

Kilometer-scale E3SM Land Model Development, Integration, and Applications

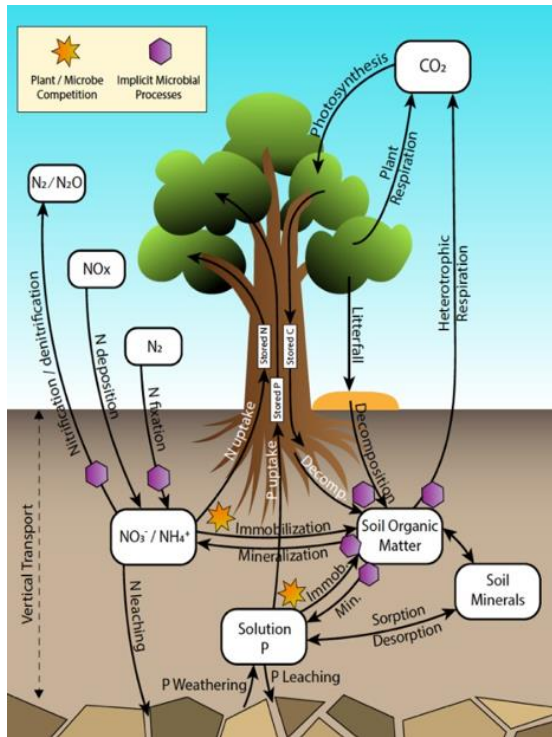
Dali Wang, Peter Schwartz, Fengming Yuan,
Shih-Chieh Kao, Michele Thornton, Daniel
Ricciuto, Peter Thornton, Qinglei Cao, Chen
Wang, and others

EESMPI2024 breakout session

August 7, 2024

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

Objective

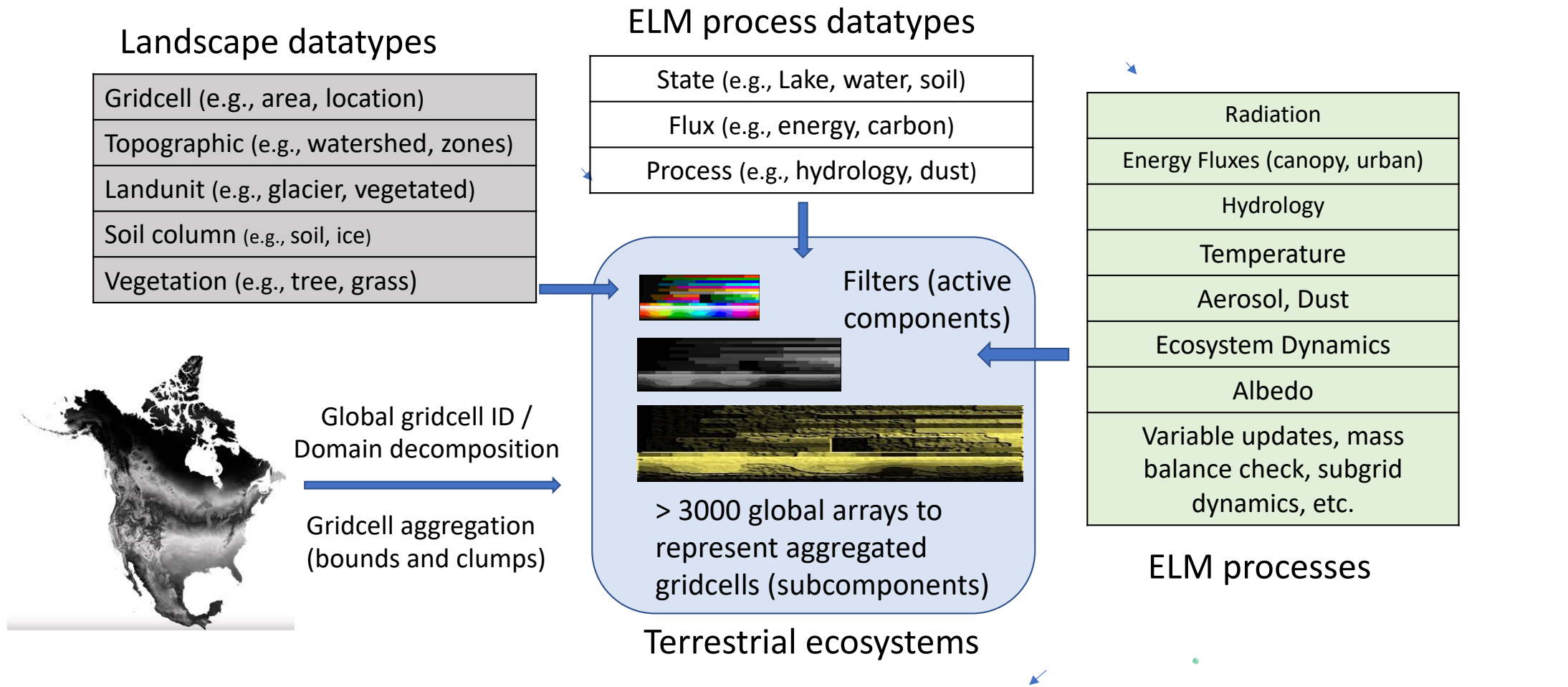


Develop ultrahigh-resolution ELM (uELM) to understand hydrologic cycles, biogeophysics, and ecosystem dynamics using start-of-the-art datasets

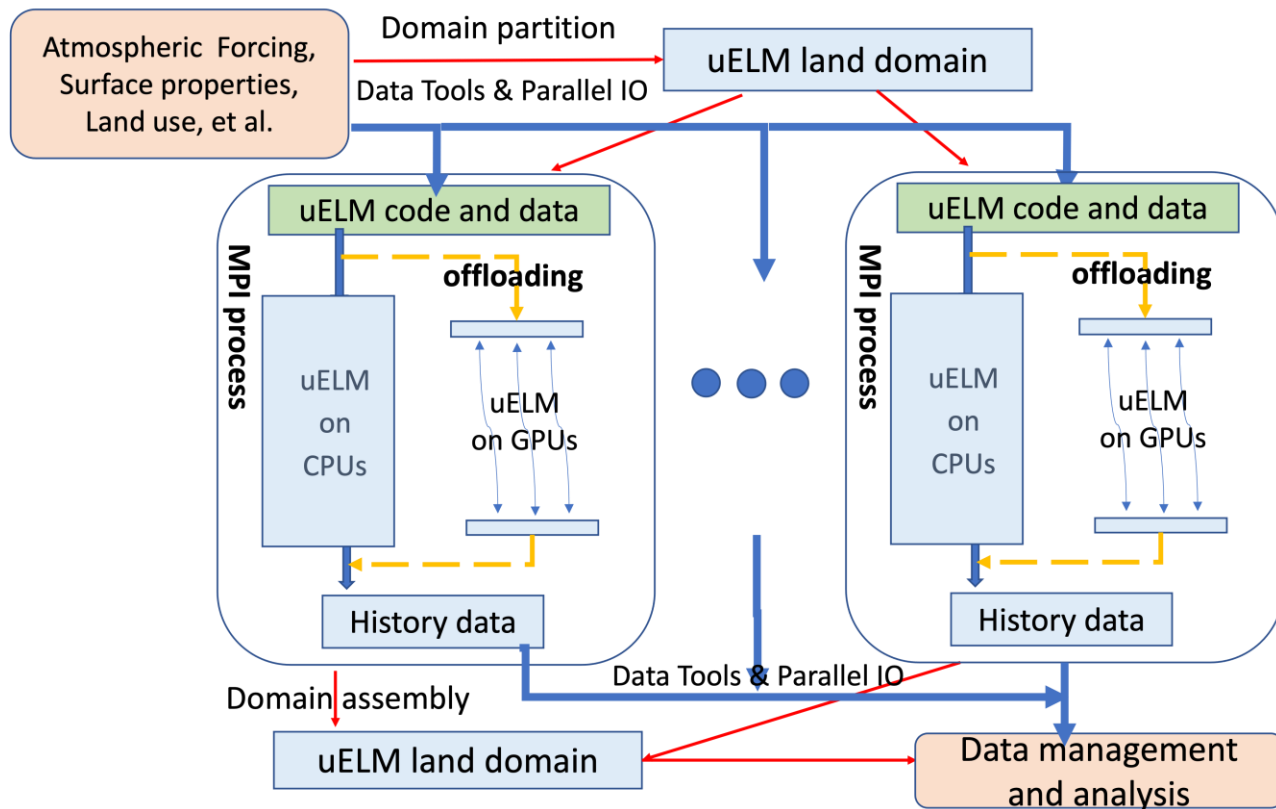
- Large-scale: continent and beyond (e.g., NA, Pan-arctic)
- Ultrahigh-resolution (1-3 km resolution)
- HPCs, including exascale computers (Summit and Frontier)



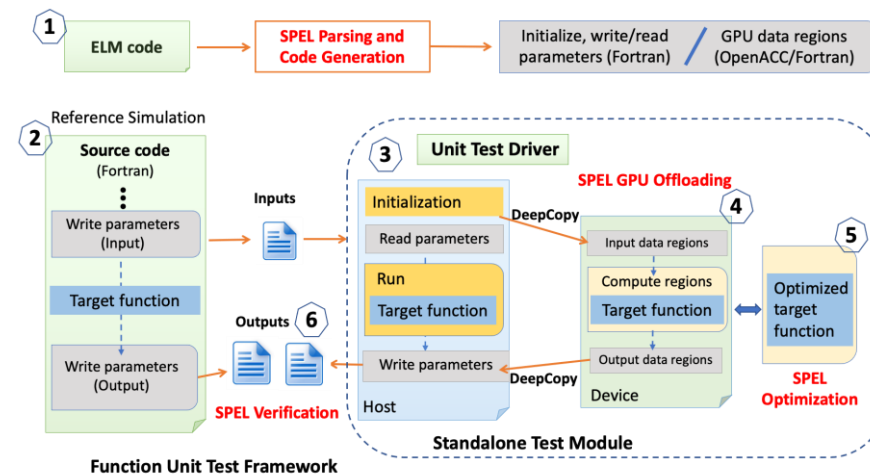
Background: ELM is a data-centric application



Background: Computational model and code porting



SPEL: Software tool for Porting ELM



High resolution datasets

- North America
 - Climate Forcing (3-Hourly Daymet (1980-present)) (50 TB)
 - Surface properties (interpolation and HR data fusion) (120GB)
 - Land use and change dataset (700 GB)
 - Others (e.g., ndep, pdep, population density, lighting frequency)

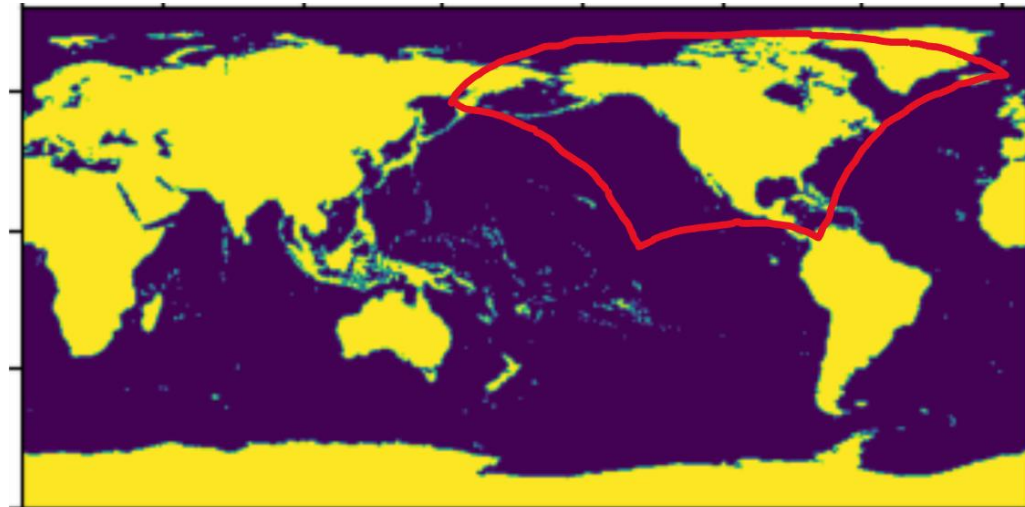


1km x 1km

Total: 63.1 M
(8075*7814)

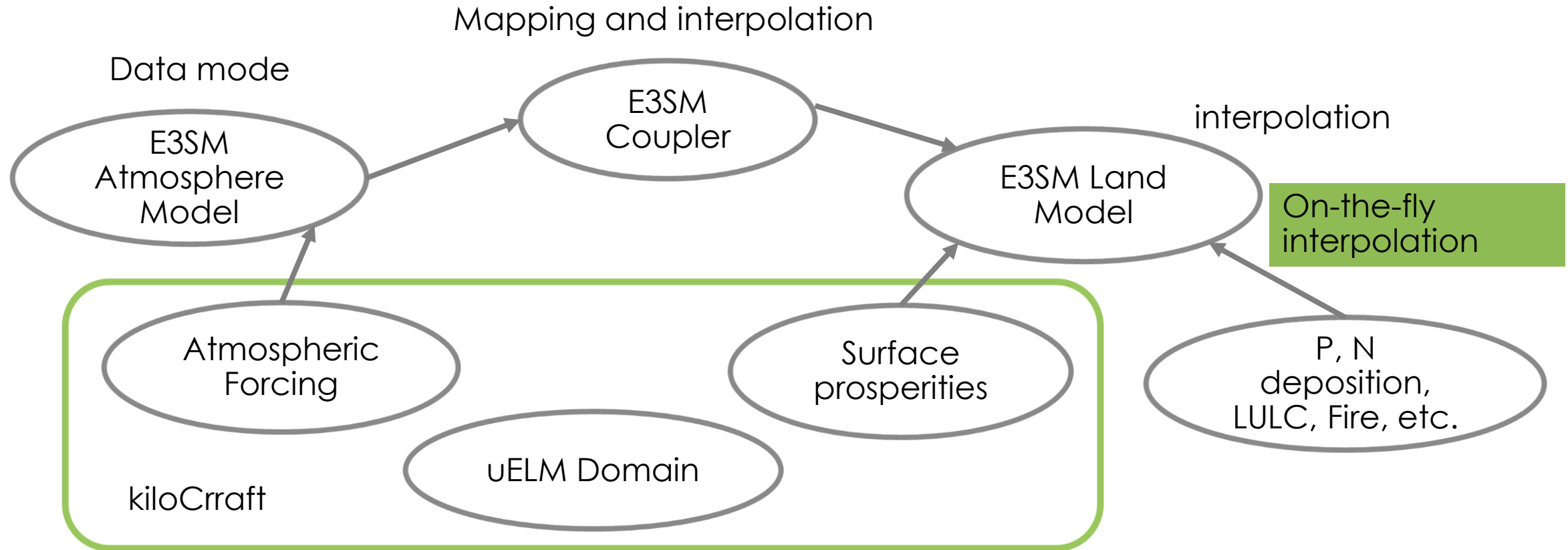
Land: 21.5 M

Lambert
Conformal Conic
/ WGS84



~ 350 times LT 62K (total land cells at a 0.5x0.5 resolution simulation)

Offline uELM simulation preparation with kiloCraft



Multi-resolution, regridding, data checking
Target large-scale simulation on high-performance computers
Easy adoption of state-of-the-art data products

Create uELM cases

- User define the AOIs (gridID or coordinates)
- Create domain/surfddata/forcing with kiloCraft scripts (interactive or batch mode)
- Link AOI data with the name convention required by CIME and DATM mode (Daymet_uELM)
- Create and run the uELM case with user-defined dataset and data ATM (ELM_USRDAT)

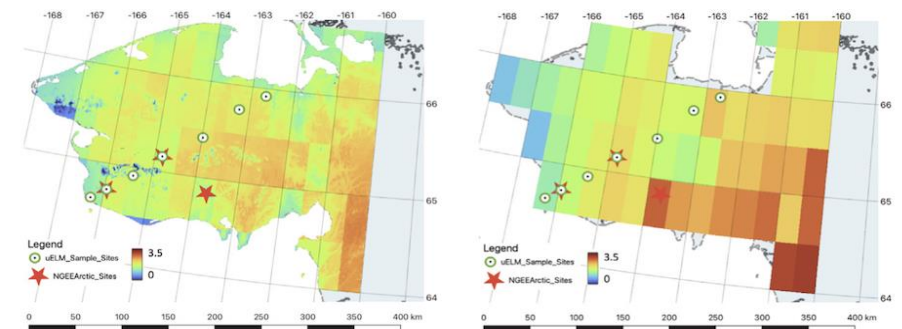
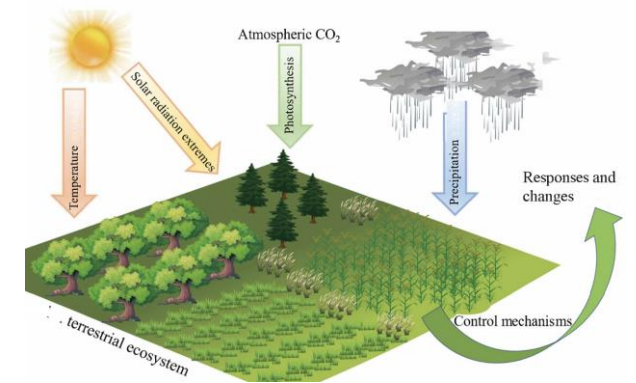
Next: Code and data development

- uELM simulation on Summit and Perlmutter
 - 72 million gridcells (CPU: 84000 cores) (* GPU code scale up to 120 GPUs)
 - Target 130 million gridcells (4000 nodes/Summit)
 - KM-scale data integration and uELM input datasets (kiloCraft)
- uELM GBC spin-up acceleration on Frontier
 - ORBIT: AI foundation model
 - Exascale Emulator: statistical inference
- uELM input data preparation:
 - Atmospheric forcing (Daymet-GSWPS, and Daymet-ERA5) (Kao, et al)
 - Surface properties data (continental level) (TESSFAs and others)

Next: Program integration and applications

- Activities

- Pan-arctic simulation (1km x 1km)
- Regional peatland simulation (4kmx4km)
- NetZero-Armada (TVA 2050 pathways, 1km x 1km)
- Urban studies (Chicago, New Orleans, Atlanta, Baltimore)
- Trustworthy AI foundation model (ORBIT)



Acknowledgement

- Energy Exascale Earth System Model (E3SM) project.
- NGEE Arctic, TES-SFA, and NASA:ESDIS/TE
- Computing resources from OLCF, NERSC, ORNL-cloud, OLCF-baseline, etc.