



131 Hartwell Avenue
Lexington, Massachusetts
02421-3126
USA
Tel: +1 781 761-2288
Fax: +1 781 761-2299
www.aer.com

FINAL REPORT

Training on Evaluation of GEOS-Chem Using Satellite Data

TCEQ Contract No. 582-19-90498
Work Order No. 582-22-31390-010
Revision 2.0

Prepared by:
Eric Hunt, Chase Calkins, and Matthew Alvarado
Atmospheric and Environmental Research, Inc. (AER)
131 Hartwell Ave.
Lexington, MA 02421-3126

Correspondence to: malvarad@aer.com

Prepared for:
Cara Scalpone
Texas Commission on Environmental Quality (TCEQ)
Air Quality Division
Bldg. E, Room 381S
Austin, Texas 78711-3087

June 30, 2022

Document Change Record

Revision	Revision Date	Remarks
0.1	14 June 2022	Internal Version for Review
1.0	15 June 2022	Draft Submitted to the TCEQ
2.0	30 June 2022	Revised Final Report Submitted to the TCEQ

TABLE OF CONTENTS

Executive Summary.....	6
1. Introduction.....	7
1.1 Project Objectives.....	7
1.2 Background.....	8
1.3 Report Outline.....	8
2 Satellite Product Selection.....	9
2.1 TROPOMI (ESA).....	9
2.2 CRIS (NOAA).....	11
3 Evaluation Script Development.....	12
3.1 Package Description.....	12
3.1.1 TropOMI.....	13
3.1.2 CrIS.....	14
3.1.3 Statistics and Plotting.....	15
3.2 Gathering Satellite Data.....	15
3.3 Quality Screening and Observation Operators.....	17
3.4 Challenges Encountered.....	17
3.4.1 Interpolation and Layering.....	17
3.4.2 TropOMI Tropospheric NO ₂	18
3.4.3 CrIS Averaging Kernels.....	18
3.4.4 Linux to Windows Transition.....	18
3.4.5 Processing Speed.....	18
4 TCEQ Training Sessions.....	19
5 Quality Assurance.....	19
6 Conclusions.....	20
7 Recommendations for Further Study.....	20
8 References.....	21

List of Figures

FIGURE 1. PROCESSING STEPS FOR THE EVALUATION SCRIPTS. 13
FIGURE 2. EXAMPLE PLOT FROM CO_AER.PY. 16

List of Tables

TABLE 1. PROJECTED SCHEDULE FOR TCEQ WORK ORDER #582-22-31390-010..... 7

List of Acronyms

AK – Averaging Kernel

CO – Carbon Monoxide

CrIS – Cross-Track Infrared Sounder

EPA – Environmental Protection Agency

ESA – European Space Agency

NASA – National Aeronautics and Space Administration

NOAA – National Oceanic and Atmospheric Administration

NO₂ – Nitrogen Dioxide

O₃ - Ozone

QAPP – Quality Assurance Project Plan

TCEQ – Texas Commission on Environmental Quality

TIR – Thermal Infrared

TropOMI – TROPOspheric Monitoring Instrument

UV-VIS – Ultraviolet-Visible

Executive Summary

The purpose of this project was to train TCEQ staff for model performance evaluation using satellite data for GEOS-Chem (Bey et al., 2001), the global 3-D model of atmospheric chemistry driven by meteorological input from the Goddard Earth Observing System (GEOS).

After reviewing the strengths and weaknesses of different satellite datasets on air pollutants, we decided to develop GEOS-Chem evaluation scripts for Tropospheric Monitoring Instrument (TROPOMI) ozone (O_3), carbon monoxide (CO), and nitrogen dioxide (NO_2), as well as Cross-track Infrared Sounder (CrIS) O_3 and CO.

The scripts developed in this project are for the evaluation of GEOS-Chem with five satellite products: TropOMI total column CO, TropOMI total column O_3 , TropOMI tropospheric column NO_2 , CrIS total column O_3 , and CrIS total column CO. The appropriate observation operators (averaging kernel and *a priori* profiles for O_3 and NO_2 ; just averaging kernel for CO) for each satellite retrieval within a GEOS-Chem grid box within ± 1.5 hours of the satellite overpass time (13:30 local solar time) are applied to the GEOS-Chem model profile for that 3-hour window independently. The resulting smoothed GEOS-Chem columns are averaged for each grid box and day, as are the corresponding satellite results. Evaluation statistics are then calculated for all GEOS-Chem grid boxes over all days analyzed, as well as for each day.

Challenges included special processing steps for the TropOMI NO_2 and CrIS codes to obtain the correct averaging kernels, and the need to map the GEOS-Chem layering to the CrIS and TropOMI levels.

We held two training sessions for TCEQ staff to review the basic principles of evaluating atmospheric chemistry models with satellite data and on the hands-on use of our evaluation scripts.

Future work should include a full evaluation of GEOS-Chem with the scripts, with the results carefully analyzed to ensure the scripts are performing as expected. Additional statistics and plots can be created using the templates we have provided in this project.

The downloading of the satellite data could also be streamlined. As many of these data sources are now available on cloud repositories, processing on the cloud may remove the need to download the data entirely.

1. Introduction

1.1 Project Objectives

The TCEQ uses chemical transport models (CTMs) to perform regulatory modeling in support of ozone and regional haze State Implementation Plans (SIPs). This project will train TCEQ staff for model performance evaluation using satellite data for GEOS-Chem (Bey et al., 2001), the global 3-D model of atmospheric chemistry driven by meteorological input from the Goddard Earth Observing System (GEOS).

The amended schedule of deliverables for this project is given in **Error! Reference source not found.**

Table 1. Projected Schedule for TCEQ Work Order #582-22-31390-010

Milestones	Planned Date
Task 1 - Work Plan	
1.1: TCEQ-approved Work Plan	January 6, 2022
1.2: TCEQ-approved QAPP	January 6, 2022
Task 2 – Progress Reports	
2.1: Monthly Progress Reports	Monthly
Task 3 – Cybersecurity Training	
3.1: List of Individuals taking Training	January 6, 2022
3.2: Complete Cybersecurity training and submit completion certification to TCEQ Contract Manager	7 days after TCEQ provides access to the training
Task 4 – Training Workshop on CTM Evaluation Using Satellite Data	
4.1: A virtual training for TCEQ staff on CTM evaluation using satellite data that includes the elements listed in the task description.	February 26, 2022
4.2: Any training material used in the presentation for future TCEQ staff reference.	February 26, 2022
Task 5 – Developing Tools to Evaluate GEOS-Chem Using Satellite Data	
5.1: List of satellite products with advantages, disadvantages, and recommendations for GEOS-Chem evaluation to help decide on a set of two to three target satellite products for the tools.	February 21, 2022
5.2: Draft scripts to process satellite data from the set of target satellite products and compare with GEOS-Chem model output to begin testing on the TCEQ system.	April 4, 2022
5.3: Final scripts to process satellite data from the set of target satellite products and compare with GEOS-Chem model output, and processed satellite data.	May 19, 2022
5.4: A User's Guide for the scripts provided in deliverable 5.3.	May 19, 2022
Task 6 – Training and on-going technical assistance to train TCEQ staff	
6.1: A virtual training for TCEQ staff on how to use the processing scripts.	By June 17, 2022
6.2: All materials used in the hands-on training for future TCEQ staff reference.	By June 17, 2022
Task 7 – Draft and Final Reports	
7.1: Draft Report	June 15, 2022
7.2: Final Report	June 30, 2022

1.2 Background

Model performance evaluation of CTMs relies on the statistical comparison of model output to observational data, which can be conducted using satellite data products in otherwise unmonitored areas. GEOS-Chem is the CTM used by the TCEQ to simulate atmospheric composition outside of the TCEQ modeling domain and to provide initial conditions and boundary conditions (IC/BC) for the regional photochemical modeling. Most international, non-North American emissions in the TCEQ's regional photochemical modeling are represented in the IC/BC. Therefore, assessing the accuracy of GEOS-Chem model output at the domain boundaries is necessary to ensure an appropriate characterization of international and recirculated emissions. This project will provide the TCEQ staff with the tools to evaluate GEOS-Chem performance using satellite data, which will improve SIP modeling.

1.3 Report Outline

This draft Final Report highlights major activities and key findings, provides pertinent analysis, describes encountered problems and associated corrective actions, and details relevant statistics including data, parameter, or model completeness, accuracy and precision. It satisfies Deliverable 7.2 of the Work Plan for Work Order No. 582-22-31390-010:

Deliverable 7.1:	Final Report
Deliverable 7.1 Due Date:	June 30, 2022

The major results of the project are summarized in Section 2. Section 3 discusses the quality assurance findings for this project following the procedures from the project Quality Assurance Project Plan (QAPP). Section 4 summarizes our conclusions and Section 5 makes recommendations for future work based on the results of this project.

2 Satellite Product Selection

For Deliverable 5.1, We surveyed available satellite products to use for GEOS-Chem model evaluation tools, including advantages, disadvantages, and recommendations. We focused on the official, operational products from each sensor, so unofficial research products were not discussed. Furthermore, as the goal is to evaluate a global GEOS-Chem simulation for performance outside of the contiguous US, we focused on polar-orbiting sounders and imagers that provide global coverage.

After a brief description of each sensor, the following questions are answered for each sensor:

- 1) Are the data from 2019 available and is there a cost?
- 2) What variables of interest are measured?
- 3) How much of the globe is covered on a daily basis?
- 4) What is the spatial resolution?
- 5) What is the reported precision?
- 6) Is there an accessible log or database that shows important algorithm changes and/or previous data outages?

Then the advantages and disadvantages of each sensor were summarized.

After consultation with the TCEQ project manager, it was decided that the satellite evaluation tools focus on two satellite instruments: TropOMI (Veefkind et al., 2012) and CrIS (Gambacorta and Barnet, 2012). TropOMI is the newest of the Ultraviolet-Visible (UV-VIS) trace gas instruments, and thus would give the longest effective lifetime for the tools developed in this project. It is also much more advanced than OMI and the thermal infrared sounders and provides a much higher spatial resolution.

Recommended TropOMI gases included O₃, NO₂, and/or CO. Evaluating O₃ would help to assess the CAMx boundary conditions for this important pollutant, while observations of NO₂ and CO would assess the ability of GEOS-Chem to represent both anthropogenic and natural sources of O₃ precursors (with CO as a proxy for volatile organic compounds) globally.

CrIS provides measurements of radiance spectra of 1,305 channels in thermal infrared (TIR). Flown on board the Suomi National Polar-Orbiting Partnership (S-NPP) satellite since 2011 and JPSS-1 (Joint Polar Satellite System) since November 18, 2017, CrIS provides estimates of O₃ and CO throughout the vertical column but has lower sensitivity to concentrations near the surface than TropOMI. The recommended gases for CrIS were O₃ and CO. We found that the NOAA CLASS site did not include information on the averaging kernels needed for comparison with GEOS-Chem, so we instead used the NASA CLIMCAPS CrIS retrievals (Smith and Barnet, 2020).

Further details on TropOMI and CrIS are given below.

2.1 TROPOMI (ESA)

Brief description: TROPospheric Monitoring Instrument (TROPOMI) on board Copernicus Sentinel-5 was launched in October 2017. TROPOMI measures reflected

sunlight in the ultraviolet and visible (270–500 nm), near-infrared (675–775 nm) and shortwave infrared (2305–2385 nm) spectral bands.

1) Are the data from 2019 available and is there a cost?

Yes, available in NetCDF format. Data are free from [Copernicus Data Hub](#).¹

2) What variables of interest are measured?

NO₂, SO₂, CH₄, CO, O₃, HCHO

3) How much of the globe is covered on a daily basis?

All of the globe is covered on a daily basis.

4) What is the spatial resolution?

5.6 km x 3.5 km

5) What is the reported precision?

The Product Read Me (PRF) files indicate that all data have gone through quality assurance (QA) procedures and all data have a corresponding QA value. Only data with QA > 0.5 are recommended to be used. There are also precision estimates for the variables of interest from TROPOMI. They are as follows:

NO₂ (tropospheric): 25-50%

O₃: 5%

SO₂: 50%

CH₄: 1.5%

CO: 15%

HCHO: 80%

6) Is there an accessible log or database that shows important algorithm changes and/or previous data outages?

Yes. Accessible via the PRF files listed on the [Sentinel Products and Algorithms](#) page.²

Advantages: Data exist for 2019. Many variables are available from TROPOMI to be able to evaluate GEOS-Chem. Resolution is much higher than OMI.

¹ <https://s5phub.copernicus.eu/dhus/#/home>

² <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-5p/products-algorithms>

Disadvantages: Shorter period of record. Ground pixel resolution changed from 7.2 km x 3.5 km to 5.6 km x 3.5 km on 5 August 2019, which may complicate seasonal comparisons. Uncertainty may be too high for certain variables of interest.

2.2 CRIS (NOAA)

Brief description: The Cross-track Infrared Sounder (CrIS) is a Fourier transform spectrometer that generates profiles of water vapor, temperature, and greenhouse gases in the middle and upper atmosphere. CrIS provides measurements of radiance spectra of 1,305 channels in the long-wave IR (LWIR), midwave IR (MWIR), and short-wave IR (SWIR) band.¹ Flown on board the *Suomi National Polar-Orbiting Partnership (S-NPP)* satellite since 2011 and JPSS-1 (Joint Polar Satellite System) since November 18, 2017.

1) Are the data from 2019 available and is there a cost?

The data are freely available from NASA.³

1) What variables of interest are measured?

O₃, CO, and CH₄

2) How much of the globe is covered on a daily basis?

All of the globe is covered on a daily basis.

3) What is the spatial resolution?

48 x 48 km; average sampling distance is 16 km.

4) What are the known/reported biases?

For O₃, RMSE of 18-23% with a negative bias of 9% in the troposphere (Nalli et al., 2017). Absolute accuracy and precision of CO 5% and 15% (Nalli et al., 2020). For CH₄, precision of ~1.5% (Nalli et al., 2020).

5) Is there an accessible log or database that shows important algorithm changes and/or previous data outages?

Yes.

Advantage: Several variables of interest available.

Disadvantage: TIR products have less sensitivity to the boundary layer.

³https://disc.gsfc.nasa.gov/datasets/SNDRJ11ML2CCPRET_2/summary?keywords=climc
aps

3 Evaluation Script Development

3.1 Package Description

The scripts are for the evaluation of GEOS-Chem with five satellite products: TropOMI total column CO, TropOMI total column O₃, TropOMI tropospheric column NO₂, CrIS total column O₃, and CrIS total column CO. The processing procedure is shown in Figure 1. The appropriate observation operators (averaging kernel and *a priori* profiles for O₃ and NO₂; just averaging kernel for CO) for each satellite retrieval within a GEOS-Chem grid box within ± 1.5 hours of the satellite overpass time (13:30 local solar time) are applied to the GEOS-Chem model profile for that 3-hour window independently. The resulting smoothed GEOS-Chem columns are averaged for each grid box and day, as are the corresponding satellite results. These are saved as CSV files with this format:

Column 1: Dummy Index (integer)

Column 2: Average satellite columns (TropOMI or CrIS) in GEOS-Chem grid box (float, mol/m²)

Column 3: Number of non-zero satellite points in each GEOS-Chem grid box that passed the quality filters (integer)

Column 4: GEOS-Chem column with the satellite observation operator applied and averaged over the grid box (float, mol/m²)

Column 5: Latitude of the GEOS-Chem grid box (float, degrees)

Column 6: Longitude of the GEOS-Chem grid box (float, degrees, -180 to 180)

Column 7: Date (YYYYMMDD)

1. Column 8 (CrIS Only): GEOS_noAK (mol/m²): the GEOS-Chem profile before the observation operator was applied. Used for sanity checking the AK applied values.

Evaluation statistics are then calculated for all GEOS-Chem grid boxes over all days analyzed, as well as for each day.

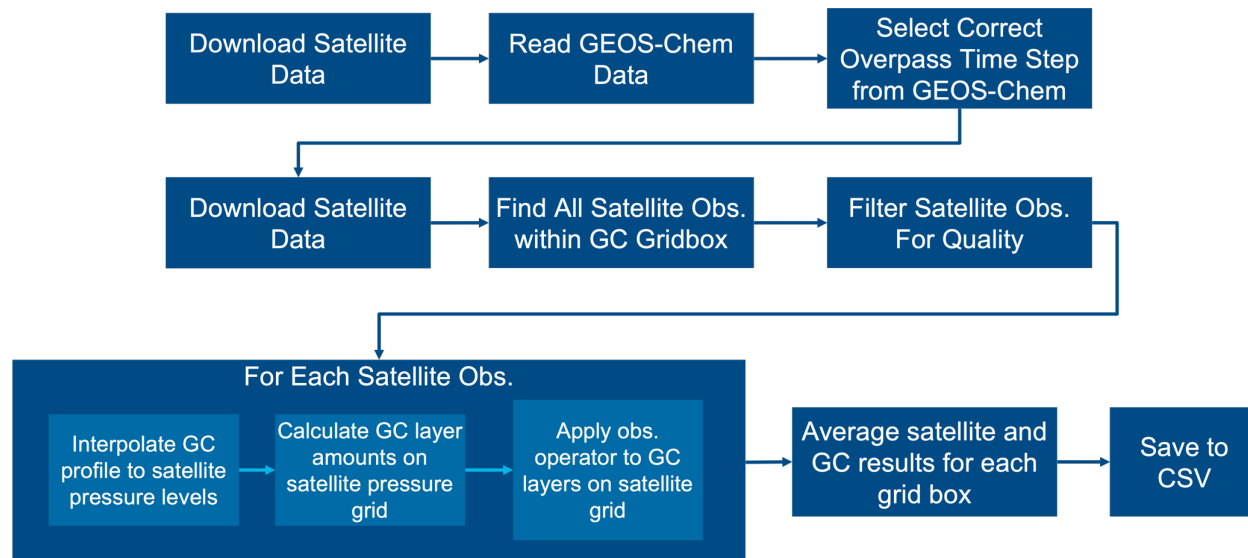


Figure 1. Processing steps for the evaluation scripts.

The software is provided as a combination of Python (v3.8.5) and bash shell scripts. Currently the shell scripts are only used for data downloading, which we expect to take place on Linux machines. The scripts themselves can be run on Windows if the appropriate paths to the Linux data directories are used. The scripts for each satellite are described below. The final statistical calculation and plotting packages are independent of satellite species, as all the scripts produce similar CSV output files.

3.1.1 TropOMI

The ten TropOMI processing scripts are:

- 1) **TROPOMI_Downloads.sh**: Shell script that downloads the TropOMI satellite data and runs the corresponding python processing script.
- 2) **dhusget_o.3.8.sh**: helper script for AER_GEOS_Download.sh that subsets and downloads TropOMI data
- 3) **TM5_profile.py**: This is a utility supplied by the TropOMI team to extract the TM5 data in the “AUX” file that is used for the NO₂ *a priori* profile and add it to the Level 2 NO₂ file. This must be run on all Level 2 NO₂ files before they can be used by the other scripts. Not needed for TropOMI CO or O₃.
- 4) **get_TROPOMI_O3.py**, which is the main program that produces the output CSV file containing daily average TropOMI and GEOS-Chem smoothed O₃ *total* columns for each GEOS-Chem grid box.
- 5) **get_TROPOMI_NO2.py**, which is the main program that produces the output CSV file containing daily average TropOMI and GEOS-Chem smoothed NO₂ *tropospheric* columns for each GEOS-Chem grid box.

- 6) **get_TROPOMI_CO.py**, which is the main program that produces the output CSV file containing daily average TropOMI and GEOS-Chem smoothed CO *total* columns for each GEOS-Chem grid box.
- 7) **GEOS_TropOMI_layer_interp.py**, which is the main program that takes a GEOS-Chem O₃ profile on 47 layers in mol/mol and returns the columns on the TropOMI layers in units of mol/m². This program is called in the appropriate TROPOMI functions script.
- 8) **TROPOMI_NO2_functions.py**, which contains:
 - a) **GEOS_TROPOMI_retrieval**: a subroutine that finds all TropOMI NO₂ scans that pass the quality filters for a given GEOS-Chem grid box and date
 - b) **Calc_GEOS_TROPOMI_NO2_Values**: a subroutine that applies a single TropOMI NO₂ tropospheric column AK and *a priori* profile to the GEOS-Chem NO₂ profile
- 9) **TROPOMI_O3_functions.py**, which contains:
 - a) **GEOS_TROPOMI_retrieval**: a subroutine that finds all TropOMI O₃ scans that pass the quality filters for a given GEOS-Chem grid box and date
 - b) **Calc_GEOS_TROPOMI_O3_Values**: a subroutine that applies a single TropOMI O₃ total column AK and *a priori* profile to the GEOS-Chem O₃ profile
- 10) **TROPOMI_CO_functions.py**, which contains:
 - a) **GEOS_TROPOMI_retrieval**: a subroutine that finds all TropOMI CO scans that pass the quality filters for a given GEOS-Chem grid box and date
 - b) **GEOS_CO_Interp**: a subroutine that takes a GEOS-Chem CO profile on 47 layers in mol/mol, maps it to the 50 TropOMI CO layers, and then applies a single TropOMI CO total column AK to the GEOS-Chem CO profile.

3.1.2 CrIS

The CrIS processing scripts are:

- 11) **CrIS_downselect_files.py**: The main program that reads in CrIS data and stores the lat/lon.
- 12) **CrIS_functions.py**, which contains:
 - a) **GEOS_CRIS_retrieval**: a subroutine that finds all CrIS O₃ or CO scans that pass the quality filters for a given GEOS-Chem grid box and date
 - b) **GEOS_CRIS_Interp**: a subroutine that takes a GEOS-Chem O₃ or CO profiles on 47 layers in mol/mol, maps it to the 100 CrIS O₃ or CO layers, and then applies a single CrIS O₃ or CO column AK to the GEOS-Chem O₃ (CO) profile.
 - c) **CO_ap**: a function that calculated the CO a priori based on latitude and date
 - d) **slb2fin**: A function to construct the CLIMCAPS trapezoid state function on 100 pressure levels one at a time
 - e) **calc_finv_mp**: Calculates a scene-dependent transformation matrix (F_matrix) and its inverse to allow the mapping of the 9x9 CrIS averaging kernels (for the 9

trapezoidal functions) to the 100x100 averaging kernels for the 100 retrieval pressure levels.

- 13) **get_CrIS_CO.py**: The main program that produces the output CSV file containing daily average CrIS and GEOS-Chem smoothed CO *total* columns for each GEOS-Chem grid box.
- 14) **get_CrIS_O3.py**: The main program that produces the output CSV file containing daily average CrIS and GEOS-Chem smoothed O₃ *total* columns for each GEOS-Chem grid box.

3.1.3 Statistics and Plotting

The statistics and plotting scripts are:

- 15) **AER_Stats.py**, which:
 - a) Reads in the CSV file produced by the other scripts
 - b) Calculates several statistics for the daily grid-point averages (see Appendix A)
 - c) Calculates the average statistics for each day and plots them as a timeseries
 - d) Produces a scatterplot of the daily grid-point averages (x = Satellite column, y = GEOS-Chem column with observation operator applied)
 - e) Calculates the line of best fit for the scatterplot
- 16) **AER_maps.py_**: plots a map of the smoothed GEOS-Chem columns and the difference between the satellite average columns and the GEOS-Chem smoothed columns for each grid box (e.g., Figure 2).

3.2 Gathering Satellite Data

We have included a shell script (TROPOMI_Downloads.sh) that allows you to download Level 2 TropOMI data for a specific day and longitude/latitude box for NO₂, CO, and O₃. The variable of interest (NO₂, CO, or O₃) and the year, month, and day of interest are set on L3-6. The longitude and latitude ranges are set on L14 after the -c flag, as in “-c -150,6:-45,70” for -150 to 6 degrees longitude and -45 to 70 degrees latitude. Prior to the -c flag, the variable of interest is set. Note that this portion of the script must be exactly 10 spaces. So eliminate one of the ‘_’ after NO2. CO and O3 require no such changes. The script also calls the appropriate TropOMI processing script after downloading a day of data. Note that this script requires another script in the package (dhusget_o.3.8.sh).

For TropOMI NO₂, a separate auxiliary data file is needed for input to TM5_profile.py. This file can be downloaded from <https://s5phub.copernicus.eu/dhus/#/home> as Product Type AUX_CTMANA and has the format of S5P_OPER_AUX_CTMANA*.nc

CrIS CLIMCAPS satellite data can be downloaded in NetCDF format from the NASA Goddard Space Flight Center (GSFC) DAAC. The product name is SNDRJ1IML2CCPRET. The DOI for the current Version 2 of the product is doi:10.5067/LESQUBLWS18H . The current download and sub-setting link is:

https://disc.gsfc.nasa.gov/datasets/SNDRJ1IML2CCPRET_2/summary?keywords=climcaps

The files contain data for both O₃ and CO, although no *a priori* is provided in the netcdf files for CO. Instead, the *a priori* for CrIS CO is read from a .csv file(s) we have provided. A function called 'CO_ap' in CrIS_functions.py returns the CO *a priori* profile. The data are generated from the AIRS v6 climatology ([@@@REF](#)).⁴ Note also that while the subsetting tools claims to reduce the number of files based on time and geography, in practice it returns all files (> 200) for a single day.

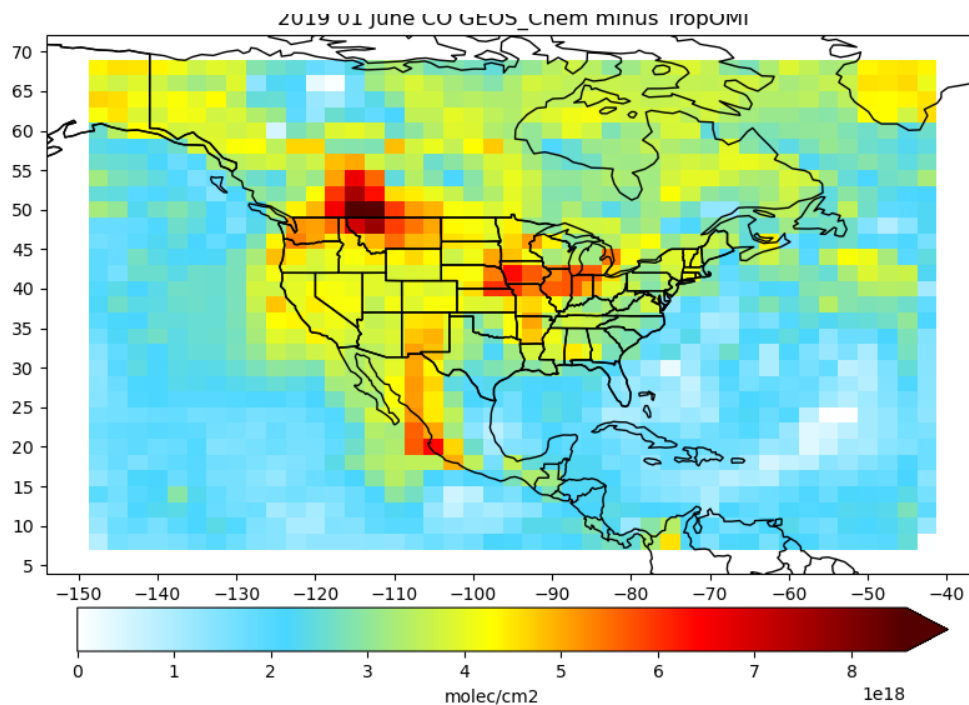


Figure 2. Example plot from AER_maps.py.

3.3 Quality Screening and Observation Operators

For TropOMI, quality flags apply to the entire retrieved column. For CO⁵ and O₃⁶, only retrievals with quality flags ≥ 0.5 are used to evaluate GEOS-Chem, while for NO₂ only retrievals with quality flags ≥ 0.75 are used.⁷

For TropOMI CO, the observation operator is matrix multiplication by the averaging kernel of the TropOMI retrieval.

$$\hat{x} = Ax_m$$

where x_m is the GEOS-Chem profile interpolated to the TropOMI CO retrieval layers and converted into subcolumns (mol/m²) and A is the averaging kernel.

For TropOMI O₃, the a priori profile also has to be included:

$$\hat{x} = A(x_m - x_a) + x_a$$

For TropOMI NO₂, the a priori is not required, but the averaging kernel needs to be rescaled based on the ratio of tropospheric and total air mass factors, and then all values outside the troposphere need to be set to 0 (See Section 8.8 of Footnote 6). Once the AK is transformed, the same equation as that for CO is used for the observation operator.

For CrIS, the quality flags apply to each retrieval layer, where 0 is best, 1 is good, and 2 is bad. Only layers with a quality flag ≤ 1 are used for comparison with GEOS-Chem, and all layers with QC of 2 are set to zero in both the satellite and GEOS-Chem profiles prior to comparison. The log observation operator is used for these products:⁸

$$\ln(\widehat{x}) = A(\ln(x_m) - \ln(x_a)) + \ln(x_a)$$

The O₃ a priori profile is provided in the CrIS CLIMCAPS data files. The CO a priori is derived by interpolating the static monthly profiles for the Northern and Southern Hemispheres according to the equations in Footnote 8.

3.4 Challenges Encountered

3.4.1 Interpolation and Layering

GEOS-Chem provides layer-average volume mixing ratios (VMR, mol/mol) on 47 pressure layers, but to apply the TropOMI and CrIS observation operators, we need layer amounts (mol/m²) on the TropOMI (variable layers) and CrIS (100 layers) pressure grids.

To do this, we:

⁵ <https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Carbon-Monoxide-Level-2-Product-Readme-File>

⁶ <https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Total-Ozone-Level-2-Product-Readme-File>

⁷ <https://sentinel.esa.int/documents/247904/2474726/Sentinel-5P-Level-2-Product-User-Manual-Nitrogen-Dioxide>

⁸

https://docs.server.gesdisc.eosdis.nasa.gov/public/project/Sounder/CLIMCAPS_V2_L2_science_guides.pdf

- Use the `interp1d` function to interpolate the GEOS-Chem VMR and temperature to the top and bottom pressures of each satellite layer
- Calculate the layer average VMR, temperature, and pressure. For VMR and temperature, this is the average of the top and bottom values from `interp1d`. For pressure this is the logarithmic average of the top and bottom values.
- Use the `metpy` hydrostatic equation function (`calc.thickness_hydrostatic`) to calculate the layer height (m)
- Use ideal gas law to calculate layer density (mol/m^3)
- Then Layer amount (mol/m^2) = (layer density)*(layer height)*(layer average VMR)

3.4.2 TropOMI Tropospheric NO₂

The a priori profile for TropOMI NO₂ is not stored in the Level 2 data files. Instead, a separate script (`TM5_profile.py`) has to be used along with an additional auxiliary data file to add the a priori profile to the Level 2 file. Instructions for this process are in the scripts User's Guide.

In addition, while TropOMI supplies the averaging kernel for NO₂ throughout the entire atmosphere, only the tropospheric column is reported. The tropospheric averaging kernel is obtained by scaling the full kernel by the ratio of the total NO₂ column to the tropospheric column (M/M_{trop}) and setting all elements of the kernel to zero above the tropopause layer.

3.4.3 CrIS Averaging Kernels

The CLIMCAPS CrIS averaging kernels are not stored on the 100 retrieval levels but are instead stored as coefficients for 9 trapezoidal functions. The Science Applications Guide provides an IDL code to create the necessary mapping matrices in text and we have converted the IDL code into Python. Refer to functions `slb2fin` and `calc_fin_mv_mp` in `CrIS_functions.py` starting at lines 365 and 471 respectively.

3.4.4 Linux to Windows Transition

We had written our scripts to use bash shell scripts to set environmental variables, which causes problems when the python environment is in Windows but the data is in Linux. We adjusted our scripts to read environmental variables from a text file (`TCEQ_env.env`) so that they will work on Windows as expected, but the data download script for TropOMI still needs to be run on a Linux machine.

3.4.5 Processing Speed

The scripts can be very slow. TropOMI CO and CrIS CO and O₃ take about 6 hours to process one day of data, but TropOMI NO₂ and O₃ can take multiple days. The CrIS processing was sped up substantially by sorting the latitude and longitude range of each file at the beginning of the processing, so file reading could be minimized later in the calculation. Running multiple days on separate cores will also speed up the time needed to process months of data.

4 TCEQ Training Sessions

AER hosted two training sessions for TCEQ staff. The first (Task 4) covered:

1. Background on Satellites: Fundamentals of remote sensing, different satellite spectral regimes (US-VIS, TIR), different satellite orbits (geostationary, polar), observation geometries (limb, nadir), and remote sensing method (active or passive).
2. Remote Sensing Theory and Satellite Observation Operators: Optimal estimation fundamentals, definition of an Averaging Kernel and Observation Operator
3. Air Quality Satellites: Strengths and weaknesses of various air quality satellites
4. Model Evaluation Examples
5. An introduction to more advanced methods, like Data Assimilation

The second training session focused on the hand-on use of the developed evaluation scripts (Task 6). We described the basic operation of the scripts and a few of the challenges we encountered in developing them (see Section 3). We assisted TCEQ staff with the steps needed to use the scripts, including setting up a Python environment with Anaconda, using Jupyter notebooks, downloading (and preprocessing for TropOMI NO₂) the Level 2 satellite data, running the evaluation scripts, and calculating statistics and making plots with the output.

5 Quality Assurance

All of the scripts provided to TCEQ have undergone extensive quality testing, including examining intermediate calculations (layer heights, layer amounts, averaging kernels) and verifying them via independent calculations, code reviews by team members who did not write the original scripts, and evaluation of the output of the scripts for consistency with expectations. These reviews satisfy the the 10% audits of data quality stipulated by the TCEQ contract and Work Order for Category III QAPPs.

As of this draft final report, all scripts appear to be giving realistic and correct results. Careful review of intermediate calculations has confirmed that these calculations are being done correctly.

As the quality of the information, including secondary data, was not evaluated by EPA, the below disclaimer applies to all project deliverables:

Disclaimer: The information contained in this report or deliverable has not been evaluated by EPA for this specific application, i.e. the identification of brown carbon aerosols and biomass burning smoke.

6 Conclusions

Here we summarize the conclusions of our project, with reference to the corresponding report section.

- After reviewing the strengths and weaknesses of different satellite datasets on air pollutants, we decided to develop GEOS-Chem evaluation scripts for TropOMI O₃, CO, and NO₂, as well as CrIS O₃ and CO (Section 2).
- The scripts developed in this project are for the evaluation of GEOS-Chem with five satellite products: TropOMI total column CO, TropOMI total column O₃, TropOMI tropospheric column NO₂, CrIS total column O₃, and CrIS total column CO. The appropriate observation operators (averaging kernel and *a priori* profiles for O₃ and NO₂; just averaging kernel for CO) for each satellite retrieval within a GEOS-Chem grid box within ± 1.5 hours of the satellite overpass time (13:30 local solar time) are applied to the GEOS-Chem model profile for that 3-hour window independently. The resulting smoothed GEOS-Chem columns are averaged for each grid box and day, as are the corresponding satellite results. Evaluation statistics are then calculated for all GEOS-Chem grid boxes over all days analyzed, as well as for each day. (Section 3.1).
- Challenges included special processing steps for the TropOMI NO₂ and CrIS codes to obtain the correct averaging kernels, and the need to map the GEOS-Chem layering to the CrIS and TropOMI levels.
- We held two training sessions for TCEQ staff to review the basic principles of evaluating atmospheric chemistry models with satellite data and on the hands-on use of our evaluation scripts (Section 4).

7 Recommendations for Further Study

Based on the results of this work, we make the following recommendations for further study:

- A full evaluation of GEOS-Chem should be performed by TCEQ with the scripts, and the results carefully analyzed to ensure the scripts are performing as expected. AER will assist with any script errors found during these evaluations.
- The scripts can be expanded to cover additional satellite observations, including those of aerosols.
- Further work to speed up the scripts without sacrificing accuracy.
- The downloading of the satellite data could be streamlined. As many of these data sources are now available on cloud repositories, processing on the cloud may remove the need to download the data entirely.
- Additional statistics and plots can be created using the templates we have provided in this project.

8 References

- Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., ... & Schultz, M. G. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research: Atmospheres*, 106(D19), 23073-23095.
- Gambacorta, A., & Barnet, C. D. (2012). Methodology and information content of the NOAA NESDIS operational channel selection for the Cross-Track Infrared Sounder (CrIS). *IEEE transactions on geoscience and remote sensing*, 51(6), 3207-3216.
- Nalli, N. R., Gambacorta, A., Liu, Q., Tan, C., Iturbide-Sanchez, F., Barnet, C. D., ... & Smith, J. W. (2017). Validation of atmospheric profile retrievals from the SNPP NOAA-unique combined atmospheric processing system. Part 2: Ozone. *IEEE Transactions on Geoscience and Remote Sensing*, 56(1), 598-607.
- Nalli, N. R., Tan, C., Warner, J., Divakarla, M., Gambacorta, A., Wilson, M., ... & Rettinger, M. (2020). Validation of carbon trace gas profile retrievals from the NOAA-unique combined atmospheric processing system for the cross-track infrared sounder. *Remote Sensing*, 12(19), 3245.
- Smith, N., & Barnet, C. D. (2020). CLIMCAPS observing capability for temperature, moisture, and trace gases from AIRS/AMSU and CrIS/ATMS. *Atmospheric Measurement Techniques*, 13(8), 4437-4459.
- Veefkind, J. P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., ... & Levelt, P. F. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote sensing of environment*, 120, 70-83.