World Data Center for Paleoclimatology

Description

The World Data Center for Paleoclimatology is one of 52 World Data Centers that house a wide range of solar, geophysical, environmental, and human dimensions data, and the only World Data Center devoted to paleoclimatic data. We follow standardized guidelines for data archive and distribution developed by the International Council for Scientific Unions (ICSU), supporting the ICSU goal to strengthen international science for the benefit of society, and making all data freely available without restriction.

Our scientific goal is to make paleoclimate data and information as useful as possible in accomplishing the NOAA mission and specifically the NOAA climate goal (understand climate variability and change). NOAA's efforts benefit a diverse audience that includes paleoclimate scientists, non-climate specialist scientists, decision-makers and resource managers, students and educators, and interested citizens. Program staff work closely with national and international science initiatives such as the Climate Change Science Plan, the Intergovernmental Panel on Climate Change Assessment, CLIVAR participants, and other stakeholders to create an archive that meets the needs of each community. NOAA also works with resource planners, decision-makers, and interested citizens to make the paleoclimatology data archive effective and relevant in environmental problem solving and inquiry.

To produce the archive of paleoclimate data needed to extend the instrumental record farther back in time, NOAA decided early on to rely on the contributions from the extramural scientific community around the world, and established the World Data Center for Paleoclimatology in 1992 under the guidance of Jonathan Overpeck. Instead of a large in-house research staff, a small team of science experts, programmers and data specialists work with scientists around the world to archive the results of their published research. This strategy has several advantages. It relies on peer-review as the fundamental basis for quality control, and engages scientists as the experts in paleoclimate data and research. It keeps the archive focused on the latest and most relevant scientific discoveries. Finally, this strategy leverages funding from agencies around the world in the development of a rich, high quality set of observations. It represents an unusual experiment in problem solving by a diverse community. So far, the vigorous use of the data, new publications resulting, and awards and recognition demonstrate the success of this approach.

Special Request for All Users

For all users of the data, we have two requests. First, please cite the original publications when using this data, because this credit is essential to the scientific endeavor. Second, please notify us with problems associated with the data, so that the corrections will benefit others.

Data Management Strategy

The Paleoclimatology Branch's efforts to provide an archive of paleoclimate data rely entirely on the effort of scientists around the world who take the time to contribute the published results of their scientific research, and on the funding agencies who support their efforts. The

users of these data come from different backgrounds, with different needs and goals. We archive information so that it can be used by paleoclimate experts, scientists from other fields, resource managers and decision-makers, students, educators, and interested citizens -- in short, by anyone interested in gaining a better understanding of long-term variations in climate and environment.

Our primary goal is to document each data set so that it can be discovered over the Internet and its suitability for use determined (based on this documentation, users should be able to determine whether a data set meets their needs). To maximize the usefulness and visibility of contributed data, we add FGDC standard content, utilize meta-tags, employ XML and other emerging technologies, and place data on a site with a high search engine ranking. The descriptive information makes data easier to interpret, however the metadata are standardized and concise and are meant to be used in conjunction with original publications, which provide a more complete description of the data.

A secondary goal is to quality-control the data to the extent possible. As data are contributed, all files are assessed to make sure metadata and formatting are correct, and files received and information added are verified with the contributor. Beyond this, a variety of approaches have been tailored for the different proxy records. For example, a program for assessing the quality of the dating in tree-ring data, developed by the Dendrochronology community, is run on incoming tree ring data and results are made available for user assessment. In the case of the paleoceanographic data, gross error checking is used to spot values that lie outside the range of possible values (i.e., negative percentages). We are currently assessing other methods for quality control, including a geographic tool to plot the coordinates listed in the metadata to ensue they match the locations described by other metadata fields. Some procedures used for instrumental climate data may eventually be applied as appropriate but users should realize that quality control for instrumental data is generally more extensive than the quality control applied to paleoclimate data, where uncertainties are generally larger and the accuracy is not always known. One of the biggest problems we face lies in determining whether two quantities were measured in the same way and can be combined as the same variable. We favor the cautious approach and document ambiguous measurements as unique variables. The result for paleoceanography data is five different variables for calcium carbonate percent. But these and other variables can be examined and combined by the user as appropriate for their application. In general, users planning to synthesize measurements must take care when combining measurements from different studies.