Supplementary Material

# EPIC Column Weighting Functions

To facilitate error analysis, Column Weighting Functions (CWFs) had been included in version 3 processing. Since UV satellite measurements have reduced sensitivity to ozone changes in the boundary layer, CWFs are aimed to help users to interpret EPIC total ozone retrievals and indicate the weight of measurements in each layer. The EPIC CWFs are calculated using the formulation described in (Rodgers, 2000):

(1)

where **A** is the Averaging Kernel matrix with 11 × 11 elements, **K** is the matrix with ozone profile sensitivity for 3 EPIC UV channels used in the ozone retrievals. The retrieval algorithm uses the logarithm of sun-normalized radiances **I** called N-values, . **Se** is the measurement error covariance matrix with dimensions 3 × 3 that has only diagonal elements equivalent to 1% error in measured sun-normalized radiances **I** (or 0.43 N-value).   is a priori covariance matrix (11 × 11) that is defined as:

(2)

where is a climatological ozone profile defined at 11 Umkehr layers, with pressure at the bottom of each layer defined as Pbottom(i)=1013.25\*0.5(i-1), and *i* and *j* are layer indexes from 1 to 11. Finally, CWF are determined by summing rows of the Averaging Kernel matrix:

(3)

The shape of CWF is determined by sensitivity of measured radiances to ozone profile changes K=. The examples of vertical profiles of CWF are shown in Supplementary Fig. 1. We recommend using values of CWFs in the lowest layer *i=*1 to determine sensitivity of EPIC total column retrievals to ozone changes in the lowest layer. However, it is important to note that we do not use the Optimal Estimation technique to retrieve EPIC ozone columns so that Averaging Kernels and Column Weighting Functions derived from Equations (1-3) should be used with caution for layers in the stratosphere (*i ≥*3).

# Boundary layer correction for EPIC tropospheric column ozone

To account for the limited sensitivity of UV satellite measurements to the ozone variability in the boundary layer (BL) between 506-1013 hPa, we used simulated tropospheric ozone data from GEOS-Replay (Strode et al. 2020) over 12-years (2005-2016). This correction represents a seasonal-cycle adjustment, since ozone variability in the troposphere including BL is largely determined by the seasonal cycle. First, we created monthly mean (MM) ozone climatology on 5° latitude × 5° longitude grid by averaging model data over 12 years. That resulted in the MM array of ozone partial columns in the BL with dimensions 12 × 72 × 36 for month, longitude, and latitude, respectively. Then, to correct daily EPIC TCO maps, we extrapolated the 12 mid-monthly climatological model values () to 365 days using a sine-cosine Fourier series expansion. The resulting daily-mean (DM) climatology has dimensions 365 × 72 × 36 for day of the year, longitude, and latitude. Similarly, 12-month zonal mean apriori climatology was extrapolated to 365 days creating a 365 by 36 array. Then the differences between daily model climatology and apriori were calculated on the 5° × 5° grid: that also has dimensions 365 × 72 × 36 for day of the year, longitude, and latitude. Finally, we extrapolated these 5° latitude × 5° longitude daily differences to the 1° latitude × 1° longitude grid to match the resolution of daily EPIC TCO for applying the adjustments. We only use the Column Weighting Function for the bottom layer (506-1013 hPa) to calculate the adjustments :

(4)

# EPIC total ozone retrieval algorithm

EPIC ozone retrieval algorithm, described in (Herman et al., 2018), uses triplets of wavelengths to retrieve total ozone column (TOZ). Two ozone absorption channels either 317.5 and 340 nm or 325 and 340 nm, depending on optical depth conditions, are combined with the 388 nm channel which has no ozone absorption. The retrieval algorithm uses the logarithm of albedo values *αλ* called N-values *αλ*The total column ozone is derived as:

(5)

where Ωn-1 is TOZ from previous iteration or initial estimate, λ1 and λ2 are ozone absorption wavelengths, ΔNλi is residuals (differences between measured and calculated albedos) in N-values, is the computed top of atmosphere albedo sensitivity with respect to the total column ozone, is the computed sensitivity to the surface reflectivity at wavelengths λ1 or λ2; λ0 is a reference reflectivity wavelength 388 nm. The important assumption in the EPIC’s ozone retrieval algorithm is that the surface reflectivity changes linearly with wavelength:

|  |  |
| --- | --- |
|  | (6) |

where *b* is a constant reflectivity slope. The 388 nm reflectivity and cloud fraction are derived once based on EPIC 388 nm measured albedos and do not change during the ozone retrieval iterations. The residuals represent differences between measured and calculated N-values:

(7)

The non-zero residuals are caused by differences between assumed ozone and true ozone and by differences in the estimated surface reflectivity at ozone absorption wavelengths λ1 and λ2 from the 388 nm reflectivity . Therefore, residuals can be also approximated by the following expression:

(8)

By combining equations (8) and (6) we can express residuals as:

(9)

The reflectivity slope can be found as:

As the TOZ value is updated with each iteration using eq. (5), top of the atmosphere albedos and sensitivities ( and ) are re-computed. It is important to note that the surface reflectivity and the cloud fraction derived from EPIC 388 nm measured albedo are not impacted by changes in ozone values and remain constant during iterations. The albedos and sensitivities are computed with the wavelength independent surface reflectivity at 388 nm . The reflectivity slope, *b*, gets updated after each iteration. We report reflectivity slope values and residuals in the Level 2 ozone files. The version 3 ozone retrieval algorithm stops iterations when ΔΩ is less than 0.5 DU. The final residuals for ozone absorbing wavelength are equal to zero.

# Supplementary Figures and Tables

|  |  |
| --- | --- |
| Shape  Description automatically generated with low confidence | Shape  Description automatically generated with low confidence |
| **Supplementary Figure 1.** Vertical profiles of EPIC CWF for two different cases: in the equatorial latitudes (left) and mid-latitudes (right). CWFs are typically close to 1 in all layers except for the boundary layer indicating that the measurements are very sensitive to ozone changes in those layers. The CWF for the lowest boundary layer ranges between 0.0 and 0.7 meaning that only a fraction of ozone variations in that layer will be retrieved. The magnitude of the CWF in the boundary layer depends on SZA/SLA, reflectivity, and scene pressure (terrain height). The 11 vertical points plotted in this figure correspond to layer centroid pressures of 716.48, 358.24, 179.12, 89.56, 44.78, 22.39, 11.19, 5.60, 2.80, 1.40, 0.70 hPa. | |

|  |
| --- |
| A picture containing text, antenna  Description automatically generated |
| **Supplementary Figure 2.** Retrieved cloud-height pressure in the equatorial region (10S-10°N) derived from EPIC A-Band and OMI OCP algorithms on April 25, 2017, as a function of cloud fraction. The uncertainty bars represent the spread (1-sigma standard deviation) in retrieved cloud pressures. EPIC A-Band cloud-height pressure agrees well with OMI OCP with differences less than 50 hPa over a wide range of cloud fraction values. A 50 hPa offset will result in ~ 1 DU error in retrieved total ozone. The differences are larger for small cloud fractions (<0.1) but have a very small effect on ozone retrievals. |

|  |
| --- |
| Diagram  Description automatically generated with medium confidence |
| **Supplementary Figure 3.** Average number of EPIC observations per day (per 24 hours) at a given fixed grid point, shown as function of latitude and month for year 2016. The EPIC observations are not possible during the polar night. Because of telemetry limitations caused by the location of the receiving antenna at Wallops Island, Virginia, the EPIC data are downloaded to the Earth more often during northern summer months than in winter. This results in more sampling in the northern hemisphere compared to the southern hemisphere. |

|  |
| --- |
| Chart, scatter chart  Description automatically generated |
| **Supplementary Figure 4.**  EPIC hourly (bi-hourly) TOZ measurements at different locations on June 20 (Day 171) and June 21 (Day 172) of 2018 as a function of SZA. Negative SZA values corresponds to morning measurements and positive to afternoon. The red arrows indicate EPIC retrievals derived from the 325 nm triplet, which is typically used at large SZA/SLA. This plot shows that the 325 nm retrievals deviate from daily mean values, and those deviations cannot be explained with the natural ozone variability. Deviations are larger for the afternoon measurements than early morning. These errors are currently under investigation. For scientific purposes, we do not recommend using the 325 nm retrievals that correspond to the algorithm flag equal to 2 or 102 or 112. |

|  |  |
| --- | --- |
|  |  |
| **Supplementary Figure 5.** A comparison of EPIC and OMI data for different averaging periods of 1 month (left) and 1 week (right). At both a 3-month and 1-month averaging periods the EPIC and OMI data agree closely, but the noise tends to increase at shorter averaging time windows. Once you go to shorter periods, the scene location and time matching become critical. In addition, natural ozone variability and instrumental noise further contribute to increased noise in comparisons. | |

|  |
| --- |
| Chart, histogram  Description automatically generated |
| **Supplementary Figure 6.** Comparison of daily EPIC TCO with coincident OMI (black solid line) and OMPS TCO (red dotted line) over 2015-2020 as a function of latitude. The vertical error bars indicate 1 standard deviation. The biases in the tropics and northern mid-latitudes tend to be positive with 1-3 DU mean biases. In the southern hemisphere, EPIC tends to have negative biases compared to both OMI and OMPS. The noise is smaller in the tropics and increases in both hemispheres. |

**Supplementary Table 1.** List of ground-based stations used in this study.

|  |  |  |  |
| --- | --- | --- | --- |
| **Station name** | **Latitude** | **Longitude** | **Type of measurements** |
| Reading | 51°N | 1°W | Brewer |
| Izana | 28°N | 16W | Brewer |
| Aosta | 46°N | 7°E | Brewer |
| Abu Dabi | 24°N | 55°E | Brewer |
| Murcia | 38°N | 1°W | Brewer |
| Hobart | 43°S | 147°E | Brewer |
| Greenbelt | 39°N | 77°W | Pandora |
| Fairbanks | 65°N | 148°W | Pandora |
| Innsbruck | 47°N | 11°E | Pandora |
| Boulder | 40°N | 105°W | Pandora |
| Buenos Aires | 35°S | 58°W | Pandora |
| Comodoro Rivadavia | 46°S | 67°W | Pandora |
| Legionowo | 52°N | 21°E | Sonde |
| Hohenpeissenberg | 48°N | 11°E | Sonde |
| Payerne | 46°N | 7°E | Sonde |
| Madrid | 40°N | 3°W | Sonde |
| Naha | 26°N | 127°E | Sonde |
| Hilo | 19°N | 155W | Sonde |
| Paramaribo | 6°N | 55°W | Sonde |
| Nairobi | 1°S | 37°E | Sonde |
| Natal | 6°S | 35°W | Sonde |
| Ascension Island | 8°S | 14°W | Sonde |
| Lauder | 45°S | 170°E | Sonde |
| Marambio | 64°S | 57°W | Sonde |