

**Minutes of the Meeting of the WMO Expert Team on GAW World Data Centres (ET-GAW WDC)**

**2<sup>nd</sup> - 3<sup>rd</sup> October 2008**

**A.Voeikov Main Geophysical Observatory, St Petersburg, Russia.**

**Participants**

|                  |           |  |                     |
|------------------|-----------|--|---------------------|
| Jörg Klausen     | <b>JK</b> | (GAWSIS – EMPA Dübendorf)                        | <i>Chair ET-WDC</i> |
| Anatoly Tsvetkov | <b>AT</b> | (WRDC – MGO St. Petersburg)                      |                     |
| Julian Wilson    | <b>JW</b> | (WDCA – JRC, Ispra)                              | <i>Rapporteur</i>   |
| Kazuto Suda      | <b>KS</b> | (WDCGG – JMA, Tokyo)                             |                     |
| Kathrin Höppner  | <b>KH</b> | (WDC-RSAT, DLR, Oberpfaffenhofen)                |                     |
| Michael Bittner  | <b>MB</b> | (WDC-RSAT, DLR, Oberpfaffenhofen)                |                     |
| Ed Hare          | <b>EH</b> | (WOUDC – Environment Canada, Toronto) – by phone |                     |
| Geir Braathen    | <b>GB</b> | (ARE/AER – WMO, Geneva)                          |                     |
| Kjetil Torseth   | <b>KT</b> | (EMEP – NILU, Kjeller, Norway)                   |                     |

**Excused**

|                     |                            |
|---------------------|----------------------------|
| Tim Coleman         | (WDCPC, SUNY Albany)       |
| Vincent-Henri Peuch | (ET-NRT-CDD, Meteo France) |
| Kenneth Masarie     | (ESRL, NOAA, Boulder CO)   |
| Steve Wilcox        | (NREL, Golden CO)          |

| Agenda Item   | Action Item |
|---|-------------|
| <p><b>1</b> Meeting opened by <b>Dr. Sergey Chicherin</b>, deputy director for research, who welcomed participants and wished the meeting every success. Dr. Chicherin mentioned that MGO is responsible for surface observations of meteorological parameters and atmospheric trace constituents, precipitation chemistry, applied meteorology and climate modeling etc. MGO has been running for 160 years. Dr. Chicherin said that the old AOD network (broadband filters) in Russia had suffered and that they were looking for new funding and modern equipment. He also mentioned MGO's plan to set-up a new global station in the Caucausus.</p> |             |
| <p><b>2</b> <b>JK</b>, chairing the meeting thank <b>AT</b> for the invitation and all his work in getting the meeting organized and expresses his satisfaction that the meeting can finally take place in SPb after many attempts. He nominates <b>JW</b> as rapporteur, proven successful and productive in the past. <b>JW</b> is unanimously elected.</p> <p><b>JK</b> conveys apologies from Volker Mohnen on behalf of WDCPC, who sincerely regrets that WDCPC was unable to be represented at this meeting,</p>  |             |

| Agenda Item  | Action Item |
|--|-------------|
| <p>and who expressed his hopes that – under a new funding scheme – WDCPC can be re-established.</p>  |             |
| <p><b>3 WRDC (AT)</b></p> <p>Reviewed operation, data collected, QA procedures</p> <p>Archive of 1118 stations that have provided data over 44 years; at present, about 400 stations regularly submitting data.</p> <p>Data are shown on the web after log-in (user account required for the conventional data), can be downloaded by cut-and-paste of html tables, also some graphics available. Current procedure is: go to web site, register, look at map and (Java) data interface, make request, receive the data.</p> <p>Differentiate between GAW and conventional stations, 402 GAW sites communicated to GAWSIS, but only 23 GAW sites are accessible via the web.</p> <p><i>Discussion:</i></p> <p>Priorities for WRDC must be to make ASCII files available. The present possibilities for users to access data are not acceptable. AT could make these data available within a couple of weeks, however the format for these file is to be decided (suggestions: CREX or one of the existing world data centre formats.)</p> <p>WRDC mentions that they face a problem regarding conventional data – in particular, it has not been decided how to deal with correction to world radiometric standard in 1981. Many NMHSs have not indicated if/when they changed to the new standard, also many other data quality issues. WRDC prefers to only publish good quality data and is therefore reluctant to make these data available. JK suggests to make all data available with caveats – disclaimers asking for feedback – because open access to data is prerequisite for users to be able to comment on the data.</p> <p>Separation of GAW and non GAW sites is considered very confusing and not helpful.</p> <p><b>WOUDC (EH) – presenting his slides via phone</b></p> <p>Working closely with IGACO Ozone and SAG Ozone as well as SAG UV</p> <p>Data usage disclaimer has been added to all data files</p> <p>Completion of ‘data passport’ (a comprehensive metadata collection for each data set) and calibration histories are a primary focus of WOUDC</p> <p>Suggests co-ordination on data delivery standards across networks</p> <p><i>Discussion:</i></p> <p>The meeting takes note of Ed’s presentation and thanks him for joining via telephone. Unfortunately, technical difficulties make a meaningful discussion impossible, and JK suggests that participants study the slides again in detail.</p> |             |

| Agenda Item  | Action Item |
|--|-------------|
| <p><b>WDCGG (KS)</b></p> <p>New web interface for searching and retrieving data and metadata at <a href="http://gaw.kishou.go.jp/cgi-bin/wdogg/search_form.cgi">http://gaw.kishou.go.jp/cgi-bin/wdogg/search_form.cgi</a> now available.</p> <p>Leaflet describing data centre issued, as well as a data submission guide (WMO GAW Report No. 174)</p> <p>Lead time for reporting is typically 1 year with an autumn maximum in data sets submitted.</p> <p>Usefulness of calibration / traceability info as metadata is stressed.</p> <p>Demand for NRT data was suggested by SAG GG.</p> <p>Data submission increasingly also from ships and mobile platforms.</p> <p>WDCGG is currently negotiating for a possible mirror for GOSAT satellite data.</p> <p><i>Discussion:</i></p> <p>WDCGG reported analyses of CO2 using model and observations as a JMA product. There are other similar products, in particular NOAA's Carbon Tracker. It is noted that the SAG GG is to advise how and where to coordinate the various products to give a best GAW product.</p> <p><b>WDCA (JW)</b></p> <p>WDCA is forced to change its simple web address to something more complicated due to the fact that it is hosted by an EC institution.</p> <p>Progress was made on the European data/metadata in collaboration with NILU, on the global data through collaboration with ESRL and Natchem (Environment Canada). The internal consistency of the data sets could be improved through these efforts. The WDCA metadata format (NARSTO DES) requires more metadata items than the NASA-AMES 1001 format used at NILU. These metadata have in the past be added manually. The submission of the Natchem data sets has significantly enlarged the WDCA data holdings, these data are presently being incorporated.</p> <p>An AOD data portal is presently being developed.</p> <p>WDCA is establishing a relational database to manage the metadata. This will also significantly facilitate the exchange of metadata for GAWSIS and is expected to be completed by end of October 2008.</p> <p><b>WDC-RSAT (KH)</b></p> <p>WDC-RSAT is hosted by the Applied Remote Sensing Cluster of DLR = DFD &amp; IMF. An ICSU WDC since 2003.</p> <p>German ICSU-WDC cluster combines WDC Climate, WDC Mare, WDC-Terra</p> |             |

| Agenda Item  | Action Item |
|--|-------------|
| <p>(candidate) and WDC-RSAT and shares the C3-grid architecture. Collectively, through the Technical Library of Germany, they also assign DOIs to data sets.</p> <p>Data management centre for the Network for the Detection of Mesopause Change (NDMC) – affiliated to GAW &amp; NDACC.</p> <p>WDC-RSAT has special relationships with and is responsible for the data management for the Environmental Research Station Schneefernerhaus.</p> <p>The MoU between WMO and DLR that makes WDC-RSAT a WMO GAW WDC is expected to be signed soon.</p> <p>Current activities include the development of a strategy plan for WDC-RSAT, to manage implementation as a WMO-WDC, develop and maintain a portal both for in-house and external sources, communication, publicize, assist WMO, participate in SAGs Ozone &amp; Aerosols, and to support the long term goals as per GAW and IGACO. First phase of implementation focuses on the GAW-parameter ozone and aerosol.</p> <p>WDC-RSAT requests ET-GAW WDC to review the draft of the strategic plan (v0). This document will also be iterated with SAGs Aerosol and Ozone and with the GAW Secretariat.</p> <p><b>GAWSIS (JK)</b></p> <p>Mechanism for automatic metadata flat file ingestion implemented for everyone except WDCPC; for WDCA, the mechanics are implemented, but the flat file is very much incomplete.</p> <p>The system to serve as a clearing house for GAW IDs has been implemented and is being used. At present, 289 out of 788 stations have an ID.</p> <p>Also, bibliographic references can now be managed by GAWSIS. This feature is primarily intended to help locate ‘grey’ literature (audit reports, annual reports, etc.) but can equally include journal paper references.</p> <p>A preliminary XML version of GAWSIS has been established as a test bed for producing ISO/WIS-compliant metadata representations.</p> <p>GAWSIS has been migrated to a new server. The new URL is <a href="http://gaw.empa.ch/gawsis">http://gaw.empa.ch/gawsis</a></p> <p>The representation of QA information is still weak in GAWSIS. This needs improvement and possibly also some interaction with ISO/TC211 when it comes to representing such information in an ISO-compliant format.</p> <p>In conclusion, it is asserted that both human and machine interactive discovery and retrieval systems are wanted but at low (or zero) cost to the users.</p> <p><b>NILU EBAS &amp; GEOMON (KT)</b></p> <p>GEOMON will have a GUI linking EBAS (the EMEP database, mainly time series information) and CDB (cal/val campaign database for ENVISAT/ESA;</p> | <p>1</p>    |

| Agenda Item  | Action Item  |
|--|--|
| <p>more complex data – HDF data with relational db index; restricted access for external user groups.)</p> <p>Data centres serve very many different purposes and must meet very many different requirements:</p> <ul style="list-style-type: none"> <li>- For funding agency: cost, compliance acknowledgement</li> <li>- For provider: control, acknowledgement and little effort</li> <li>- Data centre: maintain links with data originator, visibility, long term view</li> <li>- Data user: low effort, value added, no thought required.</li> </ul> <p>EBAS GUI data disclaimer: public data (WMO type acknowledgement) and restricted data where input from the user is required – either general request or specific request.</p> <p><i>Discussion:</i></p> <p>EBAS is organized by project which is not helpful for the user, and as a consequence, GAW (often not the primary funding source but extremely important as a structuring element) is rather invisible.</p> <p>EUSAAR is due to review the GUI. ET-WDC concerned about visibility of GAW in EBAS EUSAAR interface.</p> <p>The discussion addresses the role of NILU within the European atmospheric monitoring programs. Basically, NILU could act as the primary entry point for data submissions for all European GAW data (this is now so for aerosol data). Additionally, NILU – as the host of the EMEP database – would have a natural role as a regional data centre for precipitation chemistry and could serve as a regional node for a new (distributed) WDCPC. This needs to be discussed further, and WMO needs to give guidance on this. TFMM is the right environment for this to be discussed, WMO to ensure that all GAW member agencies not presently in TFMM are invited to a joint session of TFMM that seeks to resolve and define the link between EMEP and WMO.</p> | <p style="text-align: center;">2</p> <p style="text-align: center;">3</p> <p style="text-align: center;">4</p> |
| <p><b>4 WIGOS and WIS (GB)</b></p> <p>Objective is to improve data management, transfer, retrieval etc.</p> <p>WIGOS/WIS implementation culmination at 2011 Congress – now until then, portions of the system are developed and tested through pilot and demonstration projects.</p> <p>1<sup>st</sup> level integration standardization instruments and measurements within constituent networks</p> <p>2<sup>nd</sup> level standardised for communication and data exchange; common information infrastructure (WIS).</p> <p>3<sup>rd</sup> level standardized products, QA etc.</p> <p>WIGOS/WIS will be components of GEOSS.</p> <p>Core are a number of Global Information System Centres (GISCs) that will permit open data searching, but not necessarily open data access. Metadata</p>   |  |

| Agenda Item  | Action Item |
|--|-------------|
| <p>are held and shared among the GISCs</p> <p>Next level out are DCPCs = data collection/product centre: WDCs &amp; GAWSIS are considered as DCPCs under several conditions including compatibility with ISO19115.</p> <p>WIS builds upon and will gradually replace the present GTS.</p> <p>WIGOS/WIS requires metadata standards and has decided to comply with ISO19115. IPET-MI is responsible for developing a WMO metadata profile in compliance with ISO19115. INSPIRE is the European implementation of ISO19115.</p> <p>ISO23950 is the machine search metadata standard.</p> <p><i>Discussion:</i></p> <p>First operational GISC (the VGISC, a partnership of Germany, France and UK) includes ECMWF and should become operational in 2009.</p> <p>WIS does not provide an alternative to Internet or GEANT for communication – it is a label for existing efforts.</p> <p>HTAP uses the CF Standard for metadata.</p> <p>ISO19115 dictionary formation is not straightforward. ET-WDC and IPET-MI need to communicate to develop the necessary catalogues required to represent GAW relevant information.</p> <p>At the first meeting of the ad-hoc CBS CAS Expert Team on the Implementation of WIGOS, a number of joint GOS-GAW pilot projects were discussed. With regards to GAW, it was decided to go ahead with a pilot project that will implement near-real-time delivery of ozone and AOD data. This pilot is the responsibility of the GAW ET-NRT-CDD and others (e.g. SAG-Ozone, SAG-Aerosol, CBS bodies).</p> <p>ET-WDC is encouraged by the ad-hoc CBS CAS Expert Team (see Minutes of that meeting; Doc_5.1 of this meeting) to develop a pilot project focusing on interoperability of the WDCs: Doc_5.2 of this meeting provides a draft with data centres to express metadata in ISO19115 compliant form and create a portal giving access to data of a given station archived at different WDCs.</p> <p>Funding of this effort is an issue. Limited funding should be available through the WIGOS trust fund.</p> |             |
| <p><b>5 WDC Reviews (JK)</b></p> <p>The quality of the reviews varies. The WDCs are encouraged to clarify open questions bilaterally with their respective reviewer. Overall the data centres received fairly good grades from each other. We are weakest in areas 4 and 5 (value added products and ancillary QA data.)</p> <p><b>WOUDC:</b> Minimum that flags descriptions should be included in the core data file. Issue of the right amount of meta-data to include in the file – data passport.</p>   |             |

| Agenda Item  | Action Item   |
|--|---|
| <p><b>WDC-RSAT:</b> A few issues with the interface contact &amp; how to submit data. Overall very large site but also offers access to a lot of data.</p> <p><b>WDCGG:</b> Separation of ancillary met data from measurement data has been done but these data are still not discoverable and not part of the metadata sent to GAWSIS – they should simply be a separate category in its own right. Perennial issue of how to engage with the data users and providers. Link with providers has become stronger, however, some data users still have to be fully convinced of the value of WDCGG.</p> <p><b>WRDC:</b> Difficult to access data due to the Java applet used to run the interface. Web-site much faster but data access insufficient. Registration is an issue – have to do so before one can even find out what is available. GAW and non GAW site/data distinction is an issue – a bit like the EBAS project variable.</p> <p><i>Discussion:</i> Data catalogue for long lived archive like WRDC involves name changes, country changes etc. It is unclear to WRDC how to deal with this. The recommendation is to use an ID as a persistent attribute, but it is recognized that expert judgement is required to decide whether a station can remain the same if it is re-located (that depends on a lot of factors). WRDC is advised to seek the advice of the other WDCs and GAWSIS if in doubt. Regarding the functionality of the web site and the necessary improvements for access to data, the recommendations of the group are clearly towards open, easy access, most easily implemented by providing ftp access to ASCII versions of the data sets. These can be programmatically generated from the binary archives that are maintained at WRDC.</p> <p><b>GAWSIS:</b> For lack of time in the plenary, this audit was discussed bilaterally.</p> <p><b>WDCA:</b> For lack of time in the plenary, this audit was discussed bilaterally.</p> <p><b>WDCPC:</b> Due to the current state of WDCPC, this audit could not be completed in a meaningful way.</p> <p><i>Discussion:</i></p> <p>GAW IDs are in use and assigned by GAWSIS, WDCGG and NOAA, as well as individual stations. JW will implement the new GAWID. GAWSIS maintains the catalogue of requested IDs.</p> <p>All WDCs are encouraged to implement recommendations and to work bilaterally with their respective reviewer to clarify open questions.</p> | <p style="text-align: center;">5</p> <p style="text-align: center;">6</p> |
| <p><b>6</b> The review of action items from previous meeting was skipped due to lack of time. The group is encouraged to re-visit the minutes of the 2007 meeting and to complete outstanding actions</p>  | <p style="text-align: center;">7</p>                                      |
| <p><b>7 WIGOS/WIS pilot project (JK)</b></p> <p>Joint CAS CBS (commission basic systems) effort, but CBS is the dominant partner and would have to be the funding party.</p> <p>The proposed pilot project out of ET-WDC will focus on improving</p>   |   |

| Agenda Item  | Action Item |
|--|-------------|
| <p>interoperability of data centres and the development of a prototype data access portal.</p> <p>Pilot projects are specified by Congress and EC in several areas to speed up implementation of WIGOS/WIS.</p> <p>Rationale behind this proposal is to bring data centres forward and closer to WIS in order to maintain importance within WMO and beyond.</p> <p><i>Discussion:</i></p> <p>How does DOI relate to ISO19115 – former was developed by libraries and is commercial, latter was developed by the geographical community and use is free. DOIs represent a citable reference, ISO-compliant metadata are a different concept.</p> <p>INSPIRE? INSPIRE is a directive (=legislation) that prescribes compliance with ISO19115 for the description of geographical information.</p> <p>WMO and WIGOS trust fund is called upon to contribute to training and manpower costs of executing the project, not merely travel funds. We estimate circa 1 man-year of effort required from WIGOS to be divided across the ET-WDC members.</p> <p>Concretely, task involved in realizing this pilot project is to provide an XML representation of every dataset in a directory somewhere.</p> <p>The individual members of ET-WDC were then asked to cast a vote regarding the proposed pilot project (doc 5.2):</p> <ul style="list-style-type: none"> <li>- WDCA: favourable subject to manpower constraints, investigate support from INSPIRE unit.</li> <li>- WDCGG: basically in favour, but concerned about how the proposed data portal would operate in reality, and visibility, issues of metadata handling particularly in terms of data quality JK responds that the idea would be to include a full set of metadata with each request.</li> <li>- WDC-RSAT: yes, already doing so, would also like to include the initiative between Schneefernhaus and WDC-RSAT as that already involves merging all their metadata in ISO compatible form.</li> <li>- WOUDC (JK speaks on behalf of Ed Hare who could not be consulted during the meeting): favourable, but concerned that the pilot projects are quite ambitious and that resources will be needed to implement them.</li> <li>- WRDC: unsure, would definitely need training, advice, and some man power</li> </ul> | 8           |
| <p><b>8</b> KT offered to host one of the next meetings at NILU.</p> <p>The next meeting is foreseen as a side-meeting (at most 1 day) during GAW2009 (tentatively beginning of May 2009)</p>  |             |
| <p>The meeting closed Friday, 3 October 2008 at 16:59 LST</p>  |             |



## List of Action Items

| # | Action Item   | Responsible   | Due Date    |
|---|---|---|-------------|
| 1 | WDC-RSAT to identify points requiring detailed comments from ET-EDC in their strategic plan.  | KH, MB  | 15 Oct 2008 |
|   | Review WDC-RSAT strategic plan and provide general comments   | All   | 31 Oct 2008 |
| 2 | ET-WDC tell SAG Aerosol of concern about lack of visibility of GAW in the 'EUSAAR' EBAS GUI and ask that this be looked at during the forthcoming review of the GUI by EUSAAR as several SAG members will be involved in this.  | JW, JK  | 15 Oct 2008 |
| 3 | Ask WMO as co-chair of TFMM to organize joint EMEP-GAW RA VI session (May 2009) to analyze current situation of GAW and EMEP data flows, overlaps and differences, in Europe; and to define – inter alia – a consistent model for data flow for EMEP and GAW on a country by country and GAW program basis. | GB, JK  | 15 Nov 2008 |
| 4 | Produce a draft plan for the data flow in #3 and define the role of regional data centre as straw-man for the session.  | KT  | 31 Mar 2009 |
|   | Put this on the agenda for the next meetings of SAG Aerosol, GG, and PC, and JSC.   | JW (re: SAG Aerosol),<br>KS (re: SAG GG),<br>JK (re: SAG PC, JSC) |             |
|   | Solicit review by ET-WDC  | JK  |             |
| 5 | Consult with Steve Wilcox, SAG UV, and ET-WDC regarding an appropriate format for ASCII representations of solar radiation data sets.   | AT, Steve Wilcox, Ann Webb, JK                                    | 30 Nov 2008 |
|   | Put recommendations for the specific developments that are desirable for WRDC to implement (open catalogue, ftp access to facilitate download of ascii files, etc) in a letter of support for WMO to send to WRDC.  | JK  | 30 Nov 2008 |
| 6 | All WDCs to work bilaterally with their reviewer to clarify open questions and implement recommendations.   | All   | 31 Dec 2008 |
| 7 | Report progress on action items from previous meeting (2007) to chair ET-WDC  | All   | 31 Oct 2008 |
| 8 | Provide text block regarding the WDC-RSAT collaboration with Schneefernerhaus for pilot   | KH, MB  | 15 Oct 2008 |

| # | Action Item  | Responsible                       | Due Date  |
|---|--|-----------------------------------|---|
|   | <p>project proposal to JK.</p> <p>Submit finalized proposal to WMO.</p> <p>Communicate to the WIGOS team the general consensus to support the pilot project and the key role that the support of WIGOS (funding effort and training) has in moving it forward.</p> | <p><b>JK</b></p> <p><b>GB</b></p> | <p><b>31 Oct 2008</b></p> <p><b>31 Oct 2008</b></p> |