ET-WDC Meeting Report

WebEx Teleconference on 2 December 2015

Agenda

1. News from WDCs and GAWSIS (tour de table)
2. Recent developments on WIGOS Metadata Standard and implications for WDCs (Klausen)
3. Conclusions from the GAW Data Management Meeting at MeteoSwiss in August: Setup of a federated GAW data management infrastructure (Klausen)
4. Development of new GAW Implementation Plan, new ToR for WDCs and ET-WDC (Klausen, Schultz)
5. Development of work plan for a federated GAW data management infrastructure (Colavecchia)

Participants

Van Bowersox
Geir Braathen (WMO Secretariat)
Tony Colavecchia
Markus Fiebig
Oleg Goussev
Jörg Klausen (chair)
Hiroshi Koide
Martin Schultz
Kjetil Tørseth

Apologies

Vincent-Henri Peuch
Anatoly Tsvetkov

During the opening of the meeting, Geir Braathen said that he would like to give a brief information about Obs4MIPS, and this was inserted as a new point no. 2 on the agenda.

1. News from the WDCs and GAWSIS

GAWSIS

Jörg Klausen informed about the status of the new GAWSIS, which has been available in a beta version for a few months already. The beta version is running stably and regular integration of external data is

"work in progress". It is expected to go live in February 2016.

WDCPC & QA/SAC Americas

Van Bowersox informed about the status at the World Data Centre for Precipitation Chemistry and the QA/SAC on Precipitation Chemistry.
The Acid Rain 2015 meeting took place from 19-23 October in Rochester, New York. This is an international meeting that takes place about every 5 years. In conjunction with this meeting, the SAG on Total Atmospheric Deposition (formerly Precipitation Chemistry) had a meeting. The meeting was led by the new SAG chair, Ariel Stein. The SAG agreed on Terms of Reference. These ToRs will have to be finally approved by the OPAG EPAC SSC.
Currently about a dozen stations submit data to WDCPC. There are more stations "out there" and there is a need to go out and hunt for data. Typically one obtains only aggregated data from the stations, such as monthly or annual means but no individual data points. This is a long-standing problem. At the SAG TAD meeting there was agreement that the WDC PC will be taken over by the National Atmospheric Deposition Program (NADP), which is led by David A Gay. This transition is planned to take place over the next three years. Silvina Carou in the WMO Secretariat is involved with this from the WMO side.

WOUDC

Tony Colavecchia gave a brief overview of the status at WOUDC. The renewal of WOUDC is ongoing, investing in value-added data products, e.g. an integrated data set on UV index for the medical community. A new data submission guidebook will be ready by the end of this year. A new QA framework is on the way. The ftp site is being renewed with a more stable environment for data contributions. There will be a move towards web accessible folders, both for data submission and for data retrieval.

WDCRG (newly established)

Kjetil Tørseth informed about the newly established World Data Centre for Reactive Gases (WDCRG), which will be hosted at NILU. Most of the reactive gases that so far have been stored in WDCGG at JMA will now be taken care of by NILU, using their EBAS system. The gases in question are surface ozone, NOxy, SO2, and VOC. Gases that are relevant for the carbon cycle community will remain at WDCGG. The WDC for some gases (in particular carbon monoxide) is still under dispute. The responsible person at NILU for WDCRG will be Sverre Solberg. There was agreement that he should be invited to become an ex-officio member of the ET-WDC.

WDCGG

Hiroshi Koide informed about the latest developments at WDCGG. The move of some species from WDCGG (JMA) to WDCRG (NILU) was explained at the GGMT meeting at La Jolla, California in mid September 2015. The WMO Greenhouse Gas Bulletin was issued on 9 November at a press conference in Geneva. The rebuilding of WDCGG is progressing. Koide hopes to be able to share a prototype early in 2016. There are discussions under way with the ICOS Atmospheric Thematic Centre. Koide informed the meeting that Ken Masarie at NOAA will retire at the end of 2015. This will most likely have an effect on the GHG data collection at NOAA. Koide offered to share the WDCGG rebuild work plan, intended system architecture and interface design. Next year he will be able to share more. Feedback from ET-WDC would be welcome.

WDCA

Markus Fiebig gave an overview of the status at the World Data Centre for Aerosols. There are several projects with AeroCom on climate modeling and systematic comparison of optical and microphysical observations with AeroCom models.
Two new EU projects have been funded:

1) ACTRIS-2 is a continuation of ACTRIS. Also stations outside of Europe are included. One of the aims of this project is to streamline the interaction between users and providers. ACTRIS-2 could also be linked to the Obs4MIPS activity (see below). Within this project metadata will be in conformance with ISO19115. There will also be an interactive data submission portal where data providers will get immediate feedback until the submitted file passes the QA check. Fiebig promised to share documents once available.

2) ENVRIplus is a project about architecture and data management. Information scientists will also take part in this project and one will aim at developing a common reference model, also for data infrastructure. The aim is to obtain interoperability across data communities and domains (ocean, atmosphere, terrestrial).
In this project one will also look at how one can track the use of scientific data in the same way as is done for scientific articles. This will be done through DOIs that are assigned to data sets. For this to be effective one needs to find a suitable level of granularity when one assigns DOIs to data sets. One will have to find a middle way between the two extremes: 1) One DOI for the whole data centre and 2) One DOI for each single data point. It will be up to each data centre to define a suitable DOI granularity. Klausen expressed the wish that Fiebig can help the ET-WDC members wrt the use of DOIs.
More information can be found on the project web site: http://www.envriplus.eu/

WDC-RSAT

Oleg Goussev, standing in for Julian Meyer-Arnek, gave an orientation on work going on at WDC-RSAT. At this data centre they work on many of the same issues as at WDCA, such as the issue of DOIs. A tool for data submission of time series is under development. A renewal of the interface is also under way. Goussev offered to share software if there is interest.

At this point Klausen called for more close collaboration between the various WDCs and pointed out the possibilities of synergies. All the WDCs ought to be involved in the above-mentioned projects. Tørseth informed about a new GEOSS-related call on data interoperability under Horizon 2020. The deadline for proposals is 8 March 2016. More information about this call can be found here: https://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/2207-sc5-20-2016.html

2. Obs4MIPS

Braathen gave a brief orientation on the Obs4MIPS initiative, which had been discussed at the SPARC Scientific Steering Group meeting in Boulder in mid November. The aim of this initiative is to make observational data available in a harmonised way to simplify comparison with model output. Also in NDACC this is an issue being discussed. This is also relevant to the service oriented direction of GAW in the new Implementation Plan. Klausen encouraged everybody to study the web site on Obs4MIPS: https://www.earthsystemcog.org/projects/obs4mips/

3. Recent developments on WIGOS Metadata Standard and implications for WDCs

Klausen informed about recent developments on the WIGOS metadata standard. This is still work in progress. The WIGOS metadata standards, which is a semantic standard, is approved (Cg-17) and binding. It is being worked on to become a formal standard. WMO RA VI had a meeting last week where

INSPIRE was mentioned. The INSPIRE metadata encoding will be considered. An XSD (XML Schema Definition) for metadata is now at the Alpha-2 stage. OGC compliance and at least partial INSPIRE compliance are sought.

The WIGOS Metadata Standard WMDS will very likely be encoded using the O&M (ISO 19156) framework. O&M also links to MD metadata (ISO 19115) but extends it with specific information on the observation. So there is some overlap. WIS metadata are currently encoded using ISO 19115-1, but there is discussion to move to ISO 19115-2 which uses the MI classes instead of the MD class. There is an open issue concerning the granularity of the WIS vs WIGOS metadata records, but the recent IPET-MDRD-3 meeting concluded that WIS metadata records shall reference WIGOS metadata records in future, using the new WMO ID (aka, WIGOS ID). For the WDCs and GAWSIS, and the interoperability within GAW data infrastructures, this means that metadata will have to be encoded as XML in the specified WMDS format.

The current alpha-2 of the encoding rules are attached, together with a template for feedback. Please note that this encoding is not quite final, but a Beta version will be available in a week or two. WDCs will have to submit metadata compliant with this standard, and it might be that GAWSIS will not be completely comprehensive if the WDCs arelate. The Alpha-2 is distributed as an attachment to these minutes, and Klausen also offered to help the WDCs as best he can in providing metadata to GAWSIS. He is open to bilateral discussions. Tony Colavecchia informed that he can have such metadata in XML format by March 2016.

4. TT-ACV

Martin Schultz gave the latest news from the Task Team on Atmospheric Chemistry Vocabulary. Vocabulary lists for atmospheric composition terms shall be established and implemented in OSCAR and in support of WMDS (WIGOS MetaData Standard). A draft list of atmospheric compounds is available at https://ontology.geodab.eu/experimental/20151016/atmospheric-compounds, an ontology server running the same software as codes.wmo.int. Besides chemical compounds additional lists will be needed to map the OSCAR variable list, for example a list on physical parametersor a list on the matrix in which the compound ismeasured.
Schultz sent a mail with attached lists about this to the ET-WDC members. 90% should be ready by mid January and a draft for OSCAR must be ready by the end of January. Formal approval is needed from IPET-DRMM (Inter-Programme Expert Team on Data Representation Maintenance and Monitoring). Fiebig warned that he might have to involve the Aerosol SAG. Jörg Klausen suggested that we set up telecons to discuss the review exercise in order to end up with an official reference table and a manual on codes. ().

5. Conclusions from GAW Data Management Meeting at MeteoSwiss in August

A group of about 40 people spent two days discussing data management at the Zürich meeting in mid August. There was strong support for a federated GAW data infrastructure. The existing GAW data centres will play an important role in this context, but also data centres from contributing networks. One of the aims of such a federated data infrastructure is to take work load off users who currently have to submit data to many data centres and who also have to search for data in many different centres. An interesting idea that was proposed at the meeting was that GAWSIS should take up metadata from other networks, also on urban air pollution data. There was consensus that GAW needs a central place and

needs to have a portal to provide this. GAWSIS would definitely be the best candidate for this, but this will represent a huge additional workload. Tony Colavecchia mentioned that there was a lot of talk about contributing networks during the Zürich meeting and that we need a list of contributing networks (post telecom information: this list is available at http://www.wmo.int/pages/prog/arep/gaw/GAW_contr_networks.html; IDAF is not yet on this list). A distinction must be made between those for which an MoU exists and the others. We have to be concrete.
Martin Schultz mentioned that there is a proposal from the SSC to arrange a meeting with potential networks.
Tony Colavecchia reminded the Team that Oksana Tarasova has expressed interest in value-added products. The Team should brainstorm about possible products.

6. GAW Implementation Plan

Contributions to the next round of the IP development are due on Friday 4 December. Documents were sent out some weeks ago with request for input. One of these documents contains new Terms of Reference for GAW central facilities. After editing by the Secretariat, a clean version will be distributed on 8 December. Jörg Klausen was of the opinion that most of what is needed on data management has been provided. ET-WDC will have to be behind what has been written. Jörg Klausen will contact Greg Carmichael and Oksana Tarasova and find out if all needed input has been provided and if more input is needed.

7. Development of work plan for a federated GAW data management infrastructure

Tony Colavecchia gave an introduction on this topic. He presented a series of slides concerning a GAW Ozone and UV Radiation Data Federation with a flow diagram involving various networks and data centres that are active in the field of ozone and UV.
Van Bowersox had a question about versions. Will one get the same version if one downloads the same data set from different data centres? This issue of authoritative source was raised in Zürich and it represents one of the big problems with scattered data centres.

By the time of the next meeting Colavecchia will have examples of CSV and maybe also XML files.
Both data and metdata need exchange formats.

8. Horizon 2020 Call

It was agreed to hold a telecon on 10 December at 1300 UTC in order to discuss how one can get involved in a project under the Horizon 2020 call for a “European data hub of the GEOSS information system”.
Other opportunities for funding within this area might be available from the Research Data Association (RDA) and from WMO’s Resource Mobilisation Department.

2015-12-04/GB

Action items

#	Responsible	Action to take	Deadline / Status
1	Markus Fiebig	Share documents about the ACTRIS-2 data submission portal once available.	ASAP
2	All	Study the Obs4MIPS web site: https://www.earthsystemcog.org/projects/obs4mip	10 Dec 2015
3	Jörg Klausen	Share URL to alpha-2 of new XSD.	4 Dec 2015 / done
4	Martin Schultz	Share the Atmospheric Composition Vocabulary lists with the team.	4 Dec 2015 / done
5	Geir Braathen	Share list of officially accepted contributing net- works with team. The list is posted on the web and can be found here: http://www.wmo.int/pages/prog/arep/gaw/GAW_contr_networks.html	Done
6	All	Show more interest in value-added products.	
7	Jörg Klausen	Contact Greg Carmichael and Oksana Tarasova and extract what is available in the new GAW IP on WDCs and ET-WDC as soon as possible and also find out if additional input from ET-WDC is needed.	4 Dec 2015 / done