1. **Data and Information Management**

**6.1 Introduction**

A comprehensive DEEPWAVE project web page has been developed and is available at: <https://www.eol.ucar.edu/field_projects/deepwave>. This web page provides project documentation and access to operations and logistical information, facilities and instrumentation, mailing lists, meetings and presentations, education and outreach, and data management from the planning to the post-field phase of DEEPWAVE. The development and maintenance of a comprehensive and accurate data archive is a critical step in meeting the scientific objectives of DEEPWAVE. The overall guiding philosophy for the DEEPWAVE data management is to make the completed data set (and project documentation described above) available to the scientific community as soon as possible following the DEEPWAVE Field Phase, while providing ample time to the DEEPWAVE Investigators and Participants to process, quality control, and analyze their data before providing open access. DEEPWAVE will coordinate closely with other collaborating partners in the archival and exchange of data and associated information.

The DEEPWAVE data will be available to the scientific community through a number of designated distributed DEEPWAVE Data Archive Centers (DDAC)s, coordinated by the NCAR/EOL. This would include the main archive at EOL but also include (but not limited to) other archives such as NASA, NOAA, NIWA, NZMS, Australian BoM, and DLR. The EOL coordination activities fall into three major areas: (1) determine the data requirements of the DEEPWAVE scientific community and develop them into a comprehensive DEEPWAVE Data Management Plan through input received from the DEEPWAVE Scientific Steering Committee (SSC), project participants, and other tools such as the data questionnaire; (2) development and implementation of an on-line Field Catalog to provide in-field support and project summaries/updates for the Principal Investigators (PIs) and project participants to insure optimum data collection; and (3) establishment of a coordinated long-term distributed archive system and providing data access/support of both research and operational data sets for the DEEPWAVE PIs and the scientific community. To accomplish these goals, EOL will also be responsible for the establishment and maintenance of the DEEPWAVE Data Management Portal. These web pages provide "one-stop" access to all distributed DEEPWAVE data sets, documentation, on-line Field Catalog products, collaborating project data archives, and other relevant data links. EOL will make arrangements to ensure that "orphan" data sets (i.e. smaller regional and local networks) will be archived and made available through the DEEPWAVE archive. The EOL may also quality control and reformat selected operational data sets (e.g. atmospheric soundings or surface data) prior to access by the community as well as prepare special products or “composited” data sets (see Section 6.4).

**6.2 Data Policy**

The World Meteorological Organization (WMO) Resolutions 40 and 25 (adopted by the XII Congress on 26 October 1995) comprises the basis for the DEEPWAVE data policy and protocol to be adopted and practiced by each of the DDACs:

*"As a fundamental principle of the World Meteorological Organization (WMO), and in consonance with the expanding requirements for its scientific and technical expertise, the WMO commits itself to broadening and enhancing the free and unrestricted international exchange of meteorological and related data and products".*

In general, users will have free and open access to all the DEEPWAVE data, subject to procedures to be put into place at the various DDACs and the DEEPWAVE Data Policy. The following is a summary of the DEEPWAVE Data Policy by which all DEEPWAVE participants, data providers, and data users are requested to abide by:

1. All investigators participating in DEEPWAVE agree to promptly submit their preliminary processed data and metadata to the main DEEPWAVE Data Archive Center at EOL no later than **29 January 2015** (six months after the end of the field campaign) to facilitate initial instrument inter-comparisons, quality control checks and calibrations, as well as early interpretation of the combined data set. Individual preliminary datasets can be restricted (password protected) at the discretion of the data provider. All archived supporting operational data and products will be open and accessible by the Scientific Community during this period. **The preliminary data submission period is from 29 July 2014 to 29 January 2015.**

2. DEEPWAVE Investigators agree to submit their final research data and metadata to the EOL within the one-year period following the conclusion of the field campaign. **The final data submission period is from 29 July 2014 to 29 July 2015.**

3. During the initial data analysis period, defined as a one-year period following the preliminary data submission deadline to the DEEPWAVE archive, DEEPWAVE Principal Investigators (PIs) will have exclusive access to these research data. This initial analysis period is designed to provide an opportunity to quality control the combined data set as well as to provide the PIs, their students and collaborators ample time to analyze and publish their results. **The initial data analysis period is from 29 January 2015 to 29 January 2016.**

4. All data and metadata in the archive will be considered open to the public domain 18 months following the end of the field campaign (i.e., on **1 February 2016** and thereafter). However, any research dataset within the DEEPWAVE archive can be opened to the public domain earlier at the discretion of the responsible data provider in consultation with the DEEPWAVE SSC.

5. A list of DEEPWAVE Investigators will be provided by the project science leadership to EOL and will include the PIs directly participating in the field experiment as well as collaborating scientists and agencies who have provided guidance and data in the planning and analysis of DEEPWAVE data. All DEEPWAVE investigators will have equal access to all data. All data shall be promptly provided to other DEEPWAVE investigators on the above specified list upon request. However, the DEEPWAVE science leadership will be responsible for approving any data requests from investigators not included on the list.

6. During the initial data analysis period, the responsible data provider must be notified first of the intent to use their data, in particular if data are to be provided to a third party (e.g., journal articles, presentations, research proposals, other investigators). It is strongly encouraged that the responsible data provider(s) be invited to become collaborators and/or co-authors on any projects, publications and presentations. If the contribution of the data product is significant to the publication, the PIs responsible for generating a measurement or a data product should be offered the right of co-authorship. Any use of the data should include an acknowledgment or preferably a citation (e.g. Digital Object Identifiers or DOIs). The EOL expects to be assigning DOIs for all final datasets submitted to the main archive at EOL. In all circumstances, the responsible data provider(s) should be acknowledged appropriately.

7. All acknowledgments of DEEPWAVE data and resources should identify: (1) DEEPWAVE; (2) The providers who collected the particular datasets being used in the study; (3) The relevant funding agencies associated with the collection of the data being studied, and (4) the role of EOL or relevant data archive center, and (5) use of any relevant DOIs.

8. The EOL will be responsible for the long-term data stewardship of the DEEPWAVE archive.

**6.3 Real-Time Data**

**6.3.1 DEEPWAVE Field Catalog**

EOL will implement and maintain a web-based DEEPWAVE on-line Field Catalog that will be operational during the DEEPWAVE field phase to support the field operational planning, product display, and documentation (e.g. facility status, daily operations summaries, weather forecasts, and mission reports) as well as provide a project summary and “browse” tool for use by researchers in the post-field analysis phase. Data collection products (both operational and research) will be ingested into the catalog in near real time beginning **the week of 5 May 2014**. The Field Catalog will permit data entry (data collection details, field summary notes, certain operational data etc.), data browsing (listings, plots) and limited catalog information distribution. A Daily Operations Summary will be prepared and contain information regarding operations (aircraft flight times, major instrument systems sampling times, weather forecasts and synopses, etc.). These summaries will be entered into the Field Catalog either via a web form or through uploading of a pdf document. It is important and desirable for the PIs to contribute product graphics (e.g., plots in gif, jpg, png, or pdf format) and/or preliminary data to the Field Catalog whenever possible. Although the Field Catalog will be publically available, access to preliminary data will be restricted to project participants only. Updates of the status of data collection and instrumentation (on a daily basis or more often depending on the platforms and other operational requirements) will be available. Public access to the on-line Field Catalog is located at: <http://catalog.eol.ucar.edu/deepwave/> The Field Catalog User's Guide (with specific instructions for submitting reports and data products) is located at: <http://catalog.eol.ucar.edu/deepwave/tools/user_guide> . Help documentation for working with the Field Catalog can be found at <http://catalog.eol.ucar.edu/deepwave/tools/help> .

EOL will monitor and maintain the field catalog through the duration of the field deployment and also provide in-field support and training to DEEPWAVE project participants. Following the DEEPWAVE field phase, this Field Catalog will continue to be available on-line (as part of the long-term archive) to assist researchers with access to project products, summaries, information, and documentation. Preliminary data will not be retained as part of the Field Catalog, but will be password protected and available through the DEEPWAVE archive (Section 6.4).

**6.3.2 Field Catalog Components, Services and Related Displays**

The Field Catalog will be the central web site for all activities related to the field campaign. As such it will contain products and reports related to project operations as well as forms for entering/editing reports, uploading new products, data files, photos or reports. The Field Catalog will also provide a preliminary data sharing area, a missions table to highlight major project operations, links to related project information and help pages to familiarize users with the various features of the catalog interface. The DEEPWAVE Field Catalog front page will be customized to provide pertinent project information and rapid access to the most popular catalog features and will include access to the GIS display tools like Catalog Maps and the Mission Coordinator display. Access to project chatrooms will be provided through the Field Catalog with a link on the front page as well.

GIS Display tools

The Mission Coordinator and Catalog Maps displays are the main GIS display tools that will be provided by EOL for the DEEPWAVE campaign. The Mission Coordinator display is a real-time tool for situational awareness and decision-making aboard the NCAR GV aircraft. This display will contain a small subset of products pertinent to aircraft operations from the Field Catalog that can be displayed along with GV and other aircraft tracks. The Mission Coordinator display is also accessible to forecasters and aircraft coordinators on the ground. The Catalog Maps display is a GIS tool that is integrated into the Field Catalog and provides access to a larger number of real-time products as well as an ability to replay products from any previous day during the campaign. Project participants will be able to easily follow project operations using the Catalog Maps tool.

IRC Chat

EOL will provide IRC chat services as the primary communications tool between the various ground-based and airborne facilities. A number of logged chatrooms will be provided for DEEPWAVE including (but not limited to):

*#DEEPWAVE* – for chat related to mission coordination and real-time decision for all DEEPWAVE facilities

*#GV* – for chat related to instrument issues aboard the GV

After the DEEPWAVE campaign is completed, the logs from these chatrooms will be sanitized to remove sensitive information and will become part of the long-term data archive for the project. As many project participants will be connecting to chat via different networks including satcom and possibly cellular, dropouts may occur when the user loses their internet connection or is in a no-coverage area. In each of the logged chatrooms a replay capability is provided so that when connectivity is re-established the user can query the system to replay all messages in a given chatroom sent during the last user selectable number of minutes. The chat service also provides the capability for users to have private conversations that are not logged should they wish to move their discussion to a separate chatroom or desire to exchange sensitive information. Help documentation is available in the Field Catalog that describes all of these features along with a description of the chat interface and common chat commands.

**6.4 Data Archive and Access**

The DEEPWAVE will take advantage of the capabilities at existing DDACs to implement a distributed data management system. This system will provide “one-stop” single point access (Project Portal) at EOL using the web for search and order of DEEPWAVE data from DDACs operated by different agencies/groups with the capability to transfer data sets electronically from the respective DDAC to the user (or to EOL). Access to the data will be provided through the Data Access and Data Documentation links from the DEEPWAVE Project page (<https://www.eol.ucar.edu/field_projects/deepwave> ). These Data Management links will contain general information on the data archive and activities on-going in DEEPWAVE (i.e. documents, reports), data submission instructions and guidelines, links to related programs and projects, and direct data access via the various DDACs. Parts of the website will be password protected and access restricted, as appropriate.

In addition to providing a comprehensive archive of PI and supporting datasets, the EOL will be responsible for the collection, processing, quality assurance, and archival of all DEEPWAVE Upper Air soundings (both special research and operational releases) following the field phase. This will include the processing of highest vertical resolution datasets (all in a common format) as well as creation of a 5-mb interpolated “composite” dataset suitable for model ingest and intercomparison.

The EOL will be responsible for the long-term data stewardship of DEEPWAVE data and metadata. This includes ensuring that “orphan” datasets are properly collected and archived, verifying that data at the various DDACs will be archived and available in the long-term, and that all supporting information (e.g. Field Catalog) are included in the archive.