EOS

EOS, TRANSACTIONS, AMERICAN GEOPHYSICAL UNION

Climate Data Guide Spurs Discovery and Understanding

PAGES 121–122

Highly accurate and stable observations—beyond those provided by routine weather monitoring—are essential for understanding the behavior of the climate system, developing and validating Earth system models, and attributing extreme weather events and long-term trends to causes [*National Research Council*, 2012; *Trenberth et al.*, 2013]. In parallel with an exploding volume of climate data, ready access to data in user-friendly formats is important to an expanding number and increasing diversity of individuals worldwide across public, private, and academic sectors [*Overpeck et al.*, 2011].

Not only are these users challenged to deal with enormous data volumes, but they also must choose from the multitude of data sets that are available for many variables, from stratospheric ozone to soil moisture [*National Research Council*, 2012]. Faced with such decisions, some users may select data sets based upon convenience of access or familiarity of format rather than suitability for addressing the question at hand. Many researchers and professionals work with data from outside their core field of training and lack the knowledge, time, and motivation to thoroughly survey and synthesize the literature and documentation on unfamiliar data sets. In addition to these challenges, along with increased data accessibility comes a responsibility for users to appreciate the nuances and biases of individual data sets and for data experts to communicate these openly.

How can the hard-earned knowledge about the strengths and limitations of individual data sets best be shared with the community? Funded by the National Science Foundation's Earth System Modeling (EaSM) initiative, the Climate Data Guide (http://climatedataguide.ucar.edu) is being developed as a community-authored guide to the world of climate data sets. The Web site provides concise and reliable information on data strengths and limitations, enabling users to make more informed selections of data sets for their own research and to better understand the data and analysis methods that are encountered in numerous contexts (e.g., in publications, proposals, presentations, classes, Web sites).

Goals, Specifics, and a Case Study of the Climate Data Guide

The unique function of the Climate Data Guide (hereafter "the guide") is the gathering and dissemination of well-informed perspectives on the essential features, strengths, limitations, and utility of a broad range of data sets. The guide actively solicits commentary from both data set developers and experienced users. These perspectives, referred to as "expert-user guidance," address questions such as, "What are the typical research applications of this data set? What are the most common mistakes that users encounter when processing or interpreting these data? What are some comparable data sets, if any? How is uncertainty characterized in these data? How do I best compare these data with model output? Are there spurious (nonclimatic) features in the temporal record?"

Based at the National Center for Atmospheric Research (NCAR), the guide's Web site profiles more than 100 observational climate data sets including atmospheric reanalyses, satellite data, and climate indices. Each profile contains descriptions of a data set's key strengths, limitations, and typical uses, tailored to Earth system model evaluation and general climate research. Metadata (period of record, spatial coverage, formatting, etc.) are provided along with links to download servers, relevant peer-reviewed publications, and product Web sites. To facilitate comparisons, this database-driven Web site generates dynamic lists and tables of data sets organized by different attributes, such as by variable, Earth system component (atmosphere, cryosphere, ocean, etc.), satellite platform, utility for model diagnostics, or designation in a special category such as Climate Data Record [e.g., *National Research Council*, 2004].

To illustrate the utility of the expert-user guidance entries, consider the case of cloud fraction data from the International Satellite Cloud Climatology Project (ISCCP) [*Pincus and National Center for Atmospheric Research Staff*, 2013, and references therein]. Cloud fraction information is invaluable to climate scientists studying Earth's radiation budget, as well as to renewable energy developers interested in a region's solar energy potential and many others. As summarized under "Key Strengths" on the guide's ISCCP Web page, ISCCP is a high-value data set due to its 3-decade length and global, sub-daily sampling.

One issue, however, is that the comparison of cloud amounts across regions can be an apples-to-oranges problem because the inconsistent viewing angle and changing instrumentation of geostationary satellites create artifacts in ISCCP's spatial patterns. Further, as cloud changes are one of the main uncertainties in climate change projections, assessing the performance of climate models with data sets such as ISCCP is important in gaining confidence in the projections. However, even if analysts apply cloud simulators to the model output, ISCCP data pose special challenges due in part to the treatment of partly cloudy pixels, as highlighted in another expert-user guidance entry [*Kay and National Center for Atmospheric Research Staff*, 2013].

While issues like these are often well known among a small circle of researchers, it has been difficult to communicate them to the expanding community of users. The guide thus plays a key role in disseminating vital end user information. Before posting reviews, metadata, and links, the guide verifies their general accuracy and validity. However, users are encouraged to post comments and questions about the reviews and to report broken links and/or errors in the metadata. Users seeking more in-depth analyses and further details can follow links from the guide to supporting peer-reviewed literature or to data set–specific Web sites.

As the guide places emphasis on community input and focuses on a limited selection of data sets that are most useful for large-scale climate research and model evaluation,

By D. P. Schneider, C. Deser, J. Fasullo, and K. E. Trenberth

it serves a different purpose from typical data repositories or directories, for example, the National Climatic Data Center (http://www.ncdc.noaa.gov) or the Global Change Master Directory (http://gcmd.nasa.gov). Moreover, data set profiles in the guide are presented in a consistent, accessible style and are not limited to the products of a single agency or funding call. The guide incorporates the broad range of data sets in common usage among climate researchers.

Highlights of the Guide's First Operational Year

2012 was the first full calendar year of operation for the guide. Usage was tracked with a commonly used log analyzer (AWStats; http://awstats.sourceforge.net) as well as Google Analytics (http://www.google.com/analytics). Additional information on user demographics and interests was gathered through an optional registration process (https://climatedataguide.ucar.edu/user/register) that enables users to post comments and upload materials to be featured on the Web site. In 2012 there were in excess of 40,000 unique visitors, more than 300,000 page views, and nearly 2.3 million hits. Visitors came from more than 100 countries, with 36% of traffic from the United States. The United Kingdom, China, Germany, and Canada rounded out the top five countries contributing Web traffic. Nearly half of visitors came via search traffic, while 30% came through referrals from established programs within NCAR as well as from external programs with which the guide has made efforts to interface, for example, the Coupled Model Intercomparison Project (http://cmip-pcmdi.llnl.gov).

The guide attracted nearly 600 registered users in 2012. Their range of occupations and nationalities supports assertions of an expanding and diversifying market for climate data and knowledge [e.g., *Overpeck et al.*, 2011]. The list of registered users includes graduate students and professors, scientists and engineers in the public and private sectors, and analysts in fields such as insurance and renewable energy. Based on usage trends so far, the user base is expected to grow significantly in 2013, with a projected 50–100% growth in site visits over 2012.

Community Involvement in Improving the Guide

The guide is a cost-effective solution to calls by both scientists [e.g. *Overpeck et al.*, 2011] and funding agencies [*National Science Foundation*, 2013] to share climate data and research methods transparently, freely, and effectively. Each contributor of expert-user guidance is acknowledged, and individual pages and reviews are citable. Text, figures, publication lists, and metadata may be uploaded to the Web site or e-mailed to climatedataguide@cgd.ucar.edu.

Registered users are alerted to new content via e-mail notifications, although anyone may follow updates through Really Simple Syndication (RSS) feeds and, in the future, through social media. In addition, to better serve its diverse and growing audience, project leaders plan to update the guide by improving its Web design, search tools, and user interfaces; reaching out across social media and mobile platforms; and, most important, incorporating broad input from the climate research community.

Currently, the most comprehensive coverage on the guide is for atmospheric and surface climate variables, including those used to diagnose the performance of the Community Atmosphere Model, a component of NCAR's Community Earth System Model (CESM). Most major model diagnostic data sets used by CESM and other Earth system models will be incorporated into the guide over the next year. Other categories to be expanded are boundary condition and forcing data sets used in historical climate simulations, satellite data sets, and reprocessed data sets of broad interest, such as those considered in the Climate Data Records and the Global Energy and Water Cycle Experiment projects [see *Trenberth et al.*, 2013; http://www.gewex.org].

Through these efforts, data sets and publications profiled on the guide will become more accessible and visible. However, the success of this endeavor hinges upon community participation.

Strong participation in this vital community resource will help to ensure that researchers are well equipped to address the complex questions that pertain to past, present, and future climate. The Climate Data Guide invites scientists to add themselves to the growing list of expert contributors (https://climatedataguide.ucar.edu/experts-list).

References

Kay, J., and National Center for Atmospheric Research Staff (Eds.) (2013), The Climate Data Guide: COSP: Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package, Natl. Cent. for Atmos. Res., Boulder, Colo. [Last modified 8 January 2013. Available at http://climatedataguide.ucar.edu/guidance/cosp-cloud-feedback-model-intercomparison-project-cfmip-observation-simulator-package.]

National Research Council (2004), *Climate Data Records from Environmental Satellites*, 150 pp., Natl. Acad. Press, Washington, D. C.

National Research Council (2012), *A National Strategy for Advancing Climate Modeling*, 280 pp., Natl. Acad. Press, Washington, D. C.

National Science Foundation (2013), Dissemination and sharing of research results, in *Award and Administration Guide*, chap. VI.D.4, Washington, D. C. [Available at http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/aag_6.jsp#VID4.]

Overpeck, J. T., G. A. Meehl, S. Bony, and D. R. Easterling (2011), Climate data challenges in the 21st century, *Science, 331*, 700–702, doi:10.1126/science.1197869.

Pincus, R., and National Center for Atmospheric Research Staff (Eds.) (2013), The Climate Data Guide: ISCCP: International Satellite Cloud Climatology Project, Natl. Cent. for Atmos. Res., Boulder, Colo. [Last modified 4 January 2013. Available at http://climatedataguide.ucar.edu/guidance/isccp-international-satellite-cloud-climatology-project.]

Trenberth, K. E., et al. (2013), Challenges of a sustained climate observing system, in *Climate Science for Serving Society: Research, Modeling and Prediction Priorities*, edited by J. W. Hurrell and G. Asrar, Springer, Berlin, in press.

—David P. Schneider, Clara Deser, John Fasullo, and Kevin E. Trenberth, National Center for Atmospheric Research, Boulder, Colo.; E-mail: climatedataguide@cgd.ucar.edu