Research Statement

Md Abdullah Al Mazid
Florida International University | PhD in Computer Science, expected Spring 2027
mazid.rafee@gmail.com | mazid-rafee.github.io

Research Vision

My research develops machine-learning methods for Earth observation and environmental sensing that are not only accurate, but also physically meaningful, multimodally grounded, and useful under the imperfect conditions of real scientific data. Satellite observations are intrinsically heterogeneous: optical imagery may be obscured by clouds, radar captures structure through a different sensing mechanism, atmospheric correction depends on expensive radiative-transfer calculations, and environmental measurements are often irregular in space and time. I view these limitations not as isolated preprocessing problems, but as opportunities to design learning systems that explicitly account for the sensing process, complementary modalities, and uncertainty in the available evidence.

My doctoral research at Florida International University's High Performance Database Research Center, advised by Prof. Naphtali Rishe, has therefore centered on three connected themes: (1) multimodal learning for Earth observation, (2) physics-guided and computationally efficient scientific machine learning, and (3) spatiotemporal and interpretable models for environmental data. Across these themes, my goal is to bridge modern deep learning with the structure already present in scientific problems. This perspective has led to work on multispectral cloud segmentation, SAR-guided cloud removal, atmospheric-correction emulation, water-quality forecasting, satellite-based wildfire progression prediction, and interpretable visual modeling.

Multimodal Learning for Earth Observation

A central question in my research is how complementary sensing modalities should interact when no single observation is fully reliable. This question first motivated my work on **MSCloudCAM**, a multispectral cloud-segmentation model designed for observations from different satellite platforms. The model integrates a Vision Transformer backbone with multi-scale contextual representations and cross-attention, allowing local spectral-spatial evidence to interact with broader scene context. I evaluated the approach on Sentinel-2 CloudSEN12 and Landsat-8 L8Biome imagery for multi-class cloud segmentation. Beyond the specific task, this project shaped my interest in architectures that can combine heterogeneous Earth-observation signals without assuming that every source is equally informative in every scene.

My current dissertation research, **MSCloudr**, extends this idea from segmentation to reconstruction. The goal is to recover cloud-obscured 13-band Sentinel-2 observations using paired Sentinel-1 synthetic-aperture radar (SAR) as complementary structural guidance. SAR is attractive because it is not blocked by clouds, but it differs fundamentally from optical imagery in sensing physics, spatial statistics, and semantic content. A useful cloud-removal model therefore should not simply concatenate SAR and optical features; it should determine when, where, and how strongly the radar observation should influence the reconstruction.

I am investigating this problem through a **conditional diffusion-bridge formulation**. Rather than treating the restoration trajectory as a generic denoising process, I study bridge dynamics that connect the observed cloudy image to the desired clean optical state under auxiliary SAR guidance. Current directions include mean-reverting bridge dynamics, risk-aware trajectory design, reliability-aware conditioning, cross-modal alignment, bridge-consistency objectives, and alternative bridge geometries. The broader research question is theoretical as much as empirical: *how should a generative restoration path be shaped when the conditioning modality is informative but imperfectly aligned with the target modality?*

This work suggests a general framework for multimodal scientific reconstruction. Auxiliary observations should be incorporated according to their reliability and physical relationship to the target, rather than

treated as universally trustworthy context. I plan to extend this principle to missing-data reconstruction, multi-sensor harmonization, and event-driven Earth-observation problems in which cloud cover, revisit intervals, or sensor failures produce incomplete observations.

Physics-Guided Scientific Machine Learning

A second theme of my work is replacing expensive scientific computation with learned surrogates while preserving the structure required by downstream scientific workflows. In **pKANrtm**, published in *Remote Sensing*, I developed a physics-aware multi-fidelity surrogate for atmospheric-correction coefficients. Atmospheric correction is essential for converting top-of-atmosphere satellite measurements into surface-reflectance estimates, yet high-fidelity radiative-transfer simulation can be computationally expensive when repeated over large parameter spaces.

The key idea in pKANrtm is to use the lower-cost 6S radiative-transfer model as a low-fidelity source of physical guidance and learn a residual correction toward higher-fidelity libRadtran outputs using a Kolmogorov-Arnold Network. The system predicts path reflectance, total transmittance, and spherical albedo while retaining compatibility with standard coefficient-based surface-reflectance equations. This design differs from a purely black-box emulator: it preserves interpretable atmospheric quantities and explicitly exploits the relationship between low- and high-fidelity physical models. The work has also been submitted as FIU Invention Disclosure D2026-0057 and is currently under review.

This project established a direction I intend to pursue further: **multi-fidelity learning for scientific simulation**. Many environmental models offer a hierarchy of approximations, from fast but simplified solvers to computationally demanding high-fidelity systems. Learning should exploit that hierarchy rather than ignore it. I am interested in methods that combine physical priors, simulation fidelity, uncertainty estimation, and neural operators or structured networks to create surrogates that are both computationally efficient and scientifically traceable.

Spatiotemporal Environmental Modeling and Interpretability

Environmental observations are also challenging because they are irregular, incomplete, and distributed over space and time. In my work on **water-quality forecasting**, presented at IEEE SoutheastCon 2025, I investigated the integration of ordinary differential equations with bidirectional LSTMs to model irregular environmental time series. This work reinforced the importance of modeling observation timing and process dynamics directly instead of forcing environmental measurements into uniformly sampled assumptions.

My ongoing research on **wildfire detection and progression prediction** focuses on developing deep learning frameworks for short-term wildfire forecasting using multitemporal satellite imagery and environmental data. I am investigating spatiotemporal feature extraction, recency-aware temporal fusion, and the influence of historical observation windows to improve burned-area prediction and wildfire spread modeling. The broader question is how forecasting models should use incomplete, unevenly informative observations when environmental dynamics change rapidly.

I have also worked on **interpretable computer vision and saliency modeling**, including object-grounded text priors for saliency-ranking detection and prototype-based approaches for explaining video saliency. These projects examine how semantic or prototype-level evidence can be connected to a model's spatial predictions. Although the application domain differs from remote sensing, the underlying question is closely related to my broader agenda: can a model expose the evidence that drives its prediction rather than provide only an output map?

My research is additionally informed by applied environmental collaborations. Through **BBROIPS**, I contribute to visualization, simulation, and AI-driven analysis of environmental data in South Florida in work involving FIU, Biscayne Bay National Park, and the Florida Department of Environmental Protection. This experience keeps my methodological research connected to the realities of heterogeneous

environmental data, operational constraints, and interdisciplinary users.

Future Research Agenda

My future research program will build on these projects through three mutually reinforcing directions.

1. Reliability-aware multimodal Earth observation. I plan to develop models that explicitly estimate and use modality reliability in optical, SAR, thermal, meteorological, and in-situ data. Rather than assuming that multimodal fusion is always beneficial, these systems will learn when a modality should dominate, complement, or be ignored. I am particularly interested in spatially varying reliability fields, sensor-specific representation learning, and conditional mechanisms that can attenuate a modality when it becomes misregistered, noisy, or semantically inconsistent with the target. Generative reconstruction, sensor harmonization, and multi-temporal forecasting are natural applications. A key part of this agenda will be evaluating models under controlled modality degradation rather than only on clean benchmark pairs, so that claims about multimodal robustness are tied to interpretable failure modes. Wildfire progression forecasting is a natural testbed for this approach: I plan to investigate adaptive temporal selection, the integration of weather and terrain information, and uncertainty-aware forecasts that distinguish reliable fire-spread predictions from cases with insufficient observational support.

2. Physics-guided and multi-fidelity learning for environmental simulation. Building on pKANrtm, I will investigate efficient learned surrogates for radiative-transfer and other environmental models. I am especially interested in architectures that preserve physically interpretable variables, quantify uncertainty, and leverage multiple simulation fidelities. One direction is to learn corrections between a hierarchy of solvers while retaining conservation laws, coefficient structure, or other physically meaningful intermediate variables. Another is to couple surrogate models with inverse problems, enabling rapid retrieval of environmental parameters while propagating uncertainty from the learned forward model. This direction can support large-scale parameter exploration, rapid inversion, and operational remote-sensing pipelines without treating a neural network as an unconstrained replacement for the underlying science.

3. Interpretable and trustworthy scientific AI. Scientific users need to know not only what a model predicts, but also which observations, spatial regions, modalities, or physical assumptions support that prediction. I plan to connect prototype-based explanation, uncertainty estimation, and model intervention with multimodal Earth-observation systems. My current work on prototype-based video saliency motivates an intervention-oriented view of interpretability: if a prototype or modality is claimed to matter, deliberately suppressing or altering it should produce a measurable and semantically consistent change in the output. I want to extend this idea to Earth observation by testing whether explanations remain stable across seasons, sensors, geography, and distribution shifts. The aim is to move from post-hoc visualization toward explanations that are part of the model's computational structure and can be evaluated quantitatively.

Collaborative and Translational Research

My research style is inherently interdisciplinary. Work in Earth observation requires interaction among machine-learning researchers, remote-sensing specialists, environmental scientists, and the engineers who manage large scientific data systems. Through FIU's NSF-supported multimodal environmental research infrastructure and BBROIPS, I have seen how methodological questions change when models must operate on heterogeneous observations collected for real scientific objectives. These collaborations motivate me to design methods whose inputs, outputs, uncertainty, and computational requirements can be understood outside a machine-learning benchmark.

My prior software-engineering and competitive-programming experience also shapes how I approach computational research. Before beginning my PhD, I worked on large production software and database systems, which taught me to value reproducibility, maintainable implementations, profiling, and clear

interfaces. My undergraduate experience in competitive programming further strengthened my ability to decompose unfamiliar problems, reason about algorithmic complexity, identify suitable data structures and algorithms, and translate mathematical ideas into efficient implementations. These habits continue to influence how I design experiments, diagnose model behavior, optimize computational pipelines, and approach new research problems for which the solution is not immediately apparent.

This background also informs how I think about research training and mentorship. Competitive programming taught me that difficult computational problems become more approachable through repeated practice, discussion of alternative solution strategies, and careful analysis of both correctness and efficiency. In mentoring students, I aim to encourage the same habits of systematic problem solving, experimentation, and computational rigor, whether they are working on algorithmic problems, machine-learning systems, or scientific applications.

In future collaborations, I intend to build research software that can move from experiments to reusable scientific tools, including well-documented data pipelines, reproducible evaluations, and modular implementations that make it easier for domain collaborators and students to test models on new observations.

These directions are unified by a simple principle: environmental AI should respect how scientific observations are produced. My long-term goal is to develop machine-learning systems that combine the representational power of modern AI with physical structure, multi-sensor evidence, and interpretable reasoning. Such systems can make Earth-observation analysis more robust and computationally accessible while preserving the scientific meaning needed for environmental research and decision support.