International Journal of Remote Sensing

An assessment of long-term ozone trend uncertainties using Total Ozone Mapping Spectrometers (TOMS)

G. Jaross; S. L. Taylor; C. G. Wellemeyer; R. P. Cebula; L. -K. Huang; R. S. Stolarski; R. D. McPeters

Raytheon ITSS, Lanham, MD, USA
Code 916 NASA/Goddard Space Flight Center, Greenbelt, MD, USA

Online publication date: 26 November 2010

To cite this Article

To link to this Article: DOI: 10.1080/01431160304971
URL: http://dx.doi.org/10.1080/01431160304971

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
An assessment of long-term ozone trend uncertainties using Total Ozone Mapping Spectrometers (TOMS)

G. JAROSS, S. L. TAYLOR, C. G. WELLEMEYER, R. P. CEBULA, L.-K. HUANG
Raytheon ITSS, 4500 Forbes Boulevard, Lanham, MD 20706, USA†

R. S. STOLARSKI and R. D. McPETERS
Code 916 NASA/Goddard Space Flight Center, Greenbelt, MD 20771, USA

Abstract. We provide here an estimate of the uncertainties that would result when combining multiple Total Ozone Mapping Spectrometer (TOMS) datasets to produce a long-term column ozone record. We use calibration results and various post-launch validations to estimate calibration-related uncertainties for each of the instruments. Two basic methods for combining the individual data records are examined. We assess the combined uncertainty in the global column ozone data record for the period 1978–2000 for both methods using a Monte Carlo model with basic Gaussian statistics. In the case where TOMS data are combined without relative adjustments we estimate a trend uncertainty of slightly less than 1% per decade in the column ozone time dependence. The recently re-calibrated Solar Backscattered Ultraviolet 2 (SBUV2) datasets can be utilized to bridge a gap between TOMS datasets that occurred in the mid-1990s. With our current understanding of the sensor comparisons, we estimate a small improvement in the combined TOMS trend uncertainty. A more significant improvement comes from considering an extended TOMS data record, out to 2007. We estimate an extended ozone trend uncertainty as low as 0.5% per decade.

1. Introduction

The standard for determining the long-term trend in global column ozone amount, based upon model predictions, is 1% over a decade. This sets a rather stringent goal for the calibration of the succession of sensors which have been making ozone measurements for more than two decades. In addition, the downward ozone trend is predicted to cease followed by a slow recovery. This will place further restrictions on the accuracy of the calibration of satellite sensors measuring total ozone. We want to be able to determine, at the earliest possible time, when a recovery can be said to have begun. This requires an understanding of the sources of uncertainty in the determination of the time-dependent calibration of each instrument as well as their intercalibration.

†The authors are currently employed by Science Systems and Applications, Inc., 10210 Greenbelt Road, Lanham, MD 20706, USA.

Specific predictions regarding ozone loss and recovery and global dynamics favour the use of global rather than ‘spot’ measurements provided by ground stations and sondes. The data record of sensors utilizing the backscatter ultraviolet (BUV) technique is unique in terms of its global coverage and its length. The BUV instruments measure a normalized radiance that is proportional to the directional albedo of the Earth. The normalized radiance is the Earth radiance in the instrument field of view divided by the solar irradiance. This quantity, which is the primary input for the ozone retrieval algorithms, is invariant as the sensor throughput and solar irradiance change.

The BUV sensors, shown in figure 1, are launched into polar, Sun-synchronous orbits with local equator crossing times near noon. The Total Ozone Mapping Spectrometers (TOMS) employ a scan mirror to provide cross-track coverage of the Earth’s surface with repeated measurements every 24 hs. The solar backscatter ultraviolet (SBUV) sensors have a repeat rate of approximately three days because their 11° field of view is fixed at nadir. The differing TOMS and SBUV wavelength range leads to some small differences in column ozone retrieval algorithms, but these do not result in significant column ozone trend differences. Improvements and updates to the algorithm are generally applied to all TOMS and SBUV datasets.

The TOMS on Nimbus 7 (TOMS/N7), with a 14½ year record, seems the obvious starting point for a long-term ozone dataset. The datasets from TOMS/N7 and TOMS on Meteor 3 (TOMS/M3) have been combined (McPeters et al. 1996a) to yield a 16-year column ozone record with 1% per decade trend uncertainties over

Figure 1. The timeline of BUV sensors operated by the US Government is shown. Note the gap in the TOMS record between 1994 and 1996. That gap is bridged by the SBUV2 sensors on NOAA-9 and NOAA-11.
much of the globe. However, none of the subsequent TOMS has operated much longer than four years. More importantly, as seen in figure 1, a one and a half year gap exists between TOMS/M3 and the TOMS on Earth Probe (TOMS/EP). This gap is a major source of uncertainty in a combined TOMS data record. How best to deal with this gap is an underlying theme of this paper.

The SBUV and SBUV2 data record, which has remained unbroken since 1978, may represent a viable alternative to the combined TOMS record. However, combining their datasets is a more formidable task than for TOMS because of the slow drift of National Oceanic and Atmospheric Administration (NOAA) satellite local equator crossing times. The orbits of NOAA-9, NOAA-11 and NOAA-14 all drift later from their initial 1330 h to 1430 h crossing times. The result, even for a well calibrated sensor, is a dataset from which long-term trends are difficult to derive. Studies with TOMS/M3 (Seftor et al. 1997), which had a 212-day orbit precession period, indicated that retrieved column ozone amounts can vary as much as 5–10% as sensor viewing conditions change. The effective precession period of current NOAA satellites is approximately 24 years, making the separation of true ozone trends from orbital effects challenging. Consequently, most efforts to date have focused on using TOMS data for determining long-term trends.

We will examine two approaches to dealing with the 1995–96 TOMS data gap: first, using the TOMS datasets as they stand with no intercalibration, and secondly, normalizing the earlier and later portions via intermediary datasets. Our focus is on using the SBUV2 datasets as this intermediary. A recent re-calibration of these datasets has improved the understanding of their time-dependence. Also, the NOAA-9 and NOAA-11 orbits have drifted to the extent that the SBUV2 viewing conditions are similar to what they were at launch. Ground-based column ozone measurements from Dobson and Brewer stations can also be used to connect the TOMS data records. However, the uncertainties in comparing satellite and ground ozone measurements are considerable, and we have elected not to consider this approach.

We chose to assess the two approaches to combining TOMS datasets by comparing long-term trend uncertainties estimated with each. These estimates ignore sources of correlated error such as the solar cycle, Quasi-Biennial Oscillation and seasonal cycles, and focus on sensor measurement uncertainties. We reason that correlated error sources represent a minor fraction of the trend uncertainty for many TOMS and SBUV2 measurements. This is borne out by McPeters et al. (1996a) in their derivation of combined TOMS/N7 and TOMS/M3 global trend uncertainty. They find that instrumental uncertainty is the dominant component in all but high latitude data. By ignoring correlated errors and treating the trend as linear we will certainly underestimate the total trend uncertainty. But our estimates can be thought of as minimum uncertainties, ones which approach the combined uncertainty when global ozone trends are linear.

2. Sensor uncertainties

Sensor measurement uncertainties are made up of two components, often referred to as accuracy and precision. In the context of long-term trends, errors which vary with a frequency of less than one year can generally be categorized as precision components. Compared to sensor accuracy uncertainties, the precision components for TOMS and SBUV are small and can be ignored. Sensor accuracy can be divided between time-independent, $\sigma_a$, and time-dependent, $\sigma_t$, components. The latter are a
major source of long-term trend uncertainty. Time-independent uncertainties are dominated by initial, or absolute, calibration errors. These uncertainties are relevant when calculating trends without TOMS intercomparisons.

In a discussion of multisensor trends, the absolute measurement uncertainty is less important than the relative uncertainty between sensors. However, we adopt the use of the term ‘absolute’ to simplify the text. In our context we define an absolute error as one which would introduce a time-invariant difference in the column ozone measured by two sensors under identical viewing conditions. Table 1 (McPeters et al. 1996b, 1998) lists the components of time-invariant ozone uncertainty for TOMS/N7 and TOMS/EP. Retrievals from either sensor share the errors in the ozone absorption cross-sections, Rayleigh cross-sections, and other retrieval errors. These retrieval errors are dominated by inadequate cloud height determination and the insensitivity to variations in tropospheric ozone variations (McPeters et al. 1998). Shared errors do not lead to offsets between the column ozone measured by two sensors provided the viewing conditions are, on average, the same. Comparisons between TOMS/N7 and TOMS/M3 (Seftor et al. 1997) demonstrated that differences are small provided stringent matching criteria are applied to minimize the effects of atmospheric variability. Thus we include only radiometric and wavelength calibration uncertainties from table 1 in our estimates of $\sigma_a$. A root sum squared of these components yields an estimate for the relative TOMS/EP and TOMS/N7 calibration uncertainty of $\sigma_a \approx 2.5\%$.

Long-term sensor calibration is a topic of great concern for BUV sensors, and has led to several sensor design improvements over the years. In the BUV technique, changes in radiometric response mostly cancel by using the Sun as a calibration source (Krueger 1995). But changes in the solar diffuser reflective properties must be accurately characterized in order to achieve good long-term calibrations. In fact, knowledge of diffuser reflectance is the primary source of long-term ozone trend uncertainty for TOMS and SBUV2 (McPeters et al. 1998).

The Nimbus 7 TOMS had no direct means of measuring reflectance changes in the solar diffuser. Consequently the time-dependence calibration of that instrument did not rely upon solar measurements, but rather used observed sensor changes at wavelengths not absorbed by ozone to estimate the changes at absorbed wavelengths (Wellemeyer et al. 1996). The result was a column ozone time-dependence uncertainty of about 1\% per decade. It is likely that the uncertainty is somewhat greater (as much as 2\%) in 1992–93 due to the drifting orbit and the effects of the Mt Pinatubo eruption (mid-1991).

As a result of the Nimbus 7 experience, subsequent TOMS and SBUV2 instruments were designed with mechanisms to monitor solar diffuser reflectance in orbit.

<table>
<thead>
<tr>
<th>Component</th>
<th>TOMS/EP</th>
<th>TOMS/N7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayleigh scattering</td>
<td>&lt;0.5</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>Ozone absorption cross section</td>
<td>&lt;2</td>
<td>&lt;2</td>
</tr>
<tr>
<td>Wavelength calibration</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Radiometric calibration*</td>
<td>&lt;1</td>
<td>1</td>
</tr>
<tr>
<td>Retrieval error</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Net</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

*Includes diffuser reflectance uncertainty.
All TOMS, beginning with TOMS/M3, use multiple diffusers to detect changes (Jaross et al. 1996). The SBUV2 sensors have only one diffuser, but employ a lamp to monitor diffuser reflectance change (Hilsenrath et al. 1995). TOMS Earth Probe also has a diffuser reflectance lamp, but it has not proved useful for detecting long-term changes.

The multiple diffuser technique has proved very precise in monitoring reflectance changes in the primary solar diffuser (Jaross et al. 1998). Using four years of diffuser data from TOMS/EP, we estimate a time-dependent calibration uncertainty of $\sigma_t = 0.1\%$. This figure is validated by monitoring residues from the ozone retrieval (McPeters et al. 1998). The primary wavelength used to derive TOMS equatorial ozone is 313 nm. Column ozone is also derived simultaneously from three other wavelengths that have both greater and lesser ozone absorption cross-sections. The residues at those three wavelengths are simply the retrieved ozone differences expressed in units of normalized radiance. The residue at 322 nm, which is highly sensitive to calibration drifts due to its relatively low absorption cross-section, provides us with a time-dependence uncertainty of $<0.3\%$. A recent re-calibration of the NOAA-9 (SBUV2/N9) and NOAA-11 SBUV2 (SBUV2/N11) datasets (Huang et al. 2002, Taylor et al. 2002) relies on a technique similar to the TOMS residues described above. An alternative calibration was required for these sensors due to malfunctions in their respective diffuser monitoring mechanisms. In contrast to the TOMS validations, wavelengths were chosen with large ozone absorption cross-sections to minimize the sensitivity to radiometric calibration errors. The D-pair calibration method, a variant of the Pair Justification Method (Herman et al. 1991), relies on the large differential cross-section between 306.5 nm and 312.5 nm and the small spectral separation. Wavelength pairs are generally used in the BUV technique to minimize the effects of wavelength-independent errors and variations in scene reflectivity. Since wavelength-dependent optical degradations are the primary source of long-term calibration errors, a small wavelength separation minimizes the relative radiometric error between the two channels. The D-pair can only be used at low latitudes due to the large ozone cross-sections. But calibration adjustments are determined for B-pair channels (317.5 nm and 331.2 nm) by comparing the retrieved ozone column amount in the equatorial region with that from the D-pair. The current estimated uncertainty in the D-pair time-dependence calibrations of the SBUV2 sensors ($\sigma_{t,N9}$ and $\sigma_{t,N11}$) is about 1% over the life of each.

3. TOMS–SBUV2 comparisons

Comparisons of TOMS data with the newly re-calibrated SBUV2 data have only just begun, but the results look promising. One approach used in the initial comparisons involved the use of ozone data averaged over selected areas of the globe where radiative transfer errors and ozone variations are minimal. One such area is near the equator in the central Pacific. Retrieval variations are at a minimum when solar zenith angles are low, aerosol contamination is at a minimum, and scenes are relatively cloud-free.

Comparisons between monthly mean ozone from TOMS and SBUV2 over the Pacific Ocean are shown in figures 2 and 3. While variations of several Dobson Units (1 DU = 1 milli atm-cm, about 0.4% of column ozone at the equator) persist, some of these variations are the result of the Mt Pinatubo eruption. Consequently, we have elected not to compare with TOMS/M3 data, and comparisons with
Figure 2. The difference between monthly mean column ozone amounts measured by SBUV2 and TOMS/N7 is shown. The means are for the Pacific Ocean between 7.5° S and 7.5° N latitude. The mean solar zenith angle in that region is shown as well. Comparison periods for TOMS and SBUV2 were chosen when the solar zenith angle of SBUV2 was less than 50°. Only comparisons prior to the Mt Pinatubo eruption are used. The standard deviations of the mean differences during these periods are indicated.

TOMS/N7 are restricted to the pre-Pinatubo period. The slow orbit drift of NOAA satellites results in dissimilar TOMS and SBUV2 viewing conditions for periods of several years. The primary change is in the mean solar zenith angle, shown in figures 2 and 3 for each sensor. When SBUV2 data are restricted to below 50° solar zenith angle, variations in the comparisons are reduced. A reduced solar zenith angle differential between sensors is the primary criterion for selecting the averaging periods shown in the figures.

The standard deviations of TOMS–SBUV2 ozone differences are given in table 2. The values are reasonably consistent, and result in standard errors of the mean differences that are 0.3 DU or less. These preliminary ozone comparisons indicate that TOMS/EP measures 2–3 DU more ozone in the equatorial Pacific region than did TOMS/N7. Estimates of the TOMS differential using either SBUV2/N9 or SBUV2/N11 agree to within 1 DU of each other.

By selecting the equatorial region, ozone comparisons lack significant seasonal variation and dependence on solar zenith angle. However, both effects are apparent
Characterization and radiometric calibration

NIMBUS-7

Solar zenith angle (°)

Ozone difference (DU)

NOAA-11–TOMS/EP

σ = 0.9 DU

Figure 3. The difference between monthly mean column ozone amounts measured by SBUV2 and TOMS/EP is shown. The same comparison criteria were applied as for TOMS/N7 (figure 2).

Table 2. TOMS SBUV2 ozone difference standard deviations (Dobson Units).

<table>
<thead>
<tr>
<th></th>
<th>NOAA-9</th>
<th>NOAA-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMS/N7</td>
<td>0.9</td>
<td>0.6</td>
</tr>
<tr>
<td>TOMS/EP</td>
<td>0.7</td>
<td>0.9</td>
</tr>
</tbody>
</table>

at higher latitudes. These effects may be related to the greater cloud fractions at higher latitudes, but we do not currently understand why the variations increase with latitude. Seftor et al. (1997) demonstrated that the latitude dependence of the TOMS/N7 vs TOMS/M3 ozone comparison is 2 DU or less provided their matching criteria are selected to minimize the effects of atmospheric variability. Unlike the intra-TOMS comparisons, it is difficult to match the SBUV2 and TOMS viewing conditions. The result is that retrieval errors may play a role in the observed column ozone differences. Consequently, we estimate comparison uncertainties of $\sigma_c = \pm 1.5$ DU ($\sim 0.5%$ ozone), much larger than the 0.3 DU standard errors. These larger uncertainties reflect the range of variation observed with latitude, solar zenith angle and scene reflectivity. We expect this uncertainty to decrease as viewing condition dependences are resolved.
4. Trend uncertainty

As discussed above, we considered two approaches to forming a combined TOMS data record: with and without an intermediate dataset to bridge the data gap. The TOMS/M3 dataset is not calibrated independently from TOMS/N7, so we exclude those data from our trend calculations. In the case where only TOMS data are used, the trend uncertainty depends upon how the trend is calculated. If a situation existed where, for a given TOMS, $\sigma_t > \sigma_a$, then it might make more sense to define a trend in terms of the end points of the combined TOMS dataset rather than fit those data. However, the uncertainties presented in §2 indicate that TOMS trend calibrations are always better than absolute calibrations. We therefore conclude that the smallest combined trend uncertainty is found by fitting the data.

We have taken two approaches to fitting the data, depending upon whether or not we use comparison information from SBUV2. In either case the two TOMS records can be thought of as a single time series with an ‘intervention’ in the middle. This terminology, used by Weatherhead et al. (1998) in their description of trend uncertainties, refers to an offset in measured values somewhere in the middle of the data record. They derive trend uncertainties by assuming that the combined dataset can be fitted to a function that is linear in time with an unknown additive offset $\delta$. They also considered the case where $\delta$ is known to within a given uncertainty.

We have adopted the approach of fitting the data with interventions, but find the formulae derived by Weatherhead et al. are not applicable to TOMS and SBUV2. The treatment by Weatherhead et al. and earlier by Tiao et al. (1990) ignores uncertainties in the time-dependent calibration of a sensor. In fact, as discussed earlier, this is by far the largest contributor to uncertainty in individual TOMS and SBUV2 trend measurements. Using the formalism provided by Weatherhead et al. (1998) we estimate a $1\sigma$ trend uncertainty of less than 0.2% for the TOMS/N7 and TOMS/M3 data record. Even with a moderate autocorrelation of 0.5 (50% of the measurement variation is repeated year to year), this uncertainty grows to only 0.3%. The TOMS/N7 trend uncertainty presented in §2 is about 1% per decade, which is very close to the value estimated by McPeters et al. (1996a) for the combined TOMS/N7 and TOMS/M3 equatorial trend uncertainty. Since their treatment includes both sensor and statistical sources of uncertainty, we conclude that the sensor contributions dominate.

Rather than attempting to derive a complex analytic expression for trend uncertainty that includes contributions from time-dependent calibration uncertainty, we chose to simulate the uncertainty using a simple Monte Carlo. In our model we fixed the ozone value measured by TOMS/N7 at launch to zero and allowed the per decade linear trend in the data to vary randomly within a $\sigma = 0.01$ normal distribution. Similarly, the TOMS/EP per decade trend was varied independently with $\sigma = 0.003$. Though the ozone trend may be non-linear, sensor time-dependent calibration errors tend to be linear in time. An offset between the TOMS/N7 and TOMS/EP initial values was also introduced with $\sigma = 0.025$, the relative calibration uncertainty of the two sensors. Each sensor’s data were represented by one point per year, which was varied about the linear trend with $\sigma = 0.002$. This inter-annual variation was estimated by observing the variance in TOMS and SBUV2 data at the equator. The final results were insensitive to the magnitude of this value.

Following each of 10,000 random selections the simulated data points were fitted to a straight line with an allowed offset at the time of the TOMS/EP launch. That is, a third parameter was included in the regression that is the offset value between TOMS/N7
and TOMS/EP. The trends on either side of the offset were assumed to be the same. We found a standard deviation for the year 2000 of 2.03%, or 0.92% per decade.

Next, we introduced information from SBUV2/N11 to fill the TOMS gap. Its trend was allowed to vary randomly with $\sigma = 0.01$ from 1989 to 1999. The TOMS/N7 vs TOMS/EP absolute uncertainty was replaced by a TOMS vs SBUV2 comparison uncertainty of $\sigma = 0.005$. Each of 10,000 TOMS time series samples was fitted to a linear trend without an offset parameter. The resulting year 2000 values had a standard deviation of 1.78%, or 0.81% per decade. Comparison with the TOMS-only uncertainty, above, suggests that intercomparison offers only a slight advantage given the current SBUV2 uncertainties.

Finally, we turned our attention to the forthcoming addition of QuikTOMS (TOMS/QT) to the combined TOMS data record. Its launch is currently scheduled for spring 2001†. We anticipate a minimum one-year overlap between TOMS/EP and TOMS/QT that should yield a 0.2% comparison uncertainty (Seftor et al. 1997). In our model we assumed the TOMS/QT time-dependent uncertainty was equal to that of TOMS/EP. With a somewhat optimistic six-year lifetime, the 3-TOMS combined trend uncertainty falls to 0.64% per decade with no SBUV2 comparison and 0.53% per decade with SBUV2/N11 bridging the 1993–96 gap. The decreased uncertainties are primarily a result of the smaller relative contribution of the TOMS/N7 data to the combined time series. Trend uncertainties are summarized in table 3.

### Table 3. TOMS $1\sigma$ trend uncertainties (% per decade).

<table>
<thead>
<tr>
<th></th>
<th>Without NOAA-11</th>
<th>With NOAA-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMS/N7, /EP</td>
<td>0.92</td>
<td>0.81</td>
</tr>
<tr>
<td>TOMS/N7, /EP, /QT</td>
<td>0.64</td>
<td>0.53</td>
</tr>
</tbody>
</table>

5. Conclusions

We have used time-invariant and time-dependent uncertainties of TOMS calibrations to estimate minimum long-term trend uncertainties for a combined TOMS/N7 and TOMS/EP dataset. We estimate an uncertainty of just under 1% per decade, close to the trend uncertainty from TOMS/N7 alone and to the goal for long-term ozone monitoring. The 14$\frac{1}{2}$ year TOMS/N7 record dominates that of the much shorter-lived TOMS/EP. The three-year data gap between TOMS also contributes to the trend uncertainty, but using an intermediary dataset to bridge the gap reduces this uncertainty only marginally. A more important factor in reducing the trend uncertainty is the length of the TOMS/EP record. A long-lived TOMS/EP could substantially reduce the uncertainty. A more probable scenario, and one with nearly equal results, is a subsequent dataset (TOMS or OMI) that is well intercalibrated with TOMS/EP. We estimate that an additional six year dataset could reduce the per decade trend uncertainty by a third. We recognize that the trend uncertainty during the last quarter of the twentieth century will not improve. This may make the assessment of the predicted ozone recovery somewhat difficult since that recovery is expected to begin near the turn of the century. The recovery rate, though, should be well measured by subsequent instruments.

†Note: The September 2001 QuikTOMS launch failed to place the instrument in a useful orbit.
Acknowledgments.
The Raytheon ITSS authors were supported under National Aeronautics and Space Administration (NASA) contract NAS5-31755 with the NASA Goddard Space Flight Center.

References


