

Plan Overview

A Data Management Plan created using DMPTool

Title: Daymet Follow-On: Surface Weather Data with Uncertainty Quantification for Terrestrial Ecosystem Process Models

Creator: DMP dmpcurator

Affiliation: University of California, Office of the President (UCOP)

Principal Investigator: DMP dmpcurator

Data Manager: DMP dmpcurator

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 8278

Template: NSF-AGS: Atmospheric and Geospace Sciences

Last modified: 05-30-2014

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Daymet Follow-On: Surface Weather Data with Uncertainty Quantification for Terrestrial Ecosystem Process Models

Products of research

Describe the types of data and products that will be generated in the research, such as physical samples, space and/or time-dependent information on chemical and physical processes, images, spectra, final or intermediate numerical results, theoretical formalisms, computational strategies, software, and curriculum materials.

Meteorological station data used as input to Daymet (daymet.ornl.gov) will be obtained from the GHCN-D collection at NCDC as well as regional meteorological offices. These data will be checked for quality and formatted into a consistent comma separated values format, followed by automated ingest and storage as netCDF formatted files.

Data format

Describe the format in which the data or products are stored (e.g. hardcopy logs and/or instrument outputs, ASCII, XML files, HDF5, CDF, etc). What metadata will be part of the data sets produced?

The data generated by the proposed project will consist of daily, gridded surfaces of climatological data (temperature, precipitation, humidity, and radiation) for North America, Europe, Australia, and portions of Africa and Asia. The Daymet data will be compiled for distribution in 2 degree tiles in CF-1 compliant netCDF format in an equal area projection suitable for each specific region. The data product has been thoroughly documented (Thornton et al. 1997, 2000; Thornton and Running 1999). In addition, the individual meteorological station data used for each tile will be made available in CSV format. New data sources developed under the proposed project may impose distribution restrictions on the raw station data, but we will not engage new data sources unless permission is granted in advance for open distribution of resulting gridded surface weather products. Metadata will be comprised of two formats—contextual information about the data in a text based document and ISO 19115 standard metadata in an xml file. These two formats for metadata were chosen to provide a full explanation of the data (text format) and to ensure compatibility with international standards (xml format). Daymet will follow best practices for compiling metadata (Hook et al., 2010).

Access to data, and data sharing practices and policies

Describe your plans for providing access to data, including websites maintained by your research group and contributions to public databases. If maintenance of a web site or database is the direct responsibility of your group, provide information about the period of time the web site or database is expected to be maintained. Also describe your practice or policies regarding the release of data—for example whether data are available before or after formal publication and the approximate duration of time that the data will be kept private. Describe your policies (where applicable) for protection of propriety data, privacy and confidentiality, intellectual property, or other rights or requirements.

Data products will be generated twice during this project. For each geographical region, we will produce a test version of the data around the mid-point of the project and a final version at the end of the project. The test version will be made available to modeling and applications teams for evaluation and the final version will be provided to the ORNL Distributed Active Archive for long-term data management. During the course of this project, we will generate daily and monthly backups of the data files, which will be retained by the Environmental Data Science and Systems group at Oak Ridge National Laboratory

Policies and provisions for re-use, re-distribution and production of derivatives

Describe your policies regarding the use of data provided via general access or sharing. If you plan to provide data on a website, will the site contain disclaimers, or conditions regarding the use of the data in other publications or products? If the data or products are copyrighted, how will this be noted on the website?

The test Daymet data will be released to the public as soon as our initial quality checks have been completed and the data has been prepared, now planned for the end of Year 1. There is no period of exclusive use by the data collectors. Users can access documentation and Daymet climate data files via the Daymet Web site (daymet.ornl.gov).

Archiving of data

Describe whether and how data will be archived and how preservation of access will be handled. For example, will hardcopy logs, instrument outputs, and physical samples be stored in a location where there are safeguards against fire or water damage? Is there a plan to transfer digitized information to new storage media or devices as technological standards or practices change? Will there be an easily accessible index that documents where all archived data are stored and how they can be accessed? If the data will be archived by a

third party, please refer to their preservation plans (if available).

Our intent is that the long-term high quality data product generated by this project will be available for use by the research and policy communities in perpetuity. The investigators have made arrangements for long-term stewardship and curation at the NASA-funded Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC; see letter of support). The standardized metadata record for the Daymet data will be added to the metadata record database at the ORNL DAAC and NASA Earth Observing System Data and Information System clearinghouse (ECHO), so that interested users can discover the Daymet data along with other related Earth science data. Daymet and the ORNL DAAC have a standardize data product citation (http://daymet.ornl.gov/sites/default/files/Daymet_Reference_CitationDOI_link.pdf) which enables users to cite the source of the data, gives credit to the Daymet team, and enables journal readers to obtain a copy of that product. The ORNL DAAC will include a Digital Object Identifier, which will enable clear citation to the product as well as tracking of use of Daymet.
