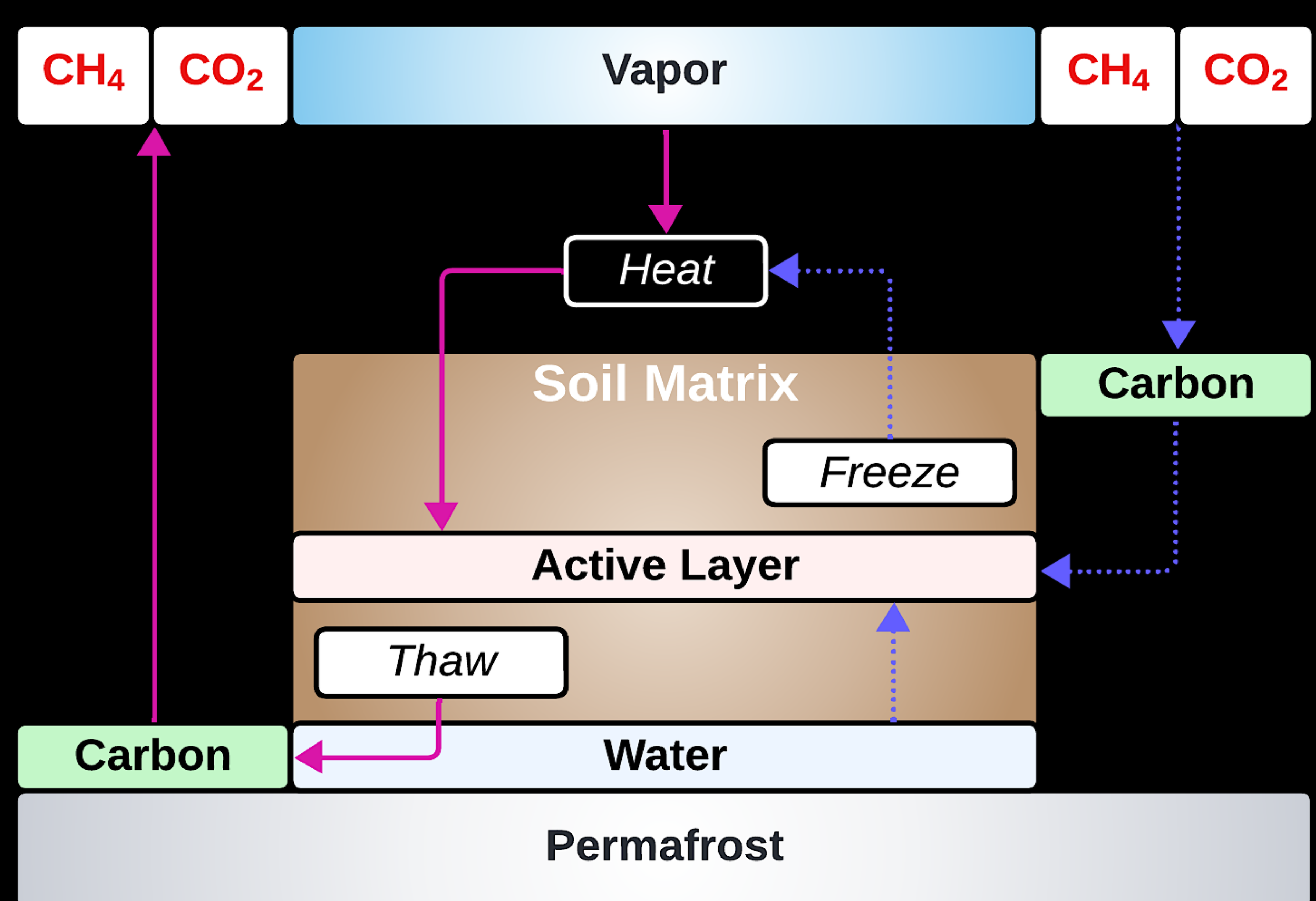
Decoding the Spatiotemporal Complexities of the Permafrost Carbon Feedback with Multimodal Ensemble Learning

Bradley A. Gay[1], Neal J. Pastick[2], Jennifer D. Watts[3], Amanda H. Armstrong[4], Kimberley R. Miner[1], and Charles E. Miller[1]

(1) NASA Jet Propulsion Laboratory, California Institute of Technology (2) United States Geological Survey, Earth Resources Observation Science Center (3) Woodwell Climate Research Center (4) NASA Goddard Space Flight Center, University of Maryland Earth System Science Interdisciplinary Center

BACKGROUND

Complex non-linear relationships exist between the permafrost thermal state, active layer thickness, and terrestrial carbon cycle dynamics in Arctic and boreal Alaska. Frozen soil and carbon-rich permafrost characterizes approximately 14 million square kilometers globally, with soil organic carbon stock estimated at 130±170 PgC (Hugelius et al., 2014). Thaw-induced carbon release is a climate change catalyst - and when coupled with anthropogenic-induced warming - can trigger, accelerate, and sustain a positive nonlinear carbon-climate feedback for hundreds of thousands of years (Schuur et al., 2015). The variability and uncertainty of thaw-induced carbon release and feedback mechanisms challenge efforts to quantify the magnitude, rate, timing, composition, and extent of the permafrost carbon feedback (PCF; Miner et al., 2021), further complicating this issue. The PCF is an emerging phenomenon resulting from rising global temperatures due to climate change, accelerating permafrost degradation, increasing exposure of ancient carbon to microbial decomposition, leading to further amplification of warming.

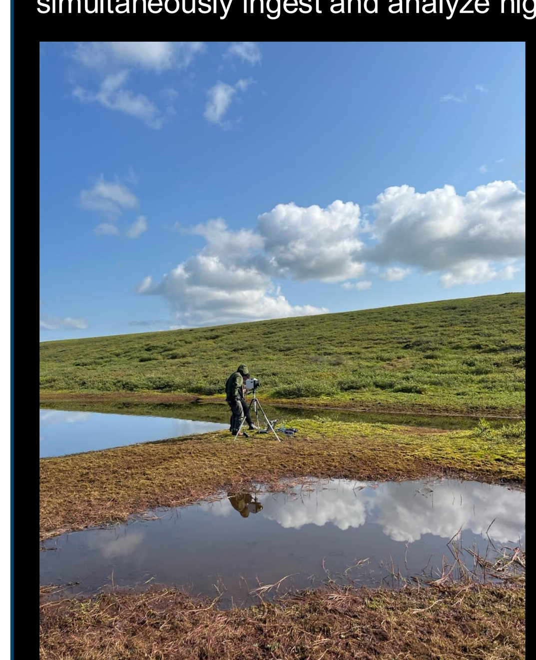


This research examines three challenges presented by the PCF: the big data problem, the remote sensing problem, and the modeling problem:

- We are operating in a space of diametrically opposing issues to store, process, and analyze information over space and time, i.e., *dearth* of field data or an *over-abundance* of data acquired from remote sensing and modeling resources).
- The ability to quantify or infer the magnitude, rate, and extent of the PCF and subsurface phenomena with high confidence across space and time is restricted with remote sensing platforms (Miner et al., 2021; Gay, et al., 2023; Esau et al., 2023). Due to spatiotemporal limitations, instrument constraints, and other challenges in the high latitudes (e.g., frequent cloud cover, short summer periods, low illumination angles).
- Subroutines and interactions governing earth system models vary widely, with many overlooking the dynamics and long-term impacts of the PCF when simulating high-latitude systems (Li et al., 2017; Randall et al., 2007). Fortunately, artificial intelligence (AI) optimizes complex earth system processes, captures nonlinear relationships, and improves model skill.

MOTIVATION

There is an *urgent* need to both understand *how* and *to what extent* thawing permafrost destabilizes the carbon balance in Alaska and to characterize the feedback involved. The objective of this research is to reconcile these challenges with AI, constrain these questions realistically in space and time, and apply these solutions at scale to simulate and disentangle the control factors and contributing drivers of the PCF signal. The study domain consists of Alaska (1.723M km^2), covering 26.92% of the Arctic Boreal Vulnerability Experiment (ABoVE) Domain (6.4M km^2) and 11.88% of the Arctic (14.5M km^2). We leverage a hybridized multimodal ensemble learning formulation (GeoCryoAI) with 13.1M site-level in situ measurements, 8.06B remote sensing observations, and 7.48B modeling outputs across the Alaskan tundra and boreal landscapes. After transformation, dimensionality reduction, trend removal, time-delayed supervision, and regression analyses, model training initializes 12.4M parameters to simultaneously ingest and analyze high-dimensional, time-variant multimodal data.

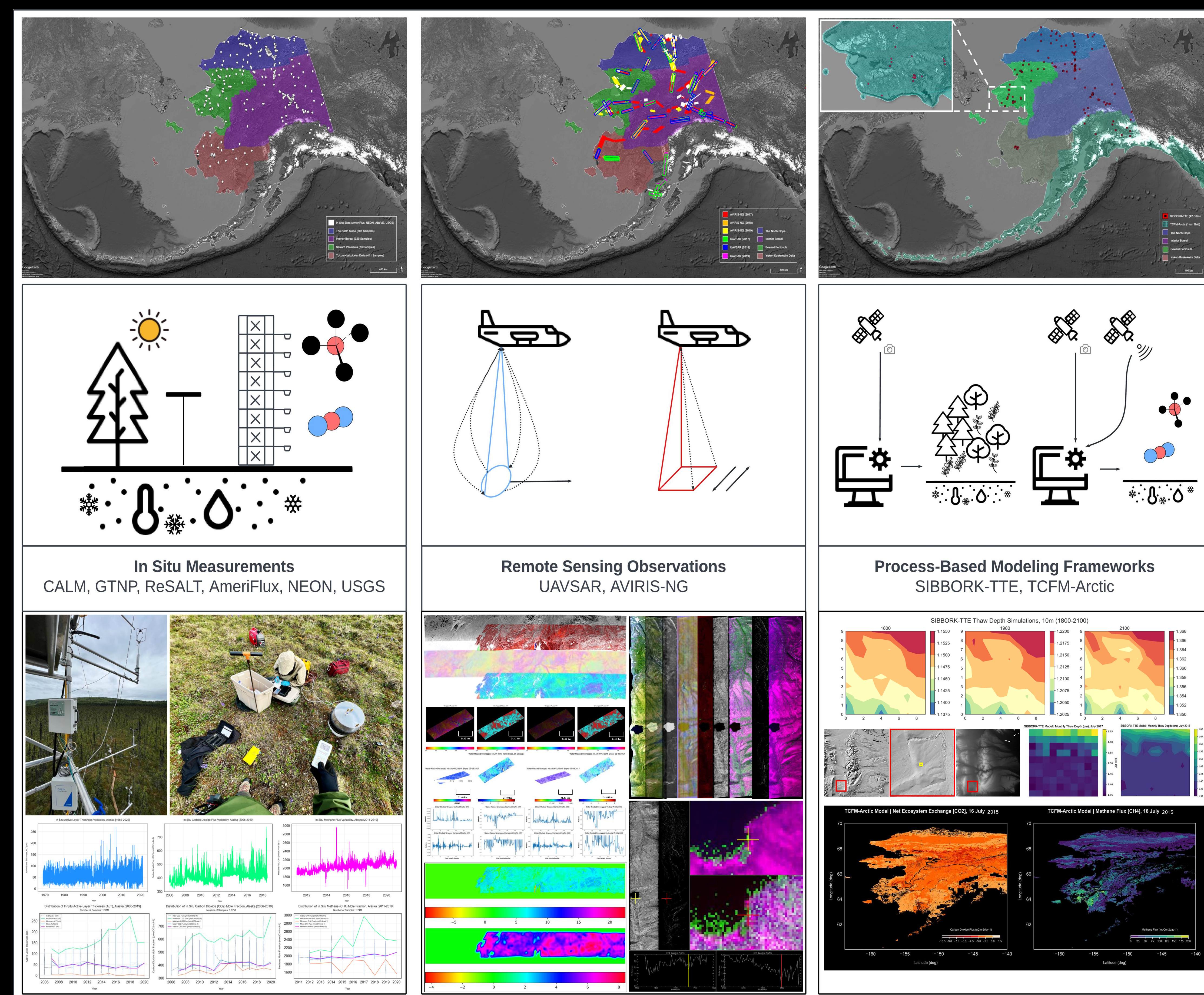


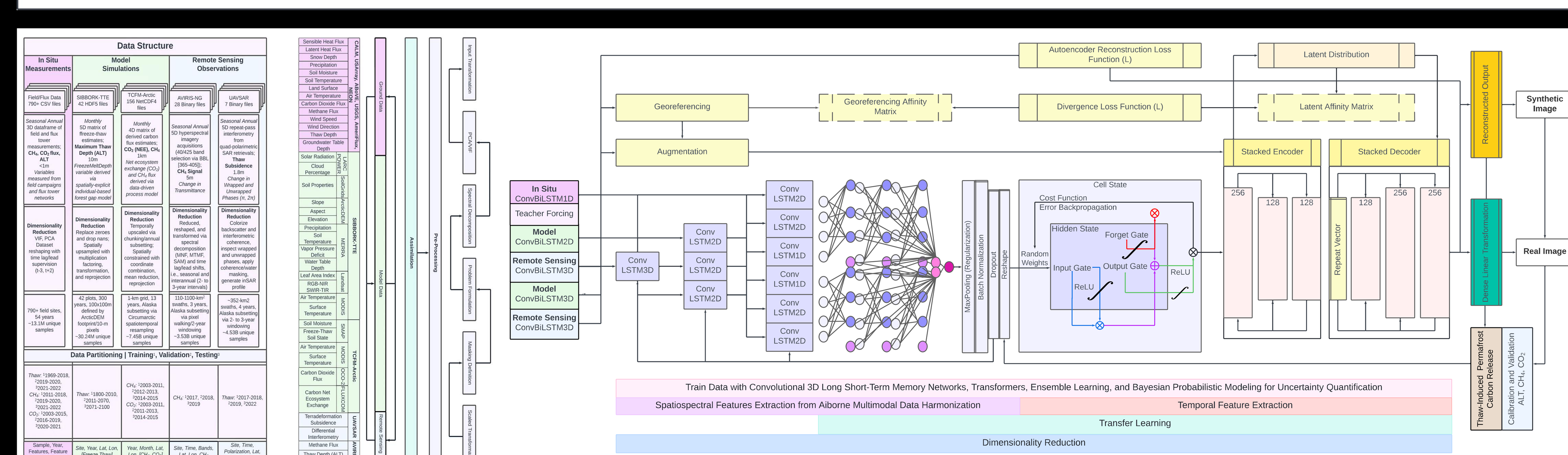
$$\mathbf{y}_{(t)} = \emptyset(\mathbf{W}_X^T\mathbf{x}_{(t)} + \mathbf{W}_y^T\mathbf{y}_{(t-1)} + \mathbf{b})$$

$$H_p = \underset{x\in X}{argmin}\, f(x)$$

The GeoCryoAI architecture is constructed with a hybridized process-constrained ensemble learning framework composed of a stacked convolutionally-layered long short-term memory-encoded recurrent neural network and optimized with a Bayesian hyperparameter optimization search algorithm. Feedback nonlinearities are emulated with ground-truth teacher forcing and module reconstruction functions (i.e., consolidated tabular time-series layer processing and sequential time-distributed convolving layers). We compared the performance of teacher forcing (i.e., in situ training), multimodal data assimilation among time-delayed naïve persistence (baseline), and GeoCryoAI simulations for ALT, CH_4, and CO_2 yielding five error metrics informed by training, validation, and testing over 10 epochs.

METHODS



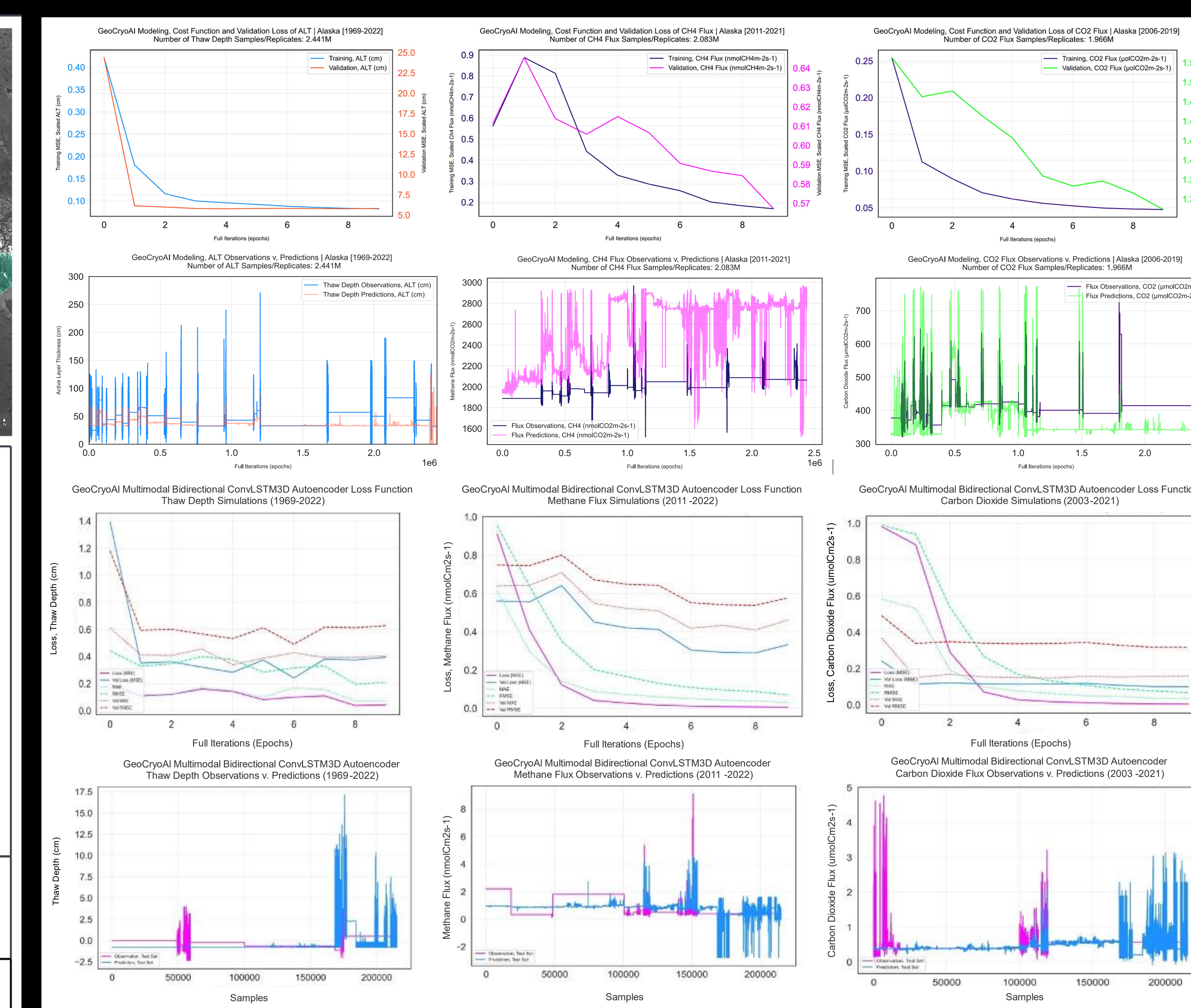


RESULTS | SIGNIFICANCE

GeoCryoAI simulations mirrored PCF dynamics across Alaska yielding promising results. With every epoch pass, validation loss is reduced. However, though validation and testing loss improved for CH_4, forecasting the CH_4 signal variability was a challenge during teacher forcing (i.e., failed to stabilize during periods of abrupt change of CH_4 and consistently overestimated CH_4 signal). By introducing more data into the framework, this discrepancy was ameliorated with limited changes to validation and testing loss. However, new changes emerged, i.e., failed to capture and predict early/initial pulses of thaw and CO_2 release. These results suggest the need for (1) more discriminant data partitioning (multitemporal coverage complexities, e.g., SAR), (2) further regularization (minimize weight aggregation), (3) more training (increase epochs), and/or (4) model is overfitting and may be resolved by simplifying network to aid generalizability. GeoCryoAI introduces ecological memory components and effectively captures and learns subtle spatiotemporal complexities as well as abrupt and persistent changes in high-latitude ecosystems by emulating permafrost degradation and carbon flux dynamics across Alaska with high precision and minimal loss (RMSE: 1.007cm, 0.694nmolCH_4m^{-2}s^{-1}, 0.213μmolCO_2m^{-2}s^{-1}).

In addition, this study underscores the significance of thaw-induced climate change exacerbated by the PCF and highlights the importance of resolving spatiotemporal variability of ALT as a sensitive harbinger of change. To our knowledge, this is the first time AI is applied to ameliorate dichotomy gaps while investigating the PCF combining ground, remote sensing, and modeling data.

DISCUSSION



	Active Layer Thickness (cm) 1800-2100	**Methane** (nmolCH_4km^{-2}month^{-1}) 1996-2022	**Carbon Dioxide** (μmolCO_2km^{-2}month^{-1}) 1996-2022
Naïve Persistence Model			
Test RMSE	1.997	0.884	1.906
GeoCryoAI | Teacher Forcing			
Test RMSE	1.327	0.715	0.697
Fractional Reduction	-33.55%	-19.12%	-63.43%
GeoCryoAI | Multimodality			
Test MAE	0.708	0.591	0.090
Test MSE	1.014	0.481	0.045
Test RMSE	1.007	0.694	0.213
Fractional Reduction	-24.11%	-2.94%	-69.44%

Our study establishes a new methodology for assimilating multimodal data across different observational platforms. We demonstrate the effectiveness of AI-driven ensemble learning frameworks in modeling complex permafrost-climate interactions. Our approach overcomes traditional model inefficiencies, bridges gaps, and resolves spatiotemporal disparities, providing a novel methodology for assimilating contemporaneous information on scales from individual sites to the Pan-Arctic. GeoCryoAI ingests ~15.8B measurements and observations to learn, simulate, and forecast the primary constituents of the PCF with prognostic and retrospective capabilities. This is important to corroborate *how* permafrost degradation and thaw subsidence *affect* carbon loss by *conserving* energy balance.

ONGOING WORK

Ongoing research elucidates on the PCF and delayed subsurface phenomena by (1) enrichment, i.e., expanding the flexibility and knowledge base of the model with current and future missions to minimize loss and improve performance (e.g., AVIRIS-3, UAVSAR, TROPOMI, PREFIRE, NISAR, CRISTAL; UAS DSMs; TIR), and (2) development, i.e., generating Circumarctic zero-curtain space-time maps to distribute to AK, First Nations/Native Corporations, and the USGS as a *JPL-led first-order* effort to engage leadership and identify cross-sector risks at local, state, and regional levels (e.g., critical infrastructure damage, cultural vulnerabilities). All source and pre-processed datasets are distributed at the ORNL DAAC while the source scripts and codebase are located on GitHub. This information is preliminary and is subject to revision. It is being provided to meet the need for timely best science. The information is provided on the condition that neither the U.S. Geological Survey nor the U.S. Government shall be held liable for any damages resulting from the authorized or unauthorized use of the information.

Gay, B.A., et al. (2025). Resolving Circumarctic Zero-Curtain Phenomena with AI-Integrated Earth Observations. Nature: Machine Intelligence. *Under Review*.
Gay, B.A., et al. (2025). Decoding the Spatiotemporal Complexities of the Permafrost Carbon Feedback with Multimodal Ensemble Learning. Journal of Geophysical Research: Machine Learning and Computation. 2, e2024JH000402.
Gay, B.A., et al. (2025). GeoCryoAI: Ensemble Learning and the Permafrost Carbon Feedback in Alaska, 1963-2022. ORNL DAAC, Oak Ridge, Tennessee, USA.
Gay, B.A., et al. (2023) Investigating Permafrost Carbon Dynamics in Alaska with Artificial Intelligence. Environmental Research Letters.
Gay, B.A., et al. (2023). Investigating High-Latitude Permafrost Carbon Dynamics with Artificial Intelligence and Earth System Data Assimilation. Virginia: George Mason University; 2023. 281 p: ProQuest LLC.
Gay, B.A., et al. (2022). Quantifying Feedback Sensitivities of Permafrost Degradation and Carbon Release with Earth Observation Data and Feedback Neural Networks. Earth and Space Open Archive.
Gay, B.A., et al. (2022). Understanding Active Layer Thickness Variability Under Changing Climatic Conditions Across the North American Taiga-Tundra Ecotone, Earth and Space Science Open Archive.
Gay, B.A., et al. (2021). Examination of Current and Future Permafrost Dynamics Across the North American Taiga-Tundra Ecotone. Earth and Space Science Open Archive.

bradley.a.gay@jpl.nasa.gov

National Aeronautics and Space Administration
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California
www.nasa.gov