introduce the datasets of COLA

1. Model NMME North Amercia MultiModel Ensemble

NMME CMC1-CanCM3 Forecast Model

Datasets of model outputs

Update on a regular bases

Download via script, python code

2. Global Modeling and Assimilation office

Reanalysis frequently used in climate model, collect all the historical observations from that time, reanalysis help fill in the missing points, focus on taking observation from arctic to form a more continuous and reasonable dataset

Full reanalysis can run to 100tb, surface side analysis, access and transfer of that dataset

3. CPC Global unified cauge-based analysis of daily precipitation

Write a script

Download CPC\_uni product

Create a index

Correlate with another fields and see how the patterns merge

Correlation Between rainfall index and other fields

Teleconnection is a type of analysis and now just run on a subsetted dataset

Useful on longer time series of datasets, like look for temperature five or ten days away

Seasonal statistics, influences more slowly varying can help predict the long term event

Two months average use GraDAS to calculate, the software is easy to use, use python to direct calculate inde, netcdf, single time series,

Facilitate analysis first with the basic steps for the teleconnection model

4. Extended reconstructred sea surface temperature (ERSST) v5

Assessment of multiple rainfall datasets is very important. The disagreement of datasets could cause complete different conclusions. Depend on sensor networks.

El Nio

Other fields like surface temperature, regions around bowl, single grid points against the other points in the domain,

Teleconnection analysis

GrADS script, need be integrated into CyberConnector

Several steps are performed with multiple scripts

Open file -> set variables -> average -> select fields I want -> compose four dimensional data -> chop out unwanted part-> integrate into one time series

Shell script

Used a lot of netcdf-related commands

Each file comes out with zero month time step, overwrite time variable with 0, resetting the time

Tell CTL that is a netcdf or grib, GrADS can detect the file type.

Average across months or other time duration is a frequently used tool.

Extensive GrADS community, a tool widely used in that group. Jennifer is the developer of GrADS.

Besides the correlation map, what else do you do with the results? One-point correlation map;

It is hard to say if it is statistical that correlating surface temperature with anything can result in high odds. Strength of correlation may tell something that is meaningful.

Python, shell, GrADS,

Consistency of models, diviations , detect where the models vary widely, look into the graphs and plots can find some pattern or interesting things.

California drought sees a obivious no-increase in rainfall, look into those associations.

El Nio and rainfall association with sea surface temperature,

Support the language that everyone is familiar with, as few changes on user habits as possible.

Python is important. GrADS is pushed to be more like python. Must interface with python.

Waste of space if multiple people writing the same model ,

Show the scripts easily and version control of the scripts,

Annual cycle VDP, fixed thing that people can call, It should be easy to diagnose and inspect

Search API client

Not necessary to download data to the laptop, it should be able to download directly from CyberConnector to the server. Only Figures to the laptop, mostly everything is on the server.

Find the data Id and download it use the API.

Input data

NASA press release, peer-reviewed papers, some data from co-workers, recommendation from other people is the major way to discover.

FTP may be the best way to transfer data

Option out is needed

Institutional data and knowledge, where the data is and how we can get it? how to process them? Why to do it like this? Path to solution that scientists find. Sitting in UCAR, no idea how to discover the datasets, discovering dataset is still significantly missing function in science community and need be promtoed.

This project is not creating new chain than the old chain, but make the workflow more efficient and repeatable. Opportunity to , seamless, flexibility,