Climatology Calculation Support in the GeoCAT Ecosystem

Part of the Pivot to Python

Heather R. Craker,
Purdue University, SIParCS/GeoCAT Intern
July 27, 2021
NCL functionality needs to be added to the GeoCAT ecosystem

• Visualization
  – Example gallery (GeoCAT-examples)
  – Wrapper classes (GeoCAT-viz)

• Computation (GeoCAT-comp)
  – Working with different kinds of model data (i.e. CAM-SE)
  – Interpolation of data
  – Handling data on non-rectangular mesh grids
  – Calculating climatological averages
What is a climatological average?

January average from 2020 to 2050

February average from 2020 to 2050
What do users need and want?

• User Stories
  – Asked for input on Zulip
  – Contacted people working on the Earth System Data Science (ESDS) Initiative

• Feedback
  – Different ways to handle the output datetimes
  – Weight monthly data when finding seasonal means

2020
December

2021
January

2021
February

Winter ‘20-'21

December 1st, 2020?
February 28th, 2020?
January 15th, 2021?
Calculates long term hourly, daily, monthly, or seasonal averages across all years in the data

• Inputs:
  − The data as an xarray.DataArray or xarray.Dataset
  − A string representing the frequency of the calculated climatological averages
    • ‘hour’, ‘day’, ‘month’, ‘season’
  − The name of the time dimension
    • Optional, will be inferred if not given

• Output:
  − Climatological averages for each period

```python
def climatology_average(
    dset: typing.Union[xr.DataArray, xr.Dataset],
    freq: str,
    time_dim: str = None) -> typing.Union[xr.DataArray, xr.Dataset]:
```
Jupyter Notebook Demo
What’s next?

Implement NCL Functions

Incorporate User Feedback

Create New Tools
Thank you for listening!

Special thanks to my mentors and collaborators this summer:
Alea Kootz, Orhan Eroglu, Max Grover, Anderson Banihirwe, and Deepak Cherian

GeoCAT Website
https://geocat.ucar.edu/

GeoCAT-comp GitHub
https://github.com/NCAR/geocat-comp