U.S. DEPARTMENT OF ENERGY

Office of Science

Exploiting Artificial Intelligence and Machine Learning for Advancing Carbon Cycle Science

Forrest M. Hoffman[1], Jitendra Kumar[1], Zachary L. Langford[1], V. Shashank Konduri[2], Auroop R. Ganguly[3], Zheng Shi[4], Elias Massoud[1], Nathan Collier[1], Min Xu[1], William W. Hargrove[5], Nicki L. Hickmon[6], Scott M. Collis[6], Charuleka Varadharajan[7], and Haruko Wainwright[7]

6th Training Course on New Advances in Land Carbon Cycle Modeling

June 16, 2023

[1]Oak Ridge National Laboratory, Oak Ridge, TN, USA
[2]National Ecological Observatory Network, Boulder, CO, USA
[3]Northeastern University, Boston, MA, USA
[4]University of Oklahoma, Norman, OK, USA
[5]US Department of Agriculture - Forest Service, Asheville, NC, USA
[6]Argonne National Laboratory, Lemont, IL, USA
[7]Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Introduction

- Observations of the Earth system are increasing in spatial resolution and temporal frequency, and will grow exponentially over the next 5–10 years
- With Exascale computing, simulation output is growing even faster, outpacing our ability to analyze, interpret and evaluate model results
- Explosive data growth and the promise of discovery through data-driven modeling necessitate new methods for feature extraction, change/anomaly detection, data assimilation, simulation, and analysis

Frontier at Oak Ridge National Laboratory is the #1 fastest supercomputer on the [TOP500](TOP500) List (May 22, 2023) and the first supercomputer to break the exaflop barrier (May 2022)

FOCUS NEXT-GENERATION SUPERCOMPUTERS

This article is the second in a two-part series. The first part, "How to Build a Hypercomputer," by Thomas Sterling, appeared in the July 2001 issue.

The Do-It-Yourself Supercomputer

By William W. Hargrove, Forrest M. Hoffman and Thomas Sterling

Photographs by Kay Chernush

Scientists have found a cheaper way to solve tremendously difficult computational problems: connect ordinary PCs so that they can work together

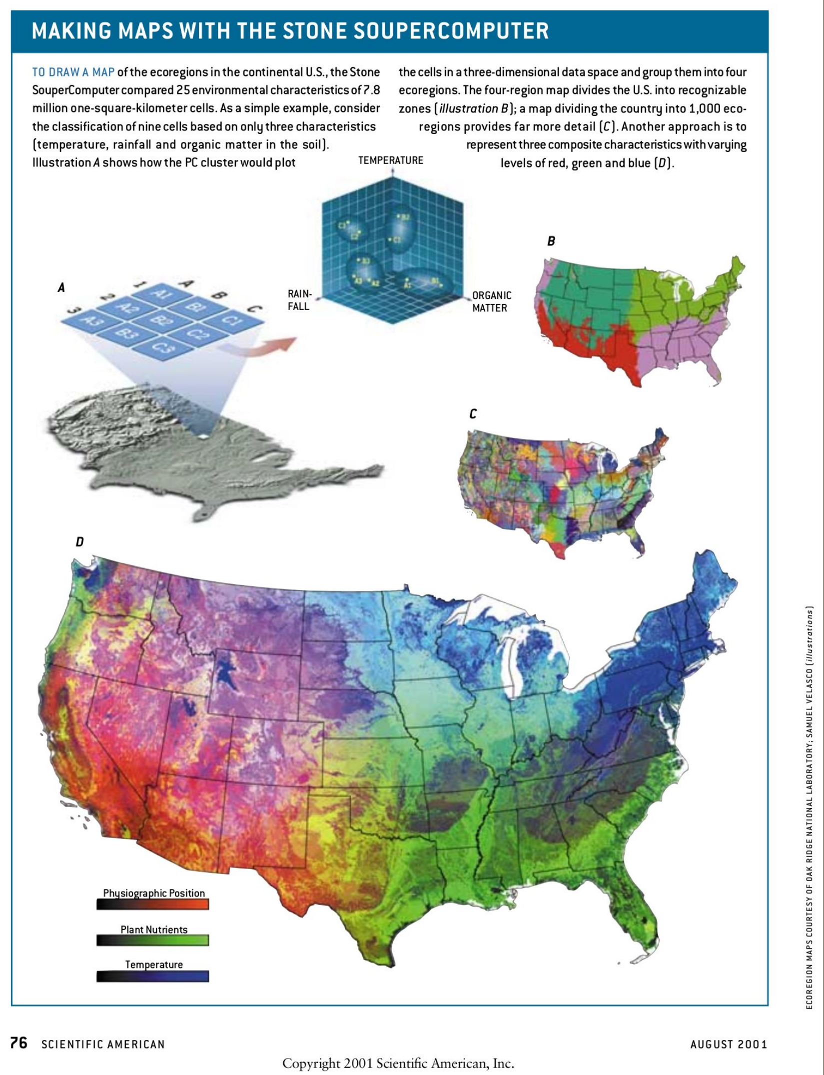
CLUSTER OF PCs at the Oak Ridge National Laboratory in Tennessee has been dubbed the Stone SouperComputer.

Hargrove, W. W., F. M. Hoffman, and T. Sterling (2001), The Do-It-Yourself Supercomputer, *Sci. Am.*, 265(2):72–79, https://www.scientificamerican.com/article/the-do-it-yourself-superc/

Multivariate Geographic Clustering

- Ecoregions have traditionally been created by experts
- Our approach has been to objectively create ecoregions using continuous continental-scale data and clustering
- We developed a highly scalable *k*-means cluster analysis code that uses distributed memory parallelism
- Originally developed on a 486/Pentium cluster, the code now runs on the largest hybrid CPU/GPU architectures on Earth

Hargrove, W. W., F. M. Hoffman, and T. Sterling (2001), The Do-It-Yourself Supercomputer, *Sci. Am.*, 265(2):72–79, https://www.scientificamerican.com/article/the-do-it-yourself-superc/



MAKING MAPS WITH THE STONE SOUPERCOMPUTER

TO DRAW A MAP of the ecoregions in the continental U.S., the Stone SouperComputer compared 25 environmental characteristics of 7.8 million one-square-kilometer cells. As a simple example, consider the classification of nine cells based on only three characteristics (temperature, rainfall and organic matter in the soil). Illustration *A* shows how the PC cluster would plot the cells in a three-dimensional data space and group them into four ecoregions. The four-region map divides the U.S. into recognizable zones (*illustration B*); a map dividing the country into 1,000 ecoregions provides far more detail (*C*). Another approach is to represent three composite characteristics with varying levels of red, green and blue (*D*).

ECOREGION MAPS COURTESY OF OAK RIDGE NATIONAL LABORATORY; SAMUEL VELASCO (*illustrations*)

Eos, Vol. 84, No. 48, 2 December 2003

EOS
EOS, TRANSACTIONS, AMERICAN GEOPHYSICAL UNION

VOLUME 84 NUMBER 48
2 DECEMBER 2003
PAGES 529–544

New Analysis Reveals Representativeness of the AmeriFlux Network

PAGES 529, 535

The AmeriFlux network of eddy flux covariance towers was established to quantify variation in carbon dioxide and water vapor exchange between terrestrial ecosystems and the atmosphere, and to understand the underlying mechanisms responsible for observed fluxes and carbon pools. The network is primarily funded by the U.S. Department of Energy, NASA, the National Oceanic and Atmospheric Administration, and the National Science Foundation. Similar regional networks elsewhere in the world—for example, CarboEurope, AsiaFlux, OzFlux, and Fluxnet Canada—participate in synthesis activities across larger geographic areas [*Baldocchi et al.*, 2001; *Law et al.*, 2002].

The existing AmeriFlux network will also form a backbone of "Tier 4" intensive measurement sites as one component of a four-tiered carbon observation network within the North American Carbon Program (NACP). The NACP seeks to provide long-term, mechanistically detailed, spatially resolved carbon fluxes across North America [*Wofsy and Harriss*, 2002]. For both of these roles, the AmeriFlux network should be ecologically representative of the environments contained within the geographic boundaries of the program. A new ecoregion-scale analysis of the existing AmeriFlux network reveals that, while central continental environments are well-represented, additional flux towers are needed to represent environmental

By William W. Hargrove, Forrest M. Hoffman, and Beverly E. Law

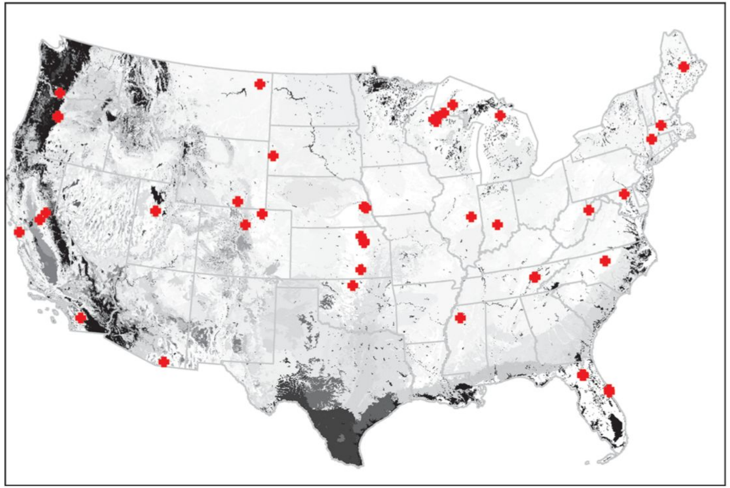


Fig. 1. The representativeness of an existing spatial array of sample locations or study sites—for example, the AmeriFlux network of carbon dioxide eddy flux covariance towers—can be mapped relative to a set of quantitative ecoregions, suggesting locations for additional samples or sites. Distance in data space to the closest ecoregion containing a site quantifies how well an existing network represents each ecoregion in the map. Environments in darker ecoregions are poorly represented by this network.

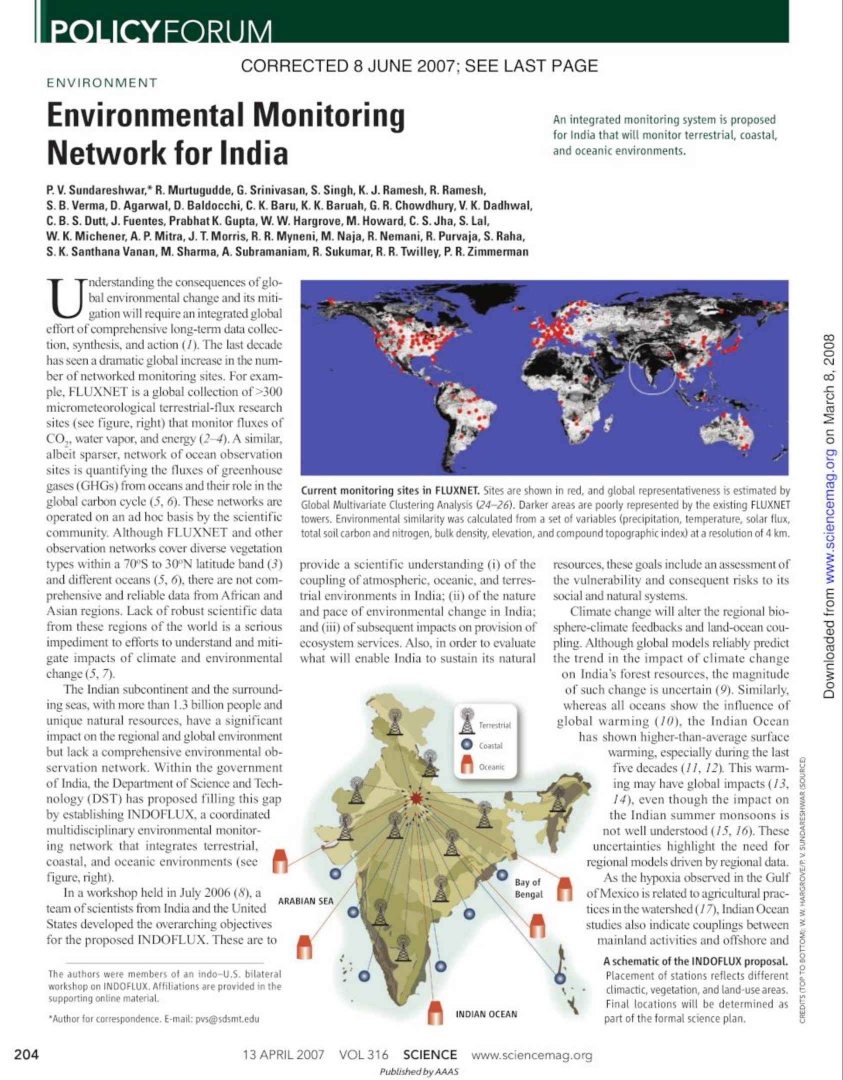
Network Representativeness

- The *n*-dimensional space formed by the data layers offers a natural framework for estimating representativeness of individual sampling sites
- The Euclidean distance between individual sites in data space is a metric of similarity or dissimilarity
- Representativeness across multiple sampling sites can be combined to produce a map of network representativeness

Hargrove, W. W., and F. M. Hoffman (2003), New Analysis Reveals Representativeness of the AmeriFlux Network, *Eos Trans. AGU*, 84(48):529, 535, doi:10.1029/2003EO480001.

Optimizing Sampling Networks

- Our group produced this network representativeness map for the authors from global climate, edaphic, and elevation and topography data
- Dark areas, including most of the Indian subcontinent, were poorly represented by the constellation of eddy covariance flux towers participating in FLUXNET in the year 2007



POLICY FORUM

CORRECTED 8 JUNE 2007; SEE LAST PAGE

ENVIRONMENT

Environmental Monitoring Network for India

An integrated monitoring system is proposed for India that will monitor terrestrial, coastal, and oceanic environments.

P. V. Sundareshwar,* R. Murtugudde, G. Srinivasan, S. Singh, K. J. Ramesh, R. Ramesh, S. B. Verma, D. Agarwal, D. Baldocchi, C. K. Baru, K. K. Baruah, G. R. Chowdhury, V. K. Dadhwal, C. B. S. Dutt, J. Fuentes, Prabhat K. Gupta, W. W. Hargrove, M. Howard, C. S. Jha, S. Lal, W. K. Michener, A. P. Mitra, J. T. Morris, R. R. Myneni, M. Naja, R. Nemani, R. Purvaja, S. Raha, S. K. Santhana Vanan, M. Sharma, A. Subramaniam, R. Sukumar, R. R. Twilley, P. R. Zimmerman

Understanding the consequences of global environmental change and its mitigation will require an integrated global effort of comprehensive long-term data collection, synthesis, and action (*1*). The last decade has seen a dramatic global increase in the number of networked monitoring sites. For example, FLUXNET is a global collection of >300 micrometeorological terrestrial-flux research sites (see figure, right) that monitor fluxes of CO_2, water vapor, and energy (*2–4*). A similar, albeit sparser, network of ocean observation sites is quantifying the fluxes of greenhouse gases (GHGs) from oceans and their role in the global carbon cycle (*5, 6*). These networks are operated on an ad hoc basis by the scientific community. Although FLUXNET and other observation networks cover diverse vegetation types within a 70°S to 30°N latitude band (*3*) and different oceans (*5, 6*), there are not comprehensive and reliable data from African and Asian regions. Lack of robust scientific data from these regions of the world is a serious impediment to efforts to understand and mitigate impacts of climate and environmental change (*5, 7*).

The Indian subcontinent and the surrounding seas, with more than 1.3 billion people and unique natural resources, have a significant impact on the regional and global environment but lack a comprehensive environmental observation network. Within the government of India, the Department of Science and Technology (DST) has proposed filling this gap by establishing INDOFLUX, a coordinated multidisciplinary environmental monitoring network that integrates terrestrial, coastal, and oceanic environments (see figure, right).

In a workshop held in July 2006 (*8*), a team of scientists from India and the United States developed the overarching objectives for the proposed INDOFLUX. These are to provide a scientific understanding (i) of the coupling of atmospheric, oceanic, and terrestrial environments in India; (ii) of the nature and pace of environmental change in India; and (iii) of subsequent impacts on provision of ecosystem services. Also, in order to evaluate what will enable India to sustain its natural resources, these goals include an assessment of the vulnerability and consequent risks to its social and natural systems.

Climate change will alter the regional biosphere-climate feedbacks and land-ocean coupling. Although global models reliably predict the trend in the impact of climate change on India's forest resources, the magnitude of such change is uncertain (*9*). Similarly, whereas all oceans show the influence of global warming (*10*), the Indian Ocean has shown higher-than-average surface warming, especially during the last five decades (*11, 12*). This warming may have global impacts (*13, 14*), even though the impact on the Indian summer monsoons is not well understood (*15, 16*). These uncertainties highlight the need for regional models driven by regional data. As the hypoxia observed in the Gulf of Mexico is related to agricultural practices in the watershed (*17*), Indian Ocean studies also indicate couplings between mainland activities and offshore and

Current monitoring sites in FLUXNET. Sites are shown in red, and global representativeness is estimated by Global Multivariate Clustering Analysis (*24–26*). Darker areas are poorly represented by the existing FLUXNET towers. Environmental similarity was calculated from a set of variables (precipitation, temperature, solar flux, total soil carbon and nitrogen, bulk density, elevation, and compound topographic index) at a resolution of 4 km.

Terrestrial
Coastal
Oceanic
ARABIAN SEA
Bay of Bengal
INDIAN OCEAN

A schematic of the INDOFLUX proposal. Placement of stations reflects different climactic, vegetation, and land-use areas. Final locations will be determined as part of the formal science plan.

The authors were members of an indo–U.S. bilateral workshop on INDOFLUX. Affiliations are provided in the supporting online material.

*Author for correspondence. E-mail: pvs@sdsmt.edu

Published by AAAS

Sundareshwar, P. V., et al. (2007), Environmental Monitoring Network for India, *Science*, 316(5822):204–205, doi:10.1126/science.1137417.

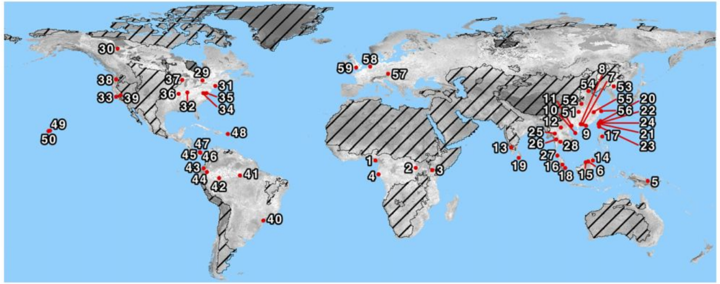


Fig. 1 Map of the CTFS-ForestGEO network illustrating its representation of bioclimatic, edaphic, and topographic conditions globally. Site numbers correspond to ID# in Table 2. Shading indicates how well the network of sites represents the suite of environmental factors included in the analysis; light-colored areas are well-represented by the network, while dark colored areas are poorly represented. Stippling covers nonforest areas. The analysis is described in Appendix S1.

Table 1 Attributes of a CTFS-ForestGEO census

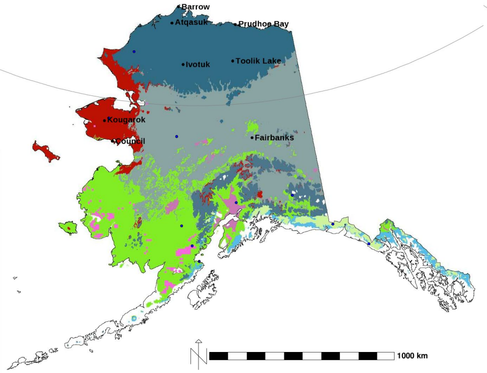
Attribute	Utility
Very large plot size	Resolve community and population dynamics of highly diverse forests with many rare species with sufficient sample sizes (Losos & Leigh, 2004; Condit et al., 2006); quantify spatial patterns at multiple scales (Condit et al., 2000; Wiegand et al., 2007a,b; Detto & Muller-Landau, 2013; Lutz et al., 2013); characterize gap dynamics (Feeley et al., 2007b); calibrate and validate remote sensing and models, particularly those with large spatial grain (Mascaro et al., 2011; Réjou-Méchain et al., 2014)
Includes every freestanding woody stem ≥1 cm DBH	Characterize the abundance and diversity of understory as well as canopy trees; quantify the demography of juveniles (Condit, 2000; Muller-Landau et al., 2006a,b).
All individuals identified to species	Characterize patterns of diversity, species-area, and abundance distributions (Hubbell, 1979, 2001; He & Legendre, 2002; Condit et al., 2005; John et al., 2007; Shen et al., 2009; He & Hubbell, 2011; Wang et al., 2011; Cheng et al., 2012); test theories of competition and coexistence (Brown et al., 2013); describe poorly known plant species (Gereau & Kenfack, 2000; Davies, 2001; Davies et al., 2001; Sonké et al., 2002; Kenfack et al., 2004, 2006)
Diameter measured on all stems	Characterize size-abundance distributions (Muller-Landau et al., 2006b; Lai et al., 2013; Lutz et al., 2013); combine with allometries to estimate whole-ecosystem properties such as biomass (Chave et al., 2008; Valencia et al., 2009; Lin et al., 2012; Ngo et al., 2013; Muller-Landau et al., 2014)
Mapping of all stems and fine-scale topography	Characterize the spatial pattern of populations (Condit, 2000); conduct spatially explicit analyses of neighborhood influences (Condit et al., 1992; Hubbell et al., 2001; Uriarte et al., 2004, 2005; Rüger et al., 2011, 2012; Lutz et al., 2014); characterize microhabitat specificity and controls on demography, biomass, etc. (Harms et al., 2001; Valencia et al., 2004; Chuyong et al., 2011); align on the ground and remote sensing measurements (Asner et al., 2011; Mascaro et al., 2011).
Census typically repeated every 5 years	Characterize demographic rates and changes therein (Russo et al., 2005; Muller-Landau et al., 2006a,b; Feeley et al., 2007a; Lai et al., 2013; Stephenson et al., 2014); characterize changes in community composition (Losos & Leigh, 2004; Chave et al., 2008; Feeley et al., 2011; Swenson et al., 2012; Chisholm et al., 2014); characterize changes in biomass or productivity (Chave et al., 2008; Banin et al., 2014; Muller-Landau et al., 2014)

Optimizing Sampling Networks

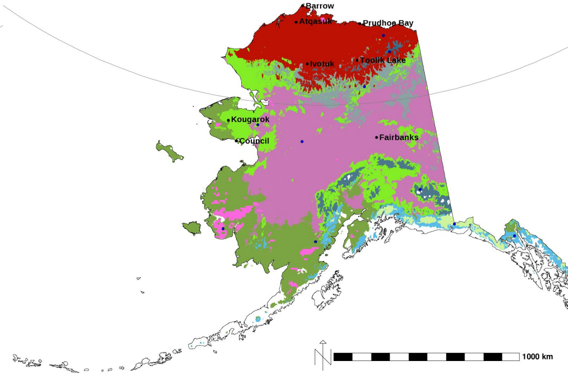
- The CTFS-ForestGEO global forest monitoring network is aimed at characterizing forest responses to global change
- The figure at left shows the global representativeness of the CTFS-ForestGEO sites in 2014
- Non-forested areas are masked with hatching, and as expected, they are consistently darker than the forested regions, which are represented to varying degrees by the monitoring sites

Anderson-Teixeira, K. J., et al. (2015), CTFS-ForestGEO: A Worldwide Network Monitoring Forests in an Era of Global Change, *Glob. Change Biol.*, 21(2):528–549, doi:10.1111/gcb.12712.

Sampling Network Design

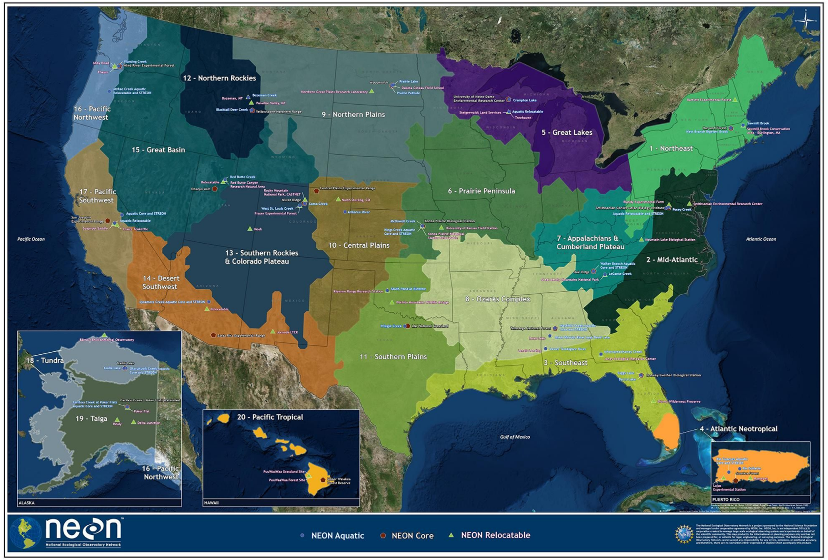


2000–2009



2090–2000

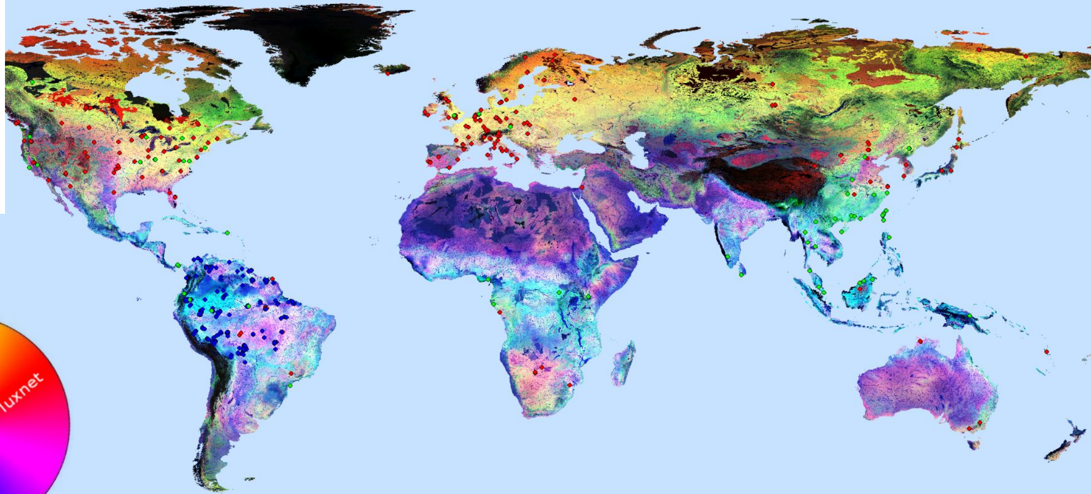
(Hoffman et al., 2013)

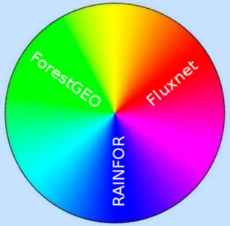


NSF's NEON Sampling Domains

Gridded data from satellite and airborne remote sensing, models, and synthesis products can be combined to design optimal sampling networks and understand representativeness as it evolves through time

Triple-Network Global Representativeness

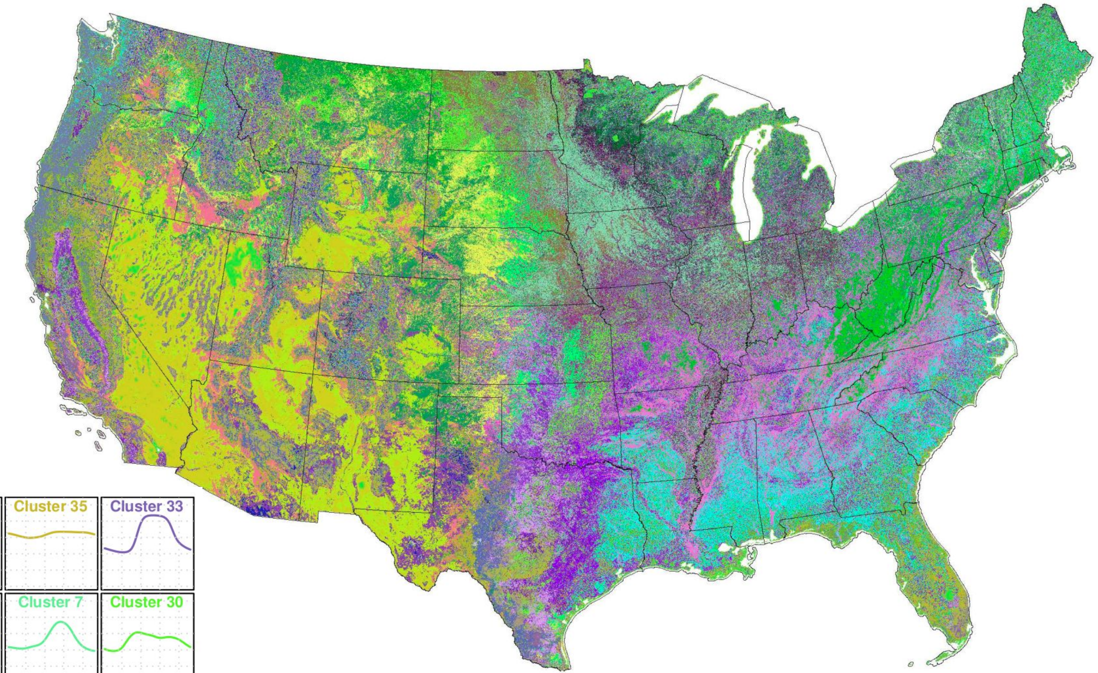


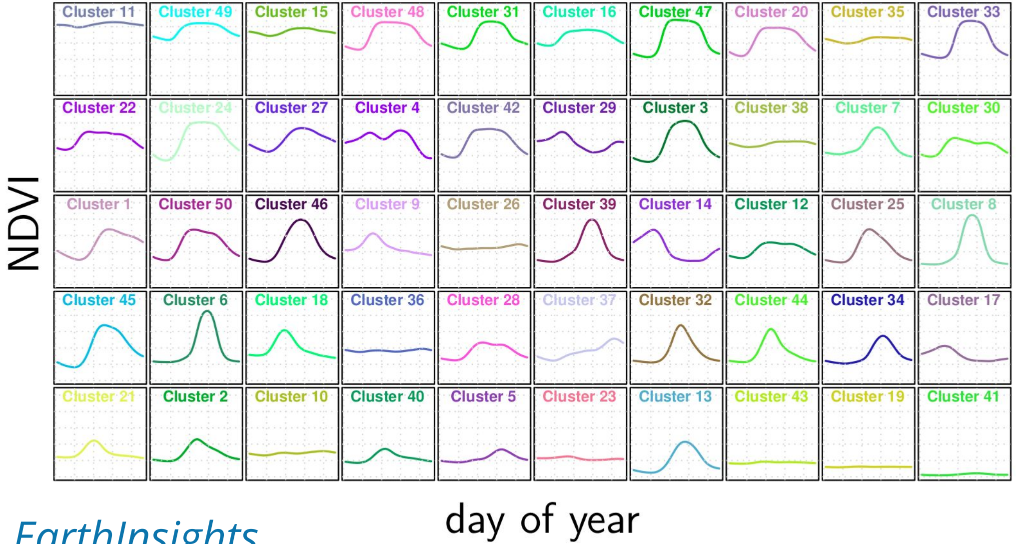


(Maddalena et al., in prep.)

50 Phenoregions for year 2012 (Random Colors)

250m MODIS NDVI
Every 8 days (46 images/year)
Clustered from year 2000 to present



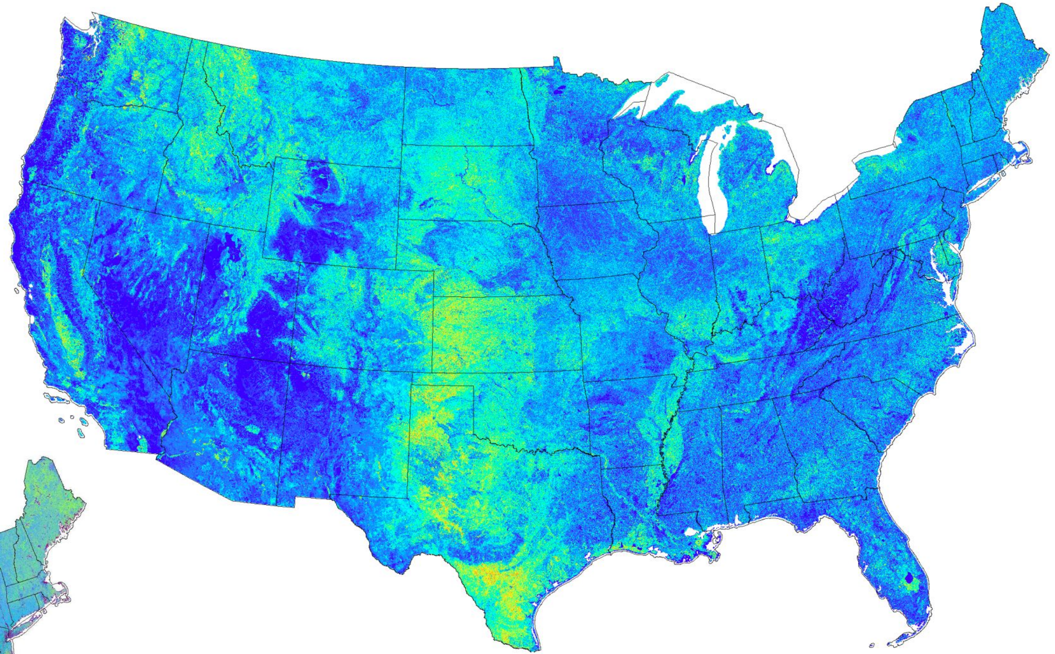


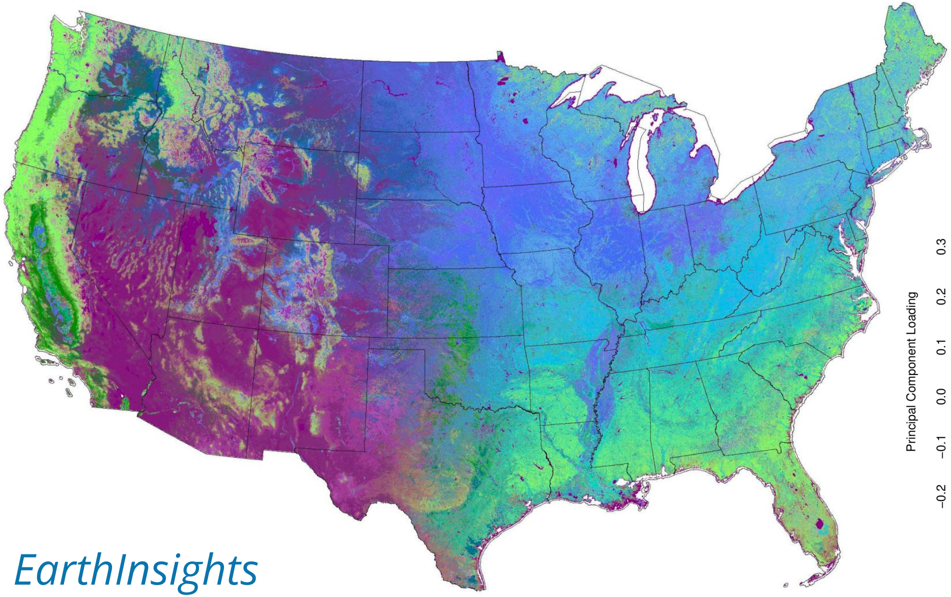
50 Phenoregion Prototypes (Random Colors)

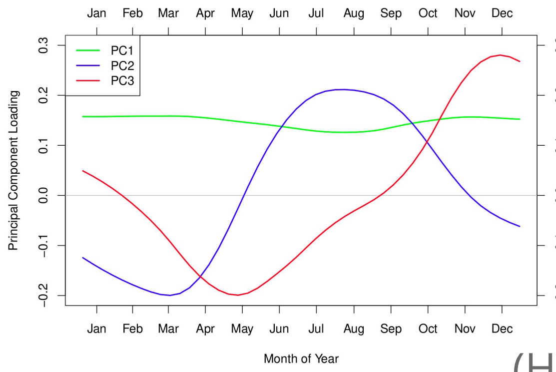
EarthInsights

(Hargrove et al., in prep.)

50 Phenoregions Persistence
and
50 Phenoregions Max Mode
(Similarity Colors)







Principal Components Analysis

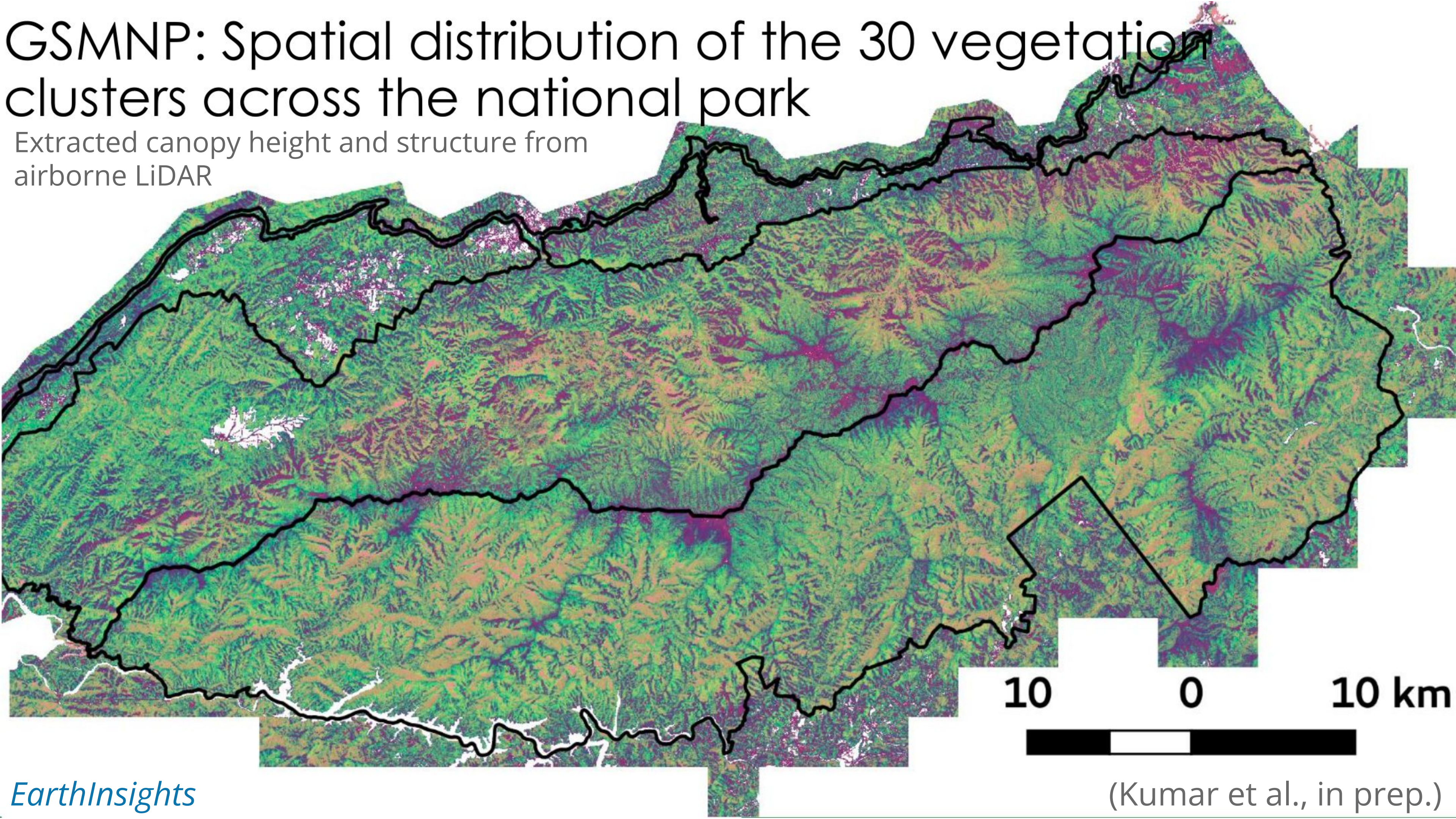
PC1 ~ Evergreen
PC2 ~ Deciduous
PC3 ~ Dry Deciduous

EarthInsights

(Hargrove et al., in prep.)

GSMNP: Spatial distribution of the 30 vegetation clusters across the national park

Extracted canopy height and structure from airborne LiDAR



EarthInsights

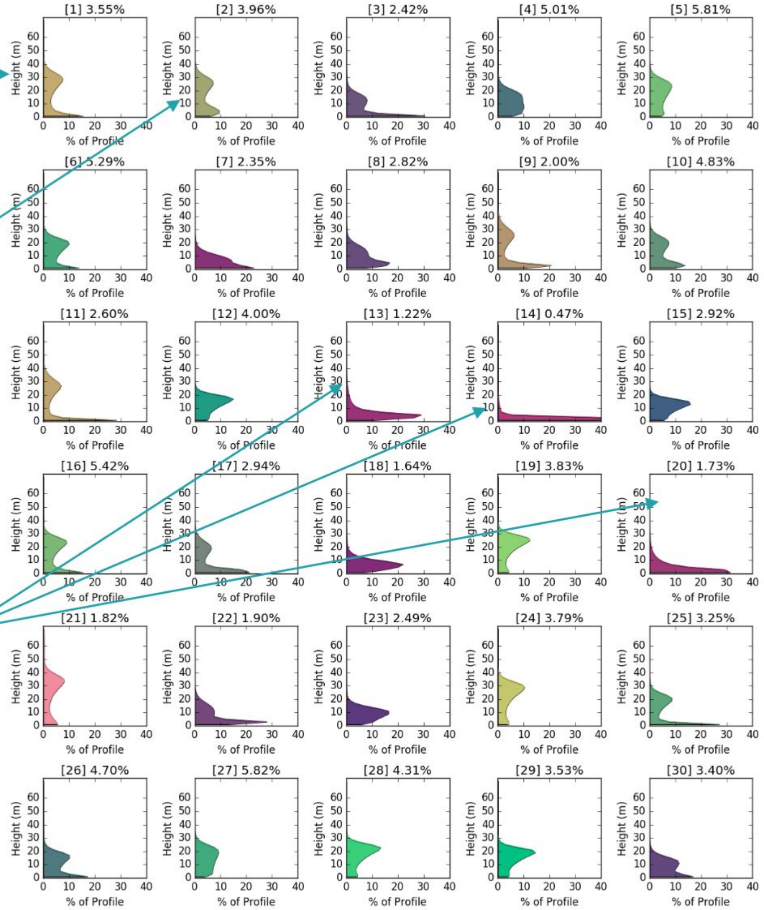
(Kumar et al., in prep.)

GSMNP: 30 representative vertical structures (cluster centroids) identified

tall forests with low understory vegetation

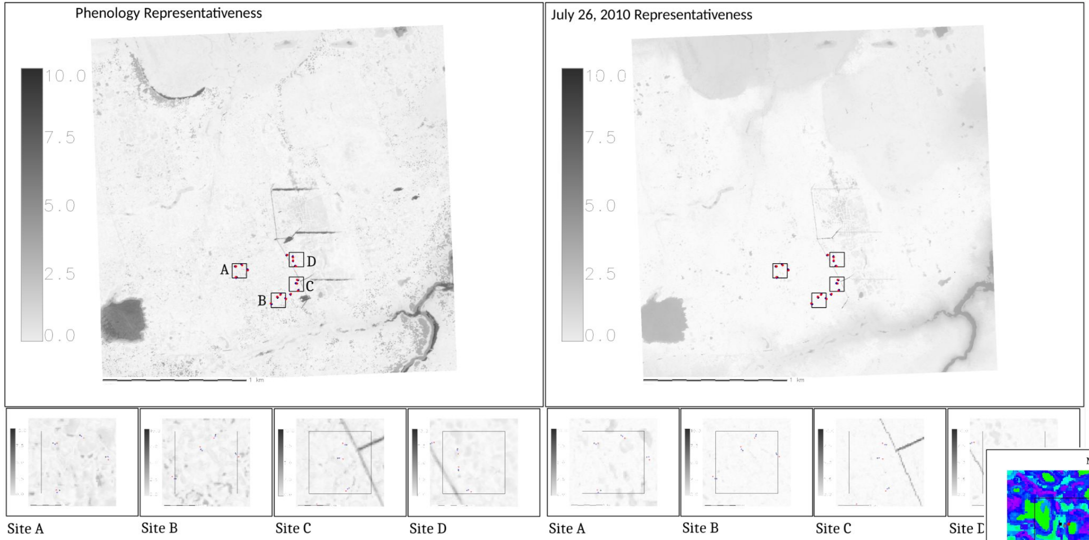
forests with slightly lower mean height with dense understory vegetation

low height grasslands and heath balds that are small in area but distinct landscape type



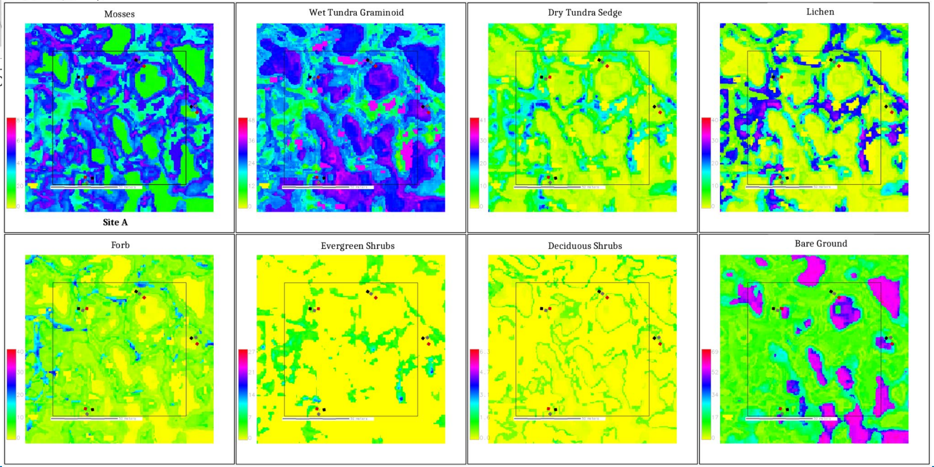
(Kumar et al., in prep.)

Vegetation Distribution at Barrow Environmental Observatory



Representativeness map for vegetation sampling points in sites A, B, C, and D with phenology (left) and without (right) from WorldView2 multispectral imagery for the year 2010 and LiDAR data

Example plant functional type (PFT) distributions scaled up from vegetation sampling locations

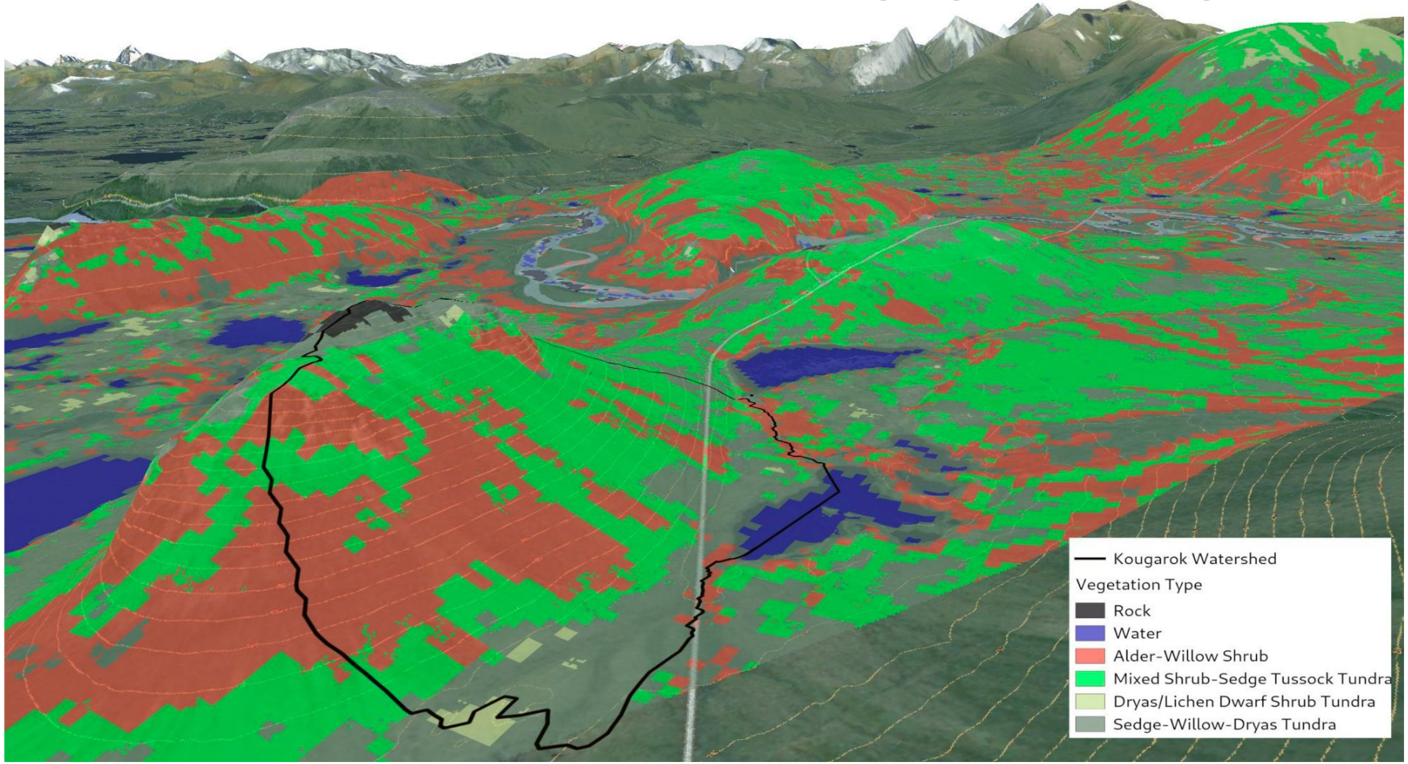


In situ data from field measurement activities inform the development of wide-scale maps of vegetation distribution through inference using remote sensing data as surrogate variables, and relationships with environmental controls can be extracted

Langford, Z. L., et al. (2016), Mapping Arctic Plant Functional Type Distributions in the Barrow Environmental Observatory Using WorldView-2 and LiDAR Datasets, *Remote Sens.*, 8(9):733, doi:10.3390/rs8090733.

Arctic Vegetation Mapping from Multi-Sensor Fusion

Used Hyperion Multispectral and IfSAR-derived Digital Elevation Model, applied cluster analysis, and trained a convolutional neural network (CNN) with Alaska Existing Vegetation Ecoregions (AKEVT)



Langford, Z. L., et al. (2019), Arctic Vegetation Mapping Using Unsupervised Training Datasets and Convolutional Neural Networks, *Remote Sens.*, 11(1):69, doi:[10.3390/rs11010069](https://doi.org/10.3390/rs11010069).

Satellite Data Analytics Enables Within-Season Crop Identification

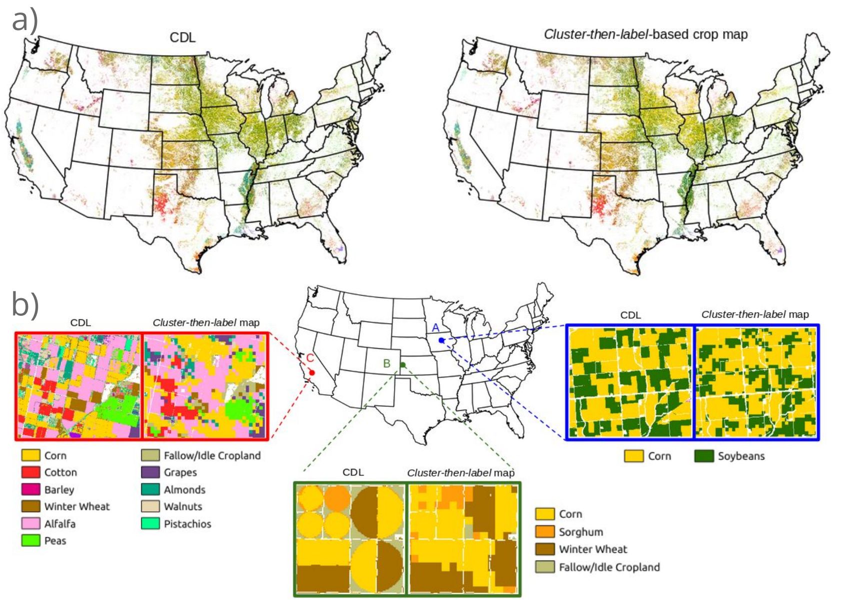
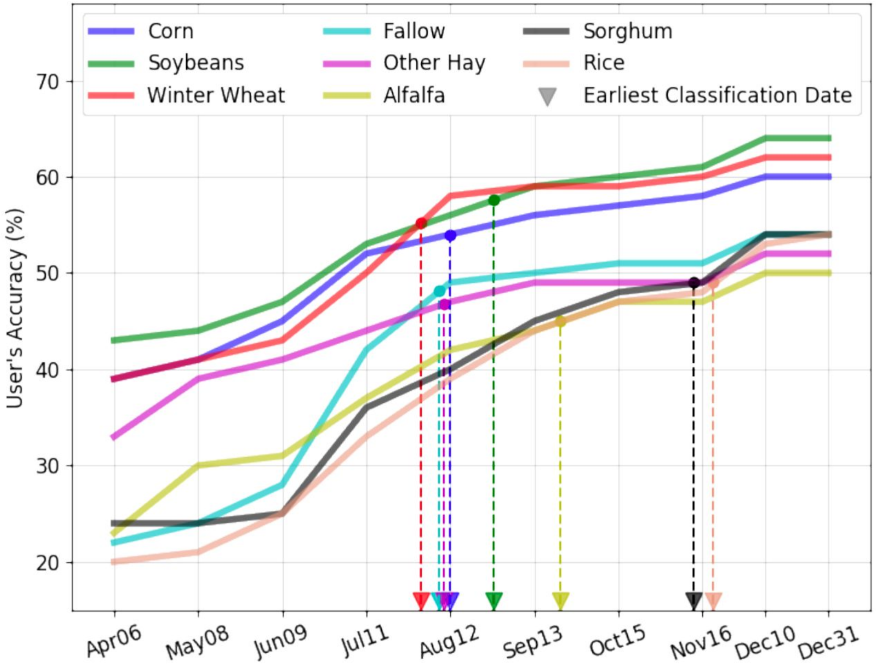


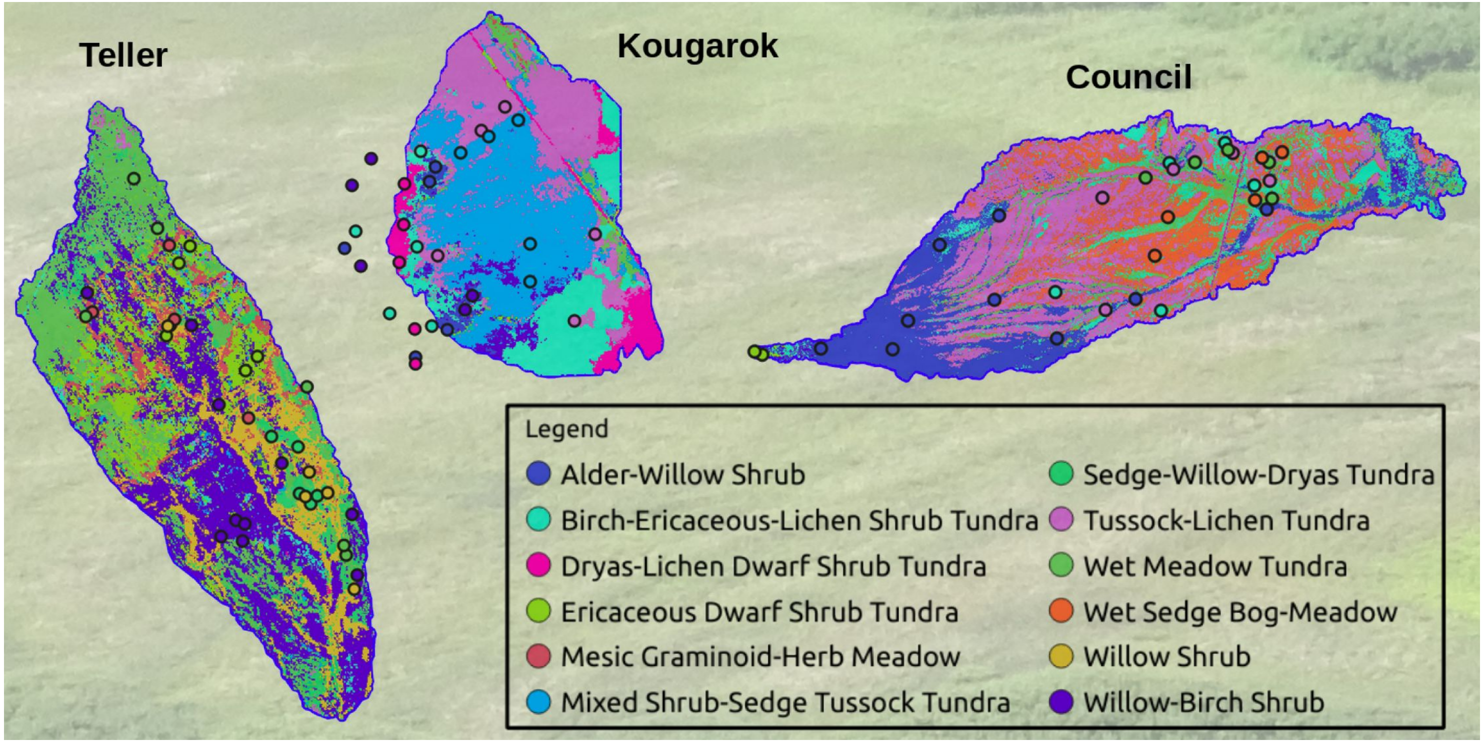
Figure: a) Comparison of cluster-then-label crop map with USDA Crop Data Layer (CDL) shows similar patterns at continental scale. b) Good spatial agreement is found at three selected regions, but cluster-then-label crop maps lack sharpness at field boundaries due to coarser resolution of MODIS data.

Earliest date for crop type classification



Konduri, V. S., J. Kumar, W. W. Hargrove, F. M. Hoffman, and A. R. Ganguly (2020), Mapping Crops Within the Growing Season Across the United States, *Remote Sens. Environ.*, 251, 112048, doi:[10.1016/j.rse.2020.112048](https://doi.org/10.1016/j.rse.2020.112048).

Watershed-Scale Plant Communities Determined from DNN and AVIRIS-NG



At the watershed scale, vegetation community distribution follows topographic and water controls.
At a fine scale, nutrients limit the distribution of vegetation types.

(Konduri et al., in prep.)

Leveraging Advances in Machine Learning for Earth Sciences

Existing machine learning techniques can improve understanding of biospheric processes and representation in Earth system models

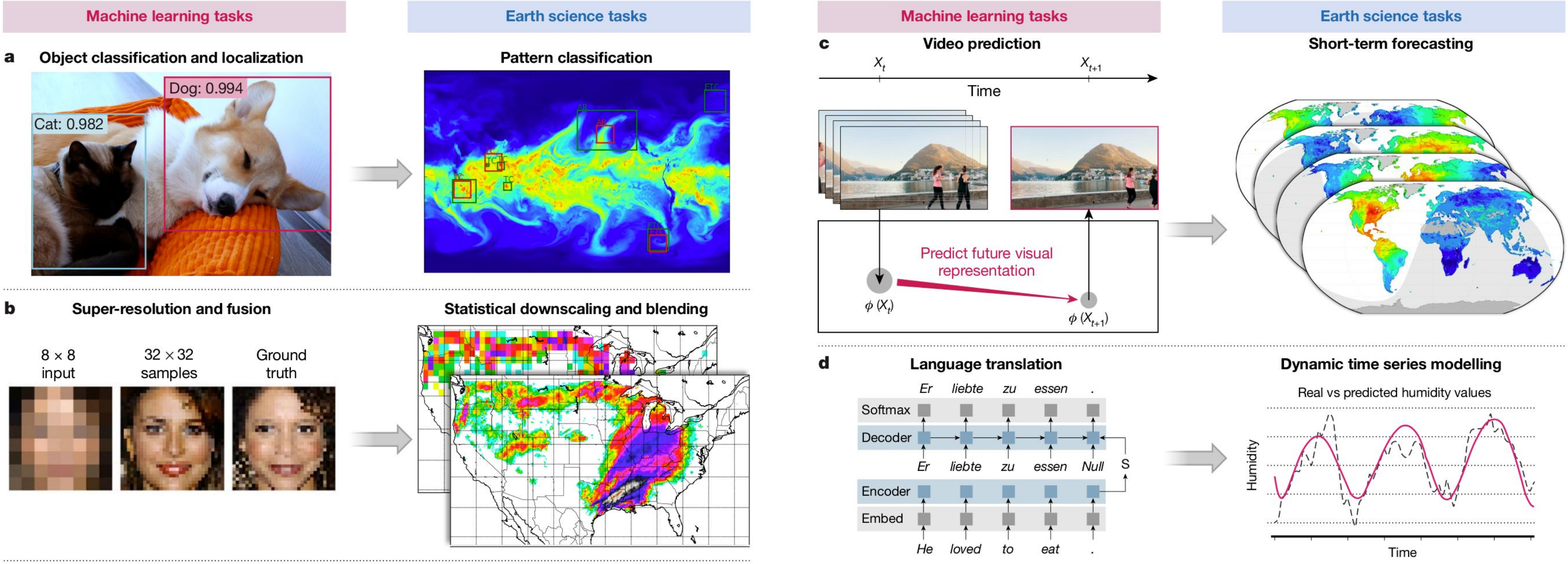
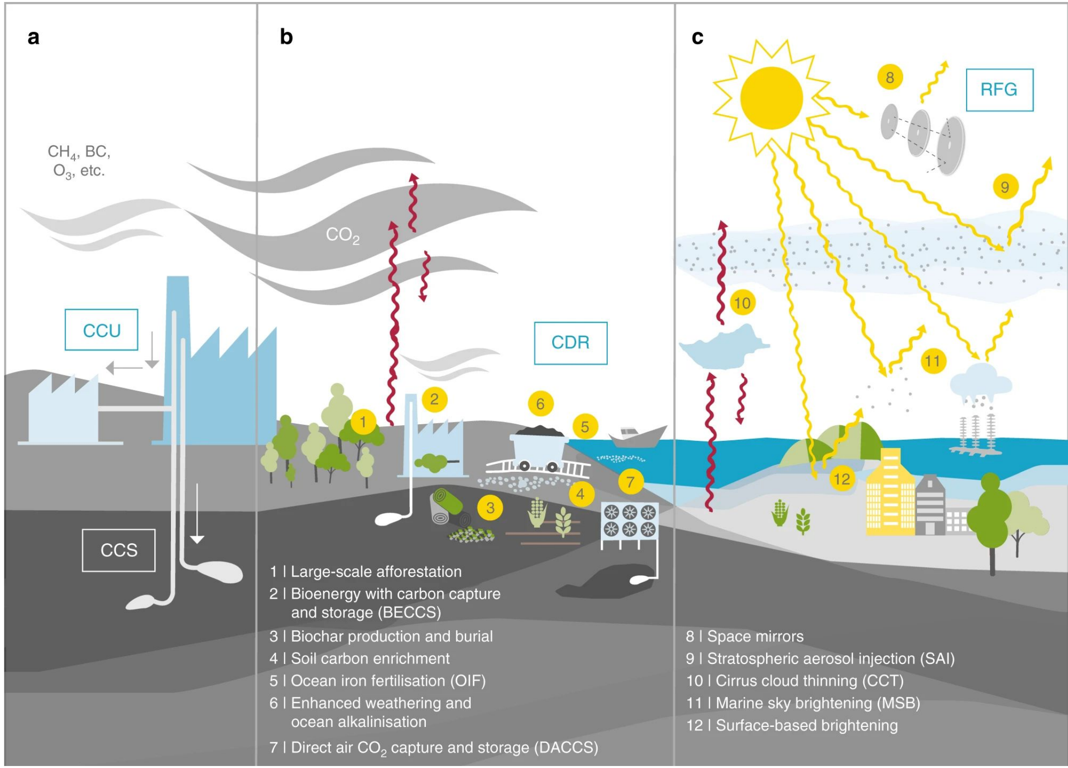


Figure 2 in Reichstein et al. (2019)

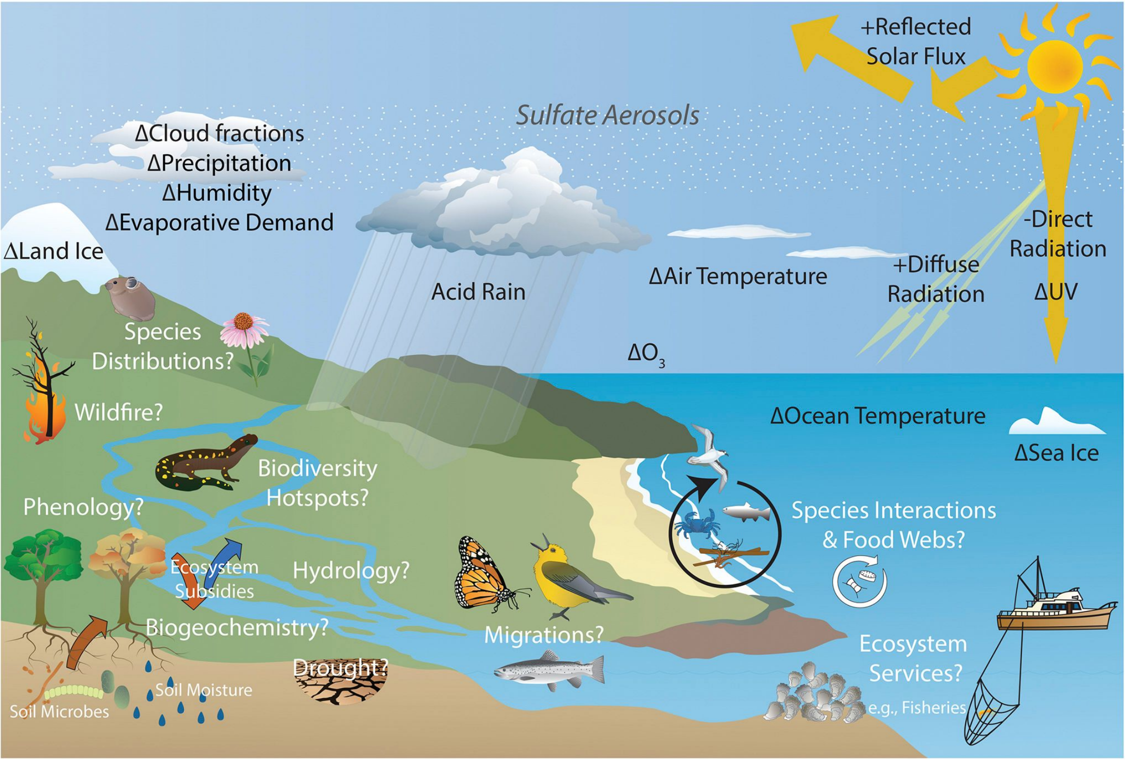
Climate Change Mitigation through Climate Intervention

- The increasing severity of extreme events and wildfire is threatening utilities, built infrastructure, and economic & national security
- Loss of life and property is motivating consideration of *climate intervention* or *geoengineering*
- In addition to *carbon dioxide removal (CDR)* through *direct air capture (DAC)* and other means, interest is growing in reducing or stabilizing Earth's surface temperature
- *Solar radiation management (SRM)* is an approach to partially reduce warming, and *stratospheric aerosol intervention (SAI)* by injecting sulfur into the lower stratosphere is considered the most feasible scheme



A wide variety of natural solutions and geoengineering techniques are proposed for mitigating the effects of climate change. Adopted from Lawrence et al. (2018).

Potential Ecological Impacts of Climate Intervention



Although some effects of SRM with SAI on climate are known from certain SAI scenarios, the effects of SAI on ecological systems are largely unknown. Adopted from Zarnetske et al. (2021).

- While climate research has focused on predicted climate effects of SRM, few studies have investigated impacts that SRM would have on ecological systems
- Impacts and risks posed by SRM would vary by implementation scenario, anthropogenic climate effects, geographic region, and by ecosystem, community, population, and organism
- A transdisciplinary approach is essential, and **new modeling paradigms are required**, to represent complex interactions across Earth system components, scales, and ecological systems

Geoengineering Increases the Global Land Carbon Sink

Objective: To examine stratospheric aerosol intervention (SAI) impacts on plant productivity and terrestrial biogeochemistry.

Approach: Analyze and compare simulation results from the Stratospheric Aerosol Geoengineering Large Ensemble (GLENS) project from 2010 to 2097 under RCP8.5 with and without SAI.

Results/Impacts: In this scenario, SAI causes terrestrial ecosystems to store an additional 79 Pg C globally as a result of lower ecosystem respiration and diminished disturbance effects by the end of the 21st century, yielding as much as a 4% reduction in atmospheric CO_2 mole fraction that progressively reduces the SAI effort required to stabilize surface temperature.

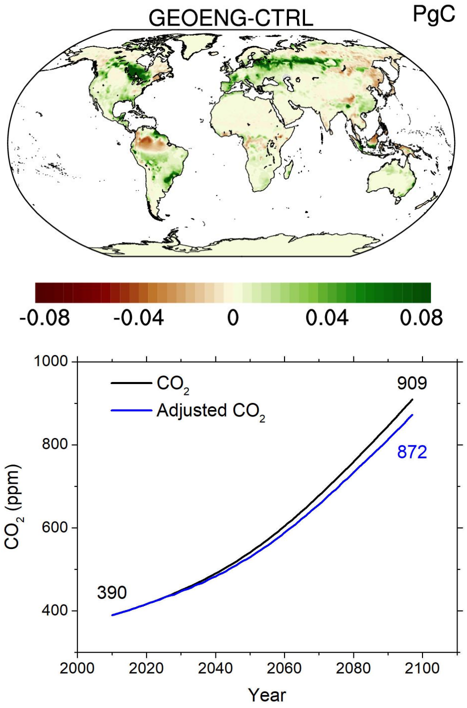
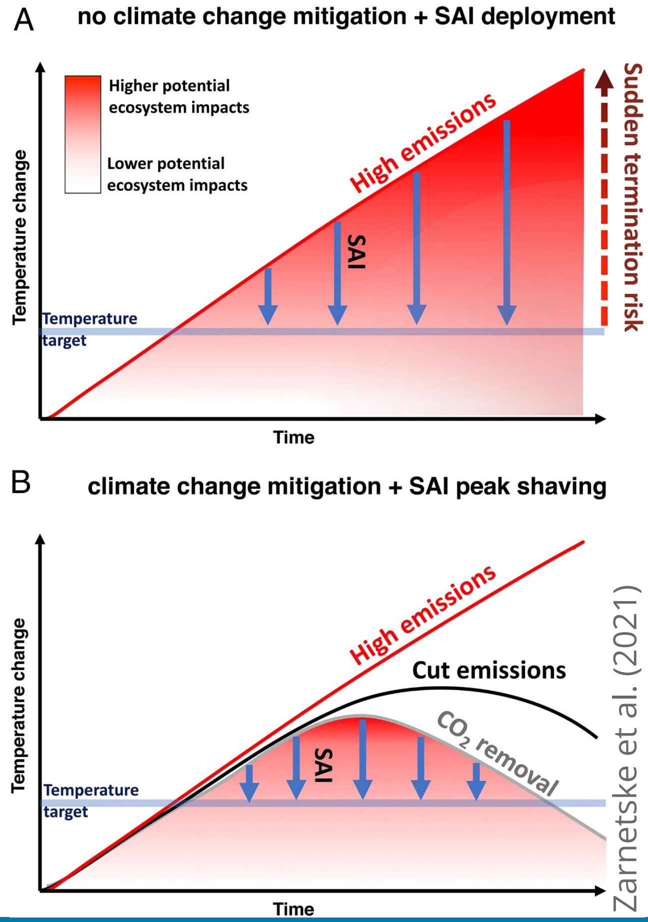


Figure: The larger sink under SAI increased land C storage by 79 Pg C by 2097, which would reduce the projected atmospheric CO_2 level.

Yang, C.-E., **F. M. Hoffman**, D. M. Ricciuto, S. Tilmes, L. Xia, D. G. MacMartin, B. Kravitz, J. H. Richter, M. Mills, and J. S. Fu (2020), Assessing Terrestrial Biogeochemical Feedbacks in a Strategically Geoengineered Climate, *Environ. Res. Lett.*, doi:[10.1088/1748-9326/abacf7](https://doi.org/10.1088/1748-9326/abacf7).

Exploring Feedbacks of SAI

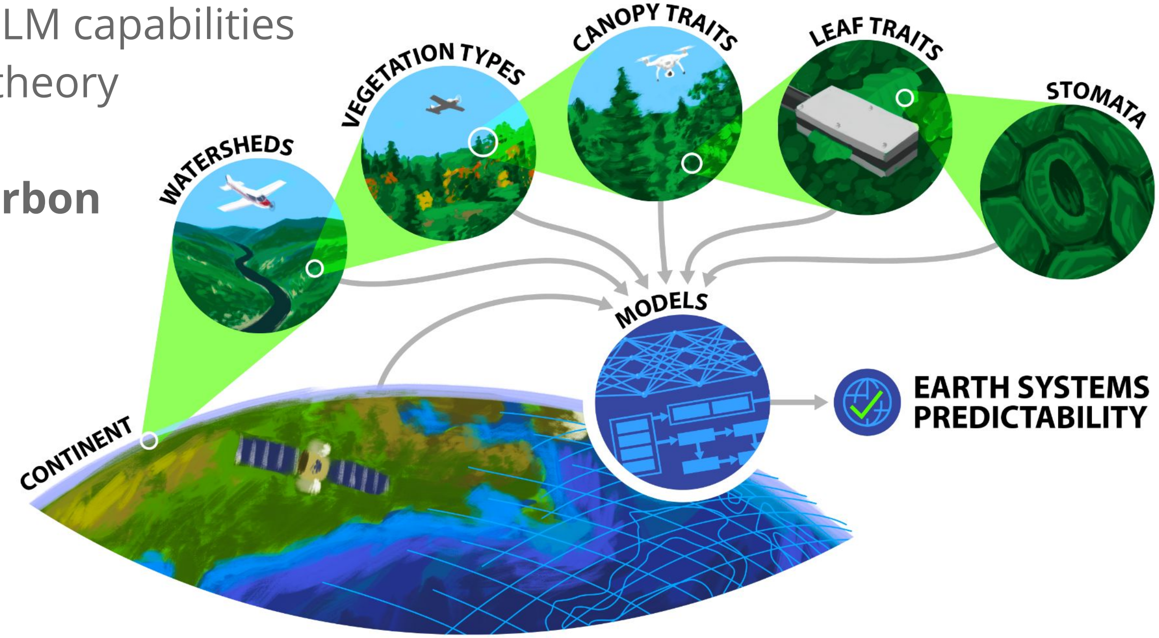
- To fill research gaps in understanding Earth system feedbacks of SAI on ecosystems, we are conducting a series of increasingly complex geoengineering simulations with **DOE's Energy Exascale Earth System Model (E3SM)**
- **Simulations will mimic effects of CDR, SAI, and CDR plus SAI**
- Start with SSP5-3.4-OS mid-range overshoot CO_2 trajectory from CMIP6, which prescribes a drawdown of CO_2
- Global surface temperatures will rise by >2.5°C around 2040, **above the 2°C threshold that may induce irreversible impacts**
- Next, introduce SAI to simultaneously cool the surface until drawdown is sufficient to assure < 2°C warming, called **temperature "peak shaving"**
- To quantify feedbacks from reducing, not increasing, atmospheric CO_2, **but may not capture all the as yet unobserved processes**



Zarnetske et al. (2021)

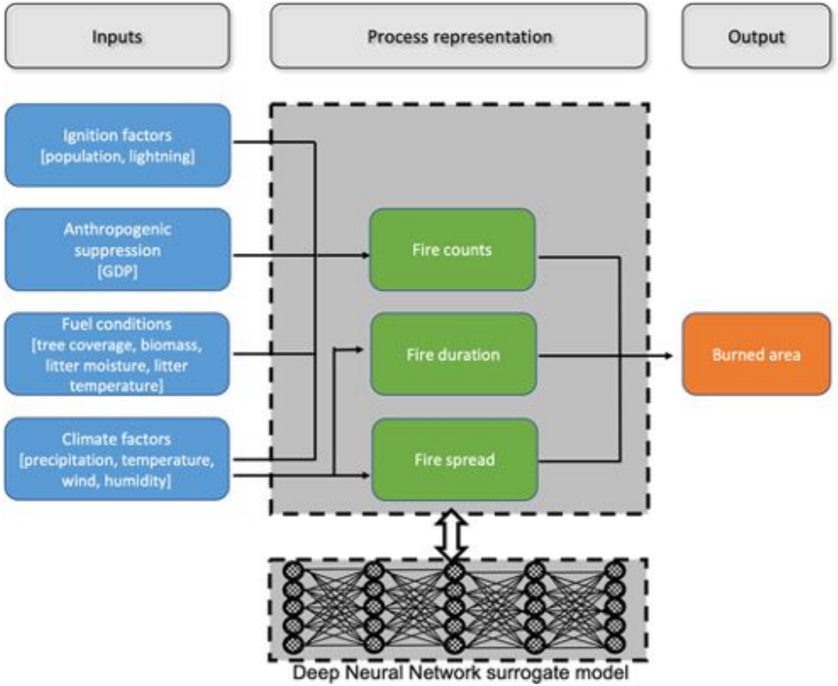
Machine Learning for Understanding Biospheric Processes

- Widening adoption of deep neural networks and growth of climate data are fueling interest in AI/ML for use in weather and climate and Earth system models
- ML potential is high for improving predictability when (1) *sufficient data are available for process representations* and (2) *process representations are computationally expensive*
- Example methods for improving ELM capabilities by exploring ML and information theory approaches:
  - **Soil organic carbon & radiocarbon**
  - **Wildfire**
  - **Methane emissions**
  - **Ecohydrology**
- All of these applications involve unresolved, subgrid-scale processes that strongly influence results at the largest scales



Hybrid Modeling of Wildfire Activities

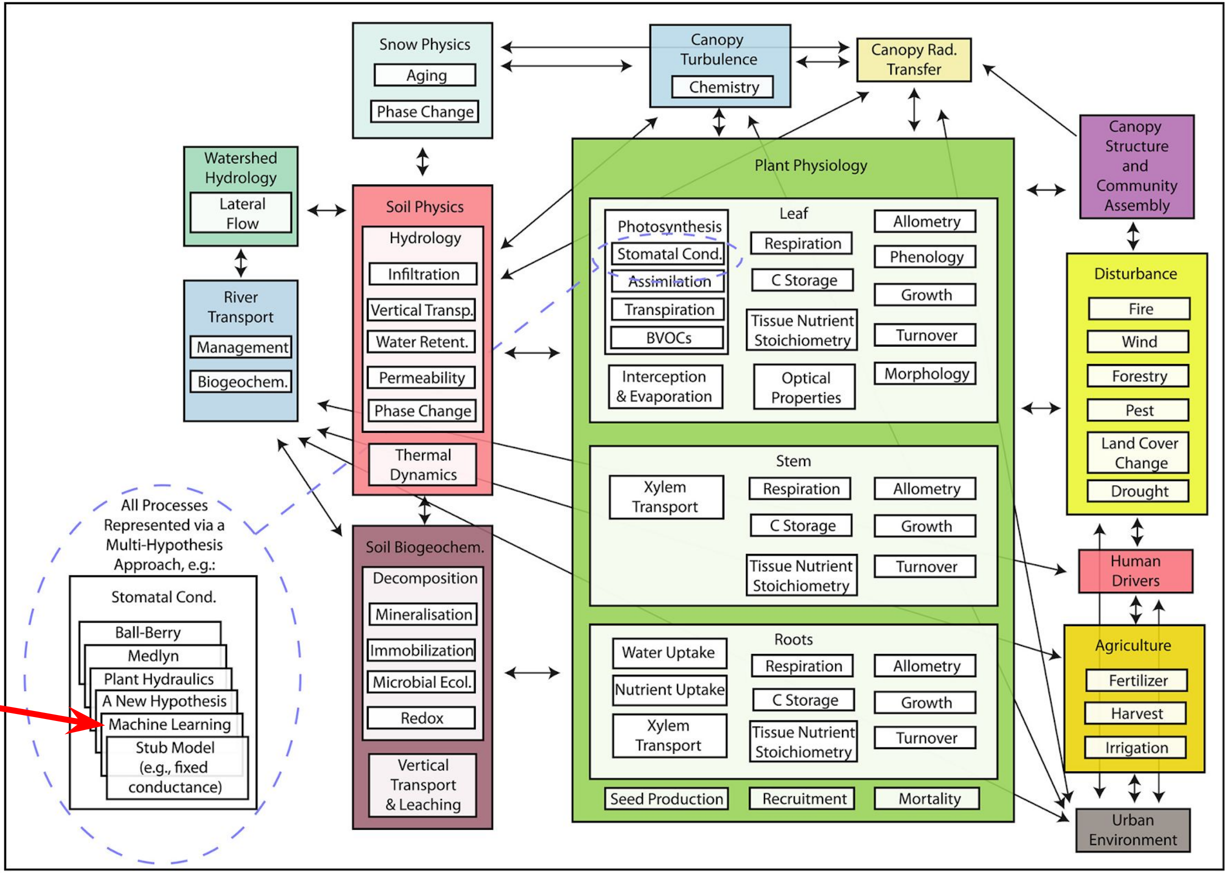
- Improve model simulations of **wildfire processes**, including ignition, fire duration, and spread rate with Deep Neural Network models
- Improve simulated **wildfire emissions** and their impacts on atmospheric properties, including aerosols, greenhouse gases, phosphorus transport, and pollutants
- Improve the projection of near-future and long-term dynamics of wildfire activities
- Accelerate E3SM coupled land–atmosphere modeling activities for wildfire research
- Explore online ML training/validation strategy for E3SM coupled model simulations



Zhu et al. (2022)

Hybrid ML/Process-based Modeling for Terrestrial Modeling

Individual processes can be represented in a multi-hypothesis approach, and ML provides an opportunities for (1) a model surrogate module or (2) a data-derived module that can be further explored or used to calibrate other hypotheses, when sufficient data are available.



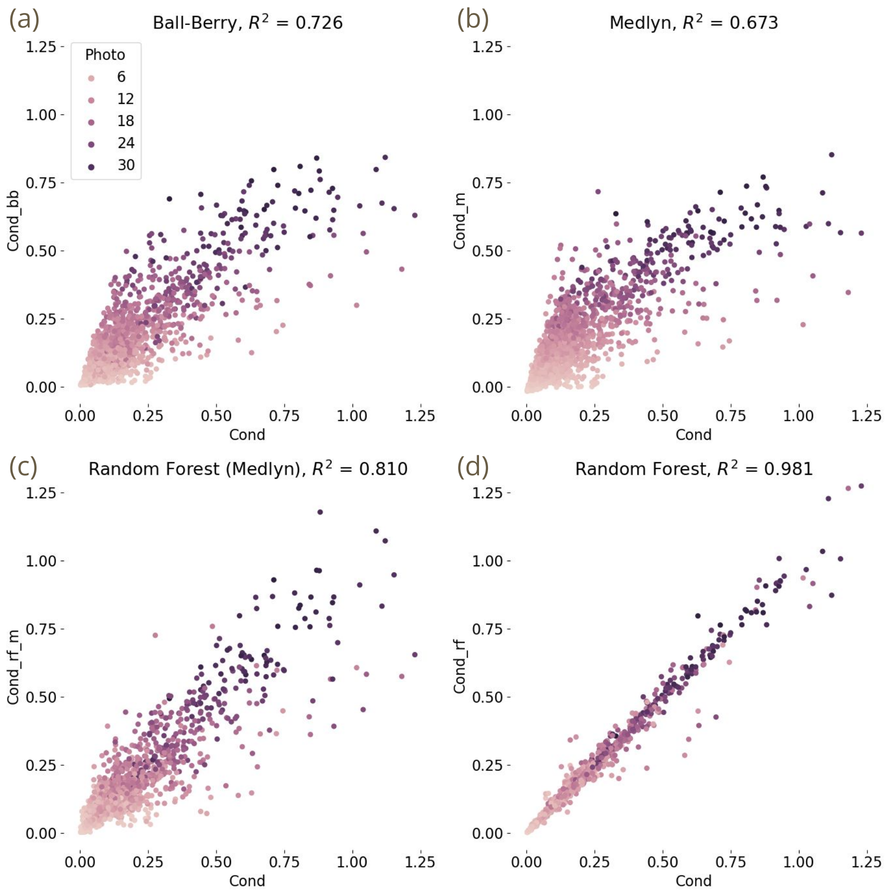
(a) Process Schematic of a Possible Full-Complexity Configuration of a Land Surface Model

(Fisher and Koven, 2020)

Hybrid Modeling of Photosynthesis and Ecohydrology

- Significant leaf-level data may be used to train ML parameterizations to **improve accuracy** and **computational performance**
- **Estimated stomatal conductance** vs. measured stomatal conductance for (a) Ball-Berry, (b) Medlyn, (c) Random forest (with Medlyn inputs), and (d) Random forest with all inputs from Lin et al. (2015)
- Inputs to the Medlyn parameterization are leaf-level CO_2, photosynthesis, and vapor pressure deficit
- Random forest trained on these three inputs (c) performs slightly better than Medlyn
- Random forest trained on more variables (d) achieves an R^2 of 0.98

(Massoud, Collier, et al. in prep)

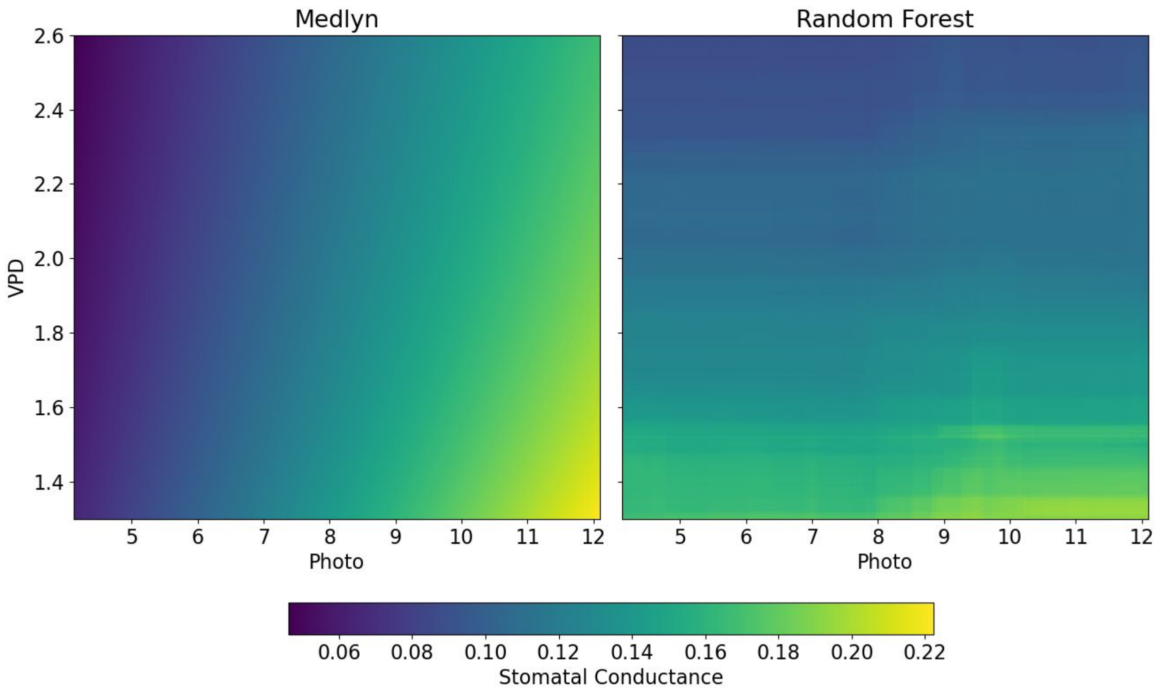


Hybrid Modeling of Photosynthesis and Ecohydrology

- Most process-based or empirical formulations are continuous
- But ML formulations may exhibit discontinuities in the multi-dimensional space of inputs because of out-of-sample data or artifacts of sampling or precision

- For example, we can see such discontinuities at right for Random Forest in the VPD vs. photosynthesis heat map for stomatal conductance

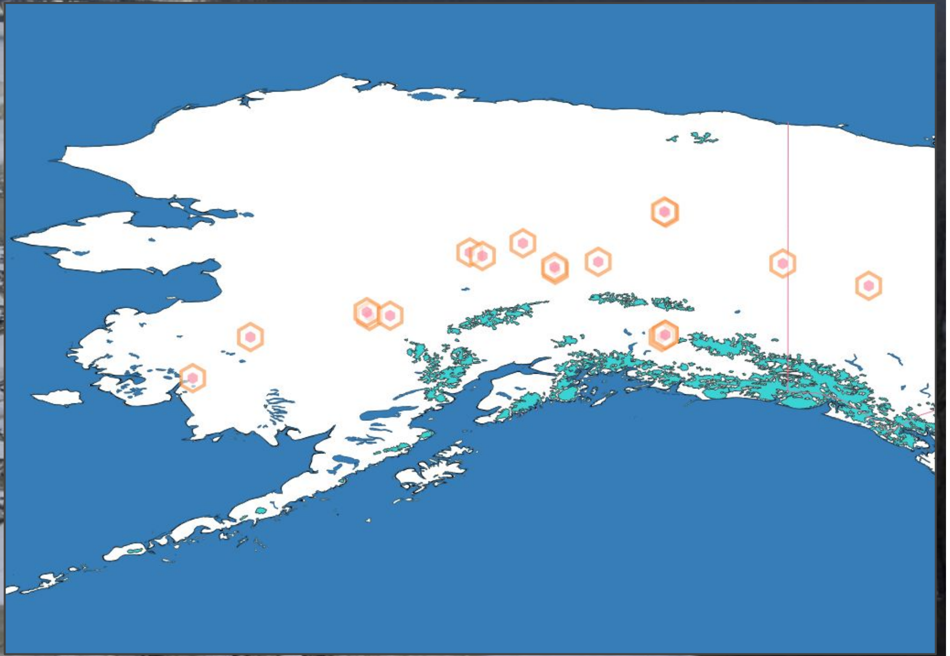
- These discontinuities are likely to have numerical consequences when attempting to couple a ML parameterization into a hybrid empirical / ML Earth system model



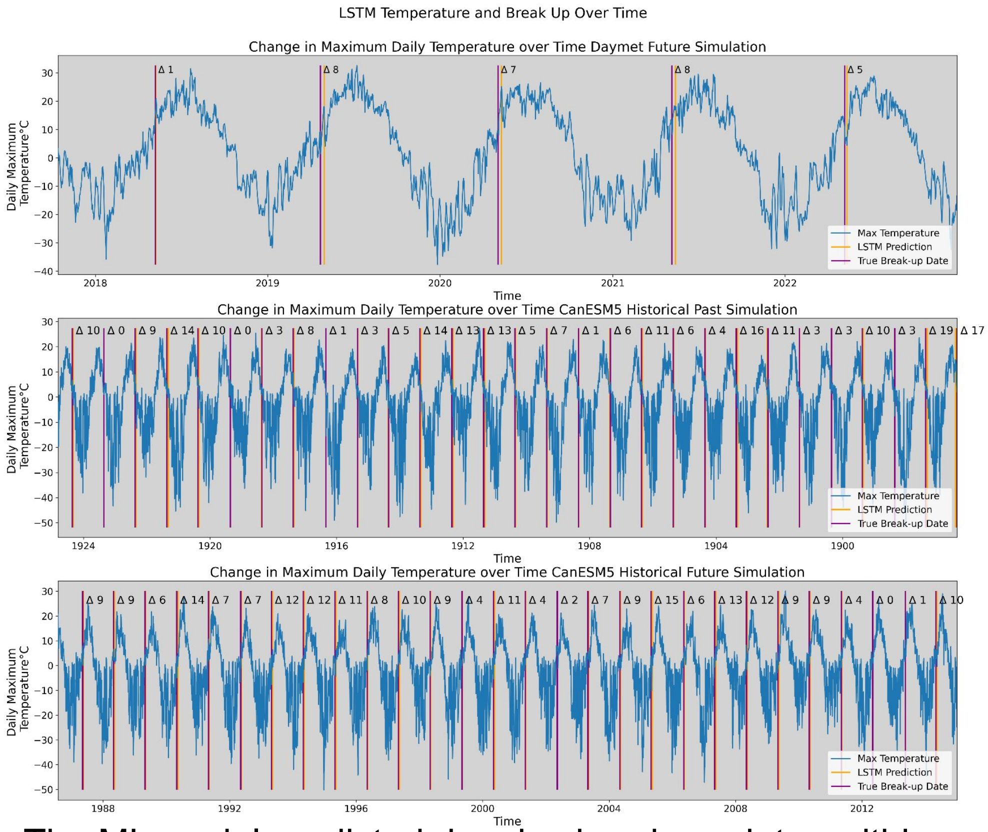
(Massoud, Collier, et al. in prep)

Forecasting River Ice Breakup using LSTM

- Study sites were selected at long term river ice monitoring stations in the Yukon river basin
- We developed Long Short Term Memory (LSTM) models to predict river ice breakups
- Primary predictor variables: daily min/max air temp., precipitation, snow water equiv., shortwave radiation
- Datasets: DAYMET, CanESM5 (Historical, SSP119, SSP370, SSP585, SSP534-over)

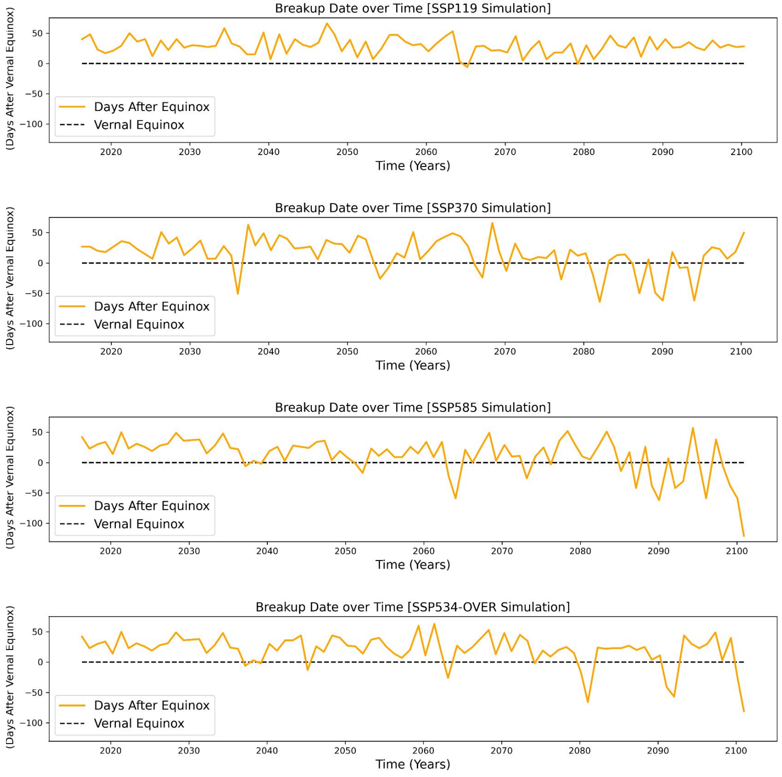


Break-up date predictions for historical period



The ML model predicted river ice break-up dates within 1–14 days of observed dates

Break-up date predictions under future scenarios



Projections suggested increasingly early break-up of river ice under warming scenarios https://ai4esp.org/ https://ai4esp.slack.com/

AI4ESP | Artificial Intelligence for Earth System Predictability

A multi-lab initiative working with the Earth and Environmental Systems Science Division (EESSD) of the Office of Biological and Environmental Research (BER) to develop a new paradigm for Earth system predictability focused on enabling artificial intelligence across field, lab, modeling, and analysis activities.

White papers were solicited for development and application of AI methods in areas relevant to EESSD research with an emphasis on quantifying and improving Earth system predictability, particularly related to the integrative water cycle and extreme events.

How can DOE directly leverage artificial intelligence (AI) to engineer a substantial (paradigm-changing) improvement in Earth System Predictability?

156 white papers were received and read to plan the organization of the **AI4ESP Workshop on Oct 25–Dec 3, 2021**

U.S. DEPARTMENT OF ENERGY

Earth System Predictability Sessions

- Atmospheric Modeling
- Land Modeling
- Human Systems & Dynamics
- Hydrology
- Watershed Science
- Ecohydrology
- Aerosols & Clouds
- Climate Variability & Extremes
- Coastal Dynamics, Oceans & Ice

Cross-Cut Sessions

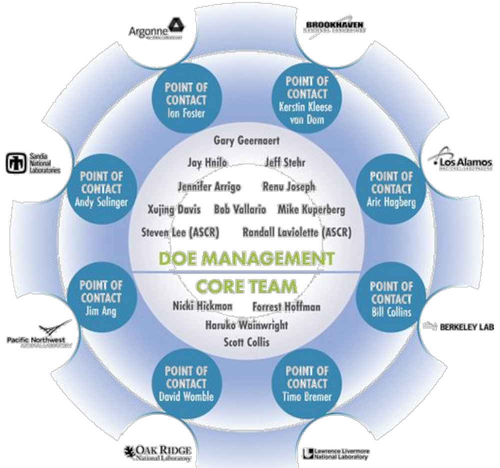
- Data Acquisition
- Neural Networks
- Surrogate models and emulators
- Knowledge-Informed Machine Learning
- Hybrid Modeling
- Explainable/Interpretable/Trustworthy AI
- Knowledge Discovery & Statistical Learning
- AI Architectures and Co-design

Workshop Report

- Posted on ai4esp.org
- Executive Summary
- Long summary
- Earth science chapters
- Computational science chapters

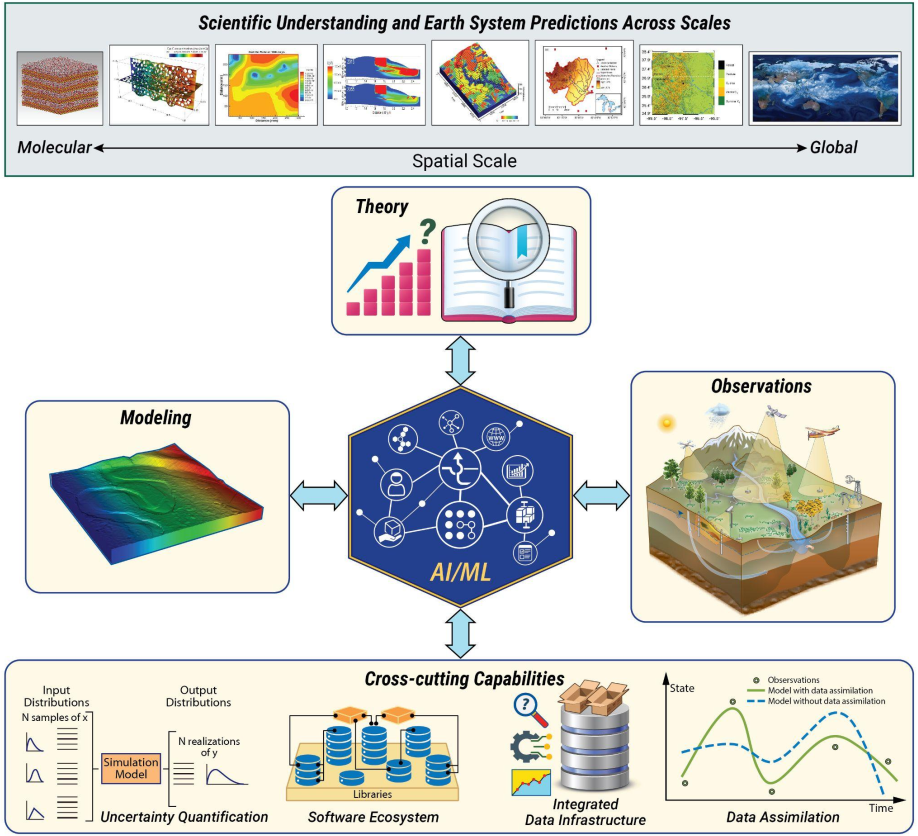
AMS Special Collection

- Open submissions for new AI for the Earth Systems journal



AI4ESP

AI4ESP WORKSHOP HIGHLIGHTS



EESA22-032

AI4ESP WORKSHOP HIGHLIGHTS

Overview of priorities emerging from the AI4ESP workshop across 3 key themes.

These priorities will help address major challenges for Earth system predictability

Earth Science Priorities

- New observations
- AI-ready data products
- Data-driven and hybrid models
- Analytical approaches
- Uncertainty quantification, model parametrization & calibration

To Tackle Challenges

- Significant data gaps
- Scaling and heterogeneity
- Extreme events
- Representation of human activities
- Knowledge discovery
- Accurate high-resolution predictions with low bias, uncertainty
- Providing actionable, timely information for decision making

Computational Science Priorities

- Hybrid models
- Fundamental math and algorithms
- Interpretable, trustworthy AI
- AI-enabled data acquisition
- Data, software, hardware infrastructure

To Tackle Challenges

- Physically consistent predictions for data-driven models
- Computational costs of process models
- Sparse data, extreme values
- Identifying causality
- Interpretable, trustworthy predictions
- Data discovery, access, synthesis
- Model development and comparison

Programmatic and Cultural Priorities

- AI research centers
- Workforce development
- Codesign infrastructure
- Common standards, benchmarks
- Seed projects, integrate AI into programs
- AI ethics and policies

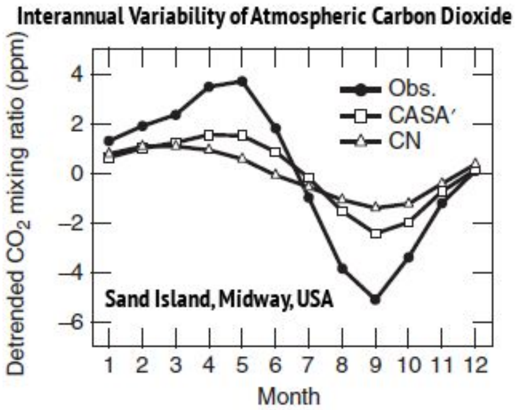
To Tackle Challenges

- Interdisciplinary scientific research
- Diverse organizational missions
- Personnel lack training in AI/ML
- Using data, communicating across research domains, organizations
- Data bias, model fairness, explainability of predictions

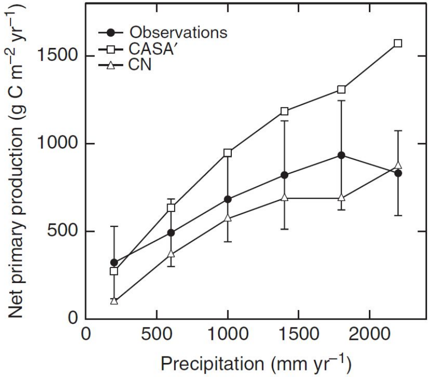
International Land Model Benchmarking (ILAMB)

What is a Benchmark?

- A **benchmark** is a quantitative test of model function achieved through comparison of model results with observational data
- Acceptable performance on a benchmark **is a necessary but not sufficient condition** for a fully functioning model
- **Functional relationship benchmarks** offer tests of model responses to forcings and yield insights into ecosystem processes
- Effective benchmarks must draw upon **a broad set of independent observations** to evaluate model performance at multiple scales



Models often fail to capture the amplitude of the seasonal cycle of atmospheric CO_2



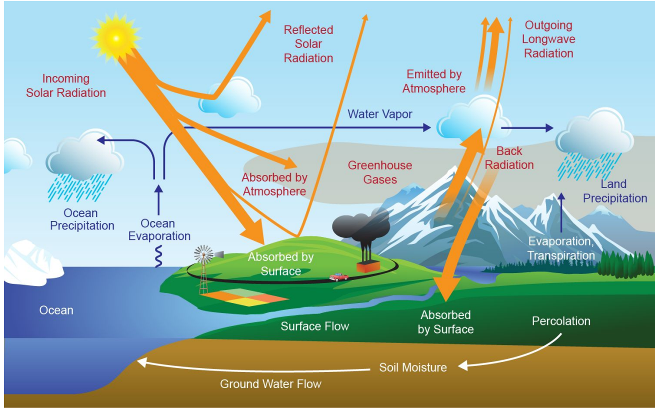
(Randerson et al., 2009)

Models may reproduce correct responses over only a limited range of forcing variables

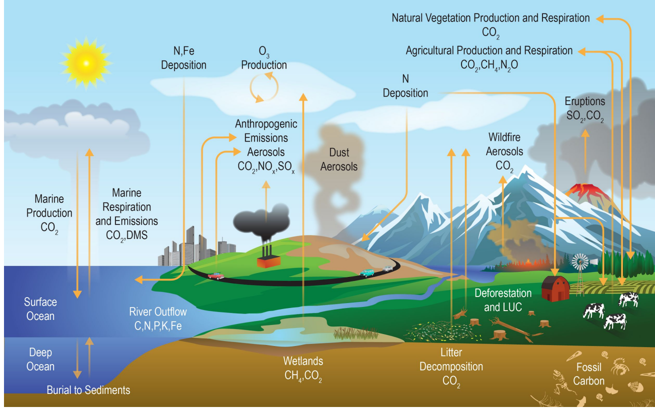
What is ILAMB?

A community coordination activity created to:

- **Develop internationally accepted benchmarks** for land model performance by drawing upon collaborative expertise
- **Promote the use of these benchmarks** for model intercomparison
- **Strengthen linkages between experimental, remote sensing, and Earth system modeling communities** in the design of new model tests and new measurement programs
- **Support the design and development of open source benchmarking tools**



Energy and Water Cycles

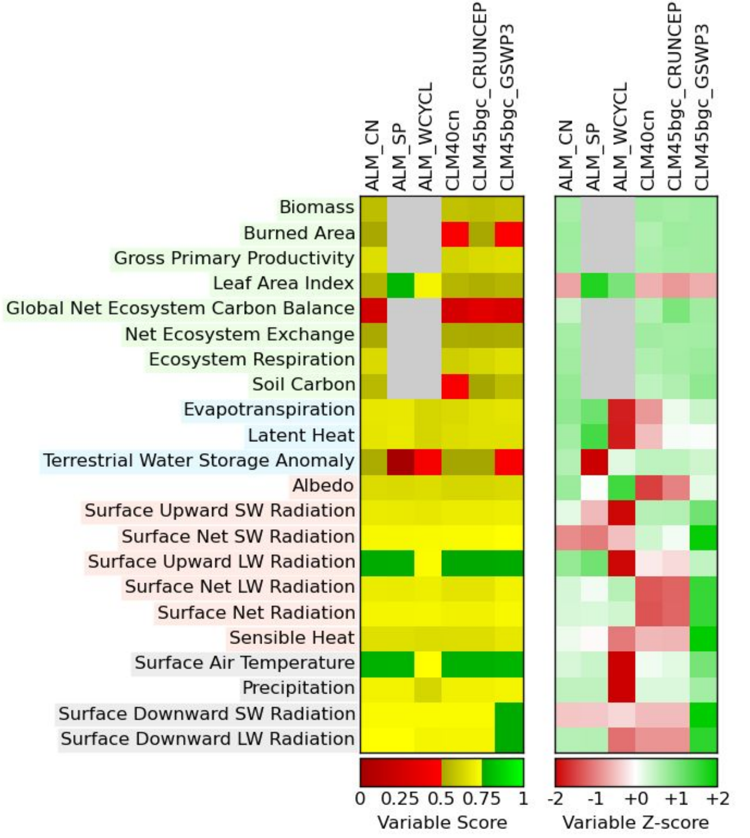


Carbon and Biogeochemical Cycles

Argonne NATIONAL LABORATORY
BROOKHAVEN NATIONAL LABORATORY
BERKELEY LAB
Los Alamos NATIONAL LABORATORY
NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH
OAK RIDGE National Laboratory
UCI
MICHIGAN

Development of ILAMB Packages

- **ILAMBv1** released at 2015 AGU Fall Meeting Town Hall, doi:[10.18139/ILAMB.v001.00/1251597](https://doi.org/10.18139/ILAMB.v001.00/1251597)
- **ILAMBv2** released at 2016 ILAMB Workshop, doi:[10.18139/ILAMB.v002.00/1251621](https://doi.org/10.18139/ILAMB.v002.00/1251621)
- **Open Source software** written in Python; **runs in parallel** on laptops, clusters, and supercomputers
- Routinely used for land model evaluation during development of ESMs, including the **E3SM Land Model** (Zhu et al., 2019) and the **CESM Community Land Model** (Lawrence et al., 2019)
- **Models are scored** based on statistical comparisons and functional response metrics



Argonne National Laboratory · Brookhaven National Laboratory · Berkeley Lab · Los Alamos National Laboratory · NCAR National Center for Atmospheric Research · Oak Ridge National Laboratory · UCI · Michigan

ILAMB Produces Diagnostics and Scores Models

- ILAMB generates a top-level **portrait plot** of models scores
- For every variable and dataset, ILAMB can automatically produce
  - **Tables** containing individual metrics and metric scores (when relevant to the data), including
    - Benchmark and model **period mean**
    - **Bias** and **bias score** (S_{bias})
    - **Root-mean-square error (RMSE)** and **RMSE score** (S_{rmse})
    - **Phase shift** and **seasonal cycle score** (S_{phase})
    - **Interannual coefficient of variation** and **IAV score** (S_{iav})
    - **Spatial distribution score** (S_{dist})
    - **Overall score** (S_{overall}) → $$S_{\text{overall}} = \frac{S_{\text{bias}} + 2S_{\text{rmse}} + S_{\text{phase}} + S_{\text{iav}} + S_{\text{dist}}}{1+2+1+1+1}$$
  - **Graphical diagnostics**
    - Spatial contour maps
    - Time series line plots
    - Spatial Taylor diagrams (Taylor, 2001)
- Similar **tables** and **graphical diagnostics** for functional relationships

RUBISCO

Argonne NATIONAL LABORATORY · BROOKHAVEN NATIONAL LABORATORY · BERKELEY LAB · Los Alamos NATIONAL LABORATORY · NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH · OAK RIDGE National Laboratory · UCI · MICHIGAN

RUBISCO

ILAMBv2.6 Package Current Variables

- **Biogeochemistry:** Biomass (Contiguous US, Pan Tropical Forest), Burned area (GFED3), CO_2 (NOAA GMD, Mauna Loa), Gross primary production (Fluxnet, GBAF), Leaf area index (AVHRR, MODIS), Global net ecosystem carbon balance (GCP, Khatiwala/Hoffman), Net ecosystem exchange (Fluxnet, GBAF), Ecosystem Respiration (Fluxnet, GBAF), Soil C (HWSD, NCSCDv22, Koven)
- **Hydrology:** Evapotranspiration (GLEAM, MODIS), Evaporative fraction (GBAF), Latent heat (Fluxnet, GBAF, DOLCE), Runoff (Dai, LORA), Sensible heat (Fluxnet, GBAF), Terrestrial water storage anomaly (GRACE), Permafrost (NSIDC)
- **Energy:** Albedo (CERES, GEWEX.SRB), Surface upward and net SW/LW radiation (CERES, GEWEX.SRB, WRMC.BSRN), Surface net radiation (CERES, Fluxnet, GEWEX.SRB, WRMC.BSRN)
- **Forcing:** Surface air temperature (CRU, Fluxnet), Diurnal max/min/range temperature (CRU), Precipitation (CMAP, Fluxnet, GPCC, GPCP2), Surface relative humidity (ERA), Surface down SW/LW radiation (CERES, Fluxnet, GEWEX.SRB, WRMC.BSRN)

Argonne NATIONAL LABORATORY · BROOKHAVEN NATIONAL LABORATORY · BERKELEY LAB · Los Alamos NATIONAL LABORATORY · NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH · OAK RIDGE National Laboratory · UCI · MICHIGAN

Multi-Model Validation Example

RUBISCO

Evaluation of CMIP5 vs CMIP6 with ILAMB and IOMB

- (a) ILAMB and (b) IOMB have been used to evaluate how land and ocean model performance has changed from CMIP5 to CMIP6
- Model fidelity is assessed through comparison of historical simulations with a wide variety of contemporary observational datasets
- The UN's Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) from Working Group 1 (WG1) Chapter 5 contains the full ILAMB/IOMB evaluation as Figure 5.22

	CMIP5 ESMs									CMIP6 ESMs										
	bcc-csm1-1	CanESM2	CESM1-BGC	GFDL-ESM2G	IPSL-CM5A-LR	MIROC-ESM	MPI-ESM-LR	NorESM1-ME	HadGEM2-ES	BCC-CSM2-MR	CanESM5	CESM2	GFDL-ESM4	IPSL-CM6A-LR	MIROC-ES2L	MPI-ESM1.2-LR	NorESM2-LM	UKESM1-0-LL	Mean CMIP5	Mean CMIP6
(a) Land Benchmarking Results																				
Land Ecosystem & Carbon Cycle	-0.72	-0.93	-1.55	-1.51	-0.13	0.60	-0.43	-1.31	0.19	-0.43	0.66	0.48	-1.09	0.22	0.60	-0.07	1.00	0.49	1.63	2.30
Biomass	0.20	-0.45	-1.52	-0.40	-1.26	-0.26	-1.07	-1.77	0.92	1.39	0.74	-0.20	-0.54	0.16	0.93	-0.96	-0.01	1.04	1.23	1.82
Burned Area			-0.87				0.10	-0.83				1.60								
Leaf Area Index	-0.20	-0.64	-1.30	-2.53	-0.01	0.30	0.01	-1.85	-0.16	0.27	0.08	0.34	-0.70	1.19	0.82	0.46	0.37	0.69	1.04	1.81
Soil Carbon	0.27	1.26	-1.46	0.07	0.75	0.47	-0.03	-1.14	0.07	0.23	1.35	-0.99	-2.04	-1.55	0.90	-0.75	-0.17	0.24	1.01	1.48
Gross Primary Productivity	0.59	-1.23	0.01	-1.81	-1.40	0.29	-0.53	-0.24	-1.04	0.77	0.04	0.59	-0.38	1.17	-1.02	-0.37	0.73	0.09	1.51	2.22
Net Ecosystem Exchange	-0.42	-1.81	-0.21	-0.65	1.10	-0.24	0.80	0.02	-1.03	-1.02	-1.19	0.59	1.69	-0.42	0.63	-0.21	1.08	-1.43	1.28	1.43
Ecosystem Respiration	0.90	-0.56	-0.86	-0.24	-1.35	0.99	-0.01	-0.94	-1.54	0.81	0.59	0.51	-0.79	0.90	-0.21	-1.24	0.43	-0.94	1.34	2.21
Carbon Dioxide		-1.54	-0.36	-2.92	-0.74	1.53	-0.00	0.37	0.85		0.42	0.26	0.39	0.59	1.10	-0.87	0.21	0.69	0.09	-0.07
Global Net Carbon Balance		-1.64	-0.88	-1.13	0.17	-0.31	-0.38	-0.50	0.24		-0.23	1.34	-1.70	0.17	-0.74	1.45	1.56	0.26	0.92	1.40
Land Hydrology Cycle	-2.65	-0.42	0.44	-0.18	-0.49	-0.52	-0.57	0.17	0.70	0.15	-0.47	1.51	-1.24	0.58	-0.72	-0.83	0.97	0.87	1.00	1.70
Evapotranspiration	-0.82	-0.99	-0.27	-1.02	0.64	-1.14	-0.62	-0.60	0.28	0.39	-1.08	1.09	0.65	0.43	-1.40	-1.01	0.82	1.05	1.41	2.20
Evaporative Fraction	-0.34	0.74	0.74	-0.14	-0.85	0.21	-1.98	0.22	-0.34	0.10	0.11	1.25	-0.88	1.29	-1.65	-1.81	1.11	-0.06	0.98	1.29
...																				
Terrestrial Water Storage Anomaly	-2.79	-0.45	0.47	0.50	-0.38	0.34	0.35	0.43	0.58	0.15	-0.08	0.95	-2.91	0.43	0.37	0.15	0.39	0.51	0.49	0.50
Permafrost	-0.88	-2.26	0.01	0.13	0.83	0.69	0.56	0.69	-0.56	-0.11	-3.02	0.83	0.74	-0.18	0.49	0.42	0.89	0.43	0.06	0.23
(b) Ocean Benchmarking Results																				
Ocean Ecosystems			2.18	0.20	-0.20		0.04		0.22		-0.37	0.83	-0.37	-0.26	-0.91	-0.67	-1.93	0.27	0.30	0.67
Chlorophyll		-1.50	2.15	0.44	1.02		0.49		0.56		-0.67	0.88	-0.21	0.10	-1.02	-0.41	-2.19	0.18	0.13	0.04
Oxygen, surface			0.73	-0.13	-1.98		-0.53	-1.53	-0.29		0.73	0.34	-0.09	-0.41	0.35	-0.30	0.40	0.49	0.64	1.57
Ocean Nutrients			-0.84	-0.10	0.91		-0.80	-1.25				-0.02	1.00	1.88		-0.90	-1.14	-0.17	-0.16	1.60
Nitrate, surface		0.21	-1.63	0.67	1.22		-0.18	-1.70	0.82		1.21	-0.90	0.29	1.21	1.02	0.39	-1.78	-0.56	-0.47	0.18
Phosphate, surface			-0.69	-0.04	0.04		-0.45	-0.43				0.39	-0.14	0.17	-0.41	-0.98	0.00	0.02	0.88	1.63
Silicate, surface			0.44	-0.71	0.24		-0.81	-0.20	-2.16			0.50	1.24	1.60		-1.21	-0.19	0.18	-0.29	1.37
Ocean Carbon											1.24	-0.23	-0.62	-0.69	-1.08	-1.12	1.31			1.19
TAlk, surface		-0.27	1.01	0.12	0.19		0.32	-2.31	-0.22		0.06	-0.36	0.85	-0.42	0.29	-2.40	1.27	0.06	1.27	0.54
...																				
Salinity, 700m	0.44	-0.35	-1.06	-0.54	0.70	0.46	-0.46	-0.80	0.32	0.36	0.25	-1.16	-0.47	0.54	0.33	-0.39	-0.87	-0.54	1.58	1.64
Ocean Relationships			-1.86	-0.36	-0.29		1.50	-0.43	0.68		-0.02	0.72	1.20	0.17	-1.86	0.02		-1.12	0.39	1.25
Oxygen, surface/WOA2018			0.27	0.23	-0.63		-0.26	-0.12	-0.38		0.29	-0.21	0.19	0.18	0.14	-0.07		0.03	-0.23	0.53
Nitrate, surface/WOA2018		-2.41	-1.38	-0.18	0.06		1.41	-0.16	0.78		0.09	0.79	1.07	0.26	-1.35	0.20		-0.74	0.52	1.04

Relative Scale

Worse Value — Better Value

Missing Data or Error

Conclusion

- Earth and environmental science data are rapidly increasing in volume, velocity, variety, voracity, and value
- Artificial intelligence approaches for data collection and machine learning methods for data management, reduction, gap-filling, extrapolation, and analysis and application are required
- For modeling, machine learning potential is high for improving predictability when (1) *sufficient data are available for process representations* and (2) *process representations are computationally expensive*
- Physics-informed machine learning and explainable artificial intelligence approaches are providing real value in improving both predictions and process understanding
- Machine learning models must be evaluated, just like process-based models, through comparison with observational data