MOPITT

(Measurements of Pollution in the Troposphere)

Version 9 Product User's Guide

MOPITT Algorithm Development Team*

Atmospheric Chemistry Observations and Modeling Laboratory

National Center for Atmospheric Research

Boulder, CO 80307

Last Revised February 3, 2022

*Correspond to Helen Worden at hmw@ucar.edu

Table of Contents

1   MOPITT Version 9 Product Highlights

This guide describes new product features in the MOPITT Version 9 (V9) Level 2 products for tropospheric carbon monoxide (CO). This comprehensive version of the User's Guide is meant to serve as a "standalone" document, and includes relevant material from earlier User's Guides. Any updates to this document will be announced on the MOPITT News webpage at https://www2.acom.ucar.edu/mopitt/news .

Features incorporated into the V9 retrieval product include:

-    **Forward model revisions. Details in Sec. 3.1.**

-    **Cloud detection revisions. Details in Sec. 3.4.**

-    **NIR calibration method revisions. Details in Sec. 3.6.**

2   Overview of Earlier MOPITT Products

Following the launch of Terra near the end of 1999, the MOPITT Version 3 (V3) product became available in 2000 [1]. This was the first satellite dataset for tropospheric CO featuring global coverage. This product was followed in 2009 by the Version 4 (V4) product [2]. Significant improvements first introduced in the V4 product included (1) temporally and geographically variable a priori for CO based on the chemical transport model MOZART, (2) representation of CO variability by log-normal statistics, (3) extension of the forward radiative transfer model MOPFAS to simulate much higher CO concentrations, and (4) improved observation-based a priori values for surface emissivity. Processing of the V4 product ended at the end of 2012. The V5 product became available in 2011 in three retrieval configurations: thermal infrared-only ("TIR-only"), near infrared-only ("NIR-only") and multispectral ("TIR/NIR"). (The previous V3 and V4 products were TIR-only products.) V5 retrieval products also benefited from a new time-dependent radiative transfer model to reduce the influence of long-term instrumental changes on retrieval biases (i.e., "bias drift"). The V5 product was validated using a variety of in-situ and satellite datasets [3,4]. The V6 product was released in 2013 and featured (1) correction of a systematic geolocation error, (2) use of the MERRA (Modern-Era Retrospective Analysis For Research And Applications) reanalysis for needed meteorological fields and (3) an updated model-based a priori for CO [5]. Features introduced in the V7 product which became available in 2016 included (1) representation of increasing atmospheric concentrations of N_2O, (2) use of meteorological fields from the MERRA-2 reanalysis and (3) an improved NIR calibration method [6]. The V8 product, which was released in 2018, included the use of (1) updated spectroscopic data, (2) the MODIS Collection 6.1 Cloud Mask (instead of Collection 6.0), and (3) a more refined radiance bias correction scheme involving time-dependent and water vapor-dependent bias correction terms.

3   Features of the MOPITT Version 9 Retrieval Algorithm

Algorithm improvements affecting the content of the V9 retrieval products are described in the following section. Significant changes in the V9 algorithms and/or products are emphasized in **bold** text.

3.1   Radiative Transfer Modeling

Improvements made to the operational radiative transfer model for V7 were described in [6]. For V8, spectroscopic data for H_2O and N_2 were each updated. Water vapor continuum absorption in the V8 operational radiative transfer model is based on the MT_CKD version 3.2 spectroscopic model from AER [7]. Water vapor continuum absorption in earlier MOPITT radiative transfer models was based on MT_CKD version 1.0. The radiative effects of molecular nitrogen (N_2), which were not previously represented, are now derived from the line-by-line model GENLN3 [8]. In addition, the globally-fixed concentration of CO_2 assumed in the operational radiative transfer model was increased to 410 ppm. The operational radiative transfer model for V9 is based on HITRAN12, which is the same version of HITRAN used for MOPITT V7 and V8 processing. **For V9, operational modeling of the MOPITT Pressure Modulated Cell (PMC) radiances (7A and 7D) now includes a time-dependent optical depth, which is adjusted monthly in MOPFAS when the PMC pressure and temperature are also updated. This is an improved description of the PMC. Revised radiance bias correction factors for Channel 7 (which will be documented in the V9 validation paper) are weaker than the corresponding correction factors for V8, indicating that the PMC model revisions have resolved a substantial component of the discrepancy between observed and model-calculated radiances for Channel 7.**

3.2   Retrieval Grid

V9 retrieved profiles are expressed on the same ten-level grid (surface, 900 hPa, 800 hPa, ..., 100 hPa) used for products since V4. For V4 products, each retrieval level was assigned to a layer centered on that level, with a weighting that tapered to zero at the adjacent retrieval levels. For all later products, in contrast, each retrieval level simply corresponds to a uniformly-weighted layer immediately above that level. For example, for retrievals where the surface pressure exceeds 900 hPa, the V9 surface-level retrieval product corresponds to the mean volume mixing ratio over the layer between the surface and 900 hPa. Thus, when comparing MOPITT profiles with model results, we suggest first applying simple (unweighted) averaging to the model results in the layers above each retrieval level. (For the topmost MOPITT retrieval level at 100 hPa, the uniform-VMR layer extends from 100 hPa to 50 hPa. Assumed VMR values in the layer from 50 hPa to TOA are based on the CAM-chem model climatology and are fixed.)

3.3   *Meteorological Fields and A Priori*

For each observation at a particular time and geographic location, the MOPITT retrieval algorithm requires temperature and water vapor profiles as well as a priori surface temperature values.   For V6 processing, meteorological profiles were derived from the NASA MERRA reanalysis product. For V7, V8 and V9 products, meteorological profiles are extracted from the improved MERRA-2 product (https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/).   Generally higher quality retrieval results are expected using MERRA-2 (instead of MERRA) due to the assimilation of more satellite datasets (e.g., IASI) and other improvements.

MOPITT retrieval products rely on variable a priori to represent the geographical and seasonal variability of "background" concentrations of CO. For the V4 and V5 products, the a priori was based on monthly CO climatologies for the years 1997–2004 as simulated with the MOZART chemical transport model. For the V6, V7, V8 and V9 products, the CO a priori is based on monthly climatologies for the years 2000–2009 simulated with the CAM-Chem model [5]. MOPITT a priori log(VMR) values vary by month, but do not vary from year to year; this simplifies the interpretation of long-term trends in the data. For all MOPITT product versions, model-based climatologies used to generate the a priori are gridded at one degree (lat/lon) horizontal resolution and monthly temporal resolution. Spatial and temporal interpolation are used to generate a priori values at each specific observation location and day.

3.4   *Cloud Detection*

Only MOPITT observations made in clear sky conditions are used in Level 2 retrieval processing. The clear/cloudy determination is based both on MOPITT's thermal-channel radiances and the MODIS cloud mask. With respect to the MOPITT thermal-channel test, the ratio of the observed MOPITT Channel 7 Average radiance and the corresponding model-calculated value is compared to a predefined global threshold value. If the radiance ratio is less than the threshold value, clouds are indicated and that MOPITT observation is flagged as cloudy by the MOPITT test. **For V9, the radiance ratio for each MOPITT retrieval is reported in the new diagnostic "MOPCld Rad Ratio".**

The overall outcome of the MOPITT cloud detection algorithm for a particular retrieval can be determined using the "Cloud Description" diagnostic in the Level 2 files. The Cloud Description diagnostic values (1-6) are defined as follows. A value of 1 means "MOPITT clear, MODIS cloud mask unavailable," a value of 2 means "MOPITT and MODIS cloud mask agree on clear," a value of 3 means "MOPITT cloudy, MODIS cloud mask clear," a value of 4 means "MOPITT clear, MODIS indicating low clouds," and a value of 5 means "polar regions, MODIS cloud mask clear." Finally, a value of 6 is assigned to observations where the MODIS cloud mask indicates that the area was cloudy but the test based on MOPITT's thermal-channel radiances finds that the area was clear. This last class was first

introduced in the V7 product and was originally applied to ocean scenes as a response to declining quality in the MODIS cloud mask [9].

For the V9 product, two significant changes were implemented in the revised cloud detection algorithm [10]. The first change is related to the interpretation of the four possible MODIS cloud mask outcomes, whereas the second change concerns the treatment of observations deemed cloudy by the MODIS cloud mask but clear by the MOPITT thermal channel test. Together, these changes significantly increase MOPITT retrieval coverage over land.

The MODIS cloud mask reports one of four possible outcomes for each MODIS 1 km pixel: Cloudy, Uncertain, Probably Clear, or Clear. Prior to V9, the MOPITT cloud detection algorithm interpreted the Probably Clear and Clear outcomes as clear and treated the Cloudy and Uncertain outcomes as cloudy. If at least 95% of the MODIS cloud mask pixels enclosed within a given MOPITT pixel indicated either Probably Clear or Clear, that MOPITT pixel was considered clear according to MODIS. For V9 processing, the MODIS cloud mask test was relaxed to treat Uncertain MODIS pixels as clear (in the same manner as Clear and Probably Clear MODIS pixels). This change was motivated by the observation that such MODIS pixels can be found in apparently cloudless but heavily polluted scenes [10].

Also in V9, observations over land are only discarded if both the MODIS cloud mask and MOPITT radiance tests indicate the presence of clouds. This change was introduced earlier for observations over the ocean, beginning with V7 products. It allows MOPITT retrievals in cases for which the MODIS cloud mask tests indicate clouds (or are ambiguous) while the MOPITT TIR radiances are consistent with clear-sky conditions. Consequently, this change should allow the retrieval of scenes for which clouds in the MOPITT field of view have a negligible effect on the MOPITT radiances. MOPITT retrievals for which the MODIS cloud mask considers the observation to be cloudy while the MOPITT thermal channel test passes the observation as clear are assigned the Cloud Description index of 6 and can therefore be analyzed separately from retrievals where MODIS determined the scene to be clear. Prior to V9, this value for the Cloud Description index was only allowed for observations over the ocean.

Finally, a minor change was also made in the revised cloud detection algorithm regarding cloud index 4 (MOPITT clear, MODIS indicating low clouds). In the revised algorithm, this index is only applied to observations over the ocean, where low clouds are more reliably detected. Retrievals over land which would have been assigned a cloud index value of 4 in the V8 algorithm are assigned a cloud index value of 6 in V9. Thresholds for the MODIS cloud mask and MOPITT thermal channel tests for V9 are unchanged relative to the values used for the MOPITT Version 8 product, i.e., the MODIS clear-sky fraction threshold was set to 0.95 and the MOPITT radiance

ratio threshold was set to 1.00.

A separate diagnostic is provided for each retrieval in the Level 2 product file to quantify the results of the various cloud tests applied to the set of MODIS Cloud Mask pixels matched to each MOPITT pixel. This diagnostic, which has been revised for V9, may be of use for analyzing potential retrieval biases associated with particular types of clouds**. For V9, elements of the 12-element "MODIS Cloud Diagnostics" floating point vector indicate:**

> **(1) Number of "determined" MODIS pixels**
> **(2) Percentage of cloudy MODIS pixels**
> **(3) Percentage of clear MODIS pixels, test 1**
> **(4) Percentage of clear MODIS pixels, test 2**
> **(5) Percentage of clear MODIS pixels, test 3**
> **(6) Average value of "sun glint" MODIS flag**
> **(7) Average value of "snow/ice background" MODIS flag**
> **(8) Average value of "non-cloud obstruction" MODIS flag**
> **(9) Average value of "IR threshold test" MODIS flag**
> **(10) Average value of "IR temperature difference tests" MODIS flag**
> **(11) Average value of "visible reflectance test" MODIS flag**
> **(12) Fraction of "determined" MODIS pixels**

Elements 3, 4, and 5 of the Cloud Diagnostics vector report the percentage of clear-sky MODIS pixels within the MOPITT pixel according to three tests with varying levels of confidence. Test 1 (vector element 3) reports the clear-sky percentage based solely on Clear outcomes for the MODIS cloud mask and is therefore the strictest test. Test 2 (element 4) reports the clear-sky percentage considering both Clear and Probably Clear MODIS cloud mask outcomes as clear, and corresponds to the cloud mask test used in prior versions of the MOPITT cloud detection algorithm. Finally, test 3 (element 5) reports the clear-sky percentage considering Clear, Probably Clear, and Uncertain MODIS cloud mask outcomes as clear. For V9, this last test actually determines whether MODIS classifies the MOPITT pixel as clear or cloudy. Elements 3 and 5 in the Cloud Diagnostics vector represent information not previously included in the MOPITT product.

3.5   *Radiance Bias Correction*

The MOPITT Level 2 processor exploits a set of radiance-bias correction factors to compensate for relative biases between (1) simulated radiances calculated by the operational radiative transfer model and (2) actual calibrated Level 1 radiances. Without some form of compensation, radiance biases produce biases in the retrieved CO profiles. Radiance bias correction factors compensate for a variety of potential bias sources including errors in instrumental specifications, forward model errors,

spectroscopy errors, and geophysical errors. For V8 and V9 products, radiance bias correction is based on a parameterization involving the date of observation and water vapor total column (derived from MERRA-2 water vapor profiles). V8 validation results [11] demonstrated that this strategy successfully reduces long-term bias drift and retrieval bias geographical variability. Coefficients used to correct radiance biases and bias drift are obtained through a series of experiments in which biases are primarily evaluated using the large set of NOAA aircraft profiles (mainly over North America). Radiance bias correction factors used in the processing of each retrieved profile are reported in the L2 diagnostic "L2 Radiance Correction Factor" (see Appendix A).
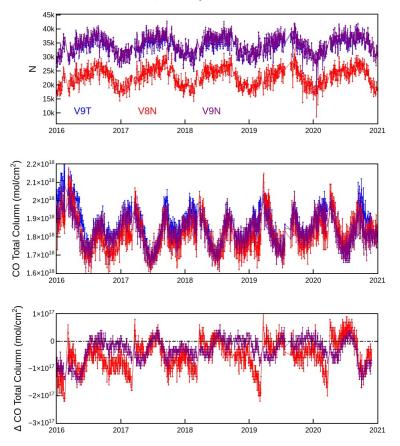
3.6   NIR Calibration

Calibration of MOPITT's NIR channels (e.g., Channel 6) relies on a two-point calibration scheme involving both cold-calibration ("cold-cal") events and hot-calibration ("hot-cal") events. Cold-cals occur many times per day, while hot-cals are performed only about once per year. Ideally, NIR channels are calibrated with gain and offset values determined by interpolating the information from hot-cals occurring both before and after the time of observation. While this method is feasible in retrospective processing mode (i.e., processing previous years of data), it is not possible in forward processing mode (i.e., when processing recently acquired observations). Thus, in forward processing mode, only information from the most recent hot-cal is used to calibrate MOPITT's NIR radiances. Comparisons of NIR-only retrieval products generated in retrospective and forward processing modes exhibit significant differences (10% to 20%) in total column results, with the retrospectively processed data in better agreement with daytime/land TIR-only total column values and time dependence. Therefore, because of the degraded quality of MOPITT products processed in forward processing mode, V8 and V9 products generated in this manner are labeled as "beta" products. These products are eventually reprocessed and replaced by standard archival files following the next hot-cal. Typically, this will occur no more than a year from the time of a particular observation (depending on the date of the most recent hot-cal). The beta products should not be considered for examining long-term records of CO although these products should still be useful for some applications. See Sec. 5.

For V9, the NIR calibration methodology has been revised for retrospective processing. Hot-cals are typically performed annually, usually in March, in conjunction with a decontamination procedure. In most years, hot-cals are executed both immediately before and after the decontamination procedure. For earlier products, NIR calibration for archival (non-beta) products relied on the closest bracketing hot-cals such that, usually, NIR radiances for a given date were calibrated using the most recent previous post-decontamination hot-cal and the immediately following pre-decontamination hot-cal. For example, for V8, NIR radiances observed between March 5, 2016 and March 5, 2017 were calibrated using information from the

post-decontamination hot-cal on March 4, 2016 and the pre-decontamination hot-cal on March 6, 2017.

However, it was recently discovered that this NIR calibration strategy often results in a growing retrieval bias in the NIR-only products over the period between the two hot-cals used for calibration. As illustrated in Fig. 1, this time-dependent bias is most obvious when comparing TIR-only and NIR-only CO products immediately before and after the hot-cal/decontamination cycle. In the top panel of the figure, timeseries plots are shown for the total number of daytime retrievals over land regions between 60° S and 60° N for the V9 TIR-only (V9T), V8 NIR-only (V8N) and V9 NIR-only (V9N) products. Numbers of daytime/land retrievals for the V9T and V9N products are substantially greater than for the V8N product due to the cloud detection revisions described in Sec. 3.4. Corresponding timeseries plots of daily-mean CO total column values are shown in the middle panel. Finally, in the bottom panel, timeseries are shown for the 'Delta' CO total columns obtained by subtracting daily-mean V9T CO total column values from the V8N and V9N products. While NIR-only and TIR-only retrievals are characterized by different vertical sensitivities, and are therefore not expected to agree precisely, V9T CO total column values are unaffected by NIR calibration issues and therefore serve as a useful reference. For each of the years shown, the Delta CO total column timeseries for V8N (plotted in red) exhibits a physically unrealistic discontinuity when comparing dates just before a pre-decontamination with dates just after the post-decontamination several weeks later. For example, in 2019, Delta CO total column for V8N increased from about -2×10^{17} molecules/cm^2 just before the pre-decontamination hot-cal to about 0 molecules/cm^2 just after the post-decontamination hot-cal. While the physical source of this discontinuity is not yet fully understood, it suggests that the pre- and post-decontamination hot-cals are not equally useful for calibration.

V9T vs V8N and V9N, 60S:60N, Land, Daytime, 2016-2020

Figure 1. Timeseries comparisons of N (daily number of daytime/land retrievals), CO Total Column, and Delta CO Total Column (as described in Sec. 3.6) for the V9T, V8N and V9N products.

Experiments were performed to identify an improved NIR calibration strategy for V9. It was found that the typical discontinuity in Delta CO total column values before and after the hot-cal and decontamination cycle was greatly reduced when only post-decontamination hot-cals were used for calibration. The Delta CO total column timeseries using this strategy, which was implemented for V9N operational processing, is plotted in purple in the bottom panel of Fig. 1. For each of the years shown, the improved stability of the V9N product compared to V8N is readily apparent. Further details regarding the specific hot-cals used for NIR calibration in V9 over the entire MOPITT mission will be reported in a forthcoming revision of the L0-L1 Algorithm Theoretical Basis Document (ATBD).

4   Product Format and Content

The contents of the V9 Level 2 and Level 3 product files are described and tabulated in Appendices [A](#)
and [B](#) respectively. Beginning with the MOPITT V6 products, the format of the archival Level 1, Level
2 and Level 3 data files switched from HDF-EOS2, based on HDF4 libraries, to HDF-EOS5, based on
HDF5 libraries. **V9 products are in the same HDF5 format as V6, V7 and V8 products and should
be readable with analysis tools developed for those earlier products.** HDF is a standard format for
large datasets (https://www.hdfgroup.org/). The HDF-EOS format is an extension of HDF developed by
the HDF-EOS group (http://hdfeos.org/index.php). Files distributed in the HDF-EOS5 format can be
accessed with a wide range of software tools including IDL, NCL and MATLAB
(https://support.hdfgroup.org/products/).

4.1   Level 1 Data

The format and contents of the Level 1 product files containing the MOPITT calibrated radiances is
unchanged for V9.

4.2   Level 2 Data

The contents of the V9 Level 2 product files are described in table form in Appendix [A](#). **The reported
fields in V8 and V9 Level 2 files are identical, with the exception of the added diagnostic
'MOPCld Rad Ratio' and added information in the diagnostic 'MODIS Cloud Diagnostics' (see
Sec. 3.4).**

Appendices [C](#), [D](#) and [E](#) provide examples of IDL, NCL and Python code, respectively, used to open and
read the contents of V9 Level 2 data files.

4.3   Level 3 Data

MOPITT Level 3 products are available as gridded daily-mean and monthly-mean files. Data averaging
is performed on a one-degree latitude/longitude grid. Compared to Level 2 products, these gridded
products typically are less affected by random retrieval errors (e.g., due to instrument noise or
geophysical noise) and are packaged into much smaller HDF5 data files. Within each type of Level 3
file, data-averaged products are separately available for daytime and nighttime MOPITT observations.
The contents of the V9 L3 files are described in Appendix [B](#); for brevity, only daytime-overpass L3
fields are listed. V9 L3 files include several diagnostics first introduced in the V7 product, including the
total column averaging kernel, dry air column, water vapor column, smoothing error covariance matrix
and measurement error covariance matrix.

V9 Level 3 processing employs the same strategy for improving information content used for V7 and
V8. The method relies on both pixel filtering and signal-to-noise ratio (SNR) thresholds for Channel 5

and 6 Average radiances (i.e., 5A and 6A). SNR values for each observation are obtained using radiance and radiance uncertainty values contained in the diagnostic "Level1 Radiances and Errors" in the Level 2 data files. Observations from Pixel 3, one of the four elements of MOPITT's linear detector array, are excluded from TIR-only and TIR/NIR L3 products due to highly variable Channel 7 SNR values. Observations with relatively low SNR tend to yield retrievals weighted heavily by the a priori and generally result in low DFS. Excluding low-SNR observations from the Level 3 cell-averaged values raises overall mean DFS values. Specific filtering rules used for both the daily-mean and monthly-mean V9 Level 3 products are:

- V9 Level 3 TIR-only products exclude all observations from Pixel 3 in addition to observations where the 5A SNR < 1000

- V9 Level 3 NIR-only products exclude all observations where the 6A SNR < 400

- V9 Level 3 *daytime* TIR/NIR products exclude all observations from Pixel 3 in addition to observations where both (1) the 5A SNR < 1000 *and* (2) the 6A SNR < 400

- V9 Level 3 *nighttime* TIR/NIR products exclude all observations from Pixel 3 in addition to observations where the 5A SNR < 1000.

5   Data Analysis Considerations

5.1   Data Quality

Compared to V7, validation results for V8 CO products indicated reduced long-term bias drift, weaker bias geographical variability and smaller biases overall [11]. Similar results are expected for the V9 product and will be documented in an upcoming validation paper.

As described in Sec. 3.6, MOPITT retrieval products processed without bracketing calibration data (obtained during both previous and subsequent annual calibration events) are considered less reliable than those processed with such data, and are therefore labeled as beta products. V9 beta products are easily identifiable as the filenames of such files include the word "beta" (e.g., "MOP02T-20210501-L2V19.9.1.beta.he5"). For example, at the ASDC http site (https://asdc.larc.nasa.gov/data/MOPITT/), beta files are stored in separate directories from archival products: the archival V9 Level 2 TIR-only products will be found in subdirectories under the main directory "MOP02T.009" while corresponding V9 beta TIR-only products will be found in subdirectories under "MOP02T.109". For the initial release of V9 in mid-2021, beta products will only be produced and delivered to the data archives for observations made after March, 2021.

Another type of retrieval anomaly has been noticed in a small number of nighttime/land scenes. This anomaly is typically manifested as a striping pattern in retrieved surface-level CO concentrations and surface temperature for one or two pixels in the four-element detector array over a series of tracks. This type of anomaly primarily affects Pixels 1 and 4 and appears to be partly the result of a calibration issue affecting Channel 5 radiances in which gain and offset values exhibit frequent sudden jumps throughout the day. These gain and offset discontinuities have been traced back to cold calibration (space-view) events. The Daily Gain Dev diagnostic (included in both Level 1 and Level 2 files) may be useful for identifying days where this type of anomaly is most likely to occur. However, analyses of "outlier" retrievals suggest that retrieval anomalies due to this effect are typically small (e.g., less than 10 ppbv).

Retrieval data quality may also be affected by forward model error, errors in assumed meteorological profiles, and instrumental degradation. For these effects, robust estimates of the associated retrieval errors are not feasible. Quantifying such errors is also complicated because they may be highly variable in space and time.

5.2   *Data Averaging*

Because of both geophysical and instrumental noise, individual MOPITT retrieved profiles are not typically useful for scientific analysis. This is especially true for the NIR-only and TIR/NIR products. Instead, users should employ statistical methods applied to ensembles of retrievals. Simple data averaging is the most common statistical method for reducing the effects of random instrument noise. The tradeoff to this benefit is reduced spatial and/or temporal resolution. When averaging over an ensemble, retrieval uncertainties associated with random radiance errors (such as geophysical noise and instrumental noise) decrease approximately as the square root of the number of retrievals in the ensemble.

The retrieval state vector represents CO volume mixing ratio (VMR) variability as a log-normal quantity. This has important implications for data averaging. For MOPITT data subsets where the main source of retrieval variability is random retrieval noise (rather than CO variability), retrieved VMR values will follow a log-normal distribution. In such cases, the appropriate method for reducing the effects of instrument noise is to (1) convert individual VMR profiles into log(VMR) profiles, (2) compute the mean, and (3) convert the mean log(VMR) profile back to a VMR profile. Because of the asymmetry of the log-normal distribution, directly computing the mean VMR profile from a set of VMR profiles in such cases will produce a positive bias relative to "truth." In contrast, when CO variability is the dominant source of retrieval variability (such as when averaging over large regions and temporal intervals), directly calculated mean VMR values are most meaningful.

5.3 Use of Averaging Kernels

For scientific analysis, users should understand the role of the a priori on the MOPITT retrieved profiles and total columns. The relative weight of the a priori profile x_a in the retrieved profile x_{rtv} can be described by the relation

$$x_{rtv} \approx x_a + A(x_{true} - x_a) = A x_{true} + (I - A) x_a \tag{1}$$

where x_{true} is the true profile, A is the averaging kernel matrix, and I is the identity matrix. For MOPITT retrievals, x_{rtv}, x_{true}, and x_a are all expressed as profiles of $\log_{10}$(VMR). Both A and x_a are highly variable and are provided as ancillary data for each MOPITT retrieved profile. Averaging kernels indicate the sensitivity of x_{rtv} to different levels of the atmosphere and exhibit significant spatial variability, both vertical and horizontal. Inclusion of a priori information, as described by Eq. 1, is associated with all remote sensing methods based on optimal estimation. The retrieved profile is most heavily weighted by the a priori profile at levels where the actual measurements are least sensitive. An important consequence of this effect is that, when comparing MOPITT retrieved profiles against in-situ data or model results, users must transform these comparison datasets using Eq. 1, so that the comparison data exhibit the same degree of smoothing and a priori dependence as the MOPITT product [1,2]. For the V9 products (like the earlier V5, V6, V7 and V8 products), users must make this transformation in terms of $\log_{10}$(VMR) rather than VMR. Users should also observe the level-layer associations described in Sec. 3.2. Users of the MOPITT total column product (rather than retrieved profiles) should refer to Sec. 5.5.

5.4 Data Filtering

Because of the potential variability of MOPITT data quality, as discussed above, users may decide to exclude particular subsets of the Level 2 data in their analyses. This is a reasonable strategy, *unless such methods could impose retrieval bias*. For example, filtering based on the "Degrees of Freedom for Signal" (DFS) parameter tends to retain high-CO profiles and discard low-CO profiles, thereby imposing a positive retrieval bias on the filtered subset. Filtering on DFS, or other characteristics of the retrieval averaging kernels, is therefore not recommended.

Alternatively, filters based on geophysical criteria which are known to provide better conditions for remote sensing retrievals may be appropriate. For example, over land, daytime conditions typically provide better thermal contrast conditions for TIR- and TIR/NIR-based retrievals than nighttime conditions.

5.5 Analysis of Retrieved CO Total Column Values

For users interested in quantitatively comparing MOPITT retrieved CO total column values with other

datasets (or model output), the V9 Level 2 and Level 3 product files include the total column averaging kernel a; this vector quantifies the sensitivity of the retrieved total column to perturbations at each level of the CO profile. Thus, if C_{rtv} is the retrieved CO total column, and x is the CO state vector comprised of CO $\log_{10}$(VMR) values, the total column averaging kernel element for level j in the CO profile is defined by

$$a_j \;=\; \frac{\partial C_{rtv}}{\partial x_j} \quad . \tag{2}$$

Given a comparison CO profile x_{cmp} (e.g., either from in-situ measurements or model output), the total column averaging kernel vector is used to simulate MOPITT total column retrievals using the equation

$$C_{sim} \;=\; C_a \;+\; a(x_{cmp} - x_a) \tag{3}$$

where C_a is the a priori total column value corresponding to the a priori profile x_a. Here, as in Eq. 1, x_{cmp} and x_a are both expressed as $\log_{10}$(VMR) profiles. Both C_a and x_a are provided in the V8 Level 2 product files for each retrieval.

For V8 and V9 products, the method of calculating a has been revised for consistency with Rodgers [12], Sec. 4.3. Specifically, the total column averaging kernel is now calculated as

$$a^T \;=\; h^T A \tag{4}$$

where h is the partial column operator (i.e., the vector of sensitivities of partial column values to $\log_{10}$(VMR)).

5.6 *Conversions Between VMR and Partial Column CO Profiles*

When comparing MOPITT data to either model-generated output or other remote sensing datasets, users may need to convert CO vertical profiles expressed in partial column amounts to profiles expressed in terms of volume mixing ratio (or vice-versa). Below we derive the scaling factor relating the CO *VMR* value in a given atmospheric layer (e.g., in ppbv) and the corresponding CO partial column in that layer (e.g., in molecules/cm^2).

The hydrostatic equation can be written

$$\frac{\Delta p}{\Delta z} \;=\; -\rho g \tag{5}$$

where (in SI units) Δp is the pressure thickness of some atmospheric layer (in Pa), Δz is the geometric thickness of that layer (in m), ρ is the air mass density within that layer (in kg/m^3), and g is the acceleration of gravity (in m/s^2) [13]. In the following we ignore the negative sign in the hydrostatic equation, which is unimportant in the context of this derivation.

The air mass density ρ can be written as the sum of separate components ρ_{dry} and ρ_W representing dry air and water vapor, respectively. These two terms are related to the dry air number density n_{dry}, dry air molecular weight M_{dry}, water vapor number density n_W and water vapor molecular weight M_W such that

$$\rho \;=\; \rho_{dry} \;+\; \rho_w \;=\; \frac{n_{dry}M_{dry}}{N_A} \;+\; \frac{n_w M_w}{N_A} \tag{6}$$

where N_A is Avogadro's number (6.0221×10^{23} molecules/mole). In SI units, n_{dry} and n_W are expressed in molecules/m^3 while M_{dry} and M_W are expressed in kg/mole. Also, since the water vapor volume mixing ratio (or 'mole fraction') x_W is given by

$$x_W \;=\; \frac{n_W}{n_{dry} \;+\; n_W} \;=\; \frac{n_W}{n_T} \quad, \tag{7}$$

where n_T is the total number density, the air mass density equation can be rewritten as

$$\rho \;=\; \frac{n_T\left[(1-x_W)M_{dry} \;+\; x_W M_W\right]}{N_A} \;=\; \frac{n_T M_{eff}}{N_A} \tag{8}$$

where M_{eff} is the effective moist-air molecular weight (i.e., the weighted mean of the molecular weights of dry air and water vapor). Substituting the RHS of this equation into the hydrostatic equation and solving for $n_T \Delta z$, we obtain

$$n_T \Delta z \;=\; \frac{\Delta p \; N_A}{g \, M_{eff}} \quad. \tag{9}$$

Equation 9 gives the moist-air partial column (number of molecules of air per unit area, including both dry air and water vapor) for a layer of pressure thickness Δp and effective molecular weight M_{eff}. For a trace gas in that layer with volume mixing ratio *VMR*, the trace gas partial column C is then

$$C \;=\; VMR \, n_T \Delta z \;=\; \frac{VMR \, \Delta p \, N_A}{g \, M_{eff}} \quad. \tag{10}$$

In Equation 10, C will have units of molecules/m^2 provided that *VMR* is dimensionless, Δp has units of Pa, g has units of m/s^2, and M_{eff} has units of kg/mole. Switching the units of *VMR* from ppv to ppbv (*VMR'*), Δp from Pa to hPa ($\Delta p'$), M_{eff} from kg/mole to g/mole (M_{eff}') and C from molecules/m^2 to molecules/cm^2 (C') yields

$$C' \;=\; 10^{-8}\frac{VMR' \, \Delta p' \, N_A}{g \, M_{eff}'} \quad. \tag{11}$$

Example 1. For dry air ($x_W = 0$), $M_{eff}' = 28.97$ g/mole [13]. For this scenario,

$$C'(molecules/cm^2) = 10^{-8} \ VMR'(ppbv) \ \Delta p'(hPa) \ (6.0221(10^{23})) \ / \ [(9.806)(28.97)]$$
$$= 2.120(10^{13})VMR'(ppbv)\Delta p'(hPa)$$

Example 2. In very warm and humid atmospheric conditions, the water vapor mole fraction x_W may reach values as high as 0.03 [13]. The corresponding effective molecular weight in this case is $M_{eff}' =$ 28.64 g/mole ($M_{eff}' = 0.97(28.97$ g/mole$) + 0.03(18.02$ g/mole$)$). In this scenario,

$$C'(molecules/cm^2) = 10^{-8} \ VMR'(ppbv) \ \Delta p'(hPa) \ (6.0221(10^{23})) \ / \ [(9.806)(28.64)]$$
$$= 2.144(10^{13})VMR'(ppbv)\Delta p'(hPa)$$

Thus, an increase of x_W from 0.00 (*Example 1*) to 0.03 (*Example 2*) increases C' only by 1.13%.

As illustrated in these two examples, the VMR to partial column scaling factor derived for extremely moist conditions is only different by the dry-air scaling factor by a little more than 1%. Consequently, if an error of this magnitude is acceptable for a particular application, there is no need to account for the effect of water vapor variability on M_{eff}'. Thus, the scaling factor derived in the first example above can be used generally.

6   References

[1] Deeter, M. N., et al. (2003). Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, J. Geophys. Res. 108, doi:10.1029/2002JD003186.

[2] Deeter, M. N., et al. (2010). The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability. J. Geophys. Res. 115, doi:10.1029/2009JD013005.

[3] Deeter, M. N., et al. (2013). Validation of MOPITT Version 5 Thermal-infrared, near-infrared, and multispectral carbon monoxide profile retrievals for 2000-2011. J. Geophys. Res. Atmos., 118, 6710–6725, doi:10.1002/jgrd.50272.

[4] Martínez-Alonso, S., et al. (2014). Comparison of upper tropospheric carbon monoxide from MOPITT, ACE-FTS, and HIPPO-QCLS. J. Geophys. Res. 119, 14,144–14,164, doi:10.1002/2014JD022397 .

[5] Deeter, M. N., et al. (2014). The MOPITT Version 6 product: algorithm enhancements and validation. Atmos. Meas. Tech., 7, 3623-3632, doi:10.5194/amt-7-3623-2014 .

[6] Deeter, M. N., et al. (2017). A climate-scale satellite record for carbon monoxide: The MOPITT

Version 7 product, Atmos. Meas. Tech., 10, 2533-2555, doi:10.5194/amt-10-2533-2017 .

[7] Mlawer, E. J., et al. (2012). Development and recent evaluation of the MT_CKD model of continuum absorption, Phil. Trans. Royal Soc. A, 370, 2520–2556, doi:10.1098/rsta.2011.0295 .

[8] Edwards, D. P. (1992). GENLN2: A general line-by-line atmospheric transmittance and radiance model, Technical Report NCAR/TN-367+STR, NCAR, P.O. Box 3000, Boulder, Colorado, 80307-3000, 147 pp.

[9] see https://modis-atmosphere.gsfc.nasa.gov/sites/default/files/ModAtmo/C6.1_Calibration_and_Cloud_Product_Changes_UW_frey_CCM_1.pdf

[10] Deeter, M. N., et al. (2021). Impacts of MOPITT cloud detection revisions on observation frequency and mapping of highly polluted scenes, accepted for publication in Rem. Sens. Env.

[11] Deeter, M. N., et al. (2019). Radiance-based retrieval bias mitigation for the MOPITT instrument: The Version 8 product, Atmos. Meas. Tech., 12, 4561–4580, doi:10.5194/amt-12-4561-2019.

[12] Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding*, Theory and Practice, World Scientific, Singapore, 2000.

[13] Wallace, J. M., and Hobbs, P. V., *Atmospheric Science: An Introductory Survey*, Academic Press, San Diego, California, 1977.

Appendices

A. Tabulated Contents of V9 Level 2 Product Files

DIMENSIONS

Dimension Name	Definition	Size
nTime	Number of retrievals	Varies with each granule
nChan	Number of channels of the instrument	8
nTwo	For data fields which include associated variability (or uncertainty), first element is the parameter and second element is variability/uncertainty, unless otherwise specified.	2
nPrs	Number of fixed pressure levels	9
nPrs2	Number of pressure levels, including the surface	10
nPix	The number of pixels (i.e., number of detector elements)	4

GEOLOCATION FIELDS ('HDFEOS/SWATHS/MOP02/Geolocation Fields/…')

Field Name	Definition	Dimension	Units	Data Type
SecondsinDay	Time of the measurement in seconds from the beginning of the day.	nTime	s	F
Latitude	Geolocation of the pixel	nTime	deg	F
Longitude	Geolocation of the pixel	nTime	deg	F
Pressure	Pressure Levels	nPrs	hPa	F
Pressure2	Pressure Levels including the surface	nPrs2	hPa	F
Time	TAI Time of the observation	nTime	s	F
F = Floating Point **I** = Integer				

RETRIEVED CO FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

Field Name	Definition	Dimension	Units	Data Type
RetrievedCOTotalColumn	CO Total Column	nTwo, nTime	mol/cm^2	F
RetrievedCOMixingRatioProfile	CO Mixing Ratio for the layer above each pressure level	nTwo, nPrs, nTime	ppbv	F
RetrievedCOSurfaceMixingRatio	CO Mixing Ratio at the surface	nTwo, nTime	ppbv	F

A PRIORI FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

Field Name	Definition	Dimension	Units	Data Type
APrioriCOMixingRatioProfile	Temporally and geographically variable a priori CO profile based on a climatology for 2000-2009 simulated with the CAM-Chem model	nTwo, nPrs, nTime	ppbv	F
APrioriCOSurfaceMixingRatio	Temporally and geographically variable a priori surface CO based on a climatology for 2000-2009 simulated with the CAM-Chem model	nTwo, nTime	ppbv	F
APrioriCOTotalColumn	Temporally and geographically variable a priori total column CO based on a climatology for 2000-2009 simulated with the CAM-Chem model	nTwo, nTime	mol/cm^2	F

AVERAGING KERNEL FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

Field Name	Definition	Dimension	Data Type
RetrievalAveragingKernelMatrix	For each retrieval, a floating point array (10 x 10) containing the matrix describing the sensitivity of the retrieved CO log(VMR) profile to the true CO log(VMR) profile. Dimensions of Retrieval Averaging Kernel Matrix are ordered (nrow, ncolumn, ntime)	nPrs2, nPrs2, nTime	F
AveragingKernelRowSums	This diagnostic provides a single scalar value for each row of the averaging kernel matrix equal to the sum of the elements in that row. Small row-sum values indicate retrieval levels heavily weighted by the a priori while values approaching unity indicate levels with relatively weak sensitivity to the a priori.	nPrs2, nTime	F
TotalColumnAveragingKernel	This diagnostic allows users to properly compare MOPITT total column retrievals with total columns derived from in-situ profiles or model simulations.	nPrs2, nTime	F
TotalColumnAveragingKernelDimless	This diagnostic allows users to properly compare MOPITT total column retrievals with total columns derived from in-situ	nPrs2, nTime	F

	profiles or model simulations based on partial columns instead of log(VMR).		

CLOUD FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

These quantities are unitless

Field Name	Definition	Dimension	Data Type
CloudDescription	0 = clear, only MOPITT thermal and cloudtop used 1 = clear, only MOPITT thermal used 2 = MOPITT and MODIS cloud mask agree on clear 3 = MODIS cloud mask only clear (when MOPITT determines cloudy) 4 = MOPITT overriding MODIS cloud mask over low clouds (MODIS test flags used) 5 = MODIS cloud mask only, clear over polar regions and daytime 6 = Nighttime ocean scenes where the MODIS cloud mask-based test finds that the area was cloudy but MOPITT's thermal-channel radiances finds that the area was clear (new for V7)	nTime	I
MODISCloudDiagnostics	(1) Number of "determined" MODIS pixels (2) Percentage of cloudy MODIS pixels (3) Percentage of clear MODIS pixels, test 1 (4) Percentage of clear MODIS pixels, test 2 (5) Percentage of clear MODIS pixels, test 3 (6) Average value of "sun glint" MODIS flag (7) Average value of "snow/ice background" MODIS flag (8) Average value of "non-cloud obstruction" MODIS flag (9) Average value of "IR threshold test" MODIS flag (10) Average value of "IR temperature difference tests" MODIS flag (11) Average value of "visible reflectance test" MODIS flag (12) Fraction of "determined" MODIS pixels	12, nTime	F
MOPCldRadRatio	Ratio of observed Channel 7 Average radiance and corresponding model-calculated value (used for cloud detection)	nTime	F

SURFACE PROPERTY FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

Field Name	Definition	Dimension	Units	Data Type
APrioriSurfaceEmissivity	A priori surface emissivity from static map	nTwo, nTime		F
APrioriSurfaceTemperature	A priori surface temperature from MERRA2 analysis	nTwo, nTime	K	F
DEMAltitude	Digital Elevation Model surface height	nTime	m	F
SurfaceIndex	0 = water 1 = land 2 = mixed (coastline)	nTime		I
SurfacePressure	Surface pressure from MERRA2 analysis	nTime	hPa	F
RetrievedSurfaceEmissivity	Surface emissivity from the retrieval	nTwo, nTime		F
RetrievedSurfaceTemperature	Surface temperature from the retrieval	nTwo, nTime	K	F

RADIANCE FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

Field Name	Definition	Dimension	Units	Data Type
DailyGainDev	This diagnostic provides the standard deviation of the gain values used to calibrate the radiances for a particular channel and pixel on one day. In this case, nTwo is the Average and Difference signal gains.	nTwo, nChan, nPix		F
Level1RadiancesandErrors	First dimension corresponds to radiances/uncertainties; second dimension corresponds to channel (in sequence 7A, 3A, 1A, 5A, 7D, 3D, 1D, 5D, 2A, 6A, 2D, 6D).	nTwo, 12, nTime	$W/m^2 Sr$	F
L2RadianceCorrectionFactor	Time- and water vapor-dependent radiance correction factors applied to each radiance (same sequence as Level1RadiancesandErrors).	12, nTime		F

RETRIEVAL DIAGNOSTICS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

These quantities are unitless

Field Name	Definition	Dimension	Data Type
DegreesofFreedomforSignal	For each retrieval, a floating point value describing the number of pieces of independent information in the retrieval, equal to the trace of the averaging kernel matrix.	nTime	F
MeasurementErrorCovarianceMatrix	Measurement error represents the retrieval log(VMR) uncertainty due to uncertainties in the measured radiances (including instrumental and geophysical 'noise').	nPrs2, nPrs2, nTime	F
RetrievalAnomalyDiagnostic	These flags are set to true when particular anomalous conditions are observed, suggesting that the retrievals should either be ignored or used cautiously. The first four flags (i.e., elements 1-4 of the Retrieval Anomaly Diagnostic array) are set to true (i.e., a value of 1) when one of the thermal channel weighting functions exhibits a sign change vertically. Elements 1-4 correspond respectively to the 5A, 5D, 7A, and 7D weighing functions. The fifth flag of the Retrieval Anomaly Diagnostic array is set to true when the retrieval averaging kernel matrix includes at least one negative element on the matrix diagonal.	5,nTime	I
RetrievalErrorCovarianceMatrix	For each retrieval, a floating point array (10 x 10) containing the a posteriori covariance matrix in base-10 log(VMR).	nPrs2, nPrs2, nTime	F
RetrievalIterations	The number of retrieval algorithm iterations needed for convergence	nTime	I
SignalChi2	A measure of the retrieval's goodness of fit.	nTime	F
SmoothingErrorCovarianceMatrix	Smoothing error represents the retrieval log(VMR) uncertainty due to the influence of the a priori and weighting functions.	nPrs2, nPrs2, nTime	F

RetrievedCOTotalColumnDiagnostics	This two-element vector contains both the estimated smoothing error and measurement error contributions (in that order) to the total retrieval error for CO total column. Smoothing error represents the uncertainty in the retrieved total column due to to the influence of the a priori and weighting functions. Measurement error describes the uncertainty due to random errors in the measured radiances.	nTwo, nTime		F

ATMOSPHERIC AND OBSERVATIONAL FIELDS ('HDFEOS/SWATHS/MOP02/Data Fields/…')

Field Name	Definition	Dimension	Units	Data Type
PressureGrid	Fixed pressure levels (i.e., 900, 800, 700, 600, 500, 400, 300, 200 and 100 hPa)	nPrs	hPa	F
SolarZenithAngle	Solar zenith angle	nTime	deg	F
SatelliteZenithAngle	Viewing angle of the satellite	nTime	deg	F
SwathIndex	For each retrieval, a three-element integer vector containing the unique 'pixel' (varies from 1 to 4), 'stare' (varies from 1 to 29), and 'track' indices. 29 stares (equivalent to the mirror pivoting out and back) define one track.	3, nTime		I
DryAirColumn	Dry-air atmospheric column to facilitate the computation of the equivalent mixing ratio of CO averaged over the atmospheric column	nTime	mol/cm^2	F
WaterVaporColumn	Moist-air atmospheric column to facilitate the computation of the equivalent mixing ratio of CO averaged over the atmospheric column	nTime	mol/cm^2	F

NOTES: Fill Values are -9999

B. Tabulated Contents of V9 Level 3 Product Files

All fields listed below reside in HDF5 directory 'HDFEOS/GRIDS/MOP03/Data Fields/…'. Find data field descriptions in Appendix A. Field names of nighttime-overpass L3 products, e.g., RetrievedCOMixingRatioProfileNight, are also included in L3 files but are not shown in tables below. All L3 fields are floating point, except for integer-valued fields SurfaceIndexDay, SurfaceIndexNight, NumberofPixelsDay and NumberofPixelsNight. L3 files also contain mean uncertainty and variability (i.e., standard deviation) statistics for fields in table marked with dagger symbol (†). For example, mean uncertainty and variability statistics for RetrievedCOMixingRatioProfileDay are included in L3 files as RetrievedCOMixingRatioProfileMeanUncertaintyDay and RetrievedCOMixingRatioProfileVariabilityDay .

GEOLOCATION FIELDS

Field Name	Definition	Dimension	Units
Latitude	Latitudes of grid cell centers	180	deg
Longitude	Longitudes of grid cell centers	360	deg
Pressure	Fixed pressure grid	9	hPa

RETRIEVED CO FIELDS

Field Name	Dimension	Units
RetrievedCOMixingRatioProfileDay†	9, 180, 360	ppbv
RetrievedCOSurfaceMixingRatioDay†	180, 360	ppbv
RetrievedCOTotalColumnDay†	180, 360	mol/cm^2

A PRIORI FIELDS

Field Name	Dimension	Units
APrioriCOMixingRatioProfileDay	9, 180, 360	ppbv
APrioriCOSurfaceMixingRatioDay	180, 360	ppbv
APrioriCOTotalColumnDay	180, 360	mol/cm^2

AVERAGING KERNEL FIELDS

Field Name	Dimension
RetrievalAveragingKernelMatrixDay	10, 10, 180, 360
TotalColumnAveragingKernelDay	10, 180, 360

SURFACE PROPERTY FIELDS

Field Name	Dimension	Units
RetrievedSurfaceTemperatureDay[†]	180, 360	K
RetrievedSurfaceEmissivityDay[†]	180, 360	
APrioriSurfaceTemperatureDay	180, 360	K
APrioriSurfaceEmissivityDay	180, 360	
DEMAltitudeDay	180, 360	m
SurfaceIndexDay (integer-valued)	180, 360	
SurfacePressureDay	180, 360	hPa

RETRIEVAL DIAGNOSTICS

Field Name	Dimension
NumberofPixelsDay (integer-valued)	180, 360
DegreesofFreedomforSignalDay	180, 360
RetrievalErrorCovarianceMatrixDay	10, 10, 180, 360
MeasurementErrorCovarianceMatrixDay	10, 10, 180, 360
SmoothingErrorCovarianceMatrixDay	10, 10, 180, 360
RetrievedCOTotalColumnDiagnosticsDay	180, 360
SignalChi2Day	180, 360

ATMOSPHERIC AND OBSERVATIONAL FIELDS

Field Name	Dimension	Units
SolarZenithAngleDay	180, 360	deg
SatelliteZenithAngleDay	180, 360	deg
DryAirColumnDay	180, 360	mol/cm^2
WaterVaporColumnDay	180, 360	mol/cm^2

NOTES: Fill Values are -9999.

C. Sample IDL Code for Opening and Reading V9 Level 2 files

```
pro example_v9_lev2

; sample IDL code for opening and reading MOPITT Version 9 Level 2 files

; note different paths for 'Geolocation Fields' and 'Data Fields'

infile = '/MOPITT/V9T/Archive/L2/200301/0101/MOP02T-20030101-L2V19.9.1.he5'

print, infile

print, 'open hdf5 file'
file_id = H5F_OPEN(infile)

; read latitudes into variable 'moplat'
print, 'read latitudes'
dataset_id = H5D_OPEN(file_id, '/HDFEOS/SWATHS/MOP02/Geolocation Fields/Latitude')
moplat = H5D_Read(dataset_id)
H5D_CLOSE, dataset_id

; read longitudes into variable 'moplon'
print, 'read longitudes'
dataset_id = H5D_OPEN(file_id, '/HDFEOS/SWATHS/MOP02/Geolocation Fields/Longitude')
moplon = H5D_Read(dataset_id)
H5D_CLOSE, dataset_id

; read retrieved profiles into variable 'rtvprofl'
print, 'read retrieved CO profiles'
dataset_id = H5D_OPEN(file_id, '/HDFEOS/SWATHS/MOP02/Data Fields/RetrievedCOMixingRatioProfile')
rtvprofl = H5D_Read(dataset_id)
H5D_CLOSE, dataset_id

; read averaging kernel matrices into variable 'avkrn'
print, 'read AK matrix'
dataset_id = H5D_OPEN(file_id, '/HDFEOS/SWATHS/MOP02/Data Fields/RetrievalAveragingKernelMatrix')
avkrn = H5D_Read(dataset_id)
H5D_CLOSE, dataset_id

H5F_CLOSE, file_id

; print lat and lon of first retrieval
print, moplat(0), moplon(0)
; print retrieved profile of first retrieval (fixed-levels only, surface retrieval missing)
print, rtvprofl(0,0:8,0)
; print AK matrix for first retrieval
print, avkrn(0:9,0:9,0)

return
end
```

D. Sample NCL Code for Opening and Reading V9 Level 2 files

```
;=====================================================;
; example-mopitt-v9-lev2.ncl
;=====================================================;
;
; Concepts illustrated:
; - For use with MOPITT CO version 9, level 2 TIR product
; - Loading CO data, AK, latitude and longitude from
;   MOPITT level 2 hdf5 file
;
; To run type:
; ncl example-mopitt-v9-lev2.ncl
;
;=====================================================;


begin

; -----------------------------
; MOPITT file location
; -----------------------------
indir = "/MOPITT/V9T/Archive/L2/200301/0101/"
fname = "MOP02T-20030101-L2V19.9.1.he5"
infile = indir+fname
print(infile)


;-------------------------------------------
; load file and extract data structures
;
; note: names of data structures
; determined from an ncl_filedump
;-------------------------------------------

print("load hdf5 file")
fin = addfile(infile, "r")

; extract longitude
moplon = fin->$"Longitude_MOP02"$

; extract latitude
```

```
moplat = fin->$"Latitude_MOP02"$

; extract profile
ret_profile = fin->$"RetrievedCOMixingRatioProfile_MOP02"$

; extract averaging kernel
avkrn = fin->$"RetrievalAveragingKernelMatrix_MOP02"$



;------------------------------------------
; print data structure information
;------------------------------------------
; Determine dimensions of data structures
printVarSummary(moplon)
printVarSummary(moplat)
printVarSummary(ret_profile)
printVarSummary(avkrn)

; print lat, lon of first retrieval
print("Latitude: "+ moplat(0)+", Longitude: "+moplon(0))

; print first retrieved profile on fixed levels - surface retrieval is in a separate data structure
print(ret_profile(0,:,0))

; print AK matrix for first retrieval
print(avkrn(0,:,:))


end
```

E. Sample Python Code for Opening and Reading V9 Level 2 files

```python
import mopittAccessFunctions as maf
import numpy
# https://docs.h5py.org/en/stable/
import h5py
import os.path
import sys

print("Using MOPITT Access Functions")
inFileName = sys.argv[1]
fName = os.path.basename(inFileName)

list = maf.listFields(inFileName)
#PRINT OUT THE LIST OF DATA FIELDS
print ("DATA FIELDS")
for field in list:
    print(field)

if (fName[0:5] == "MOP02"):
    #### LEVEL 2 CO Retrievals at 22km resolution ####
    sampleFieldName = 'SurfaceIndex'
    print("SAMPLE FIELD " + sampleFieldName)

else:
    #### LEVEL 3 CO Retrievals averaged onto a 1x1deg grid ####

    sampleFieldName = 'SignalChi2VariabilityDay'
    print("SAMPLE FIELD " + sampleFieldName)

# PULL A FIELD OUT AND USE IT
data = maf.getField(inFileName,sampleFieldName)

dataType = type(data)
print("DATA TYPE " + str(dataType))
dataShape = data.shape
print("DATA SHAPE " + str(dataShape))


#    YOU CAN USE THIS PROGRAM TO ACCESS MOPITT DATA WITH PYTHON
import numpy
import h5py
import os.path

def listFields(inFileName):
    print("READING FILE " + inFileName)
    fName = os.path.basename(inFileName)

    try:
        inFile = h5py.File(inFileName,'r')
```

```
   except:
      print("ERROR: CANNOT OPEN " + inFileName)
      return 0

   structure ='/HDFEOS/SWATHS/MOP02/Data Fields/'

   if ("MOP03" == fName[0:5]):
      structure ='/HDFEOS/GRIDS/MOP03/Data Fields/'
   try:
      fieldList = inFile[structure]
   except:
      print("UNKNOW OBJECT " + structure)
      return 0

   fields = []
   for x in fieldList:
      fields.append(x)
   try:
      inFile.close()
   except:
      print("ERROR CANNOT CLOSE " + inFileName)
      return 0
   print(inFileName + " Closed")
   return fields

def getField(inFileName,fieldName):
   print("READING FILE " + inFileName)
   fName = os.path.basename(inFileName)

   try:
      inFile = h5py.File(inFileName,'r')
   except:
      print("ERROR: CANNOT OPEN " + inFileName)
      return 0

   structure ='/HDFEOS/SWATHS/MOP02/Data Fields/'

   if ("MOP03" == fName[0:5]):
      structure ='/HDFEOS/GRIDS/MOP03/Data Fields/'


   dataName = structure + fieldName
   try:
      dataF = inFile[dataName]
      data = dataF[()]
   except:
      print("CANNOT READ " + dataName)
      return 0

   try:
      inFile.close()
   except:
```

```
    print("ERROR CANNOT CLOSE " + inFileName)
    return 0
print(inFileName + " Closed")

return data
```