Project Title: Cloud-based small area estimation based on fast, on-demand processing of large-area data sets and mid- to high-resolution geospatial auxiliary remote sensing

PIs and affiliations: Aaron Weiskittel, Professor of Forest Biometrics and Director, Center for Research on Sustainable Forests, University of Maine; Director, NSF Center for Advanced Forestry Systems

Kasey Legaard, Research Associate of Geospatial Analytics and Machine Learning, Center for Research on Sustainable Forests, University of Maine

Kenneth Bundy, Software and Machine Learning Engineer, Center for Research on Sustainable Forests, University of Maine

Michael Premer, Assistant Professor of Forest Management, University of Maine

Collaborators: Jimm Domingo, Software Engineer, Center for Research on Sustainable Forests, University of Maine; Senior Data Engineer, Institute for Social Research, University of Michigan

Jereme Frank, Forest Biometrician and Analyst, Maine Forest Service

Phil Radtke, Associate Professor of Forest Biometrics, Virginia Tech University

Jim Westfall, Research Forester, USFS Forest Inventory and Analysis, Northern Research Station

Progress for reporting period (e.g., from date of award to Jan. 31, 2024): The goal of this effort is to support collaborative development and application of SAE methods through an accessible, secure, and efficient cloud-based system that seamlessly connects users to data, algorithms, and computing resources (Figure 1). Design-considerations will address the needs of a range of stakeholders including scientists, analysts, and data end-users. Supporting objectives include: (1) Developing and hosting an cloud-based SAE system for collaborative R&D and application deployment by parties invested in the Partnership for Small Area Estimation (PSAE); (2) Online system deployment using non-sensitive data sets for software verification and testing, with varying access levels for hosting secure R&D connections alongside public-facing application deployments; and (3) Closed prototyping, testing, and evaluation to ensure the security of proprietary or restricted resources, with accountability to NCASI/FIA.

System design and implementation will satisfy multiple needs: (a) analysis toolsets and workflows to model inventory attributes from auxiliary data with support for very large data sets (i.e., “big data”); (b) Integration of 3D LiDAR and photogrammetric point cloud processing supported by fast, scalable cloud computing; (c) Support for secure hosting of proprietary inventory and auxiliary data; (d) Integration of existing, community-sourced, or user-supplied SAE

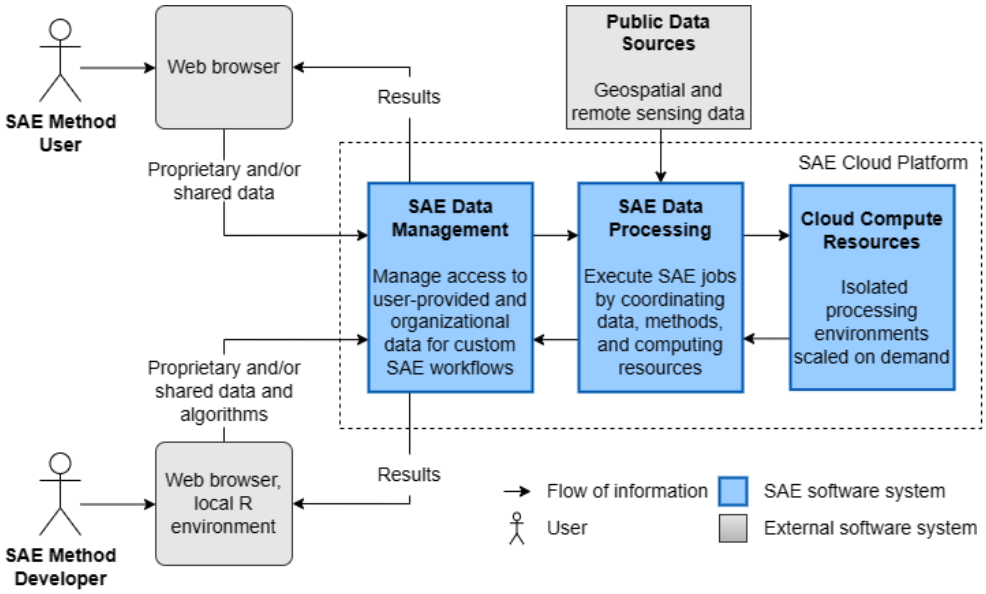

Figure 1: SAE cloud platform within the context of user types and external systems.

algorithm implementations as plugins; (e) Robust authentication and security measures to support secure, customizable SAE workflows for collaborating researchers and organizations; (f) Ability to securely integrate spatially precise, often proprietary or sensitive inventory data into development and testing of application deployments; and (g) User and administrative interfaces including tools for controlling access to data and services.

Since projection inception, our technical development team has met regularly to refine software architecture and define software development objectives. A hierarchical system has been adopted to guide system design, and progress has been made in defining the behavior and requirements of software subsystems (Figure 2). Substantial advances have been made in the development and implementation of a containerized R runtime environment for SAE algorithm execution. This effort includes the identification and management of R resources and dependencies required to deploy a broad set of SAE methods, including those implemented by FIA (e.g., within the FIESTA package). SAE jobs will be executed within secure, fully isolated, and short-lived instances of this R runtime environment (i.e., no network access at runtime). Access to the containerized project-maintained R environment will support future development of compatible SAE algorithms and workflows by collaborators and researchers outside of this project. Additionally, project collaborator Dr. Phil Radtke met with project partners from the Minnesota Department of Natural Resources about sharing available ground-based inventory and recent airborne LiDAR point cloud data. A suitable area of interest has been identified for data sharing for the purposes of system testing and evaluation.

Next period plans: In the coming months, the software developer team will prototype the data provider and/or processing subsystems (Figure 2) and identify appropriate external software services, data, and computing resources for linkage to the system. Regular meetings with project collaborators will be used to integrate early feedback on system design considerations.

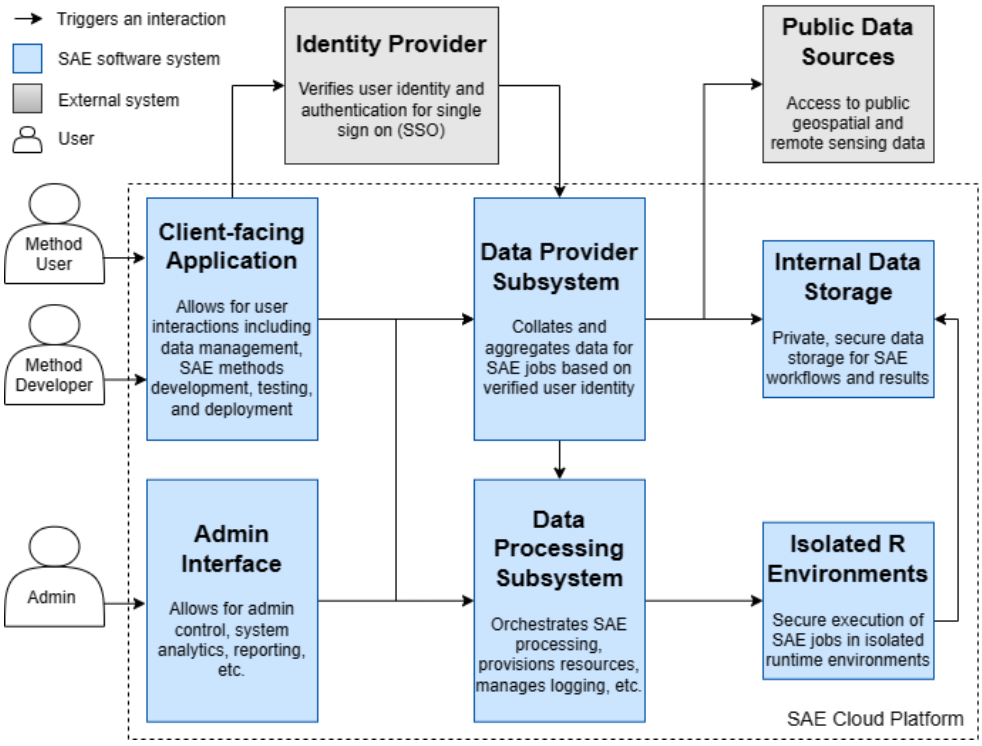

Figure 2: Primary SAE platform subsystems (blue) and key external systems (grey).

Problems or delays: The University of Maine's Office of Research Administration has yet to finalize the agreement with NCASI, so the team has no available funding to support the effort, which has delayed the processing of HR paperwork and limited project progress. The research office has been notified of these delays by the project team and NCASI.