

Data and Model Output Management

Background and Needs

The implementation of SOLAS will involve the collection of large quantities of environmental data by both nationally and internationally organised projects. These data will be collected from process studies and experiments (field and laboratory), time series studies, and large-scale surveys. Similarly, SOLAS will make use of a hierarchy of different modelling approaches. In most cases, the utility of the models and data involved in these projects will extend beyond the projects themselves and be of interest to other investigators. Further, many SOLAS data will be more useful when combined with, or compared against, other data and models, including non-SOLAS data. Scientific findings and conclusions derived from SOLAS projects should be available for assessment by independent scientists: this implies that the underlying data and/or models must be readily accessible.

Increasingly, key management issues are based on model results as well as data. Model-derived results are also used extensively as input for other models, for scientific planning and for policy decisions. SOLAS must therefore ensure that the models developed and/or used in SOLAS research, as well as the data collected, are documented sufficiently to allow independent evaluation and be readily accessible to the scientific community for assessment and interpretative purposes.

Data and model management are therefore critical logistical tasks for SOLAS and we recognise that there are significant challenges ahead: effective data management is a global problem and SOLAS will make strenuous efforts to adopt best practice and learn from past experiences (e.g. WOCE, JGOFS etc).

SOLAS science will involve collection of complex and sophisticated data sets. This will include difficult, error-prone measurement of biological, physical and chemical parameters. Because SOLAS is an international, multi investigator programme, such measurements will be made all over the world, at different times, by different groups, often using different equipment and techniques.

Attention to data quality management will be critical for the scientific integrity and success of SOLAS.

Data Management Principles

Certain basic principles should guide the development of data management procedures in SOLAS:

- Do not 'reinvent the wheel'; use existing knowledge and infrastructure, wherever appropriate.
- Use internationally agreed standards and protocols (e.g. those of ISO, W3C, IOC/ICES), wherever possible.
- Work with other projects towards an integrated data management system and policy.
- Plan ahead for rapid data assembly.
- Data managers should support data gatherers.
- SOLAS should reward excellence in data collection and data release.
- Participation in SOLAS research requires submission of data to a SOLAS approved database or centre.

Specific Steps Towards Data Management

SOLAS will establish an international data management task team with at least one staff member. This team should:

- Evaluate and document the likely data needs of SOLAS (data types, quantities and sources, metadata requirements).
- Conduct a review and have intense discussions with the WCRP (WOCE/CLIVAR) and IGBP/SCOR (JGOFS, IMBER, GLOBEC, LOICZ, GEOHAB) data management communities concerning lessons learned and present data management policies, problems and solutions. Identify the potential for SOLAS data management within existing structures and programmes.
- Host a workshop for the modelling community in order to develop and write a practical model management and documentation plan.
- Evaluate together with other programmes the feasibility and benefits of establishing a peer reviewed 'Journal of Global Change Data' and participate in constructing the IGBP Global Change Atlas.
- Write a detailed SOLAS data management policy. This policy is to include time limits, enforcement procedures, access rights, metadata requirements, reporting requirements and procedures, data quality and archiving, and linkages to data agencies. Potentially, this document will include guides to model description and data organisation.
- Initiate negotiations with data centres about possibilities for their direct support of individual PIs with respect to data management procedures and tools.

Data Quality Management

In addition to establishing clear guidelines for data management, SOLAS must also establish procedures for assessing and controlling data quality. Once again, lessons learned during WOCE and JGOFS can be used to help address such issues.

Data quality management should be addressed by the Data Management Task Team by:

- establishing clear quality targets for SOLAS data;
- documenting recommended protocols for 'standard' SOLAS measurements;
- providing support for technical workshops, training sessions and calibration and intercalibration activities.