

SHEBA Data Management Support

*J. A. Moore, R. Dirks, G. Stossmeister, and N. Gamage
UCAR Joint office for Science Support
Boulder, Colorado*

Introduction

The Surface Heat Budget of the Arctic Ocean (SHEBA) Project is designed as a multiagency sponsored, interdisciplinary project with many investigators and varied instrumentation. An integrated data management activity is central to providing a complete database that is easily accessible to all project investigators and the science community in general. The University Corporation for Atmospheric Research (UCAR) Joint Office for Science Support (JOSS) has been funded by the National Science Foundation (NSF) to work with the SHEBA Project Office (SPO) in the development and implementation of a data management plan to support Phase II SHEBA activities.

The SHEBA Project is a multiagency interdisciplinary effort to address the interaction of the surface energy balance, atmospheric radiation and clouds of the Arctic Ocean. SHEBA is part of the larger Arctic System Science/Ocean-Atmosphere-Ice Interactions (ARCSS/OAII) Program sponsored by the NSF. The primary goal of SHEBA is to characterize the surface radiation balance, mass changes of the sea ice, the storage and retrieval of energy and salt in the ocean mixed layer, the formation of clouds and their interplay with the radiation balance, and the relationships between the air-sea-ice system and the signals received by satellite-borne remote sensors.

The SHEBA field program will consist of a multi-season deployment of an ice station in the Arctic Ocean (initial point approximately 143W - 76N) north of Prudhoe Bay, Alaska. The Canadian Coast Guard Ship (CCGS) Des Groseilliers will act as the central base of operations on the Arctic ice pack. Deployment is planned for approximately 13 months, from mid-September 1997 through freeze-up (approximately mid-October) 1998. Several intensive observation periods (IOPs) will occur during the experiment, making use of aircraft, submarines, satellites, and autonomous platforms.

JOSS has primary responsibility for the implementation of the SHEBA data management strategy. The final objective is a high quality data archive that has easy and timely access by a large community of investigators. This is a large task given the diversity of participation and instrumentation

planned for SHEBA. In addition, several other national data centers will assist in this project in assuring the availability of this rich dataset.

JOSS tasks include provision of an on-line catalog at the ice station and on the mainland to provide primary access to summary data and documentation of activities during and following the field season; access to preliminary datasets and selected operational data using the on-line catalog and CODIAC during and after the 13-month field deployment; planning and coordination of data management activities among other agencies, projects, and groups to meet investigator needs at the ice station and on the mainland and coordinate the transfer of SHEBA datasets to the final archive at the National Snow and Ice Data Center (NSIDC).

General guidance will be given to JOSS and the project office by the ARCSS/OAII SHEBA Panel, NSF/ARCSS and Office of Naval Research (ONR) High Latitude Dynamics management groups.

SHEBA Data Management Policy

The following data protocols are specified in the SHEBA Science Plan, August 1996, and form the basis of the data management strategies discussed in subsequent sections of this document.

- Open access to all SHEBA datasets.
- Cooperation/coordination with several existing data centers. SHEBA will take advantage of several existing data centers to house a variety of datasets to be collected, including the NSIDC, National Center for Atmospheric Research (NCAR), National Aeronautics and Space Administration (NASA)/Langley Research Center (LaRC)/Distributed Active Archive Center (DAAC), Atmospheric Radiation Measurement (ARM) Program/Oak Ridge National Laboratory (ORNL) and UCAR/ JOSS.
- Cooperative agreements for unrestricted exchange and access of SHEBA data with related projects will be

established, including ARM, First ISCCP Regional Experiment (FIRE), and Scientific Ice Expeditions (SCICEX). In addition, informal data exchange is expected from several ancillary experiments that are utilizing SHEBA facilities. The following efforts will be made to provide access to appropriate data from National Oceanic and Atmospheric Administration (NOAA), NASA, U.S. Department of Energy (DOE)-ARM, etc. that might be conducting programs in conjunction with or of interest to SHEBA investigators:

- prompt submission of data
- complete documentation with all data
- prompt dissemination of data
- direct exchange of data among the investigators is encouraged
- use a distributed archive strategy
- ensure that the SHEBA dataset is comprehensive.

JOSS Data Management Strategy

It is important that the SHEBA data management strategy be responsive to the needs of the investigators assuring data are accurate, accessible, well documented and disseminated in a timely fashion. It is also important that the investigators know what is expected of them in this process.

Central to the JOSS data management is the on-line, interactive, catalog, archival and distribution system. The on-line catalog capability within this system allows investigators limited perusal and display of preliminary data products during the field phase. The catalog will also provide in-field project summaries (daily or otherwise as required) and summarize collection activities. offers scientists a means to identify datasets of interest, the facilities to view selected data and associated metadata, and the ability to automatically obtain data from geographically dispersed data centers via Internet file transfer protocol (FTP) or separate media (tapes, CD-ROM, disks, etc.).

The catalog will permit on-line entry (data collection details, metadata, field summary notes, certain operational and research data etc.), data browsing (listings, images) and information distribution to other locations worldwide. Daily operations summaries will be prepared that contain as much information about operations (major instrument systems

status and sampling times, satellite overpasses, submarine tracks, aircraft flight times and tracks, etc.) as desired by the investigators. The operational SHEBA catalog may be found at: <http://www.joss.ucar.edu/sheba/catalog/>

One version of the on-line catalog will be operated on the ship by SPO and/or JOSS personnel. An identical catalog will be "mirrored" (identical in form, function and content) in the mainland. In this way, scientists not at the ice station can monitor project operations, instrument status and special observations of interest to them. Both catalogs will be completely updated every 6 weeks to correspond to the transport of a CD-ROM to coincide with logistics aircraft flights. Following the field season, a complete catalog will continue to be operational for several years and be accessible from JOSS or the SHEBA World Wide Web (WWW) home page. The catalog will act as the SHEBA Project Operations Summary for years to come.

The CODIAC archive function will be used during and following the field season as the primary access point for SHEBA operational data and for preliminary and final research datasets as they become available. Links will also be provided from CODIAC to other data centers holding cooperative project data and other relevant information to SHEBA research. CODIAC will be available for receipt and redistribution of SHEBA datasets from the first day of the project. It will be possible for the principal investigators to provide preliminary and final datasets along with associated documentation they become available throughout the 13-month field season (SHEBA Archive).

Following the field season, CODIAC will be populated with any data provided by the investigators. Data will be searchable by file name and/or data type. The investigators will have complete responsibility for the processing and delivery of their data to JOSS/SPO within 1 year of the conclusion of the field phase. As data are received they will be promptly staged as "preliminary" datasets and made available to all SHEBA participants.

Information from the on-line field catalog (data collection summaries, operations reports, instrumentation attributes, graphical products, etc.) obtained during the field season will also be accessible. JOSS intends to keep the catalog available for several years following the end of the field phase as a ready summary of the ice station observations and operations. Links to the other data centers (ARM, NASA, etc.) will be in place as necessary so that users can have access to all SHEBA specific data from a single entry point. It will be possible to make data requests, via WWW/Netscape or other interface, and download files, via FTP, from the catalog or other data center.

The importance of providing complete and separate documentation (“read me” file) with every SHEBA dataset, regardless of format, cannot be over emphasized. It is critical for the long-term viability of the comprehensive data base and the easiest way to explain to everyone who might use a dataset important details that might be forgotten in years to come. The SHEBA investigators are following agreed to procedures for the submission of their datasets.

SHEBA Cooperation with Other Programs

The SHEBA Project will take advantage of existing data centers for archival, and in some cases collection, of project specific datasets and important supporting information. The principal sites and their respective roles in SHEBA data management are listed below.

1. The NSIDC is the archive for all ARCSS datasets.
2. NCAR will archive model, sounding, and C-130 aircraft data.
3. ORNL is the archive of the DOE ARM data. A memorandum of participation between SHEBA and ARM now exists to assure easy access to these important data by all SHEBA participants.
4. The NASA LaRC DAAC archives all data collected during the multiyear FIRE.

JOSS will continue to work with these data centers to make the exchange of information among the projects as easy as possible for all participants. Coordination is already well under way to assure that high quality data from all of these programs will be accessible as quickly as possible following their respective field phases.