# Enhanced Climate Science Research Proposals

*Five Improved High-Impact Projects*

Addressing Data Challenges and Operational Feasibility in African Climate Science

**Moses**
Environmental Scientist & Sustainable AI Researcher
Digital Society School, Amsterdam University of Applied Sciences

*November 2025*

# Executive Summary

This document presents five enhanced climate science research proposals that address critical challenges in African climate prediction and adaptation. Each proposal has been redesigned to explicitly tackle data quality issues, validation challenges, and operational constraints that are often overlooked in academic research but crucial for real-world implementation.

Key improvements across all proposals include:

- **Robust validation frameworks** with baseline comparisons and quantitative metrics
- **Tiered data strategies** that handle varying data quality and availability
- **Stakeholder co-development** ensuring outputs meet actual decision-making needs
- **Realistic timelines** with phased implementation and capacity building
- **Uncertainty quantification** making predictions actionable for risk management

# Proposal 1: Adaptive Multi-Scale Drought Early Warning System

*This research develops an adaptive, multi-scale drought early warning system that combines ensemble machine learning with local knowledge systems. By integrating satellite observations, climate forecasts, and community-reported indicators through a hierarchical Bayesian framework, we provide probabilistic drought forecasts at 1-6 month lead times with explicit uncertainty quantification and locally-relevant impact translations.*

## 1. Scientific Innovation and Refined Approach

We address three fundamental challenges in African drought prediction:

### *Data Integration with Quality Control*

We develop a Bayesian data fusion framework that weights observations by reliability, handles missing data through Gaussian processes, and incorporates local observations via mobile phone apps. Quality flags from satellite products guide adaptive weighting schemes.

### *Multi-Scale Prediction Architecture*

Instead of a single model, we employ a hierarchy:

- **Continental scale:** Deep learning on reanalysis data
- **National scale:** Physics-informed models incorporating local climate drivers
- **District scale:** Statistical downscaling with bias correction
- **Community scale:** Integration of indigenous indicators (animal behavior, plant phenology)

### *Dynamic Baseline Adaptation*

Unlike static models, our system continuously updates its baseline using online learning, adapting to changing climate patterns and land use without full retraining.

## 2. Robust Methodology

### *2.1 Tiered Data Strategy*

- **Tier 1:** Satellite products with uncertainty estimates (CHIRPS, MODIS, GRACE)
- **Tier 2:** National meteorological data with quality control algorithms
- **Tier 3:** Crowd-sourced observations via validated mobile platform
- **Tier 4:** Social media and market price scraping for early signals

### *2.2 Ensemble Architecture with Uncertainty Quantification*

- Base models: Gradient boosting, LSTM networks, and Gaussian processes
- Ensemble integration via Bayesian model averaging with time-varying weights
- Conformal prediction for distribution-free uncertainty intervals
- Explicit modeling of aleatory vs epistemic uncertainty

### *2.3 Validation Framework*

- **Baseline comparisons:** Standardized Precipitation Index, operational FEWS NET
- **Metrics:** Brier skill score, reliability diagrams, economic value analysis
- **Historical validation:** 2000-2020 with held-out drought events
- **Real-time validation:** Parallel running with operational systems

## 3. Stakeholder Co-Development

Partnership structure with IGAD, national meteorological services, and farming cooperatives ensures relevance:

- Monthly feedback loops with forecast users
- Participatory design of warning thresholds
- SMS/radio dissemination in local languages
- Integration with agricultural extension services

## 4. Realistic Timeline and Resources

**Year 1:** Data pipeline development, stakeholder mapping, pilot in 3 countries
**Year 2:** Model refinement, validation, expansion to 8 countries
**Year 3:** Operational transition, capacity building, sustainability planning
**Year 4:** Full deployment, local institution handover

**Budget:** €650,000 (includes local capacity building and infrastructure)

# Proposal 2: Hybrid Physics-ML Framework for Flood Prediction

*This research develops a hybrid framework combining simplified physics-based models with machine learning to predict floods in data-sparse West African rivers. By using differentiable hydrological models within neural architectures, we achieve physically-consistent predictions while learning from regional patterns, providing probabilistic forecasts with 24-72 hour lead times.*

## 1. Refined Scientific Approach

### *Differentiable Physics Core*

Instead of pure ML, we embed simplified hydrological equations (kinematic wave, Green-Ampt infiltration) as differentiable modules within neural networks. This ensures mass conservation while allowing parameter learning.

### *Hierarchical Transfer Learning*

We train on data-rich basins first, then transfer learned parameters to ungauged basins with similar characteristics, quantifying transferability uncertainty.

### *Human System Integration*

We explicitly model dam operations and irrigation withdrawals using reinforcement learning agents trained on historical operation patterns.

## 2. Enhanced Methodology

### *2.1 Hybrid Model Architecture*

- **Physics modules:** Differentiable routing, infiltration, evapotranspiration
- **ML modules:** Precipitation downscaling, parameter estimation, state correction
- **Coupling:** Alternating optimization between physics and ML components

### *2.2 Uncertainty-Aware Training*

- Multiple model initializations to capture structural uncertainty
- Dropout and ensemble techniques for prediction uncertainty
- Probabilistic loss functions accounting for observation errors

### *2.3 Comprehensive Validation*

- **Baseline:** HEC-RAS, GloFAS, persistence forecasts
- **Metrics:** Nash-Sutcliffe efficiency, peak timing/magnitude errors, false alarm rates
- **Cross-validation:** Leave-one-basin-out for transferability assessment
- **Event-based validation:** Focus on extreme events >10-year return period

## 3. Operational Implementation

- **Computational optimization:** Model distillation for edge deployment
- **Fallback mechanisms:** Simplified models when data streams fail
- **Real-time updating:** Ensemble Kalman filter for state correction
- **Warning protocols:** Integration with national disaster management systems

## 4. Realistic Timeline

**Year 1:** Physics module development, data collection, baseline establishment

**Year 2:** Hybrid model training, validation on historical events
**Year 3:** Pilot operational deployment in 2 basins, refinement
**Year 4:** Expansion to 5 major basins, technology transfer

# Proposal 3: Theory-Guided Discovery of Climate System Connections

*This research combines machine learning with dynamical systems theory to discover physically-meaningful teleconnections. By constraining pattern search with theoretical understanding of atmospheric dynamics and implementing rigorous statistical validation, we identify robust climate connections while controlling false discovery rates.*

## 1. Theory-Constrained Approach

### *Dynamical Guidance*

Search is constrained to patterns consistent with:

- Rossby wave propagation paths
- Conserved quantities (potential vorticity, energy)
- Known timescales of atmospheric/oceanic variability

### *Causal Hierarchy*

We distinguish between:

- Direct causation (momentum/heat transfer)
- Mediated connections (through intermediate regions)
- Common drivers (external forcing)

## 2. Rigorous Methodology

### *2.1 Hypothesis-Driven Search*

- Generate candidates based on dynamical theory
- Test using causal discovery algorithms
- Validate with mechanistic models

### *2.2 Statistical Validation*

- Bootstrap confidence intervals for all discoveries
- False discovery rate control using Benjamini-Hochberg
- Out-of-sample testing on independent time periods
- Ensemble agreement across multiple reanalysis products

## 3. Physical Interpretation Protocol

- Composite analysis of discovered patterns
- Energy/momentum budget analysis
- Numerical experiments with climate models
- Expert evaluation by dynamical meteorologists

# Proposal 4: Integrated Coastal Risk Framework for African Cities

*This research develops an integrated framework for compound coastal hazards that explicitly addresses data limitations through combining physics-based and statistical approaches. Using Bayesian networks to represent hazard dependencies and agent-based models for vulnerability, we provide actionable risk information despite deep uncertainties.*

## 1. Data-Adaptive Approach

### *Hierarchical Modeling*

Global data provides prior distributions, refined with any available local observations

### *Synthetic Event Generation*

Physics-based generation of plausible but unobserved compound events

### *Vulnerability Co-Assessment*

Community mapping of exposure and adaptive capacity

## 2. Uncertainty-Centric Framework

- Deep uncertainty methods (robust decision making, adaptation pathways)
- Stress-testing under multiple plausible futures
- Focus on decision-relevant metrics rather than precise probabilities

# Proposal 5: Multi-Constraint Framework for Climate Sensitivity

*This research develops a multi-constraint framework that combines paleoclimate, observational, and process-based constraints on climate sensitivity while explicitly accounting for structural uncertainties and potential non-analogues between past and future.*

## 1. Conservative Approach

- Report uncertainty reduction only when multiple independent constraints agree
- Explicit treatment of structural model uncertainty
- Recognition of state-dependence and pattern effects
- Transparent reporting of assumptions and limitations

## 2. Robust Validation

- Perfect model experiments to test constraint methodology
- Evaluation of constraint consistency across model generations
- Assessment of constraint independence

## Conclusion

These improved proposals represent a paradigm shift in climate science research for Africa, moving from purely academic exercises to operational, actionable systems. By explicitly addressing data quality issues, validation challenges, and stakeholder needs, these projects bridge the gap between scientific innovation and real-world impact.

The enhanced frameworks ensure:

- **Scientific rigor** through comprehensive validation and uncertainty quantification
- **Operational feasibility** through computational optimization and fallback mechanisms
- **Local relevance** through stakeholder co-development and capacity building
- **Sustainability** through phased implementation and technology transfer

Together, these proposals form a comprehensive research program that advances climate science while directly addressing the urgent adaptation needs of African communities facing unprecedented climate challenges.

## Implementation Strategy and Synergies

The five proposals create synergies through:

- **Shared Data Infrastructure:** Common preprocessing pipelines and quality control algorithms
- **Cross-Validation:** Teleconnections from Proposal 3 improve predictions in Proposals 1 and 2
- **Cascading Insights:** Climate sensitivity constraints (Proposal 5) refine all impact projections
- **Methodological Transfer:** Uncertainty quantification techniques apply across all proposals
- **Capacity Building:** Unified training programs for local institutions

**Total Program Budget:** €2.5M over 4 years
**Personnel:** 5 PhD students, 3 postdocs, 2 research engineers
**Infrastructure:** Cloud computing, local servers, mobile platforms
**Capacity Building:** Training workshops, technology transfer, documentation
**Dissemination:** Publications, conferences, stakeholder engagement

*--- End of Document ---*