Making Earth Science Data Records for Use in Research Environments (MEaSUREs)

README Document for TOMSN7SO2

Goddard Earth Sciences Data and Information Services Center (GES DISC) http://disc.gsfc.nasa.gov

NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA

Principal Investigators: Nickolay A. Krotkov, Pawan K. Bhartia, Code 614
Nickolay.A.Krotkov@nasa.gov
301-614-5553

Algorithm Developer: Brad Fisher, SSAI, Code 614
bradford.fisher@ssaihq.com
301-867-2188

Algorithm Support: Peter Leonard, ADNET, Code 619
peter.j.leonard@nasa.gov
Last Revised: June 23 2017
Reviewed by: Andrey Savtchenko

Name
GES DISC
GSFC Code 610.2
Date
Goddard Space Flight Center
Greenbelt, Maryland
Revision History
# TABLE OF CONTENTS

1. INTRODUCTION
   1.1 Dataset Description
   1.2 TOMS Instrument
   1.3 Science Background

2. MULTI-SATELLITE SO₂ ALGORITHM
   2.1 General Description
   2.2 Step 1 Retrieval
   2.3 Step 2 Retrieval
   2.4 Soft Calibration
   2.5 Uncertainty

3. DATASET ORGANIZATION

4. DATA CONTENTS

5. CONTACTS

6. REFERENCES
1. INTRODUCTION

1.1 TOMS Nimbus-7 MEaSUREs SO$_2$ Dataset: TOMSN7SO2

This document contains a brief description of the TOMSN7SO2 data product. TOMSN7SO2 is a Level 2, orbital track volcanic sulfur dioxide (SO$_2$) product for the Total Ozone Mapping Spectrometer (TOMS) onboard NASA’s Nimbus 7 satellite, which was launched on October 24, 1978 into polar sun-synchronous orbit and collected data from November 1, 1978 to May 6, 1993. This was the first mission to provide daily contiguous global maps of total ozone (O$_3$). As part of the NASA’s Making Earth System Data Records for Use in Research Environments (MEaSUREs) program, the Goddard Earth Science (GES) Data and Information Data Center (DISC) has released a new SO$_2$ Earth System Data Record (ESDR), TOMSN7SO2, re-processed using new 4 UV wavelength bands MS_SO2 algorithm that spans the full Nimbus 7 data period. TOMSN7SO2 is a Level 2 orbital swath product, which will be used to study the fifteen-year SO$_2$ record from the Nimbus-7 TOMS and to expand the historical database of known volcanic eruptions.

1.2 TOMS Instrument

TOMS is a fixed-grating Ebert-Fastie monochromator with photomultiplier tube detector that measures solar backscattered ultraviolet (BUV) radiances (I) at 6
narrow wavelengths bands (Full Width at Half Maximum band width ~1.1nm) in the near ultraviolet (UV) spectral region as well as the incident solar irradiances (F) (Heath, et al., 1975). The ratio of radiance to irradiance provides the spectral reflectivity parameter used in the ozone retrieval. The wavelength band centers shown in Table 1 were selected to optimize column ozone retrievals assuming that two pairs of shorter, absorbed wavelengths would be needed to cover the full dynamic range of ozone and solar zenith angles encountered globally, similar to the Dobson Spectrophotometer design. Two additional non-absorbed longer wavelengths were provided to measure the surface or cloud reflectivity (R) and its spectral dependence.

<table>
<thead>
<tr>
<th>Channel Number</th>
<th>Wavelength, vacuum [nm]</th>
<th>Used in SO$_2$ retrieval</th>
<th>MS$_{SO2}$ Retrieved Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>312.34</td>
<td>Not Used</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>317.35</td>
<td>x</td>
<td>SO$_2$</td>
</tr>
<tr>
<td>3</td>
<td>331.06</td>
<td>x</td>
<td>O$_3$</td>
</tr>
<tr>
<td>4</td>
<td>339.66</td>
<td>x</td>
<td>dR/d$\lambda$</td>
</tr>
<tr>
<td>5</td>
<td>359.99</td>
<td>Not Used</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>379.89</td>
<td>x</td>
<td>R</td>
</tr>
</tbody>
</table>
TOMS scans in the cross-track direction in $3^\circ$ steps from $51^\circ$ on west side of nadir to $51^\circ$ on the east, for a total of 35 cross-track samples. The instantaneous field-of-view (FOV) of $3^\circ \times 3^\circ$ results in a footprint varying from a 50 km x 50 km square FOV at nadir to a 125 km by 280 km diamond-shaped FOV at the scan extremes. The total swath width is 3000 km covering Earth’s surface in 14 orbits per day.

1.3 Science Background

The original design of the TOMS instrument assumed that ozone was the only significant gaseous absorber in near UV wavelengths (Dave and Mateer 1967). A second absorber was later discovered in 1982 as a high ozone anomaly observed over the El Chichon volcano in Mexico and recognized as SO$_2$ gas absorption in the volcanic cloud (Krueger 1983). Following the discovery, Krueger et al. (1995) developed a first algorithm to separate the ozone and SO$_2$ signals and to retrieve the column amount of SO$_2$. The off-line retrieval has been used on a case-by-case basis to retrieve SO$_2$ mass from explosive eruptions using TOMS measurements on Nimbus-7 and follow-up missions (Krueger et al., 2000; Carn et al., 2003). The TOMSN7SO2 data product is the first public release of the complete Nimbus-7 TOMS SO$_2$ Level 2 data, re-processed with a new 4 UV wavelength bands algorithm that is fast enough to permit production of the entire global dataset.
2. MULTI-SATELLITE SO$_2$ ALGORITHM

2.1 General Description

This section describes the new Multi-Satellite SO$_2$ (MS_SO2). This algorithm builds on the heritage of the TOMS total ozone (TO3) algorithm (Dave and Mateer, 1967, McPeters et al., 1996), but adds sulfur dioxide (SO$_2$) as a second absorber to the pre-computed BUV look-up tables. The algorithm uses four of the six spectral bands to retrieve a state vector ($\hat{X}$) with 4 parameters: SO$_2$ column amount, O$_3$ column amount, the Lambertian equivalent reflectivity (LER) at 380 nm, R, and the spectral dependence of R, dR/d$\lambda$. The algorithm relies on spectral differences in SO$_2$ and O$_3$ cross sections to simultaneously quantify the amounts of the two trace gases (see Table 1 and Figure 1). SO$_2$ is more absorbing than O$_3$ at the shorter channels (312 and 317 nm), whereas O$_3$ is more absorbing at the longer absorbing channel (331 nm). Figure 1 shows the O$_3$ and SO$_2$ spectral cross sections in the near-UV spectral range and the SO$_2$ to O$_3$ cross section ratio over the same range. It is relative differences in the cross-section ratio that allows for the separation of the two gases. The three vertical bars show the centers of the absorbing spectral bands used in the retrieval.
2.2 Step 1 Retrieval

The SO$_2$ retrieval is performed in two steps and applied iteratively for each FOV. The retrieval starts with a first guess value of total column ozone based on climatology and assuming zero SO$_2$. The LER is then computed at 380 nm (R). Because O$_3$ and SO$_2$ absorption is very weak at this wavelength, R remains fixed during the iterations. The algorithm then uses 3 shorter spectral bands to retrieve $\Delta(\partial R/\partial \lambda)$, $\Delta O_3$ and $\Delta SO_2$ adjustments to the state vector: $\Delta x = [ \Delta SO_2, \Delta O_3, \Delta(\partial R/\partial \lambda)]$, assuming a linear $R(\lambda)$ spectral dependence:

$$\Delta y_i = N_{m_i} - N_{c_i} = \frac{\partial N_{c_i}}{\partial SO_2} \Delta SO_2 + \frac{\partial N_{c_i}}{\partial O_3} \Delta O_3 + \frac{\partial N_{c_i}}{\partial R} (\lambda_i - \lambda_R) \Delta \frac{\partial R}{\partial \lambda} \quad (1)$$
where i = 1, 2, 3 corresponds to the 317, 331, and 340 nm bands and $\lambda_R = 380$ m. The left-hand side of Eq. 1 represents the difference (residual) between the measured and calculated sun normalized BUV radiances expressed in N-value units, defined in Equation 2 as:

$$N\text{-value} = -100\log_{10}(I/F) \quad (2)$$

where I is the measured radiance and F is the TOA incoming solar flux.

The partial derivatives (Jacobians) on the right-hand side form a Jacobian K-matrix. The matrix consists of the N value sensitivities associated with linear perturbations in state vector parameters: SO$_2$, O$_3$ and dR/d$\lambda$ (i.e., reflectivity spectral slope)

$$K = \begin{pmatrix}
\frac{\partial N_{317}}{\partial \text{SO}_2} & \frac{\partial N_{317}}{\partial \text{O}_3} & \frac{\partial N_{317}}{\partial \lambda} \\
\frac{\partial N_{331}}{\partial \text{SO}_2} & \frac{\partial N_{331}}{\partial \text{O}_3} & \frac{\partial N_{331}}{\partial \lambda} \\
\frac{\partial N_{340}}{\partial \text{SO}_2} & \frac{\partial N_{340}}{\partial \text{O}_3} & \frac{\partial N_{340}}{\partial \lambda}
\end{pmatrix}
(\lambda_i - \lambda_R) \quad (3)$$

The K matrix elements are computed for each spectral band from the BUV radiance look-up tables (LUT). The LUT’s nodal points correspond to surface pressure, satellite-viewing geometry (solar zenith, satellite nadir, and relative azimuthal angles), total ozone and column SO$_2$, assuming TOMS standard Ozone profiles. For SO$_2$ the algorithm assumes three Gaussian shape profiles (2-km half-width), with different peak SO$_2$ altitudes: 8 km (middle troposphere, TRM data),
13 km (Upper Tropospheric/Lower Stratospheric, TRU data), and 18 km (lower stratospheric, STL data). The RT calculations were compiled off-line using TOMRAD radiative transfer code (Dave 1965).

The state vector increment, $\Delta x$, is defined in (4). The adjustment to $\Delta x$ after each iteration is subsequently determined by inverting the 3 x 3 Jacobian matrix $K$:

$$\Delta x = K^{-1} \Delta y = \begin{pmatrix} \Delta SO_2 \\ \Delta O_3 \\ \Delta \frac{\partial R}{\partial \lambda} \end{pmatrix}$$

The differentials $\Delta x$ are added to the previous iteration state $x$ values, beginning with a first guess to obtain the column amounts of $SO_2$ and $O_3$ in Dobson units (1 DU = $2.69 \times 10^{16}$ molec. cm$^{-2}$) and $\partial R/\partial \lambda$ [nm$^{-1}$]. The largest change in $SO_2$ column amount occurs in the first iteration and the process typically converges in 2-3 iterations depending on the actual $SO_2$ amount.

2.3 Step 2 Retrieval

Volcanic ash and large $SO_2$ amounts in volcanic clouds bias high the Step 1 $O_3$ retrieval, leading to an underestimation of $SO_2$. In such cases the algorithm applies a “Step 2” correction. The correction is initiated when either of the following two criteria is met in Step 1:
1) SO$_2$ amount exceeds 20 DU, or

2) O$_3$ exceeds 2 standard deviations of the background regional mean O$_3$

(where the regional O$_3$ statistics do not include FOVs with SO$_2$ > 20 and UV aerosol index (AI) > 2).

The AI is calculated using the following equation:

$$AI = (340 - 380) \frac{\partial N_{340}}{\partial R} \frac{\partial R}{\partial \lambda}$$  \hspace{1cm} (5)

where the Jacobian $\partial N/\partial R$ is computed at 340 nm. The scale factor accounts for wavelength difference between the TOMS longer (reflectivity) spectral bands.

For each FOV meeting Step 2 criteria, the ozone is first corrected by masking the volcanic plume region and then linearly interpolating O$_3$ from outside the plume along the orbital track for each affected swath position. The interpolated O$_3$ is more uncertain for large regional plumes like Pinatubo and in mid- and high-latitudes with large O$_3$ variability. Step 2 uses only 317 nm and 340 nm spectral bands and one iteration to calculate the correction $\Delta x = [\Delta SO_2, \Delta \partial R/\partial \lambda]$ and apply it to Step 1 retrievals: SO$_2$ = SO$_2$ (Step 1) + $\Delta$SO$_2$ and $\Delta \partial R/\partial \lambda = \Delta \partial R/\partial \lambda$ (Step 1) + $\Delta \partial R/\partial \lambda$. Note that even though the corrected O$_3$ is not explicitly used in the construction of the K-matrix, it is used to select the SO$_2$, R and O$_3$ nodes in the LUTs. A Step 2 algorithm flag is provided in the product. The flag informs users whether Step 2 was applied and which of the above criteria was used in the selection process.
2.4 Soft Calibration of the 340 Channel

The mean \( \text{SO}_2 \) background far away from any \( \text{SO}_2 \) sources is assumed to be zero, but due to noise in the retrieval system (e.g., instrument, forward model), the retrieved \( \text{SO}_2 \) values fluctuate around zero (positive and negative). To correct for any residual \( \text{SO}_2 \) bias in the mean background, the algorithm inverts the K-matrix – what Rogers (2000) refers to as the gain matrix (or G-matrix) – and uses the matrix element associated with the \( \text{SO}_2 \) sensitivity at 340 nm to calibrate the 340 radiance.

An \( \text{SO}_2 \) and aerosol–free TOMS orbit over the central Pacific was selected for this procedure. A Step 1 retrieval is then performed for this orbit with no soft calibration applied. Using the \( \text{SO}_2 \) field generated for this orbit, a correction to the 340 radiance is then computed as shown in (6)

\[
dN_{340} = \text{Mean}(\frac{\partial \text{SO}_2}{\partial N_{340}})
\] (6)

where, \( \partial \text{SO}_2/\partial N_{340} \) is the (2,2) G-matrix element corresponding to the \( \text{SO}_2 \) sensitivity in the 340 spectral band. This calculation is applied at two latitude bands: 1) abs(lat) \( \leq 30^\circ \) and 2) abs(lat) \( > 30^\circ \). Figure 2 shows a plot of the N-value correction applied at both latitude bands.
Figure 2. N-Value correction of the 340 nm channel at two latitude bands: 1) 30S ≤ Latitude < 30N and 2) Latitude < 30S and Latitude > 30N.

2.5 Uncertainty

The retrieval of volcanic SO$_2$ is subject to limitations related to the sensitivity of the instrument and the skill of the forward model in simulating the BUV radiances. Noise in the system produces a random distribution in the SO$_2$ background around zero. We characterize the low-end sensitivity of the instrument by estimating the random error in the distribution of retrieved SO$_2$, in background regions where SO$_2$ concentrations are considered well below the detection limit of the TOMS instrument.
The spatial area of the individual FOVs vary, resulting in sampling differences as a function of the cross track position, which in turn affects the random error. We characterize the low-end sensitivity of the retrieval with respect to the standard deviation in the distribution of the SO₂ noise, as a function of the swath position, as shown in Figure 3. The maximum noise value of about 6 DU is attained at nadir. We set the sensitivity threshold for detecting volcanic SO₂ at 15 DU, just below the 99% confidence interval.

The uncertainty associated with the SO₂ retrieved from actual volcanic events was characterized by comparing two independent UV algorithms applied to the same BUV data set (Figure 4). In this case, MS_SO2 retrievals were compared against estimates from the independent Principal Components Algorithm (PCA). For this analysis, the PCA algorithm has been modified to use 6 TOMS channels (normally it is applied in a hyper spectral instruments). We subsequently estimated the uncertainty by computing the mean deviation and co-variance between these two sensors for several major volcanic events. Based on this analysis we estimate the uncertainty for the large SO₂ events to be about 10-15 DU.
Figure 3. a) Standard Deviation of SO₂ noise as function of cross-track swath position and b) background SO₂ noise histogram for all cross-track positions.

Figure 4. Scatterplot comparing retrievals from independent MS_SO2 and the PCA algorithms for several major eruptions spanning the full dynamic range of estimates. The statistical parameters are used to estimate the uncertainties for MS_SO2 retrievals in different volcanic clouds (sample of 1840 FOVs).
3. DATASET ORGANIZATION

The TOMSN7SO2 product provides a set of Level 2 orbital swath files produced by applying the MS_SO2 algorithm to the N7-TOMS data record.

3.1 File Naming Convention

The TOMSN7SO2 data granules (files) are named as in this example:

TOMS-N7_L2-TOMSN7SO2_1991m0817t090821-o64696_v02-00-2017m0502t064217.h5,

where the components of filename are as follows:

1. Instrument (TOMS)
2. Spacecraft (N7)
3. Process Level (L2)
4. ESDT Short Name (TOMSN7SO2)
5. Date and Time at Start of Orbit (1991-08-17 09:08:21 UTC)
6. Orbit Number (64696)
7. Product Version (02-00)
8. Production Date and Time (2017-05-02 06:42:17 UTC)
9. File Type (h5)

3.2 File Format and Structure

The TOMSN7SO2 data granules are in plain HDF5 that is netCDF4-compatible and CF-compliant. Each TOMSN7SO2 data granule contains global attributes, dimensions, an ancillary data group, a geolocation data group, a science data group, and a sensor data group.
3.3 Key Science Data Fields

The TOMSN7SO2 science product provides four key science data fields associated with the retrieval: SO₂ column amount, O₃ column amount, the Lambertian Equivalent Reflectivity at 380 nm (R), and the reflectivity spectral slope (\(\partial R/\partial \lambda\)). The other retrieved parameters provide important diagnostic information.

3.3.1 SO₂ Column Amount

ColumnAmountSO₂_TRM, ColumnAmountSO₂_TRU and ColumnAmountSO₂_STL are, respectively, the column amounts for SO₂ in assumed Gaussian distributions with 2-km standard deviations centered in the middle troposphere (8 km), the upper tropical troposphere (13 km) and the lower stratosphere (18 km).

3.3.2 O₃ Column Amount

ColumnAmountO₃_TRM, ColumnAmountO₃_TRU and ColumnAmountO₃_STL are, respectively, the O₃ total column amounts that correspond to the SO₂ retrievals for the middle troposphere (8 km), the upper troposphere (13 km) and the lower stratosphere (18 km).
3.3.3 Lambertian Equivalent Reflectivity at 380 nm

The Lambertian Equivalent Reflectivity at 380 nm (LER380) is independent of the height of the assumed SO$_2$ layer.

3.3.3 Reflectivity Spectral Dependence ($\partialR/\partial\lambda$)

dRdlambda_TRM, dRdlambda_TRU and dRdlambda_STL are, respectively, the reflectivity spectral slopes ($\partialR/\partial\lambda$) that correspond to the SO$_2$ retrievals for the middle troposphere (8 km), the upper troposphere (13 km) and the lower stratosphere (18 km).

4. DATA CONTENTS

Each TOMSN7SO2 data granule contains global attributes, dimensions, an ancillary data group, a geolocation data group, a science data group, and a sensor data group. This section provides specific details regarding these components, as shown in Figure 5.

Figure 5. Shows 4 main data groups and dimensions.
4.1 Global Attributes

There are 43 global attributes in each TOMSN7SO2 data granule as shown in Figure 6.

![Figure 6. TOMSN7SO2 Global attributes](image-url)
4.2 Dimensions

There are six dimensions in each TOMSN7SO2 data granule (Figure 5):

nCorners - The dimension representing the four ground-pixel corners.

nTimes - The dimension representing the time of observation.

nWavel4 - The dimension representing the subset of four TOMS wavelengths used in the retrievals.

nWavel6 - The dimension representing all six TOMS wavelengths.

nXtrack - The dimension representing the 35 cross-track scan positions.

4.3 Data Fields

4.3.1 Ancillary Data

There are only two fields, CloudPressure and TerrainPressure, in the ancillary data group in each TOMSN7SO2 data.

4.3.2 Geolocation Data

There are ten fields in the geolocation data group in each TOMSN7SO2 data granule as shown in Figure 7.

Figure 7. Geolocation Data Group
4.3.3 *Science Data*

There are 27 fields in the science data group in each TOMSN7SO2 data granule as shown in Figure 8.

![Tree Diagram of Science Data Fields]

**Figure 8. Science Data Group**
4.3.4 Sensor Data

There is only one field, Wavelength, in the sensor data group in each TOMSN7SO2 data granule.

5. CONTACTS

Principal Investigator: Nickolay Krotkov
Nickolay.A.Krotkov@nasa.gov
301-6114-5553

Algorithm Developer: Brad Fisher
bradford.fisher@ssaihq.com
301-867-2188

Algorithm Support: Peter Leonard
peter.j.leonard@nasa.gov
6. REFERENCES


