ABSTRACT

This paper reports on the design and validation of a neural network algorithm for tropospheric ozone retrievals from the Ozone Monitoring Instrument (OMI). The neural network combines OMI ultraviolet reflectances with temperature and tropopause pressure information to yield global estimates of the tropospheric ozone column. The algorithm was validated against ozonesonde data and Chemistry/Transport Model (CTM) simulations. The results seem to indicate that the algorithm is able to reproduce the global tropospheric ozone distribution with reasonable accuracy.

Key words: Neural Networks, Tropospheric Ozone, OMI.

1. INTRODUCTION

Ozone is a very important constituent for tropospheric chemistry. First, relatively high ozone concentrations near the Earth’s surface are detrimental for the biosphere [1]. Second, ozone in the troposphere is a precursor of the hydroxyl radical, which is the most important tropospheric oxidant, and as such removes several tropospheric pollutants [2]. Third, tropospheric ozone is an important greenhouse gas [3].

Quantifying tropospheric ozone concentrations from satellite observations is difficult. High spectral resolution measurements are required in order to separate the tropospheric ozone signal from the much stronger stratospheric contribution in the measured radiance spectra, and a pixel size of about 10 km or better is required in order to study the small scale variability of tropospheric ozone [4]. The latest atmospheric hyperspectral sounders, operating in the ultraviolet (UV), visible (VIS) and thermal infrared (TIR) spectral ranges, usually meet these requirements, and thus are suitable to study tropospheric ozone.

There are several ways to retrieve tropospheric ozone from satellite data, for example by subtracting a stratospheric ozone column measurement from a total ozone measurement performed with a different sensor [5], or by attempting to extract tropospheric ozone information from nadir reflectance spectra acquired by a single instrument [6]. The latter can be done via physical algorithms or statistical methods, like Neural Networks (NNs) [7].

Supervised NNs approximate functional relationships between variables by fitting the coefficients of a highly nonlinear model to a training dataset. Such dataset contains a number of coincidences between the input and the output variables of the algorithm. Provided that the dataset is representative of the statistical properties of the quantities to be related, a NN can approximate the relationship between such quantities. In the context of atmospheric retrievals from remotely sensed data, the use of NNs avoids the need for forward calculations, which are computationally expensive and rely on a priori assumptions on unmeasured components of the atmospheric state. After the training phase is completed, a NN can be used to perform accurate retrievals with a reduced computational effort. This is an attractive feature when time constraints on the delivery of the retrieval products are present, and when large amounts of data are to be processed, as is the case with the most recent hyperspectral atmospheric sounders, whose measurements are characterized by very high spectral resolutions, wide swaths and pixel sizes of the order of 10 km.

In this work, the development and validation of a NN algorithm for tropospheric ozone retrievals from the Ozone Monitoring Instrument (OMI) [8] will be discussed. The algorithm – named the OMITROPO3-NN – performs daily global tropospheric ozone retrievals using OMI measurements in the UV2 channel, together with ancillary information on tropopause pressure and atmospheric temperature profile. In order to validate the algorithm, the retrieved tropospheric ozone columns were compared to ozonesonde measurements over middle, tropical and polar latitudes, and daily tropospheric ozone fields simulated using the TM5 Chemistry and Transport Model (CTM) [9].
2. ALGORITHM DESIGN

The training dataset for the OMITROPO3-NN consists of OMI reflectance spectra in the 310 – 345 nm spectral range, co-located with ozonesonde profiles taken from the World Ozone and Ultraviolet Data Center (WOUDC), the Network for Detection of Atmospheric Composition Change (NDACC) Southern Hemisphere Ozonesondes (SHADOZ) network, as well as data from the Intercontinental Chemical Transport Experiment-B (INTEX-B) Ozonesonde Network Study 2006 (IONS06), the Arctic Intensive Ozonesonde Network Study (ARCIONS), and data from ozone soundings over Italy provided by the Center for Integration of remote sensing techniques and numerical modeling for the prediction of severe weather (CETEMPS), the Institute of Atmospheric Sciences and Climate (ISAC) of the Italian National Research Council (CNR) and the Italian Air Force Centre of Aeronautical Meteorological Experimentation (ReSMA). The data were acquired between 2004 and 2011.

The following criteria were followed in the co-location between sonde and OMI data: i) For each ozone sounding, the OMI orbits passing over the ozonesonde station were found; ii) for each overpassing orbits, the OMI pixel whose center was closest to the ozonesonde station was selected as a candidate for the co-location; iii) the candidate pixel was considered as co-located if its center and the station coordinates were less than or equal to 0.3 degrees apart – either in latitude or longitude – and discarded otherwise.

The dataset for the OMITROPO3-NN consists of 10,017 pairs. The data were further split into three subsets: 5,489 out of such data were used to train the NN; 1,737 data were used to determine when to stop the training through early stopping cross-validation; 2,791 data were used to evaluate the NN after the training phase.

Prior to the neural processing, all the quality flags for the OMI radiance and irradiance spectra were checked, in order to detect the spectral pixels to be discarded. A spectrum was considered eligible for further processing if the number of discarded wavelengths was less than the 5% of the total number of wavelengths. After this screening procedure, the radiance and irradiance spectra were linearly interpolated on a 0.1 nm wide common grid, and the logarithm of the reflectance spectra were computed. The log-reflectance spectra were then compressed using a Principal Component Analysis (PCA), and 20 Principal Components (PCs) were retained. The Solar Zenith Angle (SZA), the instrument View Zenith Angle (VZA) and the terrain height of the observed scene were used in the input vector to describe the acquisition geometry. The temperature profile and tropopause pressure from the National Center for Environmental Prediction (NCEP) / National Center for Atmospheric Research (NCAR) Reanalysis [10], as well as the Tropospheric Ozone Residual (TOR) monthly mean from [11] and the cloud fraction – taken from the OMI total ozone Level 2 product – were used as additional inputs. The OMITROPO3-NN performs tropospheric ozone retrievals when the cloud fraction is less than or equal to 0.3.

The NN architecture consists of 43 inputs, one hidden layer with 5 neurons and one output. The hidden and output layer have logistic activation functions.

3. COMPARISON WITH OZONESONDE DATA

Figure 1 shows the results of the comparisons between the 2,791 ozonesonde data which were not used during the NN training phase and the corresponding retrievals. The algorithm performance has been analyzed as a function of latitude by computing separate statistics over five latitude bands: i) Antarctica (90°S-60°S), ii) Southern midlatitudes (60°S-30°S), iii) Tropics (30°S-30°N), iv) Northern midlatitudes (30°N-60°N), and v) Arctic (60°N-90°N). The results are shown in Table 1, where the mean bias, the Root Mean Square Error (RMSE) and the Pearson correlation coefficient between measured and retrieved tropospheric ozone columns are reported. It can be seen that the algorithm performs reasonably well at all the latitudes, except the Arctic, where a significant negative bias and a low correlation coefficient are observed.

As an example of local validation results, measured and retrieved Tropospheric Column Ozone (TCO) at the Davis ozonesonde station (Antarctica) are shown in Figure 2. Figure 2(a) shows a scatter plot of retrieved versus in situ TCOs. In Figure 2(b), the same results are reported in the form of a time series. The OMITROPO3-NN seems able to reproduce the temporal variations of TCO with a reasonable agreement. Also the extreme high and low values in the time series are fairly well reproduced by the NN.

4. COMPARISON WITH TM5 SIMULATIONS

As an example of TCO map obtained by applying the OMITROPO3-NN to an entire day of OMI data, the
Table 1. Results of the comparison between OMITROPO3 NN and ozonesondes, stratified by latitude.

<table>
<thead>
<tr>
<th>Latitude band</th>
<th>Mean bias (DU)</th>
<th>RMSE (DU)</th>
<th>Pearson coeff.</th>
<th>N. data</th>
</tr>
</thead>
<tbody>
<tr>
<td>90°S–60°S</td>
<td>1.99</td>
<td>5.63</td>
<td>0.86</td>
<td>271</td>
</tr>
<tr>
<td>60°S–30°S</td>
<td>1.45</td>
<td>5.22</td>
<td>0.76</td>
<td>181</td>
</tr>
<tr>
<td>30°S–30°N</td>
<td>0.59</td>
<td>5.65</td>
<td>0.80</td>
<td>611</td>
</tr>
<tr>
<td>30°N–60°N</td>
<td>0.52</td>
<td>5.28</td>
<td>0.82</td>
<td>1357</td>
</tr>
<tr>
<td>60°N–90°N</td>
<td>−2.69</td>
<td>5.66</td>
<td>0.54</td>
<td>371</td>
</tr>
</tbody>
</table>

Figure 2. Retrieval results at the Davis station (Antarctica): (a) Scatter plot; (b) Time series.

Figure 3. TCO map produced by the OMITROPO3-NN algorithm on 15th October 2006.

Figure 4. Scatter plot of OMITROPO3-NN versus TM5 TCOs on 15th October 2006.

global TCO field at full OMI resolution retrieved on 15th October 2006 is shown in Figure 3. Apart from an along-track striping effect which is probably due to the daily Level 1b irradiance product, the map seems to capture the most relevant patterns in the global TCO distribution.

In order to evaluate the quality of the TCO map shown in Fig. 3, a comparison with maps of TCO obtained from TM5 simulation during the same day was performed. Figure 4 shows a scatter plot of OMITROPO3-NN retrieval versus the TM5 simulation. In order to perform the comparison, the OMITROPO3-NN TCO fields were remapped on the same grid as TM5 (3° × 2°, longitude × latitude) using a nearest neighbour resampling, and the grid points where a retrieved TCO was available were selected. The correlation coefficient, which exceeds 0.80, shows that the retrieved and simulated TCO spatial patterns are quite similar. However, the OMITROPO3-NN seems to have a positive bias with respect to TM5. A comparison with ozonesondes will be necessary to understand whether this is due to inadequacies in the OMITROPO3-NN or to an overestimation tendency in TM5.

An insight into the spatial structure of these differences is given by Figure 5, where a map of the absolute differences between OMITROPO3-NN and TM5 is shown. Although the general tendency of the OMITROPO3-NN is to estimate higher TCOs than TM5, this tendency is not observed in every part of the globe. In fact, a large area over which NN TCOs are lower than TM5 simulations can be observed around India and Southeastern Asia.

Comparisons between OMITROPO3-NN and TM5 over
5. CONCLUSIONS

A Neural Network scheme has been developed to estimate the tropospheric ozone column – up to the thermal lapse rate tropopause – from NASA Aura OMI data. The NN has been trained with a global set of ozonesonde data, co-located with OMI reflectance spectra, NCEP/NCAR Reanalysis data and a tropospheric ozone climatology.

The algorithm has been validated against ozonesonde data not used during the training phase, and against CTM simulations. The first experimental results suggest that the algorithm performs reasonably well in all the latitude bands. Further analyses will involve systematic comparisons between with CTM simulations over an extended time frame and a quantification of the contribution of the different inputs to the NN retrieval.

It is worth to point out that the proposed methodology can also be applied to other UV/VIS spectrometers, such as the Global Ozone Monitoring Experiment 2 (GOME-2), in order to obtain an extended dataset of tropospheric ozone observations.

ACKNOWLEDGEMENTS

All the PIs of the ozonesonde stations whose data were used are gratefully acknowledged. Marco Cervino (ISAC-CNR) is gratefully acknowledged for the help with San Pietro Capofiume ozonesonde data, Marco Iarlori and Vincenzo Rizi (CETEMPS) are gratefully acknowledged for their help with L’Aquila ozonesonde data, Stefania Vergari and Emanuele Vuerich (ReSMA, Italian Air Force) are gratefully acknowledged for their help with Vigna di Valle ozonesonde data.

REFERENCES