

DATA MANAGEMENT WORKING GROUP (DMWG)

Reporting Period: October 2005 – September 2006

URL: <http://www.eol.ucar.edu/projects/ghp/dm/>

Chair(s) and Term Dates: Steve Williams (NCAR/EOL)

Overview:

Objectives (Mission)

The charge of the GHP DMWG is to assist the GEWEX Hydrometeorology Panel (GHP) in the coordination and facilitation of data management activities and issues between the GEWEX Continental Scale Experiments (CSE)s, the Global Runoff Data Centre, and the International Satellite Land-Surface Climatology Project. Further, these activities have been broadened to begin to coordinate data management between GHP and other GEWEX Panels and WCRP sponsored programs such as CEOP.

Status:

1. Maintained the DMWG Web Page located at: <http://www.eol.ucar.edu/projects/ghp/dm/> . Each member was asked to update their data policy, access, and protocol. The DMWG page(s) were re-organized and additional links and information were provided this year. In addition, a DMWG e-mail alias was maintained to facilitate communication between DMWG members (ghp-dmwg@eol.ucar.edu).
2. The GAPP CSE submitted updated gridded data sets for distribution as GHP data sets. These data sets were archived and are available from the DMWG web page (under the “data access” section). However, more work is needed to define and expand additional GHP data sets (see Future Activities #3 below).
3. Discussions continued with the CLIVAR and CLIC programs to keep data activities better coordinated. A frozen precipitation questionnaire was maintained to obtain needed metadata for the CEOP existing and planned Reference Sites located in the cryosphere. Information from the responses are located at: <http://www.eol.ucar.edu/projects/ceop/dm/webresponse/snow/> . In addition, an exercise is underway to incorporate sample CEOP Reference Site metadata into the Data and Information Service for CLiC (DISC) system (<http://clic.npolar.no/disc/>). This demonstration could eventually lead to better interoperability and shared archives between future WCRP projects. Appropriate data activities web links were maintained between CLIVAR/CLiC and the DMWG web pages.
4. Discussions continued to coordinate GHP data management activities with the other GEWEX Panels. The DMWG Chair (Steve Williams) continues to coordinate with the GEWEX Working Group on Data Management and Analysis (<http://grp.giss.nasa.gov/wgdma.html>) to better improve cross reference/access to and coordination of project data sets within GEWEX. CEOP was identified as a good prototype to begin to share some coordinated or “subset” datasets. This discussion will continue at the upcoming Pan-GEWEX Meeting.
5. Most of the DMWG work this year continued the organization of CEOP data management between the various CSEs. The CEOP Data Management web page was maintained and is available directly at: <http://www.eol.ucar.edu/projects/ceop/dm/> Links are available to CEOP data/metadata, data policies, information regarding Reference Sites and Model Location Time Series (MOLTS) profiles, and other pertinent data links. This web page is linked to the CEOP, GHP, and the DMWG home pages.
6. The inventory and metadata of CEOP Reference Sites were maintained. A dynamic matrix table is located directly at: <http://www.eol.ucar.edu/projects/ceop/dm/documents/rsite/> and summarizes specific information and metadata about the individual Reference Sites (locations, descriptions, maps, site contacts, sample data sets, instrumentation, parameters measured, etc.).
7. The DMWG continued to work on the CEOP EOP-3 and 4 “composite” Reference Site dataset. Each Reference Site was requested to submit their data in a common format as described in the data submission requirements report for the CEOP Reference Sites which was approved by the CEOP SSC in July 2003 (see http://www.eol.ucar.edu/projects/ceop/dm/documents/refdata_report/). As data/ metadata were submitted, NCAR/EOL performed consistency checks and applied a final quality assurance review to the final data. Any resulting problems or issues were subsequently solved with the respective data providers. Completed datasets were then posted to the on-line CEOP archive for distribution to the scientific community located at: http://data.eol.ucar.edu/master_list/?project=CEOP/EOP-3/4 Details (and status) of the EOP-3/4

Reference Site “composite” data set processing was prepared and published in the GEWEX News (August 2006). At the time of this report, EOP-3/4 data were received from 32 Reference Sites for review and final processing. Completed on-line data are available from 28 of these Reference Sites (including data from 12 sites which contain full annual cycles for both EOP 3 and 4).

8. NCAR/EOL has been coordinating with the CEOP Data Integration efforts and maintains a DODS in-situ server for the interactive distribution of Reference Site data. It has been demonstrated that this has facilitated the data exchange between CEOP Data Centers and allowed for the seamless integration of data between these centers. This operational DODS in-situ server was made available to the scientific community in May 2005 as part of the new CEOP Distributed Data Integration System (http://jaxa.ceos.org/wtf_ceop/). NCAR/EOL continued to make improvements to this system to increase data access speeds. Coordination efforts have also been on-going with the CEOP Centralized Data Integration System at the University of Tokyo (http://monsoon.t.u-tokyo.ac.jp/ceop-dc/ceop-dc_top.htm).
9. A Soils and Land cover questionnaire (http://www.eol.ucar.edu/cgi-bin/ceop/ceop_veg) for the CEOP Reference Sites was maintained to facilitate the metadata needed to perform model/satellite data intercomparisons and site data analysis. This questionnaire requested information on standardized surface parameters (including profile information) such as porosity, texture class, infiltration, bulk dry density, and saturated hydraulic conductivity. In addition, standardized land cover (and seasonal changes) which is important to document the Site characteristics during CEOP was requested. Currently, information from 18 sites have been received and documented. Results of this questionnaire are available at: <http://www.eol.ucar.edu/projects/ceop/dm/webresponse/veg/>.
10. The DMWG Chair (Steve Williams) attended the Asian Water Cycle Symposium in Tokyo, Japan (November 2005), 5th CEOP Implementation Meeting in Paris, France (March 2006), CLIVAR VAMOS Panel Meeting in Iguacu, Brazil (April 2006), U.S. AMMA Workshop in Washington D.C. (May 2006), and the LPB Implementation Meeting in Sao Paulo, Brazil (September 2006) to report on on-going CEOP data activities as well as formulate plans for CEOP Phase 2.
11. The DMWG Chair (Steve Williams) also represented WCRP on the International Group on Earth Observations (GEO) Data and Architecture Committee and attending the first meeting in Tokyo, Japan (November 2005) as well as participated on various teleconferences.
12. The DMWG has initiated and continues to coordinate with several of the proposed CEOP Hydrology Reference Basins to add these data to the CEOP database. Further data details (e.g. formats, resolutions, etc.) still need to be worked out, but data from 3 of these Reference Site Basins are available from the CEOP archive in “native” formats. Further information is available at: <http://hydrology.princeton.edu/~luo/ceop/index.php>.

New Directions:

Each member of the DMWG was requested to provide a summary of the plans for final archive of their data. The following is a brief summary by member:

AMMA/CATCH - Plans and the data policy are being formulated at the AMMA Project Office, Paris.

BALTEX – The operational data are already archived and available in long-term European Data Centres and will be available through respective data center and BALTEX policies. The BALTEX PI data/products are not included but will be available from the respective investigators.

GAME – Distributed data will be archived and available at respective institutions for 10 years (an additional 10 year extension is being discussed). Most data are also available on CD/DVD with on-line “image” versions.

GCIP/GAPP – The GCIP data are archived at NCDC and NCAR (with subsets at ARM). GAPP data are archived at NCAR and ARM. Both project datasets are expected to be available from NCAR as plans are being developed to link these data to the NCAR Community Data Portal (CDP).

LBA – The Brazil Ministry of Science and Technology has agreed to archive the data for 20 years (probably at CPTEC). Data will be available through the LBA Data Information System (DIS).

LPB – Data archive and access plans are being formulated

MAGS – The final data archive and access plans are being formulated. Selected datasets will be archived on CD/DVD.

MDB – Data archive and access plans are being formulated.

GPCC – The gridded/point data all archived at GPCC. GPCP data are archived and will be available at NASA/GSFC and NOAA/NCDC.

GRDC – All data are archived and will be available at GRDC.

ISLSCP – Data are archived at NASA/GSFC. Some datasets are archived and will be available on Compact Disk (CD).

****NOTE**** - CEOP data (subset of GHP/CSE data) are archived and available at long-term archives at the National Center for Atmospheric Research (NCAR), the CERA Database (Max Plank Institute), and the University of Tokyo (UT).

Future:

1. The DMWG will continue the compilation of CEOP Reference Site data/metadata and begin data collection for Reference Sites identified for CEOP Phase 2. This information will be maintained on the CEOP data management web page(s), directly at: <http://www.eol.ucar.edu/projects/ceop/dm/documents/rsite/>. More interaction between the CSE and the Reference Site management is needed to expedite data submission.
2. The DMWG members will be requested to maintain updated CSE information on data policy, data inventory, data access, and data contacts. This information will continue to be linked from the DMWG Home Page.
3. Continue interaction with CSEs on what additional regional data sets and products might be useful as “GHP datasets” as well as continuing to finalize long-term archival plans.
4. NCAR/EOL has submitted a proposal to begin work on compilation of the CEOP Phase 2 “composite” Reference Site data set incorporating common parameters, format, and temporal resolution. The DMWG will work with the CSEs and reference sites to coordinate the data submission, quality assurance, and posting of data to the scientific community. Also as a component of this proposal, NCAR/EOL will reformat existing CEOP Reference Site data to improve the efficiency of its existing DODS Server.

Key Results:

- Provided access to GHP data through a distributed data interface
- Coordinated GHP data management activities with other GEWEX Panels and WCRP Programmes
- Continued to develop and maintain the CEOP Data and Metadata archives
- Continue to refine plans for long-term GHP/CSE data archival.

Issues and Recommendations:

- The DMWG needs feedback from the GHP (and in particular the other GHP working groups) regarding their data requirements (i.e. gridded data sets, “composites”, products, subsets, etc).
- Work needs to be continued to include hydrology data/metadata into the GHP database (primarily through the CEOP activities as a prototype). The CEOP Hydrology working group has the action to coordinate and provide these data.
- In the longer term, the DMWG needs to draft a GHP wide Data Policy. One suggested strategy is to use the approved CEOP Reference Site data policy as a “prototype”. GHP needs to improve the integration of GHP datasets with GEWEX and other WCRP projects. One first step might be the development of a method to cross-link CSE data bases and later other GEWEX and WCRP related data sets (i.e. data portals, OpeNDAP, GIS, XML, digital libraries, etc).

Contributions to WCRP Strategic Framework:

WCRP is currently evaluating and working to establish an overall data management strategy mainly through the Working Group on Observations and Data Assimilation, but input from other project panels. The input from GHP/DMWG data activities are critical to this effort since “prototype” data integration efforts such as CEOP and across GEWEX will demonstrate potential strategic planning for WCRP.

Contributions to Society and WCRP/GEWEX Visibility

Providing access to GHP and GEWEX related data are critical to supporting international scientific research and the interactions within the climate community. Plus such activities are important to WCRP/GEWEX strategic planning as well as Education and Outreach.

Summary:

The DMWG continued to coordinate data management activities between the various CSEs and GHP member groups. CEOP continues to serve as the “prototype” for this activity and is demonstrating progress that the GHP and GEWEX can benefit from.

Publications:

GEWEX News – Special Data Issue (August 2006)

CEOP Newsletter #9 (February 2006) entitled “CEOP Phase-1 Reference Site Data Set Availability”

Lawford, R.G., M. Bosilovich, S. Eden, S. Benedict, C. Brown, A. Gruber, P. Houser, J. Huang, W. Lau, T. Meyers, K. Mitchell, C. Peters-Lidard, J. Roads, M. Rodell, D. Tarpley, and S. Williams, 2006: U.S. Contributions to the Coordinated Enhanced Observing period (CEOP). *Submitted to the Bulletin of the American Meteorological Society.*

Yang, K., M. Rasmy, S Rauniyar, T. Koike, K. Taniguchi, K. Tamagawa, M. Bosilovich, and S. Williams, 2006: CEOP-based Diagnosis of Prediction Skill of Current Operational General Circulation Models and Land Data Assimilation System. *Submitted to the Journal of the Meteorological Society of Japan (JMSJ).*

Meetings and Workshops:

No specific DMWG meetings were held or are planned for the upcoming year. Most work is done via e-mail and teleconferences. Smaller group meetings are held whenever possible at larger venues such as the GHP, GRP, Pan-GEWEX, and CEOP Meetings.

List of Members:

During the 2005-2006 period, the DMWG membership remained stable (No term dates have been established since most of these members are data managers within their respective CSEs and are the appropriate long-term contact for CSE and GEWEX data issues):

Hans-Jörg Isemer (BALTEX)
Thierry Lebel (CATCH/AMMA)
Kiyotoshi Takahashi (GAME)
Steve Williams (GCIP/GAPP)
Luiz Mattesco Horta (LBA)
Jose Meitin (LPB)
Robert Crawford (MAGS)
Alan Seed (MDB)
Udo Schneider (GPCC)
Thomas Maurer (GRDC)
Eric Brown de Colstoun (ISLSCP)

Individual contact information can be found at:

<http://www.eol.ucar.edu/projects/ghp/dm/people/members.html>