CESM Diagnostics Lab

Turning numbers into pictures

Jesse Nusbaumer
*Climate & Global Dynamics, NSF NCAR*

Aug 9, 2024

- Learn about (geoscientific) python.

- Successfully run Jupyter notebooks on Jupyterhub.

- Visualize CESM model output

- Learn about diagnostic software being developed at NCAR, specifically CUPiD, which will be CESM3's default post-processing and diagnostics package.

Python

The analysis of CESM output is mostly done in a python ecosystem here at CGD.

Python is an interpreted, multi-paradigm language used for a myriad of purposes (for example the command line tools like "create_newcase" are actually python scripts).

There is "standard" python, and then various libraries and modules which have been developed for specific purposes, like working with scientific datasets and visualizations. Some of the extra libraries we will be using are shown to the right.
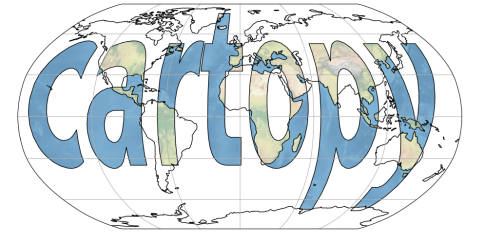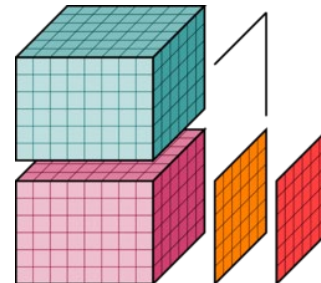
Jupyterhub

Jupyter notebooks are a way to organize text, figures, and cells with executable code. They are increasingly used by scientists here in CGD to develop and share codes and scientific analyses.

Jupyterhub is a system provided by NCAR CISL to create and run Jupyter notebooks on CISL machines (Casper and Derecho). Anyone with derecho or casper access should also have access to the Jupyterhub system and are free to use it.

1. Go to the Diagnostics section of the tutorial (https://ncar.github.io/CESM-Tutorial/notebooks/diagnostics/diagnostics.html)

1. Go to the Diagnostics section of the tutorial (https://ncar.github.io/CESM-Tutorial/notebooks/diagnostics/diagnostics.html)

1. Login to Casper or Derecho, and perform **Step 1** (git clone).

Jupyterhub login process

1. Go to the Diagnostics section of the tutorial (https://ncar.github.io/CESM-Tutorial/notebooks/diagnostics/diagnostics.html)

1. Login to Casper or Derecho, and perform **Step 1** (git clone).

1. At the beginning of **Step 2**, go to https://jupyterhub.hpc.ucar.edu/ and then WAIT.

Jupyterhub login process

1. Go to the Diagnostics section of the tutorial (https://ncar.github.io/CESM-Tutorial/notebooks/diagnostics/diagnostics.html)

1. Login to Casper or Derecho, and perform **Step 1** (git clone).

1. At the beginning of **Step 2**, go to https://jupyterhub.hpc.ucar.edu/ and then WAIT.

1. Once a lab helper has given your table/section the ok, click "production", add in your username then enter your password + "," + your Duo Pin code (DO NOT ENTER IN JUST YOUR PASSWORD)!

For example, if my CIT password was "abcd", and my DUO pin on my phone was "123456", then I would enter: abcd,123456

Jupyterhub login process

1. Go to the Diagnostics section of the tutorial (https://ncar.github.io/CESM-Tutorial/notebooks/diagnostics/diagnostics.html)

1. Login to Casper or Derecho, and perform **Step 1** (git clone).

1. At the beginning of **Step 2**, go to https://jupyterhub.hpc.ucar.edu/ and then WAIT.

1. Once a lab helper has given your table/section the ok, click "production", add in your username then enter your password + "," + your Duo Pin code (DO NOT ENTER IN JUST YOUR PASSWORD)!

1. Once you have successfully logged in you can continue on with the tutorial at any pace you choose.

CUPiD

The CESM Unified Postprocessing and Diagnostics (CUPiD) system is a new, push-button diagnostics system which will automatically produce diagnostic output for all CESM3 components for any given CESM3 run.

There is a CUPiD notebook available to try in the official diagnostics exercises. You can also learn more about CUPiD on Github here:

https://github.com/NCAR/CUPiD.git

For CESM3 the default grids for the atmosphere and land components will be unstructured, and so having analysis and visualization tools that natively handle or regrid unstructured data will be critical.

One project to help with unstructured data is UXarray. UXarray is a new python package/extension of Xarray that allows one to work with unstructured grids. It is currently under development here at NCAR (via CISL). You can learn more about it online here:

https://uxarray.readthedocs.io/en/latest/

One can try an example workflow with UXarray under the "UXarray" section, which is itself under "Additional Topics".
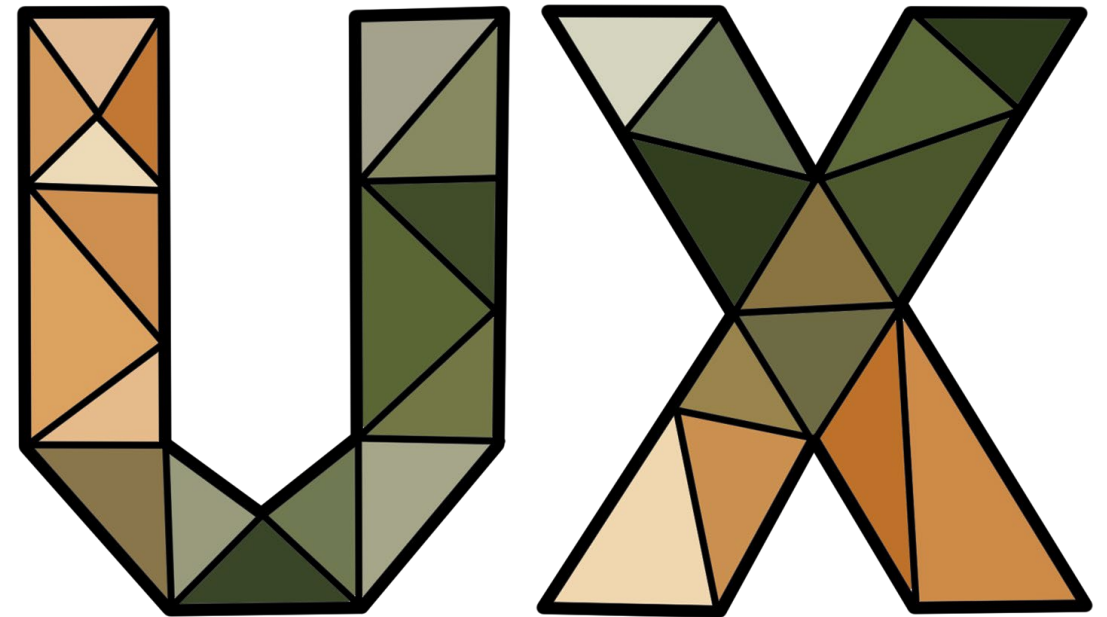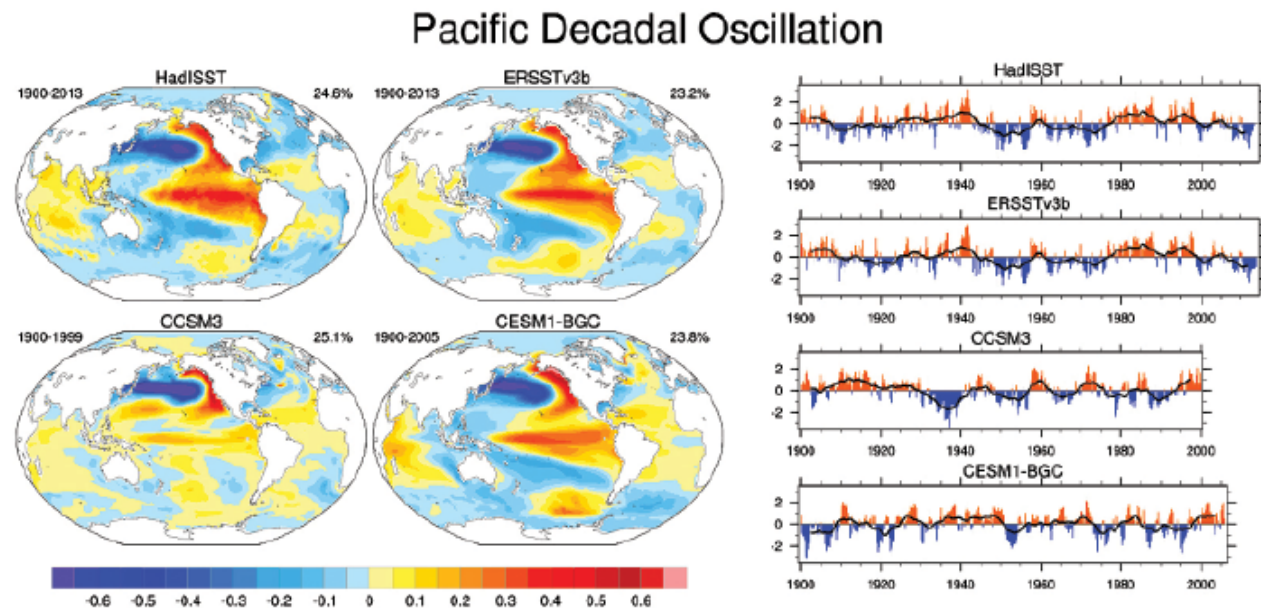
The Climate Variability Diagnostics Package (CVDP) is a diagnostics tool developed in CGD's Climate Analysis Section that is designed to help automatically generate diagnostics related to various modes of variability in a CESM run. There is also a large ensemble version (CVDP-LE) that can do the same for an ensemble of CESM simulations. Both packages are currently written in NCL, but are actively being moved to Python.

Information on the CVDP can be found online here:

https://www.cesm.ucar.edu/projects/cvdp

While information on the CVDP-LE can be found online here:

https://www.cesm.ucar.edu/projects/cvdp-le

Pacific Decadal Oscillation

The AMWG Diagnostics Framework (ADF) is a new python-based diagnostics system designed to replace the older AMWG NCL diagnostics, or in other words to post-process CAM (atm) output and generate a large collection of plots and analyses that are organized into a website for you.

The ADF can also run the CVDP as well, and has been coupled to CUPiD.

Additional information, including examples to try, can be found on the ADF tutorial page online here:

https://justin-richling.github.io/ADF-Tutorial/README.html

Thanks for listening!

Any questions?