CICompass

Funded by the National Science Foundation, Grant #2127548

Data CI Pilot: NCAR and NEON cyberinfrastructure collaborations to enable convergence research linking the atmospheric and biological sciences

by Gordon Bonan

Cyberinfrastructure for Major Facilities: Challenges and a Path Forward

March 2, 2022
CI Compass Cyberinfrastructure for NSF Major Facilities Workshop

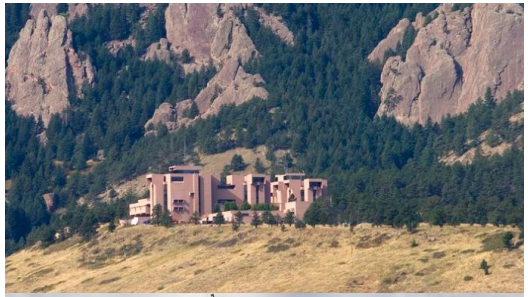
USC Viterbi School of Engineering Information Sciences Institute
INDIANA UNIVERSITY
UNIVERSITY OF NOTRE DAME
THE UNIVERSITY OF UTAH
TEXAS TECH UNIVERSITY
renci

Data CI Pilot: NCAR and NEON cyberinfrastructure collaborations to enable convergence research linking the atmospheric and biological sciences

A collaboration among NSF's CISE, BIO, and GEO directorates

Gordon Bonan
National Center for Atmospheric Research
Boulder, Colorado

NSF Cyberinfrastructure for Major Facilities Workshop
March 2, 2022



NCAR is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977

Bringing **New Voices** to Climate Science

> “An example of a synergy among facilities is that identified by the National Center for Atmospheric Research (NCAR) and the National Ecological Observatory Network (NEON) to advance the capability of Earth system prediction to include terrestrial ecosystems and biological resources.”
>
> *Next Generation Earth Systems Science at the National Science Foundation* (p. 68)

NCAR

- Models the land and its terrestrial ecosystems
- Observes the land-atmosphere system and PBL
- Provides model and observational data to the scientific community

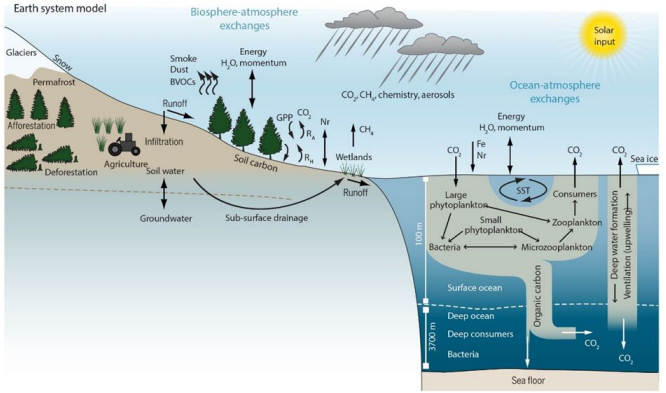
NEON

- Observes and monitors the land and its terrestrial ecosystems
- Provides observational data to the scientific community

NCAR UCAR | Research Data Archive
Computational & Information Systems Lab

New tool integrating **NCAR** models and **NEON** observations

1. Supported, portable computing environment to enable scientific discovery and innovation by university researchers with diverse backgrounds
2. Further develop and improve NCAR models; and
3. Enhance NEON measurements and data distribution



Funded by NSF CISE, BIO, and GEO directorates

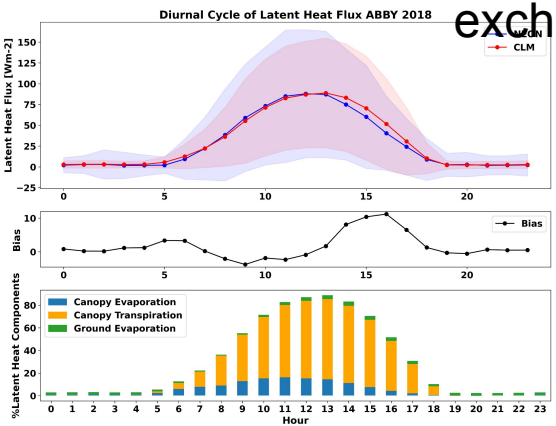
Bonan, G.B. (NCAR; PI) and Michael SanClements, David Durden, and Dawn Lenz (NEON; co-PIs) “Data CI Pilot: NCAR and NEON cyberinfrastructure collaborations to enable convergence research linking the atmospheric and biological sciences,” $600K, award 2039932 (11/20 - 10/22)

Cyberinfrastructure **Needs**

Portable computing environment from laptops to high performance supercomputers

Near-real time NCAR-NEON data exchanges

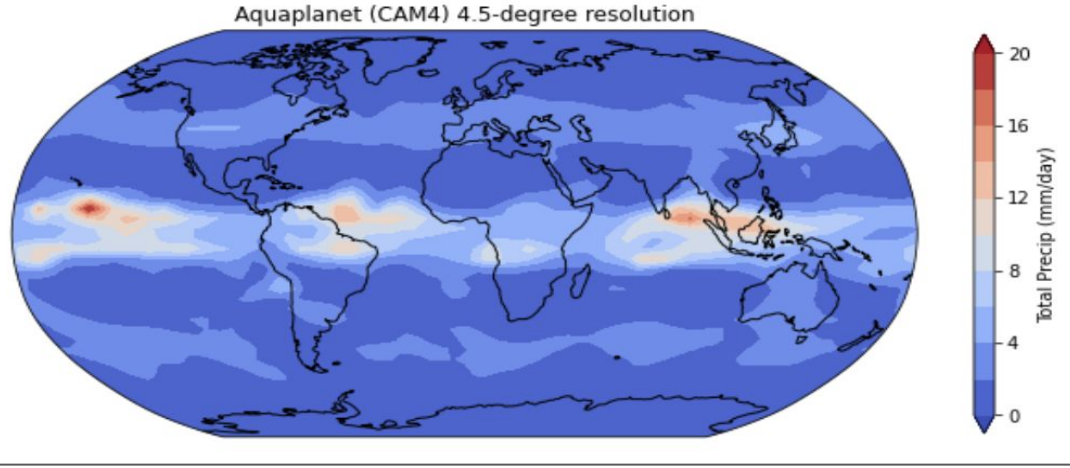
Interactive data visualization and exploration



Containerized CESM (Community Earth System Model)

- Runs on Mac/Windows/Linux systems, from laptops to servers
- Fully preconfigured – no installation or porting required

```
[14]: create_newcase --case ~/test1 --compset QPC4 --res f45_f45_mg37 --run-unsupported
      ...[snip]...
      plt.title("Aquaplanet (CAM4) 4.5-degree resolution");
```



An excellent tool for training, development & reproducible science

(Brian Dobbins, NCAR)

What is CESM-Lab?

CESM-Lab: *Containerized* environment with CESM + Jupyter Lab

CESM

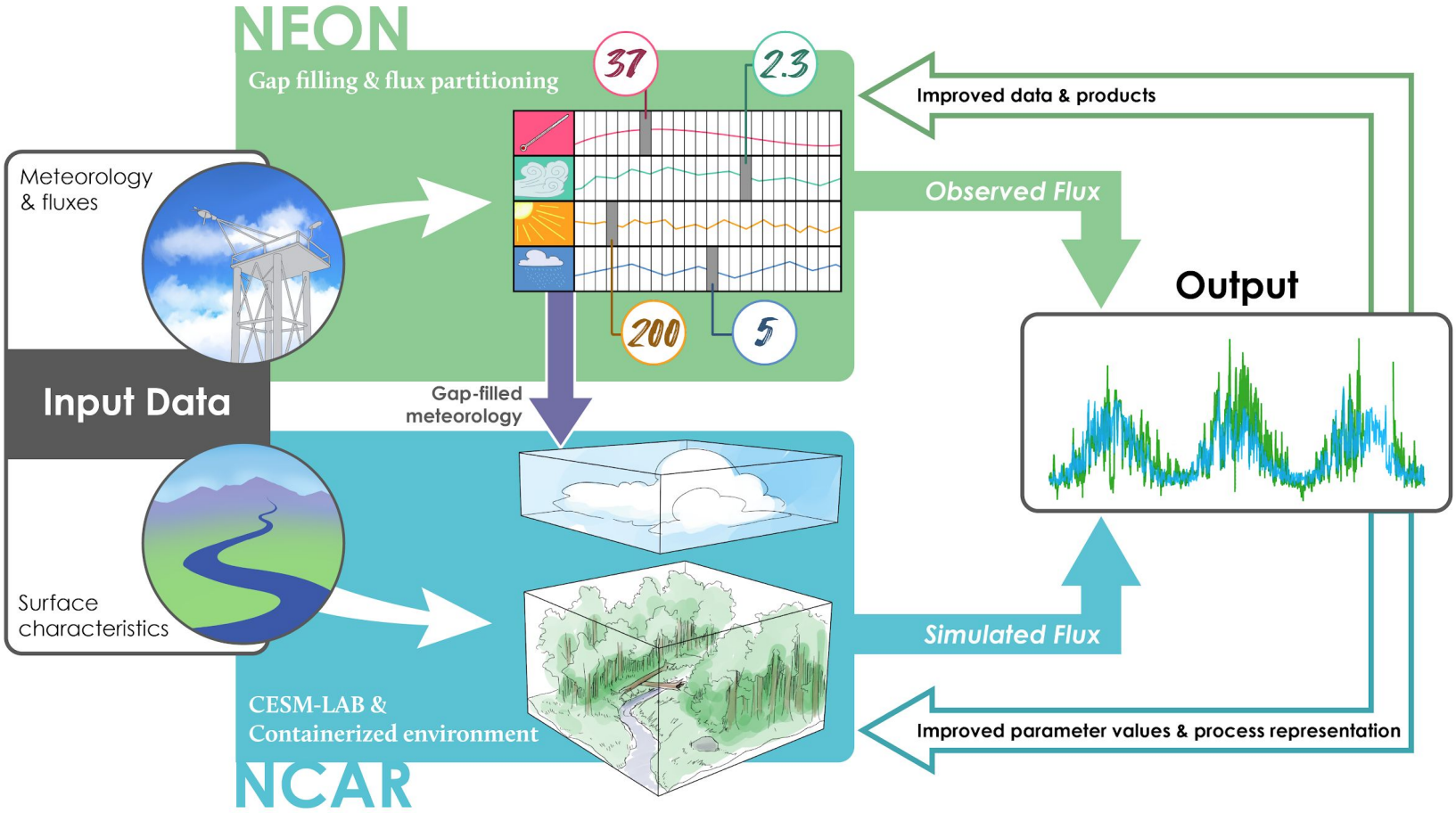
+

What is a *container*?	
Portable	Run *almost* anywhere (Mac, Windows, Linux)
Preconfigured	No installation or porting required!
Standardized	All users get the *same environment*

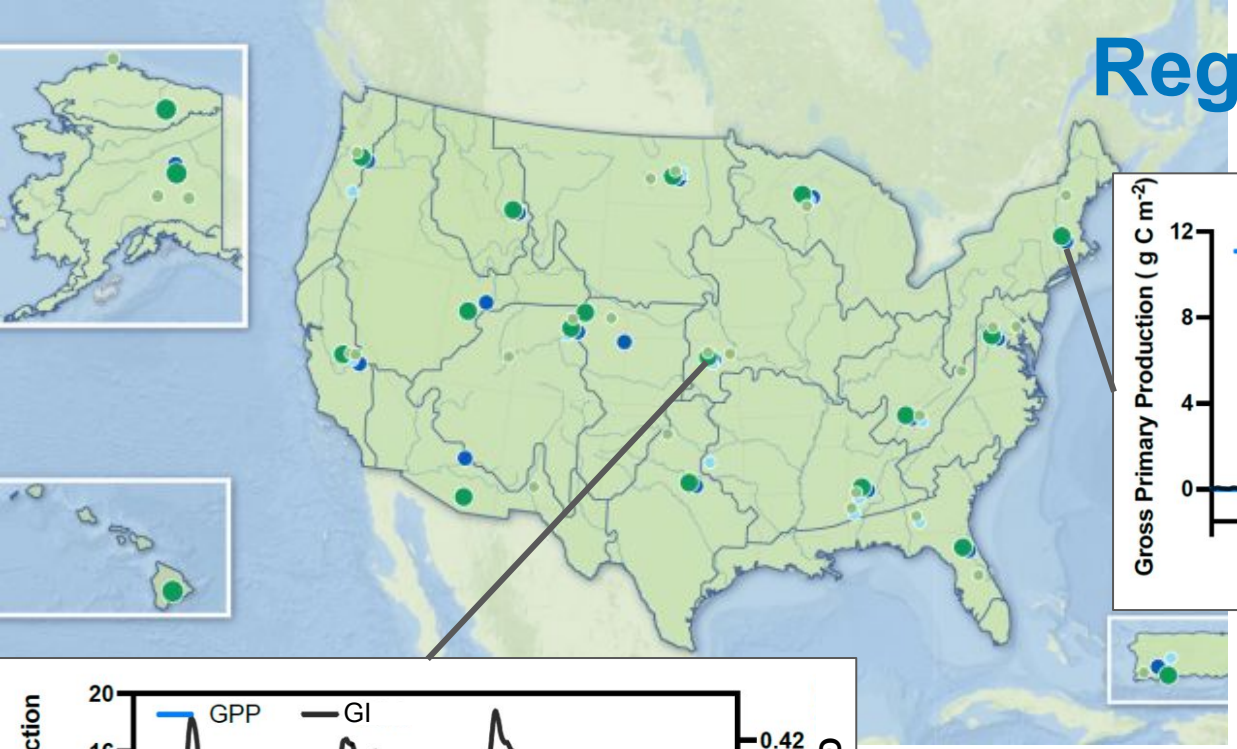
A simple, easy-to-use platform for CESM teaching & research

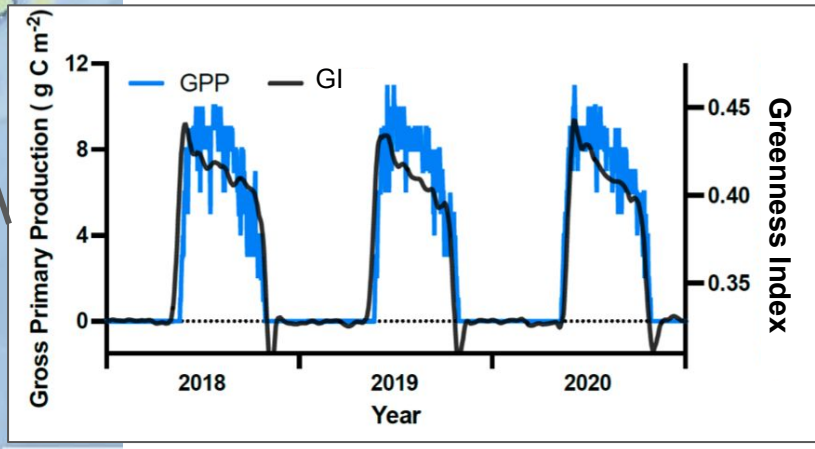
(Brian Dobbins, NCAR)

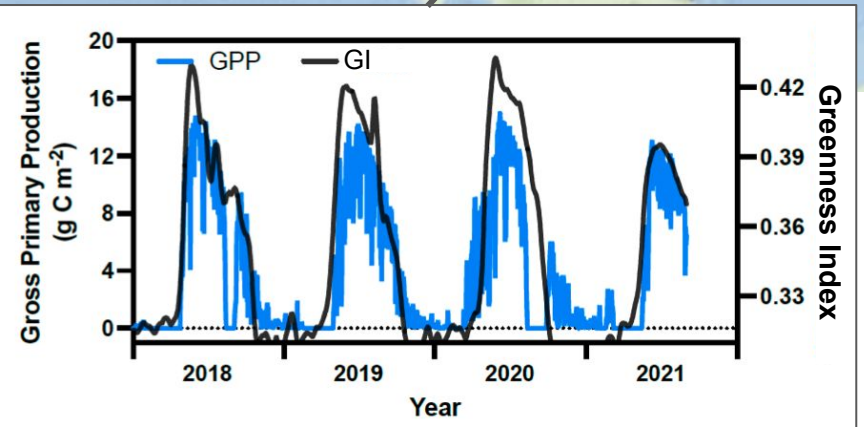
NCAR-NEON **Data Exchange**



Regional predictions





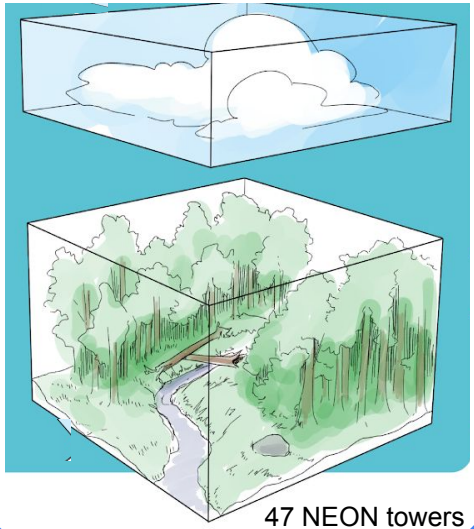


NCAR-NEON-University collaborations:

- Model improvements
- Initialized predictions
- Machine learning and data assimilation opportunities

Simple & Accessible Tutorials

CESM-Lab Docker Container

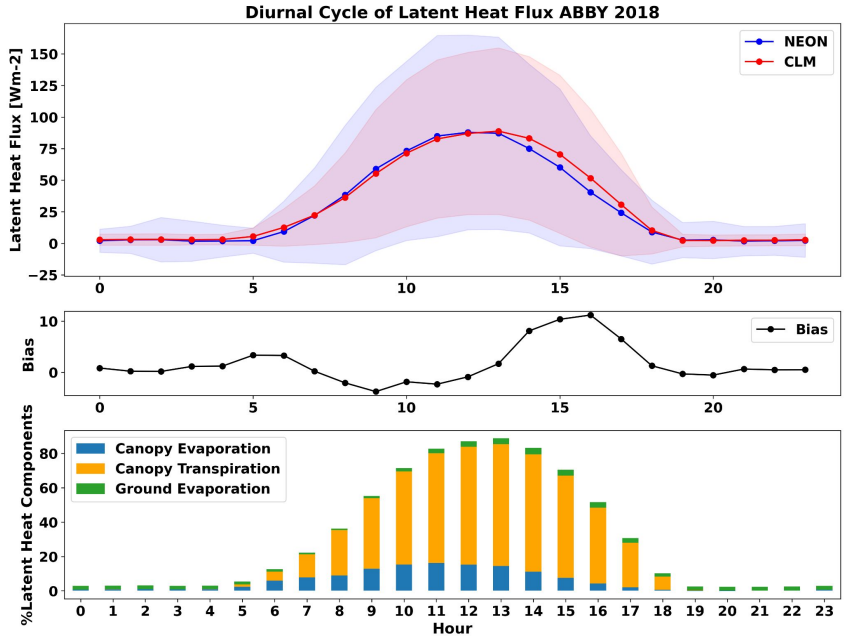


Excerpt from CLM-NEON Visualization Tutorial

How do transpiration and evaporation components of latent heat flux change throughout the day?

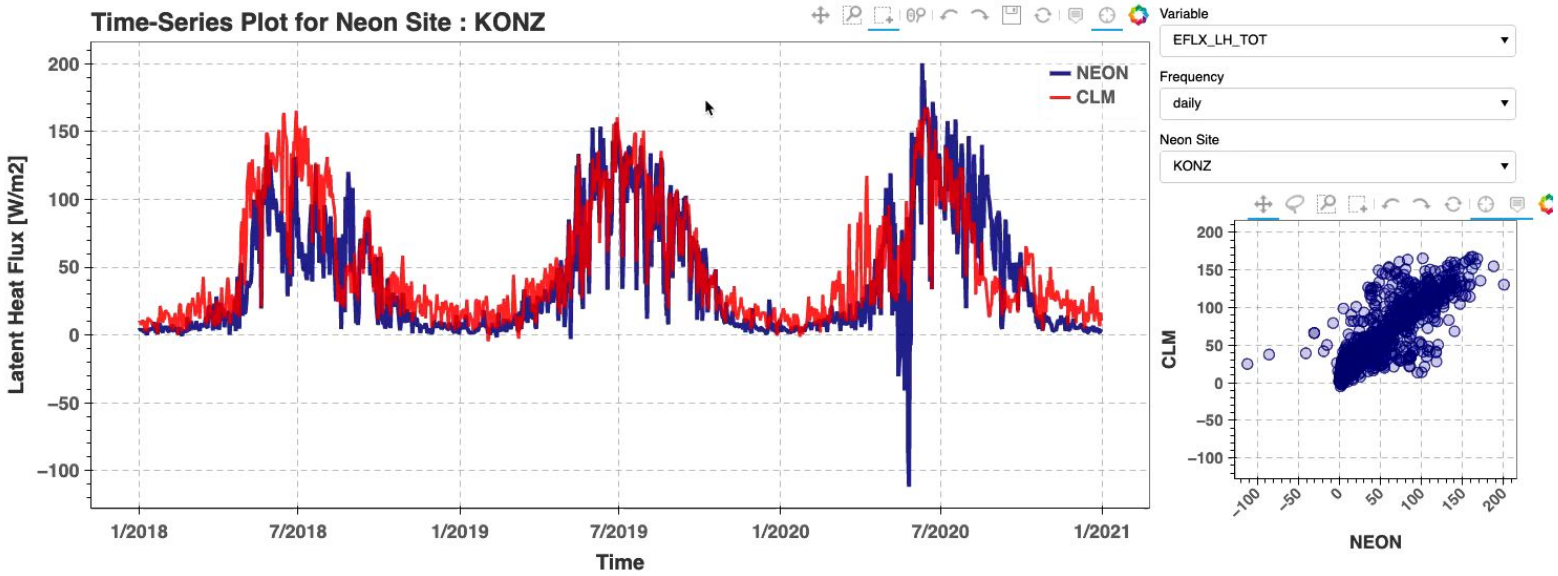
Run the cells below to add a panel to the plot

```
#Calculating diel cycle flux proportions
diel_df_mean['sim_FCEV_perc']= diel_df_mean ['sim_FCEV']/(diel_df_mean ['sim_FCEV']+diel_df_mean ['s
diel_df_mean ['sim_FCTR_perc']= diel_df_mean ['sim_FCTR']/(diel_df_mean ['sim_FCEV']+diel_df_mean ['
diel_df_mean ['sim_FGEV_perc']= diel_df_mean ['sim_FGEV']/(diel_df_mean ['sim_FCEV']+diel_df_mean ['
```



What is a container? A preconfigured, portable application that can reliably run on any computing infrastructure.

Interactive Data Explorer



promote community **Engagement,** expand **Access,** & broaden **Participation**

The CLM-NEON project offers a low-cost, high impact scientific tool that opens opportunities for communities without access to advanced computing resources or with little modeling experience

The Need:

A solution:

"The CLM-NEON simulation tool… [will] open a new frontier in data, modeling, and understanding our changing environment" –Sanjiv Kumar, Auburn University College of Forestry & Wildlife Services