TROPESS Deuterated Water Vapor Level 2 Standard Data Product User Guide

Authors: John Worden[1], Susan Kulawik[2], Robert Herman[1], Kristal Whitten[1], and Kristen Fahy[1]

[1]Jet Propulsion Laboratory, California Institute of Technology
[2] NASA Ames Research Center

1 Table of Contents

1 Introduction

1.1 Overview and Document Scope

This document is to be used as a user guide for using the TROPESS AIRS and CRIS Level 2 Standard Product Files

1.2 Dataset Description

This user guide describes the TROPESS AIRS and CRIS Level 2 Standard (useful for model/aircraft comparisons and process studies) and Summary Product Files (contains integrated quantities useful for process studies / quick look) for the δ-Deuterium content of water (HDO) and H_2O.

Product Information	Description
Parameters	HDO, H_2O
Data Product Provenance	MUSES Version 1.11
Approximate file size	~15 MB
Spatial coverage	Regular collections have global coverage: Nominal latitude range: 70 N to 50 S Nominal longitude range: -180 to 180 Special collections: Spatial coverage varies by collection
Temporal coverage	Each L2 Standard file contains 1 day of data
File format	netcdf
Vertical sensitivity	Estimates of the deuterium content of water vapor are most sensitive to deuterium variability between 825 and 400 mb.
Data quality	The data have undergone a pre-quality check, which involves checks for retrieval convergence and measurement sensitivity. There are no checks for clouds or land versus ocean as we do not find these to substantively affect the quality of the retrieval as long as the retrieval has converged.
Uncertainty	Profiles have approximately 3-5% uncertainty (or ~ 40 per mil). Integrated values between 825 and 400 mb have approximately 3% (or 30 per mil) uncertainty
Validation Stage	AIRS HDO data are at validation Stage 3 and CRIS HDO data are at validation stage 2 according to NASA guidelines https://science.nasa.gov/earth-science/earth-science-data/data-maturity-levels”

	See validation section summary (Section 5) for rationale
Retrieval Levels	17 levels from surface to top-of-atmosphere
FillValues	-999

Revision History

The current MUSES version (release March 2021) has a known issue that predicted error is overestimated. The corrected predicted error will be included in the next release.

1.3 Filename

The Level 2 Standard Products adhere to the following filename convention:

[ProjectID]_[Instrument-Platform]_[ProductLevel]_[ProductType]_[ProductName]_[DateStamp]_ [AlgorithmName]_[AlgorithmVersion]_[ProcessingStrategy]_[FormatVersion]**.**nc

Example:

TROPESS_AIRS-Aqua_L2_Standard_HDO_20200912_MUSES_R1p11_FS_F0p1.nc

2 Product File Contents and Parameter Description

Please see the README for a list of variables included in the L2 Standard product Files

2.1 Variables included in the L2 Standard Product

The variables are listed in the README document

3 References

Please cite the following references if you intend to use these data:

1. Worden, J. R. et al. (2019), Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor, *Atmospheric Measurement Techniques*, *12*(4), 2331–2339, doi:10.5194/amt-12-2331-2019.
2. Herman, R. L., J. Worden, D. Noone, D. Henze, K. Bowman, K. Cady-Pereira, V. H. Payne, S. S. Kulawik, and D. Fu (2020), Comparison of optimal estimation HDO/H2O retrievals from AIRS with ORACLES measurements, *Atmospheric Measurement Techniques*, *13*(4), 1825–1834, doi:10.5194/amt-13-1825-2020.
3. Worden, J. et al. (2006), Tropospheric Emission Spectrometer observations of the tropospheric HDO/H2O ratio: Estimation approach and characterization, *Journal of Geophysical Research-Atmospheres*, *111*(D16), doi:10.1029/2005JD006606.

The following references were also used in the development of this documentation and should be cited for model/data comparisons.

1. Risi, C. et al. (2012), Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopologues: 1. Comparison between models and observations, *J. Geophys. Res*, *117*(D5), D05303, doi:10.1029/2011JD016621.
2. Risi, C., D. Noone, C. Frankenberg, and J. Worden (2013), Role of continental recycling in intraseasonal variations of continental moisture as deduced from model simulations and water vapor isotopic measurements, *Water Resour. Res.*, *49*(7), 4136–4156, doi:10.1002/wrcr.20312.

4. Extended User Guide: L2 Standard Product

This extended user guide shows how to compare these fields to aircraft or model fields, calculate uncertainties, and calculate the dd_004 variable.

4.1 How to Compare TROPESS AIRS deuterium to aircraft or model fields

All comparisons between remotely sensed data an *in situ* measurements or high resolution model fields must account for the sensitivity of the remotely sensed measurement and any regularization used with the remotely sensed measurements. Otherwise *large* errors will be incurred with the comparisons.

Here is the general approach for comparison of the TROPESS AIRS data to models or vertical profile measurements such as from aircraft:

1) Calculate the HDO/H_2O profile using the model or aircraft fields (for the purpose of this demonstration we will call this $\boldsymbol{x\_}_{true}$).
2) Apply the following operation to the model

$$ln(\hat{x}) = ln(xa) + A\left(ln(x_{true}) - ln(xa)\right)$$

In the netcdf product file and in the equation xa is the constraint vector used to regularize the retrieval. A is the Averaging kernel matrix and must be matrix multiplied by $ln(x_{true}) - ln(xa)$.

After this operation one can compare the modified HDO/H_2O profile ($\hat{x}$) to the “x” variable in the netcdf product file.

Note that the averaging kernel matrix (A) is not symmetric so getting the row/column order of A correct is critical. You can check your work by using the variables x_test in each netcdf file in the above equation and using the zeroth entries for the averaging kernel, xa and using the zeroth entry for x as x_{true}. If your test operation matches the variable x_test then the operation is correct.

4.2 Calculation of the Degrees of Freedom For Signal (DOFS)

The Averaging kernel matrix (see previous section) is the sensitivity of the estimate to the true state. The DOFS is defined as the trace of the averaging kernel matrix. To determine if the estimate is sensitive to a particular altitude a good place to check is the DOFS. We only include measurements where DOFS for the entire profile is ~1. A value of 1 means the estimate can quantify the full range of variability. A value of zero means no sensitivity to the measurement. A good rule of thumb is to begin to trust results in which DOFS > 0.7. However, other checks such as through model comparisons or re-analysis are needed to ensure that other effects are not imparting errors that are larger than 50% of the variability (corresponding to DOFS = 0.5).

4.3 Error Derivation Used for Process Studies and Model or Aircraft Comparisons

Next is how the errors are calculated based on the derivation in Worden *et al*. (2006), We use the nomenclature in Worden *et al*. (2006) first and then match the variables in the product files to this nomenclature after.

The estimate for the deuterium content of water vapor can be described by:

$$\hat{x}_R = x_a^R + \left(A_{DD} - A_{HD}\right)\left(x_R - x_a^R\right) + \left(A_{DD} - A_{HH} - A_{HD} + A_{DH}\right)\left(x_H - x_a^H\right) +$$

$$G_R n + G_R \sum_i K_i^b (b_i - b_i^a)$$

Where $\hat{x}_R$ is the (log) of the ratio of HDO and H_2O, or log(HDO/H_2O), as a function of pressure. The x_R is the “true state” which is either generated from a model or in situ profile. The matrices A describe the averaging kernel and its cross-terms for the joint deuterium and H_2O estimate and the last two terms describe the uncertainty due to noise and other radiative interferences (e.g. temperature, clouds, methane, etc.) that affect the HDO and H_2O estimate.

The TROPESS AIRS and CRIS product files provides the averaging kernel (or avg_kernel) $\mathbf{A_R = (A_{DD} - A_{HD})}$ in order to compare the AIRS estimates with either aircraft data or models or some other constructed profile (***see previous section***)

To reduce file size and because we find it is not needed, we do not provide $\mathbf{A_H = (A_{DD} - A_{HH} - A_{HD} + A_{DH})}$.

The uncertainties can be calculated by taking the expectation of the difference between the estimate and the true state.

$$E||\hat{x}_R - x_R|| = \left(I - A_R\right)S_R\left(I - A_R\right)^T + \left(A_H\right)S_H\left(A_H\right)^T + S_R^{meas} + S_R^{int}$$

The covariances, $S_R^{obs} = S_R^{meas} + S_R^{int}$ are provided with the AIRS product product files. The $\left(I - A_R\right)S_R\left(I - A_R\right)^T + \left(A_H\right)S_H\left(A_H\right)^T$ terms are not needed if you follow the model / data recipe in the next section as the first term disappears in the comparison and the second term is typically negligible.

The product files provide the following

A_r = AVG_KERNEL

x_a^R = XA

S_R^{obs} = OBS_ERROR

$\hat{x}_R$ = X

4.4 Example of error calculation for the dd_825_421 measurement

Using the variables defined above we can provide an example of how to calculate the error in the dd_825_421 measurement. There are 5 pressure levels between 825 and 421 hPa used for the HDO_H2O retrieval. The value for the mean of the isotopic composition between 825 and 421 hPa is therefore:

$$\delta D = 1000(\frac{\frac{M\hat{x}_R}{M\hat{x}_{H2O}}}{SMOW} - 1)$$

Where M is a vector with the values of 1/5 corresponding to the 5 pressure levels between 825 and 421 hPa and 0 for all other pressure levels and δD corresponds to the dd_825_421 variable in the product file and $\hat{x}_{H2O}$ is the variable corresponding to the H2O_PROFILE variable. Note that the above assumes a vector multiplication between the two quantities. One can then show that the uncertainty in δD, in units of per mil relative to SMOW, is therefore:

$$\sigma_{\delta D} = (\frac{1000\delta D}{SMOW})(MS_R^{obs}M^T)^{1/2}$$

5 Validation Summary

Based on the NASA Validation criteria listed subsequently, we list the validation stage of the AIRS HDO/H_2O product as Level 3 (see subsequent section). Our rationale is that calculated errors have been evaluated by comparing retrievals against independent satellite and aircraft data sets in the tropics, mid-latitudes, and high-latitudes. CRIS data are at validation stage 2 as they can be compared against AIRS data for a range of latitudes; see subsequent figures.

The references for this validation are:

1) Worden, J. R. et al. (2019), Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor, *Atmospheric Measurement Techniques*, *12*(4), 2331–2339, doi:10.5194/amt-12-2331-2019.

2) Herman, R. L., J. Worden, D. Noone, D. Henze, K. Bowman, K. Cady-Pereira, V. H. Payne, S. S. Kulawik, and D. Fu (2020), Comparison of optimal estimation HDO/H2O retrievals from AIRS with ORACLES measurements, *Atmospheric Measurement Techniques*, *13*(4), 1825–1834, doi:10.5194/amt-13-1825-2020.

3) Herman, R. L., J. E. Cherry, J. Young, J. M. Welker, D. Noone, S. S. Kulawik, and J. Worden (2014), Aircraft validation of Aura Tropospheric Emission Spectrometer retrievals of HDO / H_2O, *Atmospheric Measurement Techniques*, *120*, 3127–3138, doi:10.5194/amt-7-3127-2014.

5.1 Validation of HDO/H_2O

The TROPESS project only reports data where the species retrieval quality is sufficient for science analysis. For HDO/H_2O the degrees of freedom for signal are 0.8 or larger for all profiles. The uncertainties for an integrated quantity between 825 to 400 hPa is ~3% or 30 per mil. The data are bias corrected based on comparison with the aircraft measurements (Herman *et al*. 2014, 2020) and account for the regularization and vertical resolution associated with each measurement.

AIRS HDO/H_2O retrievals have been validated (stage 4) by global comparisons with TES and detailed regional comparisons with in situ HDO/H_2O from the WISPER instrument during the ORACLES Earth Venture Suborbital (EVS-2) aircraft campaign (Herman *et al.*, 2020). Figure 1 below shows AIRS minus WISPER δD for 110 matches that have a spatial coincidence within 0.3 degrees latitude and longitude for same-day measurements. The δD notation is defined above in Section 4.4. The AIRS observation operator has been applied to the WISPER in situ profiles.

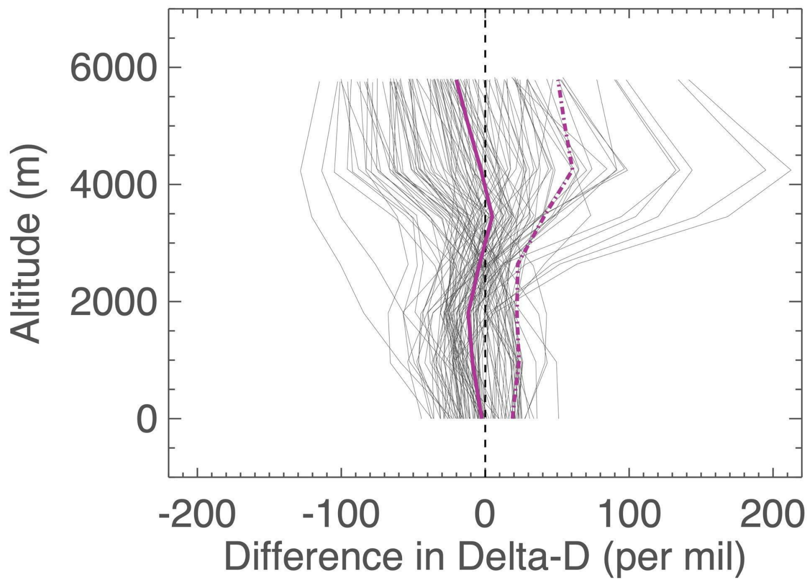

Figure 1. AIRS minus in situ WISPER δD (per mil) matched within 0.3 degrees latitude and longitude (grey lines). For these 110 matches, the mean bias (red solid line) and rms (red dashed line) are also shown.

It is seen that AIRS δD has a mean bias (red solid line) of -7 per mil relative to aircraft. The rms (red dashed line) increases from 21 per mil in the lower troposphere (1000 to 800 hPa) to 45 per mil in the free troposphere (800 to 500 hPa), largely due to spatial variability in the region where the aircraft measured HDO/H_2O.

Figure 2 compares AIRS estimated error (observation error, c.f. Section 4) with empirical error calculated from the statistics of the AIRS-WISPER comparisons. The empirical error is comparable to the optimal estimation AIRS error, increasing from 19 per mil in the lower troposphere to 26 per mil in the free troposphere.

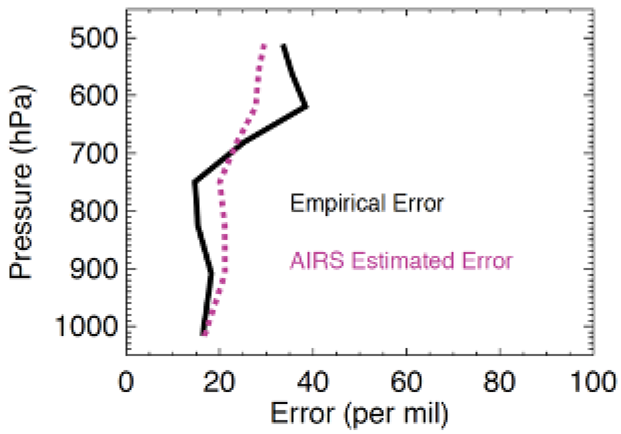

Figure 2. AIRS estimated error (red dashed line) from Optimal Estimation is validated by comparison to the AIRS-WISPER empirical error (black solid line).

5.2 Validation of CrIS HDO/H_2O

CrIS HDO/H_2O has been validated at stage 2 by limited comparisons to AIRS HDO/H_2O, only one day at limited geographical locations. For one day of data, 1 October 2019, CrIS and AIRS δDwithin 1 degree latitude and longitude (DOFS>1.0) are compared in Figure 3 below. Averaged over tropospheric pressures 825 to 421 hPa, CrIS is biased slightly low compared to AIRS: CrIS minus AIRS δDmean bias is -11 per mil with an empirical rms of 38 per mil. By comparison, the global AIRS observation error for this day is approximately 30 per mil.

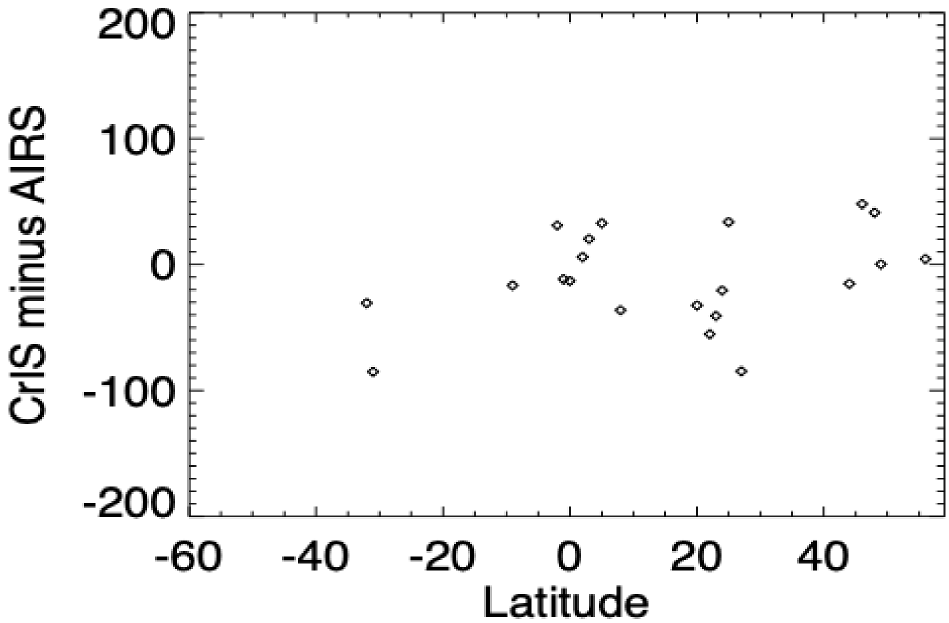

Figure 3. CrIS minus AIRS δD matched within one degree latitude and longitude for October 1, 2019, DOFS greater than 1.0. The mean difference is -11 per mil.

Appendix A. Retrieval levels

The table below contains the nominal retrieval levels. For each sounding, the surface pressure level is inserted into the retrieval levels set. Any retrieval levels below the surface pressure level are omitted.

Index	Pressure [hPa]
1	1040.0000
2	1000.0000
3	908.5140
4	825.4020
5	749.8930
6	681.2910
7	618.9660
8	510.8980
9	421.6980
10	348.0690
11	287.2980
12	237.1370
13	177.8290
14	133.3520
15	74.9896
16	28.7299
17	0.1000

Acknowledgement

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).