

Data Policy
North American Carbon Program
February 12, 2007

Introduction

The North American Carbon Program (NACP) is a multidisciplinary research program to obtain scientific understanding of North America's carbon sources and sinks and of changes in carbon stocks needed to evaluate societal concerns and to provide tools for decision makers. Successful execution of the NACP will require an unprecedented level of coordination among observational, experimental, and modeling efforts regarding terrestrial, oceanic, atmospheric, and human components of the carbon cycle. The Program is supported by several U.S. agencies through a variety of intramural and extramural funding mechanisms and award instruments. These federal agencies are actively coordinating their contributions to the NACP under the auspices of the Carbon Cycle Interagency Working Group (CCIWG).

At the heart of the NACP research strategy is an integrated data and information management system that will enable researchers to access, understand, use, visualize, and analyze large volumes of diverse data at multiple thematic, temporal, and spatial scales. Managing and integrating data for NACP requires an overarching data policy to ensure that participants have full, open, and timely access to data, to promote the exchange of quality controlled / quality assured data, to protect intellectual property rights, and to ensure that proper credit is given to data originators through authorship, citation, or acknowledgement. The data and information system will be an important asset enabling informed decision making and policy formulation.

This policy pertains to the life-cycle of data during NACP – from data collection, through quality checking and analysis, to distribution to NACP participants, and to depositing finalized products in a long-term archive. When NACP has been completed, this Data Policy will no longer be in effect.

NACP Data Policy

1. NACP Data

For the purposes of this Policy, the term "NACP data" shall mean the primary observations, monitoring data, site characterization information, model output, remotely sensed products, and ancillary data specifically supported by U.S. agencies to meet the goals outlined in the NACP Science Implementation Strategic Plan (<http://www.nacarbon.org/nacp/documents.html>). All data must be accompanied by quantified estimates of uncertainty.

Metadata are defined as the descriptive information such as content, quality, and conditions that characterize a set of measurements.

Data sets from other sources (e.g., forest inventory data, satellite products, global meteorological analysis, etc.) may be needed to meet the goals of NACP, but were not funded specifically for NACP. There may also be data from overlapping programs (e.g. Ocean Carbon and Climate Change) that are part of NACP but also fall under other program management policies. The CCIWG will broker agreements with officials and institutions in order to obtain such data sets required for use by NACP scientists.

2. Sharing Data

Full, open, and timely sharing of the full suite of North American data sets for all NACP researchers is a fundamental objective. All data and associated metadata are to be made available to every NACP investigator as soon as feasible but no later than 12 months after acquisition.

For those programs in which selected principal investigators have a period of exclusive data use longer than 12 months, data must be made openly available as soon after that period as possible. In each case, the agency sponsoring the data collection will explicitly define the duration of any exclusive use period.

Data will be analyzed cooperatively by all scientists involved in obtaining them. In addition, NACP data will be made available in preliminary form to NACP investigators to enable quality assurance through preliminary analysis and intercomparison with other data sets. Corrections and refinements to data products will be made as the analysis proceeds. Revisions will be noted, and investigators and the NACP data system will maintain version control.

3. Data management based on standard metadata and metadata

International metadata standards (e.g., ISO19115) should be used to facilitate discovery, sharing, and understanding data products among the diverse investigators within NACP. In addition, standard data formats, parameter names, and units are encouraged.

4. Credit to Data Collectors

When data are used by others in publications or in modeling or integrating studies during the course of NACP, the scientist collecting the data will be credited appropriately, either by co-authorship, citation, or acknowledgement. For data that have not been published, the data originators must be informed of analysis and publication plans well in advance of submission of a paper, given an opportunity to read the manuscript, and, if appropriate, be offered co-authorship. In cases where unpublished data from other investigators are a minor contribution to a paper, the data are to be referenced by a citation or acknowledgement. Users of the data must state the primary source of the data as well as the version number.

5. Protecting the rights of students

Universities require that key data collected by graduate students should not be published prior to submitting a thesis or dissertation because to do so can jeopardize the student's academic interests. As a result, data collected specifically for a student's thesis or dissertation are not subject to the data sharing obligations of this policy until after the student has completed their degree. This special consideration will be limited to 2 years after data collection for a masters student and 3 years after collection for a doctoral candidate.

6. Archive

NACP metadata and data products, including value-added products and model input and output generated by the program, need to be archived when the data sets are finalized or at the completion of NACP. The CCIWG will formulate a strategy for archiving data and products developed by NACP project activities, including establishing a process and timeline for archiving key NACP data. A data archive plan developed early in NACP will ensure survival, integrity, long-term stewardship, and access to NACP data and metadata.

Long-term data archives must include easily accessible information about the data holdings, including quality assessments, version of the data or model input and output, supporting ancillary information, and guidance and aids for locating and obtaining the data.

7. Sharing Models

Model source code, when made publicly available, can be used to understand the uncertainty of model results relative to results from other models or observations, enable others to see how models treat individual processes, and ultimately serve to improve models.

NACP investigators will make numerical models and key algorithms used in publications available to program participants and the broader user community no later than upon publication. As a model evolves during NACP, newer model versions associated with publications will also be made available to the research community upon publication. Each model submitted should include model source code along with version number, documentation in terms of a technical note or carefully documented code, example input and output, and a history of changes. Model archiving will follow the best-practice requirements and recommendations proposed in Thornton et al. (2005).

Credit will be given to model developers through citation when information from archived models is used in publications.

Certain model codes may be subject to copyright or international agreements that are more restrictive. The CCIWG will be responsible for brokering agreements with officials and institutions in order to obtain specific models required for use by NACP scientists.

8. Acknowledging NACP

NACP investigators will include an acknowledgement in each publication or presentation arising from participation in NACP. The wording shall be similar to the following:

“This study was part of the North American Carbon Program.”

Data providers and funding agencies may request additional acknowledgements.

Upon publication of results, investigators should send the NACP Office an electronic copy of the publication.

9. Resolving conflicts over data and the data policy.

Conflicts over the interpretation of this Data Policy, or its implementation, will be resolved at the lowest level possible within the NACP organization. Direct resolution of issues between investigators is preferred; sponsoring agencies may have to become involved if resolution cannot be reached.

References

Thornton, P.E., R.B. Cook, B.H. Braswell, B.E. Law, W.M. Post, and B.T. Rhyne. 2005. Archiving numerical models of biogeochemical dynamics. *Eos, Transactions of the Am. Geophysical Union* 86: 431 – 432.