Progress Report

Multi-regional evaluation of new machine learning algorithms for mapping tree species distribution and abundance

CAFS.20.79

Kasey Legaard[1], Aaron Weiskittel[1], Ken Bundy[1], Erin Simons-Legaard[2]

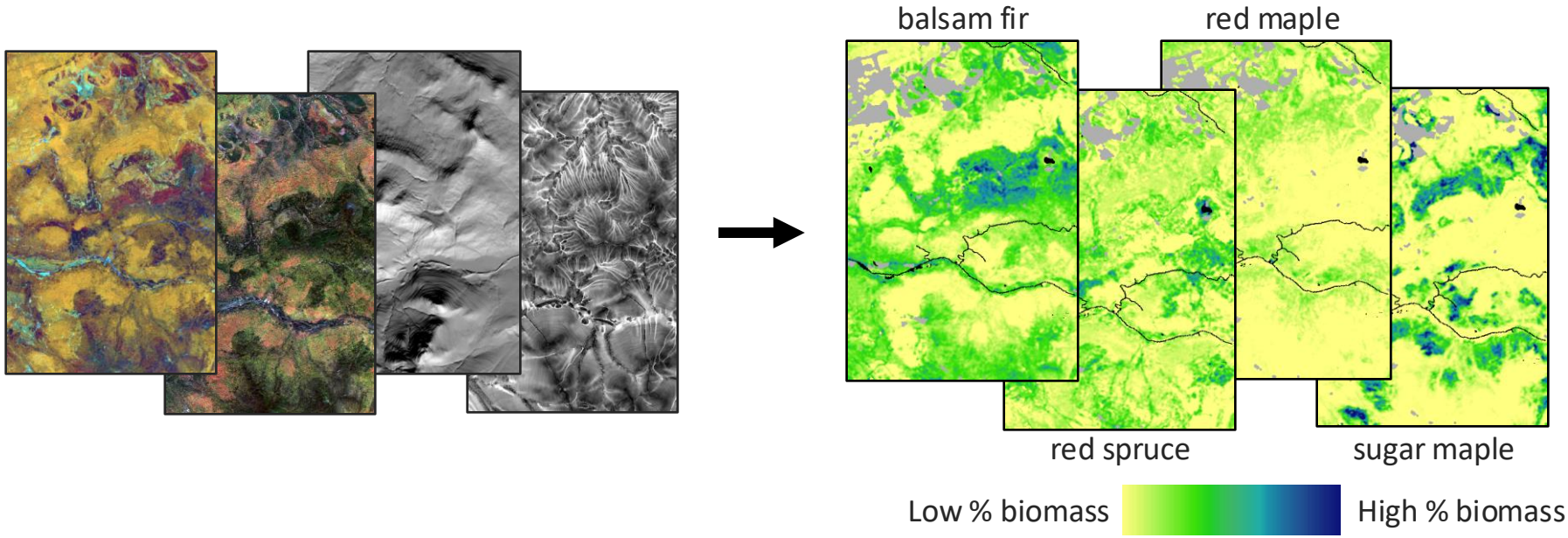
[1] University of Maine, Center for Research on Sustainable Forests
[2] University of Maine, School of Forest Resources

Presenter: Aaron Weiskittel

Justification

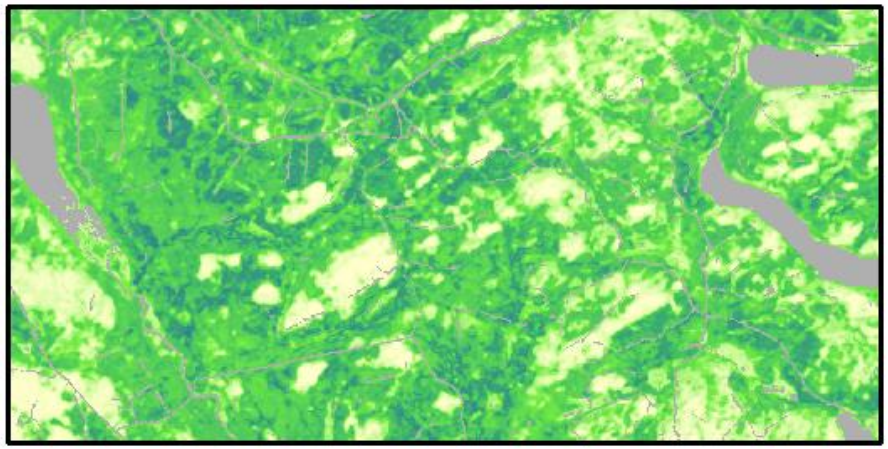
Forest remote sensing applications have tended to support pixel-level inference, with estimates of uncertainty that apply to pixels

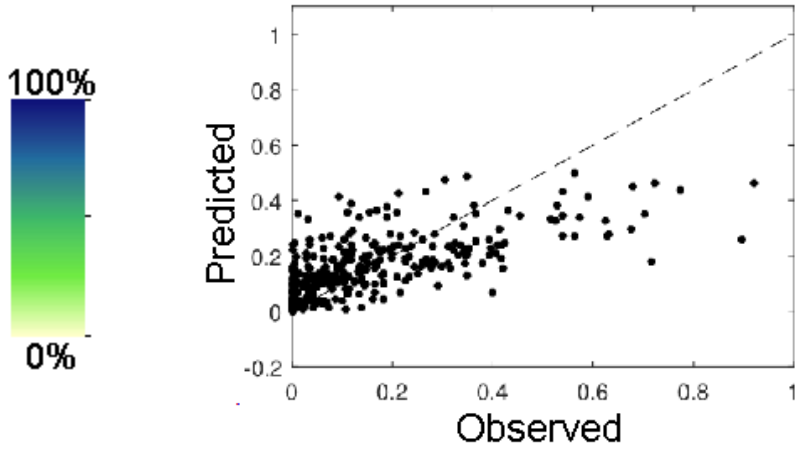
Species distributions from plot/pixel level machine learning:



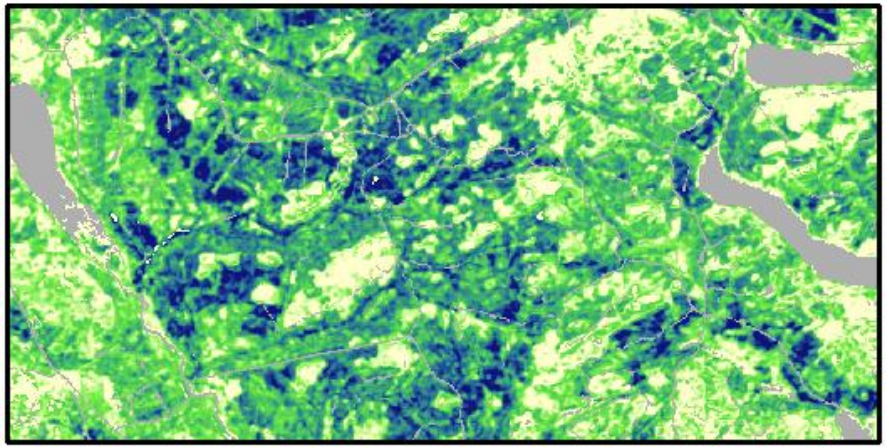
Justification

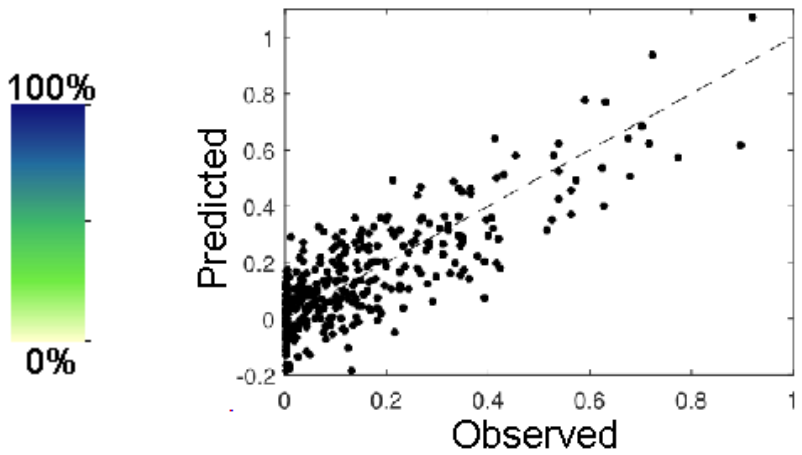
Balsam fir %AGLB, Random Forest:





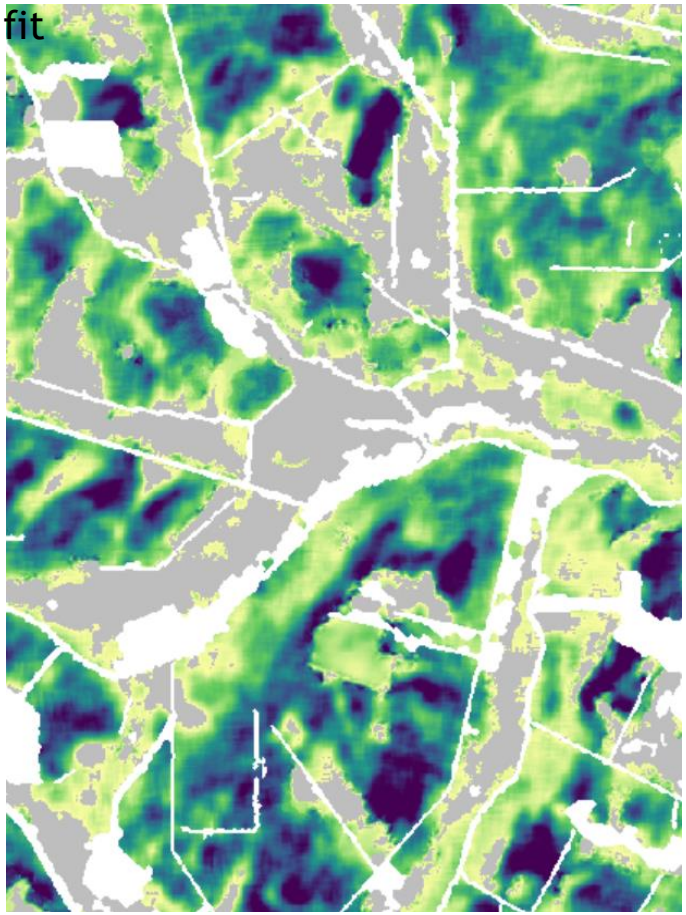
Balsam fir %AGLB, multi-objective ML:



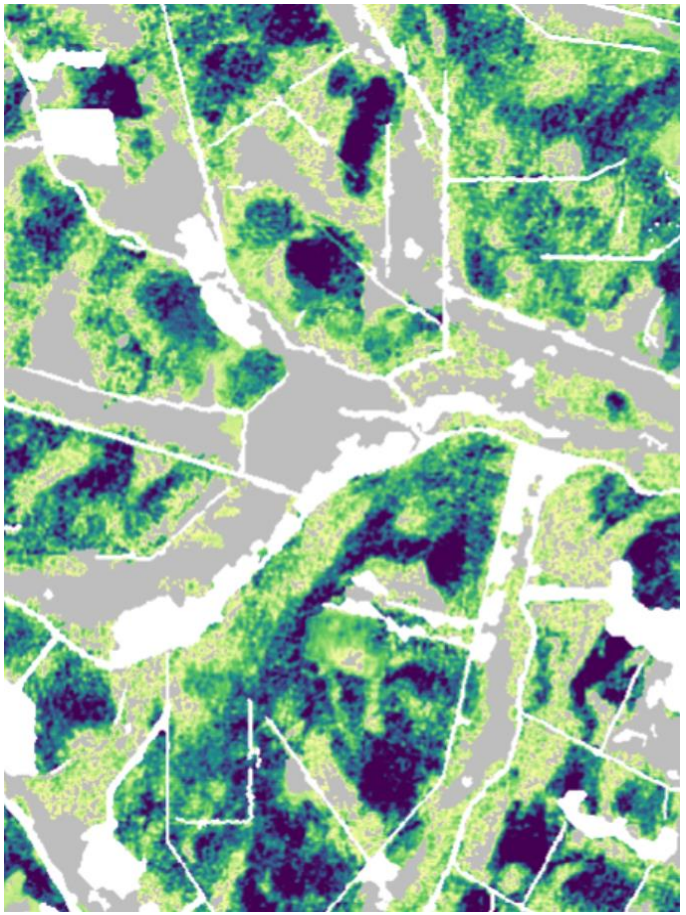


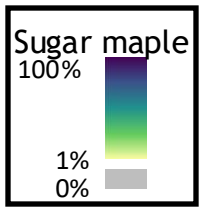
0 1 2 4
km

Former FIA workflow: Spatial smoothing used to improve ML model fit



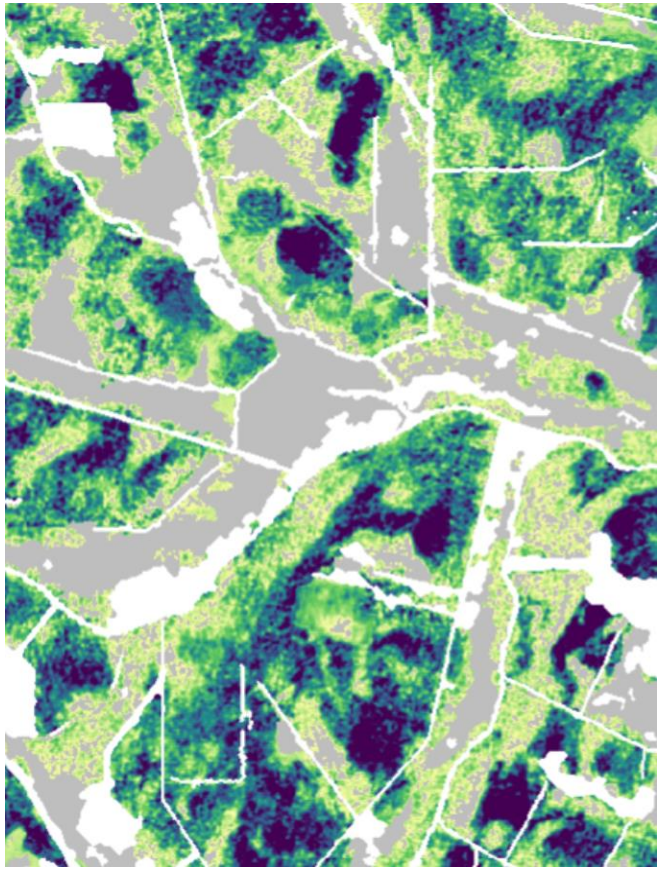
Current FIA workflow: Similar ML training process, but not smoothing ML predictions



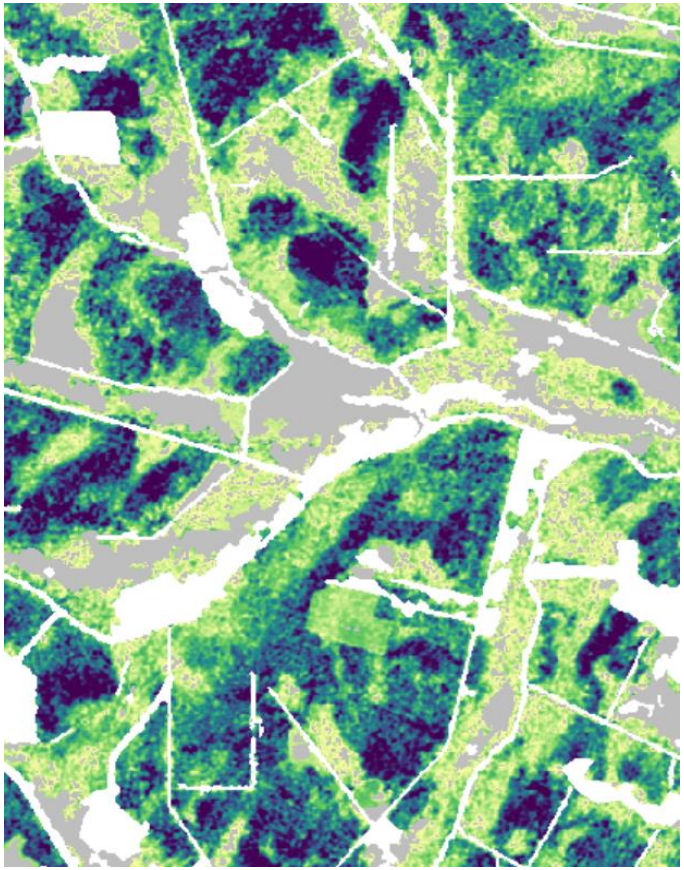


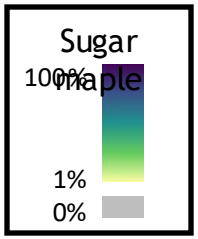
Compared species maps produced using FIA data to maps produced using fixed radius plot data provided by a Maine cooperator

FIA data:



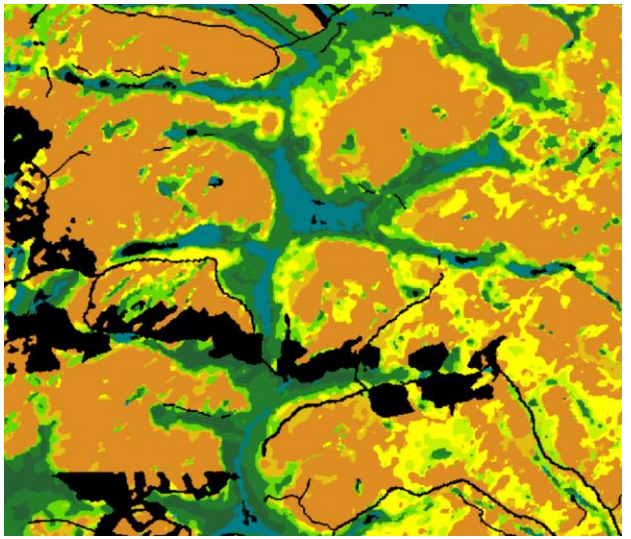
Company data:



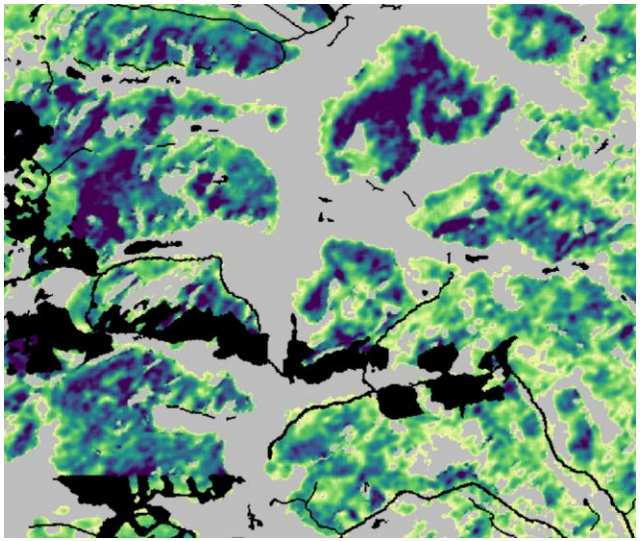


Multi-objective ML used for species (and forest type) mapping in Maine:

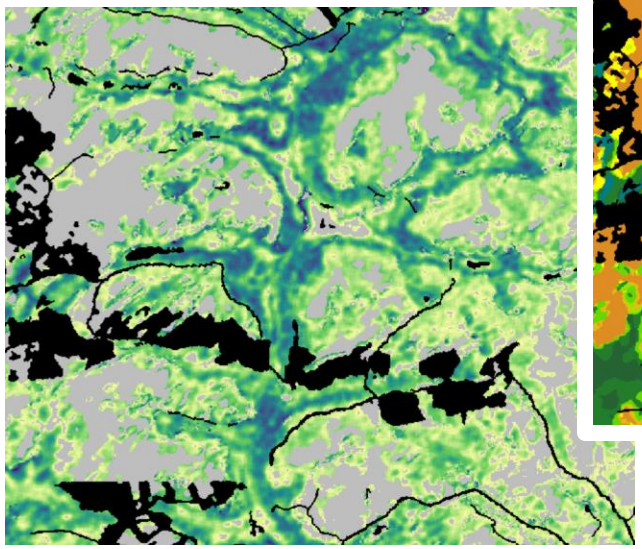
Justification



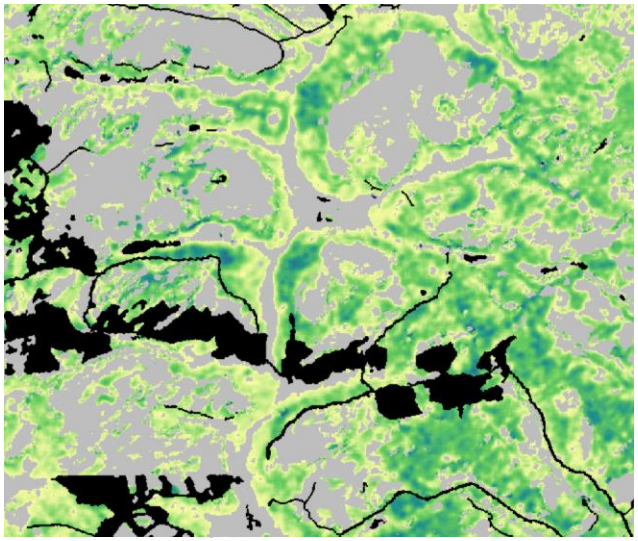
Sugar maple



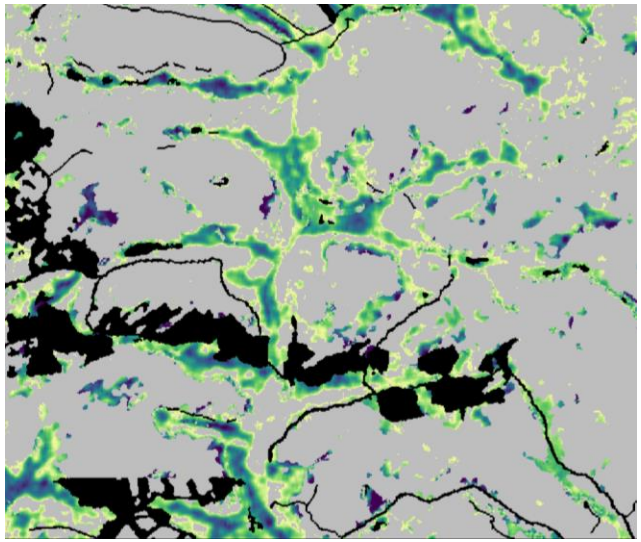
Balsam fir



Paper birch



Northern white cedar



- Oak
- Maple-Beech-Birch
- Red Maple
- Aspen-Birch
- Aspen-Birch Mixedwood
- Spruce-Pine Mixedwood
- Fir-Spruce Mixedwood
- Hemlock Mixedwood
- Hemlock
- Spruce-Pine
- Fir-Spruce
- Cedar-Black Spruce
- Hardwood Forested Wetland
- Mixedwood Forested Wetland
- Softwood Forested Wetland

Center for Advanced Forestry Systems

Methods

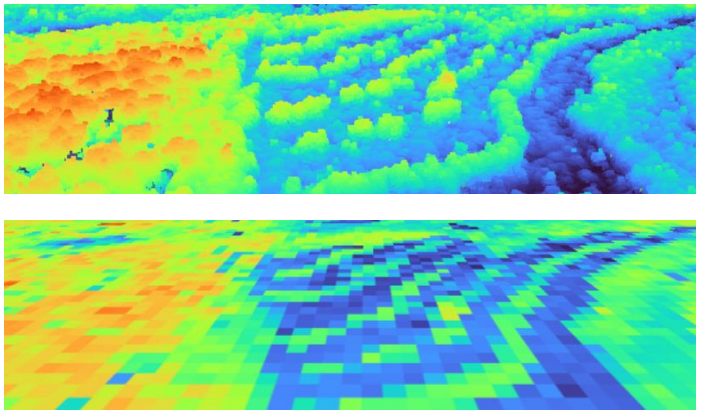
Testing alternative auxiliary data:

Currently use multi-temporal Sentinel-2 to map overstory species, and additional covariates haven't helped much at the pixel level

Can new auxiliary data improve stand-level species predictions?

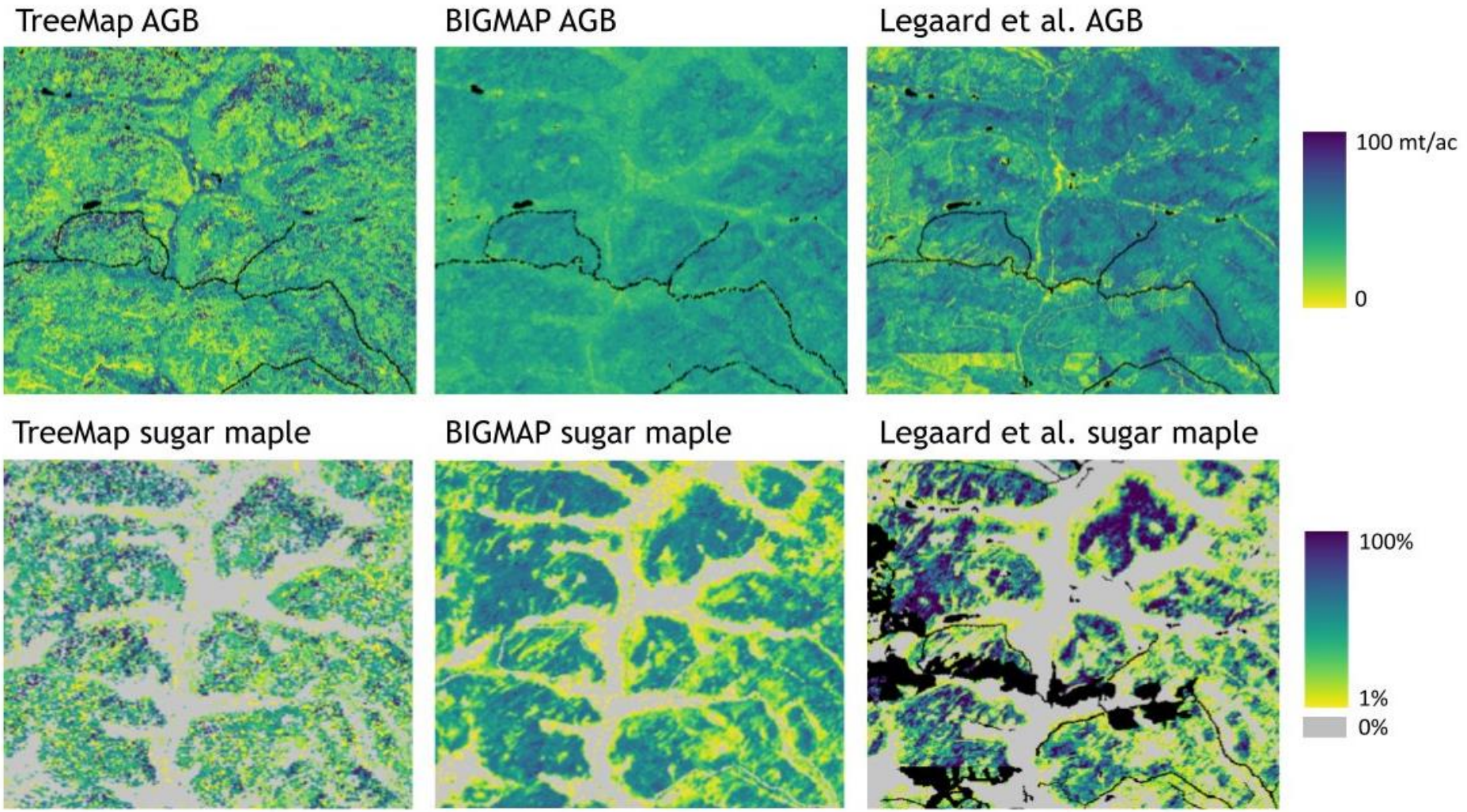
- Digital soil data
  - 7 variables available for the state of Maine at ~20 m resolution
- Gridded vegetation height metrics from 3D NAIP, computed using fast and scalable computational methods
- 3D vegetation change from multi-temporal 3D NAIP
  - 2018, 2021, 2023-24 collections in Maine

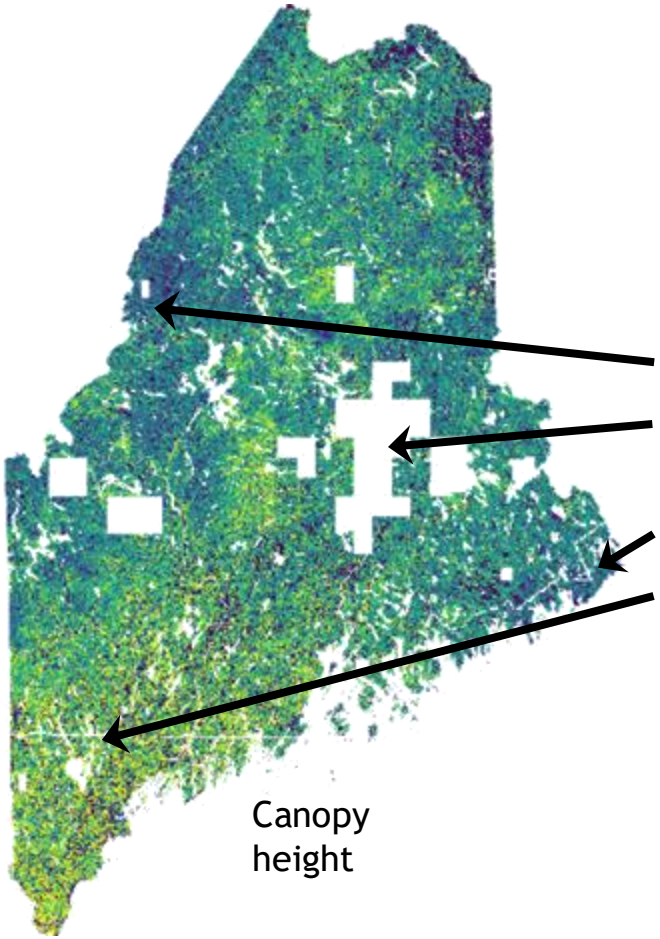
2021 NAIP DSM



95th height percentile, 10 m grid

Comparison to Other Products





Major Findings

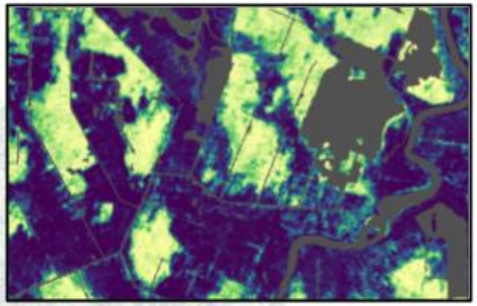
Remaining challenges in statewide NAIP processing pertain to data access and quality, not data processing

- Corrupted NAIP tiles
- Tribal data sovereignty restricts access to LiDAR-derived DEM tiles
- Gaps in LiDAR-derived elevation data
- Gaps between NAIP flight lines

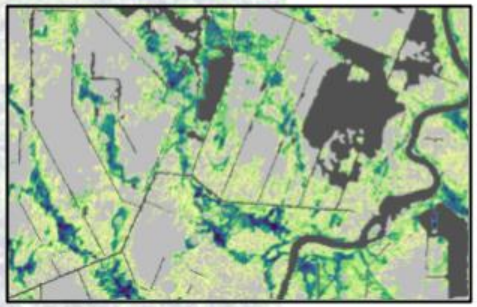
NAIP DSMs distributed only to those organizations that contributed to original buy-ups

Forest mapping progress

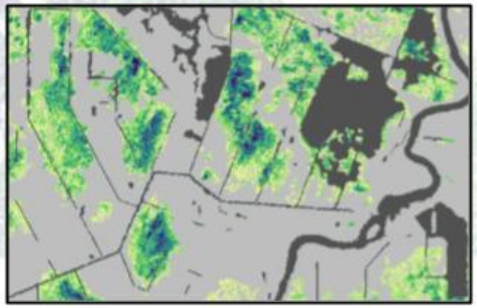
- Paper birch
- Aspen species
- Oak species
- Sugar maple
- White ash
- Yellow birch
- American beech
- Red maple
- Eastern hemlock
- Balsam fir
- Pine species
- Red spruce
- White spruce
- Black spruce
- Northern white cedar
- Softwood species



Total softwood

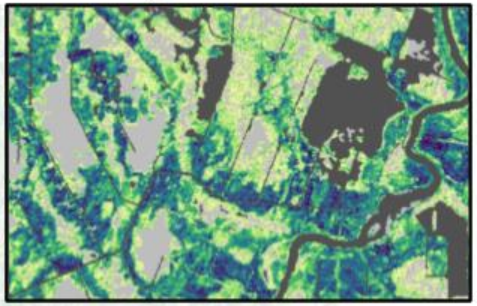


Cedar

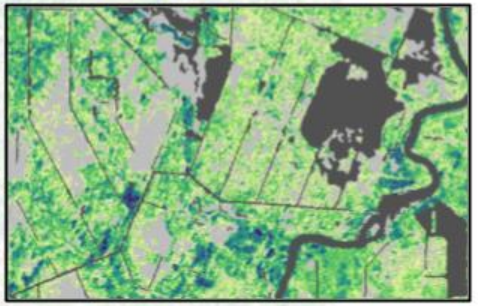


Sugar maple

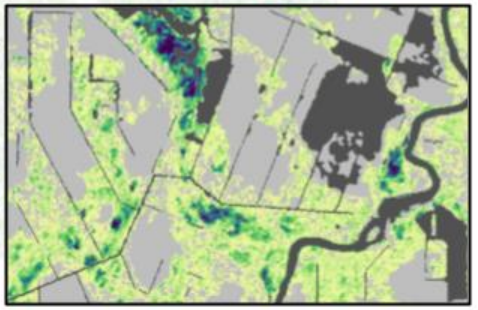
Balsam fir



Red spruce

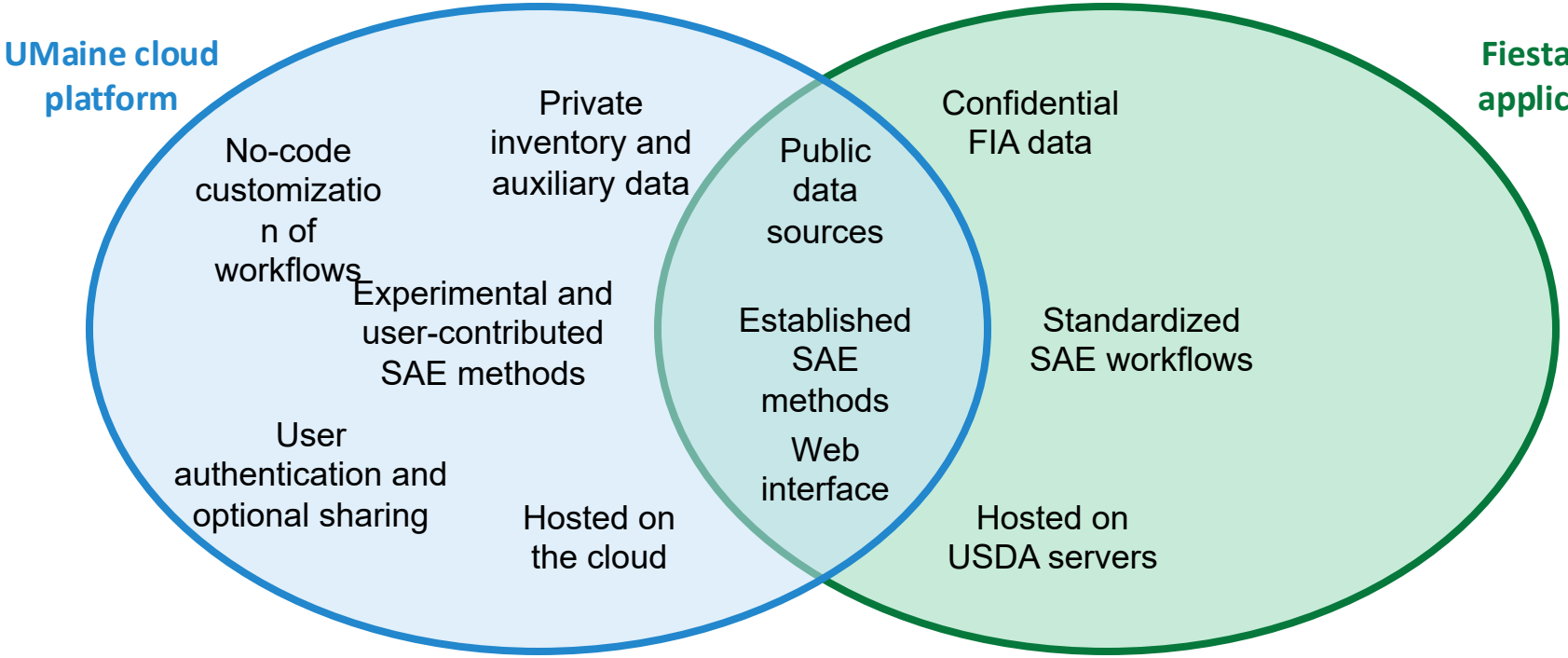


Black spruce



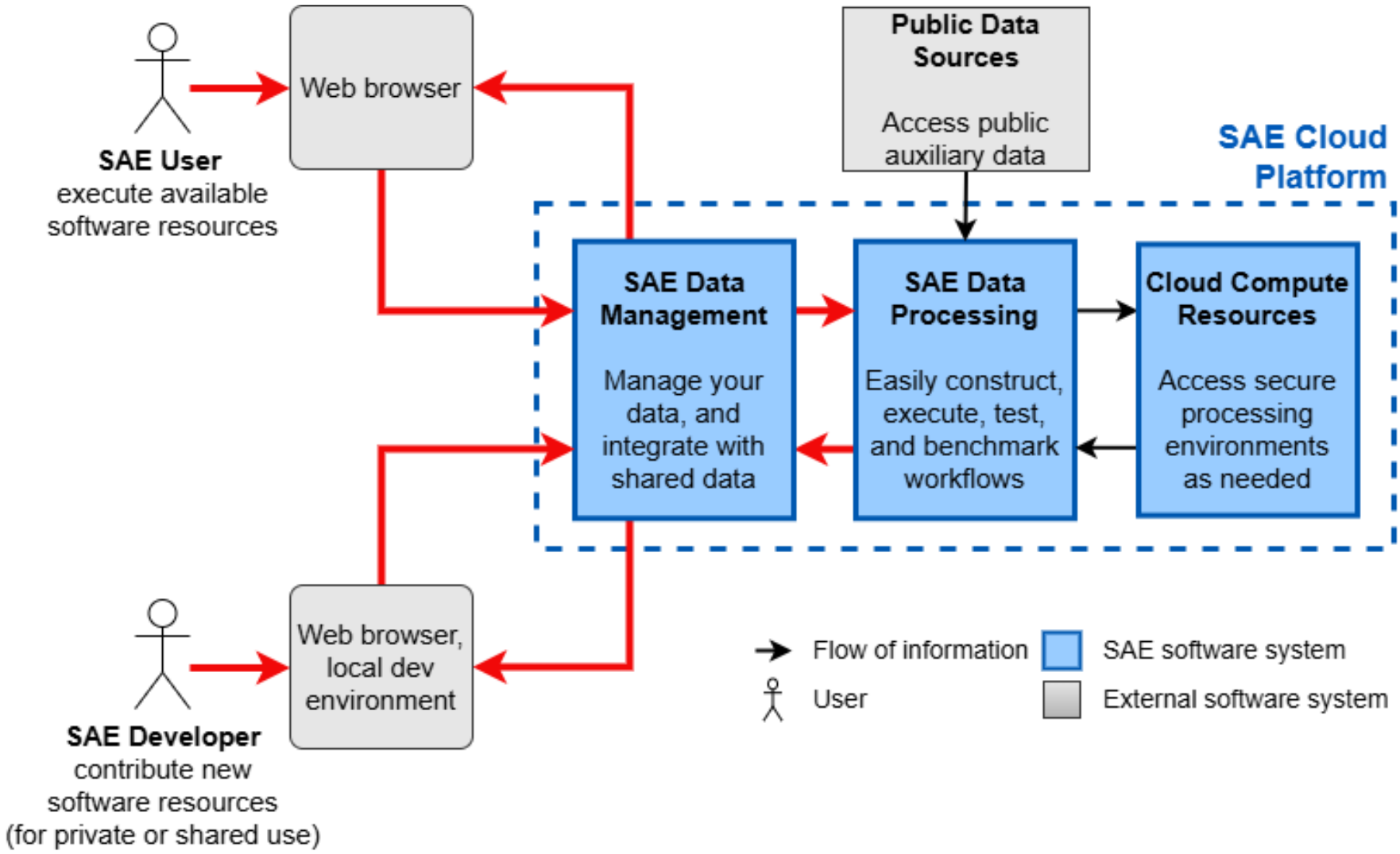
A cloud-based platform for accessible, secure, and efficient SAE

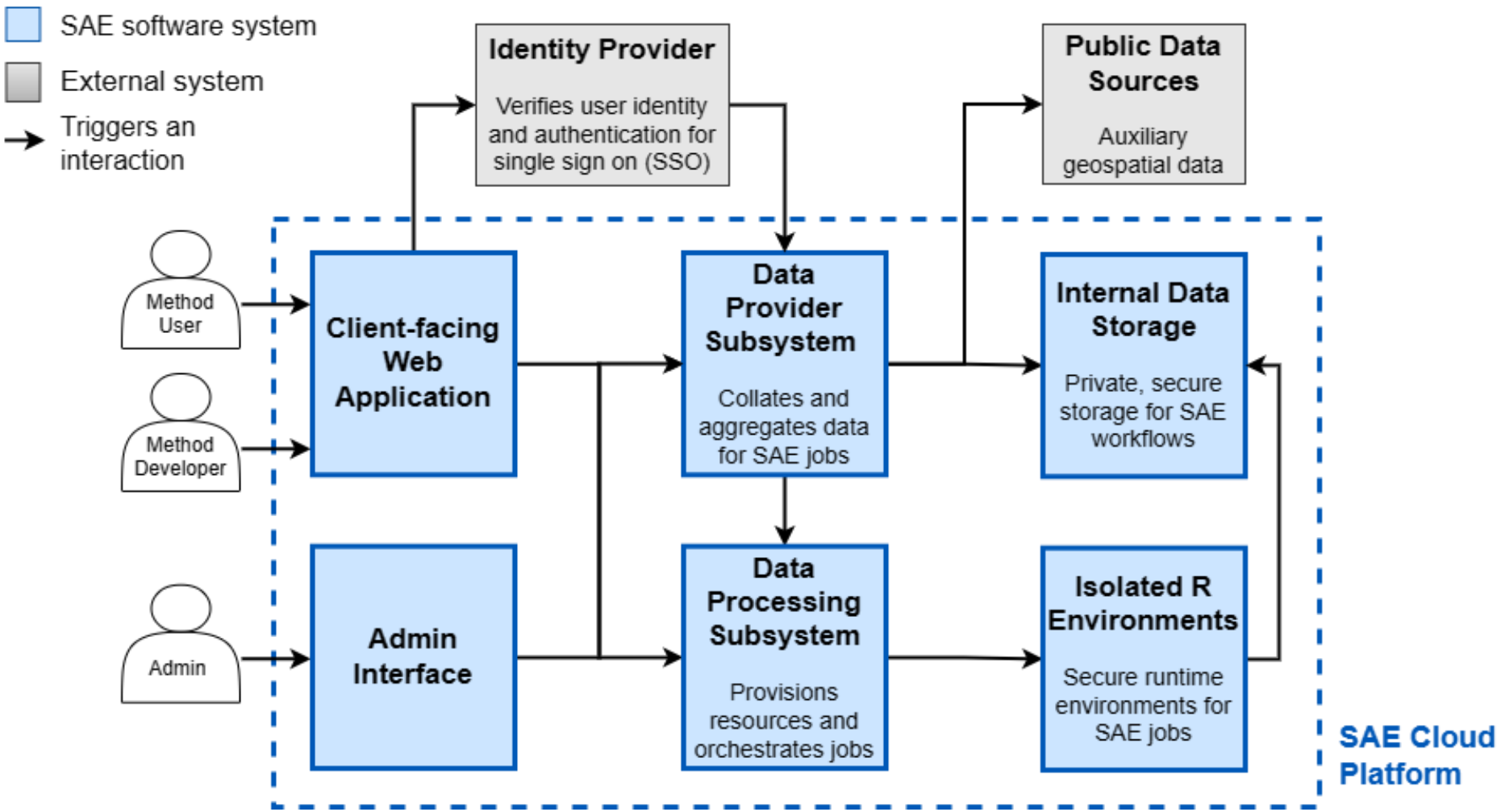
Project Goal: Support collaborative development, testing, and application of SAE methods by simplifying access to data, algorithms, and cloud computing resources.



Aaron Weiskittel, Kasey Legaard, Ken Bundy, Mike Premer, Jimm Domingo, Jereme Frank, Phil Radtke, Jim Westfall

Goal: Design and prototype a cloud-based data processing platform to expand collaborative SAE research and development within the PSAE.

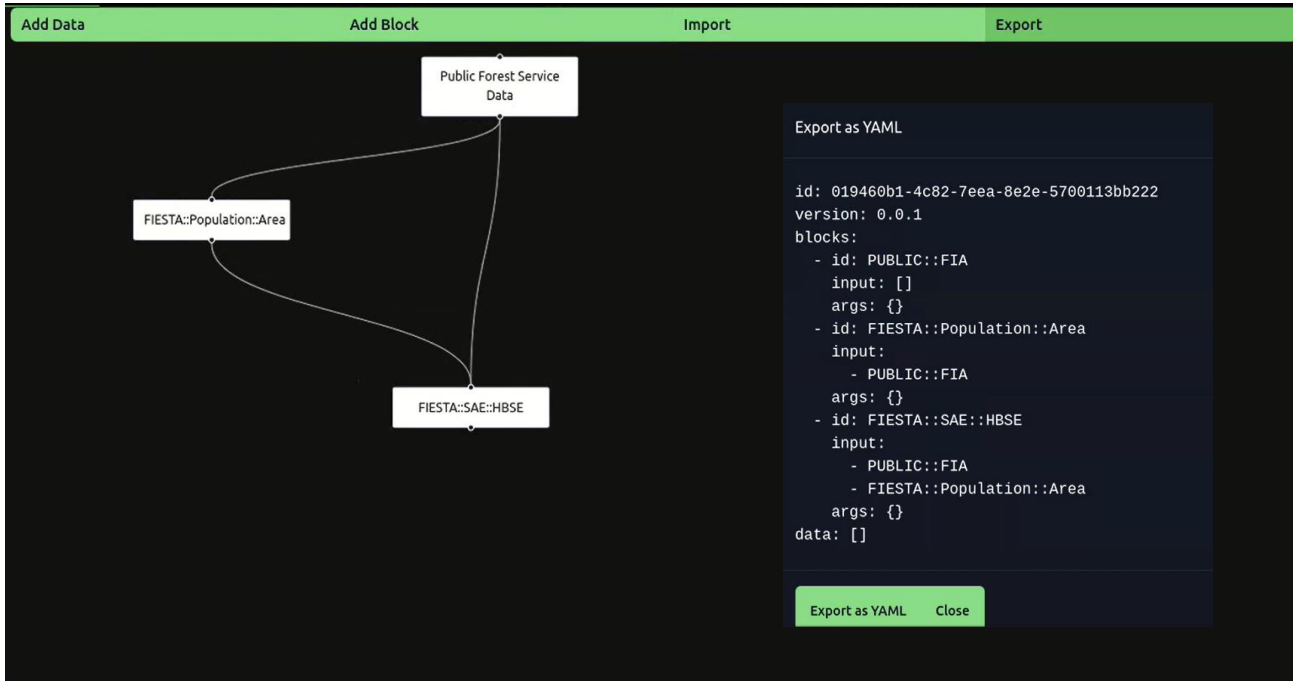



SAE software system
External system
Triggers an interaction
Identity Provider
Verifies user identity and authentication for single sign on (SSO)
Public Data Sources
Auxiliary geospatial data
Method User
Method Developer
Admin
Client-facing Web Application
Admin Interface
Data Provider Subsystem
Collates and aggregates data for SAE jobs
Data Processing Subsystem
Provisions resources and orchestrates jobs
Internal Data Storage
Private, secure storage for SAE workflows
Isolated R Environments
Secure runtime environments for SAE jobs
SAE Cloud Platform

A cloud-based platform for accessible, secure, and efficient SAE

Testing ideas for web-based, no-code interface for pipeline construction and management:

- Add data / code blocks and connect to construct a new pipeline
- Export / share constructed pipeline as YAML file
- Or import and customize an existing pipeline
- Eventual integration with data management system and cloud services



Aaron Weiskittel, Kasey Legaard, Ken Bundy, Mike Premer, Jimm Domingo, Jereme Frank, Phil Radtke, Jim Westfall

Summary

- Continued progress on large-scale mapping from remote sensing
  - Total AGB
  - Species
  - Additional metrics (volume)

- Significant refinement of ML algorithm speed and predictive capacity since project inception

- Development of a cloud-based online SAE interface for on-demand, custom jobs

- Extension to new regions a priority