The East African Climate Paradox

[redacted]

University of Oxford

A project submitted for

Topics in Computational Biology

2024

Contents

Chapter 1

Introduction

Due to relatively consistent daily temperatures and sunlight exposure, seasons in equatorial climates are often delineated by rainfall. Climatic seasons in East Africa, for instance, are characterized by a bimodal rainfall pattern, in which the region experiences the long rains from March to May (masika) and the short rains from October to December (vuli), with alternating dry seasons in between[1]. With an economy highly dependent on the agricultural sector, predictable rainy seasons are paramount to food security, economic stability, and health outcomes in the region. The disagreement between model predictions and observed trends in rainfall patterns is therefore concerning; as climate models predict longer, wetter rainy seasons [43], reality has experienced historic droughts and unreliable rains [28]. Though this discrepancy–known as the East African Climate Paradox–is of growing interest to climate scientists and policymakers, the reasons for its existence are not yet well understood.

1.1 Rainfall Dependence in East Africa

We begin with the nontrivial problem of defining what constitutes East Africa. The East African Community (EAC), an intergovernmental institution designed to uphold the political and economic interests of member countries, officially consists of eight countries, while the most expansive definition of East Africa consists of 18, defining the region from the southern tip of Mozambique to the northern tip of Sudan [4]. Here we centralize our discussion on the areas most frequently cited as part of East Africa: Tanzania, Kenya, Uganda, Burundi, and Rwanda, as well as the Greater Horn of Africa, including Somalia, Ethiopia, South Sudan, Djibouti, and Eritrea.

[1] We adopt here the widely used Swahili terminology, though these seasons are also referred to as kiremt/belg (Amharic), gu/deyr (Somali), and ganna/bona (Oromo), with slightly different month splits depending on the relevant subregion.

As climate patterns have no interest in political borders, we only use a country-based definition for convenience. As such, we also recognize portions of additional countries that deserve the same categorization on the basis of climatic zones, including eastern Democratic Republic of the Congo (DRC); northern Malawi, Mozambique, and Zambia; and southern Sudan.

Figure 1.1: Boundaries of East Africa adopted from [8]

As cited by the EAC, over 80% of its nearly 300 million+ citizens are employed in agrarian industries [1]. With such a high dependence on agriculture, there is an accordingly high dependence on climate and weather patterns. By hindering agricultural production, unreliable weather patterns can drastically alter downstream economic, health, and development outcomes [8, 34, 38, 5, 6] in addition to the direct negative impacts of extreme weather events. Most recently, historic droughts have plagued the region, leading to food insecurity, malnutrition, disease, flooding, population displacement, and conflict exacerbation among other harmful effects. In Somalia, arguably the hardest hit by recent droughts, over 250,000 people died in the 2011 famine after three failed rainy seasons, with an additional drought in 2017 resulting in near-famine for 6.2 million [5, 6, 3]. Clearly, there is no dearth of examples indicating the importance of weather and climate patterns on human well-being, particularly in East Africa.

1.2 Why Climate Modeling?

We have thus far failed to concretely distinguish between climate and weather. Noting that rainfall–the most frequent culprit of climate/weather-induced harm in East Africa–is the product of the complex interplay between climate and weather, it is important to map both to more precise definitions. In climatology, the distinguishing factor is usually scale, both of time and space. Climate typically refers to larger scale spatiotemporal patterns, whereas weather refers to the shorter, day-to-day conditions of a particular area. A region's climate essentially refers to the average of its weather patterns, inextricably linking these notions together.[2] Climate modeling is thus an endeavor not only to better understand large-scale, dynamic physical systems, but also to develop expectations for daily weather patterns. Accurately predicting weather patterns in advance, particularly unfavorable weather patterns, enables early intervention when mitigation strategies are necessary. Given the complex polycrises facing regions such as East Africa, we thus turn to climate modeling as an imperative to adequately respond to a changing climate.

1.3 Overview of Discussion

We first provide a technical background to climate modeling in Chapter 2, taking a first-principles approach to understanding coupled models before exploring the modern landscape of CMIP models. Armed with these foundations, chapters 3 and 4 develop and critique our current understanding of the East African Climate Paradox, both exploring and suggesting potential avenues for improvement. We conclude with a brief acknowledgment of related research directions, re-emphasizing the significance of this work in protecting human well-being.

[2] Though the terms *climatic weather* and *climate weather* are occasionally used in the literature, we do not use them here to avoid ambiguity.

Chapter 2

Technical Background

For a full treatment of climate modeling, the reader is advised to consult [31, 17, 22], which trace decades of progress and historical considerations in understanding modern approaches. We focus here on the essential ingredients needed to analyze the East African Climate Paradox.

2.1 Coupled Climate Models

The simplest climate model, deemed an Energy Balance Model (EBM), focuses on "balancing the planetary radiation budget" [31], equalizing the radiative energy input and output of the Earth. It has a surprisingly elegant form,

$$(1-\alpha)\frac{S}{4} = \sigma T_e^4 \tag{2.1}$$

where S is the solar constant, T_e is the effective temperature of the Earth, α is an albedo constant, and σ is the Stefan-Boltzmann constant. Equation 2.1 provides a course-grain overview of the Earth's energy balance, and although useful in some contexts, fails to provide the resolution or specificity desired for many climate modeling applications. When focusing on smaller spatiotemporal regions of interest or particular components of the Earth system, we require finer-grain details. We can achieve this goal by conceptualizing the Earth as a collection of interrelated submodels; by independently modeling both the physical systems that comprise the global climate and the interactions between such systems, we can answer a wider array of questions within climate modeling.

This "divide and conquer" mentality is the fundamental insight at the heart of coupled models, most of which focus on the four major components of our climate system: the atmosphere, oceans, ice sheets, and land. The earliest coupled models, deemed general circulation models (GCMs)[1], focused primarily on characterizing the circulating fluid processes in the atmosphere and oceans. Modern coupled models are more complex, incorporating increasingly sophisticated representations of the atmosphere, oceans, ice sheets, and land beyond just circulating fluids.

We now briefly introduce the core mathematics underpinning the major components of modern coupled models.

Preliminary 1: Discretization

The equations we seek to solve in climate modeling generally do not have analytical solutions; we therefore rely on spatiotemporal discretizations and numerical methods to solve for variables of interest. Our choice of discretization scheme can vary greatly depending on the desired resolution and application. In general, we divide the Earth into a set of discrete chunks via a spatial grid and step forward/backward in time by discrete time-steps. While the choice of time-step is usually a function of computational capacity and target resolution, the choice of spatial grid is typically more involved. Common spatial discretization schemes include Arakawa grids, linearized latitude-longitude grids, cubed spherical units, icosahedral grids, and the yin-yang grid.

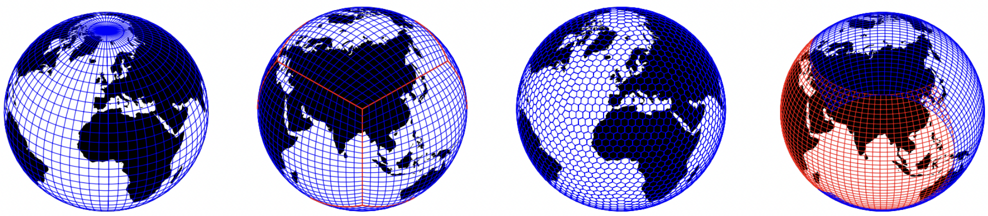

Figure 2.1: Spatial grid schemes [26]. From L-R, we show lat-long, cubed sphere, icosahedral, and yin-yang spatial discretization schemes.

Preliminary 2: Conservation Laws

Whether working with the atmosphere, oceans, land, or ice sheets, conservation laws establish the basic tenants of physics from which we derive our models. We first posit a general vectorized conservation equation for some arbitrary quantity ϕ

[1] Overtime, GCM gradually was interpreted as global coupled model as well, deviating slightly from its original meaning. We simply use GCM as an acronym with no critical need to distinguish between the two.

$$\frac{\partial \phi}{\partial t} = -\nabla \cdot (\mathbf{F} + \phi \mathbf{V}) + H \tag{2.2}$$

where $\mathbf{F}$ represents non-transport flux, $\phi\mathbf{V}$ represents transport flux, and H represents a source/sink of substance ϕ. Equation 2.2 describes an inertial frame in which ϕ is conserved within a volume V defined by surface S. In most three dimensional climate models, V and S are defined by the choice of spatial grid, while ϕ typically takes on the meaning of mass, energy, or momentum.

Mass: When dealing with mass, ϕ becomes density, typically denoted ρ. We recover a general conservation of mass equation by recognizing that mass can neither be created nor destroyed ($H = 0$) and the only flux of interest is the transport flux ($\mathbf{F} = 0$).

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \rho \mathbf{V} \tag{2.3}$$

Energy: For conservation of energy, we typically focus on heat per unit volume. ϕ becomes $\rho c_p T$, where c_p is specific heat and T is temperature. We have both conductive and transport flux in this case, which requires the introduction of a new variable, k, that corresponds to thermal conductivity. No longer assuming a zero-valued H, as we allow for heat sinks/sources, we arrive at

$$\frac{\partial \rho c_p T}{\partial t} = -\nabla \cdot (-k\nabla T + \rho c_p T \mathbf{V}) + H \tag{2.4}$$

Momentum: Lastly, turning towards momentum (per unit volume), ϕ becomes $\rho\mathbf{V}$. In terms of flux, we consider here advection and stress, the latter of which we denote using σ. External forces (e.g. gravity) can result in a nonzero H. For the case in which we are only concerned with gravity, we achieve

$$\frac{\partial \rho \mathbf{V}}{\partial t} = -\nabla \cdot (-\sigma + \rho \mathbf{V}\mathbf{V}) + \rho g \tag{2.5}$$

In practice, many of the assumptions that allow for the above conservation laws are broken by a number of factors, including but not limited to properties of compressible/non-compressible fluids, phase changes, external forces, and grid cells at critical boundaries (e.g. a volume defined over the ocean-atmosphere boundary). As a result, the generalized formulations simply serve as the backbone from which more specific governing equations can be derived, with the Navier-Stokes equations serving as a prime example.

Preliminary 3: Navier-Stokes Equations

Though a number of important relationships can be derived from conservation laws, we would be remiss not to mention arguably the most important of those: the famous Navier-Stokes equations, which arise from applying the conservation of momentum to Newtonian fluids. In vectorized form, we write a concise formulation for an incompressible fluid with constant viscosity as follows:

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} = -\frac{1}{\rho}\nabla p + \nu \nabla^2 \vec{v} + f \tag{2.6}$$

Here we relate change in the fluid velocity vector ($\vec{v}$) to advection ($\vec{v} \cdot \nabla$), kinematic viscosity (ν), pressure (p), and some collection of external forces f, usually taken to be Coriolis forces and/or gravity. The Navier-Stokes equations are critical to understanding the flow of viscous fluids[2], making them essential to understanding atmospheric/oceanic circulatory processes, glacial flow, and the behavior of substances that exhibit flow at relevant boundaries (e.g. land-atmosphere, atmosphere-ocean, ice sheet - ocean boundaries) among other applications. Excluding degenerate cases, the Navier-Stokes equations generally do not have exact solutions, so we rely on numerical techniques (e.g. finite element, spectral, or operator splitting methods) instead.

Additional Considerations

While Preliminaries 1-3 are undoubtedly necessary to lay the groundwork for subsequent discussions of climate modeling, they paint far from the whole picture. Atmospheric models don't merely characterize fluid motion–they must also consider (among other things) chemical aerosols, radiative absorption/emission/scattering, and anthropogenic greenhouse gases. Ocean models must similarly account for biogeochemical cycles, heat transfer, and thermohaline circulation, while ice-sheet models are inextricably linked to boundary processes associated with stress, strain, and thermodynamic balance. Even land surface models, which at first glance might appear the most straightforward, must include dynamic vegetation, carbon sequestration, and land-use changes to approach a comprehensive model that accounts for both human and nature-induced dynamics [17, 22, 31].

At the global scale, modeling for any one of these processes incurs great computational expense; when considering them all in tandem, there is therefore a tradeoff

[2] Accounting for viscosity is one of the main distinguishing factors between the Navier-Stokes and Euler equations, though we acknowledge that the Navier-Stokes can be appropriately modified to model inviscid flow as well.

between computational cost and resolution. One way global coupled models can modulate this balance is by manipulating the size of the spatial grid cells; increasing cell-size can reduce computational cost, but can simultaneously lose the granularity needed to study sub-processes or sub-regions. From the sub-process perspective, many physical phenomena that occur at scales smaller than that defined by a grid cell must be approximated (i.e. *parameterized*) in order to be incorporated into climate models. When grid cells are so large as to encompass entire island nations or areas of significant import, we thus lose the ability to characterize regional climate and weather patterns [2]. As a result, we turn towards downscaling to develop higher resolution models for specific regions of the world.

2.2 Regional Climate Models

Regional climate models (RCMs) are generated using largely the same concepts as global coupled models, but with a smaller area of interest. We can feed in information from global climate models to boundary conditions, while reserving the majority of our computational power for higher resolution modeling within the target area. This is known as dynamical downscaling in contrast to statistical downscaling [41], which involves characterizing the statistical relationship between global and regional observations in order to create regional climate projections. Dynamical and statistical downscaling do not stand in direct opposition to one another, rather they are more accurately conceptualized as complementary approaches to the same problem. A more rigorous treatment of global coupled models, RCMs, and weather forecasting is given in [41].

2.3 Coupled Model Intercomparison Project

Having introduced the fundamental concepts, we now turn towards climate modeling in practice. In 1995, the World Climate Research Program came together to standardize the process for climate model evaluation. Most evaluation metrics focus on hindcasting, wherein we compare historically observed data with model predictions– an intuitive process designed to align model expectations with what is/has been observed. The Coupled Model Intercomparison Project (CMIP) thus seeks to create a centralized repository for climate modelers around the world to establish and open-source benchmark results in a standardized fashion to facilitate better cross-model comparison. With its seventh edition currently in development, we rely on the sixth edition (CMIP6) as our reference for state-of-the-art global climate models. While many CMIP6 models agree on a wide-variety of metrics, they often do not agree when it comes to rainfall projections [2], both amongst themselves and with observed data. As discussed in Chapter 1, this can prove detrimental to rain-dependent regions that would benefit greatly from reliable precipitation prediction. We further explore one such region–East Africa–in Chapter 3, transitioning this discussion to an in-depth commentary on the East African Climate Paradox.

Chapter 3

Main Results

The East African Climate Paradox more pointedly refers to the failure of the masika rains between March and May, the effects of which are worsened by variable vuli rains from October to December. We first characterize the discrepancy between modeled and observed rainfall before exploring leading explanatory theories for this paradox.

3.1 The Paradox

The East African Paradox has been called as such since at least phase 3 of the CMIP project. As shown in [43], not only did CMIP3 models predict higher mean precipitation rates of both masika and vuli rains, they concurrently suggested less intense droughts. The CORDEX project, a collection of 10 RCM models, similarly overestimated rainfall in portions of East Africa (e.g. Ethiopia and the Congo basin) with eight out of ten of the included models showing wet bias [18].

Moving from CMIP3 to more modern CMIP editions, it was further shown in [36, 35] that CMIP5 models failed to capture the peak of the masika rains while simultaneously indicating an expected moisture influx. Ongoma et al. [33, 34] likewise concludes that CMIP5 models point towards an increase in both extreme rainfall events and mean precipitation during both masika and vuli. While a growing number of studies are evaluating the degree to which CMIP6 models contribute towards resolving the paradox, early results are mixed. A handful of studies show CMIP6 outperforms CMIP5 across a broad range of rainfall indices [10, 7], yet it has also been shown that CMIP6 overpredicts masika rains while failing to capture the interannual variability of the vuli rains [9, 30, 42].

In light of these established findings, we now attempt to understand why the East African Climate Paradox exists beyond its descriptive characterization.

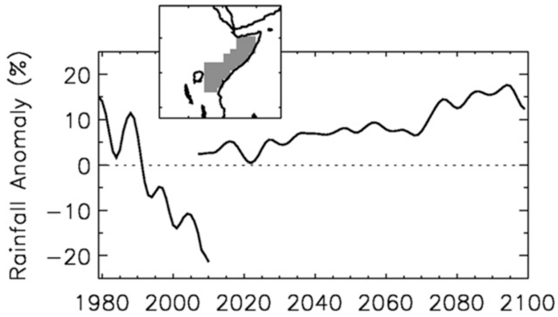

Figure 3.1: Rainfall anomalies from observations (left line) and model projections (right line). Averaged over CMIP5 models under the RCP8.5 scenario relative to 1901-2000 historical runs. Adapted from [40].

3.2 Leading Theories

There is general agreement as to which processes are linked to East African rainfall, allowing us to identify sensible starting points in explaining rainfall discrepancies. Among such starting points are variabilities in sea surface temperatures (SSTs) of the Indian Ocean, the shifts of the intertropical convergence zone (ICTZ), Walker and Hadley cells, the El Niño Southern Oscillation (ENSO), and fluctuating jet streams (e.g. the Kenyan Turkana jet stream) [16]. The question we seek to resolve is not necessarily how these influence rainfall, but rather what is driving change in these underlying causal factors and, importantly, how can we better account for it?

3.2.1 Anthropogenic Forces

The first theory worthy of discussion is that anthropogenic forcings are contributing to changes in underlying causal factors. We focus here on two key anthropogenic drivers of change: emissions (i.e. carbon and aerosols) and land-use. In either case, the picture is unclear. While some studies explicitly leave anthropogenic explanatory mechanisms as an open-question [29, 46], those that provide more in-depth analysis call for further research after arriving at mixed results [37, 40]. Namely, Rowell et al. [40] analyzes simulations isolating the effect of anthropogenic aerosol emissions. By comparing a subset of CMIP5 models fed with fixed preindustrial aerosol levels (CSIRO Mk3.6.0, HadGEM2-ES, and IPSL-CM5A-LR) against models that include historically accurate aerosol levels (Can-ESM2, CCSM4, CSIRO, Mk3.6.0, GFDL, CM3, and GISS-E2-R), no statistically significant impact of anthropogenic aerosols emerges above the naturally expected variations. Despite this, existing research on aerosol driven SST-variability suggests that perhaps the signals attributable to aerosol levels are not adequately propagated throughout the model, resulting in dampened observations of statistical significance [40] . Similar conclusions are drawn for land-use effects as well–though the underlying hypothesis may merit further study, particularly given large-scale expected land-use shifts in the Lake Victoria and Congo basins, there is no threshold for statistical significance reached by experiments isolating variable land-use [40].

Even for greenhouse gas emissions, perhaps the most frequently discussed anthropogenic force in this context, we face both a lack of dedicated research studies attributing carbon emissions to East African rainfall and a lack of robust evidence supporting/rejecting hypotheses in those studies that do exist. It is well known that greenhouse gases affect some of the underlying causal factors of rainfall variability, such as the southern shift of the ITCZ [13] and Indian Ocean Dipole [11], but all papers reviewed universally cite the downstream effects on East African rainfall as an open problem.

3.2.2 Natural Variability

The natural variability theory posits that the observed droughts and changes in masika/vuli rains are explained by the natural fluctuation and progression of the East African climate system. In the absence of external forcings (both anthropogenic and natural), the rainy season changes observed since the 1990s [40] would be unsurprising if natural variability is the main mechanism of change. To some degree, this notion is challenging to evaluate–even if the variability in an unforced climate model allows for observed rainfall trends, we exist in a reality where external forcings have, do, and will continue to exist. It is a theory nonetheless worth exploring, and, accordingly, Rowell et al. [40] demonstrated that hindcasted rainfall shifts are not statistically significantly explained solely by natural variability. This result, however, was predicated on the notion of reliable decadal models of variability, an assumption for which a number of studies have provided contrasting results.

For instance, several studies [11, 27, 37, 47] have attempted to categorize 21st century drought patterns in the context of both historical and projected droughts, as there is precedent for extended drought periods in East Africa in unforced climate models. During an era in which unforced climate models would be appropriate due to the negligible levels of both volcanic and anthropogenic emissions, there are at least two documented cases of multi-decade drought patterns (1821-1835, 1879-1902) [19, 40]. This further supports the idea that natural variability could, at least in part, be responsible for explaining recent drought patterns; however, this does not resolve the question as to why current models are unable to identify such patterns.

Following the approach from [39], we assume the variability in climate projections is comprised of both natural variability (σ_n^2) and inherent model variability (σ_m^2). In evaluating the natural variability theory, we can thus compare the fraction of the total variability (σ_{tot}^2) that is explained by σ_n^2 after propagating uncertainty throughout our model. Rowell et al. [40] does exactly as such, showing that long-term rainfall projections ($\frac{\sigma_{\text{tot}}^2}{\sigma_n^2} = 9.9$ for 2070-2099) are far less attributable to natural variability than short-term projections ($\frac{\sigma_{\text{tot}}^2}{\sigma_n^2} = 4.2$ and 2.2 for 2035-2064 and 2015-2044 respectively). This brings into question the results regarding current observations, but supports the larger idea that natural variability alone is not enough to explain the discrepancies in the long run.

3.2.3 Unreliable Models

The final and most consequential theory questions the reliability of the models used to forecast East African rainfall. There is a fundamental difference between asserting that the models need to be improved by incorporating more sophisticated representations of the complex drivers of rainfall (e.g. Indian Ocean Dipole, anthropogenic forcings, SST variations) and asserting that the models are unreliable. The latter is straightforwardly rejected by Rowell et al. [40], who instead argues that no claim of unreliability can be made before systematically and rigorously understanding all drivers of bias and uncertainty that constrain model projections. Given that different CMIP models not only tend to produce inter-compatible (albeit erroneous) results when it comes to East African rainfall projections, but also perform well on other, non-East African related applications, there is strong evidence that modern models are not fundamentally unreliable. Rather, there is a seemingly universal call for further research–to improve bias, variability, and uncertainty traces; to increase the sophistication of subprocess modeling; and to enhance our understanding of anthropogenic drivers of change.

Chapter 4

Discussion

We now suggest promising avenues of research to expound upon and potentially resolve the open problems identified by the existing literature. At a broad-level, we seek to increase resolution, improve our submodel components, or better account for external forcings. Our dominant critique of current work is that though every study reviewed encourages further research, all remain relatively ambiguous as to how novel approaches should be designed to achieve better, higher-performing models. Here we situate ourselves on the other end of the ambiguity spectrum by offering specific and concrete suggestions to resolve unsolved problems.

4.1 AI-enabled Climate Models

4.1.1 Computational Efficiency

As outlined in [14], one way in which neural networks can aid climate/weather forecasting is in freeing additional compute to be dedicated towards higher resolution modeling. This could prove particularly apt in improving parameterization schemes from a computational efficiency perspective, a proof of concept for which has already been established by [12, 15].

In particular, cloud fields and cloud formations are often parameterized, yet are key precursors for precipitation to occur. We thus recommend improving computational efficiency of cloud parameterization via neural networks, following an approach similar to [20] in which a neural network maps temperature, humidity, pressure, and ice mixing ratio variables to fine-resolution cloud properties (the cloud cover, the liquid condensate mixing ratio, and the ice condensate mixing ratio). To the best of our knowledge, downscaled CMIP5 and CMIP6 models, as well as dedicated CORDEX

RCM models, do not implement a neural network approach to cloud parameterization. The saved computational cost (which could even be boosted via neural pruning) could thus be "reinvested" into modeling the (suspected) more complex drivers of East African rainfall variability.

4.1.2 Physics Informed Neural Networks

As an addendum to the previous section, it is worth specifically citing Physics Informed Neural Networks (PINNs) as a high potential approach to incorporating machine learning into climate modeling. By using bespoke loss functions or imposing physical-constraint boundaries on network parameters, PINNs seek to model real-world phenomena while both reaping the benefits of deep learning and obeying laws of physics. For instance, we might design a loss function

$$\mathcal{L}_{tot} = \mathcal{L}_{NN} + \lambda \mathcal{L}_P$$

where $\mathcal{L}_{NN}$ represents a standard neural network loss function (e.g. MSE, Cross-Entropy, Log Loss), $\mathcal{L}_P$ represents a physics-inspired loss function, and λ represents a scaling parameter to adjust the relative importance of $\mathcal{L}_{NN}$ and $\mathcal{L}_P$. While the former evaluates how well a model predicts data within a training set, the latter evaluates how well a model's predictions conform to the laws of physics. As we minimize this hybrid loss function, we can push a model to learn a set of parameters that is consistent with physical constraints. A potential design for $\mathcal{L}_P$ is shown in [21] when applied to cloud-precipitation reactions. Much is left to be done in this space; however, preliminary results demonstrate promise in aiding uncertainty and bias analyses [12, 23]. Given uncertainty/bias attribution is one of the primary challenges cited in resolving the natural variability theory, we suggest PINNs as an initial exploration path to deepen our understanding of the uncertainty/bias encapsulated by East African climate projections.

4.2 Teleconnection Network Approach

To the best of our knowledge, a network approach has not been implemented to resolve issues surrounding the East African Climate Paradox. We outline here a framework for identifying potentially obscured teleconnections that impact regional rainfall via community detection and modularity analysis in a network-embedded climate system.

We define a network as a collection of nodes (vertices) and edges, wherein an edge e_{ij} exists if some domain-specific relationship exists between nodes i and j. A useful analytical framework in network science is community detection, which seeks to partition a network into distinct communities. Under a good partition, nodes within a given community are more densely connected to each other and sparsely connected to members of other communities. The performance of a given partition is usually empirically measured via a modularity score Q, a popular definition for which is as follows:

$$Q = \frac{1}{2m} \sum_{i,j} (A_{ij} - \frac{k_i k_j}{2m}) \delta_{c(i)c(j)} \tag{4.1}$$

where A denotes the adjacency matrix, $k_i = \sum_j A_{ij}$ denotes the degree of node i, m denotes the number of edges, and $c(i)$ denotes the community of node i in the Kronecker's delta term. We thus have an optimization problem: we seek to maximize Q over the set of possible partitions in order to identify distinct communities within a network.

Returning to climate modeling, early work has sought to apply network science techniques to climate networks [45]. We recommend generating a climate network following the approach in [44]. Nodes represent gridded climate data, with corresponding node states $\mathbf{w}_i \in \mathbb{R}^m$, where m indicates the number of variables of interest encoded by a grid cell. Following from [44], we recommend, at minimum, $m = 7$, to capture sea surface temperature, sea level pressure, geopotential height, precipitable water, relative humidity, and horizontal/vertical wind speed. Node states are then de-seasonalized to account for the inherent periodicity in many climate signals. We finally construct edges based on correlations between grid cells over the relevant time windows[1]. Here we deviate slightly from the recommendation in [44]; rather than linking two cells with a weighted edge if their correlation exceeds some threshold as measured by a linear (Pearson's) correlation coefficient, we instead recommend using a correlation measure that captures both linear and nonlinear relationships. Namely, we suggest computing the distance correlation, defined as

$$d_c^2(X, Y) = \frac{\mathrm{dCov}^2(X, Y)}{\sqrt{\mathrm{dVar}^2(X)\mathrm{dVar}^2(Y)}}$$

where $\mathrm{dCov}^2(X, Y)$ represents the distance covariance of random variables X and Y, while $\mathrm{dVar}^2(X)$ represents the distance variance. We can express distance covariance

[1]It is worth noting that this method of edge creation does not encapsulate physical space in its structure. In other words, neighboring grid-cells in physical space might not be neighbors in the generated network

in terms of the traditional Pearson covariance (pCov), while distance variance is seen as a special case of covariance in which X and Y are identical. Here we use (X', Y') and (X'', Y'') to indicate i.i.d copies of X and Y.

$$\mathrm{dCov}^2(X, Y) = \mathrm{pCov}(\|X - X'\|, \|Y - Y'\|) - 2\mathrm{pCov}(\|X - X'\|, \|Y - Y''\|)$$

We justify the choice of the distance correlation measure over other nonlinear correlation measures based on its robustness to noise and inherent normalization. The choice of a nonlinear correlation measure in general is justified due to the nonlinear relationships in climate models that may arise naturally from statistical physics. Importantly, the distance correlation still captures the linear relationships otherwise identified by Pearson's correlation coefficient.

Once an appropriate climate network has been generated, we suggest community detection (via one of the many community detection algorithms, e.g. the Louvain method) as a method for identifying potentially novel teleconnections between network communities (i.e. between regional East African cells and other regional/global patterns). Similar work has already been conducted for other regions, showing some promise in, at minimum, bolstering support for already established teleconnections. By identifying collections of highly correlated grid cells, we may uncover latent organizational structure in community-community interactions. We hypothesize that applying this community detection framework could further our understanding of the causal factors of rainfall variability in East Africa.

4.3 Turkana Jetstream

As a final point of recommendation, we find the emerging body of research characterizing the Turkana jetstream as particularly compelling. First identified in 1982 [25], the Turkana jetstream cuts across the Ethiopian highlands into the East African highlands of Kenya and Tanzania. Though it is the principle vehicle for moisture flux in the region (and therefore inherently linked to both arid and rainy seasons), it is relatively understudied. In fact, CMIP5 and CMIP6 models are unable to explicitly account for the low level jetstream (LLJ) through the Turkana channel; RCMs (i.e. CORDEX models) similarly suffer from nonstandard representations of LLJs [24].

Only in the last five years have we begun to develop a better understanding of the Turkana jetstream on daily, annual, and decadal timescales. Explained in more-depth by Munday et al. [32], we know from a dedicated field campaign (RIFTJet) in 2021 that the jetstream is diurnal, strengthening during the night. On annual scales, the LLJ through the Turkana channel peaks in boreal spring and autumn, closely aligning with masika and vuli. Most models identify an annual peak in March [32]. Notably, a statistically significant negative correlation (via Pearson's correlation coefficient) has been identified between LLJ strength and precipitation in a large portion of East Africa (i.e. Tanzania, Kenya, and southern Ethiopia) [16, 25]. On the decadal timescale, it is unclear whether broader trends are affecting the Turkana jetstream. While two of the three reanalyses conducted in [24] indicate a weakening of the Turkana over recent decades, results from the RIFTJet campaign indicate that the same models could underestimate its strength by as much as 75% [32].

The Turkana jetstream is thus an intriguing phenomenon worthy of additional study, potentially unveiling crucial insights in resolving the East African Climate Paradox.

Chapter 5

Conclusion

Knowing the humanitarian implications of unexpected rainfall variability in East Africa, high performance climate modeling is an imperative to enhance mitigation and prevention strategies. Given the continued failure of the masika rains despite a predicted increase in intensity and duration, there is a pressing need to resolve the East African Climate Paradox. Our work provides an overview of the current paradox while simultaneously identifying promising research avenues grounded in novel application of mathematical theory (e.g. deep learning, complex networks) and deeper probes into understudied physical phenomena (e.g. the Turkana jetstream). Undoubtedly, it is crucial to continue researching this paradox, as its resolution could prove beneficial to millions that call East Africa home.

Bibliography

[1] Agriculture & Food Security. URL: https://www.eac.int/agriculture.

[2] Q&A: How do climate models work? - Carbon Brief. URL: https://www.carbonbrief.org/qa-how-do-climate-models-work/.

[3] UNICEF helps avert a famine in Somalia in 2017 | UNICEF Somalia. URL: https://www.unicef.org/somalia/stories/unicef-helps-avert-famine-somalia-2017.

[4] United Nations Statistics Division- Standard Country and Area Codes Classifications (M49), July 2011. URL: https://web.archive.org/web/20110713041240/http://millenniumindicators.un.org/unsd/methods/m49/m49regin.htm#africa.

[5] Why Somalia is facing a catastrophic famine | International Rescue Committee (IRC), July 2022. URL: https://www.rescue.org/article/why-somalia-facing-catastrophic-famine.

[6] What is drought? Causes, impact & countries most affected | International Rescue Committee (IRC), July 2023. URL: https://www.rescue.org/uk/article/what-drought-causes-impact-and-how-we-can-help.

[7] Akintomide Afolayan Akinsanola, Victor Ongoma, and Gabriel J. Kooperman. Evaluation of CMIP6 models in simulating the statistics of extreme precipitation over Eastern Africa. *Atmospheric Research*, 254:105509, June 2021. URL: https://www.sciencedirect.com/science/article/pii/S0169809521000612, doi:10.1016/j.atmosres.2021.105509.

[8] J. Araujo, J. Marsham, D. Rowell, N. Zinyengere, A. Ainslie, A. Clenaghan, R. Cornforth, G. De Giusti, B. Evans, D. Finney, D. Lapworth, D. MacDonald, C. Petty, J. Seaman, F. Semazzi, and C. Way. East Africa's Climate: Planning for an Uncertain Future. pages 11–17. Climate & Development Knowledge

Network, South Africa, November 2016. URL: https://cdkn.org/resource/africas-climate-helping-decision-makers-make-sense-climate-information/?loclang=en_gb.

[9] Brian Ayugi, Victor Dike, Hamida Ngoma, Hassen Babaousmail, Richard Mumo, and Victor Ongoma. Future Changes in Precipitation Extremes over East Africa Based on CMIP6 Models. *Water*, 13(17):2358, January 2021. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/2073-4441/13/17/2358, doi:10.3390/w13172358.

[10] Brian Ayugi, Jiang Zhihong, Huanhuan Zhu, Hamida Ngoma, Hassen Babaousmail, Karim Rizwan, and Victor Dike. Comparison of CMIP6 and CMIP5 models in simulating mean and extreme precipitation over East Africa. *International Journal of Climatology*, 41(15):6474–6496, 2021. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.7207. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.7207, doi:10.1002/joc.7207.

[11] M. T. Blau and K.-J. Ha. The Indian Ocean Dipole and its Impact on East African Short Rains in Two CMIP5 Historical Scenarios With and Without Anthropogenic Influence. *Journal of Geophysical Research: Atmospheres*, 125(16):e2020JD033121, 2020. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD033121. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020JD033121, doi:10.1029/2020JD033121.

[12] N. D. Brenowitz and C. S. Bretherton. Prognostic Validation of a Neural Network Unified Physics Parameterization. *Geophysical Research Letters*, 45(12):6289–6298, 2018. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018GL078510. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078510, doi:10.1029/2018GL078510.

[13] Stefan Brönnimann, Andreas M. Fischer, Eugene Rozanov, Paul Poli, Gilbert P. Compo, and Prashant D. Sardeshmukh. Southward shift of the northern tropical belt from 1945 to 1980. *Nature Geoscience*, 8(12):969–974, December 2015. URL: https://www.nature.com/articles/ngeo2568, doi:10.1038/ngeo2568.

[14] Matthew Chantry, Hannah Christensen, Peter Duben, and Tim Palmer. Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI. URL: https://royalsocietypublishing.org/doi/epdf/10.1098/rsta.2020.0083, doi:10.1098/rsta.2020.0083.

[15] F. Chevallier, F. Chéruy, N. A. Scott, and A. Chédin. A Neural Network Approach for a Fast and Accurate Computation of a Longwave Radiative Budget. *Journal of Applied Meteorology and Climatology*, 37(11):1385–1397, November 1998. Publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology. URL: https://journals.ametsoc.org/view/journals/apme/37/11/1520-0450_1998_037_1385_annafa_2.0.co_2.xml, doi:10.1175/1520-0450(1998)037<1385:ANNAFA>2.0.CO;2.

[16] Kerry H. Cook, Rory G. J. Fitzpatrick, Weiran Liu, and Edward K. Vizy. Seasonal asymmetry of equatorial East African rainfall projections: understanding differences between the response of the long rains and the short rains to increased greenhouse gases. *Climate Dynamics*, 55(7):1759–1777, October 2020. doi:10.1007/s00382-020-05350-y.

[17] John B. Drake. *Climate modeling for scientists and engineers*. Mathematical modeling and computation. Society for Industrial and Applied Mathematics SIAM Market Street, Floor 6, Philadelphia, PA 19104, Philadelphia, Pennsylvania, 2014.

[18] Hussen Seid Endris, Philip Omondi, Suman Jain, Christopher Lennard, Bruce Hewitson, Ladislaus Chang'a, J. L. Awange, Alessandro Dosio, Patrick Ketiem, Grigory Nikulin, Hans-Jürgen Panitz, Matthias Büchner, Frode Stordal, and Lukiya Tazalika. Assessment of the Performance of CORDEX Regional Climate Models in Simulating East African Rainfall. *Journal of Climate*, 26(21):8453–8475, November 2013. Publisher: American Meteorological Society Section: Journal of Climate. URL: https://journals.ametsoc.org/view/journals/clim/26/21/jcli-d-12-00708.1.xml, doi:10.1175/JCLI-D-12-00708.1.

[19] Chaochao Gao, Alan Robock, and Caspar Ammann. Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models. *Journal of Geophysical Research: Atmospheres*, 113(D23), 2008. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2008JD010239. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010239, doi:10.1029/2008JD010239.

[20] Brian Henn, Yakelyn R. Jauregui, Spencer K. Clark, Noah D. Brenowitz, Jeremy McGibbon, Oliver Watt-Meyer, Andrew G. Pauling, and Christopher S. Bretherton. A Machine Learning Parameterization of Clouds in a Coarse-Resolution Climate Model for Unbiased Radiation. *Journal of Advances in Modeling Earth Systems*, 16(3):e2023MS003949, 2024. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2023MS003949. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2023MS003949, doi:10.1029/2023MS003949.

[21] Alice V. Hu and Zbigniew J. Kabala. Predicting and Reconstructing Aerosol–Cloud–Precipitation Interactions with Physics-Informed Neural Networks. *Atmosphere*, 14(12):1798, December 2023. Number: 12 Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/2073-4433/14/12/1798, doi:10.3390/atmos14121798.

[22] H. G. Kaper. *Mathematics and climate.* Society for Industrial and Applied Mathematics SIAM Market Street, Floor 6, Philadelphia, PA 19104, Philadelphia, Pennsylvania, 2013.

[23] K. Kashinath, M. Mustafa, A. Albert, J-L. Wu, C. Jiang, S. Esmaeilzadeh, K. Azizzadenesheli, R. Wang, A. Chattopadhyay, A. Singh, A. Manepalli, D. Chirila, R. Yu, R. Walters, B. White, H. Xiao, H. A. Tchelepi, P. Marcus, A. Anandkumar, P. Hassanzadeh, and null Prabhat. Physics-informed machine learning: case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200093, February 2021. Publisher: Royal Society. URL: https://royalsocietypublishing.org/doi/full/10.1098/rsta.2020.0093, doi:10.1098/rsta.2020.0093.

[24] James A. King, Sebastian Engelstaedter, Richard Washington, and Callum Munday. Variability of the Turkana Low-Level Jet in Reanalysis and Models: Implications for Rainfall. *Journal of Geophysical Research: Atmospheres*, 126(10):e2020JD034154, 2021. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD034154. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020JD034154, doi:10.1029/2020JD034154.

[25] J. H. Kinuthia and G. C. Asnani. A Newly Found Jet in North Kenya (Turkana Channel). *Monthly Weather Review*, 110(11):1722–1728, November 1982. Publisher: American Meteorological Society Section: Monthly Weather Review. URL: https://journals.ametsoc.org/view/journals/mwre/110/11/1520-0493_1982_110_1722_anfjin_2_0_co_2.xml, doi:10.1175/1520-0493(1982)110<1722:ANFJIN>2.0.CO;2.

[26] Peter Hjort Lauritzen. Discretization strategies.

[27] Rachel L. Lupien. Past climate unravels the eastern African paradox. *Nature*, 620(7973):279–280, August 2023. Bandiera abtest: a Cg type: News And Views Publisher: Nature Publishing Group Subject term: Climate change, Climate sciences, Hydrology. URL: https://www.nature.com/articles/d41586-023-02297-y, doi:10.1038/d41586-023-02297-y.

[28] Bradfield Lyon. Seasonal Drought in the Greater Horn of Africa and Its Recent Increase during the March–May Long Rains. *Journal of Climate*, 27(21):7953–7975, November 2014. Publisher: American Meteorological Society Section: Journal of Climate. URL: https://journals.ametsoc.org/view/journals/clim/27/21/jcli-d-13-00459.1.xml, doi:10.1175/JCLI-D-13-00459.1.

[29] Bradfield Lyon and N. Vigaud. Unraveling East Africa's Climate Paradox: Patterns and Mechanisms. pages 265–281. June 2017. doi:10.1002/9781119068020.ch16.

[30] Dawido Magang, Moses Ojara, Lou Yunsheng, and Henry Philemon. Future climate projection across Tanzania under CMIP6 with High-Resolution Regional Climate Model, January 2024. URL: https://www.researchsquare.com/article/rs-3862295/v1, doi:10.21203/rs.3.rs-3862295/v1.

[31] Kendal McGuffie and Ann Henderson-Sellers. The Climate Modelling Primer. In *The Climate Modelling Primer*. John Wiley & Sons, Incorporated, United Kingdom, 2014.

[32] Callum Munday, Sebastian Engelstaedter, Gilbert Ouma, Geoffrey Ogutu, Daniel Olago, Dennis Ong'ech, Thomas Lees, Bonface Wanguba, Rose Nkatha, Clinton Ogolla, Roba Ali Gàlgalo, Abdi Jillo Dokata, Erick Kirui, Robert Hope, and Richard Washington. Observations of the Turkana Jet and the East African Dry Tropics: The RIFTJet Field Campaign. *Bulletin of the American*

Meteorological Society, 103(8):E1828–E1842, August 2022. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society. URL: https://journals.ametsoc.org/view/journals/bams/103/8/BAMS-D-21-0214.1.xml, doi:10.1175/BAMS-D-21-0214.1.

[33] Victor Ongoma, Haishan Chen, and Chujie Gao. Projected changes in mean rainfall and temperature over East Africa based on CMIP5 models. *International Journal of Climatology*, 38(3):1375–1392, 2018. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5252. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/joc.5252, doi:10.1002/joc.5252.

[34] Victor Ongoma, Haishan Chen, and George William Omony. Variability of extreme weather events over the equatorial East Africa, a case study of rainfall in Kenya and Uganda. *Theoretical and Applied Climatology*, 131(1):295–308, January 2018. doi:10.1007/s00704-016-1973-9.

[35] Vincent O. Otieno and R. O. Anyah. CMIP5 simulated climate conditions of the Greater Horn of Africa (GHA). Part II: projected climate. *Climate Dynamics*, 41(7):2099–2113, October 2013. doi:10.1007/s00382-013-1694-z.

[36] Vincent O. Otieno and Richard O. Anyah. CMIP5 simulated climate conditions of the Greater Horn of Africa (GHA). Part 1: contemporary climate. *Climate Dynamics*, 41(7):2081–2097, October 2013. doi:10.1007/s00382-012-1549-z.

[37] Paul I. Palmer, Caroline M. Wainwright, Bo Dong, Ross I. Maidment, Kevin G. Wheeler, Nicola Gedney, Jonathan E. Hickman, Nima Madani, Sonja S. Folwell, Gamal Abdo, Richard P. Allan, Emily C. L. Black, Liang Feng, Masilin Gudoshava, Keith Haines, Chris Huntingford, Mary Kilavi, Mark F. Lunt, Ahmed Shaaban, and Andrew G. Turner. Drivers and impacts of Eastern African rainfall variability. *Nature Reviews Earth & Environment*, 4(4):254–270, April 2023. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/s43017-023-00397-x, doi:10.1038/s43017-023-00397-x.

[38] Katy Richardson, Roger Calow, Florence Pichon, Stacey New, Rebecca Osborne, Richard Jones, Jane Strachan, Kirsty Lewis, Mandar Trivedi, and Leah Mwai. Climate risk report for the East Africa region. 2022.

[39] David P. Rowell. Sources of uncertainty in future changes in local precipitation. *Climate Dynamics*, 39(7):1929–1950, October 2012. doi:10.1007/s00382-011-1210-2.

[40] David P. Rowell, Ben B. B. Booth, Sharon E. Nicholson, and Peter Good. Reconciling Past and Future Rainfall Trends over East Africa. *Journal of Climate*, 28(24):9768–9788, December 2015. Publisher: American Meteorological Society Section: Journal of Climate. URL: https://journals.ametsoc.org/view/journals/clim/28/24/jcli-d-15-0140.1.xml, doi:10.1175/JCLI-D-15-0140.1.

[41] Markku Rummukainen. State-of-the-art with regional climate models. *WIREs Climate Change*, 1(1):82–96, January 2010. URL: https://wires.onlinelibrary.wiley.com/doi/10.1002/wcc.8, doi:10.1002/wcc.8.

[42] Kevin Schwarzwald and Richard Seager. Revisiting the East African Paradox: CMIP6 models also fail to simulate observed drying trends in the Horn of Africa Long Rains. Technical Report EGU24-4687, Copernicus Meetings, March 2024. Conference Name: EGU24. URL: https://meetingorganizer.copernicus.org/EGU24/EGU24-4687.html, doi:10.5194/egusphere-egu24-4687.

[43] Mxolisi E. Shongwe, Geert Jan van Oldenborgh, Bart van den Hurk, and Maarten van Aalst. Projected Changes in Mean and Extreme Precipitation in Africa under Global Warming. Part II: East Africa. *Journal of Climate*, 24(14):3718–3733, July 2011. Publisher: American Meteorological Society Section: Journal of Climate. URL: https://journals.ametsoc.org/view/journals/clim/24/14/2010jcli2883.1.xml, doi:10.1175/2010JCLI2883.1.

[44] Karsten Steinhaeuser, Nitesh V Chawla, and Auroop R Ganguly. COMPLEX NETWORKS IN CLIMATE SCIENCE: PROGRESS, OPPORTUNITIES AND CHALLENGES. *Conference on Intelligent Data Understanding*, 2010.

[45] A. A. Tsonis and P. J. Roebber. The architecture of the climate network. *Physica A: Statistical Mechanics and its Applications*, 333:497–504, February 2004. URL: https://www.sciencedirect.com/science/article/pii/S0378437103009646, doi:10.1016/j.physa.2003.10.045.

[46] Caroline M. Wainwright, John H. Marsham, Richard J. Keane, David P. Rowell, Declan L. Finney, Emily Black, and Richard P. Allan. 'Eastern African Paradox'

rainfall decline due to shorter not less intense Long Rains. *npj Climate and Atmospheric Science*, 2(1):1–9, September 2019. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/s41612-019-0091-7, doi:10.1038/s41612-019-0091-7.

[47] Dean P. Walker, John H. Marsham, Cathryn E. Birch, Adam A. Scaife, and Declan L. Finney. Common Mechanism for Interannual and Decadal Variability in the East African Long Rains. *Geophysical Research Letters*, 47(22):e2020GL089182, 2020. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL089182. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL089182, doi:10.1029/2020GL089182.