A study on harmonizing total ozone assimilation with multiple sensors

Yves J. Rochon[1], Michael Sitwell[1] and Young-Min Cho[2]

[1]Atmospheric Science and Technology Directorate, Environment and Climate Change Canada, Toronto, M3H 5T4, Canada
[2]Centre for Research in Earth and Space Science, York University, Toronto, M3J 1P3, Canada

Correspondence to: Yves J. Rochon (yves.rochon@canada.ca)

Abstract. Bias estimations and corrections of total column measurements are applied and evaluated with ozone data from satellite instruments providing near-real time products during Summer 2014 and 2015 and Winter 2015. The developed standalone bias correction system can be applied in near-time time chemical data assimilation and long-term reanalysis. The instruments to which these bias corrections were applied include the Global Ozone Monitoring Experiment-2 instruments on the MetOp-A and MetOp-B satellites (GOME-2A and GOME-2B), the total column ozone mapping instrument of the Ozone Mapping Profiler Suite (OMPS-NM) on the Suomi National Polar-orbiting Partnership (S-NPP) satellite, and the Ozone Monitoring Instrument (OMI) instrument on the Aura research satellite. The OMI dataset based on the TOMS version 8.5 retrieval algorithm was chosen as the reference used in the bias correction of the other satellite-based total column ozone datasets. OMI data was chosen for this purpose instead of ground-based observations due to OMI's significantly better spatial and temporal coverage, as well as interest in near-real time assimilation. Ground-based Brewer and Dobson spectrophotometers, and filter ozonometers, as well as the Solar Backscatter Ultraviolet satellite instrument (SBUV/2), served as independent validation sources of total column ozone data. Regional and global mean differences of the OMI-TOMS data with measurements from the three ground-based instrument types for the three evaluated two month periods were found to be within 1 %, except for the polar regions where the largest differences from the comparatively small dataset in Antarctica exceeded 3 %. Values from SBUV/2 summed partial columns were typically larger than OMI-TOMS on average by 0.6 to 1.2 %, with smaller differences than with ground-based over Antarctica. Bias corrections as a function of latitude and solar zenith angle were performed for GOME-2A/B and OMPS-NM using colocation with OMI-TOMS and three variants of differences with short-term model forecasts. These approaches were shown to yield residual biases of less than 1 %, with the rare exceptions associated with bins with less data. These results were compared to a time-independent bias correction estimation that used colocations as a function of ozone effective temperature and solar zenith angle which, for the time period examined, resulted in larger residual biases for bins whose bias varies more in time.

The impact of assimilating total column ozone data from single and multiple satellite data sources with and without bias correction was examined with a version of the Environment and Climate Change Canada variational assimilation and forecasting system. Assimilation experiments for the July-August 2014 show a reduction of global mean biases for short-term forecasts relative to ground-based Brewer and Dobson observations from a maximum of about 2.3 % in the absence of bias correction to less than 0.3 % in size when bias correction is included. Both temporally averaged and time varying mean

differences of forecasts with OMI-TOMS were reduced to within 1 % for nearly all cases when bias corrected observations are assimilated for the latitudes where satellite data are present.

1 Introduction

Total column ozone biases from satellite measurements are typically within a few percent. Changes of a few percent over time or between instruments are significant in affecting the correct identification of long-term trends. Near-global reductions in column ozone have been -1.8 % per decade from 1980 to the mid-1990s and increases over the past two decades are at only 0.4 to 0.6 % per decade (Steimbrecht et al., 2018). A requirement on the long-term stability of corrected total column ozone observations of 1-3 % per decade was specified by the Ozone_cci project of the European Space Agencies' Climate Change Initiative program in Table 5 of Van Weele et al. (2016). This table also indicates accuracy requirements on total column ozone measurements of 2 % for facilitating research on the evolution of the ozone layer from radiative forcing and 3 % for studies on short-term, seasonal, and interannual variability. As an example, for an accuracy requirement of 2 % and measurement precisions between 1.0 and 1.7 %, biases need to be no larger than about 1.7 to 1.0 %, respectively. The comparison of column ozone data from different instruments allows for the identification of the level of agreement between datasets, potentially under various conditions, and can highlight cases and conditions with small to large relative biases. As such, sources that provide accurate and stable long-term datasets can potentially be used to provide corrections for other sources.

The validation of satellite remote sounding products usually includes a comparison to ground-based measurements, which provide a long-term reference record. For satellite instruments measuring column ozone, this typically consists of comparisons to Brewer and Dobson spectrophotometers, and potentially filter ozonometers. The main advantage of ground-based versus satellite total column ozone measurements is that they can view the sun directly as oppose to relying on the backscatter of solar radiation, reducing the complexity and error sources of retrievals. The final resulting systematic errors of the calibrated ground-based total column ozone daily averages for well-calibrated and maintained Brewer and Dobson instruments are no larger than ~1.5-2 %, excluding sites with outlier characteristics (considering Fioletov et al., 1999 and 2008). Much of the ground-based total column ozone data may be available soon after the measurements, with the original calibration usually being sufficient. For exceptional cases where the original calibration may have been faulty, a final calibration for the ground-based total column

ozone may lag by one to two years from near-real time. Previous studies have examined the dependence of the differences between the satellite-based and ground-based total column ozone measurements on latitude, solar zenith angle, viewing zenith angle, seasonal dependence, cloud cover, reflectivity, and the ozone effective temperature, as well as other factors, for instruments such as for the Ozone Monitoring Instrument (OMI; Balis et al., 2007a; Viatte et al., 2011; Koukouli et al., 2012; Bai et al., 2016), the Global Ozone Monitoring Experiment-2 (GOME-2; Balis et al., 2007b; Antón et al., 2009b, 2011; Loyola et al., 2011; Koukouli et al., 2012 and 2015; Lerot et el.,, 2014; Hao et al., 2014; Garane et al., 2018), and the Ozone Mapping Profiler Suite (OMPS; Bia et al., 2013, 2016; Flynn et al., 2014), as well as studying their long-term stability (van der A et al., 2010 and 2015). In this paper, an observation dataset that serves as a reference in a bias estimation is referred to as the *anchor*. Reanalysis studies covering many years, such as van der A et al. (2010 and 2015), have directly used ground-based data as the anchor. A limitation in the use of ground-based observations as an anchor in bias estimation is that these observations are only available for certain locations, leaving many areas uncovered, especially in the Southern Hemisphere and over oceans. For the Southern Hemisphere, the applied bias parameterization may not necessarily capture as much of the spatial or instrument-to-instrument variations of the bias as compared to using observations from a satellite-borne instrument that covers a larger domain. If a satellite-based anchor is employed, it should ideally be in good agreement with ground-based measurements. Considering the limited projected lifetimes and possible deteriorations or failures of satellite-based instruments, transitions to new references would also be required in an operational setting.

Total column ozone bias estimation for observations can be performed in different ways and depend on different factors, such as the solar zenith angle (SZA), latitude, and season, among others. Seasonal and related latitudinal changes in biases may result from limitations in retrieval algorithms. For example, the retrieval algorithm might not adequately account for the temperature dependence of the ozone absorption coefficients. Differences and limitations in accounting for clouds and surface albedos may also contribute to errors in total column ozone (e.g., Antón et al., 2009a). Bias parameterizations may range from being spatially and temporally global to more local.

The harmonization of different datasets through bias correction can be applied for standalone analyses, reanalyses, and in near-real time data assimilation. The assimilation process consists of introducing information from observations into model forecasts through the generation of analyses, the statistical blend of earlier forecast and observations, which serve as the initial conditions for subsequent forecasts. The assimilation of column and stratospheric ozone measurements for ozone-layer forecasting has been conducted mostly as of about twenty five years ago, ultimately culminating with operational ozone-layer and UV-index forecasts (e.g., Lahoz and Errera, 2010; Inness et al., 2013). This typically involves the application of measurements from single to multiple satellite remote sounding instruments with the use of ground-based and other remote sounding data for independent verifications and, occasionally, bias correction.

Traditionally, the assimilation process assumes that both the model forecasts and observations are statistically unbiased following an initial spin-up time (unless biases are estimated within the analysis step). Unremoved biases or systematic errors in the observations or forecasting model can potentially impact the quality of the analyses and forecasts (e.g., Dee, 2005; Dragani and Dee, 2008). This is important for total column ozone when it comes to monitoring for multi-decadal trends, as

referenced in van der A et al. (2010), for both trends inferred from just the observations themselves or from their use within a data assimilation system. Generally, while the effectiveness of bias correction schemes in removing biases is constrained by limited knowledge of the truth, their impact in reducing relative biases between different assimilated observations and/or correlated fields can potentially be just as significant for improved forecasting. An example of the later is in multivariate assimilation, where ozone and meteorological assimilation can be coupled (e.g., Dee, 2008; Dee et al., 2011).

Ideally, the anchor used within a bias correction scheme should be accurate, have a wide range of coverage in both space and time, and for near-real time applications be available within a few hours or less after measurements are taken. The summed partial columns from SBUV/2 satellite instruments have been recommended as an anchor for long-term studies (Labow et al. 2013). This is due to the long-term coverage provided by the series of SBUV/2 instruments, combined with the low variations in time of the differences between these instruments and ground-based data (usually within ±1 %, but reduced for recent years). Labow et al. (2013) also show differences over time of SBUV/2 with OMI data remaining within about 1 and 2 % for the Northern Hemisphere (based on the Total Ozone Mapping Spectrometer (TOMS) version 8.5 total column retrieval algorithm, an enhancement of the version 8 algorithm described by Bhartia and Wellemeyer, 2002). McPeters et al. (2015), showing similar magnitudes and stability of differences in time, concluded that OMI-TOMS data could be used in trend studies. The merging of OMI with SBUV/2 and earlier TOMS instrument data for this purpose was performed by Chehade et al. (2014).

The focus of this study is bias estimation and correction of column ozone for multiple satellite sensors, towards eventual use in near-real time data assimilation. The bias estimation and correction methods developed in this study may be integrated into an assimilation scheme, and so can be applied in near-real time, and could be utilized for other constituents. In this paper, we evaluate several different bias estimation schemes used to correct observations of column ozone from satellite-borne instruments. Many of these methods utilize colocated observation sets for bias estimation. From this consideration, OMI-TOMS was chosen as the anchor for bias estimation and correction, as its dense spatial coverage allows for more colocations with measurements from other instruments. As part of this work, the OMI-TOMS column ozone data were evaluated using ground-based Brewers, Dobsons, and filter ozonometers observations, as well as compared to SBUV/2 column ozone, for the limited time periods of interest in this study. For these datasets, a target maximum residual bias of 1 % following bias corrections was selected. This satisfies the column ozone 2 % accuracy requirement from European Space Agencies' Climate Change Initiate program (Van Weele et al., 2016) for random error levels of up to 1.7 %.

In this paper, we examine several bias correction methods that use a discrete binning in latitude and solar zenith angle that, unlike a functional parameterization, allows for arbitrary nonlinear dependencies. In addition, an alternative estimation involving the dependency on the ozone effective temperature (the mean temperature weighted by the ozone profile), as employed in van der A et al. (2010 and 2015), was explored. However, as discussed later in the paper, dependencies on factors such as changes in cloud cover and viewing zenith angle, were not examined.

Following bias estimation, data assimilations of column ozone observations from individual and multiple satellite instruments were conducted with and without bias correction. The impacts on the resulting six-hour forecasts were then assessed. The assimilations were conducted with the Environment and Climate Change Canada (ECCC) meteorological

assimilation system adapted for constituent assimilation. These assimilations were univariate ozone assimilations and utilized operational ECCC meteorological analyses. The data sources assimilated in this study and correspondingly involved in the bias estimation analysis are the GOME-2 instruments on the European MetOp-A and MetOp-B satellites (Munro et al., 2016; Hassinen et al., 2016), the total column measuring instruments of OMPS (Dittman et al., 2002a,b; Flynn et al., 2006) on the Suomi National Polar-orbiting Partnership (S-NPP) satellite, and OMI aboard the Aura research satellite (Levelt et al., 2018).

This paper is organized as follows: Section 2 describes the utilized ozone observations covering July-August 2014 and 2015 and January-February 2015. Following a general quality assessment of the OMI data based on available literature, Section 3 evaluates the OMI column ozone data for these periods against ground-based measurements. Having assessed the quality the OMI data for these specific periods, Section 4 describes and applies three different bias estimation approaches with the column ozone measurements of different satellite instruments relative to OMI. The impact of column ozone assimilation on six-hour forecasts for individual and multiple sensors with and without bias corrections is examined in Section 5 for July-August 2014 using comparisons to both OMI-TOMS and ground-based data. Conclusions are provided in Section 6. The Supplemental material document for this paper provides additional figures and tables supporting and complementing the discussed and presented results; only the tables are directly referenced in the paper.

2 Observations

In this section, we give a brief description of the column ozone observations involved in the implementation and evaluation of bias correction as well as the observations used for the validation of short-term forecasts. Observational data set were obtained for the periods of July-August of 2014 and 2015, and January-February 2015. The main data sources of interest are those specifically intended to provide satellite-based column ozone allowing near-real time (NRT) assimilation. These consist of OMI, GOME-2, and OMPS-NM (total column Nadir Mapper) instruments that rely on optical solar backscatter of ultraviolet radiation in the nadir or near-nadir and provide data only during daytime. Ground-based Brewer, Dobson, and ozonometer filter instruments and additional satellite-based data from OMPS-NP (partial column Nadir Profiler) and SBUV/2 are included for evaluation and validation purposes.

2.1 OMI

The Ozone Monitoring Instrument (OMI) aboard the Aura research satellite has been in operation since August 2004. The instrument stems from a collaboration between the Netherlands Agency for Aerospace Programmes (NIVR), now called the Netherlands Space Office (NSO), and the Finnish Meteorological Institute (FMI). The OMI instrument provides a cross-track width of about 2600 km on the ground and total column ozone mapping at a spatial resolution of 13 km along, and 24 km across, the orbit ground track at nadir (e.g., Bhartia and Wellemeyer, 2002; OMI Data User's Guide, 2012). Some strips of the

OMI measurement tracks were removed due to the row anomaly of the OMI instrument, which for the time period under consideration, effects 23 of the 60 rows.[1]

Two different total column ozone products are derived from the OMI radiances, one processed by NASA based on the Total Ozone Mapping Spectrometer (TOMS) total column retrieval algorithm (versions 8 and 8.5) and the other made by the Royal Netherlands Meteorological Institute (KNMI) using the Differential Optical Absorption Spectroscopy (DOAS) algorithm. The OMI-TOMS algorithms (Bhartia and Wellemeyer, 2002) principally utilizes only two different wavelengths, one with strong and one with weak ozone absorption, to estimate the total column ozone and surface reflectivity. In the DOAS algorithm (Veefkind and de Haan, 2002; Veefkind et al., 2006), first the slant column density is retrieved from a spectral least squares fit to the measured ratio between the Earth radiance to solar irradiance using 25 wavelengths spanning 331 to 337 nm. The slant column density is then converted to a vertical column using the air mass factor. Overall, the two different retrievals agree to a high degree, with the global average falling within 3 % of one another, with the largest differences occurring for cloudy conditions and in the polar regions (Kroon et al., 2008).

This study employs the OMS-TOMS V8.5 standard science data column ozone products which are close to, but can differ slightly from, the OMI-TOMS NRT data (OMI NRT Data User's Guide, 2010; Durbin et al., 2010). The OMI NRT Data User's Guide (2010) and Durbin et al. (2010) indicate a daily maximum percentage difference of 2.6 % between the standard science and NRT products, with a weekly average maximum difference of 1.4 %. Further comparisons by the authors show mean differences generally between 0.02-0.04% July-August 2016 and January-February 2017. The OMI-TOMS column ozone has estimated root-mean-squared errors of 1-2 % (OMI Data User's Guide, 2012).

2.2 GOME-2

Global Ozone Monitoring Experiment-2 (GOME-2) instruments are on the MetOp-A (GOME-2A) and MetOp-B (GOME-2B) polar orbiting satellites, launched in October 2006 and September 2012, respectively, and are operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). As of July 15 2013, GOME-2A has been operating with a swath width of 960 km and a 40 km × 40 km spatial resolution, while GOME-2B has a larger swath width of 1920 km and a 40 km × 80 km spatial resolution (e.g., GOME-2 ATBD, 2015; ATBD stands for the Algorithm Theoretical Basis Document).

Total column ozone retrievals are available from EUMETSAT relying on the DOAS approach (Loyola et al., 2011) and from the National Environmental Satellite, Data, and Information Service (NESDIS/NOAA) with retrievals based on the TOMS V8 algorithm (e.g., Zhand and Kasheta, 2009). The DOAS total column ozone products have estimated accuracies of better than 3.6-4.3 % (for clear to cloudy conditions) and 6.4-7.2 % for SZA below and above 80°, with precisions of under 2.4-3.3 % and 4.9-5.9 % (GOME User Manual, 2012; GOME-2 ATBD, 2015). The GOME-2 NRT products used here, as well

[1] See http://projects.knmi.nl/omi/research/product/rowanomaly-background.php for more information and updates regarding the OMI row anomaly.

as those form OMPS and SBUV/2, were acquired from NESDIS/NOAA and stem from the TOMS approach. This study provided an opportunity to evaluate the biases of the GOME-2 TOMS products.

2.3 OMPS

The Ozone Mapping Profiler Suite (OMPS) on the Suomi National Polar-orbiting Partnership (S-NPP) satellite, launched October 2011, consists of a combined nadir mapper (OMPS-NM) and nadir profiler (OMPS-NP), and a separate limb profiler (OMPS-LP), which provide total column, partial column profile, and limb profile products, respectively. A second suite was placed onboard the Joint Polar Satellite System JPSS-1 satellite (Zhou et al., 2016), renamed NOAA-20 and launched in November 2017. The retrieved data used in this study are from the OMPS S-NPP nadir measurements and are considered to be at a provisional product maturity level. They do not include improvements from the various corrections, calibration adjustments, and retrieval algorithm updates performed since the original near-real time acquisition for the July-August 2014 period (personal communication from L. Flynn, NOAA, 2016). The OMPS-NM and OMPS-NP ozone retrievals from the SBUV V8.6 retrieval algorithms (Bhartia et al., 2013; as referred by Bai et al., 2016) became available after the completion of the assimilation experiments conducted for this work.

The OMPS-NM retrievals, summarized by Flynn et al. (2014), were made at the NOAA Interface Data Processing Segment using the ratio of the measured Earth radiances to solar irradiances at multiple triplets of wavelengths. The nadir mapper has a cross-track width of about 2800 km and a 50 km × 50 km resolution at nadir. Flynn et al. (2014) provides total column ozone accuracy and precision requirements of ~3.5-4 % and ~2 %, respectively, for SZA up to 80° and found average biases of -2 to -4 % with respect to the OMI-TOMS and the Solar Backscatter Ultraviolet SBUV/2 satellite instrument products.

The results of the evaluations from Bai et al. (2015, 2016) for the more recent OMPS-NM total column ozone products based on the SBUV V8 and V8.6 retrieval algorithms, respectively, are consistent with Bai et al. (2013). Bai et al. (2015) indicate global mean differences of OMPS-NM with ground-based data of 0.59 % for Brewer measurements and 1.09 % for Dobson measurements, with standard deviations close to 3 % for the same period as Bai et al. (2013). As a reference, Bai et al. (2016) provide a distribution of OMPS-NM minus OMI-TOMS values with a mean of 7.6 DU (~2.5 % for a total column of 300 DU) and a standard deviation of 5.8 DU at the Tsukuba station (36.1° N, 140.1° E) covering the period of 2012 to early 2015.

OMPS-NP profiles, each with a 250 km × 250 km field of view on the ground, were provided from an implementation of the Version 6 SBUV/2 instrument algorithm (Bhartia et al., 1996) with the a priori profiles derived from the OMPS-NM. This version of the OMPS-NP data provide profiles on 12 layers. See Flynn et al. (2104) for a description of the accuracy and precision of the OMPS-NP V6 products. While only the nadir mapper data were assimilated in Section 5, both the nadir mapper and the summed partial columns of the nadir profiler were evaluated during bias correction.

2.4 Independent verification sources

Ground-based and satellite-based column ozone data serve as independent verifications of the OMI-TOMS measurements, with the former also used for validation of the forecasts resulting data assimilation. These data are described below.

2.4.1 Ground-based data

The ground-based data consist mostly of Brewer, Dobson, and filter ozonometer total column ozone measurements (Fioletov et al., 1999 and 2008; Staehelin et al., 2003) from the World Ozone and Ultraviolet Radiation Data Center (WOUDC) and, secondly, of Brewer and Dobson measurements from the Global Monitoring Division of the NOAA Earth System Research Laboratory (see Coldewey-Egbert et al., (2015) for various references on the validation of column ozone satellite data with ground-based Brewers and Dobsons). Only direct sun, clear-sky daily daytime averages from these instruments were used. The error standard deviations for Brewer and Dobson direct sun data are no larger than ~1.5 to 2.0 % for well-calibrated and well-maintained instruments and about 1.5 to 2 times larger for filter ozonometers (Fioletov et al., 1999, and references therein; Fioletov et al., 2008). Consistent with the above, an overall precision of 4.6 DU has been obtained by van der A et al. (2010) for Brewer and Dobson direct sun daily averages, excluding outlier data. As in van der A et al. (2010) and Koukouli et al. (2016), the Dobson ozone values were adjusted following the correction of Komhyr et al. (1993; see also van Roozendael et al., 1998) as a function of ozone effective temperature (-0.13 % K^{-1} about 227 K; The ozone effective temperature is the average value of the ozone-weighted temperature profile.). This correction is not applied to Brewer data in this study following van der A et al. (2010) based on Kerr (2002). The results of Redondas et al. (2014) support neglecting the small sensitivity to ozone effective temperature for Brewer measurements but accounting for the larger sensitivity for Dobson values. Avoiding the correction for Dobsons results in a seasonal dependence of the Dobson total column ozone errors. The calculated seasonal variations of differences of OMI-TOMS and OMI-DOAS with Brewer and Dobson instruments in Balis et al. (2007a) further support neglecting corrections to the Brewer data (if we exclude consideration of results at the equator and in Antarctica which rely only on one station each), while favouring including the corrections for Dobsons. Leaving Dobson measurements uncorrected would introduce a 1-2 % seasonal variation (for example, Bai et al. (2016) showed a fairly consistent seasonal variation of the differences with the Tsukuba Dobson measurements with an amplitude of about 2 % for both OMPS-NM and OMI-TOMS).

2.4.2 SBUV/2

Data from the Solar Backscatter Ultraviolet instrument (SBUV/2) were used for verification purposes. The ozone data from SBUV/2 for the period of interest are from the NOAA 19 satellite (Flynn, 2007; Bhartia et al., 2013; McPeters et al., 2013). Two versions of the total column ozone data are used here: The first is from the SBUV V8.6 profile retrieval using wavelengths in the range of 250 to 310 nm (Bhartia et al., 2013; summarized by McPeters et al., 2013; see also Flynn, 2007) for which the total column ozone is the sum of the partial column layers, and second is from the SBUV V8 total column retrieval using two

wavelengths between 310 and 331 nm (Flynn, 2007; Flynn et al., 2009). The ozone measurements cover 170 km × 170 km field of views at the ground and have separations along the satellite orbit tracks of about 170 km. Labow et al. (2013) found the agreement between total column ozone data of SBUV instruments from the summed partial columns and the Northern Hemisphere ground-based data to be better than 1 %. Bhartia et al. (2013) has indicated that the total column ozone values resulting from the V8.6 algorithm can be used for solar zenith angles up to 88°.

3 Evaluation of OMI-TOMS total column ozone with ground-based data

Differences between OMI-TOMS and ground-based Brewer and Dobsons data have shown long-term term stability and relatively little solar zenith angle and latitude dependence (Balis et al., 2007a; Koukouli et al., 2012; Labow et al., 2013; McPeters et al., 2008 and 2015). Comparisons of OMI-TOMS V8.5 total column ozone with Northern Hemisphere ground-based data by Labow et al. (2013), and McPeters et al. (2015) based on multiple years indicate an average underestimation of OMI-TOMS of about 1.5 %. Figure 2 of McPeters et al. (2015) shows variations of weekly mean differences about the long-term average underestimation mostly within about ±1 %. With OMI-TOMS V8, McPeters et al. (2008) found positive average differences with Northern Hemisphere Brewers and Dobsons covering 2005 and 2006 of 0.4 % with a stations-to-station standard deviation of 0.6 %. Also, OMI-TOMS total column ozone data show no to little dependency on cloud fraction, reflectivity, or cloud top pressure (<1 %; but up to ~2% for cloud top pressure) (Balis et al., 2007b; Antón et al., 2009a; Antón and Loyola, 2011; Koukouli et al., 2012; Bak et al., 2015; Bai et al., 2015). The papers by van der A et al (2010 and 2015) indicate negligible variation with viewing zenith angle. For these reasons, and the near-global daily spatial coverage of its measurements, the OMI-TOMS total column ozone product was selected as the anchor in the applied bias correction schemes described in Section 4.

To further examine the acceptability of using OMI-TOMS as a reference for bias correction, a mean differences comparison of OMI-TOMS V8.5 with near-colocated ground-based data at available sites over the periods of study was conducted. The colocation requirements are the same as those specified in Section 4.1 for the inter-comparison of satellite sensors. Summary results are shown in Table 1 and Fig. 1 (see also Tables S1 to S3). Bimonthly mean differences over regions, globally, and for the individual stations were produced for the three periods of Table 1 based on totals of 53 Brewer, 40 Dobson, and 20 filter ozonometer stations. Fig. 1 shows the station locations and mean differences for the July-August 2014 period. The sizes of the global mean differences over the different periods are in the approximate ranges of 0.0 to -0.1 % for Brewer, -0.2 to 0.4 % for Dobson, and -0.8 to -0.7 % for filter ozonometer instruments. The differences for filters are however confined to upper northern latitudes. These global and regional averages exclude stations with mean differences larger than two standard deviations of the initial mean differences, corresponding to between 3 and 4 %; this outlier removal process was also applied to each station in determining the mean differences at the stations. The total number of outlier stations per time period ranges from 0 to 5 (Tables S1 to S3), some of which are stations at high elevation or in Antarctica. While excluded from contributing to the global averages

of Table 1, the outlier station mean differences were not included as part of the regional mean differences for 60-90° S in Table 1, and the Antarctic outlier stations were retained for further evaluation later in this section.

The global mean differences, and most regional values, are typically smaller than earlier studies mentioned in the first paragraph. Possible contributors to this might be differences in time periods, region specifications, ground-based observation sets, or colocation conditions.

The regional mean differences are within 1 %, with the exceptions being Antarctica for both Brewer and Dobson instruments and the north polar region for Dobson and filter ozonometers instruments. Table 1 shows small positive biases of less than 0.7 % over the region encompassing Canada, the continental United States, and Greenland, as compared to small negative biases of up to -0.4 % over Europe and Northern Africa. The mean differences for the north polar region of -0.3 to -0.6 % for Brewers are under the 1 % target, while the mean differences for Dobsons are -1.2 to -1.6 % and -1.4 to -1.1 % for filter ozonometers. The results for Dobsons and filters are similar despite error levels for the filter instruments being about 1.5 to 2 times larger (Section 2.4.1) and the small datasets. The values over the three seasons are in good agreement despite the small to moderate (≤361) number of colocations. The mean differences covering 2007-2010 from Koukouli et al. (2012) for this region have the same sign, with values of -1.5 % for Brewers and -0.5 % for Dobsons. The average solar zenith angles for stations in the north polar region, while higher than for the middle latitude region, were less than 70° for all instruments and periods except for some Brewer instruments during the January-February 2015 period reaching at most ~76°. Koukouli et al. (2012) determined standard deviations of the differences of 2.4 and 4.3 % for SZA ranges of 25-70° and above 70°, respectively, indicating an increased variability at higher SZAs. Considering the respective periods of this study and of Koukouli et al. (2012), their differences for this region may stem from differences in the range of SZAs. The mean differences for both polar regions are all negative indicating an underestimation of OMI-TOMS column ozone in these regions for these periods relative to ground-based data, which is likely related to high SZAs.

Table 1. Regional and global relative mean differences (%) of total column ozone between OMI-TOMS and the specified ground-based instument types over July-August 2014/2015 and January-February 2015. The averaging excludes stations having outlier station mean diffferences for each period (see Supplement tables S1 to S3 and the text of Section 3) except for the two rows for the latitude region 60-90° S as described in the text. The standard deviations (S.D.) are for the inter-station variation of the mean differences about the regional or global mean differences. Unavailable S.D. values for available mean differences imply the presence of only one station. The Dobson total column ozone measurements for the two July-August periods were adjusted as a function of the ozone effective temperature (see Section 2.4); those for the January-February period were not adjusted in the absence of the ozone effective temperature for the period. The impacts of the Dobson July-August period corrections on the global mean differences were reductions between 0.0 and 0.4 %.

Instrument type	Region	Regional and global mean differences (%) [# of colocations]					
		July-Aug. 2014		July-Aug. 2015		Jan.-Feb. 2015	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
Brewer	Latitude range: 60-90° N	-0.3 [258]	0.8	-0.6 [361]	1.0	-0.6 [9]	1.3
	Latitude range: 30-60° N	0.1 [1773]	1.4	0.6 [1384]	1.6	0.0 [865]	1.1
	Latitude range: 30° S - 30° N	0.4 [296]	1.9	-0.5 [165]	0.7	-0.3 [314]	1.3
	Latitude range: 30-60° S	-	-	0.1 [38]	0.0	0.2 [55]	0.0
	Latitude range: 60-90° S	-5.5 [13]*	-	-	-	-2.5 [152]#	2.0
	North America and Greenland	0.7 [669]	1.1	0.8 [1020]	1.5	0.3 [492]	1.1
	Europe and Africa	-0.3 [1346]	1.4	-0.3 [742]	1.2	-0.5 [454]	1.1
	East Asia and Other	0.6 [312]	1.6	-0.8 [186]	0.9	-0.2 [398]	1.4
	Global	0.1 [2327]	1.4	0.3 [1948]	1.5	-0.1 [1344]	1.2
Dobson	Latitude range: 60-90° N	-1.4 [39]	1.5	-1.2 [29]	0.0	-	-
	Latitude range: 30-60° N	0.3 [331]	0.8	0.6 [301]	1.3	0.8 [167]	1.0
	Latitude range: 30° S - 30° N	-0.3 [240]	2.4	-1.0 [188]	1.3	0.1 [120]	1.4
	Latitude range: 30-60° S	-0.5 [150]	0.9	-1.0 [111]	0.4	-0.0 [136]	1.3
	Latitude range: 60-90° S	-3.3 [6]+	0.1	-4.3 [2]^	-	0.0 [102]$	1.7
	North America and Greenland	-0.5 [125]	0.7	-0.6 [57]	1.1	0.3 [53]	0.5
	Europe and Africa	-0.6 [327]	1.4	0.2 [293]	1.6	0.7 [135]	1.1
	East Asia and Other	0.3 [314]	1.8	-0.6 [279]	1.2	0.1 [337]	1.4
	Global	-0.2 [766]	1.5	-0.2 [629]	1.4	0.3 [525]	1.2
filter ozonometer	Latitude range: 60-90° N	-1.4 [47]	0.8	-1.0 [16]	1.0	-	-
	Latitude range: 30-60° N	-0.3 [54]	1.6	-0.5 [62]	2.0	-0.7 [7]	1.2
	Global	-0.8 [101]	1.4	-0.6 [78]	1.8	-0.7 [7]	1.2

* Outlier mean difference from the Marambio station. # Includes the Amundsen-Scott, Marambio and outlier Zhongshan stations.
+ Includes Marambio and Syowa stations. ^ Outlier Syowa station only. $ Amundsen-Scott, Marambio and Syowa stations.

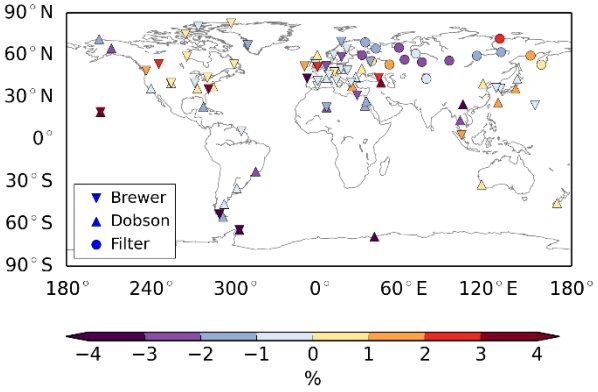

Figure 1. Mean total column ozone differences (%) between OMI-TOMS and Brewer, Dobson, and filter ozonometer measurements over July-August 2014. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

More severe underestimations of OMI-TOMS relative to ground-based observations of 3-6 % occur during July-August in Antarctica, which is associated with SZAs close to or greater than 80° and possibly a strong latitudinal gradient associated to the winter South Pole polar vortex. While the small size of the dataset of 1-3 stations in this region restricts the statistical significance of these results, the level of consistency between the instruments and sites suggest that it is worthwhile to consider this data and so were retained in Table 1 for the rows of the 60-90° S region. The following two paragraphs present reasons that may contribute to either increasing or decreasing the differences in this region.

Bernhard et al. (2005) noted an underestimation potentially exceeding 2 % for SZAs larger than 80°, reaching 4% in the ozone hole region for a SZA of 85°, that could result from the standard Dobson retrieval method assuming the ozone layer being at a specific height (however, adjusting for this would increase the differences with OMI-TOMS). Another factor that would increase the differences with OMI-TOMS for measurements at high solar zenith angles, especially Dobsons, is stray light (e.g., Moeini et al., 2018; Evans et al., 2009), which results in an underestimation of the total column ozone up to at least 5-7 %. The stray light sensitivity also depends on the total column ozone itself, with the effect being smaller under ozone hole conditions than over normal conditions. The Brewer measurements in Antarctica are from double-monochromatic instruments and so only slightly sensitive to stray light as compared to the Dobsons.

At high SZAs, in the vicinity of the polar vortex, the horizontal differences in location between the station and the average of the observed ozone would be sensitive to the strong horizontal gradients in total column ozone. Small differences in observed locations, as well as small differences in solar zenith angles of the colocation pairs at high SZAs, can imply notable differences of observed air masses. For example, approximately accounting for a latitudinal displacement of slightly more than 1° resulted in reducing the July-August 2014 mean difference from the Brewer at Marambio from the -5.9 % in Table 1 to -2.7 %. The discussion of the polar regions and high solar zenith angles is extended in Section 4.1 with a comparison to SBUV/2 total column ozone data.

While the OMI-TOMS data could be underestimating total column ozone in the polar regions for these periods, there is some uncertainty as to the actual OMI-TOMS bias considering factors that could affect the reliability of the comparison with the ground-based data at high solar zenith angles for Antarctica, this even beyond the low number of ground-based observations. van der A et al. (2015) included an adjustment to OMI-TOMS total column ozone data based on the ozone effective temperature in addition to a constant offset of 3.3 DU, which used a comparison to Brewer and adjusted Dobson data. This would increase the OMI-TOMS total column ozone in Antarctica by about 10.5 DU for the two July-August periods. Including the second-order dependence on SZA of that paper would reduce this change to 9 DU. This adjustment would have improved the agreement in the 60-90° S latitude band of Table 1, while adding to the mean differences in the other regions by less than ~1 %, except possibly in the January-February 60-90° N region.

Excluding the uncertainty in quantifying corrections in the south polar region, the low mean differences of the OMI-TOMS V8.5 data with the ground-based data for most regions supports not having to adjust the data before serving as anchor in the bias estimation and correction of the other satellite sources for the limited period covered in this study. While not done here, a correction specifically at high SZAs based on differences for the north polar and the 30-60°S regions could be envisaged.

4 Bias estimation and evaluation using OMI-TOMS as reference

Observation biases can be examined as a function of various factors. In this study, the bias correction applied in the assimilation experiments used bias estimates for discrete SZA/latitude bins as a function of time. Different bias estimation methods based on observation colocations and observation differences with forecasts will be examined. Solar zenith angle dependence is specifically included considering the varying sensitivities between the different instruments as shown in Koukouli et al. (2012). Latitude and time dependences were introduced to capture other data processing biases as well as instrumental changes over time. The alternative method of using the dependence on the ozone effective temperature instead of latitude and time (e.g., van der A et al., 2010) was also explored. Any bias impact due to differences in spatial resolutions of the instruments or model forecasts would be part of the residual biases and associated representativeness errors. Part of the effect of differences in resolution between instruments would be mitigated from bias estimation relying on local averages of differences in space in addition to time. While the dependency on other factors such as cloud cover and viewing zenith angle can vary with the instrument and retrieval algorithm, they are not included here as predictors. Their impact would then be reflected in the estimated standard deviations derived for observations. The bias correction target is to reduce residual biases as a function of SZA and latitude relative to OMI-TOMS generally to within 1 %.

Both July-August and January-February periods are considered for a comparison of bias estimates between seasons within a yearly cycle. The two sets of SBUV/2 total column ozone values obtained from the two wavelengths retrieval (SBUV/2-TC) and the sum of the retrieved partial column profiles (SBUV/2-NP) are included in the comparisons to OMI-TOMS. These have been added to extend the evaluation of the OMI-TOMS data conducted in Section 3.

4.1 Colocation approach

This method estimates the bias as the mean differences of colocated observations with OMI-TOMS. Separate bias estimations are conducted for each distinct instrument-platform. Here, the criteria for observations to be considered to be colocated are for the points to be within 200 km and ±12h, and have solar zenith angle differences smaller than 5° for SZA under 70° and smaller than 2° for SZA between 70° and 90°. The latitude and solar zenith angle bins have a size of 5° each for total column ozone measurements, and 10° each for summed partial column ozone profiles, except for solar zenith angles above 70°, where bin sizes are reduced to 2°. In any case, only data with SZA under 84° are used in assimilation considering the larger uncertainties at higher SZA. The smaller bins at high SZA were chosen since stronger gradients in the differences between instruments arise for these values. The larger bin sizes for summed partial column ozone profiles are in consideration of the smaller density of profile measurements. The resultant bias corrections are assigned to the midpoint of each bin with a two dimensional piecewise linear interpolation applied to points at intermediate SZA and latitude values; data that would require corrections from extrapolation are instead discarded.

Mean differences for each latitude/SZA bin are generated for individual six-hour intervals with, as a precaution, the removal of outliers beyond two standard deviations about the initial mean when there are at least 100 points per bin. Instead of monthly mean bias estimation, a moving window using the previous two weeks of data was applied to better capture variations in time. The six-hour mean differences over the two-week moving window were weighted in time with a Gaussian weighting function with a half width at half maximum of 4.7 days. The six-hour mean differences were generated starting two weeks prior to the start of assimilations to provide data over the full window at the start of the assimilation. Another two standard deviation outlier removal was applied, this time according to the variability of the six-hour mean differences over the two-week period. A minimum of 25 total contributing differences originating from at least four six-hour intervals is imposed for valid bias estimates for each bin.

The time mean differences with OMI-TOMS for July-August 2014 and January-February 2015 are shown in Figs. 2 and 3, respectively. The figures indicate global averaged biases in the range of -3.5 to 2 % (Table 2). The maximum time mean biases per bin reach to sizes of ~5-9 % for some datasets. These mean differences are in general larger than the mean differences of OMI-TOMS with ground-based data. The mean differences typically vary by roughly 3 % over the ranges of bins for SZA values lower than 70°, while larger variations of up to ~7 % can be seen at higher SZA values. The mean differences from SBUV/2 typically vary less between bins as compared to the other instruments. GOME-2A and GOME-2B give the largest and smallest mean differences globally, respectively. The standard errors of the mean differences shown in Figs. 2 and 3 are below 0.1 % for most bins, except for some bins at high solar zenith angles (above 70 °), due to the smaller number colocations, where the maximum standard errors found over all datasets is 0.6 %.

The discontinuity appearing at 70° in SZA for both GOME-2 instruments, as seen in Figs. 2 and 3, may be associated with the switch in the wavelength for reflectivity retrieval between lower and higher SZA from 331.3 nm to 360.1 nm (Table 1.13 from Zhand and Kasheta, 2009). As such, when bias corrections were applied for GOME-2, no interpolation was applied over

the SZA value of 70°. For the DOAS retrieval products, Hao et al. (2014) showed mean differences with Norther Hemisphere ground-based data that varied seasonally between roughly zero and 4 % over the period 2007 to Summer 2013. For the TOMS retrieval products used in this study, the seasonal variation can be seen from Figs. 2 and 3 and Tables 2 and 3, with larger differences for GOME-2A of up to about 3 %. Hao et al. (2014) also showed differences between GOME-2A and GOME-2B of less than 1 % covering December 2012 to November 2013, except in the south pole region and in the Southern Hemisphere for May to September where it reaches at least 2 %. This differs for the TOMS-based GOME-2 retrieval products used here that typically showed larger differences between the two instruments for the times studied.

The pattern about the equator in Fig. 3 (Jan-Feb) appears inverted as compared to Fig. 2 (July-August) for SBUV/2 and OMPS-NP (which can also be seen in Table 3) suggests the possibility of some seasonally dependent differences with OMI-TOMS for these instruments. The results for the provisional OMPS-NM data are smaller than the roughly 2.5 % determined at the Tsukula station for the more recent product version (Bai et al., 2016). However, this is only for a single station. The overall variations in longitude of the mean differences with OMI-TOMS are notably weaker than that in latitude. As such, one would expect the remaining spatially varying residual biases to be small. The percentage of non-empty bins with time mean differences exceeding 2 % in magnitude for the six datasets range from 0 % for SBUV/2-NP to 69 % for GOME-2A (Table 2), both for Jan-Feb 2015. For a 1% threshold, these percentages increase by factors of 1.2 to 4, depending on the instrument and season.

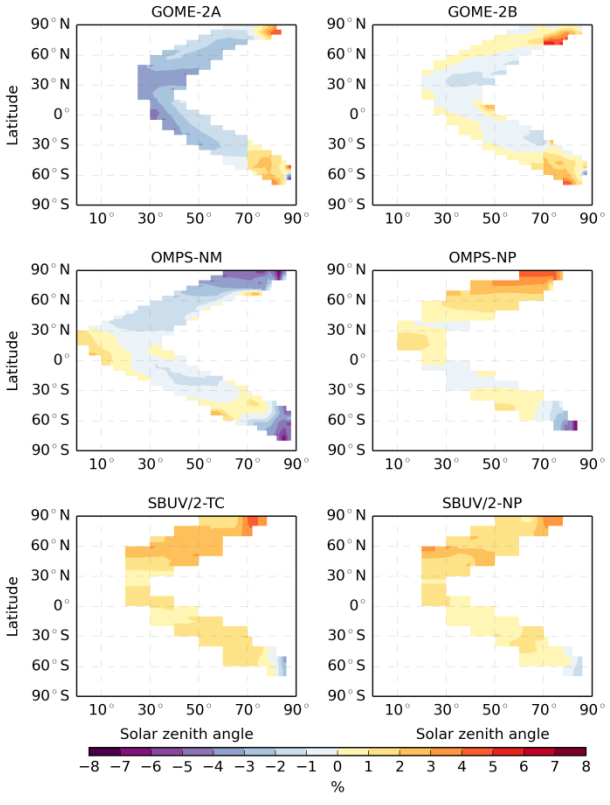

Figure 2. Mean total column ozone differences (%) between GOME-2A/B, OMPS-NM/NP, SBUV/2-TC/NP and colocated OMI-TOMS data for the period of July-August 2014. The SBUV/2-TC total column ozone values stem from the two wavelength retrieval, while those for SBUV/2-NP are the sums of the retrieved 21-layer partial columns. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

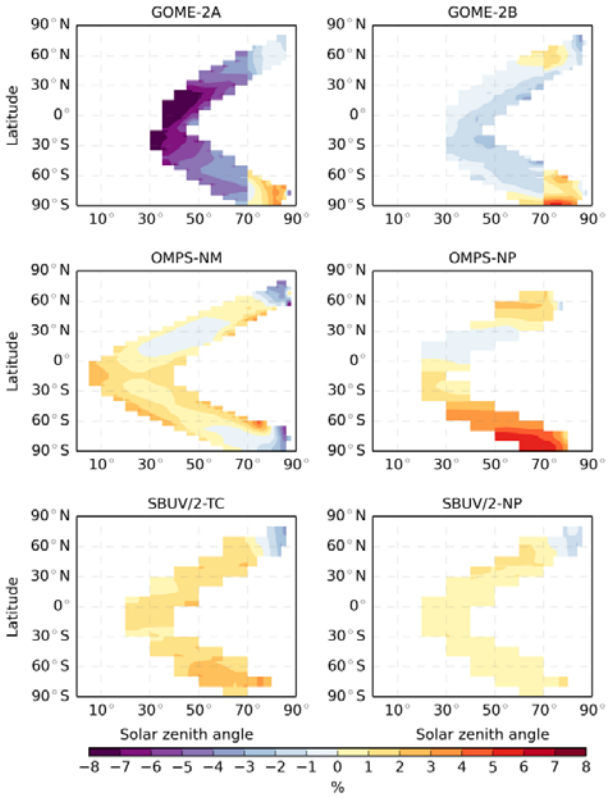

Figure 3. Same as Fig. 2 for January-February 2015.

Table 2. Global diagnostics of differences in total column ozone between satellite instuments and OMI-TOMS for July-August 2014 and January-February 2015. The diagnostics consists of global mean differences and percentages of non-empty SZA/latitude bins with mean differences exceeding 2 % in magnitude.

Instrument	Mean difference (%)		Percentage of bins with |mean differnces| > 2 %.	
	July-Aug. 2014	Jan.-Feb. 2015	July-Aug. 2014	Jan.-Feb. 2015
GOME-2A	-1.8	-3.5	50	69
GOME-2B	0.1	-0.5	14	13
OMPS-NM	-1.3	0.l	28	19
OMPS-NP	1.1	2.0	23	47
SBUV/2-TC	1.5	1.3	30	22
SBUV/2-NP	1.2	0.6	16	0

Table 3. Mean differences of the total column ozone (%) between satellite instuments and OMI-TOMS for July-August 2014 and January-February 2015 for Northern and Southerm Hemispheres, for solar zenith angles below and above 70°.

Instrument	July-Aug 2014				Jan-Feb 2015			
	SZA < 70°		SZA > 70°		SZA < 70°		SZA > 70°	
	NH	SH	NH	SH	NH	SH	NH	SH
GOME-2A	-2.3	-1.8	0.3	1.7	-5.1	-4.5	-1.1	0.9
GOME-2B	-0.1	-0.3	1.3	1.6	-0.7	-1.1	0.4	1.7
OMPS-NM	-1.6	-0.6	-4.1	-1.1	-0.1	0.6	-0.6	-0.6
OMPS-NP	1.5	0.1	3.8	-1.1	0.3	3.1	1.6	4.5
SBUV/2-TC	1.8	1.2	4.1	0.3	1.4	1.6	-0.5	2.8
SBUV/2-NP	1.5	0.7	3.6	0.2	0.8	0.6	-0.5	0.7

The SBUV/2-NP dataset could have been an alternative candidate as the anchor considering the temporal stability in the quality of the data and its level of agreement with ground-based data indicated in earlier studies. The comparisons of the SBUV/2 products with OMI-TOMS in Figs. 2 and 3 and Tables 2 and 3 suggest that OMI-TOMS may be generally closer to the ground-based data for these two periods (Table 1). OMI-TOMS also appears to be in better agreement with SBUV/2 in the Antarctic region than with the ground-based data. The agreement between OMI-TOMS and SBUV/2-NP was usually found to be slightly better than the agreement between OMI-TOMS and SBUV-TC, with the agreement being more notably better in the Jan-Feb 2015 Antarctic region.

The variations in time of the bias corrections for a selected single bin in shown in Fig. 4 for the July-August 2014 period. This figure displays the bin with the latitude, solar zenith angle centered on (52.5°N, 37.5°) for instruments with 5° wide bins and the bin centered on (55°N, 35°) for instruments with 10° wide bins. The time variations for many bins are most often within ±1 % from the time mean, but some bins can vary by ~3 % in time. The variations in time for different instruments can differ not only in size but also in tendency within the short 1-2 month periods. While the resulting moving averages usually change gradually in time, the random variation of the individual six-hour means about the moving averages can be small (within ~1 %) to more significant (reaching at least ~3 %) as can been seen in Fig. 4. The number of colocations per bin for each six-hour interval ranges from a few to a few hundred, while the number of colocations for each two-week moving window typically exceeds a thousand but can be below one hundred for a few bins. As, the number of colocations are significantly reduced for OMPS-NP and SBUV/2 measurements, as it would be for other profilers, the averaging might benefit from longer time windows, a wider Gaussian filter, or larger bin sizes.

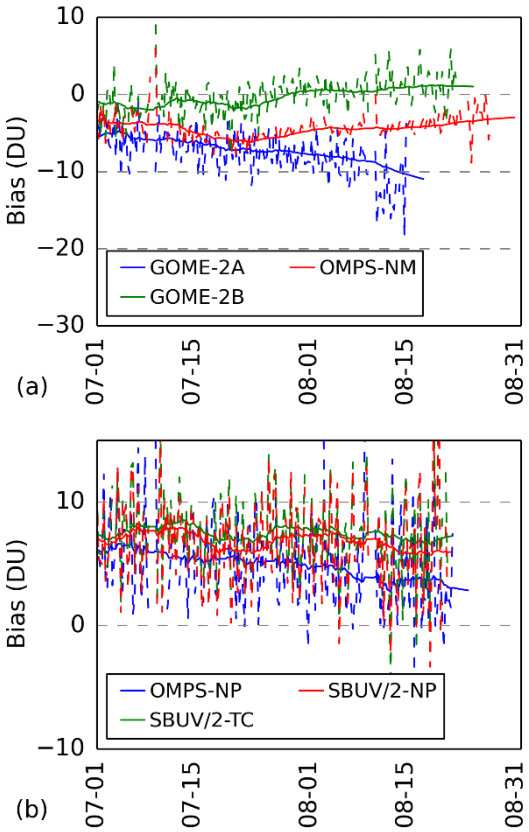

Figure 4. Time series of total column ozone bias corrections (DU) for July and August 2014 for GOME-2A/B, OMPS-NM/NP, and SBUV/2-TC/NP as derived from the colocation method described in Section 4.1. Dashed vertical lines show individual six-hour mean differences with OMI-TOMS, while the solid curves of the same colour show the two-week moving average bias corrections. The particular (latitude, solar zenith angle) bins plotted are 5° wide bins centred on (52.5°N, 37.5°) for GOME-2A/B and OMPS-NM and a 10° wide bin centred on (55°N, 35°) for OMPS-NP and SBUV/2-TC/NP. Time coverage for individual bins do not necessarily cover complete months.

4.2 Bias estimation involving differences with forecasts

An alternative bias estimation approach utilizes the differences of the original retrieved observation data with short-term model forecasts, with the same binning in latitude and solar zenith angle over a two-week moving window. This would be applicable for near-real time or reanalysis data assimilations. These bias estimates can be constructed by considering observation (O) differences with forecasts (F). Bias estimates can be obtained by taking the differences of O-F (innovations in an assimilation context) between an instrument and a reference (O denotes retrieved observations prior to bias correction), which may be done with or without colocation requirements. We identify three different options for this case:

a) $\langle (O-F) - (O-F)_{\mathrm{ref}} \rangle$ with the same colocation requirements as Section 4.1,
b) $\langle O-F \rangle - \langle O-F \rangle_{\mathrm{ref}}$ without the above colocation requirements, or

c) $\langle O - F \rangle$

where the angular brackets denoting averages and the subscript 'ref' denoting differences for observations of the anchor set (OMI-TOMS for our case). Option (a) provides the potential benefit of accounting for spatial differences between paired colocation points, while options (b) and (c) bring the potential advantage of bias correction in the absence of sufficiently close colocation pairs. If previous observations of the reference or other bias corrected instruments were assimilated into the system that produce the short-term forecasts *F*, then option (c) provides a bias correction method for times or locations where the reference is not available. In this work, options (b) and (c) become successive fallback approaches to (a) in the absence of colocated anchor measurements for a bin, with option (b) automatically reducing to option (c) in the absence of the OMI or anchor data. For option (c), innovations would be of more benefit when the forecasts more strongly reflect the influence of the anchor data from previous analyses than that of the model and initial condition errors. In addition, a cutoff criterion for the use of option (c) can be imposed by requiring reference data to have been assimilated within a certain past time period to ensure that these data sets have adequate influence over the forecasts. The same binning and time averaging as done in Section 4.1 are used in this section. As options (b) and (c) are able to use more data than option (a), the extension of (a) to use (b) and (c) as successive fallbacks can increase the number of usable bins in the bias estimation, which would be more evident at high SZAs.

All three of the above options for total column ozone bias estimation were performed and compared to the estimates from Section 4.1. Mean differences with forecasts would normally be determined and applied for bias estimation during the assimilation and forecasting cycle. For convenience, here we instead used the differences with six-hour forecasts from a separate assimilation and forecasting run (the 'OMI' assimilation run summarized in Table 5), which is described in more detail in Section 5. In practice, the forecasts used for this approach, if applied in a near-real time setting, would come from runs that assimilate the bias-corrected observations using the correction method considered in this section. In this section, all observational data sets used for bias estimation are thinned to 1°.

Bias estimates using the options (a) to (c) above for July-August 2014 are shown in Fig. 5, which also shows the colocations only method of Section 4.1 for comparison, and are summarized in Table 4. Differences between the biases resulting from options (a) to (c) and colocation alone are within 1 % over the two-month period except for a few bins, which are mostly at high SZA. The standard errors of the mean differences for all cases are mostly less than 0.1 %, but can as high as 1 % for the options (a) to (c) cases at very high SZA for bins with little data. The time evolution of these bias estimates from the two-week moving window for two different bins is shown in Fig. 6. All bias estimates (both those that do and do not use forecast differences) follow the same general evolution in time, varying within 1 % of one another. The top and bottom panels of Fig. 6 show examples of a bins that have a larger and smaller evolution in time, respectively, where for these bins the bias estimates change by ~10 DU and ~2-3 DU (~3 % and ~1 % for a total column of 300 DU), respectively.

The bias estimates that use differences with forecasts are largely consistent with estimates that use colocation alone. The estimates that utilize differences with forecast can provide additional benefits over using colocations alone if the forecasts well

represent the spatial variation in total column ozone for options (a) and (b), or if the forecasts have been sufficiently de-biased for option (c).

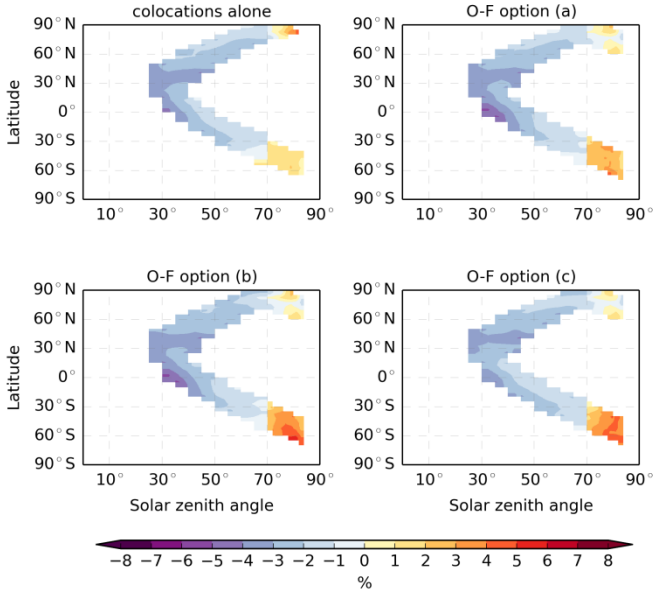

Figure 5. Time mean total column ozone biases (%) between GOME-2A and OMI-TOMS for July-August 2014 from colocation alone and for the options (a), (b), and (c) of Section 4.2 that use observation-minus-forecast differences. For options (a), (b), and (c), the forecasts were taken from the 'OMI' assimilation run (see Table 5). The bias in the 'colocations alone' panel was computed using the thinned observation data set to compare to the other cases that use thinned observations. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Table 4. Mean differences in total column ozone (%) between satellite instuments and OMI-TOMS for July-August 2014 using the options (a), (b), and (c) from Section 4.2, for Northern and Southerm Hemispheres and solar zenith angle below and above 70°.

Instrument	Colocation alone				O-F option (a)				O-f option (b)				O-F opton (c)			
	SZA<70°		SZA>70°		SZA<70°		SZA>70°		SZA<70°		SZA>70°		SZA<70°		SZA > 70°	
	NH	SH	NH	SH	NH	SH	NH	SH	NH	SH	NH	SH	NH	SH	NH	SH
GOME-2A	-2.3	-1.8	0.4	1.7	-7.4	-1.8	-0.1	2.5	-2.6	-1.7	0.0	3.5	-2.3	-1.8	-0.2	3.3
GOME-2B	-0.1	-0.3	1.3	1.6	-0.2	-0.3	1.1	1.5	-0.3	-0.3	1.2	1.9	-0.1	-0.3	1.0	1.8
OMPS-NM	-1.6	-0.6	-4.9	-1.1	1.4	-0.5	-4.7	-1.2	-1.3	-0.4	-4.6	-1.7	-1.3	-0.5	-4.8	-1.9

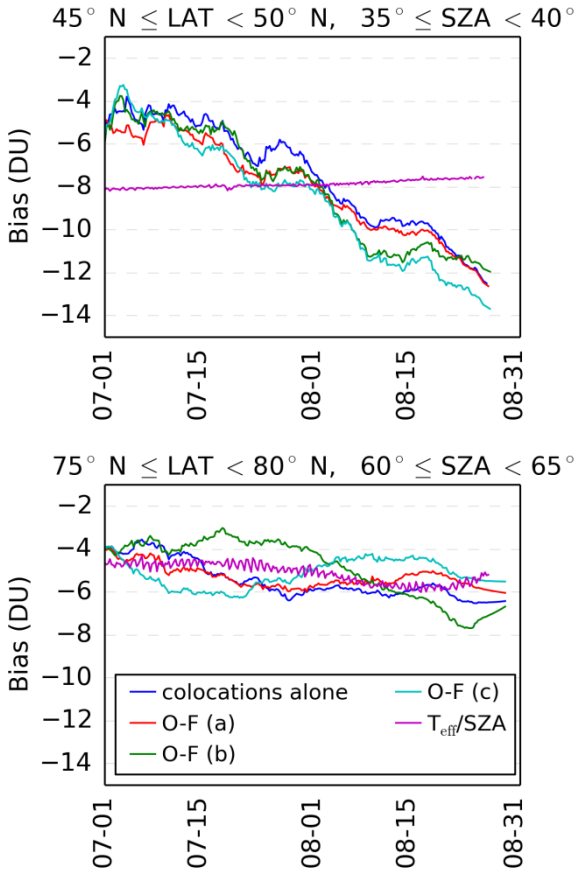

Figure 6. Time series of total column ozone bias corrections (DU) for two latitude/SZA bins covering July-August 2014 for GOME-2A using different bias correction methods. All cases that include colocation methods use thinned observation sets. The 'O-F' curves additionally use the differences of forecasts described in Section 4.2 following the assimilation of OMI-TOMS. The 'colocations alone' and 'O-F' curves were calculuated using the Gaussian two-week moving average with HWHM of 4.7 days. The 'T_{eff}/SZA' curves, described in Section 4.3, result from mapping each observation that falls within the latitude/SZA bin onto the ozone effective temperature/SZA bias estimate for July-August 2014 (shown in Fig. 7), followed by taking the average of these bias estimate values for each time.

4.3 Variation with ozone effective temperature

An alternative parameterization for the bias estimation consists of using ozone effective temperature and solar zenith angle, as done in van der A et al. (2010). A motivation for a dependency on ozone effective temperature is to compensate for any unaccounted temperature sensitivity of the ozone absorption coefficients used in retrievals. In this case, bias estimation is implicitly dependant on time through temporal changes of the ozone effective temperature (and solar zenith angle). This captures at least the seasonal variations of biases associated with changes in temperature in addition to constant offsets. In this section, we briefly consider such a parameterization.

Ozone effective temperatures were calculated from ECCC's GEM meteorological model, with short-term ozone forecasts driven by the LINOZ model and launched from ozone analyses from the assimilation of total column ozone, all of which are described in more detail in Section 5. For these estimates, we return to the methods of Section 4.1, in which mean differences with OMI-TOMS are computed using only colocated observations (i.e. no use of forecasts).

Bias estimates for GOME-2A and OMPS-NM for July-August 2014 and 2015 using an effective temperature parameterization can be seen in Fig. 7. By comparing the bias estimates for the same months from different years, we see that these bias estimates can differ notably for different time periods. With this parameterization, the bias estimate for GOME-2A differs by roughly 3-4 % between 2014 and 2015 for SZAs less than 70°. These differences are larger than the long term trends of about -2.2 DU, or roughly -0.6 to -0.8 %, per year estimated by van der A et al. (2010) for GOME-2A (DOAS), although we note that all GOME-2 data used in this study were retrieved using the TOMS method. Differences in retrievals methods and time periods might be factors in explaining these differences. For both time periods shown in Fig. 7, applying their respective corrections result in time averaged residual biases as a function of latitude and solar zenith angle typically within 1 %, with only a few bins over 2 %.

An equivalent time evolution of a latitude/SZA bin can be made from the time-averaged effective temperature/SZA bias estimate shown in Fig. 7: First, the ozone effective temperature of each observation falling within a selected latitude/SZA bin is used to map that observation onto the ozone effective temperature/SZA bias estimate (Fig. 7), then the bias estimate at at each observed ozone effective temperature/SZA point is averaged for each six-hour time period. The resulting curves are shown Fig. 6 for the latitude/SZA bins selected. The small temporal evolutions of these curves (typically well within 1 %) reflects the slight changes in the ozone effective temperature/latitude relationship in time. The greater the variation in time of the bias estimates based on the time varying latitude/SZA parameterization, the larger the differences with the estimates based on the time independent temperature and SZA parameterization (an example of which is illustrated by comparing the top and bottom panels of Fig. 6). Adding an explicit sub-seasonal to seasonal dependency on time to the ozone effective temperature/SZA bias estimate would compensate for these otherwise unaccounted for time variations. Overall, this supports the use of an ozone effective temperature parameterization as an alternative to latitude (and time) parameterization, with the stipulation that one accounts for any remaining notable temporal changes in some fashion when necessary.

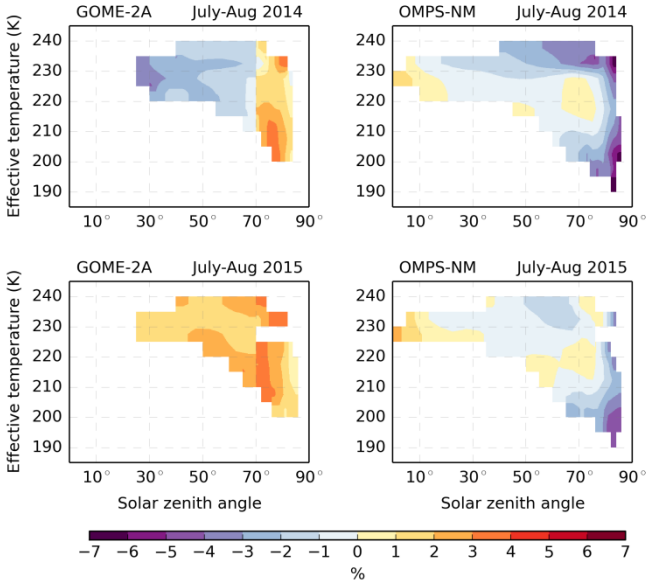

Figure 7. Mean total column ozone differences (%) between GOME-2A, OMPS-NM and colocated OMI-TOMS data as a function of ozone effective temperature (degrees Kelvin) and solar zenith angle (degrees) for the periods of July-August 2014 and July-August 2015. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

5 Assimilation system and results

In this section, we examine the effects of bias correction on global ozone assimilation and compare the six-hour forecasts launched from these analyses to ground-based observations and to OMI-TOMS. Corrections of observation biases were updated every six-hours using a two-week moving window from colocations with OMI-TOMS. Assimilation experiments were conducted for July-August 2014, with a start date of 28 June 2014, 18 UTC, with and without bias correction. All bias corrected observations used in assimilation used the colocation approach without use of forecast differences (Section 4.1) to obtain bias estimates.

The forecasting model used was the Global Environmental Multiscale (GEM) numerical weather prediction model (Côté et al., 1998a and 1998b; Charron et al., 2012; Zadra et al., 2014a,b; Girard et al., 2014) of Environment and Climate Change Canada coupled to a linearized ozone model (LINOZ) (McLinden et al., 2000; de Grandpré et al., 2016). The LINOZ model uses pre-computed coefficients generated as monthly mean climatologies for calculating the ozone production and sink contributions throughout the stratosphere and upper troposphere down to 400 hPa. A relaxation towards the climatology of Fortuin and Kelder (1998) was imposed between the surface and 400 hPa to constrain deviations away from the climatology, with a relaxation time scale of 2 days. The GEM model was executed with a 7.5 min time step with a uniform 1024×800 longitude-latitude grid and a Charney-Phillips vertically staggered grid (Charney and Phillips, 1953; Girard et al., 2014) with 80 thermodynamic levels extending from the surface to 0.1 hPa. The horizontal grid corresponds to a resolution of ~0.23° in

latitude and ~0.37° in longitude, representing a 25 km resolution at latitude 49°. In assimilation, inconsistencies stemming from the differences in resolutions between the model forecasts and the observations would usually be reflected by some corresponding increase of applied observation error variances. This is not explicitly done here. The vertical resolution in the upper-troposphere/lower-stratosphere (UTLS) region is in the range of 0.3 to 0.6 km with the resolution gradually changing to ~1.6 km at 10 hPa and 3 km at 1 hPa.

Assimilation was done using an incremental three-dimensional variational (3D-Var) approach with first guess at appropriate time (FGAT; Fisher and Andersson, 2001). This assimilation system uses components of the ECCC Ensemble-Variational data assimilation system (Buehner et al., 2013 and 2015) adapted by the authors and P. Du (ECCC) for constituent assimilation and being run without ensembles. Successive short-term three to nine hour forecasts were generated from analyses provided for 00, 06, 12, 18 UTC synoptic times. The analyses are a composite of the already available ECCC operational meteorological analysis and the ozone analyses generated from this assimilations study. Assimilation runs were compared to runs without ozone assimilation but that used the same meteorological analyses as employed by the ozone assimilation runs. The initial ozone field used was an analysis from an earlier assimilation.

The background error covariances used have latitude varying error standard deviations with values at sample vertical levels of 1, 10, 50, and 300 hPa in the ranges of ~6-12 %, ~3-5 %, ~5-15 %, and ~15-24 %, respectively. The vertical correlations have half width at half maximum values between 0.5 and 1 km between the top of the boundary layer and 100 hPa, and are nearly equal the model vertical resolutions above 100 hPa with values ranging from ~0.5 km at 100 hPa to ~3.5 km at 1 hPa. The horizontal correlation half widths at half maximum are ~125 km near the surface and increase from ~165 km at 100 hPa to just under 750 km at 1 hPa. The applied observation error standard deviations assigned to all total column measurements of all sources for the conducted assimilations were set to a constant of 2 %.

As assimilating column ozone data from two or more sources ensures that data is continually available in the event of occasional to permanent interruption of data availability from specific instruments, both individual and combined observation datasets were assimilated. For near–real time assimilation, the interruption of the availability of the anchor dataset implies the need for contingency planning for transitions of bias correction references. One might opt to assimilate data from some sensors and monitor the data from others through comparisons with the assimilation analyses. While not necessarily negating the need for bias correction, one could always select to assimilate data from sensors with retrieval products having the smallest initial biases as compared to other products. The effects of bias correction on assimilation when assimilating both individual and multiple sensors will be examined.

The applied evaluation metrics consist of mean differences, standard deviations, and anomaly correlation coefficients (ACC), i.e.,

$$\text{Mean differences, } m(O-F) = N^{-1}\sum_{i=1}^{N}(O_i - F_i) \quad (1)$$

$$\text{Standard deviations, } \sigma(O-F) = \sqrt{(N-1)^{-1}\sum_{i=1}^{N}[(O_i - F_i) - m(O-F)]^2} \quad (2)$$

$$\text{Anomaly correlation coefficients, } ACC = \frac{(N-1)^{-1}\sum_{i=1}^{N}[(O_i-C_i)-m(O-C)][(F_i-C_i)-m(F-C)]}{\sigma(O-C)\sigma(F-C)} = \frac{\text{cov}(O-C,F-C)}{\sigma(O-C)\sigma(F-C)} \quad (3)$$

where O_i, F_i, and C_i denote observations, forecasts, and climatological values at the observation locations, respectively. The ACC (e.g. WMO, 1992) provides a measure of the spatio-temporal correlation between the deviations of forecasts and a verifying dataset (observations or analyses) from a reference (often a climatological field). For this study, the mean forecast values for the no assimilation case over July-August 2014 were used the reference *C* instead of a climatology. It was verified that choosing the reference in the ACC as the 2D ozone climatology of Fortuin and Kelder (1998) instead does not significantly change the results. As anomaly correlation coefficients in assimilation typically compare forecasts with analyses instead of observations, OMI data in this case, it was also verified that both give similar results. In the tables and legends of the figures referred to in this section (Table 5 and Figs. 8 and 9), the short labels that denote the different assimilation runs are described in Table 5.

We first examine the global differences of Brewer and Dobson total column ozone measurements with six-hour forecasts following assimilation with and without bias correction. The mean and standard deviations of these differences are shown in Table 6. Note that assimilating GOME-2A observations alone without bias correction actually increases the absolute size of the global mean differences relative to the no assimilation case to over 2 %. The smaller value for the no assimilation case stems specifically from the cancellation of larger positive and negative mean differences in the tropical and extra-tropical regions, respectively (Fig. 8). Runs assimilating GOME-2A and OMPS-NM alone, as well as GOME-2A/B and OMPS-NM, have the global mean biases from both Brewers and Dobsons reduced from above to well below 1% when bias correction is introduced. Bias correction reduces the global mean differences to less than 0.3 % in size for all cases. For the south polar region, the inclusion of bias correction in the assimilation of OMPS-NM reduced mean difference from ~4-5 % to ~1-2 % (less reduction is seen for GOME-2 since it does not reach as far south). Introducing assimilation reduces the standard deviations from 3.4-3.8 % to ~2.3-2.9 %, while bias correction further reduces the standard deviations modestly to ~2.3-2.6 %. The standard deviations obtained from the assimilation of uncorrected observations incorporates the effect of the latitude and SZA variation of the biases of the different instruments. This contribution would be reduced when assimilating bias corrected observations. The small reductions in standard deviations from introducing bias correction indicate that the effect from the reduction of bias variability on the variances is small as compared to the sum of the other variance contributions. These contributions could include the variation of inter-station ground-based instrument calibration errors and/or representativeness errors associated to the model resolution, in addition to the forecast errors and the ground-based instruments random errors.

Table 5. List of assimilation experiments and their corresponding identifiers. In the second column, an asterisk (*) next to the instrument denotes that the bias-corrected observations (using the colocation method of Section 4.1) were assimilated.

Assimilation experiment identifier	Instruments assimilated
CTRL	None
OMI	OMI
GOME2A	GOME-2A
GOME2B	GOME-2B
OMPSNM	OMPS-NM
G2AB+NM	GOME-2A/B, OMPS-NM
ALLTC	GOME-2A/B, OMPS-NM, OMI
GOME2A bc	GOME-2A*
GOME2B bc	GOME-2B*
OMPSNM bc	OMPS-NM*
G2AB+NM bc	GOME-2A*/B*, OMPS-NM*
ALLTC bc	GOME-2A*/B*, OMPS-NM*, OMI

*denotes bias-corrected observations

Table 6. Global mean differences (%) between Brewer and Dobson total column ozone measurements and short-term forecasts for July-August 2014. Bias-corrected observations from the colocated observation bias correction scheme (Section 4.1) were applied in the assimilations. The Dobson measurements used were adjusted as a function of the ozone effective temperature (see Section 2.4). The uncertainties denote the standard error of the mean differences or standard deviations of the differences. The data from the two Antarctic stations have been included here even though their mean differences with OMI are outliers relative to most mean differences (Tables S1 and S2).

	Assimilated instruments	Mean difference (%)		Difference std. dev. (%)	
		No bias correction	Bias correction	No bias correction	Bias correction
Brewers	None	-1.73 ± 0.08	-	3.85 ± 0.05	-
	OMI	-0.03 ± 0.05	-	2.34 ± 0.03	-
	GOME-2A	2.33 ± 0.05	0.13 ± 0.05	2.62 ± 0.04	2.45 ± 0.03
	GOME-2B	0.19 ± 0.05	-0.07 ± 0.05	2.43 ± 0.03	2.36 ± 0.03
	OMPS-NM	1.22 ± 0.05	-0.14 ± 0.05	2.59 ± 0.04	2.44 ± 0.03
	GOME-2A/B + OMPS-NM	1.20 ± 0.05	-0.02 ± 0.05	2.51 ± 0.03	2.36 ± 0.03
	GOME-2A/B + OMPS-NM + OMI	0.89 ± 0.05	0.01 ± 0.05	2.49 ± 0.03	2.33 ± 0.03
Dobsons	None	-0.91 ± 0.12	-	3.43 ± 0.08	-
	OMI	0.20 ± 0.08	-	2.36 ± 0.05	-
	GOME-2A	2.22 ± 0.10	0.20 ± 0.08	2.94 ± 0.07	2.59 ± 0.06
	GOME-2B	0.47 ± 0.08	0.03 ± 0.08	2.54 ± 0.06	2.44 ± 0.05
	OMPS-NM	1.30 ± 0.08	0.27 ± 0.08	2.45 ± 0.06	2.43 ± 0.05
	GOME-2A/B + OMPS-NM	1.23 ± 0.08	0.14 ± 0.07	2.51 ± 0.06	2.36 ± 0.05
	GOME-2A/B + OMPS-NM + OMI	0.97 ± 0.08	0.17 ± 0.07	2.46 ± 0.06	2.32 ± 0.05

Comparisons of OMI-TOMS measurements with forecasts for the various experiments with and without bias correction and without any assimilation are shown in Figs. 8 and 9 for the July-August 2014 period. For assimilation of only GOME2-A, in most of the topics and northern extra-tropics, the reduction of the mean differences from assimilation without bias correction as compared to the no assimilation case is roughly the same order of magnitude as the reductions resulting from introducing bias correction as compared to the no bias correction case. However, assimilation of the other instruments show that the first

order improvements stem from assimilation in general, while bias corrections result in second order changes. Both the temporally averaged and time varying mean differences of forecasts with OMI-TOMS were reduced to within 1 % over the latitude ranges where satellite data are assimilated for the all cases with bias correction, with the results for GOME-2A only assimilation being the exception, slightly exceeding 1 % in some places. The GOME-2A and OMPS-NM datasets show the largest reductions in mean differences from bias correction, as would be expected from Fig. 2, where these biases are reduced from levels of ~1-3 % when no bias correction is performed to well within 1 % for bias correction cases (excluding latitudes below 60°S). The assimilation of bias corrected observations from multiple sensors (labelled as 'ALLTC bc') does not notably reduce the mean differences as compared to the assimilation of individual bias corrected sensors. Considering the earlier comparisons of forecasts with ground-based data and these results, the reduction of biases to the 1 % target appears to be achieved for the short-term forecasts in most regions with assimilated data.

Assimilation of total column observations improves the standard deviations of differences between the six-hour forecasts and OMI-TOMS across all latitudes, as seen in Fig. 8, although relatively little impact is seen for the GOME-2A/B assimilations in the southern extra-tropics where relatively few observations are available. The impact of bias correction on standard deviations of forecast is not very significant. The large mean differences and standard deviations for GOME-2A/B assimilations below 60°S stem from these datasets not reaching much further south during this period. This reflects the importance of observations near the winter poles in the absence of heterogeneous chemistry in LINOZ.

In the absence of assimilation, the mean differences between the forecasts and OMI-TOMS observations in the extra-tropics have opposite sign to those in the tropics, as seen in the top panels of Fig. 8. Also, notice that in Fig. 9 the mean differences in the extra-tropics diverge in the opposite direction as compared to the tropics. The drift of the mean biases in time in the absence of assimilation are due to the tendency of the forecast to move toward the ozone model equilibrium state. For the GEM-LINOZ model, this results in a long spin-up period in which ozone field moves from the initial ozone field, based on an earlier assimilation, toward the ozone model equilibrium state. Beginning with an initial ozone field at the model equilibrium state would have increased its mean observation minus forecast differences and would likely not have improved the ACC of the control case, as implied by Fig. 9. Also from Fig. 9, we can see that the error of the total column ozone forecast increases by less than 5 % over the course of fifteen days, reflecting the high predictability of ozone medium range forecasts. This limited deterioration would not deter, for example, in properly forecasting the movement of low column ozone regions during these periods and the corresponding changes in clear-sky UV Index.

For the ACC, forecasts from the assimilation of GOME-2B in the tropics appear better than from the assimilation of OMI-TOMS when compared to the OMI-TOMS observations. This is likely due to the larger volume of GOME-2B data (when comparing the thinned dataset from spatial sampling at 1° resolution), in addition to its low bias. Furthermore, the ACC demonstrates a more marked improvement in multiple sensor assimilation in the tropical region as compared to OMI-TOMS assimilation alone, which is not seen in the mean differences. The advantage of multiple sensor assimilation is, therefore, more notable in increasing the quality of the pattern and variation of the forecast fields.

The deterioration of the ACC with time in Fig. 9, as well as the low time mean ACC in the tropics in Fig. 8, in the absence of assimilation reflects an increase in the spatio-temporal variations of the observation-minus-forecast differences as compared to the cases with assimilation. To examine this further, we can rewrite the expression for the anomaly correlation coefficient in observation space as

$$ACC = \frac{\mathrm{cov}(O-C,F-C)}{\sigma(O-C)\sigma(F-C)} = \frac{1}{2}\,\frac{\left[\sigma^2(O-C)+\sigma^2(F-C)-\sigma^2(O-F)\right]}{\sigma(O-C)\sigma(F-C)} \quad (4)$$

where σ are the standard deviations of the quantity in its brackets. As shown in Fig. 10, when assimilation is not performed, during the time period when the ACC deteriorates rapidly in the tropics, *σ(O-C)* and *σ(F-C)* do not change substantially (roughly at 14 DU and 9 DU, respectively), while *σ(O-F)* increases from about 10 to 20 DU, illustrating the temporal deterioration in the tropics from the model. Similar increases in *σ(O-F)* are seen in the other regions as well for the no assimilation case. Introducing assimilation rapidly and substantially reduces the values of *σ(O-F)* to around 5-7 DU while pushing the values of *σ(F-C)* up closer to that of *σ(O-C)*, so that the first two terms in Eqn. 4 are roughly the same size and much larger than the third term in regions where measurements are assimilated. This results in ACC values notably closer to unity.

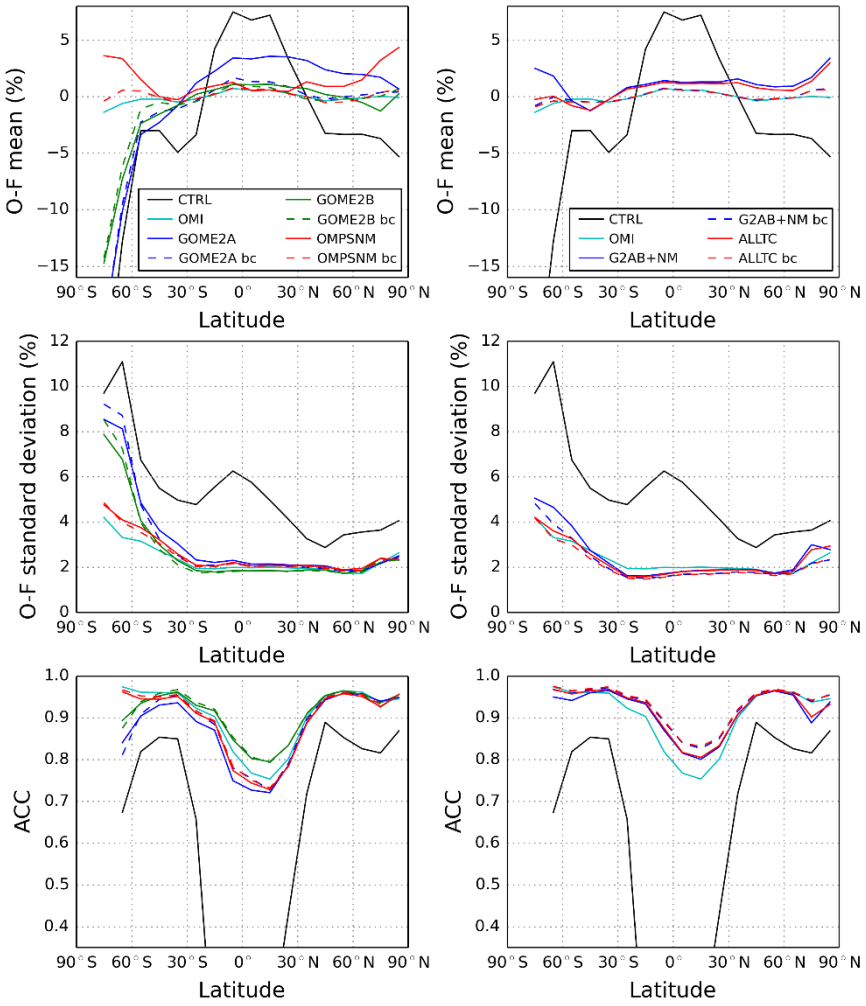

Figure 8. Zonal mean total column ozone statistics of mean differences (%), standard deviations (%), and anomaly correlation coefficients (ACC; unitless) as a function of latitude (degrees) for the comparison between OMI-TOMS measurements and short-term forecasts for July-August 2014. The legends in the top plots indicate the assimilation runs (see Table 6 for description) and apply to all plots in the same column.

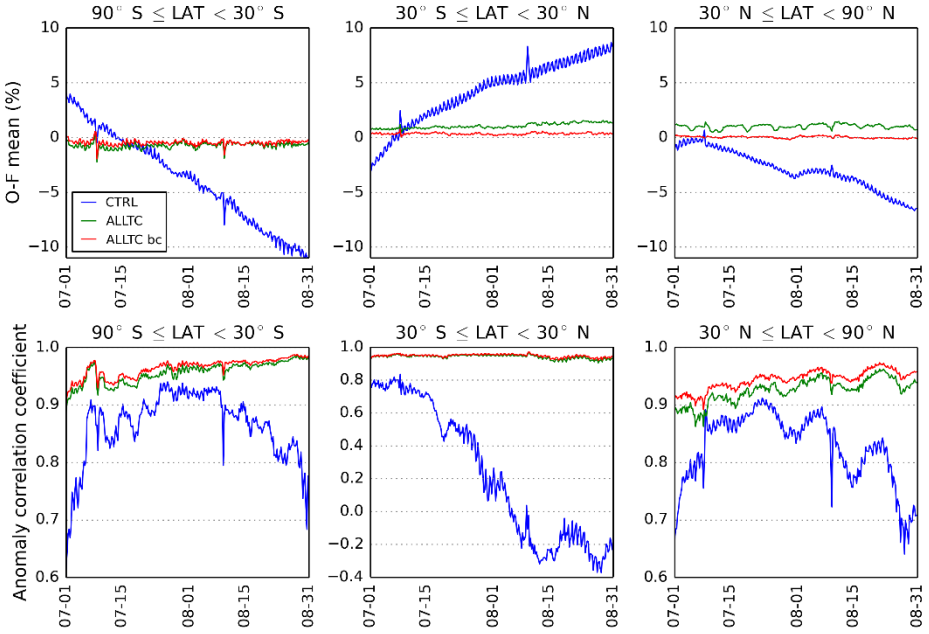

Figure 9. Zonal mean differences (%) and anomaly correlation coefficients (unitless) for total coumn ozone between OMI-TOMS observations and short-term forecasts as a function of time (date). Results are shown for the case without assimilation as well as with the assimilation of OMI, GOME-2A/B, and OMPS-NM (both with and without bias correction). The legend indicates the assimilation run (see Table 6 for description). Each value plotted was calculated using a 24 hour time window.

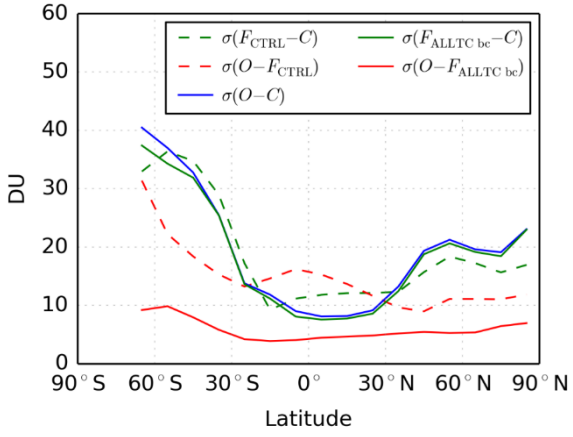

Figure 10. Zonal mean standard derivations (DU) of the differences between short-term forecasts *F*, OMI-TOMS observations *O*, and climatological values *C* over July-August 2014. For the short-term forecasts, the assimilation run from which the forecast was launched is indicated in the subscript, with the labels described in Table 5.

6 Conclusions

Bias correction of total column ozone data from satellite instruments was performed using three different approaches. Two of the methods parameterized the bias estimation as a function of latitude, solar zenith angle binning, and time, while the remaining method used the ozone effective temperature instead in place of latitude and time. These approaches consisted of using observation colocation between satellite-borne instruments and a reference, referred to in this paper as the anchor. Different variants of the bias estimation scheme were explored, including examining the effect of including short-term forecasts within the estimation. Differences between bias estimations from different methods that used the latitude/solar zenith angle parameterization were generally within 1 %. While the two month time-averaged bias estimates from the ozone effective temperature parameterization were similar to those from the other approaches, the lack of an explicit time dependence caused departures on shorter time scales between these estimates and those from methods that include an explicit time dependence, where these estimates were different by ~2-3 % for some instruments.

The anchor used in the bias estimation schemes was chosen as the OMI-TOMS data product, due to its wide coverage in both time and space, and its good agreement with ground-based instruments. In this study, for the time periods examined, OMI-TOMS was found to have global and regional mean differences with ground-based Brewer and Dobson spectrophotometers, and filter ozonometers within 1 %, except in the polar regions. Similar to larger mean differences of OMI-TOMS with SBUV/2 data were found, with OMI-TOMS generally being in better agreement to the ground-based data for the examined periods.

For the July-August 2014 and January-February 2015 periods, the observations based on TOMS retrievals for the GOME-2A instrument were found to have the largest mean differences with OMI-TOMS, which could be as high as 8 % in some regions of the parameter space for solar zenith angles below 70°. The GOME-2B instrument showed much better agreement with OMI-TOMS, with mean differences generally confined to ~1-2 %, excluding at very high solar zenith angles. The provisional OMPS ozone column products, both the total column and summed partial column profile, typically had mean differences somewhere between the two GOME-2 instruments, with mean differences generally confined to ~3-4 % (again excluding high solar zenith angle regions). As the quality of the different versions of OMPS retrieved data may differ, one might expect a reduction in bias of the more recent version of the OMPS products based on the SBUV V8.6 retrieval algorithms.

It was demonstrated that the assimilation of total column ozone observations that include bias corrections as derived in this study can improve the agreement between short-term forecasts and ground-based measurements. Using a three-dimensional variational assimilation system, the assimilation of GOME-A without bias correction gives global and time mean differences between ground-based observations and short-term ozone forecasts of ~2.3 %. The assimilation of uncorrected OMPS-NM measurements reduced these mean differences slightly to ~1.3 %. Assimilating instead the bias corrected observations brought these mean differences to well within 1 %. As minimal bias was found for GOME-2B, the assimilation of both corrected and

uncorrected GOME-2B observations yielded mean differences within 1 %. The benefit of including total column satellite data, even without bias correction, was most notable in the tropics, in addition to the polar vortex region.

The aforementioned results indicate that the reduction of biases to the 1 % target was achieved for most regions and cases, a likely exception being for conditions with high solar zenith angles. For the assimilation of two of more satellite sensors, while it is possible that the cancellation of errors from different instruments could reduce forecast biases, harmonizing the different datasets through bias correction better ensures target reductions in residual biases are achieved. The assimilation of bias corrected observations from multiple sensors does not notably reduce the mean differences as compared to the assimilation of individual bias corrected sensors. However, a notable improvement in multiple sensor assimilation was seen in the tropical region as compared to OMI-TOMS assimilation alone with the anomaly correlation coefficients metric. This improvement implies an increase the quality of the pattern and variation of the forecast fields.

Code and data availability. The bias estimation and correction software with related shell scripts can be provided with the understanding that users will need to adapt the code to their preferred input/output data file formats. The observations can be obtained from the different centres identified in the text and the acknowledgments Section below. The assimilation and forecasting system relies on ECCC computing environment tools and file conventions. As well, the computing hardware used for these assimilation cycles has since been replaced at ECCC with accompanying changes to the cycling package. References of the system components are provided in this paper. The large sets of model analyses and forecasts, and the observation minus forecast datasets, are saved with an in-house binary file format. Subsets could potentially be made available from the authors upon request. In addition to also containing a few complementary figures, the Supplement provides tables of station by station mean differences of OMI-TOMS with ground-based data related to Table 1 and Fig. 1.

The Supplement related to this article is available online at https://doi.org/XXXX-supplement.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments. The authors wish to thank Lawrence E. Flynn (NOAA) and Vitali Fioletov (ECCC) for information and advice regarding use of the different satellite datasets and the Brewer ground-based data, respectively, Jean de Grandpré and Irena Ivanova (ECCC) for the availability and assistance in use of the version of GEM with LINOZ, Ping Du and Mark Buehner (ECCC) for contributions in extending the variational assimilation code for use with constituents, Jose Garcia (ECC) and Vaishali Kapoor (NOAA), among others associated to NOAA for contributions, in facilitating data acquisition and the various instrument teams having generated the different observation sets. We also gratefully acknowledge the following centres for access to the observations used in this paper: the National Environmental Satellite, Data, and Information Service of the National Oceanic and Atmospheric Administration (NESDIS/NOAA), the Earth Observing System Data and Information

System of the National Aeronautics and Space Administration (EOSDIS/NASA), the World Ozone and Ultraviolet Radiation Data Center (WOUDC), and the Global Monitoring Division of the NOAA Earth System Research Laboratory.

References

Andersson, E., and Järvinen, H.: Variational quality control. Q. J. R. Meteorol. Soc., 125, 697–722. doi: 10.1002/qj.49712555416, 1999.

Antón, M., López, M., Vilaplana, J. M., Kroon, M., McPeters, R, Bañón, M., and Serrano, A.: Validation of OMI-TOMS and OMI-DOAS total ozone column using five Brewer spectroradiometers at the Iberian peninsula, J. Geophys., Res., 117, D14307, doi:10.1029/2009JD012003, 10 pp., 2009a.

Antón, M., Loyola, D., López, M., Vilaplana, J. M., Bañón, M., Zimmer, W. and Serrano, A.: Comparison of GOME-2/MetOp total ozone data with Brewer spectroradiometer data over the Iberian Peninsula, Ann. Geophys., 27, 1377-1386, 2009b.

Antón, M., Loyola, D., Clerbaux, C., López, M., Vilaplana, J.M., Bañón, M., Hadji-Lazaro, J., Valks, P., Hao, N., Zimmer, W., Coheur, P. F., hurtmans, D., and Alados-Arboledas, L.: Validation of the Metop-A total ozone data from GOME-2 and IASI using reference ground-based measurements at the Iberian Peninsula, Remote Sens. Environ., 115, 1380-1386, 2011.

Bai, K., Chang, N. B., Yu, H. and Gao, W.: Statistical bias correction for creating coherent total ozone record from OMI and OMPS observations, Remote Sens. Environ., 182, 150-168, 2016.

Bak, J., Liu. X., Kim, J. H., Chance, K., and D. P. Haffner, Validation of OMI total ozone retrievals from the SAO ozone profile algorithm and three operational algorithms with Brewer measurements, Atmos. Chem. Phys., 15, 667–683, doi:10.5194/acp-15-667-2015, 2015

Balis, D., Kroon, M., Koukouli, M. E., Brinksma, E. J., Labow, G., Veefkind, J. P., and McPeters, R. D.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, J. Geophys. Res., 112(D24), 2007a.

Balis, D., Lambert, J.-C., Van Roozendael, M., Spurr, R., Loyola, D., Livschitz, Y., Valks, P., Amiridis, V., Gerard, P., Granville, J., and Zehner, C.: Ten years of GOME/ERS2 total ozone data—The new GOME data processor (GDP) version 4: 2. Ground-based validation and comparisons with TOMS V7/V8, J. Geophys. Res., 112(D7), 2007b.

Bhartia, P. K., McPeters, R. D., Flynn, L. E., Taylor, S., Kramarova, N. A., Frith, S., Fisher, B. and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm. Atmos. Meas. Tech., 6(10), 2533–2548. doi:10.5194/amt-6-2533-2013, 2013

Bhartia, P. K. and Wellemeyer, C. W.: TOMS-V8 total O3 Algorithm, Chapter 2 in OMI Algorithm Theoretical Basis Document, Volume2, OMI Ozone Products, edited by P. K. Bhartia, 15-31, 2002. [Available from http://projects.knmi.nl/omi/documents/data/OMI_ATBD_Volume_2_V2.pdf]

Bhartia, P. K., McPeters, D., Mateer, C. L,. Flynn, L. E., and Wellemeyer, C.: Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique, J. Geophys. Res., 101, 18,793-18,806, 1996.

Bernhard, G., Evans, R. D., Labow, G. J. and Oltmans, S. J.: Bias in Dobson total ozone measurements at high latitudes due to approximations in calculations of ozone absorption coefficients and air mass, J. Geophys. Res., 110, D10305, doi:10.1029/2004JD005559, 2005.

Buehner, M., Morneau, J. and Charette, C.: Four-dimensional ensemble–variational data assimilation for global deterministic weather prediction. Nonlinear Processes Geophys., 20, 669–682, doi:10.5194/npg-20-669-2013, 2013.

Buehner, M., McTaggart-Cowan, R., Beaulne, A., Charette, C., Garand, L., Heilliette, S., Lapalme, E., Laroche, S., Macpherson, S .R., Morneau, J., and Zadra, A.: Implementation of Deterministic Weather Forecasting Systems based on Ensemble-Variational Data Assimilation at Environment Canada. Part I: The Global System Mon. Wea. Rev., 143, 2532-2559. doi: 10.1175/MWR-D-14-00354.1, 2015.

Cariolle, D. and Déqué, M.: Southern hemisphere medium-scale waves and total ozone disturbances in a spectral general circulation model. J. Geophys. Res., 91D, 10825–10846, 1986.

Cariolle, D. and Teyssèdre, H.: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations. Atmos. Chem. and Phys. Disc., 7, 1655–1697, 2007.

Charney, J. C. and Phillips, N. A.: Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows. J. Meteor., 10, 17–29, 1953.

Charron, M., Polavarapu, S., Buehner, M., Vaillancourt, P. A., Charette, C., Roch, M., Morneau, J., Garand, L., Aparacio, J. M., MacPherson, S., Pellerin, S., St-James, J., and Heilliette, S.: The stratospheric extension of the Canadian Global Deterministic Medium-Range Weather Forecasting System and its impact on tropospheric forecasts, Mon. Wea. Rev., 140, 1924–1944, doi:10.1175/MWR-D-11-00097.1, 2012.

Coldewey-Egbers, M., Loyola, D. G., Koukouli, M., Balis, D., Lambert, J.-C., Verhoelst, T., Granville, J., Van Roozendael, M., Lerot, C., Spurr, R., Frith, S., and Zehner, C.: The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative, Atmos. Meas. Tech. Discuss., 8, 3923-3940, doi:10.5194/amt-8-3923-3940-2015, 2015.

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC–MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. Mon. Wea. Rev., 126, 1373–1395, doi:10.1175/1520-0493(1998)126,1373:TOCMGE.2.0.CO;2, 1998a.

Côté, J., Desmarais, J.-G., Gravel, S., Méthot, A., Patoine, A., Roch, M., and Staniforth, A.: The operational CMC–MRB Global Environmental Multiscale (GEM) model. Part II: Results. Mon. Wea. Rev., 126, 1397–1418, doi:10.1175/1520-0493(1998)126,1397:TOCMGE.2.0.CO;2, 1998b.

de Grandpré, J., Ménard, R., Rochon, Y. J., Charette, C., Chabrillat, S., and Robichaud, A. : Radiative impact of ozone on temperature predictability in a coupled chemistry–dynamics data assimilation system, Mon. Wea. Rev., 137, 679-692, doi:10.1175/2008MWR2572.1, 2009.

de Grandpré, J., Tanguay, M., Qaddouri, A., Zerroukat, M. and McLinden, C. A.: Semi-Lagrangian Advection of Stratospheric Ozone on a Yin-Yang Grid System, Monthly Weather Review, 144, 1035--1050, doi:10.1175/MWR-D-15-0142.1, 2016.

Dee, D. P.: Bias and data assimilation, Q. J. R. Meteorol. Soc., 131(613), 3323-3343, 2005.

Dee D. P.: Importance of satellites for stratospheric data assimilation, In Proceedings of the ECMWF Seminar on Recent developments in the use of satellite observations in numerical weather prediction, 3–7 September 2007. ECMWF: Reading, UK, 2008.

Dee, D. P. and Uppala, S.: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, Q. J. R. Meteorol. Soc., 135(644), 1830-1841, 2009.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S.B., Hersbach, H., Hòlm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart. F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137, 553–597. doi:10.1002/qj.828, 2011.

Dittman, M. G., Ramberg, E., Chrisp, M., Rodriguez, J. V., Sparks, A. L., Zaun, N. H., Hendershot, P., Dixon, T., Philbrick, R. H., and Wasinger, D.: Nadir ultraviolet imaging spectrometer for the NPOESS Ozone Mapping and Profiler Suite (OMPS), Proc. SPIE 4814, Earth Observing Systems VII, 25 September 2002, Seattle, WA, edited by Barnes, W. L., 111–119, 2002a.

Dittman, M. G., Leitch, J. Chrisp, M., Rodriguez, J. V., Sparks, A., McComas, B., Zaun, N., Frazier, D., Dixon, T., Philbrick, R., and Wasinger D.: Limb broad-band imaging spectrometer for the NPOESS Ozone Mapping and Profiler Suite (OMPS), Proc. SPIE 4814, Earth Observing Systems VII; 25 September 2002, Seattle, WA, edited by: Barnes, W. L., 120–130, 2002b.

Dragani, R. and. Dee D. P.: Progress in ozone monitoring and assimilation, ECMWF Newsletter 116, 35–42. 2008

Dupuy E., and Coauthors: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys., 9, 287–343, 2009.

Durbin, P., Tilmes, C., Duggan, B., and Das B.: OMI near real time data processing, 2010 IEEE Intern. Geoscience and Remote Sensing Symp., Honolulu, Hawaii, USA, 586-588, July 25-30, 2010.

Evans, R., McConville, G., Oltmans, S., Petropavlovskikh, I., and Quincy, D.: Measurement of internal stray light within 30 Dobson ozone spectrophotometers, Int. J. Remote Sens., 30(15), 4247–4258, doi:10.1080/01431160902825057, 2009.

Fioletov, V. E., Kerr, J. B., Hare, E. W., Labow, G.J., and McPeters, R.D.: An assessment of the world ground-based total ozone network performance from the comparison with satellite data, J. Geophys. Res., 104, 1737–1747, 1999

Fioletov, V. E, Labow, G., Evans, R., Hare, E. W., Köhler, U., McElroy, C. T., Miyagawa, K., Redondas, A., Savastiouk, V., Shalamyansky, A. M., Staehelin, J., Vanicek, K., and Weber, M.: Performance of the ground-based total ozone network assessed using satellite data, J. Geophys, Res., 113, D14313, 19 pp, doi:10.1029/2008JD009809, 2008.

Fisher, M., and Andersson, E.: Developments in 4-D-Var and Kalman filtering, in: Technical Memorandum Research Department, 347, ECMWF, Reading, UK, 2001.

Flynn, L., Seftor, C., Larsen, J., and Xu, P.: Introduction to the Ozone Mapping and Profiler Suite (OMPS), in Earth Science Satellite Remote Sensing Vol.1: Science and Instruments, edited by J. J. Qu et al., Springer, Berlin, ISBN: 978-3-540-35606-6, 2006. [Also http://npp.gsfc.nasa.gov/omps.html.]

Flynn, L. (editor): Solar Backscatter Ultraviolet Instrument (SBUV/2) Version 8 Ozone Retrieval Algorithm Theoretical Basis Document (V8 ATBD), 45 pp., 2 February 2007 [Available from: https://ozoneaq.gsfc.nasa.gov/media/docs/SBUV2_V8_ATBD_020207.pdf]

Flynn, L. E., McNamara, D., Beck, C.T., Petropavlovskikh, I., Beach, E., Pachepsky, Y., Li, Y. P., Deland , M., Huang, L.-K., Long ,C. S., Tiruchirapalli R., and Taylor, S.: Measurements and products from the Solar Backscatter Ultraviolet (SBUV/2) and Ozone Mapping and Profiler Suite (OMPS) instruments, International Journal of Remote Sensing, 30,4259-4272, 2009.

Flynn, L., Long, C., Wu, X., Evans, R., Beck, C. T., Petropavlovskikh, I., McConville, G., Yu, W., Zhang, Z., Niu, J., Beach, E., Hao, Y., Pan, C., Sen, B., Novicki, M., Zhou, S., and Sefton, C.: Performance of the ozone mapping and profiler suite (OMPS) products, J. Geophys. Res., 119(10), 6181-6195, doi: 10.1002/2013JD020467, 2014.

Fortuin, J. P., and Kelder, H.: An ozone climatology based on ozonsonde and satellite measurements, J. Geophys. Res., 103, 31 709-31 734, 1998.

Garane, K., Lerot, C., Coldewey-Egbers, M., Verhoelst, T., Elissavet Koukouli, M., Zyrichidou, I., Balis, D., S., Danckaert, T., Goutail, F., Granville, J., Hubert, D., Keppens, A., Lambert, J.-C., Loyola, D., Pommereau, J.-P., Van Roozendael, M., and Claus Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 1: Ground-based validation of total ozone column data products, Atmos. Meas. Tech., 11, 1385-1402, doi:10.5194/amt-11-1385-2018, 2018.

Gauthier, P., Chouinard, C., and Brasnett, B.: Quality control: Methodology and applications. In Data Assimilation for the Earth System, R. Swinbank et al. (eds.), NATO Science Series IV: Earth and Environmental Science Vol. 26, Kluwer Academic Publishers, 177-187, 2003.

Girard, C., Plante, A., Desgagné, M., McTaggart-Cowan, R., Cote, J., Charron, M., Gravel, S., Lee, V., Patoine, A., Qaddouri, A., Roch, M., Spacek, L., Tanguay, M., Vaillancourt, P., Zadra, A.: Staggered vertical discretization of the Canadian Environmental Multiscale (GEM), Model using a coordinate of the log-hydrostatic-pressure type, Mon. Wea. Rev. 142, 1183-1196, doi: 10.1175/MWR-D-13-00255.1, 2014

GOME User Manual: Product User Manual for GOME Total Columns of Ozone, NO2, tropospheric NO2, BrO, SO2, H2O, HCHO, OClO, and Cloud Properties (O3M-SAF OTO and NTO), Doc. No. DLR/GOME/PUM/01, Issue 2/E, Deutsches Zentrum für Luft und Raumfahrt e.V. – DLR, Oberpfaffenhofen, Germany, 47 pp., 8 August 2012. [Available from: https://earth.esa.int/c/document_library/get_file?uuid=1bca1fc9-0525-4e67-b211-b23141ce83b7&groupId=10174.]

GOME-2 ATBD: Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, NO2, BrO, HCHO, SO2, H2O and Cloud Properties (GDP 4.8 for O3M-SAF OTO and NTO), Doc. No. DLR/GOME-2/ATBD/01, Issue 3/A,

Deutsches Zentrum für Luft und Raumfahrt e.V. – DLR, Oberpfaffenhofen, Germany, 58 pp., 30 March 2015. [Available from: https://wdc.dlr.de/sensors/gome2/DLR_GOME-2_ATBD_2A.pdf.]

Hao, N., Koukouli, M. E., Inness, A., Valks, P., Loyola, D. G., Zimmer, W., Balis, D. S., Zyrichidou, I., Van Roozendael, M., Lerot, C., and Spurr, R. J. D.: GOME-2 total ozone columns from MetOp-A/MetOp-B and assimilation in the MACC system, Atmos. Meas. Tech., 7(9), 2937-2951, 2014.

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, Jønch-Sørensen, J., H., Kalakoski, N., Kauppi, A., S. Kiemle, S., Kins, L., Koukouli, M. E., Kujanpää., J., Lambert6, 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., and Zyrichidou I.: Overview of the O3M SAF GOME-2 operational atmospheric composition and UV radiation data products and data availability, Atmos. Meas. Tech., 9, 383-407, doi:10.5194/amt-9-383-2016, 2016.

Inness, A., Baier, F., Benedetti, A., Bouarar, I., Chabrillat, S., Clark, H., Clerbaux, C., Coheur, P., Engelen, R. J., Errera, Q., Flemming, J., George, M., Granier, C., Hadji-Lazaro, J., Huijnen, V., Hurtmans, D., Jones, L., Kaiser, J. W., Kapsomenakis, J., Lefever, K., Leitão, J., Razinger, M., Richter, A., Schultz, M. G., Simmons, A. J., Suttie, M., Stein, O., Thépaut, J.-N.,Thouret, V., Vrekoussis, M., Zerefos, C., and the MACC team: The MACC reanalysis: an 8-yr data set of atmospheric composition, Atmos. Chem. Phys., 13, 407304109, doi:10.5194/acp-13-4073-2013, 2013.

IFS Documentation CY41R1: Part I: Observation, book chapter, European Centre for Medium-Range Weather Forecasts,72 pp., 2015. [Available from: http://www.ecmwf.int/sites/default/files/elibrary/2015/9208-part-i-observation-processing.pdf.]

Kerr, J. B.: New methodology for deriving total ozone and other atmospheric variables from Brewer spectrophotometer direct sun spectra, J. Geophys., Res., 107(D23), 4731-4747, doi:10.1029/2001JD001227, 2002.

Komhyr, W. D., Mateer, C. L., Hudson, R. D.: Effective Bass-Paur 1985 ozone absorption coefficients for use with Dobson ozone spectrophotometers, J. Geophys. Res., 98(D11), 20451-20465, 1993.

Koukouli, M., Balis, D., Loyola, D., Valks, P., Zimmer, W., Hao, N., Lambert, J. -C., Van Roozendael, M., Lerot, C., and Spurr, R. J. D.: Geophysical validation and long-term consistency between GOME-2/MetOp-A total ozone column and measurements from the sensors GOME/ERS-2, SCIAMACHY/ENVISAT and OMI/Aura. Atmos. Meas. Tech., 5(5), 2169-2181, 2012.

Koukouli, M. E., Lerot, C., Granville, J., Goutail, F., Lambert, J.-C., Pommereau, J.-P., Balis, D., Zyrichidou, I., Van Roozendael, M., Coldewey-Egbers, M., Loyola, D., Labow, G., Frith, S., Spurr, R., and Zehner, C.: 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, doi:10.1002/ 2015JD023699, 2015.

Koukouli, M., Zara, M., Lerot, C., Fragkos, K., Balis, D., van Roozendael, M., Allart, M., and Rvan, A. R.D: The impact of the ozone effective temperature on satellite validation using the Dobson spectrophotometer network, Atmos. Meas. Tech. 9, 5, 2055-2065, doi:10.5194/amt-9-2055-2016, 2016.

Kroon, M., Veefkind, J. P., Sneep, M., McPeters, R. D., Bhartia, P. K., and Levelt, P. F.: Comparing OMI-TOMS and OMI-DOAS total ozone column data, J. Geophys. Res., 113(D16S28), doi:10.1029/2007JD008798, 2008.

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

Lahoz, W., and Errera, Q: Constituent Assimilation, in Data Assimilation: Making Sense of Observations with eds W, Lahoz, B. Khattatov and R. Ménard, Springer-Verlag Berlin Heidelberg, 449-490, doi:10.1007/978-3-540-74703-1_18, 2010.

Lerot, C., Van Roozendael, M., Spurr, R., Loyola, D, Coldewey-Egbers, M., Kochenova, S., van Gent, J., Koukouli, M., Balis, D., Lambert, J.-C., Granville, J., and Zehner, C.: Homogenized total ozone data records from the European sensors GOME/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOp-A, J. Geophys. Res. Atmos., 119, 1639–1662, doi:10.1002/2013JD020831, 2014.

Levelt, P. F., Joiner, J., Tamminen, J., Veefkin, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B., N., Streets, D. G. Eskes, H., van der A, R, McLinden, C., Fioletov, V., Carn, S., de Latt, J. DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, S., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kelipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Gijsbert Tilstra, L., Torres, O., Wang., H., andf Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmos. Meas. Phys., 18, 5699-5745, doi:10.5194/acp-18-5699-2018, 2018.

Loyola, D. G., Koukouli, M. E., Valks, P., Balis, D. S., Hao, N., Van Roozendael, M., Spurr, R. J. D., Zimmer, W., Kiemle, S., Lerot, C., and Lambert, J. -C.: The GOME-2 total column ozone product: Retrieval algorithm and ground-based validation, J. Geophys. Res., 116(D7), 2011.

MACC-II Final Report: MACC-II Monitoring Atmospheric Composition and Climate - Interim Implementation, project funded from the European Union's Framework Programme, 137 pp., October, 2014. [Available from: http://atmosphere.copernicus.eu/sites/default/files/repository/MACCII_FinalReport_0.pdf.]

McLinden, C. A., Olson, S. C., Hannegan, B., Wild, O., Prather, M. J., and Sundet, J.: Stratospheric ozone in 3-D models: A simple chemistry and the cross-tropopause flux, J. Geophys. Res., 105, 14653–14665, 2000.

McPeters, R, D,, Frith, S., and Labow, G. J.: OMI total column ozone: extending the long-term data record, Atmos. Meas. Tech., 8, 4845–4850, doi:10.5194/amt-8-4845-2015, 2015.

McPeters, R. D., Bhartia, P. K., Haffner, D., Labow, G. J., and Flynn, L.: The version 8.6 SBUV ozone data record: An overview, J. Geophys. Res., 118., 1–8, doi:10.1002/jgrd.50597, 2013.

McPeters, R., Kroon, M., Labow, G., Brinksma, E., Balis, D., Petropavlovskikh, I., Veefkind, J. P., Bhartia, P. K., and Levelt, P. F.: Validation of the AURA Ozone Monitoring Instrument total column ozone product, J. Geophys. Res., 113(D15), 2008.

Moeini, A., Vaziri, Z., McElroy, C. T., Tarasick, D. W., Evans, R. D., Petropavlovskikh, I., and Feng, K.-H.: The effect of instrumental stray light on Brewer and Dobson total ozone measurements, Atmos. Meas. Tech. Discuss., 29 pp., doi:10.5194/amt-2018-2, 2018.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., and Eisinger, M.: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview, Atmos. Meas. Tech., 9, 1279-1301, doi: 10.5194/amt-9-1279-2016, 2016.

Ozone Monitoring Instrument (OMI) Near Real Time Data User's Guide, 2010. [Available from: https://ozoneaq.gsfc.nasa.gov/media/docs/OMI-NRT-DUG.pdf.]

Ozone Monitoring Instrument (OMI) Data User's Guide, produced by the OMI Team, 62 pp.,January 5, 2012. [Available from: https://acdisc.gesdisc.eosdis.nasa.gov/data/s4pa/Aura_OMI_Level2G/OMTO3G.003/doc/README.OMI_DUG.pdf.]

Redondas, A., Evans, R., Stuebi, R., Köhler, U., and Weber, M.: Evaluation of the use of five laboratory-determined ozone absorption cross sections in Brewer and Dobson retrieval algorithms, Atmos. Chem. Phys., 14, 1635-1648, doi:10.5194/acp-14-1635-2014, 2014.

Staehelin, J., Kerr, J., Evans R., and Vanicek, K.: Comparison of total measurements of Dobson and Brewer spectrometers and recommended transfer functions, Report No. 149, Meteorol. Org., Global Atmosphere Watch, Geneva, 2003.

Steinbrecht, W., Hegglin, M. I., Harris, N., and Weber, M.: Is global ozone recovering?, C. R. Geoscience, 350, 368-375, doi:10.1016/j.crte.2018.07.012, 2018.

Tereszchuk, K. A., Rochon, Y. J., McLinden, C. A., and Vaillancourt, P. A.: Optimizing UV Index determination from broadband irradiances, Geosci. Model Dev., 11, 1093–1113, doi:10.5194/gmd-11-1093-2018, 2018.

van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012, Atmos. Meas. Tech., 8(7), 3021-3035, 2015.

van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Multi sensor reanalysis of total ozone, Atmos. Chem. Phys., 10(22), 11277-11294, 2010.

van Roozendael, M., Peeters, P., Roscoe, H. K., De Backer, H., Jones, A. E., Bartlett, L., Vaughan, G., Goutail, F., Pommereau, J.-P., Kyro, E., Wahlstrom, C., Braathen, G., and Simon, P. C.: Validation of ground-based visible measurements of total ozone by comparison with Dobson and Brewer spectrophotometers, J. Atmos. Chem., 29, 55-83, 1998.

Van Weele, M., Dameris, M., Braesicke, P., and van der A, R.: User Requirements Document Ozone_cci, Issue 3, Ozone_cci Phase-2 (Project Deliverable D1.1), ESA-ESRIN, Reference: Ozone_cci_URD_3, Ozone-CCI, 2016, {Available from: http://www.esa-ozone-cci.org/?q=webfm_send/175, 2017.]

Veefkind, J. P. and de Haan, J. F.: DOAS Total O3 Algorithm, Chapter 3 in OMI Algorithm Theoretical Basis Document, Volume, 2, OMI Ozone Products, edited by P.K. Bhartia, Version 2.0, 33-50, 2002.

Veefkind, J. P., de Haan, J. F., Brinksma, E. J., Kroon, M., and Levelt, P. F.: Total ozone from the Ozone Monitoring Instrument (OMI) using the DOAS technique, IEEE Trans. Geosci. Remote Sens., 44(5), 1239-1244, 2006.

Viatte, C., Schneider, M., Redondas, A., Hase, F., Eremenko, M., Chelin, P., Flaud, J.M., Blumenstock, T., and Orphal, J.: Comparison of ground-based FTIR and Brewer O 3 total column with data from two different IASI algorithms and from OMI and GOME-2 satellite instruments, Atmos. Meas. Tech., 4(3), 535-546, 2011.

Waters, J. W., and Coauthors: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, IEEE Trans. Geosci. Remote Sens., 44, 1075–1092, doi:10.1109/TGRS.2006.873771, 2006.

WMO (World Meteorological Organization): Manual on the global data processing system, Volume I (Annex IV to the WMO Technical Regulations), Global Aspects, Report WMO-No. 485, Secretariat of the World Meteorological Organization, Geneva, Switzerland, 1992.

Zadra, A., McTaggart-Cowan, R., Vaillancourt, P. A., Roch, M.Bélair, S., and Leduc,A.-M.: Evaluation of tropical cyclones in the Canadian Global Modeling System: Sensitivity to moist process parameterization. Mon. Wea. Rev., 142, 1197–1220, doi:10.1175/MWR-D-13-00124.1, 2014a.

Zadra, A., Antonopoulos, S., Archambault, B., Beaulne, A., Bois, N., Buehner, M., Giguère, A., Marcoux, J., Petrucci, F., Poulin, L., Reszka, M., Robinson, T., St-James, J., and Rahill, A.: Improvements to the Global Deterministic Prediction system (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0). Canadian Meteorological Centre Tech. Note, 88 pp., 2014b [Available from: http://collaboration.cmc.ec.gc.ca/cmc/CMOI/product_guide/docs/lib/op_systems/doc_opchanges/technote_gdps300_20130213_e.pdf.]

Zhand, Z., and Kasheta, Z.: Global Ozone Monitoring Experiment-2 (GOME-2) Operational Ozone Product System Version 8: Interface Control Documnent (Documentation Version 1.0), June 2007, prepared for the National Environmental Satellite, Data, and Information Service (NESDIS), revision of July 2009.

Zhou, L., Divakarla, M., and Liu, X.: An overview of the Joint Polar Satellite System (JPSS) science data product calibration and validation, Remote Sens., 8, 139, 13 pp, doi:10.3390/rs8020139, 2016.

Table captions

Table 1. Regional and global relative mean differences (%) of total column ozone between OMI-TOMS and the specified ground-based instument types over July-August 2014/2015 and January-February 2015. The averaging excludes stations having outlier station mean diffferences for each period (see Supplement tables S1 to S3 and the text of Section 3) except for the two rows for the latitude region 60-90° S as described in the text. The standard deviations (S.D.) are for the inter-station variation of the mean differences about the regional or global mean differences. Unavailable S.D. values for available mean differences imply the presence of only one station. The Dobson total column ozone measurements for the two July-August periods were adjusted as a function of the ozone effective temperature (see Section 2.4); those for the January-February period were not adjusted in the absence of the ozone effective temperature for the period. The impacts of the Dobson July-August period corrections on the global mean differences were reductions between 0.0 and 0.4 %.

Table 2. Global diagnostics of differences in total column ozone between satellite instuments and OMI-TOMS for July-August 2014 and January-February 2015. The diagnostics consists of global mean differences and percentages of non-empty SZA/latitude bins with mean differences exceeding 2 % in magnitude.

Table 3. Mean differences of the total column ozone (%) between satellite instuments and OMI-TOMS for July-August 2014 and January-February 2015 for Northern and Southerm Hemispheres, for solar zenith angles below and above 70°.

Table 4. Mean differences in total column ozone (%) between satellite instuments and OMI-TOMS for July-August 2014 using the options (a), (b), and (c) from Section 4.2, for Northern and Southerm Hemispheres and solar zenith angle below and above 70°.

Table 5. List of assimilation experiments and their corresponding identifiers. In the second column, an asterisk (*) next to the instrument denotes that the bias-corrected observations (using the colocation method of Section 4.1) were assimilated.

Table 6. Global mean differences (%) between Brewer and Dobson total column ozone measurements and short-term forecasts for July-August 2014. Bias-corrected observations from the colocated observation bias correction scheme (Section 4.1) were applied in the assimilations. The Dobson measurements used were adjusted as a function of the ozone effective temperature (see Section 2.4). The uncertainties denote the standard error of the mean differences or standard deviations of the differences. The data from the two Antarctic stations have been included here even though their mean differences with OMI are outliers relative to most mean differences (Tables S1 and S2).

Figure captions

Figure 1. Mean total column ozone differences (%) between OMI-TOMS and Brewer, Dobson, and filter ozonometer measurements over July-August 2014. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Figure 2. Mean total column ozone differences (%) between GOME-2A/B, OMPS-NM/NP, SBUV/2-TC/NP and colocated OMI-TOMS data for the period of July-August 2014. The SBUV/2-TC total column ozone values stem from the two wavelength retrieval, while those for SBUV/2-NP are the sums of the retrieved 21-layer partial columns. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Figure 3. Same as Fig. 2 for January-February 2015.

Figure 4. Time series of total column ozone bias corrections (DU) for July and August 2014 for GOME-2A/B, OMPS-NM/NP, and SBUV/2-TC/NP as derived from the colocation method described in Section 4.1. Dashed vertical lines show individual six-hour mean differences with OMI-TOMS, while the solid curves of the same colour show the two-week moving average bias corrections. The particular (latitude, solar zenith angle) bins plotted are 5° wide bins centred on (52.5°N, 37.5°) for GOME-2A/B and OMPS-NM and a 10° wide bin centred on (55°N, 35°) for OMPS-NP and SBUV/2-TC/NP. Time coverage for individual bins do not necessarily cover complete months.

Figure 5. Time mean total column ozone biases (%) between GOME-2A and OMI-TOMS for July-August 2014 from colocation alone and for the options (a), (b), and (c) of Section 4.2 that use observation-minus-forecast differences. For options (a), (b), and (c), the forecasts were taken from the 'OMI' assimilation run (see Table 5). The bias in the 'colocations alone' panel was computed using the thinned observation

data set to compare to the other cases that use thinned observations. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Figure 6. Time series of total column ozone bias corrections (DU) for two latitude/SZA bins covering July-August 2014 for GOME-2A using different bias correction methods. All cases that include colocation methods use thinned observation sets. The 'O-F' curves additionally use the differences of forecasts described in Section 4.2 following the assimilation of OMI-TOMS. The 'colocations alone' and 'O-F' curves were calculuated using the Gaussian two-week moving average with HWHM of 4.7 days. The 'T_{eff}/SZA' curves, described in Section 4.3, result from mapping each observation that falls within the latitude/SZA bin onto the ozone effective temperature/SZA bias estimate for July-August 2014 (shown in Fig. 7), followed by taking the average of these bias estimate values for each time.

Figure 7. Mean total column ozone differences (%) between GOME-2A, OMPS-NM and colocated OMI-TOMS data as a function of ozone effective temperature (degrees Kelvin) and solar zenith angle (degrees) for the periods of July-August 2014 and July-August 2015. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Figure 8. Zonal mean total column ozone statistics of mean differences (%), standard deviations (%), and anomaly correlation coefficients (ACC; unitless) as a function of latitude (degrees) for the comparison between OMI-TOMS measurements and short-term forecasts for July-August 2014. The legends in the top plots indicate the assimilation runs (see Table 5 for description) and apply to all plots in the same column.

Figure 9. Zonal mean differences (%) and anomaly correlation coefficients (unitless) for total coumn ozone between OMI-TOMS observations and short-term forecasts as a function of time (date). Results are shown for the case without assimilation as well as with the assimilation of OMI, GOME-2A/B, and OMPS-NM (both with and without bias correction). The legend indicates the assimilation run (see Table 3 for description). Each value plotted was calculated using a 24 hour time window.

Figure 10. Zonal mean standard derivations (DU) of the differences between short-term forecasts *F*, OMI-TOMS observations *O*, and climatological values *C* over July-August 2014. For the short-term forecasts, the assimilation run from which the forecast was launched is indicated in the subscript, with the labels described in Table 5.