Discussions

- NDACC Data Agreement, Documents & Protocols
- DOI: NDACC & IRWG
- HITRAN 2016 & IRWG Linelist
- NDACC Vision, future

NDACC Data Agreement, Documents & Protocols

- Data Documents
  - ➢ Public, Proprietary, Rapid Delivery Data description…
  - Data Use Agreement
  - M & A
  - Data Formats
  - Protocols
- Data Use Agreement
  - ➢ Abbreviated Data Protocols for Providers & Users
  - Data Documents
  - M & A
  - Data Formats
  - Protocols
- Protocols
  - Introduction
  - Steering Appointments & Elections
  - Data Protocols for Providers & Users
  - Measurement…
  - Validation…
    - By WG
  - Intercomparisons, Theory & analysis & Cooperation Networks

Web site directory

Most descriptive document for data users
BUT combined
AND User section secondary to Provider

NDACC Data Agreement, Documents & Protocols

Data Use Agreement (landing page)

Whenever NDACC data is used in a publication the authors agree to acknowledge both the NDACC data center and the data provider as follows:

"The data used in this publication were obtained from *institute or PI name* as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available (see **http://www.ndacc.org**)."

Section from Data Protocol document

If substantial use is made of NDACC data in a publication an offer of co-authorship will be made through personal contact with the data providers or owners.

Users of NDACC data are expected to consult the online documentation and reference articles to fully understand the scope and limitations of the instruments and resulting data and are encouraged to contact the appropriate NDACC PI (listed in the data documentation on the web page) to ensure the proper use of specific data sets.

Those using NDACC data in a talk or paper are asked to to inform the Theory and Analysis Working Group PIs and **J. Wild** so that we may add the publication to our publications list.

NDACC Data Agreement, Documents & Protocols

"NDACC Protocol for Data Providers and Data Users "

Data Users

As stated above, NDACC data products will be archived at the DHF for public availability within one year after acquisition. In addition, some instrument PIs have approved the public release of their verified data on a shorter timescale, while others provide preliminary data to the Rapidly Delivery section of the DHF shortly after acquisition. Within the first year after acquisition, any NDACC data not authorized for early public release are proprietary and their use is only possible via collaborative arrangement with the appropriate PIs. Co-authorship shall be offered on publications resulting from the use of such proprietary data sets.

Accounts on the NDACC DHF are only available to current NDACC investigators. New parties interested in general access to NDACC data should use the "Contact Us" feature on the NDACC web site <www.ndacc.org> indicating whether they wish to become affiliated with a particular Instrument Working Group or to become a member of the NDACC Theory and Analysis Working Group. They will then be directed to the appropriate Working Group Co-Chairs who can guide them in the preparation and submission of a suitable proposal.

All NDACC investigators and data users must acknowledge both the NDACC data center and the data provider in any publication, as follows:

> "The data used in this publication were obtained from *institute or PI name* as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available through anonymous ftp at <ftp.cpc.ncep.noaa.gov/ndacc/station>"

If substantial use is made of NDACC data in a publication (i.e., the data are critical to the study and its conclusions) an offer of co-authorship must be made through personal contact with the data providers or owners. ***In all cases***, scientists using NDACC data are encouraged to contact the relevant instrument PI in order to receive additional information about the data product that could be useful in their studies.

NDACC Data Agreement, Documents & Protocols

NDACC Data Use: Case 1

[Wespes et al., 2016] Wespes, C., Hurtmans, D., Emmons, L. K., Safieddine, S., Clerbaux, C., Edwards, D. P., and Coheur, P.-F. (2016). Ozone variability in the troposphere and the stratosphere from the first 6 years of IASI observations (2008–2013). **Atmospheric Chemistry and Physics**, 16(9):5721–5743.

Table 4. Ozone trends and associated uncertainties (95 % confidence limits), given in DU year^{-1} over NDACC (Network for the Detection of Atmospheric Composition Change) stations in the NH based on daily medians of IASI (within a grid box of 1° × 1° centred on stations, two first rows) and FTIR observations (successive rows for different time intervals). Italic values (second row) refer to trends inferred from subsampled IASI data and bold values refer to statistically significant trends. Values marked with a star (*) refer to trends which are rejected by the iterative backward elimination procedure[a].

DU year^{-1}		Data periods	No. days	25–3 hPa (US)	Total columns
Ny-Ålesund (79° N)	Mar–Sep	2008–2013	1239	0.56 ± 0.73	**5.26 ± 4.72**
		Subsamp.	82	*−0.29 ± 4.58*	*6.26 ± 18.11*
		2008–2012	84	−3.58 ± 4.58	2.24 ± 20.78*
		2008–2012	168	−0.17 ± 0.70*	**−4.84 ± 3.01**
		2003–2012	288	**0.64 ± 0.60**	−1.02 ± 2.40*
		2000–2012	320	**0.62 ± 0.55**	**−2.35 ± 1.40**
		1999–2012	383	**1.03 ± 0.66**	1.31 ± 2.39*
		1995–2012	167	**1.25 ± 1.05**	3.33 ± 3.41
		1995–2003			
Thule (77° N)	Mar–Sep	2008–2013	1094	**1.24 ± 1.09**	**4.97 ± 4.72**
		Subsamp.	*231*	*1.31 ± 2.69*	*0.10 ± 7.36*
		2008–2012	340	−2.10 ± 2.89	0.39 ± 11.59*
		2008–2012	697	0.86 ± 0.89	−2.77 ± 2.99
		2003–2012	776	**1.33 ± 0.86**	−1.29 ± 1.73
		2000–2012	779	**1.69 ± 0.88**	−1.25 ± 1.74
		1999–2012	138	**3.73 ± 2.90**	4.86 ± 10.13*
		1999–2003			
Kiruna (68° N)	Mar–Sep	2008–2013	1236	0.21 ± 1.42	**4.41 ± 4.00**
		Subsamp.	226	*0.97 ± 4.05*	*3.78 ± 6.03*
		2008–2012	254	−1.97 ± 6.04*	−3.75 ± 6.64*
		2008–2012	678	0.15 ± 0.67*	2.26 ± 3.68
		2003–2012	913	**1.60 ± 1.29**	3.69 ± 4.20
		2000–2012	984	**1.10 ± 0.98**	−0.43 ± 1.64*
		1999–2012	1183	**1.11 ± 0.54**	1.82 ± 1.77
		1996–2012	596	**1.26 ± 1.21**	1.12 ± 3.77*
		1996–2003			
Jungfraujoch (47° N)		2008–2013	1580	**2.95 ± 0.61**	**5.64 ± 3.15**
		Subsamp.	*524*	***3.72 ± 1.14***	***5.61 ± 5.11***
		2008–2012	565	1.60 ± 1.80	**5.28 ± 4.82**
		2008–2012	1582	0.10 ± 0.35	−0.28 ± 0.86*
		1998–2012	1771	0.02 ± 0.33*	**0.85 ± 0.79**
		1995–2012			
Zugspitze (47° N)		2008–2013	1729	**3.17 ± 0.56**	**5.53 ± 2.92**
		Subsamp.	*538*	***3.56 ± 1.63***	***5.99 ± 4.49***
		2008–2012	597	0.71 ± 1.22	3.46 ± 3.79
		2008–2012	1472	0.08 ± 0.32*	0.81 ± 0.98
		1998–2012	1525	0.23 ± 0.32	**1.36 ± 1.01**
		1995–2012			
Izana (28° N)		2008–2013	1803	0.56 ± 0.65	**1.28 ± 0.77**
		Subsamp.	*380*	*0.32 ± 1.28*	*0.11 ± 1.95*
		2008–2012	443	0.24 ± 0.80*	0.91 ± 2.44*
		2008–2012	1257	**0.46 ± 0.25**	0.20 ± 0.33*
		1999–2012			

[a] The trend values result from the adjustment of the regression model where the linear term is kept whatever its p value calculated during the iterative process is.

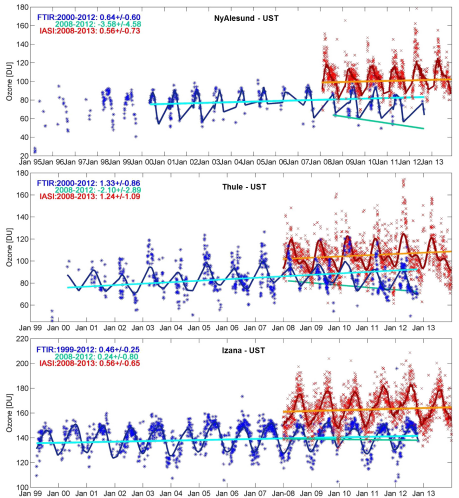


Figure 10. Daily time series of O_3 FTIR (blue symbols) and IASI (red symbols) measurements in the UST at Ny-Ålesund (top), Thule (middle) and Izana (bottom), covering the 1995–2012 and the 1999–2012 periods, respectively (given in DU). The fitted regression models (dark blue and dark red lines, for FTIR and IASI, respectively) and the linear trends calculated for periods starting after the turnaround over 1999/2000–2012 and over 2008–2012 for FTIR (light blue and green lines), and the 2008–2013 period for IASI (orange line) are also represented (DU year^{-1}). The trend values given in DU year^{-1} are indicated.

NDACC Data Agreement, Documents & Protocols

NDACC Data Use: Case 1

Data availability
The ground-based FTIR measurements are accessible from http://www.ndsc.ncep.noaa.gov/data/data_tbl/.

Acknowledgements. IASI has been developed and built under the responsibility of the Centre National d'Etudes Spatiales (CNES, France). It is flown onboard the MetOp satellites as part of the EUMETSAT Polar System. The IASI L1 data are received through the EUMETCast near-real-time data distribution service. Ozone data used in this paper are freely available upon request to the corresponding author. We acknowledge support from the O3-CCI project funded by ESA and by the O3M-SAF project funded by EUMETSAT. P.-F. Coheur and C. Wespes are, respectively, Senior Research Associate and Postdoctoral Researcher with F.R.S.-FNRS. The research in Belgium was also funded by the Belgian State Federal Office for Scientific, Technical and Cultural Affairs and the European Space Agency (ESA Prodex IASI Flow and BO3MSAF). The National Center for Atmospheric Research is funded by the National Science Foundation.

Edited by: M. Van Roozendael

NDACC Data Agreement, Documents & Protocols

NDACC Data Use: Case 2

[Müller et al., 2017] Müller, J.-F., Stavrakou, T., Bauwens, M., George, M., Hurtmans, D., Coheur, P.- F., Clerbaux, C., and Sweeney, C. (2017). Top-Down CO Emissions Based On IASI Observations and Hemispheric Constraints on OH Levels. **Geophysical Research Letters**, 45(3):1621–1629.

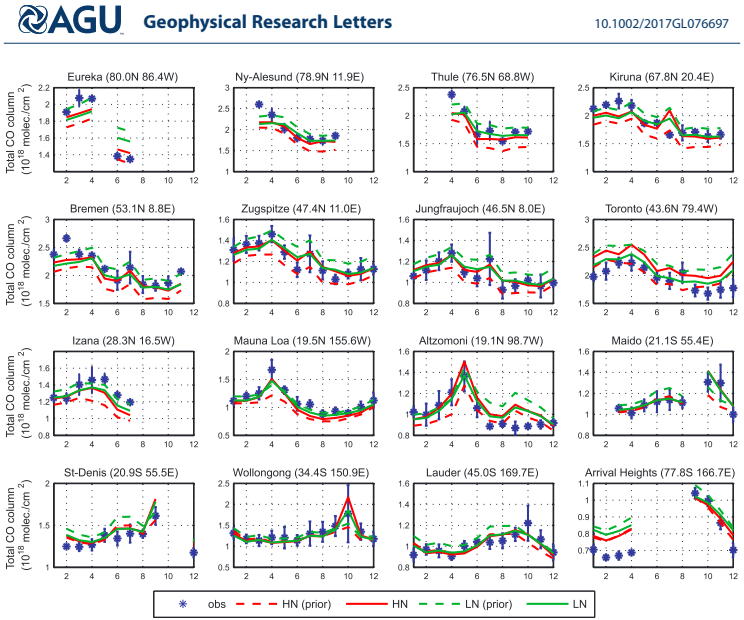


Figure 3. Modeled and observed CO columns at Fourier-transform infrared spectroscopy (FTIR) stations. In dark blue: FTIR monthly means with their standard deviations; dashed lines: modeled values using a priori fluxes; solid lines: modeled values using optimized fluxes. HN run in red, LN in green.

Acknowledgments

This work was supported by the European Commission under the Horizon 2020 GAIA-CLIM project (grant 640276) and the FP7 project QA4ECV (grant 607405) and by the Belgian Science Policy office and the European Space Agency (ESA-PRODEX) through the TROVA and IASI.flow projects. The IASI sensor has been developed and built under the responsibility of the Centre National d'Etudes Spatiales (CNES, France). It is flown as part of the EUMETSAT Polar System. The IASI L1 data are received through the EUMETCast near real-time data distribution service. We gratefully acknowledge all the PIs and data managers of the measurement networks used in this study. We thank Q. Errera for the BASCOE reanalysis of Microwave Limb Sounder data.

The authors acknowledge the unlimited and without charge access to the NOAA/GMD, GAW, and NDACC measurement networks. The data used in this study are listed in the references and supporting information.

NDACC Data Agreement, Documents & Protocols

Next Steps

What is substantial use?

- Graph or plot or table?
- Should it be completely at the discretion of the data user?
- ?

Suggestions:

- Keep Data User policy separate form Data Provider policy on website
- Data user more obvious – most strict statements at top of text
- All members should be on 'look out' for possible issues
- ?

Note to data provider:

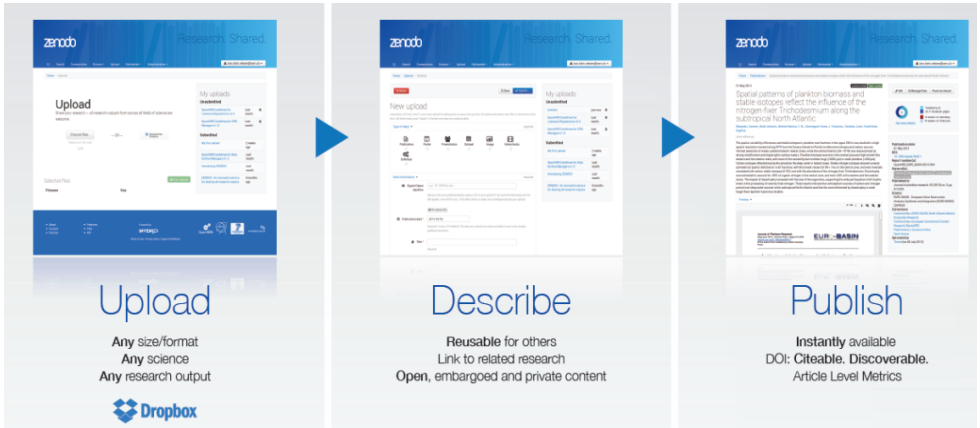
- Contact PI? ->> keep meta data file up to date!

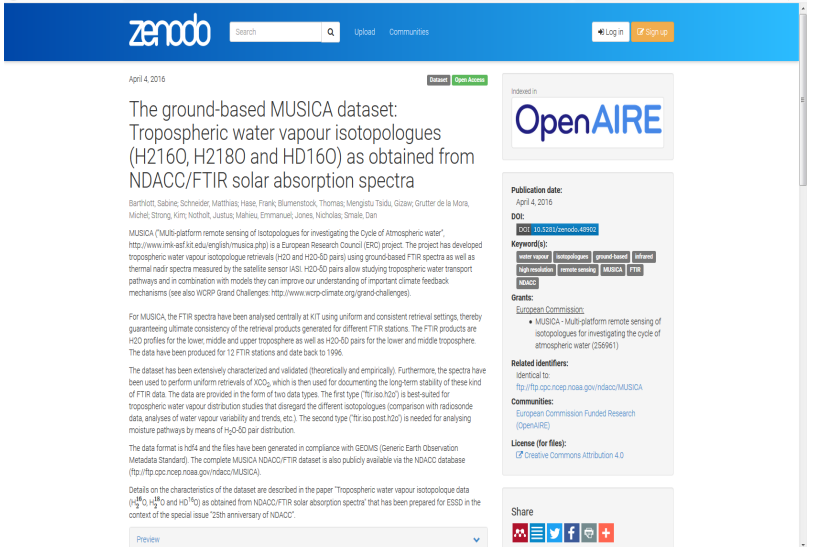
DOI 1/

- NDACC DHF (NOAA) will never issue DOI for NDACC data – it is not sufficiently constrained
    - It may be possible for a WG – still highly constrained
        - Single code written and documented to their specific standard submitted
        - Technical report, flow charts documenting processing
        - Raw data submission
- DOI obtained elsewhere would
    - Require data archived there and DOI points there
    - Links likely provided to NDACC DHF – could be issue for NDAACC DHF
- MUSICA can be found here: doi:10.5281/zenodo.48902.

Zenodo – A fast way of getting DOIs?

Sabine Barthlott





Source: http://doi.org/10.5281/zenodo.8428

From Sabine B.

About Zenodo

- **Where**: https://zenodo.org/
- **Who**: everyone from everywhere regardless of funding source
- **What**: every research output from every field of science (e.g. publications, posters, presentations, datasets, images, software, videos,….)
- **Format:** every format allowed
- **Size limits:** 50GB per dataset, multiple dataset allowed (in case of larger files, please contact zenodo)
- **Is the data safe?** Yes. Data is stored in CERN Data Center. Both data files and metadata are kept in multiple online and independent replicas.
- **Who can access my upload?** Depends on the license and access level you choose for your dataset.
- **What do I need for the upload?** metadata informations & file(s)
- **What if dataset has to be changed?** Zenodo supports DOI versioning: you automatically get a DOI for the specific version and one for all versions → for publications, usually specific version DOI should be used (Changes in metadata information possible without version change)

For more informations, see: http://help.zenodo.org/

From Sabine B.

Generating a DOI with Zenodo

- **File:** choose your file, still possible to remove again before publishing
- **Communities**: Any user can create community collection (e.g. NDACC-IRWG‘ ☺) community collection owner has to accept upload to this collection; multiple communities possible; New communities can be created here: https://zenodo.org/communities/new/
- **Upload Type**: Choose the type of your upload (Publication, Poster, Dataset, …)
- **Basic Information**: Author, Title, Description, Keywords,….
- **License**: Access rights (open, restricted, closed,..) & License
- **Funding**: OpenAIRE-supported projects only (e.g. EU grants). For other funding acknowledgements, please use the Additional Notes field
- **Related/alternate identifiers**: related publications and datasets (e.g. DOI, URL [e.g. if same dataset has been uploaded to NDACC database],..)
- **References**
- …

From Sabine B.

DOI 2/

Pros
- Well defined dataset across all sites
- Referenceable / linkable data / citeable
- Easier access for users
- Specific defined versioning

This will demand of the IRWG
- Documented analysis, likely beyond current papers
- Possibly a specific retrieval code?
    - But **certainly** more defined and adhered to retrieval parameters
- Processing reliant on all groups -> deadlines
- Reanalysis by everyone

Simplifying Issues
- Central processing…
- Obtain DOI by species, so one at a time…

Discussion on spectroscopic data set

- In NDACC we still use Hitran 2008.
- Switch to HITRAN 2016/Geoff‘s line list/individual line lists for each species or keep Hitran 2008?
- Pro: Better spectr.oscopic data are needed, in particular for CH4
- Contra: Huge effort: Reprocessing of entire data set necessary
- Pro: Maybe time to revisit WACCM apriori
- Contra: Data archiving status may degrade in the meantime
- Discussion: To decide species by species and
- To assign responsible group for each species to check which data set is best and whether a switch is recommended.
- NDACC standard species: O3, HCl, HF, ClONO2, HNO3, N2O, CH4, CO, C2H6, HCN, starting with CH4

Discussion on spectroscopic data set of each NDACC standard species:

- CH4: New line list available including line mixing and Hartmann Tran line shape
- O3:
- HCl:
- HF:
- ClONO2:
- HNO3:
- N2O:
- CO:
- C2H6:
- HCN:
- And interfering species: H2O, …!
- H2CO
- OCS

Jim – Gather initial listing from HITRAN 2016, Geoff ATM for CO2 & O3 (already reviewed)

Responsible persons for each species

Species	Person / Group
CH4	
O3	
HCl	
HF	
CLONO2	
HNO3	
N2O	
CO	
C2H6	
HCN	
H2O	
H2CO	
OCS	

Visions for NDACC Future

- Would it be useful to have an NDACC Strategy & Implementation Plan?
  - Water Vapour strategy document could be part of it?

- There is a fairly open discussion at the SC about NDACC future directions. The request is from the SC to all members for input on any topic that might affect his or her view of business as usual of the NDACC.
- There was not an expectation that we should change or concern that about the current direction of the NDACC. Rather extending an invitation for new ideas, inclusion, not missing issues of import from the experience of the membership.
- Instrumentation, data protocols or policies, membership, affiliated groups etc.

Visions for NDACC Future: IRWG

- The effect of operational demands, becoming a small component in a large infrastructure project

- Attracting new young scientists
- Retaining experienced scientist
- Requires interesting challenging science questions & progress
- Add new species to ‘required’ list?

Visions for NDACC Future: IRWG Data Quality

What could we do to improve
- data quality
- visibility

Visibility

- Paper describing data products & uncertainties across stations

IRWG Data version identifier
- Tied to specific code(s) version
- And Linelist
- And WACCM
- And criteria below

Data quality

For Ten Archival species

- Define Sa for each
- HBr / N2O cells
- Require H2O pre-retrieval / sequence of retrievals
- Restrict use of other retrieval parameters (shifts, phase, channels, ils…
- Require use of ILS

- Require QC on spectra
  - Specify minimum SNR

- Require QC on retrieval
  - Fit RMS maximum
  - DOFS minimum / maximum
  - Expand Bavo's effort for CO…

Visions for NDACC Future: IRWG Data Quality

Update our retrieval parameter criteria

- Recipe for 3020-3040cm-1 O_3
- Recipe for H2O per-retrievals

- WACCM v6 ~2010
  - 1980 – 2020 nearing end!

Gas	Required μ-w(s) [cm^{-1}]	Optional μw	OPD [cm]	Interfering species to be fit (pre- or simul-)	a Priori Linelist	Column or Profile	Note
O3	1000.0-1005.0		250	H2O, CO2, C2H4, O668, O686	WACCMV5 HIT08	P	a,e
		782.56-782.86					
		788.85-789.37					
		993.30-993.80					
HCl	2727.73-2727.83		>180	O3, HDO	WACCMV5	P	
	2775.70-2775.80			N2O, O3	HIT08		
	2925.80-2926.00			O3, CH4, NO2			
HF	4038.81 4039.07		>180	H2O, HDO, CH4	WACCMV6	P	c
		4000.86-4001.10		H2O, O3	HIT08		
		4109.77-4110.07		H2O, HDO, CH4			
ClONO2	780.10-780.35		>50	H2O	WACCMV5	C	d,g
		780.0-781.3		CO2, O3	HIT-XC/PL		
		779.0-780.0		H2O			
HNO3	867.05-870.00				WACCMV5	P	
		872.25-874.00			HIT08		
N2O	2481.30-2482.60		250		WACCMV5	P	
	2526.40-2528.20				HIT08		
	2537.85-2538.80						
	2540.10-2540.70						
CH4	2613.70-2615.40		250	HDO, CO2	WACCMV5	P	b
	2835.50-2835.80			HDO	HIT00		
	2921.00-2921.60			HDO, NO2, H2O			
		2650.60-2651.30		HDO, CO2			
		2903.60-2904.03		NO2			
		2611.60-2613.35		HDO, CO2			
		2613.70-2615.40		CH4, CO2, HDO			
		2914.70-2915.15		CH4, NO2,H2O,HDO			
		2941.23-2942.23		CH4,H2O,O3			
CO	2057.70-2058.00		250	O3, CO2, OCS	WACCMV5	P	
	2069.56-2069.76			O3, CO2,OCS	HIT08		
	2157.50-2159.15			O3, CO2, N2O, H2O			
C2H6	2976.66-2976.95		250	H2O, O3, CH4	WACCMV5.1	by site	d,f
	2983.20-2983.55			H2O, O3, CH4	HIT-XC/PL		
		2986.50-2986.95		H2O, O3, CH4			
HCN	3268.05 - 3268.40		250	H2O, C2H2	WACCMV5.1	by site	d
	3287.10 - 3287.35			H2O, CO2, C2H2	HIT08		
	3299.40 - 3299.60			H2O, H218O			
		3277.775 – 3277.950		H2O			
		3286.168 – 3288.482		H2O			
		3331.400 – 3331.800		H2O, H217O, CO2, N2O			
		3301.030 – 3301.300		H217O			
		3304.825 – 3305.600		H2O, H218O, H217O, C2H2			

Discussions

Done

NDACC Data Agreement, Documents & Protocols

Data Documents

1. **Public data**
2. **Proprietary data**
3. **Rapid delivery data**

NDACC Data Agreement, Documents & Protocols

Data Documents

1. Public data

All NDACC data more than one year old are public data. Additionally some PIs have authorized their data for early release. These data are available as soon as they are cataloged in the database. This public record is available through anonymous ftp at **ftp.cpc.ncep.noaa.gov/ndacc** or through a **station search**.

NDACC Data Agreement, Documents & Protocols

Data Documents

2. Proprietary data

All NDACC data newer than one year from acquisition and not authorized for early release are proprietary data and available only through direct accounts on the NDACC database. Parties interested in general access to this data should prepare a proposal to the NDACC Steering Committee for consideration as NDACC collaborators. For more information contact the **NDACC Steering Committee co-Chairs**. Otherwise, the use of any individual NDACC data set prior to its being made publicly available (i.e., for use associated with field campaigns, satellite validation, etc.) is possible via collaborative arrangement with the appropriate PI(s).

Data Documents

3. Rapid delivery data

For some projects it is of value to offer data to the scientific community with a maximum delay of one month. If these rapid delivery data are of less quality of traditional NDACC certified data, if the data have not yet been quality controlled, or if the data is less complete (e.g. missing uncertainty estimates) then these data must be identified as 'Rapid Delivery (RD)'. These are available separately on the NDACC public website at **ftp.cpc.ncep.noaa.gov/ndacc/RD**. As soon as the standard verified version is available, the RD data will be removed and the fully verified version will be archived in the NDACC archive.

NDACC Data Agreement, Documents & Protocols

Data Delinquency?

NDACC Data Agreement, Documents & Protocols

All NDACC PIs shall submit their certified NDACC data to the NDACC DHF, at latest one year after data acquisition, in the appropriate data format, which is the NASA Ames format dedicated to the measurement technique, or the GEOMS HDF format, following the template approved by the corresponding Instrument Working Group – see https://avdc.gsfc.nasa.gov/GEOMS).

A more rapid data publication is encouraged. To this end, the Rapid Delivery data directory in the NDACC DHF accepts NDACC data before complete NDACC certification. Data in the RD directory are public.

Compliance with the data archiving policy is important for the reputation of the Network.

In case of non-compliance with the above data archiving policy,

a. In first instance, the corresponding Instrument Working Group (IWG) Co-Chairs will communicate with the PI to remind the PI of the data submission requirements.

b. If the data submission issue is not resolved after 2 reminders by the IWG in a 6-months period, without any acceptable justification by the PI, the Co-Chairs will send a follow-up letter to the PI's institute to inform the appropriate parties about the data delinquency, and to ask for support to the PI in order to enable compliance with the NDACC policy.

c. In the case where neither of the two above actions in a 1-year timeframe lead to compliance with the NDACC data archiving protocol, the IWG co-chairs will inform the NDACC co-chairs as well as the DHF manager, and the instrument will be declared as 'not providing data' in the Measurements and Analyses Directory, and will be flagged as 'suspended data provision' on the map and in data searches in the DHF

d. In case the measurements are no longer performed, the station will be flagged as 'inactive' in the Measurements and Analyses Directory, on the map and in data searches in the DHF.