

A number of critical points were addressed (see Girardin and Mudelsee, in press) to evaluate the robustness of these findings. First, proxy representativeness is not perfect. For example, influences of spring and autumn DC values have been ignored. Tree ring series pre-processing (removal of long-term physiological growth signals) may also hide long-term climate trends. Second, feedbacks from vegetation, regional climate systems or other forest disturbances, were not accounted for. Third, emission scenarios are technically “guesses” with unquantified, presumably large, error bars. Fourth, extrapolation errors may have affected the FireOcc calibration as well as the climate model. The confidence bands should capture a portion of these uncertainties. On the other hand, selection of statistical parameters in figure 2 is rather uncritical, as a sensitivity analysis showed.

We deliberately use the word “prediction” instead of “projection” because the latter’s connotation of an assumption-free analysis is misleading. Every analysis

of future situations necessarily makes assumptions (actualism). The task is rather to quantify and include the error sources into the analyses. In this regard, it is important to evaluate climate model uncertainties. For this objective, the ensemble method is insufficient because it ignores parameterization uncertainties (e.g., hydrological cycle). The bootstrap simulation method (Challenor, 2004) has the potential to include parameterization and also emission uncertainties but its full implementation requires a leap in computing power.

The predicted increases in fire risk may lead to considerable increases in wildfire management costs, offset the influences of elevated temperatures and atmospheric CO₂ concentrations on forest and tree productivity, and affect the availability of harvestable trees (Girardin and Mudelsee, in press). More frequent, large wildfires may also become a major factor in our changing climate, owing to greater carbon losses that could feed the warming (Kurz et al., 1995; Flannigan et al., 2006).

Acknowledgements

We thank Travis Logan for providing ECHAM4 data and Brian Stocks for the photo. Financial support from Canadian Forest Service funds (M.P.G.) is acknowledged. Fire data, software and risk estimates are available from Climate Risk Analysis (www.mudelsee.com).

References

- Girardin, M.P. and Mudelsee, M., in press: Past and future changes in Canadian boreal wildfire activity, *Ecological Applications*.
- Girardin, M.P., Bergeron, Y., Tardif, J.C., Gauthier, S., Flannigan, M.D. and Mudelsee, M., 2006: A 229-year dendroclimatic-inferred record of forest fire activity for the Boreal Shield of Canada, *International Journal of Wildland Fire*, **15**(3): 375–388.
- Mudelsee, M., 2006: CLIM-X-DETECT: A Fortran 90 program for robust detection of extremes against a time-dependent background in climate records, *Computers and Geosciences*, **32**(1): 141–144.
- Mudelsee, M., Børgen, M., Tetzlaff, G. and Grünwald, U., 2003: No upward trends in the occurrence of extreme floods in central Europe, *Nature*, **425**(6954): 166–169.
- Stocks, B.J. et al., 2003: Large forest fires in Canada, 1959–1997, *Journal of Geophysical Research*, **108**(D1): 8149, doi:10.1029/2001JD000484.

For full references please consult:

www.pages-igbp.org/products/newsletter/ref2008_2.html



What can data tell us about past climate that is useful for the future? Data management in paleoclimatology

NICOLAS DITTERT¹, D.M. ANDERSON² AND H. GROBE³

¹MARUM, University of Bremen, Germany; ndittert@uni-bremen.de

²NOAA World Data Center for Paleoclimatology, Boulder, USA

³Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

During the last decades, the number of scientific publications has increased exponentially, as has the number of scientific data sets. No end to the increase is in sight. Under the constraints of an efficient cost-benefit publication system, editors have had to tighten the measures for acceptance of manuscripts (e.g., maximum number of pages, and rigorous selection criteria). Data tables have often been relegated to appendices or supplementary material. With increasing availability of publications through the internet, unique and persistent digital object identifiers (DOIs) have been invented to keep up with the flood of publications. For a long time, scientific data have not been considered within this design strategy. However, the increasing amount of scientific data calls for appropriate archiving and international availability.

Today, this concept is forcefully supported by the Recommendations of the Commission on Professional Self Regulation in Science (1998), the Berlin Declaration on Open Access to Knowledge in the

Sciences and Humanities (2003), the white paper on the free access to scientific data by the Association of Learned and Professional Society Publishers (ALPSP, 2006), and the OECD Principles and Guidelines for Access to Research Data from Public Funding (2007). The vision of the World Data Centre system is the optimized exchange of scientific information leading towards significant and sustainable support of future scientific achievements, as scientific data are perceived as the backbone of discovery. Combined with methods, results, interpretation, and discussion, data sets create a useful scientific publication. However, data have stand-alone value, and the unconfined availability of (and access to) the data allows for broader use and novel scientific debate.

The value of paleoclimate data to societally relevant problems in climate and environmental change is now well known. To make these data widely available, two remarkable archives have been developed: (1) the PANGAEA® Publishing Network for Geoscientific and Environ-



WDC for Paleoclimatology



mental Data (Bremerhaven, Germany), which serves as the technical archive for the World Data Center for Marine Environmental Sciences (WDC-MARE), and (2) the NCDC World Data Center for Paleoclimatology WDC-Paleoclimatology in Boulder, USA. WDC-Paleoclimatology describes its mission as providing “the paleoclimate data and information needed to understand and model inter-annual to centennial-scale environmental variability” PANGAEA® campaigns with its long-term and secured archiving structure, highly efficient editorial system, and extensive interoperability with other international data centers and portals. WDC-Paleocli-

matology evolved to become a source of data and information for specific themes, such as drought and abrupt climate change PANGAEA® has “special emphasis on environmental, marine and geological basic research”, with the potent notion of being an exemplary publication and library system for scientific, geo-referenced data. Whether you prefer the one or the other World Data Centre may depend on your personal needs, or your fondness for (1) more focused data collections (WDC-Paleoclimatology), or (2) data mining in a billion data points warehouse (PANGAEA®/WDC-MARE).

To realize the digital library-of-data concept, data sets are perceived as data entities. A data entity consists of meta-information and data. Meta-information is any information describing a data set. Data is the pure, primary scientific information, which can be numbers, text, graphics, logging, audio and video recording and reproduction, etc. Where a dendrochronological record provides a few bytes of data only, CTD profiles can deliver some kilobytes of data, and satellite information measure beyond the megabyte border. Whether a single data point is recorded or a gigabyte mass data stream, it is not size that matters. It is rather the data set's scientific value and its unconfined availability (cf., panFMP, Schindler and Diepenbroek, 2008), and the standard assignment of one unique

persistent, bibliographic identifier per data set to turn a plain data set into an autonomous data publication, cross-referenced with its original scientific paper. Data archiving is carried out in close coordination with the principal investigator. Owing to networking with other systems, database contents can be tracked down by means of search engines, portals, and online library catalogs. The technique of data citation gives a strong motivation for scientists to publish their data, which in the long range will improve data quality and availability.

Some data publications relate to individual papers or studies, while others are vast compilations and syntheses that receive periodic updates. For example, the Climatological Database for the World's Oceans (CLIWOC; Gallego et al., 2005) with climatological and meteorological observations from ship logbooks between 1750 and 1854, contains some 5000 data sets. Based on 36 original scientific papers, Anderson and Mulitza (2001) compiled a 7791 digit set of $\delta^{18}\text{O}$ data from planktic foraminifers in surface sediments. As part of the Paleoclimate Database of the Quaternary (PKDB), Frenzel et al. (2001) document the atlas of paleoclimates and paleoenvironments of the northern hemisphere. The International Tree Ring Database is continually updated, and provides the field with not only ten thousand raw ring-width measurements but also stan-

dardized chronologies and data quality statistics. Some data publications are remarkable for their unusual content. Müller et al. (2005) have archived harmonic tremor signals of the so-called “singing iceberg”.

Paleo data give us the window on the past. Beyond their presentation in individual publications, paleo data form a rich tapestry of the four-dimensions we inhabit (time, latitude, longitude, and elevation). They tell us when, for how long, and for what reasons, the climate has changed. Digital data libraries make it all possible.

References

- Anderson, D.M. and Mulitza, S., 2001: Compilation of delta ^{18}O data from planktonic foraminifera in surface sediment, *Publishing Network for Geoscientific and Environmental Data*, doi: 10.1594/PANGAEA.60896.
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, 2003: Berlin, 4.
- Müller, C., Schindler, V., Eckstaller, A. and Miller, H., 2005: Audio record of the 'singing iceberg' from the Weddell Sea, Antarctica, *Publishing Network for Geoscientific and Environmental Data*, doi: 10.1594/PANGAEA.339110
- Organisation for Economic Co-operation and Development, OECD, 2007: *Principles and Guidelines for Access to Research Data from Public Funding*, OECD Publishing, Paris, 24.
- Schindler, U. and Diepenbroek, M., in press: Generic framework for metadata portals, *Computer and Geosciences*.

For full references please consult:

www.pages-igbp.org/products/newsletter/ref2008_2.html



PMIP (Paleoclimate Modelling Intercomparison Project)

PASCALE BRACONNOT¹ AND SANDY P. HARRISON²

¹Laboratory for Climate and Environment Sciences, Gif-sur-Yvette, France; pascal.braconnot@lsce.ipsl.fr

²School of Geographical Sciences, University of Bristol, UK; sandy.harrison@bristol.ac.uk

The Paleoclimate Modelling Intercomparison Project (PMIP) is a long-standing initiative endorsed by both the WCRP/CLIVAR (World Climate Research Programme/Climate Variability and Predictability) Working Group on Coupled Modelling (WGCM) and PAGES. It has provided an efficient mechanism for coordinating paleoclimate modeling activities, which provide valuable information on the mechanisms of climate change, the identification of key feedbacks operating in the climate system, and the capability of climate models to reproduce climates different from today. Thanks to the production of data syntheses and to rigorous model-data comparisons, the mid-Holocene (ca. 6 kyr BP) and the Last Glacial Maximum (LGM; ca. 21 kyr BP) are now recognized as benchmark periods for climate models. Although the main focus is on model-model intercomparison and evaluation, PMIP has

acted as an important discussion forum, promoting the understanding of past climate changes as a necessary basis for confidence in future predictions. As a result, PMIP has contributed significantly to the last two IPCC assessments.

In the last 10 years, climate models have moved from atmosphere-only to coupled ocean-atmosphere models and ocean-atmosphere-vegetation models. Models that include the coupling between the physical climate and biogeochemical cycles, such as the carbon cycle, have also been developed. These couplings, and the corresponding feedbacks, shape the response of the climate system to external variations. They are required to enable understanding of how climate has evolved through time and how it will evolve in the future in response to human activities. The second phase of PMIP (PMIP2) was launched in 2002 and addresses the

role of the different feedbacks using these coupled models (Harrison et al., 2002; Crucifix et al., 2005).

All the information to run a PMIP2 simulation is available on the PMIP2 website (<http://pmip2.lsce.ipsl.fr/>; see Braconnot et al., 2007 for an overview). Model results are stored in a common database hosted at LSCE on raid disks and the data is distributed through a Linux file server. Guidelines, file format convention, variable names and structures, and utilities follow the requirements of the WCRP CMIP3 multi-model dataset. Participation in PMIP analyses is an open process. About 80 sub-projects are now registered and most of them have already produced publications in high-level international journals (see, e.g., Fig. 1, and Otto-Bliesner and Brady, p. 18-20 this issue). Several data syntheses have also been released through the website, as well as a subset of maps showing