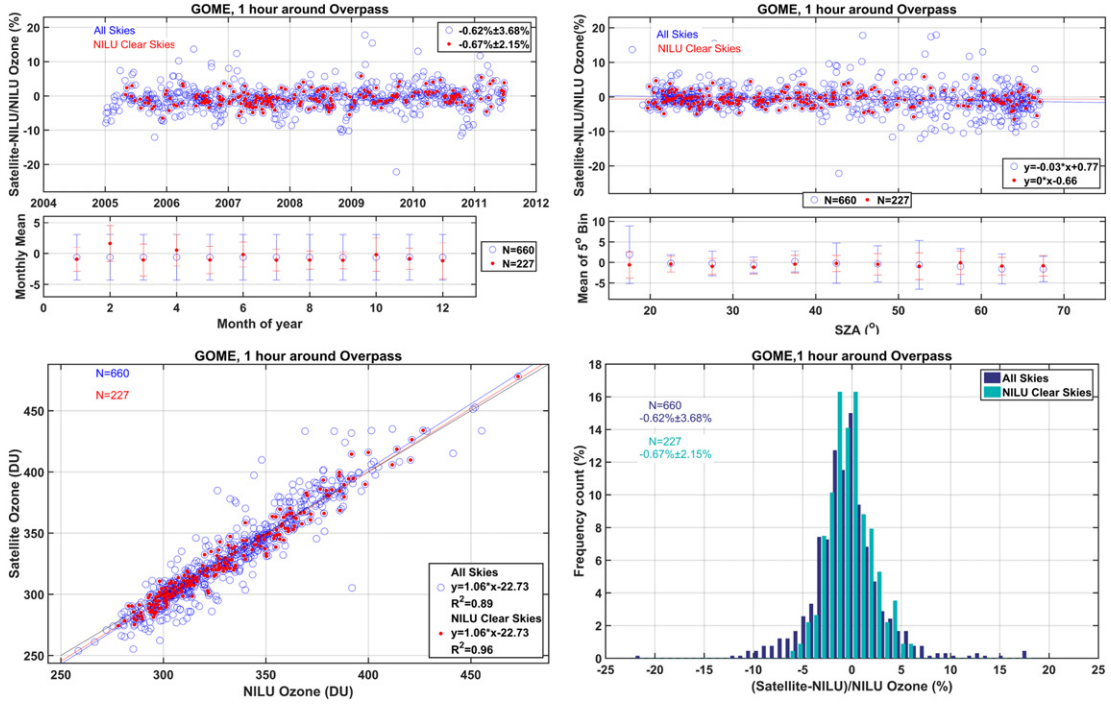


Fig. 5. Relative percentage difference (RD) statistics of the 1 h mean coincidences between GOME and NILU TOC instruments; upper left: time series and seasonal means; upper right: SZA dependency; lower left: scatter plot and regression statistics; lower right: histogram of the distribution of RDs.

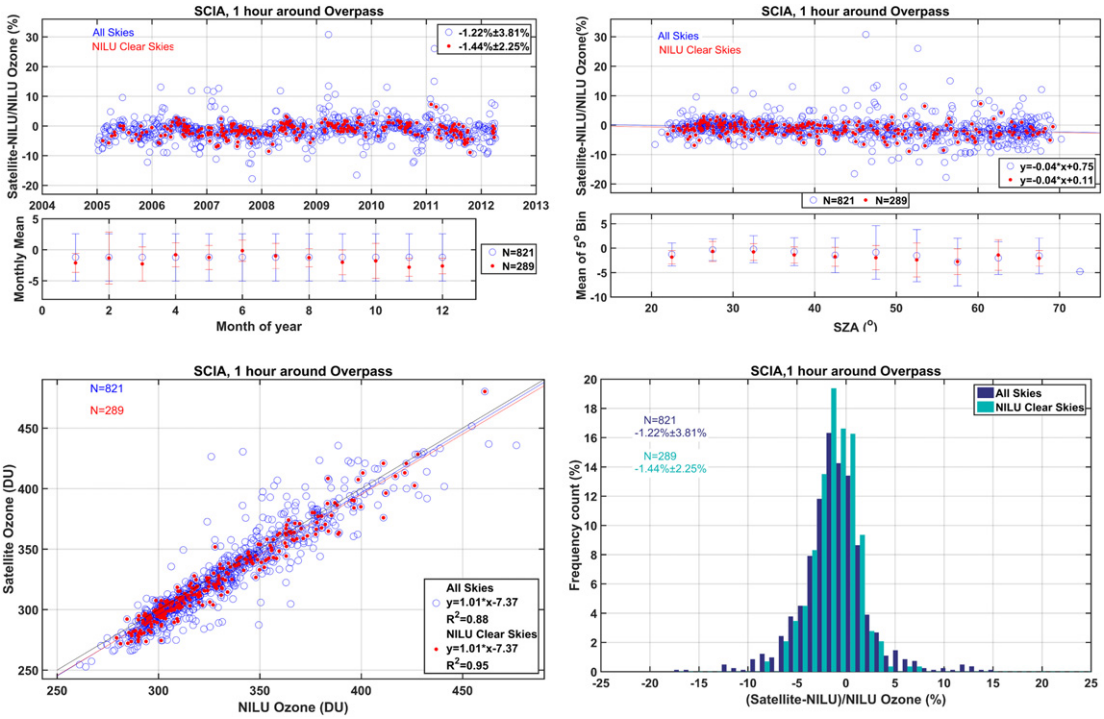


Fig. 6. Relative percentage difference (RD) statistics of the 1 h mean coincidences between SCIAMACHY and NILU TOC instruments; upper left: time series and seasonal means; upper right: SZA dependency; lower left: scatter plot and regression statistics; lower right: histogram of the distribution of RDs.

at around −1.22% to −1.44% in accordance with the slight underestimation mentioned above.

4.3. OMI/AURA comparisons

OMI provides the longest time series of NILU TOCs (January 2005–December 2014). Due to server failures in October 2013 and 2014, a total time gap of almost 6 months (4 months in 2013 and 2 months in 2014) is present in the time series of NILU. A set of 2869 common measurements around the overpass time (~11:45 UT) resulted in a mean difference value of −2.05% with a correspondent standard deviation of 3.34% for all sky cases (Fig. 7, upper left panel). Based on the monthly mean values, no seasonal trend was detected. For clear sky conditions, OMI underestimates the TOC by −2.09% ± 2.06% for this dataset of 1138 coincidences. Once again the month of June is the one containing the best agreement with the ground-based data (i.e. the closest to zero value) while November and December are the months with the highest underestimation of TOC as retrieved by the OMI instrument. The examination of the SZA dependence of the TOC retrievals presented in the upper right panel in Fig. 7, reveals a very small negative trend (−0.05 and −0.03% for cloudy and cloudless conditions) but these are statistically significant (*t*-test *p*-value < 0.025). The scatter plots in the lower left panel in Fig. 7, shows that the R^2 values are very high both for all skies and clear skies (0.9 and 0.95 respectively), the slopes are close to unity (0.98 and 0.97 respectively) and that the y-intersects are low (0.29 and 3.57 DU respectively for all skies and clear sky conditions). The histogram of RDs is symmetrically distributed around the mean value (−2.05% and −2.09% for all skies and clear skies) but is slightly positively skewed for clear skies (suggesting that the satellite underestimation is less) as presented in the lower right panel of Fig. 7.

4.4. GOME2/MetOp-A comparisons

GOME2 onboard Metop-A was found to have 1680 days where NILU TOCs were coincident for the common period January 2007 to December 2014. For clear skies, this set was reduced by 60%; resulting in 658 coincidences around the overpass time ~08:30 UT. As seen in the upper left panel of Fig. 8 the TOC estimates are in good agreement. Although the statistics of the comparison are extremely good (mean values of −0.55% and −0.85% and standard deviations of 3.56% and 2.19% for all skies and cloudless conditions respectively), a seasonal pattern is noticeable in the monthly means of the clear sky differences as shown in the upper left panel of Fig. 8. June and December are the months when under cloudless conditions, GOME2 provides the smallest underestimation while February, March and November are months when the highest underestimation of the satellite sensor is taking place. On the contrary, no monthly dependence can be detected for the monthly means of the all skies differences. A SZA pattern is also observed at the upper right panel of the figure for both all skies and clear skies. Although this pattern is well identified, the min-max range never exceeds 5% with slopes that are statistically insignificant (*p*-values < 0.025). In the lower left panel, the scatter plot of the two TOCs is provided. As for the other sensors, slopes are close to unity (1.03 and 0.99 for all skies and clear skies respectively), y-intersects are low to moderately low (−0.33 DU for cloud free cases and −12.74 DU for all skies) and the R^2 values are reasonably high (0.95 and 0.89 respectively). These findings suggest that the two TOC products are in good overall agreement. Despite the observation of some indication of seasonality in the mean monthly differences and SZA bins, the RDs appear to be symmetrically distributed around near-zero mean values as seen in the histogram in the lower right panel of Fig. 8. While the tails reach up to ±20%, correspond to isolated (outlier) cases possibly associated with problems in cloud screening where the direct beam might be blocked.

4.5. Summary statistics

In Table 4 we summarize the mean relative percentage difference (MRD) and corresponding standard deviation (STD), bias (BIAS), root-mean-squared-error (RMSE), Pearson coefficient (R) and the coefficient of determination (R^2) between each satellite TOC and NILU TOC for both

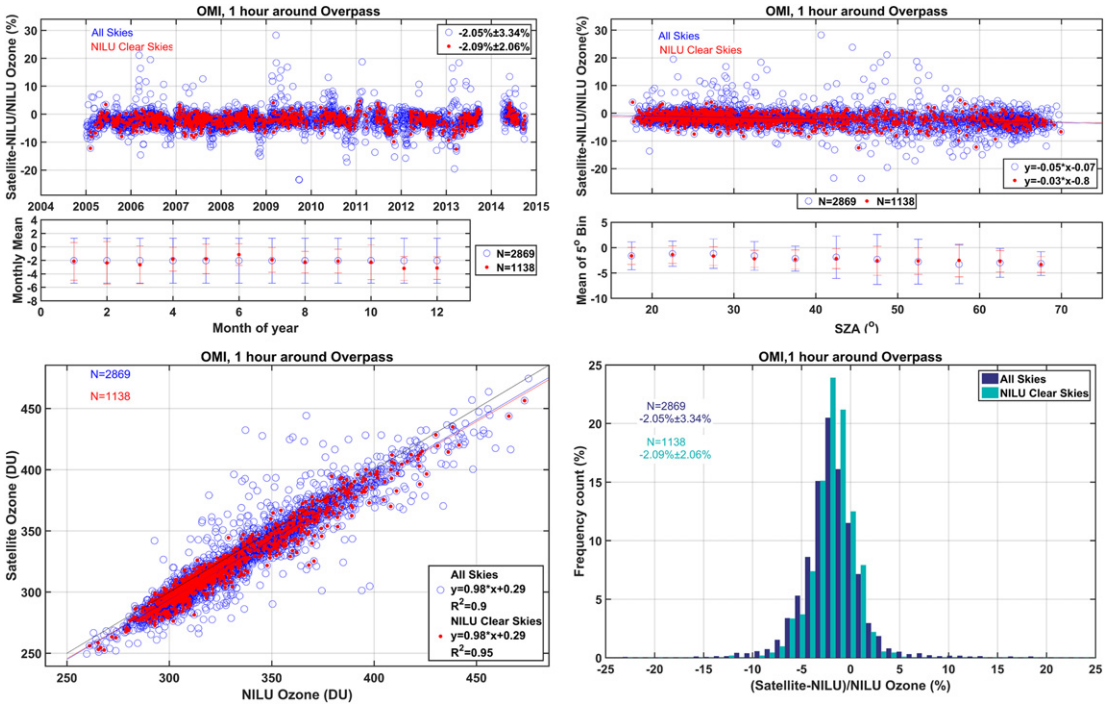


Fig. 7. Relative percentage difference (RD) statistics of the 1 h mean coincidences between OMI and NILU TOC instruments; upper left: time series and seasonal means; upper right: SZA dependency; lower left: scatter plot and regression statistics; lower right: histogram of the distribution of RDs.

all skies and clear skies cases. The bias is a metric that denotes the tendency of the satellite TOCs to overestimate (positive bias) or to underestimate (negative bias) the NILU TOC values. The RMSE was chosen as a statistical metric to measure the differences between the satellite and NILU TOC values expressed in DU. The R value measures the strength and the direction of a linear relationship between the two variables, while the R^2 value allows us to quantify the variance of satellite TOCs that is explained by the NILU TOC values.

As seen in Table 4, all satellite TOC retrievals present a high degree of homogeneity when compared with ground-based NILU TOCs. All time series are reduced by 60–65% when the comparisons are limited to clear skies. This limitation has only a small impact on the agreement between the satellite TOC and NILU TOC datasets for all satellite sensors. The mean relative percentage differences (MRD) do not alter significantly; GOME presents a mean difference from the NILU data of −0.62% that is decreased by 0.05% for the cloud free cases, while OMI

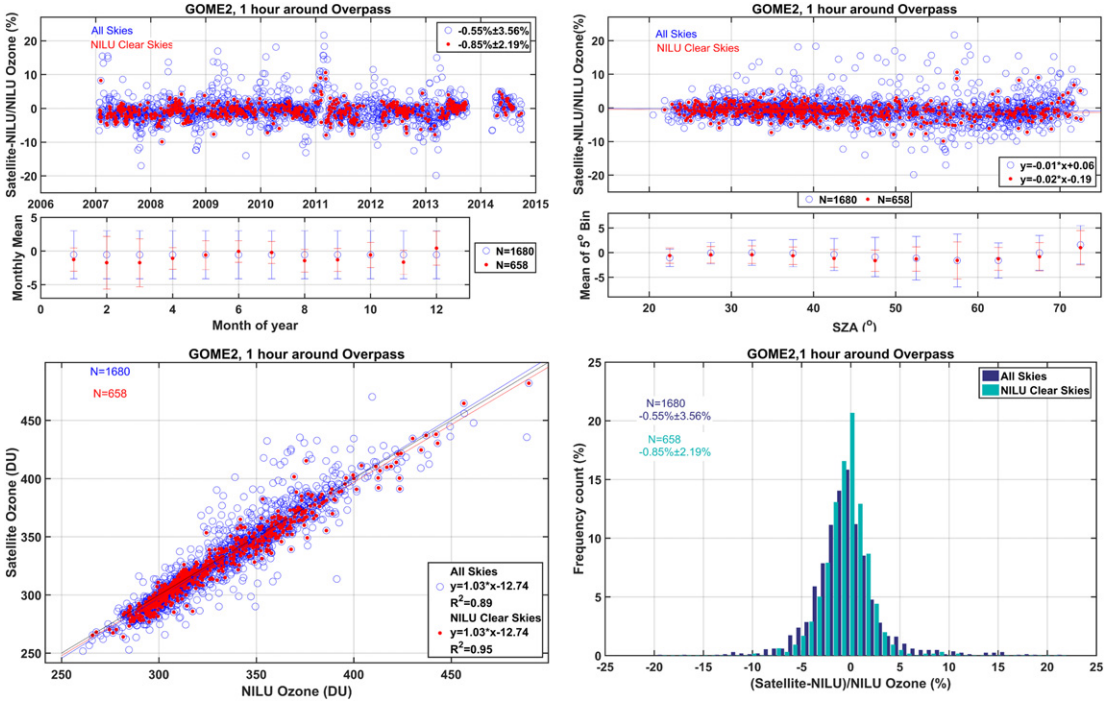


Fig. 8. Relative percentage difference (RD) statistics of the 1 h mean coincidences between GOME2 and NILU TOC instruments; upper left: time series and seasonal means; upper right: SZA dependency; lower left: scatter plot and regression statistics; lower right: histogram of the distribution of RDs.

Table 4
Statistical overview of the relative percentage differences (RD) between all 4 satellite TOCs from GOME, SCIAMACHY, OMI and GOME2 and coincident NILU TOCs (1 h means around overpass). The number of paired data (N) is provided together with the MRD, STD, BIAS, RMSE, R and R^2 values. The cloud free instances were identified based on the PAR cloud detection algorithm when 70% of the coincidences in the ±30 min period around overpass were characterized as cloudless.

	GOME ERS-2		SCIAMACHY Envisat		OMI AURA		GOME2 MetOp-A	
	All skies	Clear skies	All skies	Clear skies	All skies	Clear skies	All skies	Clear skies
N	660	227	821	289	2869	1138	1680	658
MRD (%)	−0.62	−0.67	−1.22	−1.44	−2.05	−2.09	−0.55	−0.85
STD (%)	3.68	2.15	3.81	2.25	3.34	2.06	3.56	2.19
BIAS (DU)	−1.83	−2.10	−3.96	−4.69	−6.85	−6.87	−1.68	−2.78
RMSE (DU)	12.64	7.55	13.66	8.97	13.40	9.81	12.30	7.92
R	0.94	0.98	0.94	0.98	0.95	0.97	0.94	0.97
R^2	0.89	0.96	0.88	0.95	0.9	0.95	0.89	0.95

agreement decreases from −2.05% to −2.09%. SCIAMACHY versus NILU comparisons revealed a MRD of −1.22% which is decreased to −1.44% for cloud free cases, and for GOME2 whose mean difference is −0.55%, had this reduced by 0.3% for cloudless skies. All satellite TOC retrievals present a low and well-constrained range of standard deviation of the differences both for all and for cloudless conditions. For all skies, the observed STDs range from 3.34% (OMI) up to 3.81% (SCIAMACHY). All STDs are reduced for the clear skies, ranging from 2.06% (OMI) up to 2.25% (SCIAMACHY).

What is clear though in the MRDs between satellite and ground-based TOC data, is that the satellite retrievals tend to all underestimate the ground-based NILU TOC values. This is also clear from the negative values in the biases obtained in Table 4. OMI represents the instrument that has the larger bias of −6.85 and −6.87 DU for all and clear skies respectively. The SCIAMACHY bias is −3.96 DU for all skies and −4.69 for cloudless conditions, while GOME underestimates less by only 1.83 and 2.1 DU for the cloudy and cloudless cases. GOME2 is found to be the instrument that presents the lowest difference with the NILU TOCs for all sky conditions with a bias of −1.68 DU, increased to −2.78 DU when the coincidence data is limited to clear sky cases. The systematic negative bias found in satellite retrievals can be partially attributed to the fact that tropospheric ozone columns in most cases are higher than those retrieved from the standard profiles (e.g. Kourtidis et al., 2002; Galani et al., 2003; Varotsos, 2005) thus the satellite retrievals may underestimate the TOC values.

In order to provide a more quantitate metric in DU for the general agreement of the sources of TOC retrieval used in this study, the RMSE was calculated. Cloud free cases have a RMSE ranging from 7.55 DU (GOME) up to 9.81 DU (OMI). To provide a context for the size of these biases, while the NN was trained based on Brewer#005 ozone data that are derived from measurements of the direct solar component (and thus mostly valid for clear skies), its performance was measured by a RMSE 8.21–9.78 DU. for the training and validation data respectively.

The excellent performance of the different sensor retrievals of TOC using the GODFIT_v3 algorithm with respect to ground-based NILU TOC is verified by high R and R^2 values in the RDs. The high R values of 0.94–0.98 assure that the compared data have strong and positive linear correlation, while R^2 values greater than 0.88 for all skies and 0.95 for clear skies suggest that there is a consistently high goodness of fit between the space-borne and ground-based TOC datasets.

4.6. Towards a unified total ozone records validation

The current trend in total ozone measurement studies, as extracted from satellite observations, is to provide unified, long time TOC records. The recently released Solar Backscattered Ultra Violet (SBUV) Level-3 Monthly Zonal Mean data products form part of the “Creating a Long Term Multi-Sensor Ozone Data Record” from NASA/GSFC (McPeters et al., 2013) providing a consistent forty-year long TOC dataset between 1970 and 2010 (Labow et al., 2013). Likewise, the ESA Ozone-CCI project aims to provide a similar dataset starting in 1996, with GOME/ERS2, and moving forward in time as long as the MetOp, and their successor, satellites fly (Lerot et al., 2014; Coldewey-Egbers et al., 2015). For this reason, in Fig. 9 we present the time series of RDs for all four satellite sensors together and for exact overpass time coincidence. The satellite instruments are listed according to the number of coincidences with ground-based NILU TOC retrievals and not by launch period. The time series reveal the spectacular homogeneity in the comparison of the satellite TOC retrievals as compared to the NILU TOC values. The scatter plot of the satellite TOC versus NILU TOC even more strongly emphasizes the uniformity of the four space-born sensors, with high R^2 values (0.95 to 0.96), low to moderately low y-intersects (−11.29 to +8.35 DU) and slopes close to unity (0.97 to 1.03). Even though restricting the temporal criterion for collocation along with the cloudiness limitation decreases severely the coincidences found, even down to ~25% for OMI, the statistics remain virtually un-altered and stable, yet again pointing to the high quality satellite and ground-based observations.

We can safely hence conclude that regarding the RDs and the corresponding standard deviations, the differences between the two temporal averaging approaches do not provide any noticeable change of the statistic results.

5. Conclusions

In this study, we have shown how high frequency surface spectral irradiance measurements from a NILU-UV multi-filter radiometer can be used, in conjunction with high accuracy TOC measurements from a Brewer spectrophotometer, to train a NN model capable of generating accurate and high temporal resolution estimates of the TOC at the same temporal resolution as the NILU-UV instrument (~1 min) at the mid-latitude UV and Ozone monitoring station in the Laboratory of Atmospheric Physics of the Aristotle University of Thessaloniki, Greece. The NILU TOC estimated by the NN model showed good agreement with Brewer TOC observations (R^2 = 0.94, RMSE = 8.21 DU and bias = −0.15 DU) for the training data with 95% of the coincident data differing by less than 13 DU. The methodology proposed demonstrates the ability of NNs to accurately describe the nonlinear relationship between the inputs (in our case spectral irradiances, SZA and temporal parameters) and output (TOC), independently of the station as long as the inputs are within its validity range.

In the context of space-based TOC retrieval with the GODFIT_v3 algorithm, four satellite instruments on board of different platforms were compared with the NILU TOC estimates both for the exact minute time match and for the ±30 min average around the satellite overpass within 150 km, a common technique applied in such studies. Although these time integration approaches did not result in significant different statistical analysis and patterns for the ground- and satellite-derived TOCs comparisons, the 1 h average coincidences provided a much higher number of resulting coincidences and was the focus of our analysis.

Based on our findings a series of valuable comments can be made and are summarized below.

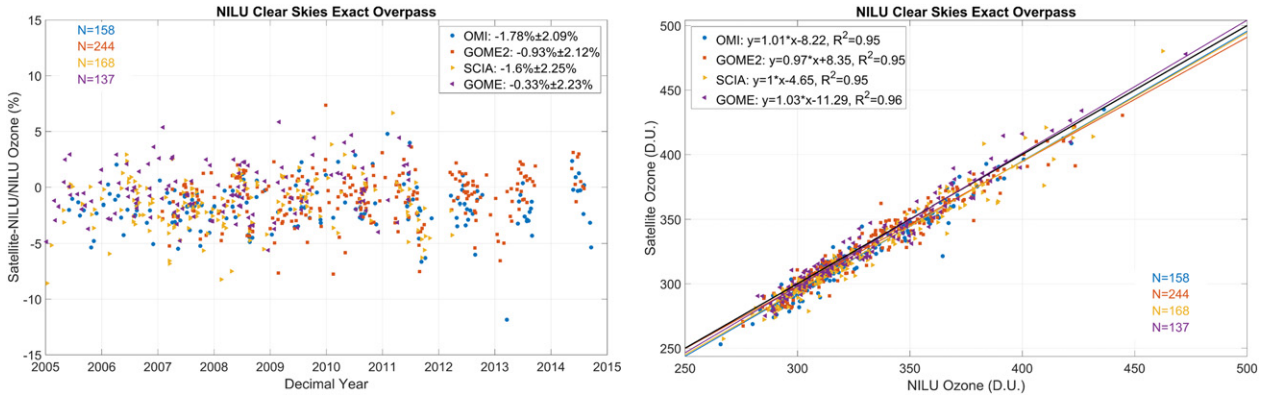


Fig. 9. Left panel: Relative percentage differences (RD) between satellite TOC and NILU TOC for all four sensors as a time series at the exact overpass minute. Right panel: scatter plot of the satellite TOC versus NILU TOCs for all four sensors around at the exact overpass minute. Statistics are presented in the figure.

- The time series reveal that there is homogeneity in the comparison of the satellite TOC retrievals as compared to the NILU TOC values. The scatter plot of the satellite TOC versus NILU TOC even more strongly emphasizes the uniformity of the four space-born sensors, with high R^2 values (0.95 to 0.96), low to moderately low y-intersects (−12.71 to +3.57 DU) and slopes close to unity (0.97 to 1.03).
- All time series are reduced by 60–65% when the comparisons are limited to clear skies. This limitation has only a small impact on the agreement between the satellite TOC and NILU TOC datasets for all satellite sensors. The MRDs do not alter significantly; GOME presents a mean difference from the NILU data of −0.62% that is decreased by 0.05% for the cloud free cases, while OMI agreement decreases from −2.05% to −2.09%. SCIAMACHY versus NILU comparisons revealed a MRDs of −1.22% which is decreased to −1.44% for cloud free cases, and for GOME2 whose mean difference is −0.55%, had this reduced by 0.3% for cloudless skies.
- All satellite TOC retrievals present a low and well-constrained range of standard deviation of the differences both for all and for cloudless conditions. For all skies, the observed STDs range from 3.34% (OMI) up to 3.81% (SCIAMACHY). All STDs are reduced for the clear skies, ranging from 2.06% (OMI) up to 2.25% (SCIAMACHY).
- What is clear though in the mean differences between satellite and ground-based TOC data, is that the satellite retrievals tend to all underestimate the ground-based NILU TOC values. This is also clear from the negative values in the obtained biases. OMI represents the instrument that has the larger bias of −6.85 and −6.87 DU for all and clear skies respectively. The SCIAMACHY bias is −3.96 DU for all skies and −4.69 for cloudless conditions, while GOME underestimates less by only 1.83 and 2.1 DU for the cloudy and cloudless cases. GOME2 is found to be the instrument that presents the lowest difference with the NILU TOCs for all sky conditions with a bias of −1.68 DU, increased to −2.78 DU when the coincidence data is limited to clear sky cases.
- In order to provide a quantitate metric in DU for the general agreement of the sources of TOC retrieval used in this study, the RMSE was calculated. Cloud free cases have a RMSE ranging from 7.55 DU (GOME) up to 9.81 DU (OMI). To provide a context for the size of these biases, while the NN was trained based on Brewer#005 ozone data that are derived from measurements of the direct solar component (and thus mostly valid for clear skies), 95% of the coincident data were found to agree by less than 13 DU.

In conclusion, this comprehensive study has demonstrated that ground-based time series of TOC can be expanded to higher temporal frequency and with high accuracy thanks to a NN model of NILU TOC based on NILU-UV irradiances and their temporal variables in conjunction with Brewer TOC observations used for training. Furthermore, the homogeneity of the four satellite TOC products and the excellent performance of the different sensor retrievals using the GODFIT_v3 algorithm, have been verified using NILU-UV ground-based TOC, with all satellite instruments providing cohesive results when compared with the NILU TOC retrievals.

It is hoped that this methodology can help local ozone monitoring stations where NILU-UV multi-filter radiometers are in operation to expand their records and to help support ongoing global validation of satellite TOC retrievals.

Acknowledgements

The authors would like to acknowledge the National Network for the Measurement of Ultraviolet Solar Radiation, uvnet.gr. MEK, DB, CL and MvanR would like to acknowledge the European Space Agency Ozone-CCI project, http://www.esa-ozone-cci.org/. MMZ acknowledges Drs. Natalia Kouremeti and Stelios Kazadzis for their valuable remarks during the calibration of NILU-UV 04103. MT would like to thank LAP/AUTh for their support, kind hospitality and training in the field.

References

Bais, A.F., Zerefos, C.S., Ziomas, I.C., Zoumakis, N., Mantis, H.T., Hofmann, D.J., Fiocco, G., 1985. Decreases in the ozone and the S02 columns following the appearance of the El Chichon aerosol cloud at midlatitude. In: Zerefos, C.S., Ghazi, A. (Eds.), Atmospheric Ozone. Springer Netherlands:pp. 353–356 http://dx.doi.org/10.1007/978-94-009-5313-0_72.

Bais, A.F., Zerefos, C.S., Meleti, C., Ziomas, I.C., Tourpali, K., 1993. Spectral measurements of solar UVB radiation and its relations to total ozone, SO_2, and clouds. J. Geophys. Res.-Atmos. 98, 5199–5204.

Bais, A.F., Zerefos, C.S., McElroy, C.T., 1996. Solar UVB measurements with the double- and single-monochromator Brewer ozone spectrophotometers. Geophys. Res. Lett. 23, 833–836.

Bais, A.F., Gardiner, B.G., Slaper, H., Blumthaler, M., Bernhard, G., McKenzie, R., Webb, A.R., Seckmeyer, G., Kjeldstad, B., Koskela, T., Kirsch, P.J., Gröbner, J., Kerr, J.B., Kazadzis, S., Leszczynski, K., Wardle, D., Josefsson, W., Brogniez, C., Gillotay, D., Reinen, H., Weihs, P., Svenoe, T., Eriksen, P., Kuik, F., Redondas, A., 2001. SUSPEN intercomparison of ultraviolet spectroradiometers. J. Geophys. Res.-Atmos. 106, 12509–12525.

Bais, A.F., Kazadzis, S., Garane, K., Kouremeti, N., Gröbner, J., Blumthaler, M., Seckmeyer, G., Webb, A.R., Koskela, T., Görts, P., Schreder, J., 2005. Portable device for characterizing the angular response of UV spectroradiometers. Appl. Opt. 44, 7136–7143.

Balis, D.S., Zerefos, C.S., Kourtidis, K., Bais, A.F., Hofzumahaus, A., Kraus, A., Schmitt, R., Blumthaler, M., Gobbi, G.P., 2002. Measurements and modeling of photolysis rates during the Photochemical Activity and Ultraviolet Radiation (PAUR) II campaign. J. Geophys. Res.-Atmos. 107 (PAU 5-1–PAU 5-12).

Bass, A.M., Paur, R.J., 1985. The ultraviolet cross-sections of ozone: I. The measurements. In: Zerefos, C.S., Ghazi, A. (Eds.), Atmospheric Ozone: Proceedings of the Quadrennial Ozone Symposium held in Halkidiki, Greece 3–7 September 1984. Springer Netherlands, Dordrecht.

Beale, M.H., Hagan, M.T., Demuth, H., 2015. Neural Network Toolbox: User's Guide, The MathWorks, 824 Inc. Natick, MA (USA).

Bishop, C.M., 1995. Neural Networks for Pattern Recognition. Oxford University Press, New York, USA.

Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V., Chance, K., Goede, A., 1999. SCIAMACHY: mission objectives and measurement modes. J. Atmos. Sci. 56, 127–150.

Burrows, J., Weber, M., Buchwitz, M., Rozanov, V., Ladstaetter-Weissenmeyer, A., Richter, A., de Beek, R., Hoogen, R., Bramstadt, K., Eichmann, K.-U., Eisinger, M., Perner, D., 1999. The Global Ozone Monitoring Experiment (GOME): mission concept and first scientific results. J. Atmos. Sci. 56, 151–175.

Chubarova, Ne, Yurova, Au, Krotkov, N., Herman, J., Bhartia, P.K., 2002. Comparisons between ground measurements of broadband ultraviolet irradiance (300 to 380 nm) and total ozone mapping spectrometer ultraviolet estimates at Moscow from 1979 to 2000. Opt. Eng. 0001 41 (12):3070–3081. http://dx.doi.org/10.1117/1.1516819.

Cleveland, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Assoc. 83 (403), 596–610.

Coldewey-Egbers, M., Loyola, D.G., Koukouli, M., et al., 2015. The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative. Atmos. Meas. Tech. Discuss. 8:4607–4652. http://dx.doi.org/10.5194/amtd-8-4607-2015.

Cybenko, G., 1989. Approximation by super-positions of a sigmoidal function. Math. Control Sig. 2, 303–314.

Dahlback, A., 1996. Measurements of biologically effective UV doses, total ozone abundances and cloud effects with multichannel, moderate bandwith filter instruments. Appl. Opt. 35 (33):6514–6521. http://dx.doi.org/10.1364/ao.35.006514.

de Forster, F., Piers, M., Shine, Keith P., 1997. Radiative forcing and temperature trends from stratospheric ozone changes. J. Geophys. Res. - Atmos. 102 (D9):2156–2202. http://dx.doi.org/10.1029/96JD03510.

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

Fan, L., Li, W., Dahlback, A., Stamnes, J.J., Stamnes, S., Stamnes, K., 2014. New neural-network-based method to infer total ozone column amounts and cloud effects from multi-channel, moderate bandwidth filter instruments. Opt. Express 22 (16), 19595–19609.

Farman, J.C., Gardiner, B.G., Shanklin, J.D., 1985. Large losses of total ozone in Antarctica reveal seasonal ClO X/NO X interaction. Nature 315 (6016), 207–210.

Fountoulakis, I., Bais, A.F., Fragkos, K., Meleti, C., Tourpali, K., Zempila, M.M., 2016. 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.

Fowler, D., et al., 2009. Atmospheric composition change: ecosystems-atmosphere interactions. Atmos. Environ. 43 (33), 5193–5267.

Fragkos, K., Bais, A.F., Meleti, C., Fountoulakis, I., Tourpali, K., Balis, D.S., Zerefos, C.S., 2014. Variability of a thirty-year record of total ozone derived from a Brewer spectrophotometer at Thessaloniki and the SBUV version 8.6. E-Proceedings of the XII EMTE National-International Conference of Meteorology-Climatology and Atmospheric Physics 1 (ISBN-978-960-524-430-9, May 28–31, Heraklion, Greece, http://comecap2014.chemistry.uoc.gr/COMECAP-ISBN-978-960-524-430-9-vol.%203.pdf).

Fragkos, K., Bais, A.F., Fountoulakis, I., Balis, D., Tourpali, K., Meleti, C., Zanis, P., 2015. Extreme total column ozone events and effects on UV solar radiation at Thessaloniki, Greece. Theor. Appl. Climatol. 2015:1–13. http://dx.doi.org/10.1007/s00704-015-1562-3.

Galani, E., Balis, D., Zanis, P., Zerefos, C., Papayannis, A., Wernli, H., Gerasopoulos, E., 2003. Observations of stratosphere-to-troposhere transport events over the eastern Mediterranean using a ground-based lidar system. J. Geophys. Res. 108, 8527.

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

Ghil, M., Allen, M.R., Dettinger, M.D., et al., 2002. Advanced spectral methods for climatic time series. Rev. Geophys. 40 (1:3), 1–41.

Glandorf, M., Arola, A., Bais, A., Seckmeyer, G., 2005. Possibilities to detect trends in spectral UV irradiance. Theor. Appl. Climatol. 81, 33–44.

Hassinen, S., Balis, D., Bauer, H., Begoin, M., Delcloo, A., Eleftheratos, K., Gimeno Garcia, S., Granville, J., Grossi, M., Hao, N., Hedelt, P., Hendrick, F., Hess, M., Heue, K.-P., Hovila, J., Jønch-Sørensen, H., Kalakoski, N., Kauppi, A., Kiemle, S., Kins, L., Koukouli, M.E., Kujanpää, J., Lambert, J.-C., Lang, R., Lerot, C., Loyola, D., Pedergnana, M., Pinardi, G., Romahn, F., van Roozendael, M., Lutz, R., De Smedt, I., Stammes, P., Steinbrecht, W., Tamminen, J., Theys, N., Tilstra, L.G., Tuinder, O.N.E., Valks, P., Zerefos, C., Zimmer, W., Zyrichidou, I., 2016. Overview of the O3M SAF GOME-2 operational atmospheric composition and UV radiation data products and data availability. Atmos. Meas. Tech. 9:383–407. http://dx.doi.org/10.5194/amt-9-383-2016.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2, 359–366.

Intergovernmental Panel on Climate Change (IPCC), 2013. Climate Change 2013 The Physical Science Basis: Contribution of the Working Group I to the Fifth Assessment Report of the IPCC. Cambridge University Press, New York.

Kazadzis, S., Bais, A., Arola, A., Krotkov, N., Kouremeti, N., Meleti, C., 2009. Ozone monitoring instrument spectral UV irradiance products: comparison with ground based measurements at an urban environment. Atmos. Chem. Phys. 9, 585–594.

Kazantzidis, A., Bais, A., Zempila, M.M., Meleti, C., Eleftheratos, C., Zerefos, C., 2009. Evaluation of ozone column measurements over Greece with NILU-UV multi-channel radiometers. Int. J. Remote Sens. 30 (15–16):4273–4281. http://dx.doi.org/10.1070/01431160902825073.

Kerr, J.B., McElroy, C.T., Olafson, R.A., 1981. Measurements of ozone with the Brewer spectrophotometer. Natl. Cent. for Atmos. Res., Boulder, Colorado 74–79.

Kolehmainen, M., Martikainen, H., Ruuskanen, J., 2001. Neural networks and periodic components used in air quality forecasting. Atmos. Environ. 35 (5), 815–825.

Koukouli, M.E., et al., 2015. Evaluating a new homogeneous total ozone climate data record from GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A. J. Geophys. Res. Atmos. 120:12,296–12,312. http://dx.doi.org/10.1002/2015JD023699.

Kourtidis, K., Zerefos, C., Balis, D., Kosmidis, E., Rapsomanikis, S., Perros, P.E., Simeonov, V., Melas, D., Thompson, A., Witte, J., Alpini, B., Rappenglueck, B., Isaksen, I., Papayannis, A., Hofzumahaus, A., Gimm, H., Drakou, R., 2002. Regional tropo-spheric ozone over eastern Mediterranean. Journal of Eophysical Research 107, 8140.

Labow, G.J., McPeters, R.D., Bhartia, P.K., Kramarova, N., 2013. A comparison of 40 years of SBUV measurements of column ozone with data from the Dobson/Brewer network. J. Geophys. Res. Atmos. 118:7370–7378. http://dx.doi.org/10.1002/jgrd.50503.

Lerot, C., Van Roozendael, M., Spurr, R., et al., 2014. Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A. J. Geophys. Res. Atmos. 119. http://dx.doi.org/10.1002/2013JD020831.

Levelt, P.F., van den Oord, G.H.J., Dobber, M.R., Mälkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J., Saari, H., 2006. The ozone monitoring instrument. IEEE Trans. Geo. Rem. Sens. 44 (5):1093–1101. http://dx.doi.org/10.1109/TGRS.2006.872333.

MacKay, D.J.C., 1992. Bayesian interpolation. Neural Comput. 4 (3), 415–447.

Mäder, J., Staehelin, J., Peter, T., Brunner, D., Rieder, H., Stahel, W., 2010. Evidence for the effectiveness of the Montreal Protocol to protect the ozone layer. Atmos. Chem. Phys. 10 (24), 12161–12171 1992.

Matsumoto, M., Nishimura, T., 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans. Modeling and Comp. Sim. 8 (1), 3–30.

McPeters, R.D., Bhartia, P.K., Haffner, D., et al., 2013. The version 8.6 SBUV ozone data record: An overview. J. Geophys. Res. Atmos. 118:8032–8039. http://dx.doi.org/10.1002/jgrd.50597.

Meleti, C., Bais, A.F., Kazadzis, S., Kouremeti, N., Garane, K., Zerefos, C., 2009. Factors affecting solar ultraviolet irradiance measured since 1990 at Thessaloniki, Greece. Int. J. Remote Sens. 30, 4167–4179.

Meleti, C., Fragkos, K., Bais, A.F., Tourpali, K., Balis, D., Zerefos, C.S., 2012. Thirty years of total ozone measurements at Thessaloniki with a MKII Brewer spectrophotometer. Quadrennial Ozone Symposium 2012, August 27–31, 2012, Toronto, Canada.

Müller, M.D., Kaifel, A., Weber, M., Burrows, J.P., 2002. Neural network scheme for the retrieval of total ozone from global ozone monitoring experiment data. Appl. Opt. 41 (24), 5051–5058.

Ramaswamy, V., Schwarzkopf, M.D., Randel, W.J., 1996. Fingerprint of ozone depletion in the spatial and temporal pattern of recent lower-stratospheric cooling. Nature 382 (6592), 616–618.

Redondas, A., Carreño, V., Hernández, B., Berjón, A., López-Solano, J., León-Luis, S.F., May 2016. Brewer Ozone Calibration at X RBCC-E/EUBREWNET Calibration Campaign. Brewer Ozone Spectrophotometer/Metrology Open Workshop, IPMA, Ponta Delgada, Sao Miguel, Azores, Portugal, pp. 17–20.

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

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. J. Am. Stat. Assoc. 63, 1379–1389.

Shepherd, T., Jonsson, A., 2008. On the attribution of stratospheric ozone and temperature changes to changes in ozone-depleting substances and well-mixed greenhouse gases. Atmos. Chem. Phys. 8 (5), 1435–1444.

Shindell, D., Kuylenstierna, J.C.I., Vignati, E., van Dingenen, R., Amann, M., Klimont, Z., Anenberg, S.C., Muller, N., Janssens-Maenhout, G., Raes, F., Schwartz, J., Faluvegi, G., Pozzoli, L., Kupiainen, K., Höglund-Isaksson, L., Emberson, L., Streets, D., Ramanathan, V., Hicks, K., Oanh, N.T.K., Milly, G., Williams, M., Demkine, V., Fowler, D., 2012. Simultaneously mitigating near-term climate change and improving human health and food security. Science 335 (6065), 183–189.

Shine, K.P., Briegleb, B.P., Grossman, A.S., Hauglustaine, D., Mao, Huiting, Ramaswamy, V., Schwarzkopf, M.D., Van Dorland, R., Wang, W.-C., 1995. Radiative forcing due to changes in ozone: a comparison of different codes, atmospheric ozone as a climate gas. NATO ASI Series I 32, 373–396.

Sitch, S., Cox, P.M., Collins, W.J., Huntingford, C., 2007. Indirect radiative forcing of climate change through ozone effects on the land-carbon sink. Nature 448 (7155), 791–794.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., Miller, H., 2007. IPCC, 2007: climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Stamnes, K., Slusser, J., Bowen, M., 1991. Derivation of total ozone abundance and cloud effects from spectral irradiance measurements. Appl. Opt. 30 (30):4418–4426. http://dx.doi.org/10.1364/ao.30.004418.

Taylor, M., Kazadzis, S., Tsekeri, A., Gkikas, A., Amiridis, V., 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. http://dx.doi.org/10.5194/amt-7-3151-2014.

Taylor, M., Retalis, A., Flocas, H.A., 2016. Particulate matter estimation from photochemistry: a modelling approach using neural networks and synoptic clustering. Aerosol Air Qual. Res. 16 (9):2067–2084. http://dx.doi.org/10.4209/aaqr.2015.07.0481.

Theil, H., 1950. A rank-invariant method of linear and polynomial regression analysis. Nederl. Akad. Wetensch. Proc. 53, 386–392.

Van Roozendael, M., et al., 2012. Sixteen years of GOME/ERS-2 total ozone data: the new direct-fitting GOME Data Processor (GDP) version 5—algorithm description. J. Geophys. Res. 117, D03305. http://dx.doi.org/10.1029/2011JD016471.

Varotsos, C., 2005. Airborne measurements of aerosol, ozone, and solar ultraviolet irradiance in the troposphere. J. Geophys. Res.-Atmos. 110, D09202.

Vyushin, D.I., Shepherd, T.G., Fioletov, V.E., 2010. On the statistical modeling of persistence in total ozone anomalies. J. Geophys. Res. - Atmos. 115 (D16), D16306.

WHO, 2003. Health Aspects of Air Pollution With Particulate Matter, Ozone and Nitrogen Dioxide. Bonne.

Williams, P.M., 1995. Bayesian regulation and pruning using a Laplace prior. Neural Comput. 7 (1), 117–143.

WMO, 2015. Scientific Assessment of Ozone Depletion: 2014, Global Ozone Research and Monitoring Project-Report No. 55, 416 pp., Geneva.

Zempila, M.M., Taylor, M., Bais, A., Kazadzis, S., 2016a. Modelling the relationship between photosynthetically active radiation and global horizontal irradiance using singular spectrum analysis. J. Quant. Spectrosc. Radiat. Transf. 182:240–263. http://dx.doi.org/10.1016/j.jqsrt.2016.06.003.

Zempila, M.M., Giannaros, T., Bais, A., Melas, D., Kazantzidis, A., 2016b. Evaluation of WRF shortwave radiation parameterizations in predicting Global Horizontal Irradiance in Greece. Renew. Energy 86:831–840. http://dx.doi.org/10.1016/j.renene.2015.08.057.

Zempila, M.M., Koukouli, M.E., Bais, A., Fountoulakis, I., Arola, A., Kouremeti, N., Balis, D., 2016c. OMI/Aura UV product validation using NILU-UV ground-based measurements in Thessaloniki, Greece. Atmos. Environ. http://dx.doi.org/10.1016/j.atmosenv.2016.06.009.

Zerefos, C.S., 2002. Long-term ozone and UV variations at Thessaloniki, Greece. Physics and Chemistry of the Earth, Parts A/B/C 27, 455–460.

Zerefos, C.S., Tourpali, K., Eleftheratos, K., Kazadzis, S., Meleti, C., Feister, U., Koskela, T., Heikkilä, A., 2012. Evidence of a possible turning point in solar UV-B over Canada, Europe and Japan. Atmos. Chem. Phys. 12, 2469–2477.