# NDACC/WOUDC Interoperability: Progress, Status, Next Steps

**Jeannette Wild**
**NDACC DHF Curator**
**ESSIC/UMD, NOAA/NWS/NCEP/CPC**
**http://www.ndacc.org**

**Tom Kralidis**
**Senior Systems Scientist**
**Meteorological Service of Canada**
**https://woudc.org**

**12 September 2018**

# Overview

- Requirement
  - WMO Ozone SAG 2018 meeting request to increase interoperability of NDACC data into WOUDC
- Purpose
  - Provide history of prior work
  - Outline issues, pros/cons
  - Provide options for future to support requirement

# 2002-2005

- Early NDSC data protocol establishing WOUDC/NDSC relationship
- 2002: Testing (20) Ozonesonde files with 2160 convertor
  - Run annually
- 2004/2005: semi-annual/annual transfer schedule (OzoneSonde) setup
  - initial conversions resulted in 40% success rate
  - microwave data submitted

- Lessons learned
  - versionitis
  - re-submissions
  - PI interpretation

# 2010: WMO ET-WDC meeting, Toronto

- WOUDC, NDACC, SHADOZ, AURA, WMO/SAG, EC, etc.
- Discussed:
  - data integrity between copies between archives
  - version sync between data centres
  - data removal
  - unified metadata / vocabularies
  - proper citation
  - data rights

Lots of issues discussed, but not sure solutions were devised

# 2008-2011 DCIO (ESA)

- Data Center Interoperability Project)/ GECA  Transitions to GEOMS (Generic Earth Observation Metadata Standard)
  - AVDC, EVDC, GEOMON, NDACC, ESA, WMO, GAWSIS
  - Single point login
  - Data format standards
  - Ability to search for data from many databases in one tool (XML)
  - Exchange of data/metadata using OAI-PMH (metadata harvest API)
  - Licensing/data protocols/data ownership
  - data resided at Data centres (copying was discussed)
  - birth of GEOMS as a result
  - Results: exchanged metadata in XML for use by search engines.  Each database using tool will refer back to originating database.  Preserves integrity of data and ownership.  Reduces versionitis

http://earth.esa.int/fringe09/proceedings/papers/s1_5meij.pdf
http://evdc.esa.int/documentation/oai-pmh/

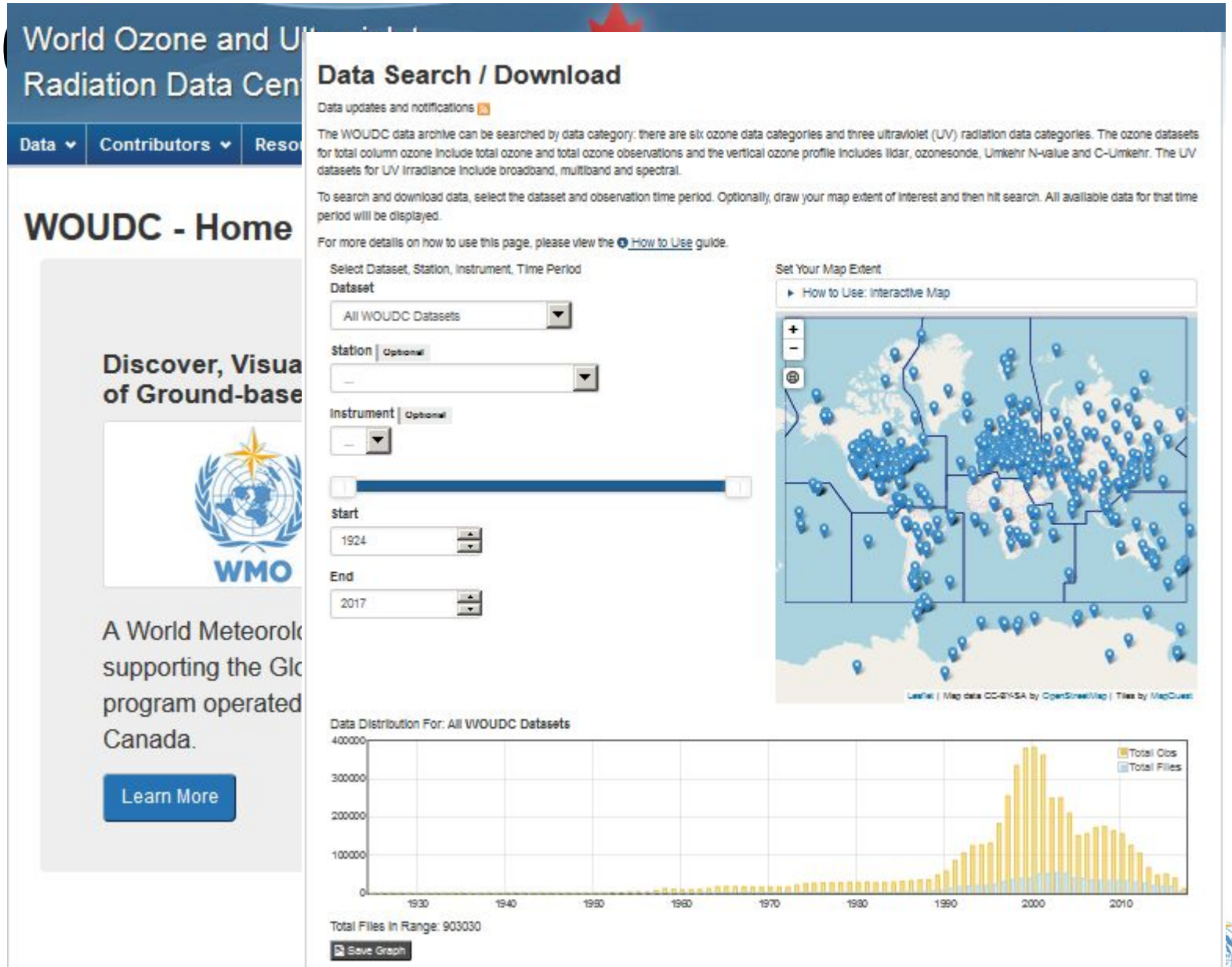
# 2012-2015

- 2nd AMES 2160 conversion effort (WMO SAG Ozone interest)
- WMO SAG Ozone interest to search NDACC on WOUDC
- serve on WOUDC FTP
- issues
  - station lookups
  - instrument information
  - multiple versions of AMES format
- revisit of approach required

# 2015: WOUDC Renewal

- focus on standards, data access, interoperability
- enhance access to Ozone and UV data and metadata
- streamline data centre processing workflow
- set course for GAW federation
- alignment with national/international Open Data efforts
  - Canada Open Data
  - WMO WIS (DCPC compatibility)
  - GEOSS
- data centre reporting metrics
- interactive maps/graphs

2

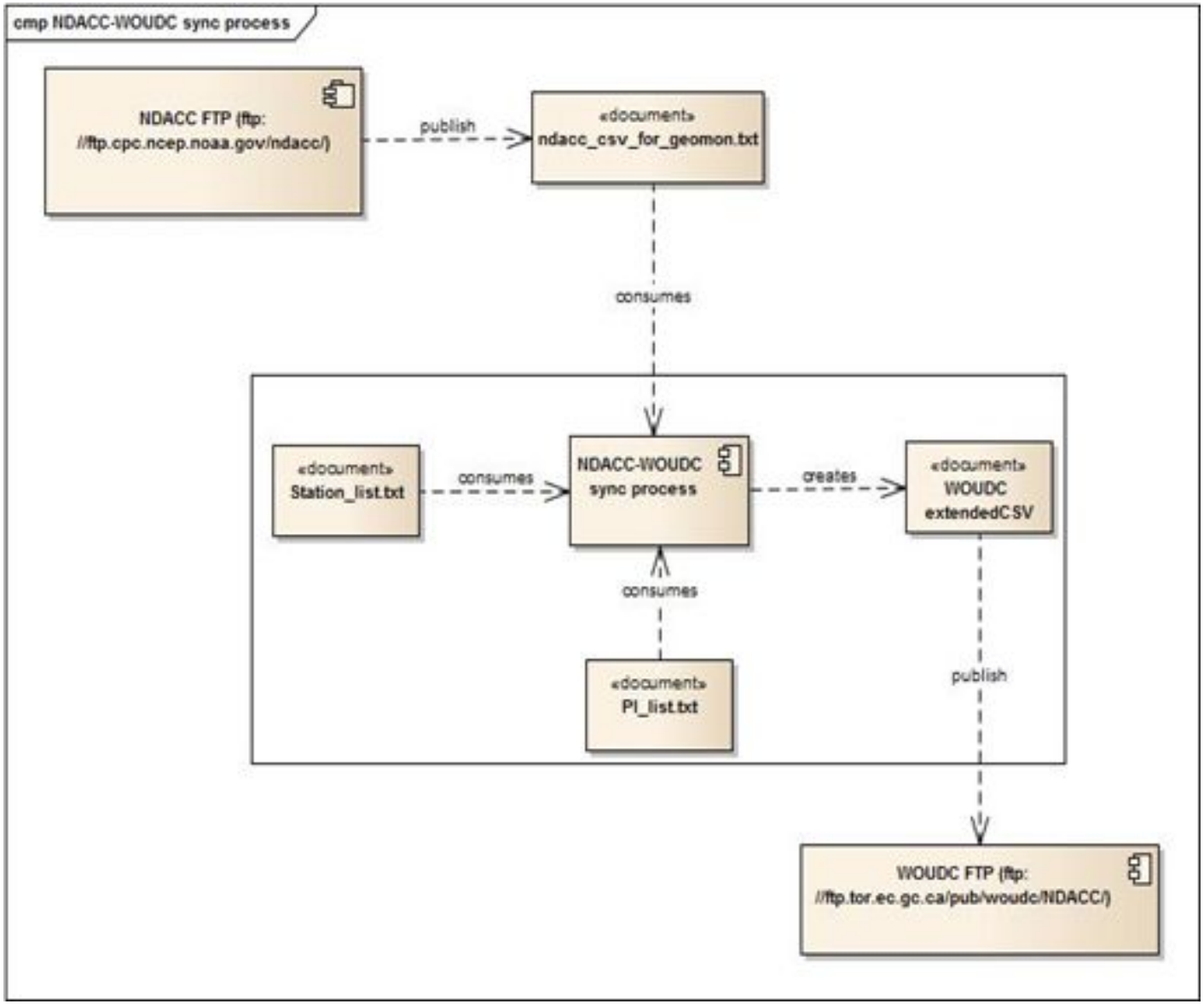

World Ozone and U
Radiation Data Cen
Data
Contributors
Reso
WOUDC - Home
Discover, Visua
of Ground-base
WMO
A World Meteorol
supporting the Gl
program operated
Canada.
Learn More
Data Search / Download
Data updates and notifications
The WOUDC data archive can be searched by data category: there are six ozone data categories and three ultraviolet (UV) radiation data categories. The ozone datasets for total column ozone include total ozone and total ozone observations and the vertical ozone profile includes lidar, ozonesonde, Umkehr N-value and C-Umkehr. The UV datasets for UV irradiance include broadband, multiband and spectral.
To search and download data, select the dataset and observation time period. Optionally, draw your map extent of interest and then hit search. All available data for that time period will be displayed.
For more details on how to use this page, please view the How to Use guide.
Select Dataset, Station, Instrument, Time Period
Dataset
All WOUDC Datasets
Station Optional
Instrument Optional
Start
1924
End
2017
Set Your Map Extent
How to Use: Interactive Map
Data Distribution For: All WOUDC Datasets
400000
300000
200000
100000
0
1930
1940
1950
1960
1970
1980
1990
2000
2010
Total Obs
Total Files
Total Files in Range: 903030
Save Graph



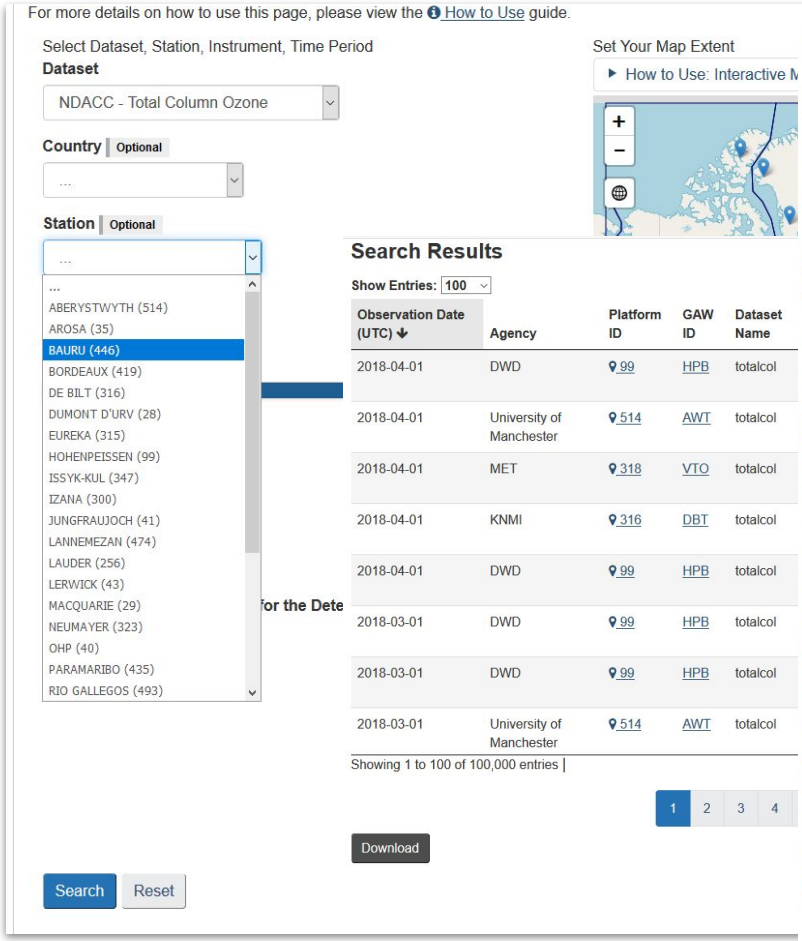
WMO GAW

# 2015-2016

- Distributed search
- Implementation in the spirit of DCIO (and we didn't know it)
  - Total column ozone
  - Vertical Ozone Profiles
  - UV irradiance
- Search on WOUDC, download from NDACC
- Results
  - download metrics at source data centre
  - data closest to source
  - no data distribution / metrics
  - incompatible formats between data centres


cmp NDACC-WOUDC sync process
NDACC FTP (ftp://ftp.cpc.ncep.noaa.gov/ndacc/)
publish
«document» ndacc_csv_for_geomon.txt
consumes
«document» Station_list.txt
consumes
NDACC-WOUDC sync process
creates
«document» WOUDC extendedCSV
consumes
«document» PI_list.txt
publish
WOUDC FTP (ftp://ftp.tor.ec.gc.ca/pub/woudc/NDACC/)

# Current WOUDC / NDACC Interoperability




```
VAUGHAN G.          UV/VIS      ABERYSTWYTH TOTALCOL    01-APR-2018 00:00:0030-APR-2018 23:59:590403
58  1010
Geraint Vaughan
University of Manchester, United Kingdom
SAOZ 1024 UV/Visible Spectrometer - Aberystwyth, Wales, UK
NDSC
1    1
2018 04 01  2018 05 01
0
julian day of current year
20
1e+13  1e+14  1e+12  0.01  1e+16  1e+17  1e+15  0.01  .001  .001
1e+13  1e+14  1e+12  0.01  1e+16  1e+17  1e+15  0.01  .001  .001
9999  9999  9999  9999  9999  9999  9999  9999  9999  9999
9999  9999  9999  9999  9999  9999  9999  9999  9999  9999
NO2 vertical column density (430 nm) ; molecules/cm**2 Morning
NO2 column density, slant path (430 nm) ; molecules/cm**2 Morning
NO2 error bar (430 nm) ; molecules/cm**2 Morning
Air mass factor for NO2 (430 nm) Morning
O3 vertical column density (510 nm) ; molecules/cm**2 Morning
O3 column density, slant path (510 nm) ; molcules/cm**2 Morning
O3 error bar (510 nm) ; molecules/cm**2 Morning
Air mass factor for O3 (510 nm) Morning
colour index: F550/F350 at 90. Morning normalised
colour index: F550/F350 at 93. Morning normalised
NO2 vertical column density (430 nm) ; molecules/cm**2 Evening
NO2 column density, slant path (430 nm) ; molecules/cm**2 Evening
NO2 error bar (430 nm) ; molecules/cm**2 Evening
Air mass factor for NO2 (430 nm) Evening
O3 vertical column density (510 nm) ; molecules/cm**2 Evening
O3 column density, slant path (510 nm) ; molecules/cm**2 Evening
O3 error bar (510 nm) ; molecules/cm**2 Evening
Air mass factor for O3 (510 nm) Evening
colour index: F550/F350 at 90. Evening normalised
colour index: F550/F350 at 93. Evening normalised
10
1.0  1.0  1.0  0.01  0.01  .001  1.0  1.0  1.0e+13  1.0e+16
9999  9999  9999  9999  9999  9999  9999  9999  9999  9999
year ; UT
month number ; UT
day of month ; UT
latitude of observation ; decimal degrees ( north is positive )
longitude of observation ; decimal degrees ( east is positive )
solar zenith angle ; degrees
air temperature ; degrees Kelvin
type of observation
NO2 residual column in reference ; molecules/cm**2
O3 residual column in reference ; molecules/cm**2
0
9
0 - On direct sun
1 - On direct moon
2 - On blue zenith sky
7 - On zenith cloud
8 - On zenith sky   ( no weather observation )
STATION : Aberyswyth
The morning and evening values for ozone and NO2 are derived from
a Langley plot method for solar zenith angles between 86 and 91 degrees.

   91 2018    4    1 5242 -407 8791  292    8 1090 1713
      296  336   69 1505 1073 1271   51 1344 9999 9999
     9999 9999 9999 9999 1166 1176   65 1150 9999 9999
   93 2018    4    3 5242 -407 8764  297    8 1090 1713
     9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
```

# 2016-2017

- Boulder data pilot: WOUDC to provide AMES format output for harvest by NDACC
- Issues of data coverage between WOUDC and NDACC
  - architecture of participation of where to submit needed
  - similar questions with EUBREWNET
- Result: Ended with an exploration of delivering Real Time data directly from Boulder to WOUDC. Initial prototyping started
- Initiated development of WOUDC / NDACC conversion utility

# WOUDC Format Library

- woudc-formats
  - Converts various formats into WOUDC Extended CSV format
  - SHADOZ (using pyshadoz)
  - NASA AMES (via nappy from CEDA), supports NDACC variation
  - Vaisala
  - Free and Open Source
  - Built for reuse

- making calls that need PI involvement
- many variations

# GEOMS

- GEOMS/HDF as unified data format
  - Data contributions
  - Data access

# Issues, Pros/Cons

- Mirroring
  - ↑ Consistent data formats
  - ↑ Contributes to WOUDC metrics / reporting
  - ↓ Station lookup variations
    - WOUDC standardized on GAW/OSCAR
  - ↓ Derivations, calculations, interpretations, PI involvement
  - ↓ Managing updates to data (versionitis)
- Distributed Search
  - ↑ Data managed closer to data centre source
  - ↑ Reduces data centre data duplication
  - ↓ Inconsistent data formats
  - ↓ Integration effort still on the user

# Future Options

- GEOMS/HDF formats
- NDACC provides export of WOUDC CSV as a service
- NDACC and WOUDC provide same format exports
- WOUDC enhances distributed search with integrated distributed metrics
- Investigate architecture of contribution between/across data centres
  - submit data once

# Conclusion

- Interoperability doesn't happen by accident *
- It's all about the format[s]

* Cliff Kottman: http://www.opengeospatial.org/ogc/honors/kottman