Environment and Climate Change Canada (ECCC) CrIS Fast Physical Retrieval (CFPR) Ammonia (NH_3) Database

1. Data Usage

The CrIS ammonia product from the CrIS Fast Physical Retrieval (CFPR) is a research product, therefore please work closely with the developers listed below as you analyze and use the data. **Appropriate use of the data (e.g., quality flags, averaging kernels, errors estimates) requires collaborating with the current developers, while feedback from users leads to a better product**. If the data are used in a publication or presentation co-authorship is to be offered to the main developers. The data are to be used for research purposes and are provided as is without any express or implied warranties. The use of this data is subject to the Open Government License – Canada (https://open.canada.ca/en/open-government-licence-canada).

Main developer and contact:

Mark Shephard
Environment and Climate Change Canada
4905 Dufferin Street, Toronto, Ontario, Canada M3H 5T4
Mark.Shephard@ec.gc.ca

Data Access: The CrIS Fast Physical Retrieval (CFPR) ammonia dataset created by Environment and Climate Change Canada (Shephard et al., 2020) is publicly available at: https://hpfx.collab.science.gc.ca/~mas001/satellite_ext/cris with login information obtained by request from Mark Shephard (Mark.Shephard@ec.gc.ca).

Development support provide by:

Shailesh Kharol
AtmoAnalytics, Inc.
6 Murphy Rd.
Brampton, Ontario, Canada, L6S 6L2
shailesh.kharol@atmoanalytics.com

Enrico Dammers
TNO, Climate Air and Sustainability,
Utrecht, The Netherlands
enrico.dammers@tno.nl

Karen Cady-Pereira
Atmospheric and Environmental Research (AER), Inc.
131 Hartwell Ave, Lexington, MA, USA
kcadyper@aer.com

Selected References for the CrIS Fast Physical Retrievals (CFPR):

White, E., M. W. Shephard, K. E. Cady-Pereira, S. Kharol, S. Ford, E. Dammers, E. Chow, N. Thiessen, D. Tobin, G. Quinn, J. O'Brien, J. Bash, Accounting for Non-detects in Satellite Retrievals: Application Using CrIS Ammonia Observations, Remote Sensing, 15, 2610, https://doi.org/10.3390/rs15102610, 2023.

Shephard M. W., E. Dammers, K. E. Cady-Pereira, S. K. Kharol, J. Thompson, Y. Gainariu-Matz, J. Zhang, C. A. McLinden, A. Kovachik, M. Moran, S. Bittman, C. Sioris, D. Griffin, M. J. Alvarado, C. Lonsdale, V. Savic-Jovcic, and Q. Zheng, Ammonia measurements from space with the Cross-track Infrared Sounder (CrIS): characteristics and applications, Atmos. Chem. Phys., 20, 2277–2302, https://doi.org/10.5194/acp-20-2277-2020, 2020.

Shephard, M.W. and K.E. Cady-Pereira, Cross-track Infrared Sounder (CrIS) Satellite Observations of Tropospheric Ammonia, *Atmos. Meas. Tech.,* 8, 1323–1336, doi:10.5194/amt-8-1323-2015, 2015.

2. User's Guide

2.1. Quality Flags

CFPR currently generates a Quality_Flag (QF) variable for each observation, ranging from -1 (did not converge) to 5, which is determined by applying a series of quality checks on the retrievals. Each QF value can be considered as a step, with higher values implying that all the conditions for lower QF values have also been met. The QF values are listed below:

Quality Flag	Description	Details
-1	Retrieval did not converge	Flag indicating that the retrieval did not converge. Often these are not written into the product files.
≥ 1	Retrieval converged	Flag indicating that the retrieval converged is specified number of iterations.
≥ 2	Large outlier flag	Quality flag 1 & profile retrieved value less than 200 ppbv.
≥ 3	Retrieval quality flag	Quality flag 2 & chi-squared less than 20
≥ 4	More conservative retrieval quality flag	Quality flag 3 & signal-to-noise ≥ 1 and thermal contrast > 0
≥ 5	More conservative retrieval quality flag with information content	Quality flag 4 & also filtered to have a minimum degree-of-freedom for signal of 0.1

The choice of which QF to use as a selection criteria depends on the users' objectives. We recommend including QF≥3 or larger. If the observational operator is being applied (e.g. model inversions) then one likely does not want to only use values with QF≥5 as the averaging kernel will already take into account the impact of cases with limited information (DOF <0.1). When generating maps to look for pattern, QF≥5 will return the most intense or persistent features. QF=5 should also be used for comparisons against in situ data, as these observations will have the lowest uncertainty.

Note, if using version 1.6 or above, for the highest quality data the pixels with Cloud_Flag=1 can be remove from your analysis. As noted below, pixels with Cloud_Flag = 1 are retrievals that are performed under thin cloud conditions and are not as reliable as retrievals performed under cloud-free conditions. Please refer to White et al., (2023) for more details.

2.2. Cloud_Flags : Identifies cloudy and clear sky retrievals, and non-detect pixels

This variable is available starting in version 1.6. A CrIS ammonia cloud detection algorithm (CACDA) uses the University of Wisconsin's CIMG product that that provides VIIRS cloud products on the CrIS footprint, plus an additional capacity by ECCC to handle smoke plumes the my otherwise be identified as clouds.

Cloud Flag	Descriptor	Comment
-1	No cloud information	Corresponding VIIRS cloud information was missing.
0	Clear retrieval	Retrieval under cloud-free conditions.
1	Cloudy retrieval	Retrieval performed was identified by VIIRS as a containing cloud.
2	Smoke plume	CrIS pixels that were initially identified by VIIRS as cloudy, but are smoke plumes.
3	Survival non-detect	Cloud-free CrIS pixels below the detection limit of sensor.
Note: Cloudy pixels with no signal (thick clouds) are not retrieved or included in dataset		

In the provided dataset the CACDA is applied to retrieval pixels (that have an NH_3 spectral signal), and used to identify non-signal cloud-free pixels (non-detects) that were not retrieved (Cloud_Flag = 3). Thick cloudy pixels with no NH_3 spectral signal are not reported in this dataset. Retrieval performed that was identified by the CACDA as containing clouds (mostly thin cloud conditions) are indicated with Cloud_Flag=1. These retrievals can be less reliable and can be rejected for better quality control. There are conditions in which the part of the CACDA algorithm using just the VIIRS CIMG product would identify a cloud, but there is also a very strong NH3 spectral signature. These are conditions of thick smoke plumes, and the CACDA will indicate these as "smoke plume" with a Cloud_Flag=2. Caution that this is not an indicator of all pixels containing smoke, but rather just the pixels in which retrievals were performed and were also initially identified as clouds using the VIIRS information. There are other retrievals in smoky conditions (e.g. smaller emission source) that are not initially identified as clouds that would have a clear flag of Cloud_Flag=0.

2.3. Comparing measurement or model data against CrIS profiles

There are a number of approaches for comparing in situ observations or model values with remotely sensed satellite data. One rigorous approach applies the observation (or instrument)

operator (see equations below) to the measurement or model values. Applying the observational operator allows for the removal of the influence of the apriori in the comparison (only compares the measurement component of the retrieved quality) when comparing the profile with the satellite profile.

$$x_o = x_a + AK_{linear} * (x_m - x_a)$$

Where x_o = in situ profile with instrument operator applied
x_a= a prior profile
AK_{linear} = linear averaging kernel
x_m= in situ or model profile (on the same profiles levels as the satellite)
The same equation above can be written if the retrievals are performed in natural log of volume mixing ratio (vmr), as is done in the CrIS NH3 retrievals:

$$\ln x_o = ln\ x_a + AK_{log} * (\ln x_m - ln\ x_a)$$

$$x_o = exp^{\left(ln\ x_a + AK_{log} * (\ln x_m - ln\ x_a)\right)}$$

Where AK_{log} = the averaging kernel in natural log [*ln*(vmr) / *ln*(vmr)] as reported in the CrIS NH_3 product file.

The argument for this approach is that x_o is what the satellite would observe if x_m is the true profile. This is a useful technique for evaluating if the retrieval algorithm is performing as designed, i.e., is it unbiased and the RMS is within the expected variability. It is also valuable for model inversions and assimilations of the CrIS NH3 products.

There might be applications in which a direct comparison how well the retrieved CrIS NH_3 product (apriori + measurement components combined) compare against observations. This is fine as long as it is clearly stated that in the comparison that the satellite retrieved product contains apriori information, and the comparison would change if the apriori in the retrieval changes. This is often not as good of a metric for satellite measurement validation as using the observation operator stated above.

See any specific caveats in the log statements below for the various versions of data.

3. Version Log

3.1. Version 1.50 (2018/2019)

Caveats:

- The averaging kernel is stored in [col,row], thus, the transpose of the typical mathematical notation of [row,col].

- In this version there are times when the instrument (measurement) error covariance matrix was not written correctly into the netcdf output file due to precision errors. This has been corrected for the next release. The total error covariance values are correct.
- Presently we are investigating retrieval values over ocean and elevated concentration values over some deserts and high elevations wintertime conditions (e.g. North American Rockies). We are also validating CrIS NH3 nighttime observations against available ground-based observations. Thus, care need to be taken if data is used over these regions, and it is recommended that these values not be used at this time.

3.2. Version 1.6 (1.6.0) (July, 2021)

This version 1.6 and higher versions contain the same retrievals as version 1.5 (see caveats above), but includes non-detects pixels that are below the detection limit of the sensor. This version explicitly identifies and accounts for cloud-free pixels below the detection limit of the sensor to provide more representative gridded averaged (level 3) values. The additional variable added to the files is the "Cloud_Flag" as described above. Note: this version had a small issue with the reported Quality_Flag for the non-detect pixels that was fixed in v1.6.1. as stated below. Also, in this version the reported apriori (xa) values for the non-detect pixels only were inserted into the output files as ppbv instead of ppmv, which was fixed in v1.6.2 as stated below.

3.3. Version 1.6.1 (August, 2021)

This version is the same as v1.6.0, but with a small correction in that the Quality_Flag = 5 for all non-detect pixels. In previous version the post-processing had set some non-detect pixel with a Quality_Flag = 3 by error. Note: in this version the reported apriori (xa) values for the non-detect pixels only were inserted into the output files as ppbv instead of ppmv, which was fixed in v1.6.2 as stated below.

3.4. Version 1.6.2 (November, 2021)

This version is virtually the same as v1.6.1, except the apriori profile (xa) units for the non-detect pixels are now in ppmv to be consistent with the other retrieval apriori profile values. In the previous version (v1.6.0 & v1.6.1) the apriori (xa) values for the non-detect pixels only were inserted as ppbv instead of ppmv. All the xa values are now consistent in units of ppmv. In addition, there were a very limited number of additional retrievals included, which were mostly over a few regions in 2020. Note: there were a few regions where the non-detects were added in twice, which was fixed in v1.6.3 below.

3.5. Version 1.6.3 (January, 2022)

This version is the same as v1.6.2, except there were a few regions in v1.6.2 were the non-detects we added in twice. The duplicate values have been removed in v1.6.3.

There were also a few small regions, mostly in 2021, where additional retrievals were add for a couple of months that were missing in v1.6.2.

3.6. Version 1.6.4 (May, 2023)

This version is the same as v1_6_3 with the following updates. There are retrievals for both CrIS SNPP (May 2012 to May 2021) and CrIS NOAA20 (March 8, 2019 to Present) are available for this version. The measurement error statistics reported in the files were corrected. The representative averaging kernels, error covariance matrices, and values for non-detect pixel over water were updated. There is a new global attribute named “satellite” to distinguish between CrIS ammonia retrievals from SNPP and NOAA20. Note: due to a small processing error some of the estimated total column errors for the non-detect pixels have values of zero instead of a small number. These have been fixed for the next release. Also, there is a new global attribute “version” that replaces the old “Version” variable. The old “Version” variable was not updated and should be ignored.

Important: Please refer to this document often for updates and use the most recent version available as it will have updates and fixes to previous versions. Don’t combine versions in your analysis. Please contact developers for any data caveats (i.e. over deserts) when using this product.