Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


Atmospheric Chemistry and Physics Discussions

# 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.** The impact of assimilating total column ozone datasets from single and multiple satellite data sources with and without bias correction has been examined with a version of the Environment and Climate Change Canada variational assimilation and forecasting system. The assimilated and evaluated data sources 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. 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 for total column ozone. 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 with the largest differences from the comparatively small dataset in Antarctica exceeding 3 %. Values from SBUV/2 summed partial columns were typically larger than OMI-TOMS on average by 0.6 to 1.2±0.7 %, with smaller differences than with ground-based over Antarctica. The OMI-TOMS dataset was chosen as the reference used in the bias correction instead of the ground-based observations due to OMI's significantly better spatial and temporal coverage and interest in near-real time assimilation. Bias corrections as a function of latitude and solar zenith angle were performed with a two-week moving window using colocation with OMI-TOMS and three variants of differences with short-term forecasts. These approaches are 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 changes in residual biases as a function of time for some cases. Assimilation experiments for the July-August 2014 period show a reduction of global and temporal mean biases for short-term forecasts relative to ground-based Brewer and Dobson data 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 are reduced to within 1 % for nearly all cases when bias corrected observations are assimilated for the latitudes where satellite data are present. The impact of bias correction on the standard deviations and anomaly correlation coefficients of



forecast differences to OMI-TOMS is noticeable but small compared to the impact of introducing any total column ozone assimilation. The assimilation of total column ozone data can result in some improvement, as well as some deterioration, in the vertical structure of forecasts when comparing to Aura-MLS and ozonesonde profiles. The most significant improvement in the vertical domain from the assimilation of total column ozone alone is seen in the anomaly correlation coefficients in the tropical lower stratosphere, which increases from a minimum of 0.1 to about 0.6. Nonetheless, it is made evident that the quality of the vertical structure is most improved when also assimilating ozone profile data, which only weakly affects the total column short-term forecasts.

**Copyright statement**



## 1 Introduction

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 use of ground-based and other remote sounding data for independent verifications and, occasionally, bias correction.

The assimilation process consists of introducing information from observations into the model forecasts through the generation of analyses, the statistical blend of the earlier forecast and observations in model space, serving as initial forecast conditions. Traditionally, this blending process assumes that both the model forecasts and observations are statistically unbiased following an initial spin-up time. Unremoved biases or systematic errors of 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), whether this is using measurements with or without data assimilation. Generally, while the effectiveness of bias correction schemes in removing the actual biases is constrained by the limited knowledge of the truth, their impact in reducing relative biases between different assimilation sources and/or correlated variables can potentially be as significant, if not more so, for improved forecasting. An example of this is in multivariate assimilation, where ozone and meteorological assimilation can be coupled (e.g., Dee, 2008; Dee et al., 2011), in the presence of differing systematic differences between forecasts and observations for different variables.



Total column ozone biases from satellite measurements are typically within a few percent. Changes of a few percent over time or between sources are significant in affecting the correct identification of long-term trends but not so much otherwise, possibly unless applying multivariate assimilation, where ozone and meteorological increments are statistically coupled, with biases large enough to negatively impact the meteorological fields; the latter assumes a correct vertical distribution of the total column ozone change and would extend to model biases. Total column bias correction is included in this ozone assimilation study considering comparatively short two month periods for two reasons other than the future potential extension to long-term trend monitoring and or multivariate assimilation. One is the conventional preference, that any residual measurement biases of the data affecting the analyses be generally less than their random errors. Second, it provides a mechanism to identify and correct bias occurrences which exceed typical levels even if not frequent or continuous. Both will contribute to the quality of UV index forecasting at least for clear sky conditions. The developed bias correction infrastructure can be applied to other constituents and data types as needed.

Total column ozone bias estimation for observations can be performed in different ways as a function of recognized dependency factors such as the solar zenith angle (SZA), latitude, and seasonal variation among others, potentially with related bias error covariances. Seasonal and related latitudinal changes in biases may result from limitations in retrieval algorithms such as, for example, not accounting well enough for the variation of the ozone layer temperatures in specifying 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). The bias parameterization may range from being spatially and temporally global to more local. As well, bias estimation and removal can be performed prior to generating the analyses during assimilation (as done in van der A et al. 2010 and 2015 and in this work for ozone) or during the objective analysis step (e.g., Dee, 2005; Dee and Uppala, 2009; Innes et al., 2013). The latter requires that the objective analysis setup is able to incorporate this task and directly provides the potential of reducing systematic differences between the model's short-term forecasts and the observations. This approach is being applied operationally by the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS, 2015) for both meteorology and ozone using a variational bias correction scheme. For ozone assimilation with the IFS, partial column profiles from the Solar Backscatter Ultraviolet instrument (SBUV/2; e.g., Bhartia et al., 2013; McPeters et al., 2013) on satellites of the National Oceanic and Atmospheric Administration (NOAA) are used as an anchor for the other assimilated ozone datasets. The anchor is the observation dataset serving as reference, or serving to establish the reference, about which bias estimation and correction are performed. This system is also applied with the Monitoring Atmospheric Composition and Climate project (MACC; e.g., Inness et al., 2013; and MACC-II Final Report, 2014) with profile measurements from both SBUV/2 and the Microwave Limb Sounder (MLS) on the Aura satellite serving as anchor. An intermediate version consists of applying two objective analysis steps: The first step assimilates the anchor data, followed by a separate bias estimation and correction for the initially unassimilated data that considers the differences between observations and analyses from the first step.

The validation of satellite remote sounding products often 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 are in the neighbourhood of roughly 1 % or better, excluding sites with outlier characteristics (considering Fioletov et al., 1999, 2005, 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. 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 for instruments such as for the Ozone Monitoring Instrument (OMI; Balis et al., 2007b; Viatte et al., 2011; Koukouli et al., 2012; Bai et al., 2016), the Global Ozone Monitoring Experiment-2 (GOME-2; Balis et al., 2007a; Antón et al., 2009b, 2011; Loyola et al., 2011; Koukouli et al., 2012; Hao et al., 2014), 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). A limitation in the use of ground-based observations as reference 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. In such cases, the applied bias parameterization may not necessarily capture as much of the absolute spatial or instrument-to-instrument relative bias variations as using observations from a satellite instrument covering a larger domain. If such a satellite-based reference is employed it should ideally be in good agreement with the ground-based measurements. Considering the limited projected lifetime and possible deterioration or failures of satellite-based instruments, transitions to new references also need to be envisaged in operational settings.

The focus of this study is specific to the assimilation of retrieved total column ozone observations from satellite remote sounding instruments using the Environment and Climate Change Canada (ECCC) meteorological assimilation system adapted for constituent assimilation. This work is being performed toward improving ECCC UV Index forecasts (Tereszchuk et al., 2018) and toward contributing to the monitoring of the ozone layer and to improving temperature forecasts through the radiative impact of ozone analyses (e.g., de Grandpré et al., 2009). While not unique to this work (e.g., van der A et al., 2010 and 2015; Innes et al. 2013), a major aspect of this undertaking is the harmonizing of the observations from multiple sensors through bias correction.

In this paper, we examine the effects of assimilating different individual or combined total column ozone satellite remote sounding observations with and without bias correction on short-term forecasts of ozone. These assimilations will be univariate ozone assimilations and utilize operational ECCC meteorological analyses. Six-hour forecasts are used when directly comparing forecasts from different experiments. The assimilated data sources used in this study are OMI aboard the Aura research satellite (Levelt et al., 2006), the GOME-2 instruments on the European MetOp-A and MetOp-B satellites (Callies et al., 2000; Munro et al., 2006), and the total column measuring instrument of OMPS (Dittman et al., 2002a,b; Flynn et al., 2006) on the Suomi National Polar-orbiting Partnership (S-NPP) satellite. In view of interest in near-real time



assimilation, criteria applied in this paper for the selection of an anchor for total column ozone bias correction of the different instruments include the availability of large near-real time daily datasets with good spatial coverage over multiple years, and having differences from well-calibrated ground-based data being stable over time and close to zero on average. Other studies, such as van der A et al. (2010 and 2015) for reanalyses covering many years, have instead directly used ground-based data. Summed partial columns from SBUV/2 satellite instruments have been recommended as the anchor for long-term studies (Labow et al. 2013) considering the long-term time coverage provided by the series of SBUV/2 instruments combined with the variation over time of overall differences from ground-based data remaining usually within ±1 %, with reduced differences for recent years. As such, SBUV/2 data, which also satisfy the above mentioned criteria, could serve as anchor for bias correction. Labow et al. (2013) also show differences over time of SBUV/2 with OMI data based on the Total Ozone Mapping Spectrometer (TOMS) version 8 total column retrieval algorithm (Bhartia and Wellemeyer, 2002) remaining between about 1 and 2 %. McPeters et al. (2015), showing similar magnitudes and stability of differences in time, conclude that OMI-TOMS data can 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). Since OMI is a mapping instrument, thus allowing for more colocations with measurements from other instruments over a few days, the OMI-TOMS dataset was instead selected as the anchor in this work. Verifications are performed with the OMI dataset and other independent measurements. In turn, the OMI-TOMS dataset is compared to the data from Brewers, Dobsons, filter ozonometers, and SBUV/2 over the examined time period.

A time varying bias dependent on both latitude and solar zenith angle is estimated that uses the differences with colocated total column ozone OMI measurements. Other dependencies not included in this study, such as changes in cloud cover and viewing zenith angle, would contribute to the resultant observation error standard deviations estimated in this work and in contributions to the far wings of the observation error probability distribution, and for some cases might result in outliers whose influence would be screened out or reduced through the assimilation quality control. A discrete binning in latitude and solar zenith angle will be used instead of a functional parameterization to allow for arbitrary nonlinear dependencies. These bias estimates will be compared to bias estimates that use differences between observations and short-term forecasts (innovations) in addition or as an alternative to colocation differences. The dependence on the ozone effective temperature is briefly evaluated in comparison to the temporal dependence with latitude.

A description of the assimilation and forecasting system, the observations, and the bias correction approaches is provided in Section 2. Section 3 consists of an evaluation of the OMI-TOMS data over the different two month time periods studied in this paper, the examination of bias estimations and corrections with a few approaches, and the analysis of relative impacts of the different assimilation scenarios. Conclusions are provided in Section 4. The Supplemental material document to this paper provides additional figures and tables supporting and complementing the discussed and presented results, the figure and table numbers of which are identified beginning with the letter ‘S’.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018




## 2 Experimental setup and methodology

### 2.1 The assimilation system

Global ozone assimilation experiments are performed with the operational Global Environmental Multiscale (GEM; Côté et al., 1998a and 1998b; Charron et al., 2012; Zadra et al., 2014a,b; Girard et al., 2014) numerical weather prediction (NWP) model coupled to a linearized ozone model (LINOZ) (McLinden et al., 2000; de Grandpré et al., 2016) using incremental three-dimensional variational (3D-Var) assimilation with first guess at appropriate time (FGAT; Fisher and Andersson, 2001). The 3D-FGAT assimilation with GEM-LINOZ 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. 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 toward the GEM ozone climatology (Fortuin and Kelder, 1998) is 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 is executed with a 7.5 min time step for 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°. 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. 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 separately from assimilations in the study. The same process of updating the meteorological fields every six hours is also applied for the reference case without ozone assimilation.

Incremental assimilation is conducted at each six-hour interval with T280 spectral space background error covariances, an FGAT time resolution of 45 min, and increments at the model vertical levels on a Gaussian grid of 300 latitudes and 600 longitudes with final increments interpolated to the model grid. As part of the background error covariances, the third order autoregressive correlation model (TOAR; Gaspari and Cohn, 1995) was applied in specifying spectral space globally homogeneous, isotropic and separable horizontal and vertical error correlations. Initial vertical correlations were calculated in physical space on the model vertical grid using monthly six-hour time differences of the free running GEM-LINOZ model as proxy to six-hour forecast errors (Polavarapu et al., 2005; Jackson et al., 2008; Section 5.5 in Bannister, 2008). The resulting static monthly error correlations are only rough approximations as compared to more realistic non-homogeneous, non-isotropic and state dependent error correlations which could be estimated through ensembles; the latter is not done and not essential for this study. Fits to the TOAR correlation model were applied to the correlations for each level to give localized non-negative correlations (see Figs. S1 and S2 in the Supplement). Smoothing was applied to the vertical level-dependent set of derived e-folding error correlation lengths of the TOAR functions, defined as distances for correlations of



$e^{-1}$. The vertical error correlation lengths are similar to those implied by the correlation length scales shown in Massart et al. (2009; accounting for approximate scaling from d[ln p] to km). The vertical error correlation lengths above 100 hPa correspond to half width at half maximum values (multiplying correlation lengths by ~1.2) that nearly equal the model vertical resolutions with values ranging from ~0.5 km at 100 hPa to ~3.5 km at 1 hPa. The horizontal error correlation lengths are ~125 km near the surface and increase from ~165 km at 100 hPa to just under 750 km at 1 hPa. A summary of the applied background and observation error variances is provided in Section 2.3.

In addition to the preliminary removal of data with large observation minus forecast differences (OmF; innovations) during background check, variational quality control (Andersson and Järvinen, 1999; e.g., Gauthier et al., 2003) is applied to the observations as part of the objective analysis step after the fifth iteration of the minimization. This reduces the weight of the observations associated to large increments at each minimization step in accordance to the size of the differences between the observations themselves and the latest analysis estimates from the most recent minimization iteration.

The assimilation experiments are conducted for July-August 2014, with a start date of 28 June 2014, 18 UTC, with and without bias correction for individual and combined datasets. The initial meteorological conditions for 28 June are from ECCC operational meteorological analyses and the initial ozone field is an analysis from an earlier assimilation.

## 2.2 Observations

In this section, we give a brief description of the column ozone observations made by the satellite-based instruments that are assimilated as well as the verification sources. The sources of the assimilated observations offer near-real time (NRT) column ozone products available for operational use as well as other products with additional or different processing resulting in a lag from NRT. The datasets selected for assimilation are being considered by ECCC for potential future NRT operational assimilation for contributing to both ozone-layer and UV index forecasting. All assimilated data sources in the study, namely the OMI, GOME-2, and OMPS-NM (total column Nadir Mapper) use optical solar backscatter of ultraviolet radiation in the nadir or near-nadir and provide data only during daytime. The data for GOME-2, OMPS, and SBUV/2 are being acquired in Binary Universal Form for the Representation of meteorological data (BUFR) from the National Environmental Satellite, Data, and Information Service (NESDIS/NOAA). For OMI we have used the standard science data column ozone products which are close to, but can differ slightly from, the NRT data (OMI NRT Data User's Guide, 2010; Durbin et al., 2010). The OMI data are obtained from the Earth Observing System Data and Information System of the National Aeronautics and Space Administration (EOSDIS/NASA). These differences between the two datasets are examined in Section 3.1. Thinning (reducing the pixel map density on the ground) to a resolution of ~1° is applied to all total column measurements in consideration of the background horizontal error correlation lengths and not including observation error correlations and is performed after the colocation-based bias estimation. The different sources of total column ozone observations are incorporated separately and not as an average when generating the analysis increments from multiple observation sources. The alternative of statistically averaging satellite observations from the different sources (e.g., van der A et al., 2010) is not done here and would decrease computational cost. As different nadir mapping satellite instruments



relying on solar backscatter radiation typically view similar horizontal regions over a few hours with measurements of sufficient precision for assimilation, one might argue that assimilating multiple sources is not necessary. On the other hand, assimilating data from different sources ensures that data availability gaps occurring in time and space from one instrument can be offset by the presence of other instruments and ensures temporal continuity following the interruption or end of an instrument's operation. Using multiple instruments can also serve in mutual near-real time monitoring of the different sources when it comes to temporal changes in biases and in better ensuring an automated transition to a new and acceptable reference source once the original anchor is no longer available.

The mean number of observations in a six-hour time period for the total column ozone measurement instruments under consideration are in the range of roughly 30 to 200 thousand depending on the source, with a reduction by factors of 5 to 20 from thinning (Fig. S3 of the Supplement). Through their much larger numbers (e.g., each by a factor of ~20 or more), the thinned OMI, GOME-2, and OMPS-NM total column measurement sets result in a stronger influence on the total ozone analyses and short-term forecasts as compared to profilers such as OMPS-NP and MLS, the later used here mainly for verifications. Adding the assimilation of profilers, while not done here except for MLS in one experiment, would serve to constrain the ozone field vertical structure.

Independent verifications of short-term forecasts are performed using both column and profile measurements from satellite, in-situ, and ground-based instruments. These sources are briefly identified following a description of the assimilated data sources.

While assimilation experiments are conducted only for the July-August 2014 period, bias corrections and evaluations are also done for July-August and/or January-February 2015 to illustrate possible differences and similarities in time and between seasons and successive years.

#### 2.2.1 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 are missing as a result from flagging of the row anomaly of the OMI instrument, which for the time period under consideration, effects 23 of the 60 rows.[1]

Two different level 2 total column ozone products are derived from the OMI radiances, one processed by NASA based on the Total Ozone Mapping Spectrometer (TOMS) version 8 (V8) total column retrieval algorithm and the other made by the Royal Netherlands Meteorological Institute (KNMI) using the Differential Optical Absorption Spectroscopy (DOAS)

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



algorithm. The TOMS V8 algorithm (Bhartia and Wellemeyer, 2002) uses principally 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).

Long-term term stability, relatively little solar zenith angle and latitude dependence, and small overall biases (averaging to roughly 1.5 % or better) have been observed in the OMI-TOMS retrieved data when compared to ground-based total column ozone measurements as well as being relatively stable over time (Balis et al., 2007b; McPeters et al., 2008;Koukouli et al., 2012; in addition to Labow et al., 2013, and McPeters et al,, 2015, referenced in the introduction). The OMI-TOMS total column ozone shows no to little (<1 %; but up to ~2% for cloud top pressure) dependency on cloud fraction, reflectivity, or cloud top pressure, depending on the reference when compared to ground-based Brewer and or Dobsons (Balis et al., 2007b; Antón et al., 2009a; Antón and Loyola, 2011; Koukouli et al., 2013; Bak et al., 2015; Bai et al., 2015) with increasing scatter for more overcast conditions (e.g., Balis et al., 2007b); ground-based column ozone data would also have errors associated to cloud cover.

The papers by van der A et al (2010 and 2015) indicate negligible variation with viewing zenith angle but a dependence on effective ozone temperature which reflects some dependence on latitude and time (e.g., Koukouli et al., 2016) comparable in magnitude to other satellite instruments. However, van der A et al. (2015) also include a second order correction as a function of solar zenith angle which is small as compared to most other instruments, and about the same or smaller than SBUV/2 instruments, but still non-negligible. We choose the OMI-TOMS total column ozone product as the anchor in our bias correction scheme, as already indicated in the introduction, and will examine the dependence with effective ozone temperature in Section 3. While not done here, one could generally adjust satellite-based anchors according to constant offsets, and long-term trends if needed, of differences with ground-based datasets, as done by van der A et al. (2010, 2015), for example.

The OMI-TOMS data have estimated root-mean squared errors of 1-2 % (OMI Data User's Guide, 2012). Comparisons of the standard OMI-TOMS data to ground-based measurements over the application periods for this work are provided in Sections 3.1. The corrections of van der A et al. (2015) for OMI-TOMS of a constant offset of ~3.3 DU plus a linear function of effective ozone temperature has not been applied here, nor the ~1.5 % offset determined in other studies (e.g., Labow et al., 2013; Bak et al., 2015; McPeters et al., 2015). There is also a second order non-negligible correction as a function SZA in van der A et al. (2105), even though small compared to most other instruments, which is not present in van der A et al. (2010). The implications of these corrections for the periods in the study are considered in Section 3.1.

The OMI NRT Data User's Guide (2010) and Durbin et al. (2010) indicate a daily maximum percentage difference of 2.6 % for the column ozone OMI-TOMS NRT products as compared to the OMI-TOMS standard products with a weekly



average maximum difference of 1.4 %. Time mean differences of the standard and NRT product versions as a function of latitude and solar zenith angle were calculated for July-August 2016 and January-February 2017 as the OMI-NRT observations were no longer available for the 2014-2015 time periods when this work was undertaken. All time mean differences over the entire range of bins for both periods were determined to be within roughly 0.02-0.04%, no larger than the 0.1 DU storage accuracy for the data files used to store the total column ozone observations, with the majority of individual differences being below the storage accuracy. While some individual point differences can be larger, time mean differences suggest that the near-real time and standard OMI-TOMS products would generate nearly equal bias estimates when serving as anchor.

#### 2.2.2 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 NOAA/NESDIS with the retrieval based on the TOMS V8 algorithm (e.g., Zhand and Kasheta, 2009). The DOAS total column ozone products are indicated to 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 operational NRT products were acquired from NOAA/NESDIS and stem from the TOMS approach. Their uncertainty characteristics estimated in this study are provided in Sections 2.3 and 3.2

#### 2.2.3 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 on board 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). Only the nadir mapper data is assimilated, with the summed partial columns of the nadir profiler also evaluated during bias correction. The nadir mapper has a cross-track width of about



2800 km and a 50 km × 50 km resolution at nadir. 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. Flynn et al. (2014) provide total column ozone accuracy and precision requirements of ~3.5-4 % and ~2 % 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 OMPS-NP provides profiles with a resultant 250 km × 250 km field of view. Only their summed partial column values are applied in the paper, and this only for comparison to the other datasets regarding total column biases. Profiles used were obtained with an implementation of the Version 6 SBUV/2 instrument algorithm (Bhartia et al., 1996) with the a priori profiles derived from the OMPS-NM. 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-NP V6 profiles consist of 12 layers with the profile layer bottoms (in hPa) of 1013 (surface), 253, 127, 63.3, 31.7, 15.8, 7.92, 3.96, 1.98, 0.99, 0.495, 0.247. The provisional OMPS-NP V6 product has the lowest layer extending to sea level instead of the surface. As a correction, the partial column values of the lowest layer were reduced by scaling according to the reduction in the pressure layer thickness of the lowest layer when moving the lower boundary from the sea level to the actual surface. Ozone profile accuracy and precision requirements are both 5-10 % in the middle to upper stratosphere for SZA up to 80° with larger uncertainty at lower and higher altitudes (Flynn et al., 2014). Based on the same reference, the OMPS-NP layer data would generally be consistent with corresponding SBUV/2 data processed with the V6 algorithm. The OMPS-NP data in the South Atlantic Anomaly region are recommended to be discarded (personal communications from L. Flynn) while total column measurements in this region are retained.

While the OMPS data used are not from the latest applied total column retrieval algorithms, the results obtained in Section 3 can serve in comparing to the overall quality of more recent dataset versions. The evaluation of Bai et al. (2013) for the OMPS-NP total column ozone obtained based on the TOMS Version 7 algorithm and covering January 2012 to February 2013 gives global mean differences with ground based data of 0.21 % for Brewer measurements and 0.86 % for Dobson measurements, each with standard deviations close to 3 %. The results of the evaluations from Bai et al. (2015, 2016) for the more recent OMPS-NM total column ozone products based 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 % and for the same period as Bai et al. (2013). Bai et al. (2016) provide, as reference, a distribution of OMPS-NM minus OMI-TOMS 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, with OMI being closer to the Dobson ground-based data (accounting for the seasonal variation of the differences).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018




### 2.2.4 Independent verification sources

Other than the OMI-TOMS data which serves as anchor, the verification data consist of Brewer, Dobson, and filter ozonometer total column ozone measurements and ozonesonde profiles from the World Ozone and Ultraviolet Radiation Data Center (WOUDC; for total column ozone measurements see Fioletov et al., 1999 and 2008; Staehelin et al., 2003; for ozone sondes, see references specified in Dupuy et al., 2009), Brewer and Dobson measurements from the Global Monitoring Division of the NOAA Earth System Research Laboratory, total column ozone from the Solar Backscatter Ultraviolet instrument (SBUV/2) on the NOAA 19 satellite (Flynn, 2007; Bhartia et al.,, 2013; McPeters et al., 2013); and ozone profiles from the Microwave Limb Sounder (MLS) on the Aura satellite (Waters et al., 2006; Livesey et al., 2006, 2013; Froidevaux et al., 2008; e.g., Inness et al., 2013; Adams et al., 2014). The error standard deviations are no larger than ~1.5 to 2.0 % for well-calibrated and well-maintained Brewer and Dobson direct sun data and about 1.5 to 2 times larger for filter ozonometers (based on Fioletov et al., 1999, and references therein), 5 % for ozonesondes (e.g., SPARC, 1998; Smit and Sträter, 2004), and as provided with the data for MLS (Livesey et al., 2013). The main disadvantage of filter ozonometers, relative to Brewers and Dobsons, is that the two spectral filters, i.e., the two wavelengths, used in ozone retrieval make them sensitive to atmospheric aerosols. Vertical piecewise averaging was applied to the ozonesonde profiles for final resolutions of 250 Pa up to an altitude of 25 km and 50 Pa for altitudes above 35 km with a linearly varying resolution between 25 km and 35 km.

The applied ground-based Brewer and Dobson spectrophotometer and filter ozonometer total ozone data are direct sun (clear-sky) daily daytime averages. 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, which is consistent with standard deviations being less than 2 % based on Fioletov et al. (1999 and 2008); the precision for filter ozonometers appear larger by roughly less than a factor of two. 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). 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 as well as accounting for the larger sensitivity for Dobson values. A consequence of avoiding the latter is 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. (2007b) 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 corrections for Dobsons showing a 1.5 % annual variation of the differences. As another example, Bai et al. (2016) show 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.

The ozone data from SBUV/2 for the period of interest are from the NOAA 19 satellite launched in February 2009. Two versions of the total column ozone data are used here. These are 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 the 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 agreement between total column ozone data of SBUV instruments form the summed partial columns and the Northern Hemisphere ground-based data to be better than 1 %. Bhartia et al. (2013) indicates that the total column ozone values from the summed partial column profiles allow extending the total column ozone retrievals to a solar zenith angle of 88°.

The MLS ozone profiles are from version 3.4 of the processing algorithm (Livesey et al., 2013) and are used without their averaging kernels. The along-track sampling and effective resolutions are ~160 km and, in the stratosphere, 300-450 km. The vertical levels of the retrieval solutions are in pressure with 12 levels per decade in the stratosphere, this corresponding to a vertical resolution of ~1.3 km, with the effective vertical resolution being ~2.5 km. The recommended usable vertical range for the ozone profiles is 261 hPa to 0.02 hPa with highest precisions of 2-3 % and accuracies of ~5 % in the range ~2-46 hPa.

### 2.3 Applied background and observation error variances

All the presented assimilation cases, except one, will have used an originally prescribed set of ozone observation and background error variances. An updated set of error variances considered to be closer to optimality, at least for the observations, was also derived with implications shown for one assimilation experiment. These updated variances were derived using the observations to which the colocation bias correction described in Section 2.4.1 was applied. Their derivation and the results are presented in this section, in addition to the originally applied variances, to support the target for reduction in total column observation bias indicated in Section 2.4. For this updated set, a single iteration of the Desroziers et al. (2005) approach applied to total ozone served as an initial step to adjust the error variances following a first set of assimilations. Separate additional metrics were then applied to complete the ozone background and observations error variance adjustments. The resultant new error variances are considered more representative of the actual observation error variances and so a better guide for imposing target conditions on observation bias reductions. As will be briefly shown, the overall diagnostics presented in Section 3 on the quality of the short-term ozone forecasts are largely consistent between the two sets of assimilations and so the results from the former can be taken as qualitatively applicable to the latter. Both are summarized for completeness.

The original error standard deviations assigned to all total column measurements are set to 2 % for all assimilated sources. For the updated set of statistics, the observation error standard deviations were re-estimated following the assimilation with all satellite total column datasets using two main steps: First, a single iteration of the Desroziers et al. (2015) approach was performed with the adjustments parameterized as a function of SZA. The variances of the differences between colocated observations (see Section 2.4.1) between each instrument and OMI were then computed and compared to the expected sum of the individual variances, which assumes the errors on the two different observation sets are not correlated. The variances of the differences between OMPS-NM and OMI were used to obtain an estimate of the error variance of OMI, where for this



estimate the error variances of OMPS-NM were replaced with the error variances of OMI scaled by the ratio of error variances between OMPS-NM and OMI as diagnosed by the Desroziers approach. The error variances of OMI were then used in the comparison of the variances of differences between collocated observations for the remaining instruments (GOME-2A/B, OMPS-NM) to estimate their error variances. The final updated error standard deviations are provided in Fig. 1. These error standard deviations, determined and smoothed as a function of solar zenith angle for each total column measurement source, are in the range of 1-2 % for SZA less than 70° and 1.5-3.5 % for higher SZAs, with the smaller values for both OMI and OMPS-NM. The obtained values are on the lower end of, or smaller than, the instrument error random uncertainties indicated in Section 2.2. On the other hand, global root-mean-square error levels obtained by van der A et al. (2010) for GOME-2A (DOAS) and OMI-TOMS from comparisons with ground based measurements were estimated to be 4.74 and 4.59 Dobson Units (DU), respectively. These translate to values in the range of ~1.3-1.9 % for column measurements between 250 and 350 DU. While not used for the applied updated values, tightening the colocation requirements for the differences with OMI from the criteria of Section 2.4.1 by factors of 4 and 2 respectively for the maximum separations in distance and time, and adding outlier removal, further reduces the final observation error standard deviations by roughly 10 % each, for a total reduction of 20 %. While colocation differences do not reflect model representativeness errors to the observation standard deviations, their contributions may not be that significant considering the respective sizes of model horizontal resolutions and the observation pixels on the ground. The actual standard deviations may therefore be smaller overall, say by 10-20 %, than those used as final values.

The original background error standard deviations as a function of latitude and vertical level are from a single iteration of the Desroziers et al. (2005) approach with an assimilation of MLS ozone data applied following originally assigned ozone background error standard deviations of 5 % in the stratosphere and upper troposphere. The resulting latitude varying background error standard deviations at sample vertical levels of 1, 10, 50, and 300 hPa are in the ranges ~6-12 %, ~3-5 %, ~5-15 %, and ~15-24 % respectively. Constant extrapolation in absolute value uncertainty was imposed for lower tropospheric levels with resultant percentage values being ~15-16 % of the Fortuin and Kelder (1998) climatology in volume mixing ratio. An update of the values assigned in the stratosphere and upper troposphere accompanied the update of the observation error standard deviations described above. This was done by scaling the original background error standard deviation field as a function of vertical level above 300 hPa by latitude dependent total ozone error standard deviation scaling factors. A linear tapering of the adjustment as a function of the natural logarithm of the Charney-Phillips vertical grid levels (section 2.1) was performed for lower altitudes to preserve the original surface values applicable for assimilation of surface observations even though surface observations were not used in this study. Total ozone background error standard deviation scaling factors were derived as a function of latitude from an assimilation of all four total column data sources, namely GOME-2A, GOME-2B, OMPS-NM and OMI, the first three being bias corrected relative to OMI. The latitude dependence of the scaling was set from a single iteration of the Desroziers et al. (2005) approach. A global absolute scaling was then determined by repeating the assimilation (with the updated observation error variances) a few times and successively scaling the background error variances to ultimately give a minimization chi-square (twice the cost-function



value divided by the number of observations) in the neighbourhood of unity (Fig. S4). The total ozone error standard deviations were thus reduced from the first set of values in the range of 1.8-3.4 % (with the larger values in the northern extra-tropics) to ~0.5 % at most latitudes, with an increase from 0.5 to 0.8 % between 40° S and 80° S (Fig. S5). These new values are typically factors of 2 or smaller than the observation standard deviations estimated as a function solar zenith angles. Contributors to these small values would be the large number of total column observations combined with their error standard deviations and, to some degree, having well predictable short-term column ozone forecasts. While still seeming rather small, these were applied as is. Using the cost-function for this purpose is valid for uncorrelated observation and background error variances, correct specification of spatial error correlations, and negligible residual biases under all conditions at least between observations and forecasts, which are not strictly true. Nonetheless, the change from the original to the updated set of error variances results in a density distribution function of the statistically normalized innovations (i.e., the distribution of observations minus forecasts divided by the prescribed standard deviation of the differences) being closer to the normal Gaussian probability density function with a reduction in the peak value from 0.72 to 0.35 and an increase in the full width at half maximum from ~1.1 to ~2.5; the normal Gaussian distribution has a maximum of ~0.40 and full width at half maximum of ~2.4. The final distribution of the statistically normalized innovations shows more outliers than from a normal Gaussian distribution (Fig. S6). Unaccounted observation systematic error sources such as may be associated with the treatment for cloud cover and differing viewing zenith angles may be contributing to outliers. The final distributions for the four total column datasets (OMI, GOME-2A/B and OMPS-NM) have maxima in the range 0.32 to 0.41 and full widths at half maximum in the range of roughly 2.2 to 2.6. These results do not imply that the ratios of observation to background error variances are correct. If the observations error standard deviations were potentially smaller by a scaling factor of say 0.8, retaining the same total prescribed error variances of the observation minus forecast differences would imply an increase of the background error standard deviations by a factor of roughly about 1.6 for regions of lower SZAs, to at least 2.0 for SZAs of ~70° and above (for latitudes most often near the southern pole for this period). This would result in background error standard deviation close in size to the OMI-TOMS and OMPS-NM error variances. The overall ratios of observation to background error variances would then be closer to those from the original variances,

While the background error variances would be partially dependent on the observation set being assimilated, they have not been accordingly adjusted in the work when changing the number of sources being assimilated.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


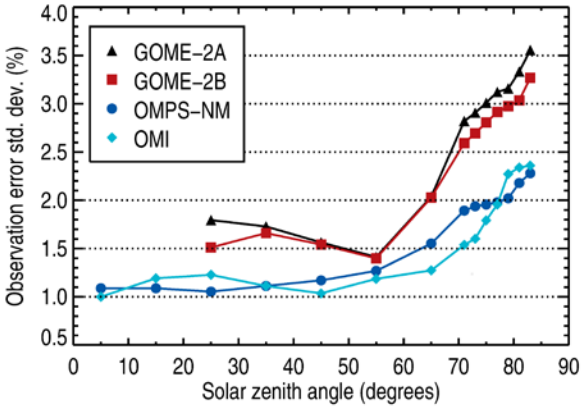


**Figure 1.** Updated set of applied observation errror standard deviation estimates (%) for GOME-2A, GOME-2B, OMPS-NM and OMI total column measurements as a function of solar zenith angle band centres (degrees), for the combined period of July and August 2014. A 2% constant served in specifying the first set of applied error standard deviation estimates.

### 2.4 Bias estimation and correction

Observation biases can be examined as a function of various factors. In this study, the bias correction applied in the assimilation experiments use 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. One of these other methods is to use the dependence on the ozone effective temperature instead of latitude and time (e.g., van der A et al., 2010). 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 are also introduced to capture other first order retrieval biases, such as potentially related to the applied a priori atmospheric state and its ozone error covariances, and the specification of the ozone absorption coefficients, as well as instrumental changes in 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 therefore be reflected in the estimated standard deviations derived for observations and outlier errors referred in Section in 2.3. The bias correction target is to reduce residual biases as a function of SZA and latitude relative to OMI-TOMS generally to within 1 %, considering total ozone measurement random error standard deviations being no smaller than ~1 % (Section 2.3). An evaluation of OMI-TOMS total column ozone as the anchor for the periods in this study is provided in Section 3.1 relative to ground-based data, with a supplementary evaluation using SBUV/2 in Section 3.2.1.

Most of the bias estimation results are presented and discussed in Section 3.2. The bias estimates and behaviours obtained in this study do not necessarily reflect the quality of the retrieval data for other periods. Some of the differences that are seen may be attributed to the colocations being approximate (where applicable) or to one or more of the other factors not taken into account here, such as bias as a function of the viewing zenith angle, which is not examined as it was not considered as



significant for instruments of interest based on the available literature. Variation as a function of ozone effective temperature (average ozone profile weighted temperature) as was applied for the van der A et al. (2010, 2015) reanalyses is separately examined in Section 3.2.2 as substitution for variations in latitude and time over weeks to months. The dependence of bias as a function of column ozone amount as another possible alternative is potentially subject to small systematic contamination from correlated uncertainties between the column ozone values and the differences and so is not applied for bias estimation.

### 2.4.1 Colocation approach

As indicated in Section 2.2.1 and in the introduction, the OMI-TOMS column ozone data were selected as anchor in bias estimation and correction. Separate bias estimations are conducted for each distinct instrument-platform pair, with the bias estimates assigned from the mean differences. The estimations can be performed either by direct comparison of observations at colocated locations or through the intermediate use of model forecasts (or analyses). While both are examined, the direct colocation approach without forecasts serves as the baseline. Here, the initial observation differences between colocated observations for use in bias estimation are taken as acceptable for points within 200 km and ±12h, with a latitude difference of no more than 3°, total column differences within 25 DU, and solar zenith angle differences smaller than 5° for SZA under 70° and smaller than 2° for SZA between 70° and 90°. A similar bias correction was applied to the equivalent total ozone of the OMPS-NP partial column profiles for comparison to OMPS-NM. The latitude and solar zenith angle bins have a size of 5° each for total column ozone measurements, and 10° each for partial column ozone profiles, except at larger solar zenith angles where bin sizes are reduced to 2° for 70-90°; larger bin sizes could have been used except likely at high solar zenith angles. 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 stem from the stronger gradients in the differences between instruments. The larger bin sizes for 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.

To provide preliminary insight on bias estimates, Fig. 2 shows the monthly and hemispheric mean colocation differences with OMI-TOMS for GOME-2 and OMPS-NM as function of SZA with separate curves for the Northern and Southern Hemispheres in August 2014. The differences between hemispheres suggest that including latitudinal variation would be more essential in some cases than others. The differences for OMPS-NM are consistent with the average biases of -2 to -4 % with respect to OMI-TOMS and SBUV/2 found by Flynn et al. (2014). The discontinuity in time mean differences with GOME-2A/B appearing at 70° in SZA may be associated to 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, no interpolation is applied over the SZA value of 70° for GOME-2A/B. The figure shows a better agreement in monthly mean differences with OMI-TOMS for GOME-2B with magnitudes under 1 % for SZA below 70° and a maximum of ~1.6 % or ~5 DU above 70°. Largest differences are found for GOME-2A and the provisional OMPS-NM data reaching at least ~4 %, or ~10-12 DU, at



some SZA values. The larger changes with SZA shown in the Southern Hemisphere for high SZA, e.g., above 80° for GOME-2, may also be found in the Northern Hemisphere depending on the month. In Section 3.1, it is shown that trends and variation over time of the differences within a month can be as significant for some latitude/SZA regions. The correlation of differences in latitude and time with effective ozone temperature is also characterized in Section 3.2

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 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 is applied to better capture variations in time. The six-hour mean differences over the two-week moving window are weighted in time with a Gaussian weighting function with a half width at half maximum of 4.7 days. The six-hour mean differences are 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 is 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.

### 2.4.2 Bias estimation involving differences with forecasts

An alternative bias estimation approach utilizes the differences of the original retrieved observation data with short-term forecasts with the same binning in latitude and solar zenith angle over a two-week moving window. These bias estimates can be obtained by considering the OMI total columns differences with forecasts (i.e., OmF), with or without colocation requirements, or simply without any direct use of OMI. The FGAT short-hour forecasts $F$ would have been influenced by the assimilation of all bias corrected data used in an experiment while the observations $O$ denote retrieved observations prior to bias correction. For each bin, the bias estimate could be obtained from moving time series window of

a) $\langle (O-F) - (O-F)_{\mathrm{ref}} \rangle$ with the same colocation requirements as Section 2.4.1,
b) $\langle O-F \rangle - \langle O-\mathrm{F} \rangle_{\mathrm{ref}}$ without the above colocation requirements, or
c) $\langle O-F \rangle$

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


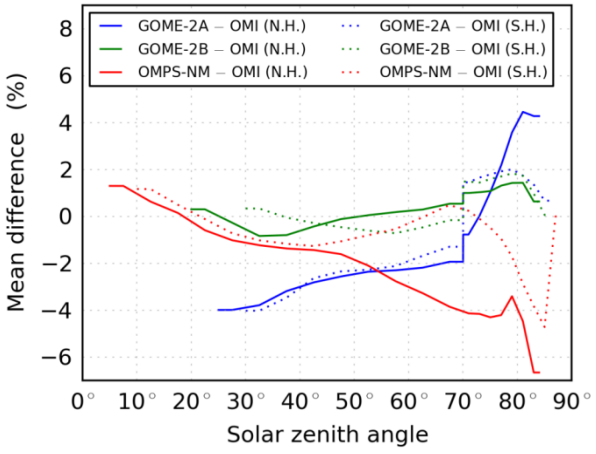


**Figure 2.** Mean total column ozone differences (%) between GOME-2A/B, OMPS-NM, and colocated OMI-TOMS data as a function of solar zenith angle (degrees) for the Northern and Sourthern Hemispheres for August 2014. Differences were computed from observation colocations. For the GOME-2 instruments, separate continuous difference regions are specified about 70° (see text in Section 2.4.1 for additional information). Constant difference extrapolation is applied from bin midpoints at the edges of distinct regions.

with the angular brackets denoting averages and the subscript 'ref' denoting differences for observations of the anchor set, this being OMI here. All three options with short-term forecasts are applied total column ozone measurements for comparison. Just as the colocation approach, these cases require the mean innovation differences being available for the SZA, latitude, and time bins. In this work, options (b) and (c) become successive fallback approaches to (a) in the absence of collocated anchor measurements for a bin, with option (b) automatically reducing to option (c) in the absence of the OMI or anchor data. 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. Option (c) provides a bias correction option for times and locations where the reference is not available. 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. The implications of having used or not used the anchor data in the assimilation are presented in Section 3.3.

## 3 Results

Assimilation runs covering July-August 2014 have been performed for individual and combined sets of total column ozone satellite data sources with and without bias correction based on the colocation approach. The resulting short-term forecasts from the assimilation runs are compared to each other and to the absence of assimilation.

Following an evaluation of OMI-TOMS for the two-month periods in Section 3.1, bias estimates are determined for the other total column datasets using the two general types of differences described in Section 2.4, one based only on colocations and the other on differences with forecasts. The dependence on the ozone effective temperature as done in van der A et al.

(2010) is examined separately as a substitution for latitude and time dependence. The bias corrections applied in assimilations are dependent on latitude and solar zenith angle with a two-week moving window. Bias estimation and correction using the observation colocation approach by itself can be done either prior to or as part of assimilation runs, the former being applied here. Mean differences with forecasts would normally be determined and applied for bias estimation during the assimilation and forecasting cycle. For convenience, here we instead use differences for spatially thinned observation sets, thinned to ~1° sampling, with six-hour forecasts of separate assimilation and forecasting sets. The estimated biases for different datasets and from different approaches are presented and discussed first in Section 3.2. Both the July-August and January-February periods are considered for a comparison of bias estimates between seasons within a yearly cycle. The related impact of total ozone assimilations on short-term forecasts is assessed in Section 3.3.

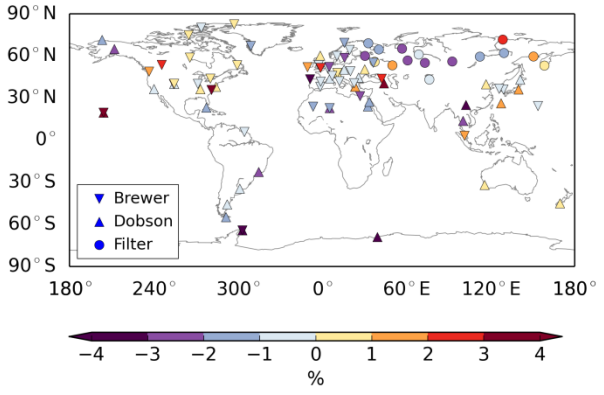


**Figure 3.** Mean total column ozone differences (%) between OMI-TOMS and Brewer, Dobson, and filter ozonometer measurements over July-August 2014. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

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

A mean differences comparison of OMI-TOMS with near-colocated ground-based data at available sites over the periods of study was conducted to support the selection of OMI-TOMS as anchor justified earlier through references identified in Section 2.2.1. Summary results are shown in Table 1 and Fig. 3 (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 52 Brewer, 40 Dobson, and 20 filter ozonometer locations; Fig. 3 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 instruments. These global and regional averages exclude stations with mean differences larger than two standard deviations, corresponding to between 3 and 4 % and obtained after removing stations with mean differences larger than 6 DU. The total number of these outlier stations per period ranges from 0 to 5 (Tables S1 to S3), with some of these being stations at high elevation or in Antarctica. The numbers of statistical



outliers are small in comparison to the much larger percentages of sites showing outlier or suspect characteristics over the 5-year period examined by Fioletov et al. (2008). Considering the few 2-3 stations in Antarctica and some consistency between station values, the outlier station mean differences are included as part of the regional mean differences for 60-90° S in Table 1.

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 instruments. Table 1 shows small positive biases less than 0.7 % over the region encompassing Canada, the continental United States and Greenland, as compared to small negative biases up to -0.4 % over Europe and Northern Africa. The global and most regional mean differences are notably smaller than the roughly 1.5 % underestimation identified by Labow et al. (2013), Bak et al. (2015), and McPeters et al. (2015).These studies have used multiple years in their analyses. McPeters et al. (2013) instead found positive average differences with Northern Hemisphere Brewers and Dobsons covering 2005 and 2006 of 0.6 % and 0.8 %, respectively, with uncertainties of 1.1 % and 1.5 %. The mean differences for both polar regions are all negative suggesting an underestimation of OMI-TOMS column ozone in these regions for these periods. 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 %. These fall within the range of the mean differences covering 2007-2010 from Koukouli et al. (2012) at -1.5±2.4 % for Brewers and -0.5±3.0 % for Dobsons. The adjustments of Dobson values applied in this study based on the ozone effective temperatures, which was not included in Koukouli et al. (2012), reduced the sizes of the mean differences in this region by less than 1 %. The average solar zenith angles for stations in the north polar 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°. As such, stray light would not significantly impact the total column ozone values (e.g., Moeini et al., 2018; Evans et al., 2009).

More severe underestimations at 3-6 % occur during July-August in Antarctica, this associated to 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 for the limited 1-3 stations in this region restricts the statistical significance of these results, the level of consistency between the instruments and sites suggest it being worthwhile to consider this data and so were retained in Table 1 for the rows of the 60-90° S region. The largest adjustments made to the July-August Dobson data occurred for the two Antarctica stations of Marambio (64.23° S, 56.62° W) and Syowa (69.01° S, 39.58° E) due to the low ozone effective temperatures, near 200 K, with increases in mean differences from near zero to -3 to -4 % (see also Koukouli et al, 2016). This brought the differences for Dobsons closer to those for the Brewer at the Marambio station. The differences for the Dobson at Ushuaia (54.85° S, 63.31° W) at the southern tip of Argentina are also negative in the range of -0.8 to -1.6 % over the three periods. Other factors may also affect the ground-based values 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. A related adjustment would increase the differences with OMI-TOMS. As well, ground-based measurements at very high SZAs

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


Atmospheric Chemistry and Physics Discussions

**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 3.1) 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 variation of the station 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 2.2.4); 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 | 0.0 [1816] | 1.4 | 0.1 [1422] | 1.6 | 0.1 [886] | 1.1 |
| | Latitude range: 30° S - 30° N | 0.4 [296] | 1.9 | -0.5 [165] | 0.7 | -0.2 [314] | 1.3 |
| | Latitude range: 30-60° S | - | - | - | - | - | - |
| | Latitude range: 60-90° S | -5.9 [13]* | - | - | - | -2.5 [152]# | 2.0 |
| | North America and Greenland | 0.7 [669] | 1.1 | 0.2 [1020] | 1.7 | 0.3 [495] | 1.0 |
| | Europe and Africa | -0.3 [1282] | 1.4 | -0.3 [780] | 1.2 | -0.3 [427] | 1.1 |
| | East Asia and Other | -0.1 [419] | 1.9 | -1.0 [148] | 0.8 | -0.3 [388] | 1.5 |
| | Global | 0.0 [2370] | 1.4 | -0.1 [1948] | 1.5 | -0.1 [1310] | 1.2 |
| Dobson | Latitude range: 60-90° N | -1.6 [64] | 1.3 | -1.2 [29] | 0.0 | - | - |
| | Latitude range: 30-60° N | 0.0 [421] | 0.7 | 0.5 [328] | 1.3 | 0.7 [187] | 1.0 |
| | Latitude range: 30° S - 30° N | 0.2 [270] | 2.7 | -0.8 [200] | 1.4 | 1.0 [149] | 2.2 |
| | Latitude range: 30-60° S | -0.4 [167] | 0.9 | -1.0 [111] | 0.4 | 0.3 [171] | 1.0 |
| | Latitude range: 60-90° S | -3.3 [6]+ | 0.1 | -4.3 [2]^ | - | -0.4 [134]$ | 1.7 |
| | North America and Greenland | -0.6 [167] | 0.6 | -0.8 [84] | 0.9 | 0.1 [73] | 0.5 |
| | Europe and Africa | -0.7 [400] | 1.3 | 0.2 [293] | 1.5 | 0.7 [135] | 1.1 |
| | East Asia and Other | 0.7 [361] | 2.0 | -0.5 [291] | 1.3 | 0.4 [433] | 1.7 |
| | Global | -0.1 [928] | 1.6 | -0.2 [668] | 1.4 | 0.4 [641] | 1.5 |
| filter ozonometer | Latitude range: 60-90° N | -1.4 [47] | 0.8 | -1.1 [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 | -0.7 [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.

require an unobscured horizon which may not always be possible throughout the day; often antennas, other buildings or structures, partially block the sun, thus producing low ozone values. 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 horizontal gradients in total column ozone. Approximately accounting for a latitudinal displacement of slightly more than 1° resulted, for example, in reducing the July-August 2014 mean difference from the Brewer at Marambio from the -5.9 % in Table 1 to -2.7 %. As another point of consideration, Figure 3 of Balis et al. (2007b) did not show significant dependence on solar zenith angles for OMI-TOMS relative to Brewers for SZA larger than 20° except above 80° with an overestimation of about 4±2 % instead of an underestimation. This difference, though, could be related to differences in time periods, stations, and dataset sizes. The discussion for the polar regions and high solar zenith angles is extended in section 3.2.1 with a comparison to SBUV/2 total column ozone data and a note regarding the assimilation with Aura-MLS in section 3.3.1.

While the OMI-TOMS data could be underestimates of 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, using a comparison to Brewer and adjusted Dobson data, which would increase the OMI-TOMS total column ozone in Antarctica by about 10.5 DU for the two July-August periods: including also the second-order dependence on SZA of that paper reduces the change to 9 DU. Its application 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.

### 3.2 Bias estimation

#### 3.2.1 Colocation approach

Time mean differences for the various instruments, GOME-2A, GOME-2B, OMPS-NM, OMPS-NP, and SBUV/2, with OMI-TOMS for July-August 2014 and January-February 2015 (Figs. 4 and 5) indicate global averaged biases in the range of -3.5 to 2 %. The two sets of results with SBUV/2 are for the SBUV/2 total column ozone from the two wavelengths retrieval (SBUV/2-TC) and the sum of the retrieved partial column profiles (SBUV/2-NP). These have been added to extend the evaluation of the OMI-TOMS data conducted in section 3.1. The averaging over all bins for July-August and January-February, with weighting by the number of colocations per bin, gives mean biases ranging from 1.5 % to -3.5 % over the different instruments in Figs. 4 and 5 (Table 2). The mean differences of OMI-TOMS with ground-based data (of opposite sign to the mean differences of the satellite data with OMI-TOMS) over the various sites are usually smaller in size than the average mean differences over all bins between the various satellite-borne instruments and OMI-TOMS. The mirroring patterns about the equator in Figs. 4 and 5 for at least SBUV/2 and OMPS-NP suggest the possibility of some seasonally dependent differences with OMI-TOMS for these data. 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; Section 2.2.3),

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


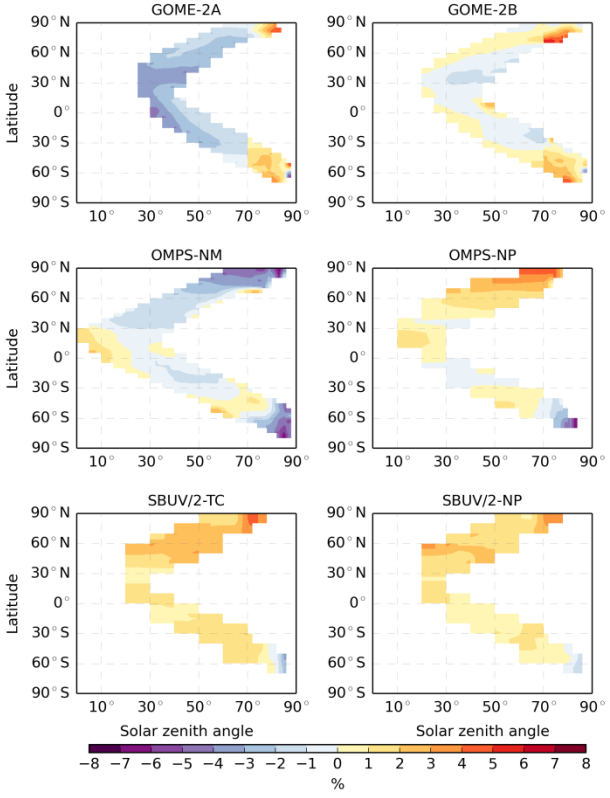


**Figure 4.** Mean total column ozone differences (%) between GOME-2A/B, OMPS-NM/NP, SBUV/2-TC/NP and colocated OMI-TOMS data for the period of July-August 2014. The SBUV/2-TC total column ozone values stem from the two wavelength retrieval while those for SBUV/2-NP are the sums of the retrieved 21-layer partial columns. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


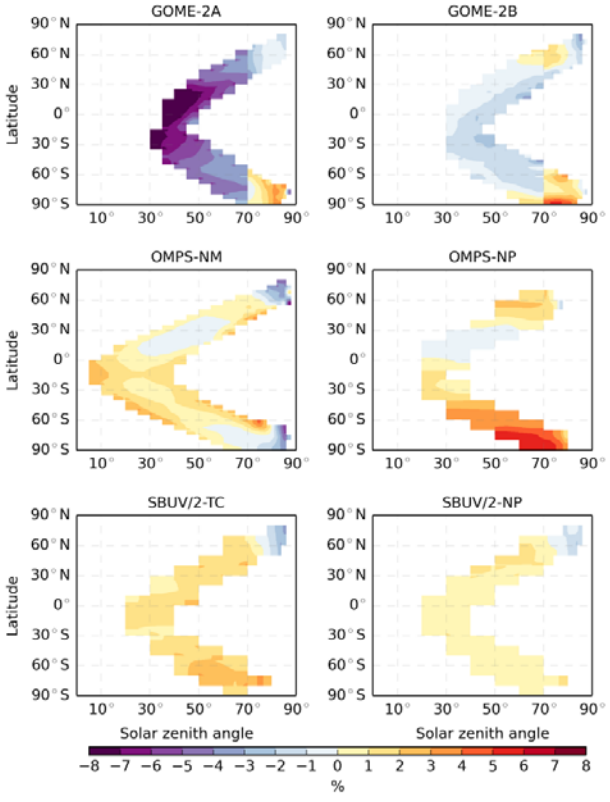


**Figure 5.** Same as Fig. 5 for January-February 2015.

**Table 2.** Global diagnostics of differences in total column ozone between separate satellite instuments and OMI-TOMS for July-August 2014 and January-February 2015 based on the non-empty bins of the data of Figs. 4 and 5. The diagnostics consists of global mean differences (%), standard deviations (%) over the bin mean differences, and percentages of SZA/latitude bins with mean differences exceeding 2 % in magnitude.

| Instrument | Mean difference (%) | | Std. dev. over bin mean differences (%) | | Percentage of bins with \|mean differnces\| > 2 %. | |
|---|---|---|---|---|---|---|
| | July-Aug. 2014 | Jan.-Feb. 2015 | July-Aug. 2014 | Jan.-Feb. 2015 | July-Aug. 2014 | Jan.-Feb. 2015 |
| GOME-2A | -1.8 | -3.5 | 1.4 | 2.4 | 50 | 69 |
| GOME-2B | 0.1 | -0.5 | 0.8 | 1.3 | 14 | 13 |
| OMPS-NM | -1.3 | 0.l | 1.3 | 0.9 | 28 | 19 |
| OMPS-NP | 1.1 | 2.0 | 1.3 | 2.0 | 23 | 47 |
| SBUV/2-TC | 1.5 | 1.3 | 0.7 | 0.9 | 30 | 22 |
| SBUV/2-NP | 1.2 | 0.6 | 0.7 | 0.5 | 16 | 0 |

however this is only for a single station. As an aside, considering the gradients over the images of Figs. 4 and 5, the bin sizes for bias estimation could be increased by factors of at least two to four for most regions, this being adaptable to each instrument.

The standard deviations of the mean differences over the bins (with removal of 2-σ outliers) range between 0.5 and 2.4 % over the different instruments (Table 2). The latter two sets of values reflect a smaller variability of the differences between SBUV/2 and OMI-TOMS. As well, the time mean biases per bin reach sizes of ~5-9 % for some datasets.

The differences between both SBUV/2 sets and OMI-TOMS are generally consistent with, and more on the lower end of, the about 1-2 % mean differences found in other studies mentioned in the introduction, with better agreement obtained here for SBUV/2-NP. They also suggest better agreement of OMI-TOMS with mean tropical and mid-latitude Northern Hemisphere ground-based data for these periods, considering Table 1. The winter polar mean differences between SBUV/2 and OMI-TOMS remain positive, reaching a maximum of ~4 % while the summer polar means for SZA<~75-80° are less than ~1 %. This seems to support the possibility of OMI-TOMS underestimation near or in the polar regions for SZA<~75-80°, but with less underestimation for the 60-90° S region than is suggested in Table. The disagreement of SBUV/2 and OMI-TOMS in the winter pole regions at even higher SZAs reaches only -1.9 % for SBUV/2-NP, while reaching -3,2 % for SBUV/2-TC, with SBUV/2-NP expected to be more accurate than SBUV/2-TC at higher SZAs. However, identifying which dataset is notably closer to the truth between particularly SBUV/2-NP and OMI-TOMS in the polar regions is somewhat obscured by the differences found in the mid-latitudes and the tropics between SBUV/2-NP, OMI-TOMS and the ground-based data. Considering results for the various datasets, during these periods only, and the results of Table 1, OMI-TOMS serving as the anchor is appropriate, with the caveat that the best reference for the polar regions and SZA>70-80° (and/or possibly ozone effective temperatures below ~200 K) may be unclear.

GOME-2B and SBUV/2-NP provide smallest differences globally and both GOME-2 instruments show strong latitudinal gradients in bias estimates for SZA>84°, which are not used in assimilation. The solar zenith angle regions below and above 70° for GOME-2A and GOME-2B give rise to an approximately bimodal distributions of differences with OMI-TOMS for July-August 2014 (and January-February 2015) centered at about -2.3 % (-4.5 %) and 1.8 % (-0.5 %) for GOME-2A and -0.5 % (-1.0 %) and 2 % (2 %) for GOME-2B. The percentages of bins with time mean differences exceeding 2 % in magnitude for the six datasets range from 0 % to 69 % (Table 2) with GOME-2B presenting the larger values; for a 1% threshold, percentages increase by factors 1.2 to 4 depending on the instrument and season. Mean differences therefore exceed the observation random error levels for a notable fraction of the datasets.

The variations in time of bias correction per bin over the two months are most often within ±1 % from the time mean, with a few bins reaching variations of ±2-3 % (e.g., Figs 6, 8, S8, S9). These 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, or scatter, of the individual six-hour means about the moving averages can be small

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


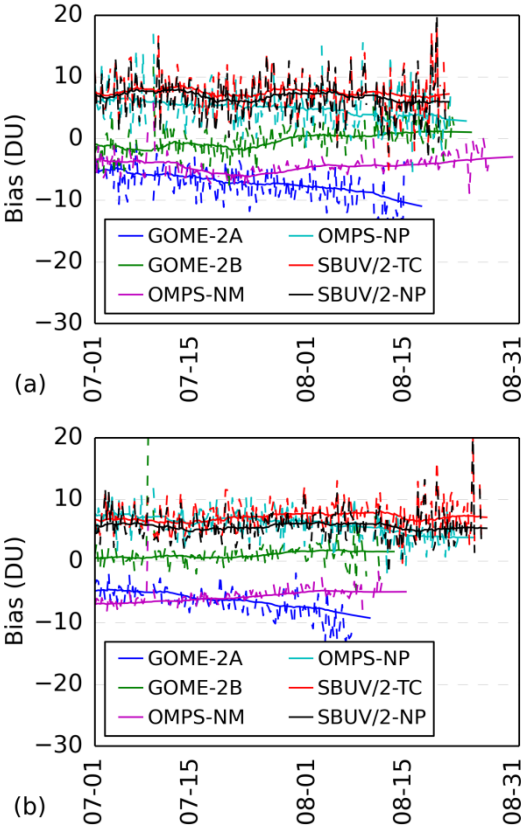


**Figure 6.** Time series of total column ozone bias corrections (DU) for July and August 2014 for GOME-2A/B, OMPS-NM/NP, and SBUV/2-TC/NP (see Fig. 5). Dashed lines show individual six-hour mean differences with OMI-TOMS, while the solid curves of the same colour show the two -week moving average bias corrections. The particular (latitude, solar zenith angle) bins plotted are (a) 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 (b) 5° wide bins centred on (62.5°N, 42.5°) for GOME-2A/B and OMPS-NM and a 10° wide bin centred on (65°N, 45°) for OMPS-NP. Time coverage for individual bins do not necessarily cover complete months.

at within ~1 % to more significant reaching at least ~3 % (Fig. 6 and Fig. S10); the number of colocations per bin ranges from a few points to a few hundred for the six-hour intervals, with larger scatter also present for large numbers per bin.

A verified cause of the larger six-hour mean differences, as seen in Fig. 6, is the occasional temporal data gaps of OMI resulting in biasing 6-hr mean differences due to a persistent time difference between colocation pair elements given the assigned ±12h colocation search window. The influence of this added bias of individual 6-hr mean differences is circumvented through outlier removal in the weighted averaging over the two-week moving window. Unaccounted biases such as from the differing handling of cloud cover are also likely contributing to the varying sizes of 6-hr mean differences about the moving averages. The number of colocated total column points applied in bias estimation for each bin over the two-week moving window typically exceeds a thousand but can be below one hundred for some few bins (e.g., Fig. S10).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


Atmospheric Chemistry and Physics Discussions

The number of colocations are significantly reduced for OMPS-NP and SBUV/2 measurements, as it would be for other profilers, and so the averaging, while not done here, could benefit from longer time windows, a wider Gaussian filter, and or larger bin sizes.

While capturing much of the differences with OMI through latitude, time, and SZA dependent corrections, one would expect remaining spatially varying residual biases (and likely temporally varying as well). The overall variations in longitude of differences with OMI-TOMS are notably weaker than those in latitude, associated to the usually larger latitudinal gradients in ozone and the solar zenith angle variation in latitude along the orbit track, with occasional longitudinally isolated areas of larger differences (e.g., such as near (10° S, 0° E) for GOME-2A/B, about (60° S, 60° W) for OMPS-NM and over Greenland for OMPS-NP as shown in Fig. S11).

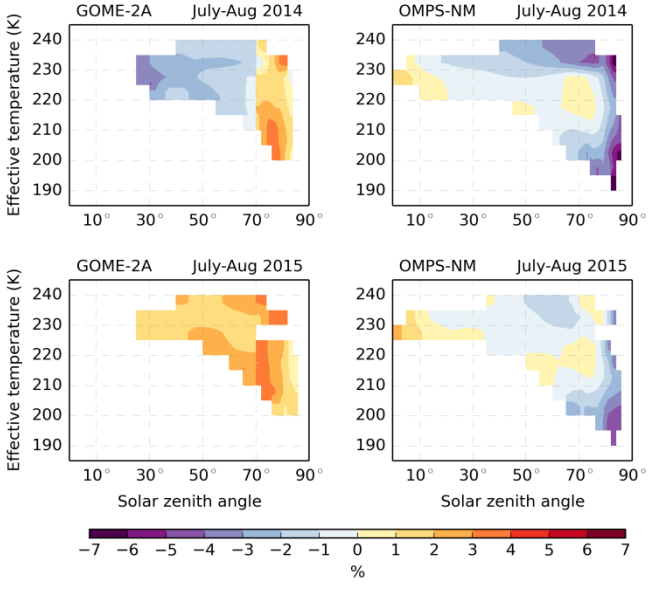


**Figure 7.** Mean total column ozone differences (%) between GOME-2A, OMPS-NM and colocated OMI-TOMS data as a function of ozone effective temperature (degrees Kelvin) and solar zenith angle (degrees) for the periods of July-August 2014 and July-August 2015. The colours blue to purple denote negative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

### 3.2.2 Variation with ozone effective temperature

An alternative parameterization for bias estimation consists in the use of time averaged differences as a function of ozone effective temperature and solar zenith angle. A motivation for a dependency on ozone effective temperature is to more directly compensate for any unaccounted temperature sensitivity of the ozone absorption coefficients used in retrievals. Here, bias estimation is made 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 to changes in temperature in addition to constant offsets. The ozone effective temperatures were calculated from the GEM with LINOZ meteorological and ozone short-term forecasts relying on the respective meteorological analyses from ECCC and ozone analyses from the assimilation of total column ozone. The resulting time mean differences from OMI as a function of ozone effective temperature and solar zenith angle can notably differ for different periods, depending on the instrument and data processing setup, as shown for GOME-2A and OMPS-NM in Fig. 7. For GOME-2A (TOMS), typical differences between 2014 and 2015 of the order of roughly 3-4 % for SZAs less than 70°. These differences are larger than the effect of gradual long term trends of about -2.2 DU, or roughly -0.6 to -0.8 %, per year for GOME-2A (DOAS) estimated by van der A et al. (2010). Differences in retrievals might be a factor in explaining these differences if not also the differing time periods. For both July-August 2014 and 2015 periods, applying their respective corrections as a function of ozone effective temperature and solar zenith angles 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 % (Fig. S12). This supports use of ozone effective temperature dependence as alternative to latitude (and time) dependence, as in other studies, with the stipulation that one accounts for otherwise remaining and notable temporal changes in some fashion where necessary. Adding an explicit sub-seasonal to seasonal dependency on time would compensate for some otherwise unaccounted time varying biases other than long-term trends. If the temporal changes may be spatially dependent, it is possible that an explicit spatial predictor (such as latitude) may be required. The variation of the bias corrections as a function of ozone effective temperature and SZA within these periods, of up to two months, is comparatively small relative to larger changes in time that can be seen from the two-week moving window bias estimations as a function of latitude and SZA as shown in Fig. 8 for GOME-2A and was also observed in some regions for other instruments. The identified differences in variation over these relatively short periods are typically within ±1 %.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


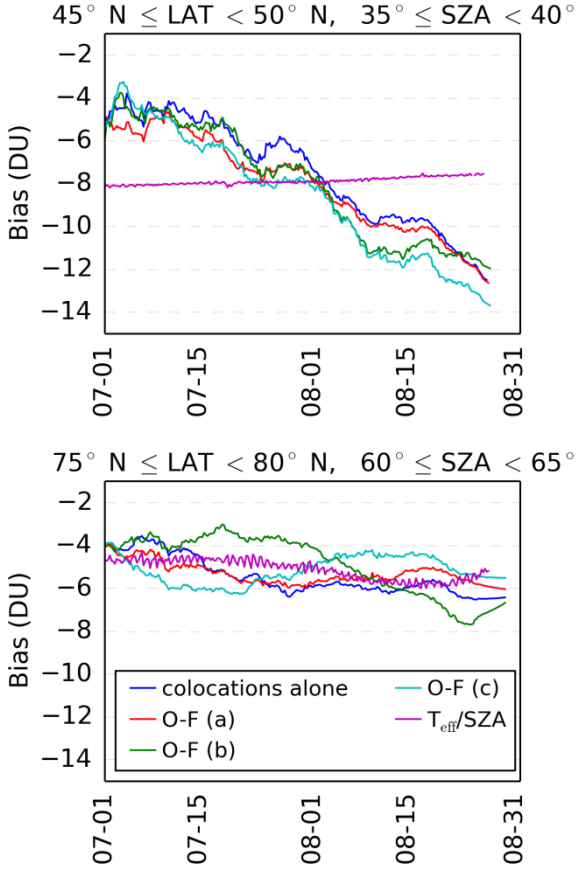


**Figure 8.** Time series of total column ozone bias corrections (DU) for two latitude/SZA bins covering July-August 2014 for GOME-2A using different bias correction methods. All cases that include colocation methods usethinned observation sets. The 'O-F' curves additionally use the differences of forecasts described in Section 2.4.2 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' curve is the result of mapping each observation that falls within the latitude/SZA bin onto the ozone effective temperature/SZA bias for July-August 2014 (shown in Fig. 7) and taking the mean of these values.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


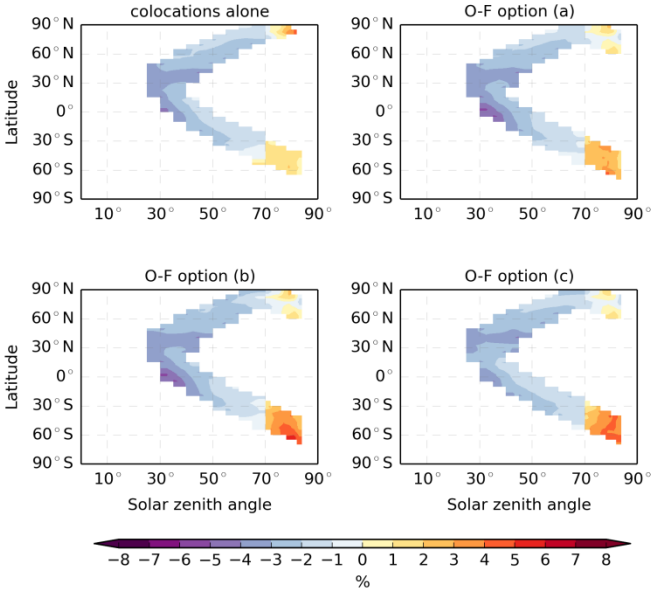


**Figure 9.** Time mean total column ozone biases (%) between GOME-2A and OMI-TOMS for the period of July-August 2014 from colocation alone and approaches (a), (b), and (c) using observation-minus-forecast differences (see Section 2.4.2). For approaches (a), (b), and (c) the forecasts were taken from the 'OMI' assimilation run (see Table 3). The bias in the 'colocations alone' panel was computed using the thinned observation data set as the thinned data set was used in the assimilation. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

### 3.2.3 Approaches based on differences with forecasts

Bias estimation and correction from differences with forecasts with the three options (a) to (c) of Section 2.4.2 was applied for GOME-2A for July-August 2014 using differences with six-hour forecasts. In this study, the forecasts used for these differences came from one of the previously run assimilation/forecasting experiments that are summarized in Table 3. However, the forecasts used for this approach could come from runs that assimilate the bias-corrected observations using the correction method considered in this section. Results from use of the forecasts from the OMI assimilation are provided in Figs. 8 and 9. Differences between the biases resulting from options (a) to (c) and colocation alone biases are within 1 % over the two month period except for a few bins, which are mostly at high SZA. For bins with larger bias estimates with the thinned data, there can be some improvement toward colocation alone results in the absence of thinning such as for option (a) at the bin centered at the latitude of 52.5° S and SZA of 65° (see Fig. 4 for the bias estimate using the unthinned data). This may be due to spatial differences being better reflected in colocation binning through additional use of the forecasts. For this period and dataset, the extension of (a) to use (b) and (c) as successive fallbacks extends the number of usable bins, this being more evident at high SZAs in the Northern Hemisphere. Results from the assimilation of non-corrected data results in differences of bias estimates from colocation alone increasing nearly always by less than 1 % for (a) and (b), with the larger increases for (b), and by roughly up to 1-2 % for (c). The latter would reflect the resultant forecast bias from the assimilated



non-corrected datasets. In the absence of ozone assimilation, the increases for (a) and (b) remain usually within 1 % for this period while increasing substantially for (c). [The implications of the other experiments are shown through the sample comparison of Fig. 8 with Figs. S13 and S14.] This demonstrates that use of differences with forecasts, and, by implication, innovations during assimilation runs, is a valid alternative to use of colocation by itself and can provide additional benefits conditional on the forecasts representing well the spatial variation in total column ozone for options (a) and (b) and being sufficiently de-biased for option (c).

**Table 3.** 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) were assimilated.

| Assimilation expt identifier | Instruments assimilated | Error std. dev. estimate |
|---|---|---|
| CTRL | None | - |
| OMI | OMI | Original |
| GOME2A | GOME-2A | Original |
| GOME2B | GOME-2B | Original |
| OMPSNM | OMPS-NM | Original |
| G2AB+NM | GOME-2A/B, OMPS-NM | Original |
| ALLTC | GOME-2A/B, OMPS-NM, OMI | Original |
| GOME2A bc | GOME-2A* | Original |
| GOME2B bc | GOME-2B* | Original |
| OMPSNM bc | OMPS-NM* | Original |
| G2AB+NM bc | GOME-2A*/B*, OMPS-NM* | Original |
| ALLTC bc | GOME-2A*/B*, OMPS-NM*, OMI | Original |
| MLS+OMI | MLS, OMI | Original |
| G2AB+NM bc us | GOME-2A*/B*, OMPS-NM* | Updated |

*denotes bias-corrected observations

### 3.3 Evaluation of ozone short-term forecasts

Successive short-term forecasts were generated from assimilations performed from 29 June to 31 August 2014 with different sets of total column ozone observations with and without bias correction, as well one experiment with assimilation of MLS ozone profiles. These are compared to each other and to the case without assimilation. All results with bias correction have bias estimates obtained from the colocation approach without use of forecast differences. All presented assimilation experiments except for one were conducted with the original background and observation error standard deviations. The short labels specified in the figure legends as of Fig. 10 identifying the different experiments are described in Table 3.

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

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

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

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018




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

where $O_i$, $F_i$, and $C_i$ respectively denote observations, forecasts and climatological values at the observation locations. These were applied to forecasts of both total column ozone and the three-dimensional (3D) ozone field over the July-August period. The ACC, computed over different regions, provides a measure of the spatio-temporal anomalies between an ensemble (or time series) of forecasts and the verifying data (or analysis) as compared to their respective differences to a reference or climatological field, specified here as the 3D temporal mean of the no assimilation case. While not applied here, analyses on the model grid can replace observations in the ACC calculation especially when evaluating the change in ACC as a function of forecast length, in which case the cosine of latitude would be included as scaling factor for the elements of each summation (WMO, 1992).

### 3.3.1 Impact on total column ozone forecasts

Global differences, over the two month period, of Brewer and Dobson total column ozone with short-term forecasts following assimilation without bias correction span the range of near zero to about 2.3 %, nearly all being positive, as compared to about -1.8 % for Brewers and -0.8 % for Dobsons in the absence of assimilation (Table 4). The mean differences shown in Table 4 for the assimilation cases without bias correction are consistent with the global mean differences from the colocation comparison of the satellite and ground-based observations for July-August 2014 provided in the first paragraph of Section 3.1. The mean differences with forecasts from the assimilation of the uncorrected provisional OMPS-NM are only slightly larger than direct differences with OMPS-NM indicated in section 2.2.3 for more recent product versions. Bias correction reduces the global differences to less than 0.3 % in size for the various cases, be it from assimilation of individual or multiple data sources. The assimilating GOME-2A observations alone without bias correction actually increases the absolute global mean differences relative to the no assimilation case. Introducing assimilation reduces the standard deviations from 3.4-3.8 % to ~2.3-2.8 %. The reduction of standard deviations from bias correction is less than 0.35 %. These smaller reductions in standard deviations are limited by the error level of the observations and forecasts and the spatio-temporal variability of the biases about the global mean differences as shown in the following sentence. Brewer and Dobson error standard deviations of 1.5 %, a forecast error standard deviation of 0.5 % (likely an underestimate) and, considering Table 1, a standard deviation over the stations time mean differences of 1.5 % give a resultant standard deviation of 2.2 % from the square root of the sum of squares, quite close to the numbers from Table 4.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


Atmospheric Chemistry and Physics Discussions

**Table 4.** 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 were applied in the assimilations. The Dobson measurements used were adjusted as a function of the ozone effective temperature (see Section 2.2.4). The uncertainties denote the square root of the sample variances of the mean differences or difference standard deviations. The data from the two Antarctic stations have been included here even though their mean differences with OMI are outliers relative to most mean differnces (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. 10 to 13 for the July-August 2014 period. Figure 10 (see also Figs. S15 and S16) displays the time mean differences which seem more prominent in the tropical region for the control case, i.e., without assimilation, reaching above 10 %. The separate assimilations of all column ozone sources with bias correction (labelled 'ALLTC bc') and of OMI yield mean differences with OMI usually within ±1 %. The absence of bias correction results in an overall time-mean reduction of column ozone by 1-3 % due to the significance of the GOME-2A and OMPS-NM biases. The results of assimilations of various individual and combined sources in terms of mean differences, standard deviations, and anomaly correlation, are provided for different latitude bands in Figs. 11 and 12. These show that the first order improvements stem from assimilation in general, while bias corrections and updating of error standard deviations (case 'G2AB+NM bc us' of Fig. 12) comparatively imply second order changes. The GOME-2A and OMPS-NM datasets shows the largest reductions in mean differences from bias correction as would be expected from Fig. 2. Figure 12 shows that the updated error variances have resulted in a reduction in impact in some areas due to the larger reduction of the background error variances relative to the change in the observation error variances. Adding the assimilation of MLS ozone profiles to that of OMI-TOMS has small varying effects within 1 % on total column ozone mean differences as a function of latitude except in Antarctica at just over 1.5 % (Fig. S17). The latter seems to support an underestimation of OMI-TOMS in Antarctica by at least that amount for July. The added MLS assimilation also suggests that any OMI-TOMS underestimation in the Arctic would be much less for that period.

For 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, in



addition to its low bias. As well, the ACC is the diagnostic that demonstrates more marked improvements in multiple sensor assimilation as compared to OMI-TOMS assimilation. The advantage of multiple sensor assimilation is, therefore, more notable in increasing the quality of the pattern and variation of the forecast fields. As anomaly correlation coefficients in assimilation typically compare forecasts with analyses instead of observations, OMI data in this case, it was verified, for completeness, that both give similar results (Fig. 11 compared to Fig. S18). It was also verified that the choice of the reference field (climatology) did not change the overall qualitative results; switching from the time average of the no assimilation case (CTRL) to the 2D ozone climatology of Fortuin and Kelder (1998) results in changes to the ACC for the case without assimilation of within ~0.l and of much less for the column ozone assimilation cases (Fig. S19).

In the absence of assimilation, the mean differences in the extra-tropics have opposite sign to those in the tropics. This stems, at least partly if not entirely, from the differences between the observations and tendency of the forecast in the absence of assimilation to move toward the ozone model equilibrium state. This is also seen by the temporal changes of the mean differences of Fig. 13 over the two month period having opposite tendencies for the extra-tropics and the tropics. The temporal changes are believed to result from a long spinup period in moving 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 likely not have improved the ACC of the control case, as implied by Fig. 13, and would have increased its mean observation minus forecast differences. The changes in time of Fig. 13 also reflect the high predictability of ozone medium range forecasts, with an increase in total column ozone forecast regional mean error of less than 5 % in ten days, this in addition to the relevance of assimilation. Both the temporally averaged and time varying mean differences of forecasts with OMI-TOMS are reduced to within 1 % over the latitude ranges where satellite data are assimilated for the various cases with bias correction, with the results for GOME-2A assimilation being the exception, slightly exceeding 1 %. Figures 11 and 13 both show the strongest improvement from bias correction of multiple sensor assimilation (‘ALLTC’ and ‘ALLTC bc’) in the northern extra-tropics.

The regions of most significant temporally averaged impact from assimilation are near the Antarctic polar vortex and in the tropics. 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 in the winter poles in the absence of heterogeneous chemistry in LINOZ.

The deterioration of the ACC with time in Fig. 13, and the low time mean ACC in the tropics in Figs. 11 and 12, 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

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

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


Atmospheric Chemistry and Physics Discussions

where *O* are observations and *C* a climatology used to evaluate forecasts *F* and $\sigma$ are the standard deviations of the quantity in its brackets. In the case when assimilation is not performed, during the time period when the ACC deteriorates rapidly in the tropics (as seen in Fig. 13), *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 also seen in the other regions though. 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 of roughly the same size and much larger than the third term in regions where measurements are assimilated, resulting in ACC values notably closer to unity, this being observed in all latitudinal regions (Fig. S20).

While the *σ(O-F)* values at mid-latitudes and in the polar regions can be smaller or larger than in the tropics when no assimilation is performed (Fig. 11), their sizes relative to both *σ(O-C)* and *σ(F-C)* are smaller in the extra-tropics (Fig. S20). The larger *σ(O-C)* and *σ(F-C)* in the extra-tropics result in larger ACCs as compared to the tropics. The smaller relative size of *σ(O-F)* in the extra-tropics stems from both the observations and forecasts being correlated to the spatio-temporal changes in weather patterns.

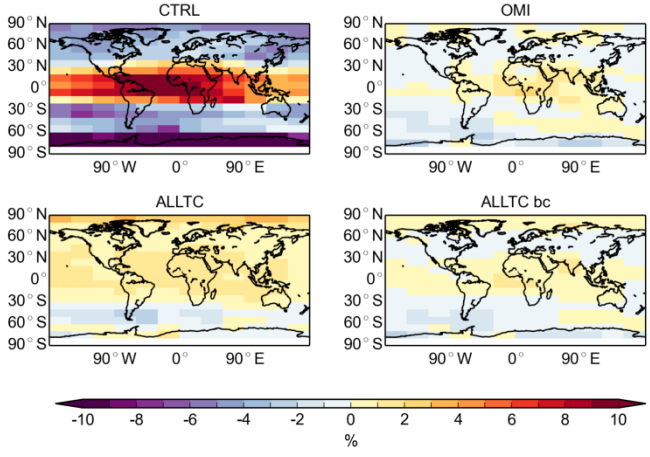


**Figure 10.** Mean total column ozone differences (%) between OMI-TOMS measurements and short-term forecasts as a function of spatial location (degrees) for July-August 2014. The plot titles indicate the assimilation run (see Table 3 for description). In the case of no assimilation, many of the values south of 60º S exceed the lower limit of the colour bar, in some instances with values as low as -30%. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


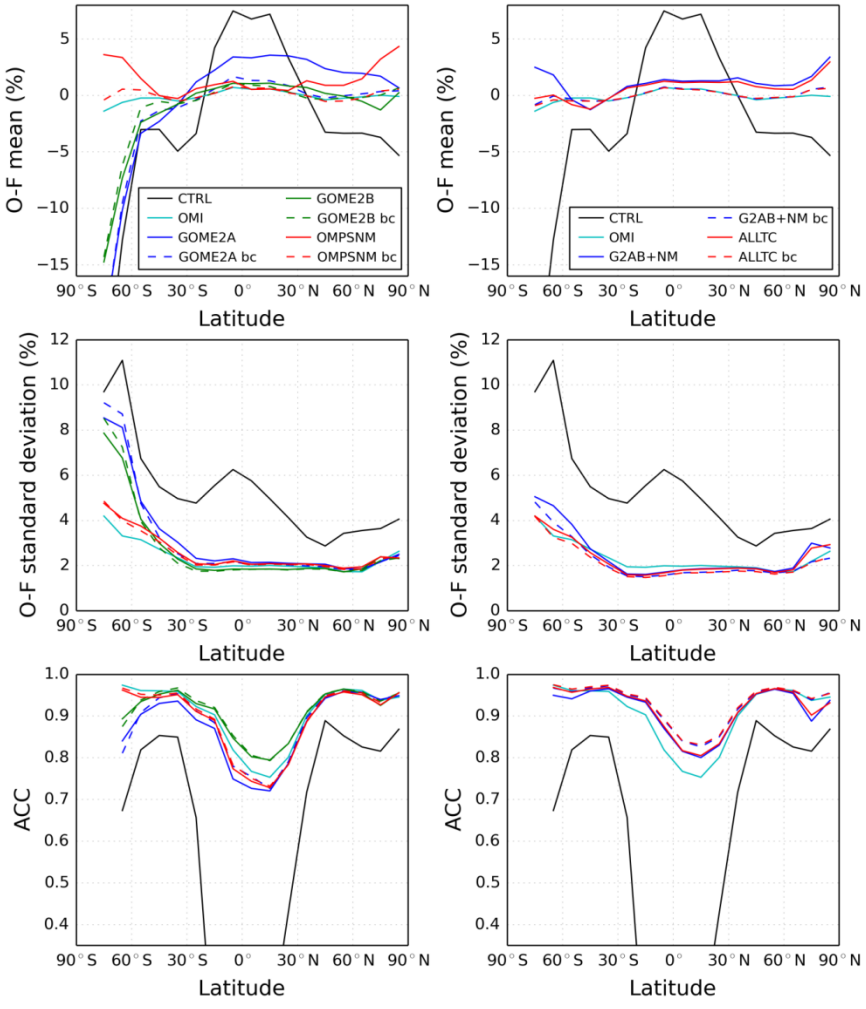


**Figure 11.** Zonal mean total column ozone statistics of mean differences (%), standard deviations (%), and anomaly correlation coefficients (ACC; unitless) as a function of latitude (degrees) for the comparison between OMI-TOMS measurements and short-term forecasts for July-August 2014. The legends in the top plots indicate the assimilation runs (see Table 3 for description) and apply to all plots in the same column.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


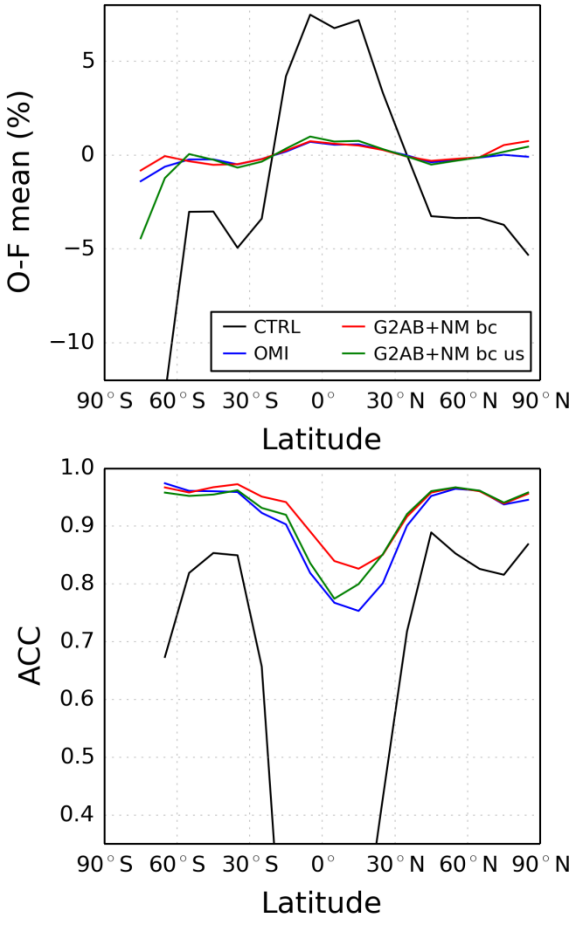


**Figure 12.** Zonal mean total column ozone differences (%) and anomaly correlaton coefficients (ACC; unitless) as a function of latitude (degrees) between OMI-TOMS observations and short-term forecasts for July-August 2014 showing the effect of using the updated sets of observation and background error variances in the assimilation. The legend indicates the assimilation run (see Table 3 for description).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


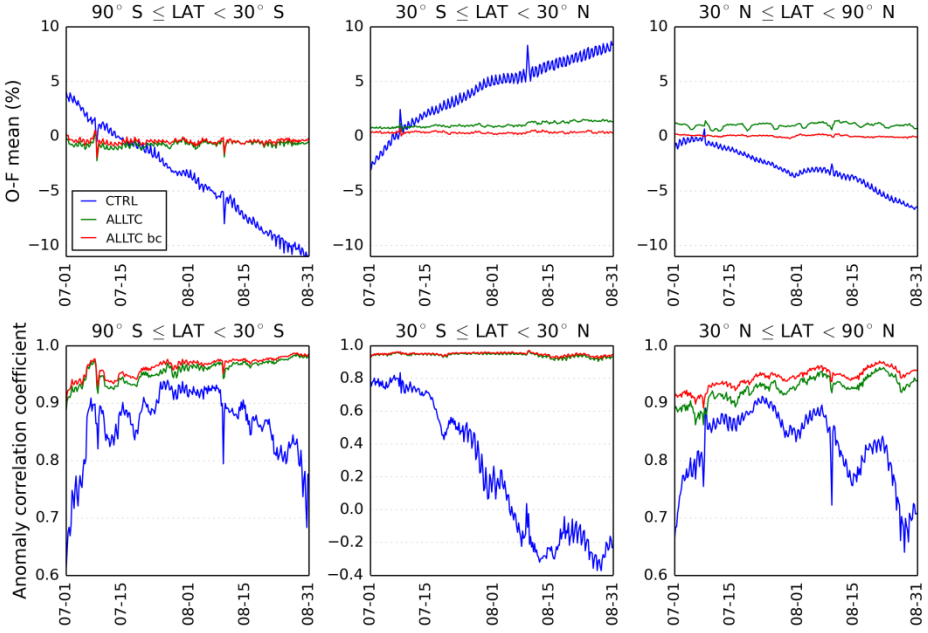


**Figure 13.** Zonal mean differences (%) and anomaly correlation coefficients (unitless) for total coumn ozone between OMI-TOMS observations and short-term forecasts as a function of time (date). Results are shown for the case without assimilation as well as with the assimilation of OMI, GOME-2A/B, 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.

### 3.3.2 Impact on the vertical structure of the ozone field

The short-term ozone field forecasts from a few assimilation experiments are compared to MLS and ozonesonde profiles using the same metrics and for the same period as above. Significant impact of total column ozone assimilation on forecasts is seen in Figs. 14 to 17, usually below 10-20 hPa, with the largest influences in the lower half of the stratosphere. A reduction in impact above 2-5 hPa stems at least partly from the short photochemical time scales at these levels reflected in the ozone forecast model. Deteriorations are also observed in the lower stratosphere and upper troposphere regions in some panels, e.g., the mean observation minus forecast differences in the Northern Hemisphere below about 70 hPa for Fig. 14. The presence of improvements and deteriorations as a function of vertical level for the mean differences of Figs. 14 (upper panels) and 15 seems to reflect, to some degree, a vertically weighted profile shift needed for the non-assimilated case to better agree with the total column ozone observations. This is consistent with not having negative background vertical error correlations. The most significant improvements from total column ozone assimilation alone are found in the tropics for the anomaly correlation coefficients throughout the lower stratosphere, as seen in Fig. 14, followed by the time mean differences in the same region. This partly stems from the relative change in vertical distribution of mixing ratio increments for each total column measurement being proportional to the product of the background error covariance matrix and the model pressure layer thicknesses vector. The former includes not only the influence of the background error variances and vertical



correlations, but also the influence of the background horizontal error correlations increasing away from the surface, which would impact local profile increments considering the proximity of neighbouring total column measurements. The relaxation time scale of two days toward climatology in the lower troposphere would also have some impact on the forecasts moving away from the influence of the observations. Figures 14 and 16 also show, in support of Fig. 12, that the updated error variances have resulted in a slight reduction in impact from assimilation in the upper troposphere and lower stratosphere.

While some improvements in the vertical structure can result from assimilation of total column ozone, the greatest improvements are seen, in Figs. 15 and 17, when adding the assimilation of ozone profile measurements as done here with Aura-MLS (see also Struthers et al., 2002, as another example), which occurs even though the horizontal density of profiles is much smaller than from total column mapping instruments (e.g., Fig. S1).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


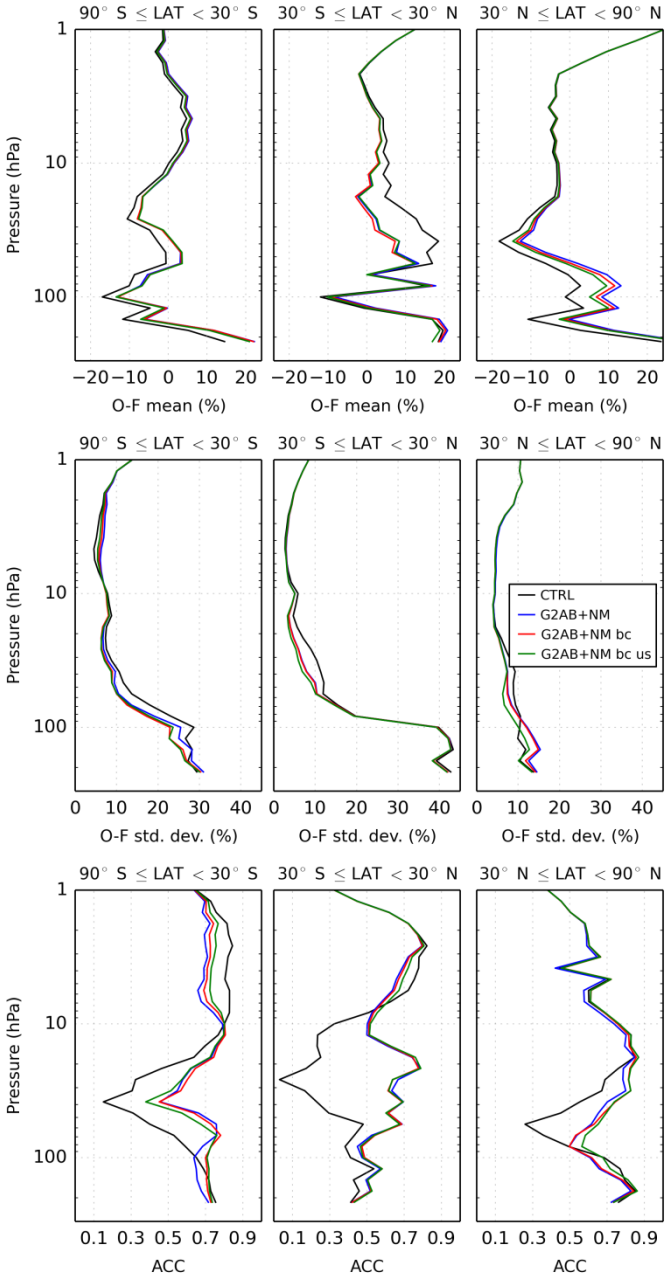


**Figure 14.** Profiles of mean differences (%), standard deviations of differences (%), and anomaly correlation coeffients (ACC; unitless) between MLS observations and short-term forecasts over July-August 2014 and as a function of pressure (hPa). The legend indicates the assimilation run (see Table 3 for description).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


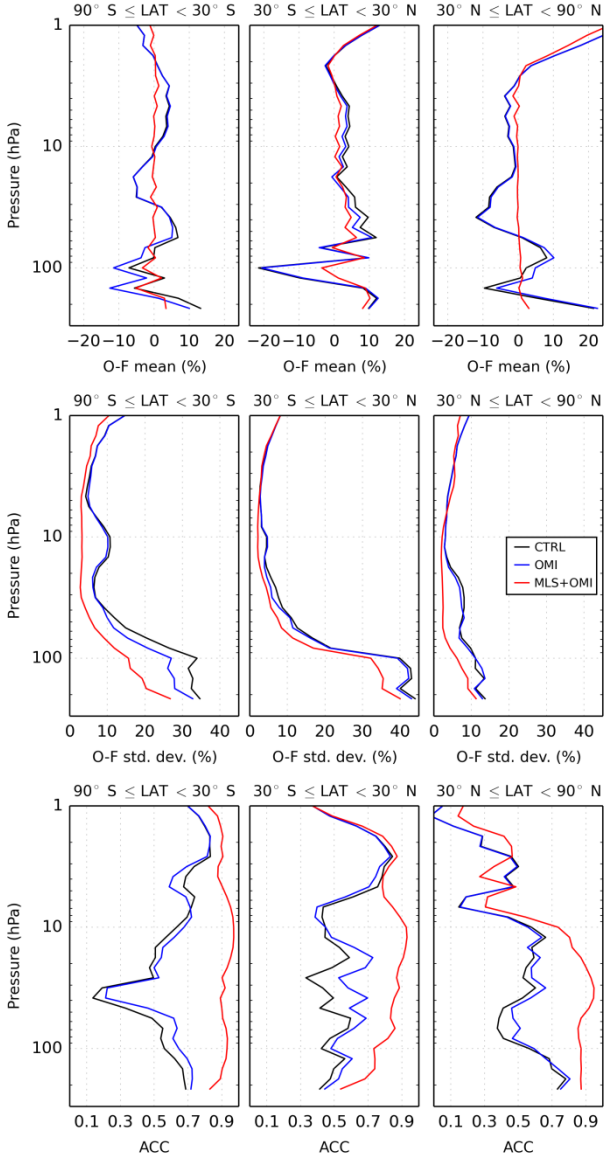


**Figure 15.** Similar to Fig. 14 but only covering 1-23 July 2014. This shorter period in this figure is imposed by the 'MLS+OMI' assimilation having been conducted only for this duration prior to local computing system changes.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


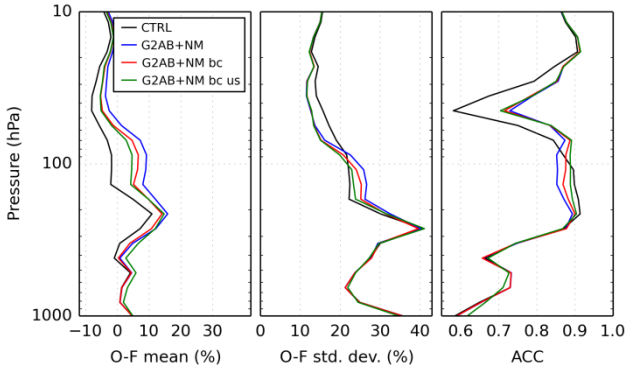


**Figure 16.** Profiles of global mean and standard deviation differences (%), and anomaly correlation coeffients (unitless) between ozonesonde observations and short-term forecasts over July-August 2014 and as a function of pressure (hPa). The legend indicates the assimilation run (see Table 3 for description).

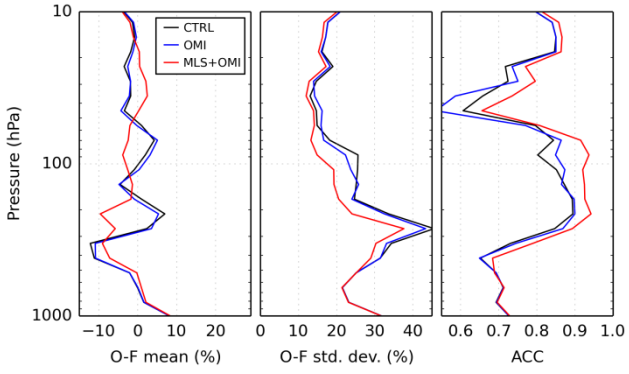


**Figure 17.** Similar to Fig. 16 but over 1-23 July 2014.

## 4 Conclusions

Bias correction of total column ozone data from satellite instruments was performed using a few different approaches relying on bias estimation as a function of latitude and solar zenith angle binning with a two-week moving time window. The approach consisting of applying colocation by itself was compared to variants that used short-term forecasts. The differences in corrected biases between these approaches were nearly always under 1 %. These were also compared to the bias estimation as function of ozone effective temperature and solar zenith angle without the two-week moving time window. While the time-averaged bias corrections from this scheme were similar to the other approaches, the relevance of explicitly including time dependence was shown to potentially matter, this likely depending on the instrument and/or retrieval process.

The anchor data source from which bias estimates were generated was the OMI-TOMS data product. The quality of the OMI-TOMS data is supported by other publications in addition to the evaluations on two month periods in this paper through comparisons to ground-based Brewer and Dobson spectrophotometers and filter ozonometers. Comparison of OMI-



TOMS with the ground-based data yielded regional and global mean differences within 1 %, except in the polar regions. Similar to slightly larger differences were found with SBUV/2 data. While no adjustments were made in this paper to the OMI-TOMS data, the choice of the anchor does not exclude the possibility of applying pre-determined corrections to its data when deemed necessary.

Prior to bias correction for the assimilated data, the GOME-2B (TOMS) products were found to be closest to OMI-TOMS, and thus to the ground-based data, with GOME-2A showing the largest differences. The differences for the provisional OMPS-NM products were typically in between these two cases.

Three-dimensional variational assimilations of total column ozone data measured from the satellite instruments have demonstrated the capability of reducing short-term forecast errors of total column ozone to within 1 % of the OMI-TOMS data in the latitude regions covered by the assimilated data when including bias corrections as applied in the paper. This applies to assimilation of single and multiple data sources, except for assimilation of GOME-2A alone with residuals slightly exceeding 1 % at some latitudes. For the assimilation of all instruments, while possible cancellation of different errors from different instruments is a factor in reducing forecast biases, harmonizing the differing datasets through bias correction better ensures achieving the target reduction in residual bias. 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; this might vary depending on the ozone forecast model. This was evident not only in improvements of mean differences, but also of anomaly correction coefficients. The most notable improvement from total column ozone assimilation without the profile data was obtained for the anomaly correlation coefficients in the tropical lower stratosphere. Based on the shorter assimilation with Aura-MLS, the addition of this profile data did not significantly alter total column ozone short-term forecasts in the regions where the total column data was assimilated; it can be surmised that it would improve the total column ozone forecasts in the winter pole regions when there are no total column ozone data available. As one would or might expect and as has been illustrated, the added assimilation of good quality profile data is important in assuring a better quality of the vertical structure of the ozone forecast field even if the spatial density of profile measurements over six hour periods is much less than that of the total column data.

*Code and data availability*. 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 Fortran-based 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 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 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. 4.



**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

Adams, C., Bourassa, A. E., Sofieva, V., Froidevaux, L., McLinden, C. A., Hubert, D., Lambert, J. C., Sioris, C. E., and Degenstein, D. A.: Assessment of Odin-OSIRIS ozone measurements from 2001 to the present using MLS, GOMOS, and ozone sondes., Atmos. Meas. Tech. 7, 49-64, doi:10.5194/amt-7-49-2014, 2014.

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., Liu, C., Shi. R. and Gao, W.: Comparison of Suomi-NPP OMPS total column ozone with Brewer and Dobson spectrophotometers measurements, Front. Earth Sci., 9(3), 369–380, doi:10.1007/s11707-014-0480-5, 2015

Bai, K., Liu, C., Shi, R., and Gao, W.: Validation of the Suomi NPP Ozone Mapping and Profiler Suite total column ozone using Brewer and Dobson spectrophotometers, Atmos. Meas. Tech. Discuss., 6, 4577-4605, 2013.

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., 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), 2007a.

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), 2007b.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Q. J. R. Meteorol. Soc. 134, 1951–1970 , doi: 10.1002/qj.339, 2008.

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.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 – Metop's second generation sensor for operational ozone monitoring, ESA Bull.-Eur. Space, 102, 28–36, 2000.

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.

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.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background, and analysis-error statistics in observation space. Q. J. R. Meteorol. Soc., 131, 3385-3396, 2005.

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., Kerr, J. B., McElroy, C. T., Wardle, D. I., Savastiouk, V., and Grajnar, T. S.: The Brewer reference triad, Geophys. Res. Lett., 32(20), 2005.

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. 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.

Froidevaux, L., and Coauthors: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, J. Geophys., Res., 113, D15S20, doi:10.1029/2007JD008771, 24 pp., 2008.

Gaspari, G., and Cohn, S.: Construction of correlation functions in two and three dimensions, Q. J. R. Meteorol. Soc., 125, 723-757, 1995.

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.

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]

Jackson, D.R., Keil, M., and Devenish, B.J.: Use of Canadian Quick covariances in the Met Office data assimilation system, Q. J. R. Meteorol. Soc., DOI:10.1002/qj, 2008.

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., 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.

Levelt, P. F., van den Oord, G. H. ., Dobber, M R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J.O. V., and Saari, H.: The Ozone Monitoring Instrument, IEEE Trans. Geosci. Remote Sens., 44, 1093–1101, 2006.

Livesey, N. J., Van Snyder, W., Read, W. J., and Wagner, P. A.: Retrieval algorithms for the EOS Microwave Limb Sounder (MLS), IEEE Trans. Geosci. Remote Sens., 44, 1144–1155, doi:10.1109/TGRS.2006.872327, 2006.

Livesey, N. J., and Coauthors: Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS): Version 3.3 and 3.4 Level 2 data quality and description document, JPL D-33509, Version 3.3x/3.4x-1.1, Jet Propulsion Laboratory, 158 pp., May 14, 2013. [Available from: http://mls.jpl.nasa.gov/data/v3_data_quality_document.pdf.]

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]

Massart, S., Clerbaux, C., Cariolle, D., Piacentini, A., Turquety, S., and Hadji-Lazaro, J.: First steps towards the assimilation of IASI ozone data into the MOCAGE-PALM system, Atmos. Chem. Phys., 9, 5073-5091, 2009.

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., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Albinana, A. P.: GOME-2 on MetOp, Proc. of The 2006 EUMETSAT Meteorological Satellite Conference, Helsinki, Finland, 2006.

Ozone Monitoring Instrument (OMI) Near Real Time Data User's Guide, 2010. [Available from: https://macuv.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.]

Polavarapu, S., Ren, S., Rochon, Y., Sankey, D., Ek, N., Koshyk, J., and Tarasick, D.: Data assimilation with the Canadian Middle Atmosphere Model, Atmos.-Ocean, 43. 77-100. 2005.

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.

Smit, M. A. H., and Sträter, W.: Jülich ozone sonde intercomparisons 2000 (JOSIE), Report No. 1225, World Meteorol. Org., Global Weather Watch, Geneva, 2004. [Available online at http://library.wmo.int/pmb_ged/wmo-td_1225.pdf]

SPARC: Stratospheric Processes And their Role in Climate (SPARC), Assessment of trends in the vertical distribution of ozone, SPARC report No. 1, in: WMO Global Ozone Res. and Monit. Project Rep. 43, edited by: Harris, N., Hudson, R., and Phillips, C., World Meteorol. Org., Geneva, 1998.

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.

Struthers, H., Brugge, R., Lahoz, W. A., O'Neil, A., and Swinbank, R.: Assimilation of ozone profiles and total column measurements into a global general circulation model, J. Geophys., Res., 10(D20), 4438-4451, doi:10.1029/2001JD000957, 2002

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.

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 online at 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.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-614
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 20 September 2018


Atmospheric Chemistry and Physics Discussions

## 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 3.1) 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 variation of the station 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 2.2.4); 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 separate satellite instuments and OMI-TOMS for July-August 2014 and January-February 2015 based on the non-empty bins of the data of Figs. 4 and 5. The diagnostics consists of global mean differences (%), standard deviations (%) over the bin mean differences, and percentages of SZA/latitude bins with mean differences exceeding 2 % in magnitude.

**Table 3.** 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) were assimilated.

**Table 4.** 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 were applied in the assimilations. The Dobson measurements used were adjusted as a function of the ozone effective temperature (see Section 2.2.4). The uncertainties denote the square root of the sample variances of the mean differences or difference standard deviations. The data from the two Antarctic stations have been included here even though their mean differences with OMI are outliers relative to most mean differnces (Tables S1 and S2).

## Figure captions

**Figure 1.** Updated set of applied observation errror standard deviation estimates (%) for GOME-2A, GOME-2B, OMPS-NM and OMI total column measurements as a function of solar zenith angle band centres (degrees), for the combined period of July and August 2014. A 2% constant served in specifying the first set of applied error standard deviation estimates.

**Figure 2.** Mean total column ozone differences (%) between GOME-2A/B, OMPS-NM, and colocated OMI-TOMS data as a function of solar zenith angle (degrees) for the Northern and Sourthern Hemispheres for August 2014. Differences were computed from observation colocations. For the GOME-2 instruments, separate continuous difference regions are specified about 70° (see text in Section 2.4.1 for additional information). Constant difference extrapolation is applied from bin midpoints at the edges of distinct regions.

**Figure 3.** 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 4.** Mean total column ozone differences (%) between GOME-2A/B, OMPS-NM/NP and colocated OMI-TOMS data for the period of July-August 2014. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

**Figure 5.** Same as Fig. 4 for January-February 2015.

**Figure 6.** Time series of total column ozone bias corrections (DU) for July and August 2014 for GOME-2A/B, and OMPS-NM/NP. Dashed 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 (a) 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 (b) 5° wide bins centred on (62.5°N, 42.5°) for GOME-2A/B and OMPS-NM and a 10° wide bin centred on (65°N, 45°) for OMPS-NP. Time coverage for individual bins do not necessarily cover complete months.



**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.** 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 usethinned observation sets. The 'O-F' curves additionally use the differences of forecasts described in Section 2.4.2 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' curve is the result of mapping each observation that falls within the latitude/SZA bin onto the ozone effective temperature/SZA bias for July-August 2014 (shown in Fig. 7) and taking the mean of these values.

**Figure 9.** Time mean total column ozone biases (%) between GOME-2A and OMI-TOMS for the period of July-August 2014 from colocation alone and approaches (a), (b), and (c) using observation-minus-forecast differences (see Section 2.4.2). For approaches (a), (b), and (c) the forecasts were taken from the 'OMI' assimilation run (see Table 3). The bias in the 'colocations alone' panel was computed using the thinned observation data set as the thinned data set was used in the assimilation. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

**Figure 10.** Mean total column ozone differences (%) between OMI-TOMS measurements and short-term forecasts as a function of spatial location for July-August 2014. The plot titles indicate the assimilation run (see Table 3 for description). In the case of no assimilation, many of the values south of 60º S exceed the lower limit of the colour bar, in some instances with values as low as -30%. The colours blue to purple denote negative differences and the colours yellow to red refer to positive differences.

**Figure 11.** 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 3 for description) and apply to all plots in the same column.

**Figure 12.** Zonal mean total column ozone differences (%) and anomaly correlaton coefficients (ACC; unitless) as a function of latitude (degrees) between OMI-TOMS observations and short-term forecasts for July-August 2014 showing the effect of using the updated sets of observation and background error variances in the assimilation. The legend indicates the assimilation run (see Table 3 for description).

**Figure 13.** 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 14.** Profiles of mean differences (%), standard diviations of differences (%), and anomaly correlation coeffients (ACC; unitless) between MLS observations and short-term forecasts over July-August 2014 and as a function of pressure (hPa). The legend indicates the assimilation run (see Table 3 for description).

**Figure 15.** **S**imilar to Fig. 14 but only covering 1-23 July 2014. This shorter period stems from the 'MLS+OMI' assimilation having been conducted only for this duration prior to computing system changes.

**Figure 16.** Profiles of global mean and standard deviation differences (%), and anomaly correlation coeffients (unitless) between ozonesonde observations and short-term forecasts over July-August 2014 and as a function of pressure (hPa). The legend indicates the assimilation run (see Table 3 for description).

**Figure 17.** Similar to Fig. 16 but over 1-23 July 2014. For this specific figure, the comparison to ozonesondes for the 'MLS+OMI' assimilation case was done using six-hour forecasts instead of nearest forecasts in the three to nine hour forecast periods at intervals of 45 minutes as the latter were not available for that comparison. This by itself would have deteriorated the results shown for 'MLS+OMI' which are still better that the other cases.