Project 1

Project Title: A Dynamic, AI-Based Susceptibility Model for Sierra Leone's Landslides, Controlling for Observational Sampling Bias

Principal Investigator: [Your Name]
Date: November 17, 2025

1. Introduction and Rationale

In regions like Sierra Leone, landslides are a lethal, recurring hazard. The 2017 Regent disaster in Freetown served as a tragic case study of the forces at play: steep topography, intense seasonal rainfall, and rapid land-use change. To mitigate future risk, a nationwide Landslide Susceptibility Map (LSM) is a critical tool for planning and early warning.

However, creating a scientifically valid LSM in a data-sparse region is plagued by a fundamental challenge: **observational sampling bias**. Existing landslide inventories are incomplete, "biased" towards events that are easily seen (e.g., near roads or cities), and miss those in remote or forested areas. A machine learning (ML) model trained on this biased data will not produce a valid "susceptibility" map; it will produce a "map of where landslides are easy to spot."

This project will develop a high-resolution LSM by *explicitly* addressing this sampling bias. Using a Random Forest (RF) model, we will train a classifier to distinguish known landslide "positive" points from a carefully constructed, environmentally-similar "Target-Oriented Background" dataset. This robust methodology will separate the signal of *true* biophysical susceptibility (slope, soil, rain) from the *noise* of human observation.

2. Literature Review and Research Gaps

ML models (RF, SVM) are now standard in LSM (Reichenbach et al., 2018). However, most studies fail to rigorously address the "absence" data problem. They often use a "random" selection of "non-landslide" points, which is statistically flawed and includes areas that are simply *unobserved* (not truly *stable*).

The **key research gap** for this region is the development of an LSM that:

1. **Controls for Sampling Bias:** Moves beyond "random sampling" to a more rigorous pseudo-absence generation technique (e.g., Target-Oriented Background sampling).
2. **Integrates Dynamic Triggers:** Properly combines high-resolution *static* data (30m slope) with low-resolution *dynamic* data (5km rainfall) in a methodologically sound way.
3. **Produces a Useful Tool:** Delivers an *operational model* for dynamic forecasting, not just a static "what-if" map.

This project addresses these gaps by building a model designed to be robust against a biased inventory and built for a real-world, dynamic forecasting application.

3. Research Questions and Objectives

Primary Objective: To develop a high-resolution (30m), bias-controlled Landslide Susceptibility Model for Sierra Leone, capable of dynamic risk assessment.

Research Questions:

1. Can a Random Forest model, trained on a biased inventory (P points) versus a Target-Oriented Background (A points), successfully distinguish the biophysical drivers of landslides?
2. What is the relative importance of static predictors (slope, LULC, distance-to-fault) versus dynamic predictors (antecedent rainfall) in triggering landslides?
3. How can the trained, validated model be operationalized to provide a *dynamic daily forecast* of landslide susceptibility, rather than a single static map?

4. Data and Methodology

A. Datasets:

- **Target Variable 1 (Positive 'P' Points):** A **Landslide Inventory** compiled from literature, government reports, and manual interpretation of Google Earth / Sentinel-2 imagery.
- **Target Variable 2 (Background 'A' Points):** A **Target-Oriented Background (TOB)** dataset. This is *not* a random sample. We will sample background points that are 1) not known landslides, but 2) share similar environmental characteristics (e.g., slope, geology, proximity to P-points) as the known inventory. This forces the model to find the *subtle*

differences, rather than just learning "landslides happen on hills."

- **Static Predictors (30m):**
    - **DEM (e.g., ALOS PALSAR 30m):** Used to derive **Slope**, **Aspect**, **Curvature**, and **Topographic Wetness Index (TWI)**.
    - **LULC (Sentinel-2 10m):** Classified (using RF) into "Forest," "Urban/Bare," "Agriculture."
    - **Infrastructure:** Distance to Roads, Distance to Rivers.
- **Dynamic Predictors (5km):**
    - **CHIRPS (Daily):** Used to derive **3-day**, **7-day**, and **15-day antecedent rainfall** for each event.

B. Methodology (Work Packages):

- **WP1: Data Collation and Bias Control (Months 1-9)**
    1. Compile the landslide inventory (P-points).
    2. Generate the TOB dataset (A-points). This is a critical, rigorous step.
    3. Process all 30m static predictors and stack them into a single data cube.
- **WP2: Data Harmonization and Training Matrix (Months 10-12)**
    1. **Addressing Resolution Mismatch (Flaw Fix):** The methodology will be explicit. For model training, the 5km CHIRPS rainfall value for the event date will be assigned to its corresponding 30m point. This assumes rainfall *within* a 5km cell is homogenous.
    2. Extract all predictor values (static and dynamic) for all P and A points to create the final training matrix.
- **WP3: Model Training and Validation (Months 13-18)**
    1. Train a **Random Forest Classifier** (sklearn) to distinguish "P" from "A" points.
    2. Perform hyperparameter tuning and k-fold cross-validation.
    3. Validate model using AUC (Area Under the Curve). An AUC > 0.8 is considered robust.
    4. Analyze feature_importances_ to rank all predictors and answer RQ2.
- **WP4: Operationalization (The *Real* Deliverable) (Months 19-24)**
    1. **Addressing "Worst-Case Map" Flaw:** We *reject* the creation of a single, arbitrary static map.
    2. The **primary deliverable is the trained, validated model file (.pkl) and its associated code.**
    3. This model is a *function*: Prob = f(Slope, LULC, ... , 3_day_rain).
    4. We will demonstrate its use by showing how a meteorological agency could feed it *tomorrow's rainfall forecast* to generate a *dynamic, daily landslide hazard map*.

5. Expected Outcomes and Scientific Contribution

- **Deliverable:** An open-source, validated AI model for landslide susceptibility in Sierra

Leone, along with a robust methodology for controlling observational sampling bias.

- **Scientific Contribution:** This research moves beyond naive LSM. It provides a scientifically-defensible *methodology* for data-sparse regions, demonstrating how to handle biased inventories and produce an *operational* tool rather than a static map.
- **Societal Impact:** This model provides the *engine* for a future national landslide early warning system, allowing for targeted alerts and saving lives.

6. Limitations

- The model's quality is still ultimately dependent on the *quality* of the (biased) P-point inventory.
- The model assumes **stationarity**—that the biophysical relationships (slope, rain) that cause landslides will remain the same in a future climate.

Project 2

Project Title: Synergistic Analysis of Urban Expansion and the Surface Urban Heat Island (SUHI) in Freetown, Sierra Leone, Using AI and a MODIS/Landsat Thermal Sharpening Approach

Principal Investigator: [Your Name]
Date: November 17, 2025

---

1. Introduction and Rationale

Freetown's rapid, often informal, urban expansion is transforming its landscape, replacing cool, vegetated hillsides with impervious, built-up surfaces. This directly creates an **Urban Heat Island (UHI)** effect, exacerbating heat stress, increasing energy demand, and posing a significant public health risk in its humid tropical climate.

Quantifying this UHI effect to inform urban planning is notoriously difficult. The standard tool, Landsat, has a 16-day revisit time, which in a tropical, cloud-covered city like Freetown, results in *almost no usable data*. Conversely, daily satellites like **MODIS** have 1km resolution, which is too coarse to see *within* the city. This project will solve this "temporal-vs-spatial" resolution problem.

We will develop a **"thermal sharpening" AI model** that fuses these two data sources. The model will be trained to learn the statistical relationship between low-res (1km) daily MODIS LST and high-res (30m) predictors (like NDVI from Landsat/Sentinel-2). This allows us to generate a *synthetic, 30m, daily* Land Surface Temperature (LST) dataset, finally enabling a high-resolution analysis of Freetown's thermal dynamics.

2. Literature Review and Research Gaps

LST analysis via remote sensing is a mature field (Voogt & Oke, 2003). However, studies in tropical cities are sparse due to the cloud-cover barrier. "Thermal sharpening" (or downscaling) is a state-of-the-art technique to address this (e.g., using TsHARP, AT-PRM). Using AI models like Random Forest (RF) for this is an emerging and powerful approach (Hassan et al., 2021).

The **key research gaps** for Freetown are:

1. **A Methodological Barrier:** No LST time-series exists because of the cloud/resolution problem. Our proposed thermal sharpening model *directly* addresses this gap.
2. **No Robust SUHI Metric:** "UHI" is often defined by an arbitrary "rural" reference point, which is unscientific. We lack a reproducible metric for Freetown's heat intensity.
3. **No UHI-LULC Linkage:** We do not know, in quantitative terms, the thermal "cost" (in °C) of converting one hectare of Freetown's forest to urban settlement.

3. Research Questions and Objectives

Primary Objective: To develop a high-resolution (30m) daily LST dataset for Freetown (2000-2024) and use it to quantify the relationship between urban expansion and the Surface Urban Heat Intensity (SUHI).

Research Questions:

1. Can a Random Forest model, trained on clear-sky data, accurately predict high-resolution (30m) LST from low-resolution (1km) MODIS LST and high-res static predictors (NDVI, LULC)?
2. What is the magnitude of Freetown's **Surface Urban Heat Intensity (SUHI)**, defined as the LST difference between "urban" and "vegetation" pixels *within* the city boundary?
3. How has the SUHI co-evolved with land-use-land-cover (LULC) change from

2000-2024?

4. What is the quantitative thermal impact (in ΔLST) of converting "forest" to "dense urban" pixels?

4. Data and Methodology

A. Datasets:

- **LST (Temporal): MODIS** Daily 1km LST (MOD11A1).
- **LST (Spatial): Landsat 8/9** 30m (resampled) LST.
- **LULC & Predictors: Sentinel-2** (10m) and **Landsat** (30m) Surface Reflectance. Used to derive **NDVI**, **NDBI** (Normalized Difference Built-up Index), **LULC classification**, and **Albedo**.
- **Topography: 30m DEM** (elevation, slope).

B. Methodology (Work Packages):

- **WP1: LULC Time-Series Classification (Months 1-6)**
  1. Using GEE or a local Python environment, create cloud-free mosaics for Freetown (e.g., 2000, 2005, 2010, 2015, 2020, 2024).
  2. Train an RF classifier to produce 30m LULC maps ("Dense Urban," "Sparse Urban," "Vegetation," "Water") for each time-period.
- **WP2: Thermal Sharpening AI Model (The Core Novelty) (Months 7-18)**
  1. **Addressing the Cloud/Landsat Flaw:** We will fuse MODIS and Landsat.
  2. Identify all clear-sky days where Landsat LST data *exists*.
  3. On these days, build a "training matrix" for the RF model:
     - **Target (y):** The high-res 30m Landsat LST.
     - **Predictors (X):** The low-res 1km MODIS LST (upsampled), and all high-res 30m predictors (NDVI, NDBI, LULC, Elevation, Albedo).
  4. Train the RF model. It will learn the complex, non-linear relationships (e.g., "a 1km MODIS pixel of 30°C that is 80% forest and 20% urban has this 30m thermal pattern...").
  5. **Application:** For *any* day, feed the model the *daily* 1km MODIS LST and the *static* 30m predictors. The model will *predict* a synthetic, high-resolution 30m LST map.
- **WP3: SUHI Quantification and Time-Series Analysis (Months 19-24)**
  1. Using the new 30m synthetic-daily LST dataset, create a stable time-series.
  2. **Addressing the "Rural Reference" Flaw:** We *reject* an arbitrary rural reference.
  3. We will calculate the SUHI as:

     $$SUHI(t) = \text{mean}(LST_{\text{urban\_pixels}, t}) - \text{mean}(LST_{\text{vegetation\_pixels}, t})$$

4. We will plot the time-series of SUHI alongside the time-series of % urban area (from WP1) to answer RQ3.
5. We will perform a change-detection analysis to quantify the ΔLST of pixels that *changed* from "Vegetation" in 2000 to "Urban" in 2020 (RQ4).

5. Expected Outcomes and Scientific Contribution

- **Deliverable:** The first high-resolution (30m), daily-estimated LST dataset for Freetown. A robust quantification of its SUHI and its link to 20 years of urban growth.
- **Scientific Contribution:** A transferable, AI-based thermal sharpening *methodology* for tropical, cloud-prone cities. This is a significant methodological advance that unlocks UHI research in previously impossible locations.
- **Societal Impact:** Provides granular "hotspot" maps for the Freetown City Council to target cooling interventions (e.g., tree planting, cool roofs) in the most vulnerable neighborhoods.

Project 3

Project Title: AI-Based Statistical Downscaling of CMIP6 Precipitation for West Africa: A Non-Stationary Approach with Satellite-Based Training

Principal Investigator: [Your Name]
Date: November 17, 2025

---

1. Introduction and Rationale

West African food security and economic stability are inextricably linked to the West African Monsoon (WAM). Understanding how the WAM will change is a primary goal of climate science. However, Global Climate Models (GCMs) like **CMIP6** are far too coarse (100-200km) to inform regional impact models (e.g., crop or hydrology models), which need 5-10km data.

Statistical Downscaling (SD) aims to bridge this gap. However, AI-based SD is plagued by two fundamental challenges:

1. **Stationarity:** Most models *assume* the statistical link between large-scale weather (e.g.,

humidity) and local rain is *stable* over time. In a warming world, this assumption is almost certainly false.
2. **Bias:** Raw CMIP6 output (e.g., its humidity values) is biased relative to observations (ERA5). Feeding this biased data into a trained AI model produces meaningless output.

This project will develop an AI-based SD model that *directly confronts* these two flaws. The core research will first **test for, and then account for, non-stationarity** in the WAM. It will then apply a rigorous **Quantile Delta Mapping (QDM)** bias-correction method, ensuring a scientifically-sound application to CMIP6 projections.

2. Literature Review and Research Gaps

AI-driven SD (e.g., Random Forest, ANNs) is a rapidly advancing field (Gutiérrez et al., 2019). Most applications, however, are in data-rich mid-latitudes and often carry the un-tested stationarity assumption. In the WAM, this is particularly dangerous, as the system is known to be highly non-linear and subject to abrupt shifts. The application of robust bias-correction (like QDM) *before* downscaling is essential but not always practiced.

Key Research Gaps:

1. **A Test of Stationarity:** Does a statistical relationship trained on 1980s WAM data still hold true for the 2010s WAM? This "test" is rarely performed and is a critical scientific gate.
2. **A Bias-Corrected, Satellite-Trained Model:** No well-documented, open-source AI-SD model for West Africa exists that is trained on high-res satellite "truth" (CHIRPS) and correctly handles CMIP6 biases using QDM.

3. Research Questions and Objectives

Primary Objective: To develop, validate, and apply a *non-stationary*, bias-corrected AI-based statistical downscaling model for daily precipitation over West Africa (to 5km).

Research Questions:

1. Are the statistical relationships between large-scale atmospheric predictors (from ERA5) and fine-scale precipitation (from CHIRPS) *stationary* over the 1981-2023 period?
2. If non-stationarity is detected, can a model that includes a non-stationarity co-variate (e.g., regional mean temperature) outperform a stationary model?
3. How do the high-resolution (5km) downscaled projections of future WAM rainfall (e.g.,

rainfall extremes, seasonal totals) differ from the raw, bias-corrected CMIP6 projections?

4. Data and Methodology

A. Datasets:

- **High-Res "Truth" (Target): CHIRPS** (0.05° / ~5km, daily) rainfall.
- **Training Predictors (Features): ERA5 Reanalysis** (0.25°, daily). Variables: 850hPa specific humidity (q), zonal/meridional winds (u/v), 500hPa geopotential height (z), MSLP.
- **Future Predictors (Features): CMIP6 Ensemble** (~10 models, daily) for the same variables, under SSP2-4.5 and SSP5-8.5.

B. Methodology (Work Packages):

- **WP1: The Non-Stationarity Test (Crucial Gate) (Months 1-9)**
  1. **Addressing the Stationarity Flaw:** We must test this assumption first.
  2. Create a "training" dataset (ERA5/CHIRPS, 1981-2000) and a "test" dataset (2001-2023).
  3. Train a Random Forest Regressor on the 1981-2000 "cool" period.
  4. Use this "cool-trained" model to predict rainfall in the 2001-2023 "warm" period.
  5. **Hypothesis:** The model's error (bias) will systematically increase over time, *proving* that the relationships are non-stationary and a simple SD model is invalid.
- **WP2: Non-Stationary Model Development (Months 10-18)**
  1. If (as expected) non-stationarity is found, we re-design the model.
  2. We will train a new RF model on the *full* 1981-2023 period, but we will add a **non-stationarity co-variate** as a predictor.
  3. This co-variate will be a 5-year running mean of regional 2m temperature. This allows the model to "learn" that the relationship between (e.g.) humidity and rain *is different* in a warmer world.
- **WP3: Bias Correction and CMIP6 Application (Months 19-30)**
  1. **Addressing the Bias Flaw:** We *cannot* use raw CMIP6 data.
  2. We will apply **Quantile Delta Mapping (QDM)** to all CMIP6 predictor variables (q, u, v, z, etc.) to bias-correct them against the ERA5 historical baseline. This adjusts their distributions while preserving their future-projected *change* (delta).
  3. Feed this bias-corrected CMIP6 data (1981-2100) into our *trained, non-stationary* RF model from WP2.
  - **Output:** A 5km, daily, bias-corrected, non-stationarily-downscaled precipitation ensemble for West Africa.
- **WP4: Analysis of Projections (Months 31-36)**
  1. Analyze this new, high-value dataset.
  2. Compare future (2070-2100) vs. historical (1981-2010) WAM totals, monsoon onset,

and (most importantly) extreme rainfall frequency (e.g., # of 50mm/day events).

5. Expected Outcomes and Scientific Contribution

- **Deliverable:** A new, 5km daily precipitation ensemble for West Africa, created using a scientifically-defensible, non-stationary, and bias-corrected methodology.
- **Scientific Contribution:** This project provides a *robust methodological framework* for AI-based downscaling in a changing climate. The non-stationarity test (WP1) is a critical contribution, moving the field from "blind application" to "critical assessment" of AI-SD models.
- **Societal Impact:** This dataset is the *direct input* needed for all subsequent impact models (hydrology, crop, health), providing the most credible high-resolution projections for the region to date.

Project 4

Project Title: Future Changes in Extreme Precipitation in Coastal West Africa: A Non-Stationary EVT Analysis of the **CORDEX-Africa** Ensemble

Principal Investigator: [Your Name]
Date: November 17, 2025

---

1. Introduction and Rationale

The coastal megacities of West Africa (Lagos, Freetown, Accra) are defined by their vulnerability to extreme rainfall and flash floods. The infrastructure in these cities (storm drains, bridges) was designed based on historical rainfall statistics. However, anthropogenic climate change is rendering these historical "return levels" (e.S. the "1-in-50-year storm") obsolete.

A scientifically-robust assessment of *future* return levels is urgently needed. This requires two things:

1. **Appropriate Data:** Coarse 100km GCMs (like raw CMIP6) are scientifically *invalid* for this task, as they do not resolve coastlines or the storms themselves. We *must* use

high-resolution Regional Climate Models (RCMs).
2. **Appropriate Methods:** The climate is **non-stationary**; it is warming. Analyzing a "future" 30-year block (e.g., 2070-2100) as if it were stable is a known flawed method that underestimates risk.

This project will use the high-resolution **CORDEX-Africa** (Coordinated Regional Climate Downscaling Experiment) ensemble and apply a rigorous **non-stationary Extreme Value Theory (EVT)** framework. This will provide the first credible, engineering-relevant estimates of how extreme rainfall risk will evolve *per degree* of global warming.

2. Literature Review and Research Gaps

EVT (Coles, 2001) is the statistical gold-standard for extreme analysis. Non-stationary EVT, where GEV parameters are modeled as a function of a co-variate (like time or global mean temp), is the state-of-the-art for climate change studies (as seen in Climatematch W2D3).

Key Research Gaps:

1. **The Scale Mismatch:** Most CMIP6-based EVT studies use GCMs that are too coarse for this region. This project *fixes* this by using the **CORDEX-Africa** (0.44°/0.22°) ensemble.
2. **The Methodology Flaw:** Many studies still use a "stationary future slice" approach. This project *rejects* this, using a *fully non-stationary* GEV fit as its core methodology.
3. **A Lack of CORDEX Validation:** The CORDEX models themselves must be validated. Does their *historical* extreme rainfall (driven by ERA-Interim) match observations (CHIRPS)? This validation is a critical, missing step.

3. Research Questions and Objectives

Primary Objective: To quantify the projected changes in the intensity and frequency of extreme daily precipitation in coastal West Africa using a non-stationary EVT framework applied to the CORDEX-Africa RCM ensemble.

Research Questions:

1. How well do the CORDEX-Africa historical runs (1981-2005) reproduce the observed EVT statistics (e.g., 20-year return levels) from the CHIRPS dataset?
2. Using a non-stationary GEV model, what is the "sensitivity" of extreme rainfall (e.g., the 50-year return level) to global warming, expressed in mm/day per °C of GMST?
3. Based on this sensitivity, what is the projected 50-year return level under SSP2-4.5 (e.g.,

+2°C) and SSP5-8.5 (e.g., +4°C) scenarios?

4. Data and Methodology

A. Datasets:

- **RCM Projections: CORDEX-Africa** Ensemble (e.g., 0.44° res). Daily precipitation (pr) for 1981-2100, for all available models and SSPs.
- **Observational Validation: CHIRPS** (0.05°) and/or **ERA5** (0.25°). Daily precipitation for 1981-2014.
- **Co-variate:** Global Mean Surface Temperature (GMST) time-series from the parent CMIP6 models.

B. Methodology (Work Packages):

- **WP1: RCM Validation (Crucial Gate) (Months 1-9)**
    1. **Addressing the Scale Flaw:** We use CORDEX, not CMIP6.
    2. Extract Annual Maxima Series (AMS) from the CORDEX *historical* runs (e.g., 1981-2005) and from *observational* (CHIRPS) data for the same period.
    3. Fit a stationary GEV to both.
    4. Spatially map the bias: (CORDEX_ReturnLevel - CHIRPS_ReturnLevel).
    5. This allows us to identify and (if necessary) discard RCMs that fail to capture historical extremes, and to understand the "wet" or "dry" bias of the ensemble.
- **WP2: Non-Stationary EVT Modeling (The Core) (Months 10-18)**
    1. **Addressing the Methodology Flaw:** We *reject* the stationary future-slice method.
    2. For each *skilled* CORDEX model (from WP1), fit a single **non-stationary GEV model** to the *entire* 1981-2100 AMS.
    3. The GEV's location parameter (μ) will be modeled as a function of GMST:

       $$\mu(t) = \mu_0 + \alpha \times \text{GMST}(t)$$
    4. This alpha (α) parameter is the **sensitivity** (RQ2). It is the (e.g.) mm/day per °C change.
    5. This is the scientifically-robust, state-of-the-art method from W2D3.
- **WP3: Return Level Projections (Months 19-24)**
    1. Using the fitted non-stationary model, we can now calculate return levels for *any* level of warming.
    2. We calculate the 50-year return level for the historical baseline (e.g., +1°C warming).
    3. We then calculate the 50-year return level for a +2°C world (SSP2-4.5) and a +4°C world (SSP5-8.5).
    4. We can also calculate the *change in frequency* (e.g., "The historical 1-in-50-year event becomes a 1-in-10-year event at +3°C warming").

5. Expected Outcomes and Scientific Contribution

- **Deliverable:** A set of maps and metrics that provide the most credible, engineering-relevant projections of future extreme rainfall risk for coastal West Africa.
- **Scientific Contribution:** This is the first known study to *both* use the high-res CORDEX ensemble *and* apply a rigorous non-stationary GEV framework to this region. It provides a methodological template for all other CORDEX domains.
- **Societal Impact:** Provides the *direct numbers* needed by engineers and urban planners to update infrastructure design codes and build resilience.

Project 5

Project Title: Development and Evaluation of an **Interpretable**, Spatially-Resolved AI Climate Model Emulator for Regional Scenario Analysis

Principal Investigator: [Your Name]
Date: November 17, 2025

---

1. Introduction and Rationale

Earth System Models (ESMs) are computationally-paralyzing. A single simulation for one future scenario can take months, creating a "computational bottleneck" that prevents policy-makers from exploring the thousands of possible futures. **AI-driven emulators** solve this, learning an ESM's behavior and replicating its output in seconds.

However, current emulators suffer from two critical flaws:

1. **The "Black Box" Problem:** Most emulators are not *interpretable*. We don't know *why* they give a certain answer. A policy-maker cannot trust a multi-billion dollar decision to a "black box."
2. **The "Global Mean" Problem:** Most emulators only predict Global Mean Temperature (GMT), which is useless for a regional planner in (e.g.) West Africa.

This project will build an emulator that *solves* these flaws. We will build an **interpretable** emulator (using **eXplainable AI - XAI**) and a **spatially-resolved** emulator (using **Dimension Reduction**). This moves the science from a "computer science" exercise to a *trustworthy and useful* tool for climate science.

2. Literature Review and Research Gaps

ML emulators (RF, NNs) are a new standard (Beusch et al., 2020), and ClimateBench (W2D4) is the key dataset. The field is now moving beyond *metric accuracy* (low RMSE) to *scientific accuracy*. This involves **XAI** techniques (like **SHAP**) to "open the black box" and see if the AI learned real physics (Gjermundsen et al., 2021).

Furthermore, regional emulation is the frontier. A naive "multi-target" model is inefficient and ignores spatial correlation. The state-of-the-art approach is to use **Dimension Reduction** (e.g., Principal Component Analysis - PCA) to find the dominant spatial *patterns* of warming and train the emulator to predict those.

Key Research Gaps:

1. **Interpretability:** Emulators are rarely validated for *physical plausibility* using XAI.
2. **Spatial Emulation:** A robust, PCA-based regional emulator has not been widely implemented or tested against simpler models.

3. Research Questions and Objectives

Primary Objective: To develop and evaluate a machine learning (RF and NN) emulator that is **interpretable** (via XAI) and **spatially-resolved** (via PCA) for regional scenario analysis.

Research Questions:

1. How does the performance (RMSE) of a simple Random Forest (RF) compare to a Deep Learning (CNN-LSTM) model for GMT emulation?
2. Using XAI (SHAP), do these models learn physically-plausible relationships (e.g., the logarithmic forcing of CO_2), or do they "cheat" using dataset-specific correlations?
3. Can a PCA-based emulator (trained to predict the scores of the dominant spatial warming patterns) skillfully reproduce regional temperature anomalies?
4. How does this PCA-based approach (RQ3) compare in accuracy and efficiency to a naive multi-target model?

4. Data and Methodology

A. Datasets:

- **ClimateBench (W2D4):** The core dataset.
  - **Inputs (X):** Emissions/Concentrations time-series (SSP scenarios).
  - **Outputs (y - Global):** GMT time-series from 3 ESMs.
- **Full Spatial Data:** The *original, gridded* CMIP6 temperature maps from the models used in ClimateBench.

B. Methodology (Work Packages):

- **WP1: Baseline Emulators (GMT) (Months 1-6)**
  1. Implement the W2D4 workflow: Train a **Random Forest** and a **CNN-LSTM** to predict GMT from emissions (X).
  2. Benchmark their performance (RMSE) on held-out test scenarios (e.g., SSP1-1.9). This answers RQ1.
- **WP2: Interpretability (XAI) (Months 7-12)**
  1. **Addressing the "Black Box" Flaw:** We open the box.
  2. Apply **SHAP (SHapley Additive exPlanations)** to the trained RF and NN models.
  3. Generate "SHAP dependence plots" for each input predictor.
  4. **Hypothesis Test (RQ2):** We will test if the SHAP plot for CO2 forcing follows a logarithmic curve, and if the plot for aerosol forcing shows a negative (cooling) relationship. This validates the *physical plausibility* of the AI.
- **WP3: Spatially-Resolved Emulator (PCA) (Months 13-24)**
  1. **Addressing the "Regional" Flaw:** We use dimension reduction.
  2. Take the *full spatial maps* of temperature anomaly (1850-2100) from the ESMs.
  3. Run **Principal Component Analysis (PCA)** on these maps.
  4. The output will be the dominant spatial *patterns* (the "EOFs" or eigenvectors) and their *time-series scores* (the "PCs" or amplitudes). PC1 will be "global warming," PC2 "land-sea contrast," PC3 "polar amplification," etc.
  5. **Re-train the AI:** The new task is *not* to predict GMT. The task is to predict the **time-series scores of the first 5 PCs** from the emissions inputs (X).
  6. To get a full map, we just multiply the predicted scores by their patterns.
- **WP4: Validation and Exploration (Months 25-36)**
  1. Validate the spatial emulator (WP3). Does it accurately reconstruct regional temperatures (RQ3)?
  2. Compare this (elegant) PCA method to a (naive) "multi-target" model trained to predict 20 regional means (RQ4).
  3. Use the final, trusted emulator to explore 10,000+ hypothetical scenarios.

5. Expected Outcomes and Scientific Contribution

- **Deliverable:** An open-source, **interpretable and spatially-resolved** AI climate emulator.
- **Scientific Contribution:** This project provides a critical *methodological* contribution by:
    1. Establishing a framework for XAI-based *validation* of emulators.
    2. Demonstrating the superiority of a PCA-based spatial emulator.

    This moves emulators from "black box" toys to trusted scientific tools.
- **Societal Impact:** Creates a tool that regional policy-makers can *actually use* and *trust*, allowing them to see the regional consequences of global emissions pathways in real-time.