Machine learning based retrieval of total ozone column amount and cloud optical depth from irradiance measurements

MILOS SZTIPANOV,[1,*] LEVENTE KRIZSÁN,[1] WEI LI,[1] AND KNUT STAMNES,[1]

[1]*Department of Physics, Stevens Institute of Technology, Hoboken, New Jersey, USA*
[*]*milostipanov@gmail.com*

Abstract: A machine learning algorithm combined with measurements obtained by the Norwegian Institute for Air Research (NILU-UV) irradiance meter enable the determination of total ozone column (TOC) amount, radiation modification factor (RMF), and cloud optical depth (COD). Operating in the New York City area, a NILU-UV instrument on the rooftop of Stevens Institute of Technology (40.74°N, -74.03°E) has been used to collect data for several years. Inspired by a previous study [Opt. Express **22**, 19595 (2014)], this research presents an updated neural network-based method for TOC and COD retrieval. The method presented in this paper provides reliable performance under heavy cloud conditions, and a convenient algorithm for simultaneous retrieval of TOC and COD. The TOC amounts are presented for 2014-2023, and compared with the look-up table method (LUT) and those obtained from the Ozone Monitoring Instrument (OMI, deployed on NASA's AURA satellite), and COD results for 2014-2019.

1. Introduction

Atmospheric ozone plays a crucial role in absorbing ultraviolet UV radiation, thus providing protection to the biosphere, which encompasses all living organisms including ecosystems and humans. While a moderate amount of UV radiation can stimulate vitamin D production in the skin, excessive exposure to UV radiation poses risks to both humans and most organisms. The UV spectral range is known to have harmful biological effects on living organisms and biomolecules. Thus, monitoring the total ozone column (TOC) is important.

Because the importance of monitoring the ozone layer, there is a constant motivation to improve and develop satellite and ground-based ozone amount retrieval techniques. In the near past machine learning had shown a great improvement and neural networks have been in the focus of different fields of natural sciences. Neural networks showed great flexibility and potential to be adopted to address different problems in cognitive science, information science, computer science, marketing, artificial intelligence, biology, chemistry and other sciences.

Radiative transfer simulations, Machine learning, and computer hardware have developed since 2013, from when we have the latest TOC and COD data from this area, derived by a radial basis function neural network [1]. The two main motivators of this study are (i) to follow up on the state and trends of the varying ozone layer in the New York City area and (ii) to create an updated machine learning based retrieval technique that is accurate and easy to use.

For radiative transfer modeling AccuRT has been used. The main instruments used in this research are the NILU-UV (No.115), which has been deployed and operated in the New York City area (40.74°N, -74.03°E) for several years and the Ozone Monitoring Instrument (OMI).

A detailed description of the radiative transfer simulations, retrieval methodology, machine learning and training methods used in this research is provided in the following Sections. Furthermore, comparisons of results obtained by the neural network and LUT method applied to the NILU-UV data and the OMI results as well as the relationship between the radiation modification factor (RMF) and COD are presented.

2. Methodology

The combination of radiative transfer simulations, machine learning methods and irradiance measurements were used to infer TOC and COD. The first step was using radiative transfer simulations to create training data for the neural network. After the set up of the neural network, the synthetic dataset with the absolute response function (of the NILU-UV) applied to it, was used for training. Once the training was done the validation of the results took place. This step was followed by preparing the collected NILU-UV raw data to be used as input to the neural network to infer the TOC and COD. Once the TOC results were obtained, they were compared to the OMI values, and statistics were calculated to quantify their similarities/discrepancies. In the following sections we describe (i) the instrumentation, (ii) the radiative transfer simulations, (iii) the neural network approach, (iv) the NILU-UV and OMI data preparation, (v) the results of our research, and finally (vi) a conclusion.c

3. Instrumentation

3.1. NILU-UV

The NILU-UV No. 115 irradiance radiometer used to collect data employed in this research. The NILU-UV radiometer measures voltages in each of six channels depending on the magnitude of the downward irradiance at the given channel an its bandwidth and response function. The channels of the NILU-UV instrument are centered at 302 nm, 312 nm, 320 nm, 340 nm and 380 nm with ∼ 10 nm at full width at half maximum (FWHM). A sixth channel takes measurements in the visible (or so-called photosyntetically active) PAR range, 400–700 nm with a bandwidth of 300 nm at FWHM. For the five UV channels the range of the absolute response function of the NILU-UV spans wavelengths from 290 nm to 387 nm. The NILU-UV instruments were optically characterized (spectral and cosine responses) at the optical laboratory of the Norwegian Radiation Protection Authority (NRPA) in 2005 [2]. Figure 1 shows the absolute response functions $R(\lambda)$ for each channel.
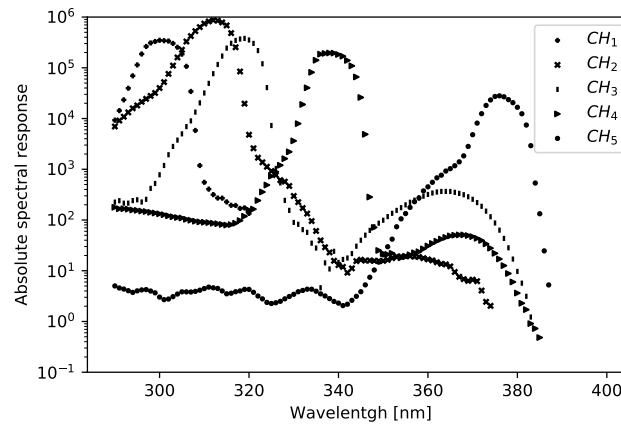

Fig. 1. The absolute spectral response of No. 115 NILU-UV instrument.

The response functions were characterized with a 1 nm resolution. Throughout this study, channels 1, 3 and 5 were used. Channel 3 is significantly less sensitive to ozone abundance than channel 1 so the ratio of the two is an ideal choice for the retrieval. At $\lambda = 380$ nm the ozone cross section is near minimal therefore channel 5 is the least sensitive to ozone but is responsive to clouds, aerosol particles and surface albedo. This makes channel 5 optimal for

RMF retrievals. The absolute response of these channels span over the whole range mentioned above dom$[R(\lambda)] = \{\lambda \in \mathbb{Z} \mid 290\,\text{nm} \leq \lambda \leq 387\,\text{nm}\}$.

The radiometer has built-in memory to store data and a temperature controller. It records data at a 1-minute time resolution. Besides the (limited capacity $\sim$ 24 day) built-in memory, data can be recorded and stored by connecting the device to a computer with a RS-232 port. The instrument is equipped with moderate bandwidth filters that tend to drift with time; hence the instrument requires a relative calibration (typically once or twice a month). The instrument also needs to be absolute calibrated before deployment. The total weight of the instrument is 3.3 kg; it is waterproof and designed to operate in harsh environments.

3.2. OMI

The Ozone Monitoring Instrument (OMI) is deployed on NASA's AURA satellite which is in a Sun-synchronous orbit. This satellite was launched on July 15, 2004. The OMI is measuring TOC amounts, UV-radiation, and aerosol abundance. OMI data are gridded at 0.25 degrees, it has a 780 × 576 (spectral × spatial) pixel CCD detector. AURA's swath is 2600 km and the nadir viewing footprint is 13 km × 24 km.

4. Radiative transfer simulations

In order to train the neural network, one needs to provide the training data to the algorithm. To create the training (or synthetic dataset), AccuRT was used. AccuRT is a unique, state of the art radiative transfer simulation package that was designed to provide a reliable, well-tested, robust, versatile, and easy-to-use radiation transfer tool for coupled (atmosphere and underlying surface) systems [3]. AccuRT uses the discrete ordinate method for the radiative transfer modeling [4].

The desired outputs from the neural network are TOC and COD at 380 nm ($\tau_{c\,(380\,\text{nm})}$). The available possible inputs from the NILU-UV are the raw measurements in the 6 channels, and an indirect product of the instrument, the so called radiation modification factor (RMF), that is described as follows:

When solar radiation passes through the ozone layer of the atmosphere, a portion of the UVB radiation will be absorbed by ozone, while the portion that penetrates the ozone layer will be multiply scattered or absorbed by air molecules, aerosols, and cloud particles [5]. To take into account the effects of clouds, aerosol particles, and surface albedo on the UV radiation a radiation modification factor (RMF) is introduced. The RMF is the measured irradiance at wavelength λ and solar zenith angle θ_0, $F_m(\lambda, \theta_0, TOC)$, divided by the calculated irradiance, $F_c(\lambda, \theta_0, TOC)$, at the same λ and θ_0 at the altitude of the site for a cloud- and aerosol-free sky and zero surface albedo.

$$RMF = \frac{F_m(\lambda, \theta_0, TOC)}{F_c(\lambda, \theta_0, TOC)} \times 100. \tag{1}$$

In this study the 380 nm channel was used to determine the RMF. The RMF defined here is relatively insensitive to the ozone abundance because the ozone absorption cross section is very small at $\lambda = 380$ nm, whereas the RMF is sensitive to clouds, aerosol particles, and the surface albedo. RMF may be larger than 100 when broken clouds are present and the direct beam from the unobscured Sun is measured by the instrument as well as diffuse sky radiation scattered by broken clouds [6].

The goal was to train the neural network in terms of the measured voltages in each channel at different atmospheric conditions for varying COD ($\tau_{c\,(380\,\text{nm})}$), TOC amounts, and solar zenith angles θ_0. Hence, the solar spectrum (290 nm - 387 nm) and the resolution of the absolute response functions (1 nm) were used in the AccuRT computations to obtain the irradiances at these wavelengths.

Cloud optical depth can not be used as a direct input to AccuRT. The cloud volume fraction ($f_{V,c}$) is used in conjunction with the cloud extinction coefficient to generate the cloud optical depth at 380 nm, $\tau_c(380\,\text{nm})$.

The input parameters and their ranges are shown in Table 1. 20,000 simulations were prepared to be computed. For each case a value was randomly selected for θ_0, O_3, and $f_{V,c}$ from the presented ranges in Table 1. A Matlab script was used to cycle through the 20000 prearranged cases and run the individual AccuRT simulations.

Parameter	θ_0	O_3	$f_{V,c}$
Minimum	17°	220 DU	10^{-11}
Maximum	70°	440 DU	10^{-6}

Table 1. Input parameters and their ranges for the AccuRT simulations.

For the simulations the US Standard Atmosphere [7] model was used and the surface albedo was chosen to be 0.14 which is typical for cities [8]. With the cloud model used (see Sec. 4.1), the limits of $f_{V,c}$ are equivalent to $\min\{\tau_c(380\,\text{nm})\} = 0.001526$ and $\max\{\tau_c(380\,\text{nm})\} = 152.596$ in terms of COD. In Table 1 $\min\{\theta_0\}$ reflects the lowest possible solar zenith angle at the measurement site, and $\max\{\theta_0\}$ the upper limit to minimize measurement error. The minimum of θ_0 was chosen to be the lowest solar zenith angle throughout the year and the maximum limit was bound to keep the measurement errors low. For $\theta_0 > 70°$ the impact of cloudiness, the vertical profile of ozone and temperature, the imperfect cosine response of the instrument, and the absolute calibration error reduce the accuracy of the results [9].

4.1. Cloud model

Clouds were assumed to consist of a collection of homogeneous water spheres having a single-mode log-normal volume size distribution with a specified volume mode radius $r_v = 10\ \mu$m and width $\sigma = 0.2$. Clouds were placed in the 13th layer in our model. This layer is between 2000 m - 4000 m altitude. A Mie code was used to compute the inherent optical properties of cloud particles.

To convert the cloud volume fraction $f_{V,c}$ to cloud optical depth further computations were done with AccuRT. Repeated simulations were done to obtain $\tau_c(380\,nm)$ at different cloud volume fractions. The result is shown in Figure 2 below.
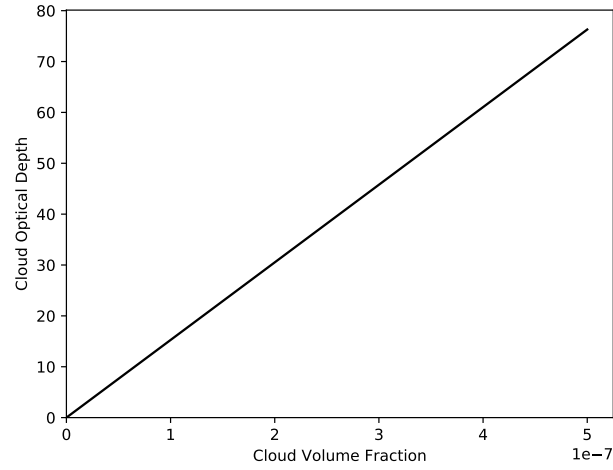
Fig. 2. $\tau_c(380\,\mathrm{nm})$ versus $f_{V,c}$.

The relationship between the RMF and $\tau_c(380\,nm)$ was investigated. Using AccuRT to calculate irradiances at 380 nm for different $f_{V,c}$ values yielded the theoretical relationship between RMF and $\tau_c(380\,nm)$ shown in Figure 3.
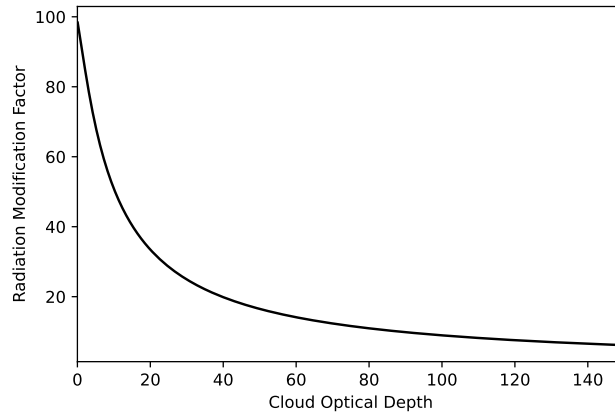

Fig. 3. Modeled relation between RMF and $\tau_c(380\,nm)$.

5. Neural network

To create a machine learning algorithm a neural network was built in Python 3 programming language with an adaptive learning rate. Three hidden layers were set up with 100, 90, 75 neurons, respectively with hyperbolic tangent activation functions.

In this section the training data preparation and the two types of validation is described below.

5.1. Training data

Besides taking the cosine of the solar zenith angle $\cos(\theta_0)$, to prepare the inputs (x_i) for the neural network training, the simulated irradiances $I_s(\lambda)$ were convolved with the absolute response

functions $R_i(\lambda)$ of the corresponding NILU-UV channels. Here i denotes the channel number. The results of this transformation represents the voltages that would be measured in the NILU-UV channels at different irradiances. For channel 1, 3 and 5 the convolutions were the following:

$$V_1 = \sum_{\lambda=290}^{320} R_1(\lambda)I_s(\lambda)$$

$$V_3 = \sum_{\lambda=290}^{383} R_3(\lambda)I_s(\lambda) \tag{2}$$

$$V_5 = \sum_{\lambda=290}^{387} R_5(\lambda)I_s(\lambda)$$

The ratio of voltages in two channels, $\frac{V_3}{V_1}$ was used for the (TOC) prediction. As a reminder, our goal is to retrieve TOC and $\tau_{c\,(380\,nm)}$ (COD at 380 nm). The inputs and outputs of the neural network are gathered in Table 2.

input	$\cos(\theta_0)$	$\frac{V_3}{V_1}$	V_5
output	TOC_{scaled}	$f_{V,c}$	-

Table 2. Input and output parameters of the neural network.

5.2. Holdout validation

Once the neural network was ready for the supervised learning, the synthetic data were applied for the training. To validate the results two procedures were done. First a 75:25 holdout, and then a K-fold validation for K=5 (see Section 5.3). For the holdout method, the mean percent error (PE), mean absolute percent error (APE) and the squared correlation coefficient (R^2) were calculated using 5000 data-points of the neural network (NN) predictions versus the modeled results. The APE and PE are defined as follows:

$$PE = \frac{1}{5000} \sum_{i=1}^{5000} \frac{NN(\underline{x}_i) - M(\underline{x}_i)}{M(\underline{x}_i)} \times 100 \tag{3}$$

and

$$APE = \frac{1}{5000} \sum_{i=1}^{5000} \frac{|NN(\underline{x}_i) - M(\underline{x}_i)|}{M(\underline{x}_i)} \times 100 \tag{4}$$

where $NN(\underline{x}_i)$ and $M(\underline{x}_i)$ are the results from the neural network and simulation (model) respectively, for the $i-th$ set of parameters (or i-th case). The results predicted by the trained neural network vs. the simulated values are plotted on Figures 4 and 5.
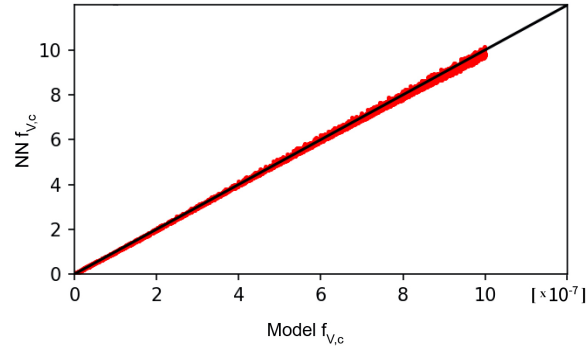
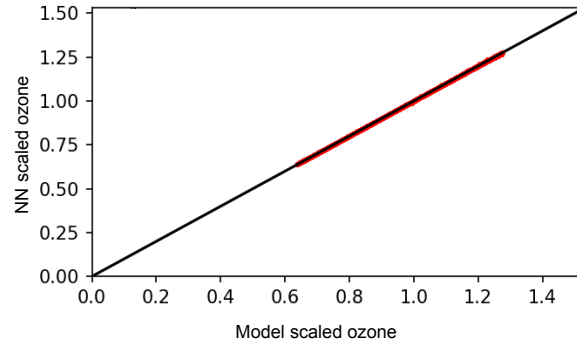Fig. 4. Cloud volume fraction results by the NN vs. modeled values.

Fig. 5. Scaled ozone amount results by the NN vs. modeled values.

In Fig. 5 the *x*-axis represents the scaled ozone amount. AccuRT uses scaled ozone amounts for the simulations. The ozone quantity is scaled to the standard US atmosphere [7], where the equivalent depth of ozone is 3.45×10^{-3} m, which in Dobson units corresponds to 345 DU. The O_3 range in Figure 5 from 220 DU to 440 DU is obtained by multiplying 345 DU by the factors 0.6377 and 1.2754, respectively. The strong correlation between the data-points is clear and the calculated statistical parameters are provided in Table 3.

	O_3	τ_c (380 nm)
R^2	1	0.999
PE	0.05 %	-0.04 %
APE	0.11 %	1.88 %

Table 3. Statistical results of the neural network validation. The sample is 5000 NN predicted vs. modeled data.

5.3. K-fold validation

A K-fold cross validation [10] was also done on the training data to see how it performs on "unseen" data. A basic cross validation technique is based on partitioning a portion of the training data and utilize it to generate predictions from the neural network. The resulting error estimation provides insight into how the model performs on unseen data (or validation set). This technique

is commonly referred to as the holdout method, a simple form of cross-validation. In this case $K = 5$ was adopted which is a common choice for this kind of validation method. K−fold cross-validation is a technique where the dataset is divided into K subsets. The holdout method is then repeated K times, with each subset being used once as the validation set, and the remaining $K − 1$ subsets being combined to form the training set. The error estimation is averaged over all K trials to determine the overall effectiveness of our model. This approach ensures that each data point is included in the validation set exactly once, and in the training set $K − 1$ times. Swapping the training and test sets further enhances the efficacy of this method. The statistical results for the K-fold cross validation are provided in Table 4.

	O_3	$\tau_c\,(380\,nm)$
R^2	0.999	0.999
PE	0.02 %	-0.08 %
APE	3.27 %	0.43 %

Table 4. Statistical results of the neural network K-fold cross validation using $K = 5$.

6. NILU-UV and OMI data preparation

Once the NN was trained and validated the preparation of the NILU-UV raw data was next. Unfortunately some of the measurement were erroneous as the instrument logged faulty data occasionally. Missing datapoints, unfinished readings and appearance of random characters were the most common. These NILU-UV data were removed from the dataset before the following steps were taken.

The NILU-UV registers the timestamps in UTC time. First the solar zenith angle θ_0 was calculated based on the location and time. The other two pieces of information needed was the ratio $\frac{CH_3}{CH_1}$ and CH_5. As mentioned in [11] section 3b, the drift factors need to be used for each channel to compensate for the instrument's Teflon diffuser's degradation. The nearest drift data to the measurement time was always used to take the instrument's drift into account. The imperfect cosine response of the instrument's channels was also accounted for.

The available data after 2013 were collected from 2014 to 2019 with some days missing, mainly from year 2016. Approximately the first third and last third of the data were lacking from 2017 and 2019, respectively. Unfortunately because of technical difficulties NILU-UV data after 2019 was significantly gappy.

Level 3 OMI data were acquired from NASA's Goddard Earth Sciences Data and Information Services Center (GESDISC) in hierarchical data format (release 5) for the years of interest.

7. Results

7.1. TOC results

The retrieved TOC amounts from the NILU-UV and OMI measurements versus day of the year are plotted for the years 2014, 2015, 2018 and 2019 in Figures 6a, 6b, 6c, and 6d. Plots for the years 2016 and 2017 are provided in the Appendix.

The plots share the same behavior, that is, the TOC values retrieved from NILU-UV tend to be lower than the corresponding OMI values. The same tendency was observed in [11] section 6, and as stated there, this tendency might indicate a solar zenith angle related error either in NILU-UV or OMI measurements.

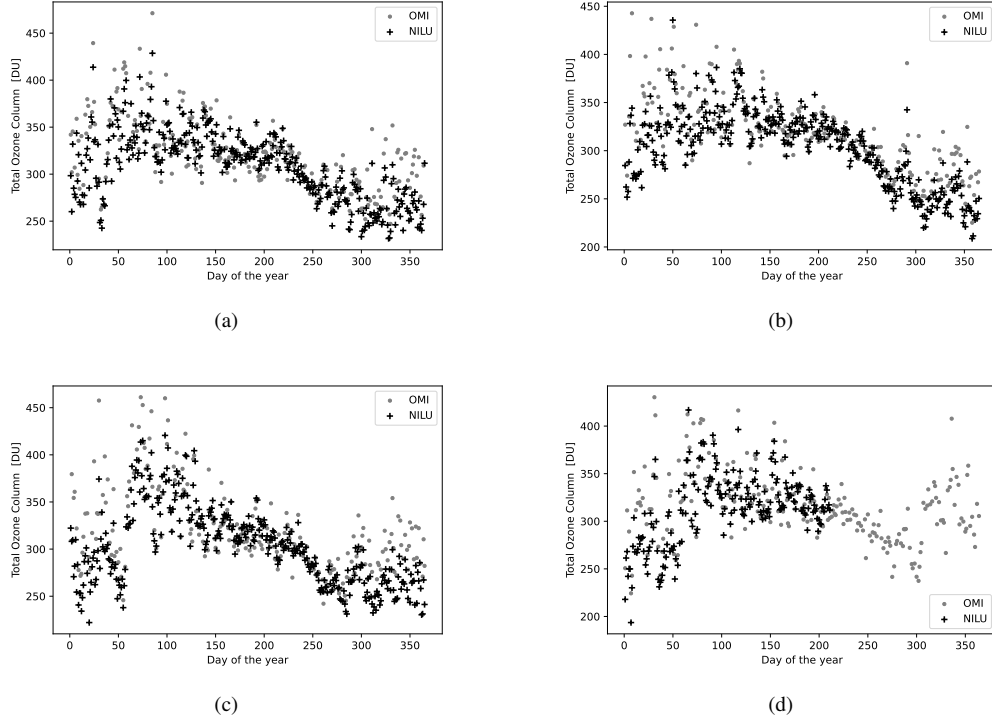For large COD values ($\tau_c\,(380\,nm)$ > 100) the NILU-UV measurements showed significant TOC

Fig. 6. TOC values from OMI (gray dots) and NILU-UV (black crosses) versus day of the year for **(a):** 2014 **(b):** 2015 **(c):** 2018 **(d):** 2019.

underestimations. This phenomenon has already been noticed by Dahlback [12], who showed that: (i) for clouds located between 2 and 4 km with an optical depth of $\tau_c = 100$ at zero surface albedo the error in the calculated TOC is less than 2 DU compared with the TOC obtained for a cloudless sky. (ii) The error will increase if the surface albedo increases. In the presence of a cloud between 2 and 4 km with optical depth $\tau_c = 100$ and a surface albedo of 0.8 the error is larger, but will not exceed 20 DU. (iii) At $\tau_c = 50$ and surface albedo of 0.8 NILU-UV underestimates the TOC by $\sim 6\%$ [6]. In our simulations the surface albedo was adopted to be 0.14 as mentioned in Section 4 and the $\tau_c(380\,\text{nm})$ occasionally exceeds 50.

Because this phenomenon occurs when the NILU-UV data analysis was done with the lookup-table method as well as with our neural network based method, it might indicate a non-linearity of the electronic components of the NILU-UV, i.e. for low irradiances (when θ_0 is large or under heavy cloud cover) the instrument registers higher voltages (in channel 3) than it is supposed to and therefore underestimates the TOC. Note that this explanation is not confirmed; it is only a speculation.

The annual TOC averages retrieved using the neural network, look-up table method, and OMI results, and the percent errors using OMI as a reference are listed in Table 5.

Daily average TOC results were obtained after large COD values ($\tau_c(380\,\text{nm}) > 100$) were filtered out. This filtering led to only very small (per mille ‰) and ~ 0.5 DU variance in the annual absolute and relative differences, but it removed the outliers. This data filtration method is an improvement compared to the suggested [13] RMF < 30 limit when using the lookup-table method. In terms of RMF, $\tau_c(380\,\text{nm}) = 100$ corresponds to a RMF value of 9.1. The relationship between RMF and $\tau_c(380\,\text{nm})$ is plotted in Figure 3.

The annual absolute differences (Δ) using OMI as the reference, the annual relative percent

Year	$\mathbf{TOC}_{NN}$ [DU]	$\mathbf{TOC}_{OMI}$ [DU]	$\mathbf{PE}_{nn}$ [%]	$\mathbf{PE}_{lut}$ [%]
2014	322.70	322.70	0	-0.23
2015	318.18	320.15	0.61	0.70
2016	-	-	-	-
2017	305.80	302.67	-0.97	-0.18
2018	320.86	319.80	-0.33	0.21
2019	323	319.45	-1.1	-0.70

Table 5. Annual TOC results and the percent error between the annual averages.

errors (PE) defined by Eq. 3, and the standard deviations (σ) are presented in Table 6 for the TOC amounts.

Year	Δ [DU]	PE [%]	σ_{OMI}	σ_{NILU}
2014	13.9	2.3	37.4	43.5
2015	12.6	2	38.9	40.1
2016	46.5	7.4	38.1	31.6
2017	16.8	2.6	29.5	37.1
2018	16.3	2.5	43.0	45.3
2019	16	2.5	38	40.2
Σ	20.4	3.2	37.5	39.6

Table 6. Annual TOC statistical results for the years 2014-2019.

As mentioned in Section 6, significant amount of data from the NILU-UV measurements were missing for the years 2016, 2017, 2019, and after. These years are marked with light gray color in Table 6 to indicate that their higher values are not representative as an annual statistic. Note, that these values were also included in the overall statistical results for all the years from 2014 to 2019 in the last line of Table 3. The available data from these years are mainly from seasons when the daily average SZA is higher, i.e. winter. The higher values in Table 6 for these years supports the observation made in regards to the SZA related error in one of the instruments.

The seasonal variations and annual periodicity of the TOC in both the NILU-UV and OMI data are clearly discernible in Fig. 7.
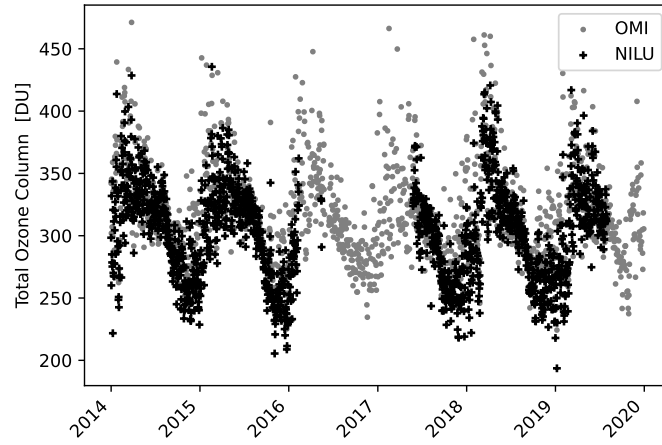
Fig. 7. The seasonal variations of TOC.

The calculated annual and semi-annual variations are caused by seasonal changes in odd-oxygen production rates, temperature-dependent ozone destruction rates, and transport by the mean circulation and by eddies [14].

Unfortunately due to technical difficulties after 2019 the available NILU-UV data significantly dropped, but to create an up to date overview of the annual TOC amounts the missing datapoints were completed by data from OMI. The yearly TOC amount results are shown in Figure 8.
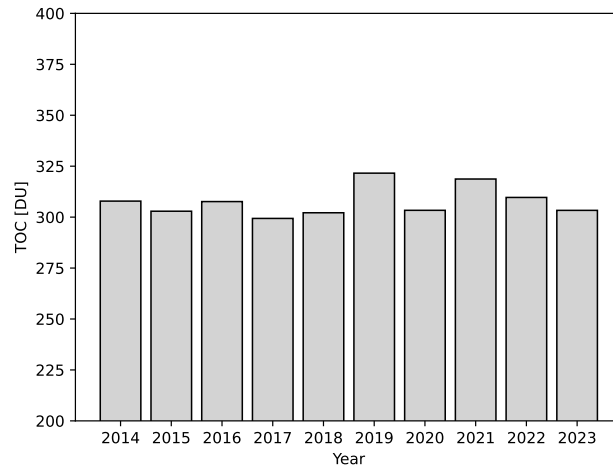
Fig. 8. Annual TOC averages from merged data of NILU-UV and OMI between 2014-2023.

7.2. COD results

As mentioned in Section 4, one of the outputs of the neural network is the cloud volume fraction, $f_{V,c}$. This output was then converted to COD at 380 nm with the relationship plotted in Fig. 2. The results for the years 2014, 2015, 2018, and 2019 are plotted in Figs. 9a–9d.
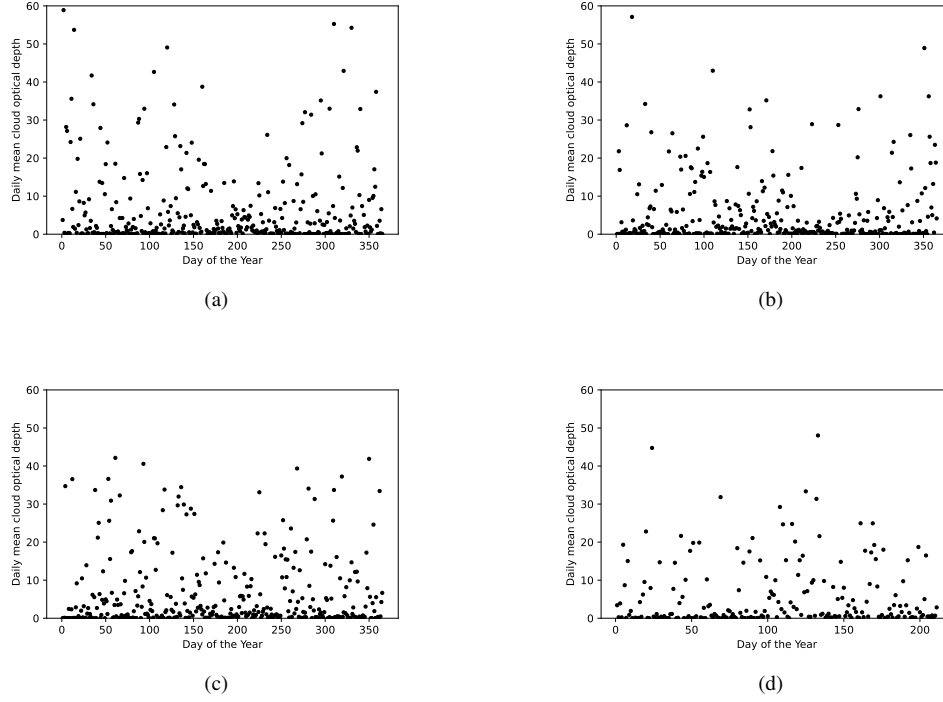
Fig. 9. Daily mean τ_c (380 nm) values versus day of the year for **(a):** 2014 **(b):** 2015 **(c):** 2018 **(d):** 2019.

The irradiance measurement in the 380 nm channel was chosen to derive τ_c (380 nm) because at this wavelength there is negligible absorption under cloud- and aerosol-free conditions, see [15] Figure 7. The estimated error due to a presence of a heavy aerosol load (i.e. aerosol volume fraction of $f_{V,a} = 10^{10}$) in the atmosphere is τ_c (380 nm) $\pm \sim 3$ using a location specific aerosol model. The main difference between RMF and τ_c (380 nm) is mainly that the RMF, derived by the lookup-table method, is unreliable under broken cloud conditions and at large surface albedos [13]. Although the absorption cross section of ozone at 380 nm is close to its minimum value, and very low ($< 10^{-24}$ cm^2), in our neural network based method τ_c is inferred simultaneously with the ozone amount and expected to give a more reliable result.

8. Conclusion

Our neural network method that retrieves TOC and COD [i.e. τ_c (380 nm)] simultaneously from a NILU-UV instrument, showed a close agreement with the OMI results. From the data shown in Figures 7 and 8 it is clear that no alarming TOC change in the trend is noticeable. The lowest annual TOC amount was in 2017 with 299.4 DU. The highest annual average of TOC was in 2019 (321.6 DU) which might have been caused by the less than typical human activity during the COVID-19 pandemic, but because the difference being relatively not very large, and the complexity of the matter, we cannot be absolutely certain how much the declined human and industrial activity contributed to this outcome.

In contrast to the lookup-table method our neural network based method takes cloud effects into account in the TOC retrieval. Even under heavily overcast conditions, τ_c (380 nm) > 100, the TOC results were consistent and showed the seasonal variation of ozone and the same patterns

as the OMI measurements. A typical choice for acceptable TOC result from the NILU-UV instrument combined with the look-up table method is when RMF > 30 for the measurement. Because if the RMF is less than 30, the cloud is deemed to be too optically thick for the NILU-UV instruments to yield reliable TOC values [16]. The neural network method showed a significant improvement in this regard as the results are in agreement with the OMI when data with RMF > 9.1 was set as a boundary.

Assessment of the error in the COD in the presence of aerosols was provided in Section 7.2. Beyond the previously discussed sources of error the large footprint and low time resolution of OMI (once a day) contribute to the discrepancies between results obtained from the NILU-UV and OMI instruments.

For the same reasons the NILU-UV is very suitable for local short- or long-term TOC, AOD, and COD monitoring.

9. Backmatter

Disclosures. The authors declare no conflicts of interest.

Data availability. The modeled synthetic data used for the neural network training and the NILU-UV raw data used in this research is available upon request from the corresponding author. All OMI data used during this study is open to the public and available from the NASA-GESDISC website [17].
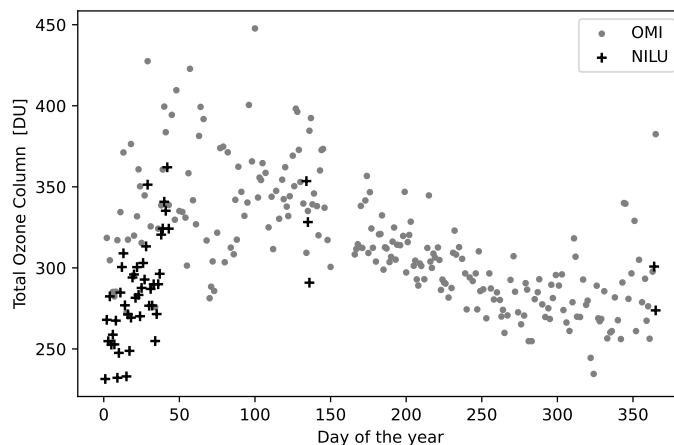
10. Appendix

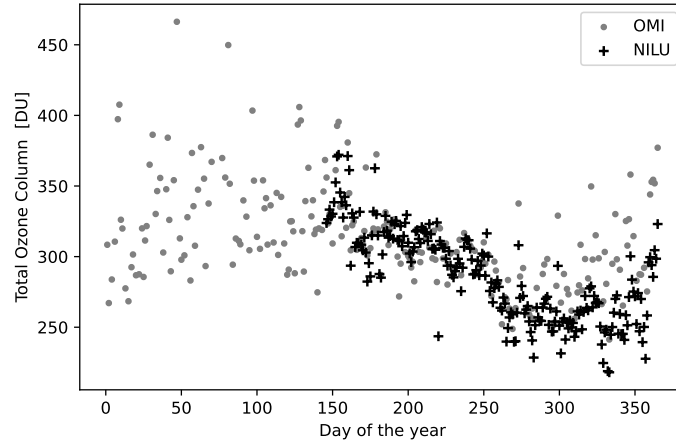Fig. 10. TOC values from OMI (gray dots) and NILU-UV (black crosses) versus day of the year for 2016.

Fig. 11. TOC values from OMI (gray dots) and NILU-UV (black crosses) versus day of the year for 2017.
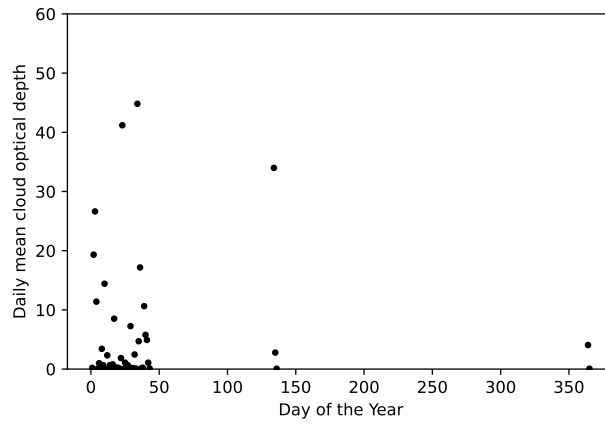

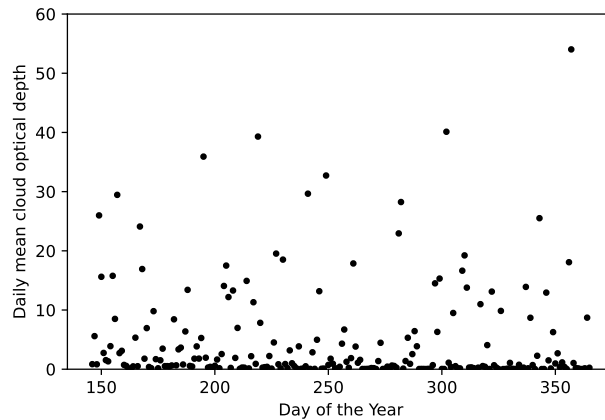Fig. 12. Daily mean $\tau_c\,(380\,\mathrm{nm})$ values versus day of the year for 2016.

Fig. 13. Daily mean τ_c (380 nm) values versus day of the year for 2017.

References

1. L. Fan, W. Li, A. Dahlback, J. J. Stamnes, S. Stamnes, and K. Stamnes, "New neural-network-based method to infer total ozone column amounts and cloud effects from multi-channel, moderate bandwidth filter instruments," Opt. Express **22**, 19595–19609 (2014).
2. T. N. Aalerud and B. Johnsen, The Norwegian UV Monitoring Network (Norwegian Radiation Protection Authority, 2006).
3. K. Stamnes, B. Hamre, S. Stamnes, N. Chen, Y. Fan, W. Li, Z. Lin, and J. Stamnes, "Progress in forward-inverse modeling based on radiative transfer tools for coupled atmosphere-snow/ice-ocean systems: A review and description of the accurt model," Appl. Sci. **8** (2018).
4. K. Stamnes, S.-C. Tsay, W. Wiscombe, and K. Jayaweera, "Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media," Appl. Opt. **27**, 2502–2509 (1988).
5. K. Stamnes, G. E. Thomas, and J. J. Stamnes, Radiative Transfer in the Atmosphere and Ocean (Cambridge University, 2017).
6. B. A. K. Høiskar, R. Haugen, T. Danielsen, A. Kylling, K. Edvardsen, A. Dahlback, B. Johnsen, M. Blumthaler, and J. Schreder, "Multichannel moderate-bandwidth filter instrument for measurement of the ozone-column amount, cloud transmittance, and ultraviolet dose rates," Appl. Opt. **42**, 3472–3479 (2003).
7. NASA, NOAA, USAF, "U. S. Standard Atmosphere," (1976).
8. H. Sugawara and T. Takamura, "Surface albedo in cities: Case study in sapporo and tokyo, japan," Boundary-Layer Meteorol. **153**, 539–553 (2014).
9. A. Kazantzidis, A. F. Bais, M. M. Zempila, C. Meleti, K. Eleftheratos, and C. S. Zerefos, "Evaluation of ozone column measurements over greece with nilu-uv multi-channel radiometers," Int. J. Remote. Sens. **30**, 4273–4281 (2009).
10. J. F. Trevor Hastie, Robert Tibshirani, The Elements of Statistical Learning (Springer, 2017), 2nd ed.
11. M. Sztipanov, L. Tumeh, W. Li, T. Svendby, A. Kylling, A. Dahlback, J. J. Stamnes, G. Hansen, and K. Stamnes, "Ground-based measurements of total ozone column amount with a multichannel moderate-bandwidth filter instrument at the troll research station, antarctica," Appl. Opt. **59**, 97–106 (2020).
12. A. Dahlback, "Measurements of biologically effective uv doses, total ozone abundances, and cloud effects with multichannel, moderate bandwidth filter instruments," Appl. Opt. **35**, 6514–6521 (1996).
13. B. A. K. Høiskar, R. Haugen, T. Danielsen, A. Kylling, K. Edvardsen, A. Dahlback, B. Johnsen, M. Blumthaler, and J. Schreder, "Multichannel moderate-bandwidth filter instrument for measurement of the ozone-column amount, cloud transmittance, and ultraviolet dose rates," Appl. optics **42**, 3472–3479 (2003).
14. L. M. Perliski, S. Solomon, and J. London, "On the interpretation of seasonal variations of stratospheric ozone," Planet. Space Sci. **37**, 1527–1538 (1989). Special Issue G.M.B. Dobson: Atmospheric Ozone Issue.
15. M. Sztipanov, W. Li, A. Dahlback, J. Stamnes, T. Svendby, and K. Stamnes, "Method for retrieval of aerosol optical depth from multichannel irradiance measurements," Opt. Express **31**, 40070–40085 (2023).
16. L. Fan, W. Li, A. Dahlback, J. J. Stamnes, S. Engelhardt, S. Stamnes, and K. Stamnes, "Comparisons of three nilu-uv instruments deployed at the same site in the new york area," Appl. Opt. **53**, 598–3605 (2014).
17. NASA, "Goddard Earth Sciences Data and Information Services Center," https://disc.gsfc.nasa.gov.