Ronak Tali

1940 Long Rd, Unit F
Ames, IA 50010

253.867.7315
r.tali@live.com

Introduction

PhD Candidate with over a decade of combined industry expertise and multi-modal geospatial machine learning, signal processing and fluid mechanics research experience involving both custom geo-transformer models, neural operators as well as geospatial foundation models, such as, OlmoEarth. I have wide ranging interests at the intersection of environment, ecology, hydrology and agronomy. The tantalizing advances in AI have fundamentally changed how we solve scientific problems today. I want to use my entire range of skills to develop meaningful scientific products in the service of our environment and sustainability.

Education

Iowa State University, Ames, IA
Ph.D. Electrical Engineering *(*December 2026)

Papers

"Live Error Correction of National Water Model Discharge Estimates in Iowa ", *Working Paper*

NOAA's National Water Model (NWM) seems to be modeling large discharge errors across various reaches in Iowa. The main reason is lack of Iowa specific calibration. In this work, we make use of IWQIS high frequency discharge data at 79 locations in Iowa, The flow physics and ensemble Kalman Filter to vastly improve the discharge accuracy between major reaches but also through a bath-tub model estimate the impact of precipitation on transient discharge through a custom back tracking algorithm.

"Geospatial Multimodal Fusion for Field-Level Yield Prediction ", *Submitted to Smart Agricultural Technology*

This work involves a large spectrum of modeling work using custom models as well as foundational models, such as OlmoEarth. Tasks include yield map corrections, custom made spectral sampling to obtain a clean time series of satellite data. Imputation of satellite imagery time series both in time as well as in space. Yield prediction across different US geographies. Specifically, our yield map predictions using both a custom U-Net model and a custom transformer model consistently beat ($R^2 > 0.8$).

"Training-Free Reconstruction of Smooth Digital Elevation Models from Sparse Tractor Trajectories via Penalized Spline Surfaces", *In-Preparation*

Survey-grade digital elevation models (DEMs), such as those meeting USGS quality standards, are critical for agricultural autonomous systems that depend on precise terrain information for tasks including drainage planning, water-flow modeling, and autonomous vehicle navigation. However, such high-fidelity DEMs are often unavailable, infrequently updated, or derived from noisy sources that lack smooth curvature. Meanwhile, sensors mounted on agricultural machinery routinely collect high-resolution but spatially sparse elevation measurements along tractor trajectories. In this work, we leverage these sparse observations to reconstruct smooth DEM estimates across an entire field tile without any training data. We represent the terrain as a tensor-product spline surface whose control points are estimated directly from the track observations by minimizing a data-fidelity term together with a curvature (bending) and slope (tension) energy. Because the surface is linear in its control points and each basis function has compact support, this estimate is the solution of a single sparse, banded linear system that is solved exactly in under a second per tile, and the resulting surface is smooth and curvature-continuous by construction. On held-out field tiles, the method achieves a mean absolute

error of 0.077 m (about three inches) and a median absolute error of 0.029 m, roughly a 3× reduction relative to learned neural reconstructions Since the method requires no training, it is not exposed to distribution shift between training and deployment fields, enabling frequent, low-cost DEM updates from routine tractor passes

"A multi-electrode electrical clamp for rapid, non-destructive estimation of internal stalk cavity volume", *Submitted to Frontiers in Plant Science*

We evaluate an electrical sensing strategy for estimating internal cavity volume using a flexible multi-electrode clamp. Applying small voltages (50kHz) across electrode pairs and measuring the resulting pairwise resistances yields a characteristic resistance pattern that reflects how current is diverted around low-conductivity, air-filled voids. Because cavity volume is mechanistically linked to structural weakening, it provides a biologically meaningful target for screening stalk strength. Using a three-dimensional finite-element model of a stalk containing an ellipsoidal cavity, we simulate current flow for thousands of virtual geometries and compute pairwise electrode resistances in the low-frequency limit of impedance. Regression models trained on this library, with gradient boosting performing best, accurately predict cavity volume for previously unseen geometries ($R^2 \approx 0.99$). Once trained, inference is performed at the millisecond scale, supporting rapid screening workflows. We also benchmarked in "Sim to Real" setting with our model significantly ($p < 0.05$) predicting the presence of cavities in corn stalks sampled from field experiments.

"Flowbench: A large-scale benchmark for flow simulation over complex geometries", *Journal of Data-centric Machine Learning Research, Apr 2025.*

We introduce FlowBench, a dataset for neural simulators with over 10K samples, which is currently larger than any publicly available flow physics dataset. FlowBench contains flow simulation data across complex geometries (parametric vs. non-parametric), spanning a range of flow conditions (Reynolds number and Grashoff number), capturing a diverse array of flow phenomena (steady vs. transient; forced vs. free convection), and for both 2D and 3D. FlowBench contains over 10K data samples, with each sample the outcome of a fully resolved, direct numerical simulation We benchmark the performance of several methods, including Fourier Neural Operators (FNO), Convolutional Neural Operators (CNO), DeepONets, and recent foundational models.

"SEARS: a lightweight FAIR platform for multi-lab materials experiments and closed-loop optimization", *Royal Society of Chemistry – Digital Discovery, Sep 2025*

The Shared Experiment Aggregation and Retrieval System (SEARS) is an open-source, lightweight, cloud-native platform that captures, versions, and exposes materials-experiment data and metadata via FAIR, programmatic interfaces. Designed for distributed, multi-lab workflows, SEARS provides configurable, ontology-driven data-entry screens backed by a public definitions registry (terms, units, provenance, versioning); automatic measurement capture and immutable audit trails; storage of arbitrary file types with JSON sidecars; real-time visualization for tabular data; and a documented REST API and Python SDK for closed-loop analysis (e.g., adaptive design of experiments) and model building (e.g., QSPR).

"MPFBench: A Large-Scale Dataset for SciML of Multi-Phase-Flows: Droplet and Bubble Dynamics", *Journal of Data-centric Machine Learning Research, Apr 2025*.

Multiphase fluid dynamics, such as falling droplets and rising bubbles, are critical to many industrial applications. However, simulating these phenomena efficiently is challenging due to the complexity of instabilities, wave patterns, and bubble breakup. This paper investigates the potential of scientific machine learning (SciML) to model these dynamics using neural operators and foundation models. We apply sequence-to-sequence techniques on a comprehensive dataset generated from 11,000 simulations, comprising one million time snapshots, produced with a well-validated Lattice Boltzmann method (LBM) framework. The results demonstrate the ability of machine learning models to capture transient dynamics and intricate fluid interactions, paving the way for more accurate and computationally efficient SciML-based solvers for multiphase applications.

"Byzantine Resilient Federated REINFORCE (GM-FedREINFORCE)", *2023 International Conference for Machine Learning Applications, December 2023*.
Recent advancements in federated learning have inspired the emergence of federated reinforcement learning, wherein multiple agents learn a policy in a federated manner without sharing trajectories. However, one of the key challenges in federated learning lies in dealing with adversarial attacks, such as Byzantine attacks. This work addresses this challenge by introducing the GM-FedREINFORCE algorithm, which extends the REINFORCE method-an established policy gradient technique-and provides robust theoretical guarantees on its convergence even in the face of adversarial attacks. We do this by replacing the mean step of GD by Geometric median. Remarkably, our algorithm remains effective when less than half of the agents are under attack, ensuring robust learning.

Teaching Assistant Experience

Undergraduate: Theoretical Foundations of Computer Engineering.
Master of Science: Deep Learning – Theory and Applications, Statistical Learning and Data Mining – II.

Tools And Techniques

Languages & Frameworks: Python (Native & Scientific computing stack), PyTorch, SQL, Next.js, React, R.
Data & Geospatial Analytics: QGIS (including Python batch processing via rasterio, rioxarray), Google Earth Engine, Power BI.
Methodologies: Scientific ML (Neural Operators), Computer Vision, Bayesian Statistics (PyMC3/Bambi), Gradient Boosting (GBMs), Cluster Analysis, Experimental Design, FAIR Data with MCP.

Professional Experience

REI, Kent, WA, 2014 – 2018
Data Scientist
At REI, I built high-impact analytics solutions across forecasting, marketing, and customer analytics. I developed demand forecasting models using the Prophet framework (GAM and STAN) that achieved up to 99% accuracy, a churn prediction model using GBMs with an AUC near 80%, and a fully in-house, behavior-based marketing attribution model that replaced an industry-leading vendor and provided a holistic, customer-level view of channel performance. I also designed a multi-level clustering framework that became the organization-wide customer segmentation standard and created innovative experimental designs that enabled zip-code-level testing despite platform limitations, reducing cost and risk

Fedex Corporation, Memphis, TN, 2013 - 2014
Senior Revenue Science Analyst
At FedEx, I developed customer profile vectors within verticals to identify future sales opportunities and measure market penetration and built a customer lifecycle and churn model for international customers that enabled early intervention and prevented losses of nearly 48,000 packages per day. I also created multivariate statistical models to optimize retail store location decisions, helping avoid estimated lifetime losses of over USD 1.2 million per store due to poor placement.

Systems Engineer, Redmond, WA, 2007 – 2011
Systems Engineer
At TCS, I developed a server monitoring tool using Lean Six Sigma methodologies that generated over USD 100,000 in cost savings for a large-scale, ongoing project valued at more than USD 1 million. I also implemented extensive software test automation, reducing manual testing effort by over 80% on average and significantly improving delivery efficiency and quality.

References

Prof. Baskar Ganapathysubramanian (Advisor)
Joseph C. and Elizabeth A. Anderlik Professor in Engineering and Professor Mechanical Engineering
Director, AI Institute for Resilient Agriculture
ISU Plant Science Institute Fellow
Department of Electrical and Computer Engineering (by courtesy)
Iowa State University
2029 Union Dr, Ames, IA 50011
Phone (515) 294 – 7442
Email – baskarg@iastate.edu

Prof. Soumik Sarkar
Walter W. Wilson Faculty Fellow in Engineering and Professor Mechanical Engineering
Director, Translational AI Center (TrAC)
ISU Plant Science Institute Fellow
Department of Computer Science (by courtesy)
Iowa State University
2029 Union Dr, Ames, IA 50011
Phone (515) 294 – 5212
Email – soumiks@iastate.edu

Prof. Adarsh Krishnamurthy
Professor of Mechanical Engineering
Department of Electrical and Computer Engineering (by courtesy)
Iowa State University
2029 Union Dr, Ames, IA 50011
Phone (515) 294 – 5568
Email – adarsh@iastate.edu