

MEETING OF THE
Expert Team on WORLD DATA CENTRES
(JMA, Tokyo, Japan, 21-23 January 2014)

Minutes

Subject:	ET-WDC, Meeting
Dates:	21-23 January 2014
Ort/Raum:	Tokyo
Chair:	Jörg Klausen
Rapporteurs:	Markus Fiebig, Van Bowersox, Tony Colavechia, Jörg Klausen, Martin Schultz
Editorial support:	Lucia Cappelletti
Attendant:	See Doc 1.2 „List of participants“
Excused:	Julian Meyer-Arneke (jma, WDC-RSAT), Vincent- Henri Peuch (vhp, ET-NRT-CDT)

Contents

1	Welcome and Approval of the Agenda, denoting the meeting Rapporteur	3
2	Report by Chair ET-WDC	3
3	Reports from WDCs and GAWSIS	4
3.1	WDCPC.....	4
3.2	WDCA	4
3.3	WRDC	5
3.4	WOUDC	5
3.5	WDCGG	6
3.6	WDC-RSAT	6
3.7	GAWSIS.....	7
4	GAW Station Types	7
5	Data Discovery and Access	7
5.1	Overview of Tokyo GISC, Status of WIS, and WMO Core Metadata Profile v1.3.....	7
5.2	GAW Metadata Profile v1.0	8
5.3	CF standard names for GAW variables	9
5.4	DOI and publication of data	9
5.5	Commercial use of GAW data and products	10
5.6	Review of GAW data policy	10
5.7	Data Access through WDCs	11
5.8	The habitat of the GAW WDCs in the sea of related data centres	11
5.9	N.R.T. data submission and what role the WDCs can/want to play	11
6	Quality and quantity of GAW data at WDCs	12
6.1	Exchange of experience on assessing data and metadata quality at WDCs	12
6.2	Documentation of data and metadata quality at WDCs (→MS)	12
7	Health of Networks (HoN)	12
7.1	Indicators and how ET-WDC can support HoN	12
7.2	How to follow up non-reporting stations?.....	13
8	Future work	13
8.1	Open source approach for development of WDC software	13
8.2	Review Tasks assigned in the GAW Strategic Plan	13
8.3	How can we support the newly added foci of GAW/CAS better?	13
8.4	A common GAW data management portal – a viable vision?	14
9	Review list of action items	14
10	Any other business	14

1 Welcome and Approval of the Agenda, denoting the meeting Rapporteur

Toru Sasaki, director of Atmospheric Environment Division/JMA offered a warm welcome ET-WDC and wished the meeting success.

Jörg Klausen (JK), chair of ET-WDC welcomed the participants and expressed his gratitude and appreciation for the skillful and flawless preparation and organization of the meeting by the staff of JMA.

[Doc 1.1 – Provisional Agenda]

Agenda is adopted by acclamation. Notes were kept by several members in turn, and the draft of the meeting minutes was prepared based on these notes.

[Doc 1.2 – List of Participants]

2 Report by Chair ET-WDC

[Doc 2: Report of the Chair of ET-WDC]

J. Klausen (JK) reminded the participants of the current ET-WDC terms of reference and informed the meeting that Øystein Hov has been elected CAS President at CAS-XVI. Greg Carmichael is the new chair of EPAC SSC (formerly 'JSC OPAG EPAC').

Since the last physical meeting in 2012, ET-WDC has worked through 5 telecons and mail exchange. The frequency and modalities of these virtual meetings were rated as useful and will continue, however, they are no replacement for physical meetings.

The walk-through of tasks in the GAW Strategic Plan involving ET-WDC led to the following remarks:

Task 5.1

- Accessibility of WRDC (and WDCPC) need to be improved.

Task 5.2:

- Submissions to WOUDC have decreased substantially, mix of decreasing networks and delayed submissions.

Task 5.3:

- Need more precise use of language in terms of metadata (discovery vs. descriptive).

Task 5.7:

- Bulletin on Greenhouse Gases that JMA produces is very successful.

Task 5.8:

- see redmine.iek.fz-juelich.de

In conclusion, progress has been made on several accounts, but GAW WDCs are facing big challenges in an ever more complex environment of atmospheric data management. GAW isn't always at the forefront of these developments. In order to be relevant and gain in recognition, WDCs must provide access to quality data, data quality must be documented, and use must be easy. Measures to improve service to the user base and to respond to recommendations by CAS should include

- Improved quality assurance at WDCs
- Documentation and transparency of internal processes
- Archiving raw and near real-time data
- Including urban environments and source regions onto the system
- Introduce / improve traceability
- Ease of data submission, discovery and access across data centers.

JK called for more coordination in the ongoing renewal of GAW data centers, and suggested that the vision of a “virtual” GAW data centre (a portal providing the services of a distributed database) should be further explored.

3 Reports from WDCs and GAWSIS

3.1 WDCPC

Precipitation chemistry assessment report is completed and currently in print:

- 2-3 years periods
- Data set will be archived by WDCPC once paper is out
- Precipitation chemistry observations, dry deposition, gridded mean model estimates

The close coordination between the QA/SAC Americas and WDCPC, including the conduct of biannual laboratory inter-comparison studies (about 90 labs) continues. On the website it is possible to view the results of the studies and keep/update lab contact information. Results are presented as "ring diagrams" (improved to show a 4-level criterion so that labs don't get all those red triangles when that is not really justified) which enable data users to visually evaluate laboratories for individual chemical species and make decisions about the appropriateness of data for their own studies.

- SAG PC doesn't consider WDCPC as a secondary distributor of data
- WDCPC links to various original data centres.

Priorities of WDCPC:

- Post assessment data
- Finalize procedures for receiving data from individual stations
- Complete the links to EBAS.

Comments/questions:

- Hiroshi Koide (HK): In Asia, EANET has role of EBAS in Europe.
- MS: requests that all data regardless of source be made available in a harmonized format, possibly in several standard formats.
- Upcoming GAWTEC course for precipitation chemistry (coordinate between SAG PC and GAWTEC).

→ **Action Item A2014-01: Streamline assessment process of applications of new GAW stations, e.g. by designating a group of experts to handle this process [JK, MS at next SSC meeting].**

3.2 WDCA

- Is working on unifying reporting standards between GAW (more extensive) and EMEP
- Documentation and standardization of the data processing level designations at WDCA:
 - submissions from EU and NOAA (about 70-75% of all submissions) consist of 3 separate files for level 0, 1, 2.
 - At EBAS levels are 0, 1, 1.5, 2 (manually QA'd)
 - MS suggest to adopt this everywhere unless there are reasons against for other types of data
- „Secondary datasets“ means frozen sets of data underlying publications.
- WDCA has plans for including GALION in WDCA under ACTRIS, building an ‘atmospheric data portal’
- Data feedback portal, mantis.nilu.no – issue tracker (data quality issues usually dealt with by email communication)
- WDCA metrics include bar plot showing for each year “data collected ever”, “data collected that year”

Discussion:

- TC called for harmonization and consistency of data metrics.
- JK: "GAW status" of EMEP still unresolved, which creates a problem for adequate treatment in GAW SIS; MoU between EMEP and GAW would clearly help.
- MF: About 40 stations reporting to WDCA, doesn't include most of the EMEP stations; About 73 ground stations. Data sharing MoU between EMEP and GAW hasn't developed over past decade or so – **recommendation to push this issue, potentially on the CAS president level**; crucial from user perspective: if you know that data from a region is in the world data center, but you find out they are incomplete, you will not use the WDC.

- **Recommendation: Submit aerosol measurements performed in Japan to WDCA as well (no station appeared on Markus' maps)**

- **Recommendation: Coordination of vocabulary among SAGs, ET-WDC, and CF**

- Ensuing discussions about "substantial use" and "DOI"
- MS: make web-page design similar on all WDCs, adapted web interface
- Other links:
 - IPPC trends papers using WDCA data
 - ESA-CCI → how to make WDCS aerosol data usable for validating satellite measurements
- Ticketing system for user feed-back appreciated
- WOUDC looking at WDCA for data submission procedures

→ **Action Item A2014-02: Clarify status of submission of Japanese PFR data to WDCA [MF, HK, bi-laterally]**

3.3 WRDC

- WRDC was presented as a supplementary data centre in GAW.
- To use WRDC a registration is necessary.
- WRDC have automatic quality checks for incoming data and report back to data provider.
- Have web-pages with station metadata for each station (http://wrdc.mgo.rssi.ru/wwwroot/Output/country_index.html)
Metadata are available in xml according to WMO recommendations, presently not accessible. A server engine on PHP converts XML to HTML webpage.

Discussion:

- JK: WRDC still maintains the same restrictions on data access, user-friendliness of WRDC needs to be improved → very good work behind the scenes, but very cumbersome to access. What is needed is a much more open display of the holdings of WRDC. Discovery of holdings must be freely accessible and searchable without prior registration, data should be downloadable as files, not displayed as html-encoded tables.

3.4 WOUDC

Renewal:

- based on feedback from users, SAGs, audit by WDCGG and OLMO (Ozone Layer Monitoring Office)/JMA.
- Websites in Canada need to meet web accessibility criteria
- Renewal objectives:
 - Enhance data submission
 - Effective data management
 - Enhanced data access
 - Enhanced web page

- Data submission:
 - look into web accessible folders for data submission (WAF), decommission FTP sites
 - New format for ozone sonde data
 - Migrate in fall 2014
 - Effort on enhancing level 0 submissions
 - New XML submission format
- Data management:
 - Move to relational database instead of file archive
 - Work on automated quality checks together with SAG
 - Data management metrics
- Data access:
 - Use of web services for data access
 - Geospatial web services
 - Discovery metadata (CSW) ... WIS-DCPC
 - Use open source software for conversion into different formats
- Website:
 - Meet international web accessibility and usability standards
 - Software in Python
 - PostgreSQL server
- Taxonomy of WOUDC data
 - Organization
 - ISO Topic category
 - Data category
 - Data type
 - Data set
 -
- Use metrics (how many files are QA'ed, how many contain errors)

Recommendation: ET-WDC should agree on a set of metrics, then implement them

Discussion:

- How was s/w chosen?
 - Requirements based
 - Uses pycsw for web services
- Eizi Toyoda (ET): WMO WIS recommends SRU and OAI-PMH and their implementation by the GISCs. Unfortunately it is too early to rely on a standard set of search indexes in SRU, but the grand trend is to recommend and standardize SRU. What WIS doesn't use is OGC CS-W, which competes with SRU as a search interface. But that doesn't mean a DCPC is discouraged to implement CS-W; it should be a good practice if the user community uses it.

3.5 WDCGG

- GEO AQ CoP metadata guidelines and NILU database structure (P. Eckhardt) were very helpful for restructuring WDCGG database
- HK: need to be careful when hosting two data sets from same site (e.g. continuous and flask samples) – document clearly to users
- Curtain plots of WDCGG data were found to be very useful by RG SAG
- New metadata database currently being established

3.6 WDC-RSAT

No attendance & no material delivered

3.7 GAWSIS

- Enforcement of registration procedures, i.e., new stations are only included if procedures outlined in Addendum to GAW Strategic Plan are followed.
- Challenges:
 - Important networks not included (AERONET, SKYNET, EANET, NADP, IAGOS, DEBITS, ...)
 - Metadata exchange and mapping
 - Controlled vocabularies
 - Migration of GAWSIS to MeteoSwiss
- “GAW performance metrics” = “Network health indicators” planned for renewal of GAWSIS

Discussion:

- MF: Google Analytics may not be best tool for web statistics as it can be blocked using “no script” setting
- MS: Ring diagrams may be useful to indicate station performance in various areas
- WDCs (should) rely on GAWSIS for station metadata, there shouldn't be duplication (and worse: inconsistencies) between station information in GAWSIS and station information in WDCs
- VB: There should be a direct link to data centres relevant for a station near the top of the station description (example investigated: BND in Illinois)

→ **Action Item A2014-03: State the data archive hosting the data next to the data series in the measurement program section of the GAWSIS station report [JK, asap]**

4 GAW Station Types

[Doc 4.1 - Additional station types and need for review of existing assignments (Proposal)]

Proposal to add “local” and “mobile” caused a lot of discussion, because it confuses a statement about operating requirements and a statement about representativeness or station setting. It is accepted that such kind of classification is needed for political reasons, but it is easy to confuse these from the user perspective (in particular, because so far this is the only station classification with controlled vocabulary).

The proposal is endorsed with the following change: rename “station type” to “WMO designation” in Doc 4.1. [Comment JK: done in updated document].

5 Data Discovery and Access

5.1 Overview of Tokyo GISC, Status of WIS, and WMO Core Metadata Profile v1.3

Eizi Toyoda, co-chair of IPET-MDRD, presented WIS and the current implementation of the GISC Tokyo. He then explained the current version of the WMO Core Metadata Profile.

Discussion:

- MF raised a question about the WIS requirements on WMO Core Profile to ISO 19115/19139. Toyoda responded that it is mandatory for every WIS metadata record to follow ISO 19115/19139, but in real practice we can find metadata records not complying to WMO Core Profile.
- Tony Colavecchia (TC) asked about the level of granularity needed for data discovery. In particular, the costs of higher levels of granularity should be weighed against user benefits. It would be helpful to define the level of granularity necessary to find a data set without going to the level necessary to identify individual data points within that data set. The broader question was how the GISC, GAWSIS, and data centres coordinate their metadata without duplicate efforts. What level of granularity is needed at each institution? How best to serve the user community is the ultimate question.
- There exists some uncertainty about the definition of a data series versus a data set.

- There was some discussion re the redistribution of data and the risks in having a secondary distributor, such as a data centre, distributing data from a provider.
- Inter-GISC metadata traffic has to be OAI-PMH compliant
- WIS: catalogue of distributed data vs WIGOS: catalogue of stations, observations, ...
- ISO19139 contains
 - 620 elements
 - 255 complex-types (that contain several elements)
 - 26 attributes
 - Communities need to specify profiles, i.e., subsets
- **File names for WIS meta data: suggests to use the variable part (after ::) with extension .xml**
- Granularity of metadata
 - Dataseries is a collection of datasets
- Schemacentral.com
- Data policy
 - WMOEssential: unrestricted
 - WMOAdditional: restricted by Res 25 and Res 40
 - WMOOther: example METAR under jurisdiction of ICAO

5.2 GAW Metadata Profile v1.0

[Doc 5.2.1 – GAW Metadata Profile v1.0 (Specification document)]

JK presented the current version of the GAW Metadata Profile, which is based on ISO19115/19139 and their WMO Core Profile. The corresponding document was approved with revisions (partly in the document, partly in the implementation) as follows:

- reduce geographic extent to boundaries of data set (point for fixed stations, area for mobile stations)
- add examples to the table listing the GAW metadata profile.
- Toyoda mentioned there are two revisions of ISO 19115: the first one is ISO 19115:2003 and the second one is ISO 19115-1:2014. Currently the 2003 revision is recommended in WIS; one reason is WMO Core Profile is based on that; another reason is the XML realization (ISO 19139) is published only for the first revision
- Verify correct implementation of fileIdentifier (for the character string, the format begins urn:x-wmo.... NOT urn_x-wmo...)
- Toyoda mentioned that it is not appropriate to use the URL of XML codelist files (for example <http://wis.wmo.int/2012/codlists/WMOCodeLists.xml>) for gmd:language. He introduced the practice in INSPIRE profile. [Later he notified the team that the CBS Task Team on Application of Metadata is discussing in a direction that we can simply use <gco:CharacterString> with only ISO language code (without URL) in <gmd:language>.]
- It would be useful to add a field for the inlet height of the instrument.

➔ **Action Item A2014-04: Revise the draft GAW metadata profile according to recommendations [JK, asap]**

[Doc 5.2.2 – Additional item(s) in WMO Core Metadata Profile v1.3 code tables (Proposal)]

The ET-WDC endorses the recommendation to add “atmosphericComposition” to the WMO Category Code list. However, the team concluded that there was no need to add another term for radiation, as the GAW observations can be adequately represented with the existing term “actinometry”. No further action is required, because IPET-MDRD has already taken note of the request to add “atmospheric-Composition”.

5.2.1 WIGOS Metadata Standard (ad-hoc agenda item)

JK presented the draft v0.0.17 of the WIGOS semantic metadata standard elaborated by TT-WMD in detail. The standard consists of 10 categories: observed quantity, purpose of observation, data quality, environment, data processing, sampling and analysis, station/platform, instrument, ownership and data policy, contact. Overall, ET-WDC supported the current standard as sufficiently comprehensive to serve the needs of GAW. Eizi Toyoda admitted that there is no overarching governance structure between WIS and WIGOS for now.

Recommendations by ET-WDC:

- /3/ data quality: should add an explicit (controlled) label to specify the QA procedures applied (e.g. “WMO QA/QC”, “NOAA”, “informal”...)
- /5/ data processing: “spatial reporting interval” should be conditional (only if observation spans spatial range = moving platform, radar, etc.); “aggregation interval” needs to be split into “temporal aggregation interval” and “spatial aggregation interval” (this needs clarification how to proceed for example with aircraft data where time and space are related).
- /4/ Environment
 - 4-02 Land cover classification scheme: need to include version number (e.g. for MODIS-based schemes) – see MACC-GFAS or EAS Fire CCI analysis
 - 4-03 topography: suggested to spell this out more explicitly for the implementation – the full text examples may be misleading people to implement this in full text rather than specified and standardized tags.

5.3 CF standard names for GAW variables

[Doc 5.3 – CF standard names for GAW variables (Proposal)]

JK introduced the table and pointed out the need to verify, complete, and formally register the names with CF. It was agreed that each data centre representative was asked to review the list of CF names and make necessary corrections.

- Tony Colavecchia: Meteo, Radiation-UV, Ozone
- Anatoly: Radiation- IR, Radiation – solar
- Hiroshi: greenhouse gases, reactive gases
- Van: Precip Chem
- Markus: aerosol, POPS

As MF was found to have the most experience with actually submitting standard names, he was requested to assist the other WDCs if need be.

➔ **Action Item A2014-05:** Validate links between parameters and CF standard names, and submit any changes to JK for inclusion in the GAWSIS name table [All WDCs, by 31 March 2014]

➔ **Action Item A2014-06:** Submit any new names to CF [All WDCs concerned, by 31 March 2014]

5.4 DOI and publication of data

In the absence of Julian Meyer-Arnek and for lack of a document to discuss, there was an informal exchange of information on the subject only.

- DOI policy requires only that DOI is referenced but that there are no other constraints on data use (such as contacting originators). Using DOIs thus appears to be in conflict with the GAW data policy, in particular that “providers be contacted when there is ‘substantial use’ of their data.”
- WDCA is considering DOI-on-demand (i.e., upon request of the data provider who wish to follow the DOI policies) for data submitted
- Annual costs for WDCA would be roughly 6000 €
- DOI always retrospective, and it is unclear how one could deal with growing time series.
- Lifetime of DOI only guaranteed as long as the annual fee is paid.

The meeting concluded that more information and an analysis of the pros and cons were needed before a recommendation on this subject could be made. An important constraint on the use of DOIs for WDC data is that the WDCs manage time series that are periodically updated; it is our understanding that a new DOI would be issued with each update. WDCs are free to offer DOIs, if specifically requested by a data provider, for specific purposes such as publication of the data in a data journal.

→ **Action Item A2014-07: Prepare doc 5.4 and submit for discussion to ET-WDC [JMA, asap]**

5.5 Commercial use of GAW data and products

[Doc 5.5 – Commercial use of GAW data and products (Position Paper)]

TC presented some background information as well as a number of examples surrounding commercial interests in GAW data. He pointed out that the Annexes to Res 40 cover commercial activities and have been around for decades for WWW. He also mentioned that NOAA has a cover letter, with a link to Res 40. For any use (commercial or non-commercial), proper acknowledgement was a minimum requirement and DOI was mentioned again as a possibility. No consensus was found on how to approach the issue, and the chair of ET-WDC was asked to defer a discussion on the matter to EPAC SSC.

→ **Action Item A2014-08: Obtain guidance on how to address requests for commercial use of data from EPAC SSC [JK, next meeting]**

5.6 Review of GAW data policy

[Doc 5.6 – GAW data policy: fit for purpose? (Discussion Paper)]

The ET-WDC recognized that there are differences in the data policy statements across the WDCs and the ET-WDC supported the need for a uniform policy. It was agreed that each of the WDCs should post the GAW data-use policy statement, verbatim, on their respective Websites. Issues that require further resolution include:

- (1) acknowledgement of providers needs to be clarified,
- (2) constraints imposed by the statement that data be used for ‘non-commercial’ purposes,
- (3) expansion of the meaning of ‘substantial use’ of the data,
- (4) the existing GAW data-use policy should be clarified for conditions of use, (a policy for NRT data needs to be clarified).

→ **Action Item A2014-09: All WDCs are required to post the current GAW data policy statement, unaltered, on their respective web sites [All WDCs and GAWSIS, 28 February 2014]**

It was also proposed to amend the official data use policy with a standard phrase that should be used in the acknowledgement sections of scientific publications.

- **Action Item A2014-10:** Draft an amendment to the GAW data policy requiring the quotation of a data acknowledgment statement for use in publications and request SSC EPAC to endorse it. [JK/MS, before next meeting of SSC EPAC]

5.7 Data Access through WDCs

5.7.1 Web Services at WOUDC

- Lots of support...Data policy concerns
- Machine-to-machine
- EC pulls from 5-6 sources already

Discussion:

- How do data policies survive?
 - Example Juelich: `join.iwk.fz-juelich.de`
- How prevent secondary access to cached data?
 - Already possible, can download data

5.7.2 Data and Metadata Access through WRDC

See statements under 3.3. The topic was not further approached for lack of time.

5.7.3 WDCGG Reform

The topic was not further approached for lack of time.

5.8 The habitat of the GAW WDCs in the sea of related data centres

[Doc 5.8 – Roles and responsibilities, existing collaborations, data flows, ... (Information paper)]

The document was presented and discussed. Several parts were still missing and need to be provided.

Discussion:

- Options: links to related data centres (requires full interoperability), ingesting data from related data centres into WDC, single portal into related data centres.
- Ingesting data from other data centres comes with a variety of risks (e.g., NDACC ingested in WOUDC, WDCGG receiving European data through EBAS)

- **Action Item A2014-11:** Provide input for empty sections in document [WDCPC, WDCA, WRDC, 31 March 2014].

5.9 N.R.T. data submission and what role the WDCs can/want to play

[Doc 5.9 – Existing n.r.t. data streams of relevance for GAW (Information paper)]

Unfortunately, no document was available to support the discussion of this topic. It was however acknowledged that the subject is of great relevance. Any in-depth consideration of the (future) role of

WDCs with regards to N.R.T. data was deferred. A brief round-table resulted in the following statements:

- WDCA involved with about 45 stations, 5 different instrument types, maybe extended with 2 more instruments; access restricted with log-in approved by SAG.
- WDCGG not involved; issue handled by MACC with DWD hosting an ftp site; MACC will be operational in 2015; IGIS also includes scope of NRT;
- WDCGG willing to explore n.r.t. data activity but resource constraints
- WOUDC not highest priority on the short term; longer- term yes

For most of the WDCs, serving the n.r.t. data needs of users is presently out of scope and deferred to MACC or other, rather ad-hoc, arrangements. However, there is a general interest in getting involved more directly, not least in order to improve the submission of fully quality-controlled data to the WDCs. It was agreed that improved user friendliness involves at least the provision of format conversion tools, both for submission but also retrieval of data.

→ **Action Item A2014-12:** Create a mechanism for users to share conversion software [MS and JK]

6 Quality and quantity of GAW data at WDCs

6.1 Exchange of experience on assessing data and metadata quality at WDCs

[Doc 6.1 – Existing QA/QC procedures at WDCs (Information paper)]

VB walked the group through the data quality assurance objectives, data quality objectives (guidelines), and why we need metadata on data quality

- meet WIGOS requirement and WDC terms of reference.
- Need to provide a high level summary of QA/QC for the lifecycle of the data and details of QA/QC procedures at each of the data centre

→ **Action Item A2014-13:** Develop a GAW Data Quality Report by end of 2014 [JK to coordinate, all WDCs to contribute]. First milestone: Develop an outline based on doc 6.1, 6.2, 5.8 [JK, June 2014]

6.2 Documentation of data and metadata quality at WDCs (→MS)

Harmonization of data flagging schemes for atmospheric composition observations

[Doc 6.2 – Data flagging schemes in use in GAW (Information paper)]

- Qualifier (quality), Selection (selection criteria) , Information (additional information) flag
- Data Quality flags...consistency

→ **Action Item A2014-14:** Develop a consistent data quality assessment scheme based on WOUDC approach [MS and TC to coordinate, all WDCs to contribute, September 2014]

7 Health of Networks (HoN)

7.1 Indicators and how ET-WDC can support HoN

→ **Action Item A2014-15:** WDC managers to share their list of metrics, ideally supported by charts, screenshots, documentation available [HK to coordinate and compile, September 2014]

7.2 How to follow up non-reporting stations?

A number of stations are not reporting in a timely fashion or no longer reporting. The meeting shared the following observations:

- WDCA: WDCA: a few usual suspects that simply don't reply to requests to submit
- WDCGG: CONTRAIL submit only limited data because of policy issues; GHG works pretty well, RG perhaps less so
- WOUDC: 10 stations report within 1 year, 17 within 2 years, 61 within 3 years, 20 never
- WDCPC: handling 12 datasets annually
- WRDC: receive data from 450 stations each year, not all GAW stations; upload data quarterly; also has problem with non-reporting station

Actions taken in the past include

- Annual reminder
- Engage SAG members with personal relationships
- Engage WMO Secretariat to send out letters to specific culprits

Suggestions for improvements were limited to reiterating the need for personal engagement through personal contact and the notion that individual letters proved to be the only way to go. A coordination of such letters was not found to be too helpful, however, there was unanimous agreement that the following was a meaningful escalation path:

1. Letter from the data centre to the PI for an observation,
2. Personal engagement of members of the relevant SAG,
3. Official letter from the Secretariat to the PR

Geir Braathen (GB) representing the Secretariat confirmed the commitment by the Secretariat to follow up if needed.

→ **Action Item A2014-16:** Confirm with Secretariat and inform ET-WDC how the Secretariat can provide support in chasing silent stations to submit their data [GB, April 2014]

→ **Action Item A2014-17:** Confirm with the Secretariat the process of maintaining GAW country contacts in GAWSIS [JK, asap]

→ **Action Item A2014-18:** Nominate outstanding contributors for follow-up (with candy) by GAW Secretariat and EPAC SSC chair [All WDCs to send nominations to JK, by April 2014]

8 Future work

8.1 Open source approach for development of WDC software

[Doc 8.1 – Open source approach for development of WDC software (Vision paper)]

The suggestion by TC of using a common set of open source tools by the WDCs was welcomed in principle. However, it was found that there were too many different corporate ICT policies / restrictions for this to be a viable approach. It was agreed to focus instead on developing an interoperability strategy for GAW.

8.2 Review Tasks assigned in the GAW Strategic Plan

See Doc 2 and section 2 in this report for more on this.

8.3 How can we support the newly added foci of GAW/CAS better?

The general consensus of the meeting was that ET-WDC contributes by facilitating exchange of information, sharing of good practice, and striving for interoperability of the archives. It was further agreed

that, even if substantial progress has already been made in the past, improving the data quality assurance procedures at the WDCs to ensure high quality of the data made available and document data quality was necessary to raise the profile of the GAW WDCs. Finally, it was acknowledged that more discussion is needed to develop strategies on how to reach out to the modeling community, on how to integrate data from regional networks and large urban complexes, and to be more present in the near-real-time data management.

8.4 A common GAW data management portal – a viable vision?

The vision with single point of access and single point of contribution along the lines of the figure in the presentation was endorsed as viable in the long-term. MS presented a web-based tool (see: <http://join.iwk.fz-juelich.de>) that supports discovery, access and retrieval from a number of data sources and was developed in the context of the GEO AQ CoP.

- **Action Item A2014-19:** Describe the vision in a short paragraph to define the scope [JK, asap]
- **Action Item A2014-20:** Develop the strategy report (internal) as input for scoping our contributions to the next GAW Implementation Plan [TC, with contributions by all, December 2014]

9 Review list of action items

It was acknowledged by all that, while the will to engage in the agreed action items exists, the resources to also act upon them are limited. A number of action items require relatively little effort, and those should be followed up upon in due course. The meeting further agreed to work on the larger action items with the following priorities:

1. Data quality and procedures → see action item **A2014-13**.
2. Interoperability → see action items **A2014-19** and **A2014-20**.

10 Any other business

- **Action Item A2014-21:** Schedule next meeting and teleconferences [JK, asap]

Meeting closed at 16:01 (sharp) on Thursday 23 January 2014.