Last Millennium Reanalysis
First Community Hackathon

October 4, 2017

Summary

Paleo data assimilation as exemplified by the LMR framework[1] is a powerful new tool enabling new lines of research. This power comes at the cost of high complexity, so it is important to train a new generation of scientists to use them effectively. This objective is all the more urgent as the latest NOAA CPO OOM funding call[2] is "*specifically soliciting projects that exploit or evaluate the CPO-funded Last Millennium Climate Reanalysis (LMR) products and the resources housed by NOAA at the World Data Service (WDS) for Paleoclimatology*". To help early-career scientists to be competitive in this competition, and more generally to accelerate the scientific transition towards paleoclimate reanalysis products, the LMR team proposes a 1-day hackathon designed to build capacity with LMR resources. The hackathon will happen on the heels of the Third Annual LMR workshop, to take place Oct 2-3 in at the National Center for Atmospheric Research's Mesa Lab in Boulder, CO, and will be primarily geared towards early career scientists and graduate students who otherwise would not be able to attend the workshop.

As with all hackathons, the training will be entirely hands-on. The morning session will consist of Python-based exercises to run LMR experiments and perform basic numerical and visual diagnostics. The afternoon session will be based on individual or small group research projects tailored to participant goals and needs, and will focus on the interpretation of LMR results. Expecting 20 participants, we request $20k to support this activity (see Budget).

Objectives

The availability of structured compilations of paleodata and new methods for the objective synthesis of proxy data with information from numerical models through data assimilation enable new lines of inquiry in low-frequency climate dynamics. The goal of this hackathon is to encourage and promote the use of LMR's data assimilation tools within the (paleo)climate community. The hackathon is intended for early-career scientists and graduate students who have some experience with high-level programming languages (e.g., Python, Matlab, R). The goal is for participants with Python experience to run their own LMR experiments after the morning session; others should be able to visualize and have basic notions of how to interpret LMR output by day's end.

Participants will be asked to make suggestions regarding the LMR documentation and configuration to improve its use by others. The afternoon sessions are less structured, and will focus on exploring the spatial patterns of temperature and precipitation through time, the temperature response following volcanic eruptions, and the construction of common climate indices. Participants will be encouraged to use their preferred language in the afternoon to produce and share analyses and visualizations. In the final session participants will compete in small teams in the visualization challenge to produce the most compelling visualization of last millennium climate. An award will be presented to the team that creates the most compelling visualization of the last millennium climate.

[1] http://onlinelibrary.wiley.com/doi/10.1002/2016JD024751/full
[2] https://www.grants.gov/web/grants/search-grants.html?keywords=climate%20program%20office

Facilitators

Julien Emile-Geay (lead) is Associate Professor of Climate Dynamics at the University of Southern California, where several of his courses emphasize data science. He is a principal investigator on the Linked Earth Project, principal investigator on the Last Millennium Project, and participant in the PAGES data stewardship initiative as well as the PAGES 2k, Iso2k, and VICS working groups.

Michael Erb is a postdoctoral fellow at the University of Southern California Climate Dynamics lab, an LMR user and contributor studying drought and climate dynamics.

Robert Tardif is a Research Scientist at the University of Washington, with expertise in data assimilation applied to weather and climate. He has played a central role in the development of the LMR code infrastructure.

Andre Perkins is a graduate student in Atmospheric Sciences at the University of Washington studying ocean-atmosphere variability using an online assimilation methodology. He has played a central role in the development of the LMR software.

Kaleb Horlick is is a graduate student in Geology and Geophysics at Oregon State University, and LMR user studying tropical circulation variability.

Jessica Badgeley is a graduate student in Earth and Space Sciences at the University of Washington, and LMR user studying past climates of Greenland.

Prerequisites

Participants should have some experience programming (Python, R, Matlab, NCL, or similar) to analyze and visualize scientific data. To take advantage of the first sessions (using the DA code) it is highly recommended that participants complete at least a basic Python tutorial[3]. We will provide a self-test exercise prior to registration. Participants lacking python experience but with experience with other languages will benefit from the afternoon session (visualization) and their participation is welcome. Python knowledge is not a requirement to make use of LMR output, though it helps.

Each participant needs a computer with a Python distribution installed to participate in the hackathon. Participants are encouraged to install and configure the Anaconda[4] Python 3.5+ and the modules required by the LMR prior to the hackathon.

Materials

Sphinx documentation of LMR code; PAGES2k 2.0.0 data documentation

Travel Support

We hope to provide travel support for 20 participants. Support will be limited to participants who intend to register for both the workshop and the hackathon. Priority will be given to early career scientists and graduate students.

[3] http://earthpy.org/tag/python-for-geosciences.html

[4] https://www.continuum.io/downloads

Proposed Schedule

Time	Activity
9:00 - 9:15	Introduction
9:15 - 9:45	**Running the LMR (basic)** : with default configuration (30min)
9:45 - 10:45	**Running the LMR (intermediate):** Changing the configuration to work with different proxy data sets, different calibration data sets, and different numerical simulations, as well as producing different climate outputs. (1h)
10:45 - 11:00	Coffee Break
11:00 - 11:30	**Verification and uncertainties** (30min) Participants will explore verification metrics of the LMR output compared to instrumental data sets. Discussion will focus on the quantification of uncertainty.
11:30 - 1:00	**Visualization** (1.5 hours). Introduction to LMR analysis focusing on two examples: 20th century temperature change and temperature response to volcanoes. [NOTE: if isotope fields are available by the hackathon, comparison of un-assimilated proxy records to LMR output would be a interesting topic to focus on, instead of one of the two examples mentioned.]
1:00 - 2:00	LUNCH at Mesa Lab cafeteria
2:00 - 4:00	**Analysis challenge** (teams) (2 hours). Participants will work in small teams to identify an aspect of the past millennium climate and develop a visualization (an image, a time series, a map, an animation) that exploits the full-spatial fields, multi-variable, uncertainty-attached aspects of data assimilation products. Participants will be encouraged to consider potential analyses prior to the workshop, so they can focus on topics that most interest them. An award will be presented to the team that creates the most compelling visualization of the last millennium climate.
4:00 - 4:30	**Afternoon Break**
4:30 - 5:30	**Rejoinder** (1 hour): teams present the result of their analysis (~5 teams @ 10 minutes each)
5:30 - 6:30	**Discussion** (1 hour): Discussion about the LMR approach and data assimilation in general, including further discussion of the user analyses presented in the last session. Discussion may also focus on future research directions and suggestions about code features, data products, and community engagement.
	DINNER at TBD + award ceremony