3-2022

The Impacts of Climate Uncertainty on Streamflow in Andes, Antioquia, Colombia

Kristen R. Roberts

THE IMPACTS OF CLIMATE UNCERTAINTY ON STREAMFLOW IN ANDES, ANTIOQUIA, COLOMBIA

THESIS

Kristen R. Roberts 2nd Lt, USAF

AFIT-ENV-MS-22-M-252

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-22-M-252

THE IMPACTS OF CLIMATE UNCERTAINTY ON STREAMFLOW IN ANDES, ANTIOQUIA, COLOMBIA

THESIS

Presented to the Faculty

Department of Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Engineering Management

Kristen R. Roberts

2nd Lt, USAF

March 2022

AFIT-ENV-MS-22-M-252

THE IMPACTS OF CLIMATE UNCERTAINTY ON STREAMFLOW IN ANDES, ANTIOQUIA, COLOMBIA

Kristen R. Roberts

2nd Lt, USAF

Committee Membership:

Lt Col Justin D. Delorit, PhD, PE
Chair

Dr. Thomas Phelan, PhD, PE
Member

Dr. Christopher Chini, PhD
Member

Abstract

Natural hazards, such as hurricanes, wildfires, floods, and droughts impact human systems that rely on predictable patterns in the natural elements with which they interact. These events threaten communities everywhere, and humanity continually seeks to adapt. Skillful prediction of the impacts of climate change on linked, human-natural systems, like surface water resources, can help ensure physical risks within vulnerable communities are mitigated, resource sustainability is maximized, and intersectoral markets continue to contribute to socio-economic stability. Due to water resources being a primary conduit through which climate uncertainty impacts people, economies, and ecosystems, its study is worthy of investigation; particularly, where those resources are uncertain and demanded by a variety of competitive users. This work evaluates a season-ahead statistical prediction model of growing season streamflow (September – December) in Andes, Antioquia, Colombia, against a suite of global and local predictor variables: precipitation, soil moisture, Niño 3.4 sea-surface temperature anomaly, and Southern Oscillation Index anomaly. Skillful results, which are defined as streamflow forecasts that outperform a specified climatological baseline, are produced for the models when analyzing extreme streamflow events ($r^2 = 0.77$, mean absolute percentage error = 21.87, ranked probability skill score = 0.21). Even a lean model, consisting of just Niño 3.4 as a predictor, produces skillful results ($r^2 = 0.37$, mean absolute percentage error = 21.98, ranked probability skill score = 0.087). Viewed cumulatively, these results suggest streamflow predictions and forecasts can identify the role of global and local climate on communities, inform how and when changes should be implemented, and justify stakeholder decisions.

For both of my families: Air Force and biological.

Acknowledgments

I want to start off by thanking my advisor for his mentorship and guidance while I have been a student at the Air Force Institute of Technology. Sir, you have demonstrated what good leadership looks like, and I hope that someday I will be able to be the type of leader you are. Next, I would like to thank Dr. S and Dr. P for letting me be a part of this research team and funding my research trips. Also, thank you for all the time and effort you have put in to develop me as a person and Air Force officer. Next, I would like to thank C.G for being my translator when I was working with community members. Your help really allowed me to conceptualize the issues surrounding water competition in the area and think of a way my model could be implemented to help members of the Andes community. Thank you, M.S. and J.V. for helping me with all of my random research questions and helping me find the answers to my questions in a timely manner. Having a team I could depend on made this entire experience a lot more enjoyable and greatly reduced my stress. Next, I would like to thank the United States Air Force for giving me this opportunity to get my master's degree. Lastly, I would like to thank my friends L.D., S.R., and G.C. for making my experience in the area better and giving me something to look forward to throughout this process.

This project is funded by the National Science Foundation (NSF) under Project Award No. NSF – 1743749. Thank you for funding my research.

Kristen R. Roberts

Table of Contents

List of Figures

List of Tables

Page

THE IMPACTS OF CLIMATE UNCERTAINTY ON STREAMFLOW IN ANDES, ANTIOQUIA, COLOMBIA.

I. Introduction

Background

Hydrology is one of a few natural processes through which climate directly impacts people, ecosystems, and economies (Sadoff and Muller, 2009). In fact, it was speculated that in 2012 over 780 million people globally, lack access to a dependable source of clean and safe water (Salaam-Blyther, 2012). Supply uncertainty is influenced by climate uncertainty, and communities due to their dependency on water resources, and water is likely to be increasingly less available as climate changes. Even under supply stress, demand continues to increase. Water is frequently demanded for use in many different sectors such as: municipal, agriculture, recreation, energy production, industry, etc.

Generally, the impacts of a changing climate are nonlinear, but have been observed across the globe, as illustrated by the intensification and increasing frequency of extreme hydrologic events such as hurricanes and other high-intensity rainfall events. Previous studies have sought to identify the indicators of a changing climate, recognize important climate thresholds, and apply adaptation strategies through intensive modeling simulations; however, few have had long-term success with the direct mitigation of climate uncertainty impacts within small communities (Smith et al., 2001; Fischer, 2002; Hitz, 2004; Johansen, 2017; Didier 2018; Cutter, 2020). Accurately predicting water availability at the community scale is one of the first steps toward adaptation planning as it pertains to preparing communities for the added stresses of long-term changes in weather patterns, and hydrologic regime shifts brought about by climate change. Water scarcity and increased demand will continue to threaten communities everywhere unless humanity adapts (Hellmuth, 2007).

Motivation

Since the turn of the century, there have been numerous instances where considerable tension existed between industry and residents over the quantity and quality of available water resources, particularly in underdeveloped countries (Hellmuth, 2007; Sadoff and Muller, 2009). Numerous studies have sought to remedy these issues through the application of adaption strategies that utilize intensive modeling simulations (Olmstead, 2014; Rod et al., 2020). However, few have had long-term success with the direct mitigation of climate uncertainty impacts within small communities. Despite complicated intersectoral demands, communities need tailored forecasts to provide decision-makers with suitable lead times to enact water supply, demand, and adaptation policies. The lead-time, which is embedded in forecasts, compliments proper decision-making time, and will only increase in value as climate-induced water resources uncertainty intensifies. A case study is developed to explore this problem, and further illustrate the value of information held in forecast leads. The community of Andes, Antioquia, Colombia is selected as there is an opportunity to test the applicability of statistically based forecasting modeling methods, and the development of water allocation policies across many competing agricultural, aquacultural, industrial, and municipal uses. A forecast-informed water policy, which does not exist in this current community, may reduce tensions between water users and ensure that the major economic uses of water, such as coffee farming and gold mining, remain successful through the changes brought forth by climate uncertainty. While this case study considers a small area, the community, water conflicts, and climate uncertainty are analogous to many locations, and as such, the forecast model constructed here could be easily adapted to any location.

Case Study: Andes, Antioquia, Colombia

The Andes community is a rural municipality of approximately 42,000 residents, located in the southwestern part of Antioquia, Colombia, South America (Información del Municipio, 2018) (Fig. 1). Even though the community itself is relatively small, water quality and quantity issues have the potential to impact nearly 4-million inhabitants downstream. Andes is located in the Aburrá Valley, which is one of the most populous in Colombia; this valley exists amongst the western part of the Andes Mountain range in rugged terrain (Bedoya, 2009). Due to its location in the high rainforests on the west end of the Amazon Basin, the human and natural communities of Andes are very sensitive to changes in climate. These changes in climate can include increased storm intensification, which can lead to a higher probability of landslide occurrence; or, decreased water availability during prolonged periods of drought.

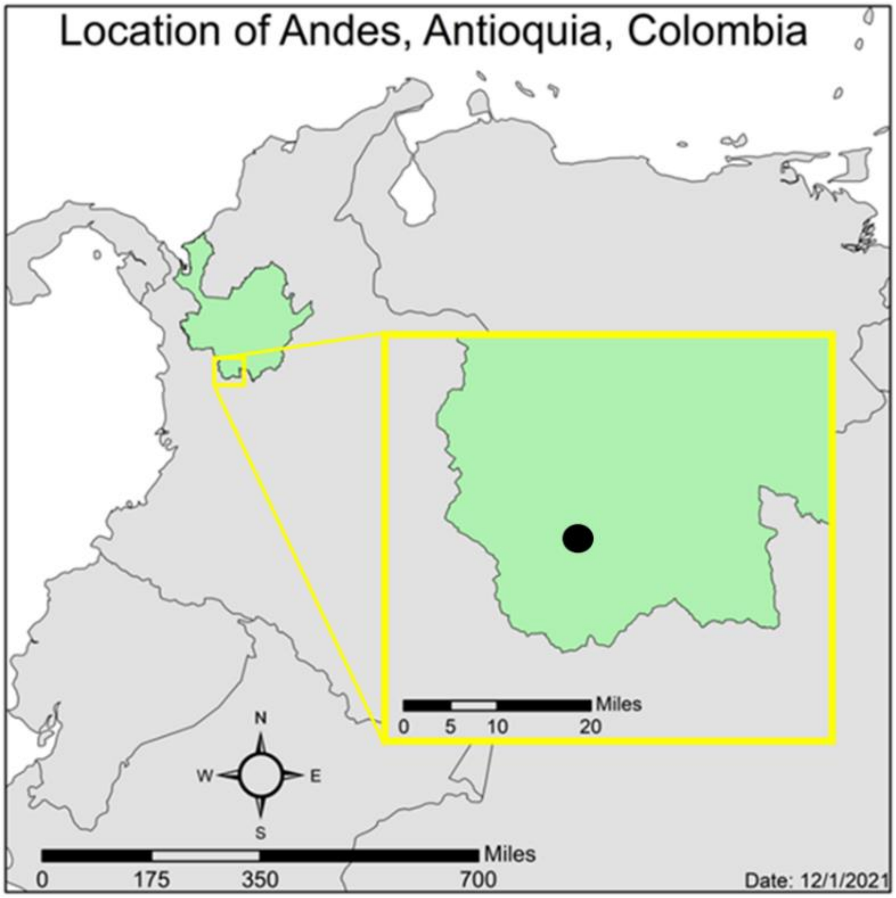

Figure 1: Geographical location of Andes, Antioquia, Colombia, South America

The main economic drivers in Andes: coffee farming, artisanal small-scale gold mining (ASGM), and fisheries are all affected by surface water availability, and could face additional risk as the climate continues to shift (Schwartz, 2021; Zapata Restrepo and Mejia Aramburo 2019; Salazar A. 2014; Información del Municipio, 2018). In Andes, coffee farmers make up approximately 90 percent of local economic activity and serve as the main source of dependable income for many families. ASGM makes up approximately 10 percent of the economy but has the perception of being responsible for most of the environmental pollution present in surface waters (O'Brien 2020). Return flows from ASGM are laden with heavy metals, the most concerning of which is mercury. The instream aquaculture (captive fish production) community is sparse, with only one local fishery being noted near the Andes municipality. Both coffee growers and aquaculturists view contaminated return flows from ASGM as destructive to their activities.

Coffee Production

Coffee is the second-most globally traded product, after oil (Davis et al. 2012). Growers contribute significantly to the socio-economic development of 120 million people in over 60 countries, in mostly equatorial regions (TCI 2016; Jayakumar et al. 2017). Generally, there are two distinct ways coffee is processed: wet or dry (Goto, 1956; Bosselmann et al., 2009). Depending on which of these methods is used to process the coffee, the characteristics of coffee waste such as biochemical oxygen demand, acidity, and nutrient content vary significantly. Unfortunately, coffee waste is not generally viewed as toxic by producers and is therefore commonly disposed of in nearby rivers or streams without regulation or regard for potential consequences (Beyene et al., 2012). Several studies suggest that untreated waste from coffee processing productions threatens surface waters around the world, and is the most severe in developing countries during harvesting seasons (Joshi and Sukumaran 1991; Beyene et al. 2012). Left untreated, these water quality indicators commonly lead to poor water conditions and depletion of aquatic life.

There are currently two distinct seasons of interest for coffee bean production in Antioquia, Colombia: March through April, and again between September and December (Fig 2). Locally, coffee wet-processing methods are used, which borrow from surface water. Coffee trees are predominantly rainfed and supplemented with stored irrigation when precipitation is insufficient to ensure suitable production.

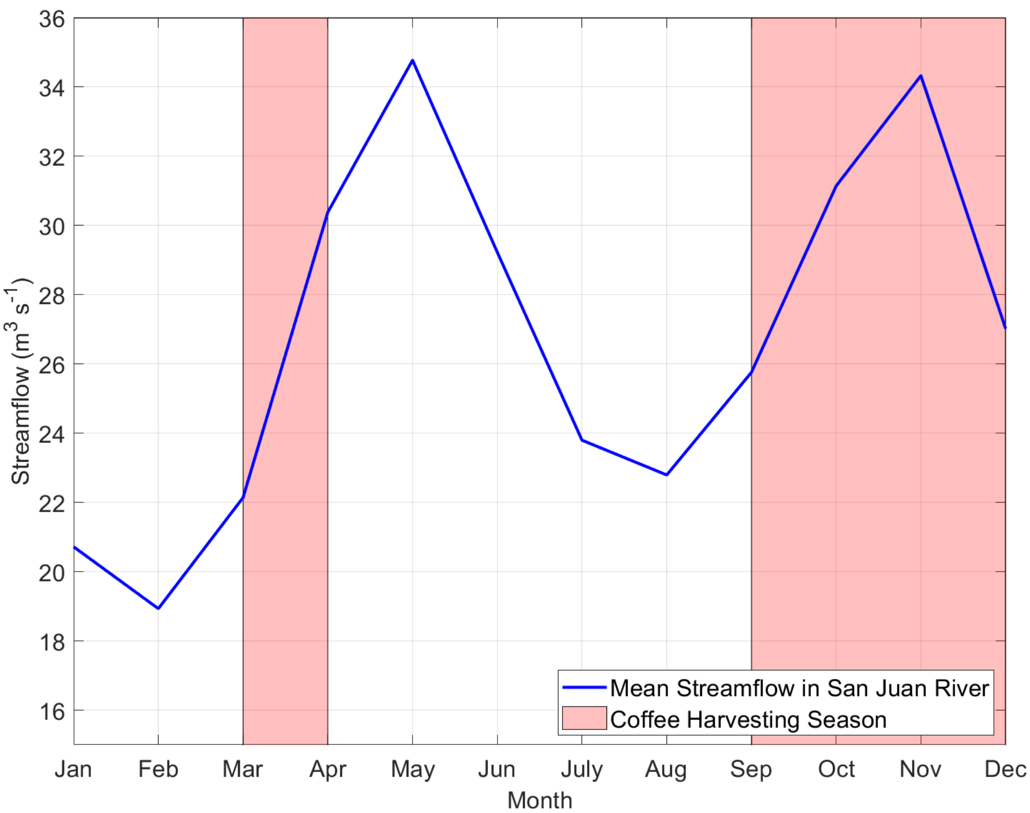

Figure 2: Identifying Seasons of Interest for Coffee Farming

During the growing seasons (February and June-September) and the harvesting seasons (March – April and September – December) farmers are very dependent on ideal temperature and precipitation conditions. Coffee farmers have expressed concern with noticeable changes in precipitation and temperature, citing longer seasons of high heat and drought, accompanied by high-intensity storm events (Frank, 2010). For this analysis, forecasts were created for the main growing season (September- December) due to its extreme importance on the local economy and the prolonged use of water from the San Juan River for wet-processing techniques.

Successful coffee production is highly sensitive to both local and global climate patterns, especially during the vegetative and reproductive phases of the plant (DaMatta and Ramalho 2006; Pham et al., 2019; Tavares et al., 2018). Shifts in temperature and precipitation surplus or shortages reduce the production of coffee by over 34% in many regions (Gay et al., 2006). Several other negative impacts, including loss of coffee-optimal areas throughout major global coffee-producing countries, growth in pest and disease, and the reduction of flowering, fruiting, and bean quality have been found (Lin, 2007). Clearly, irrigation could be used to meet crop-water demand, but only if sufficient surface water supply and quality are provided.

Artisanal Gold Mining

Although Andes is well-known for its production of coffee, it also has a long history of gold mining, which has remained largely artisanal since 1852 (Schwartz, 2021). As mentioned above, ASGM makes up approximately ten percent of local economic activity, but is widely regarded as the largest contributor to water pollution in the area, and has gained international attention as the largest human-caused source of mercury pollution in the world, credited with releasing nearly 1,220 metric tons of mercury to both water and land in 2015 alone (UN Environment, 2019). ASGM is also responsible for the release of arsenic, copper, lead, cadmium, and a wide array of additional inorganic contaminants into water, soil, sediment, and air (Schwartz, 2021; Gottesfeld et al., 2019; Rajaee et al., 2015; Tirima et al., 2016). The release of these contaminants into both soil and waterways makes ASGM a significant source of environmental contamination and is known to contribute to major health risks for locals and downstream communities. ASGM in Andes is conducted by individual miners or small enterprises with limited production capability. In the past, local miners commonly used mercury to extract gold from the ore and disposed of used mercury in local waterways; however, miners have actively worked to remove mercury from their current gold-extraction practices since Colombia's 2018 law which banned mercury use in mining activities (O'Brien, 2021). Throughout the Andes region, there are both active and abandoned mines scattered throughout the mountains.

Fisheries

Fisheries make up a very small portion of the economy in Andes, with only one local fishery noted upstream of the city center. Even as a minor use of water from the San Juan River, due to the off-river tanks' water levels being maintained with river water, it is an area of concern. Climate uncertainty has been predicted to negatively impact fresh-water fisheries due to changes in temperature, water availability, and streamflow variability (Ficke et al., 2007; Allison and Edward, 2009). If climate change leads to changes in freshwater habitats, low-income populations especially in Latin America can be regionally impacted by loss of food security (Harrod et al., 2019).

Municipal Water Use of San Juan River

The municipality of Andes uses centralized running water drawn from the San Juan River, while rural Andes populations tend to use rainwater collection systems as a source of drinking water. The municipality's supply system is reportedly treated, however the water quality in the San Juan River is known to be very poor and is avoided by concerned members of the community whenever possible (O'Brien, 2021). The municipal uses of water in Andes include cleaning, cooking, bathing, and watering vegetation around the home.

Interconnectedness of Water Use

The aforementioned water uses are largely in conflict, and there are anecdotal examples of residents interacting in multiple sectoral uses (O'Brien, 2021). Each of these communities are interconnected and depend upon appropriate water availability and quality to survive. Not only are all communities dependent upon appropriate water resources, but they all impact each other's use of water in some sense and are responsible for its quality (Fig. 3).

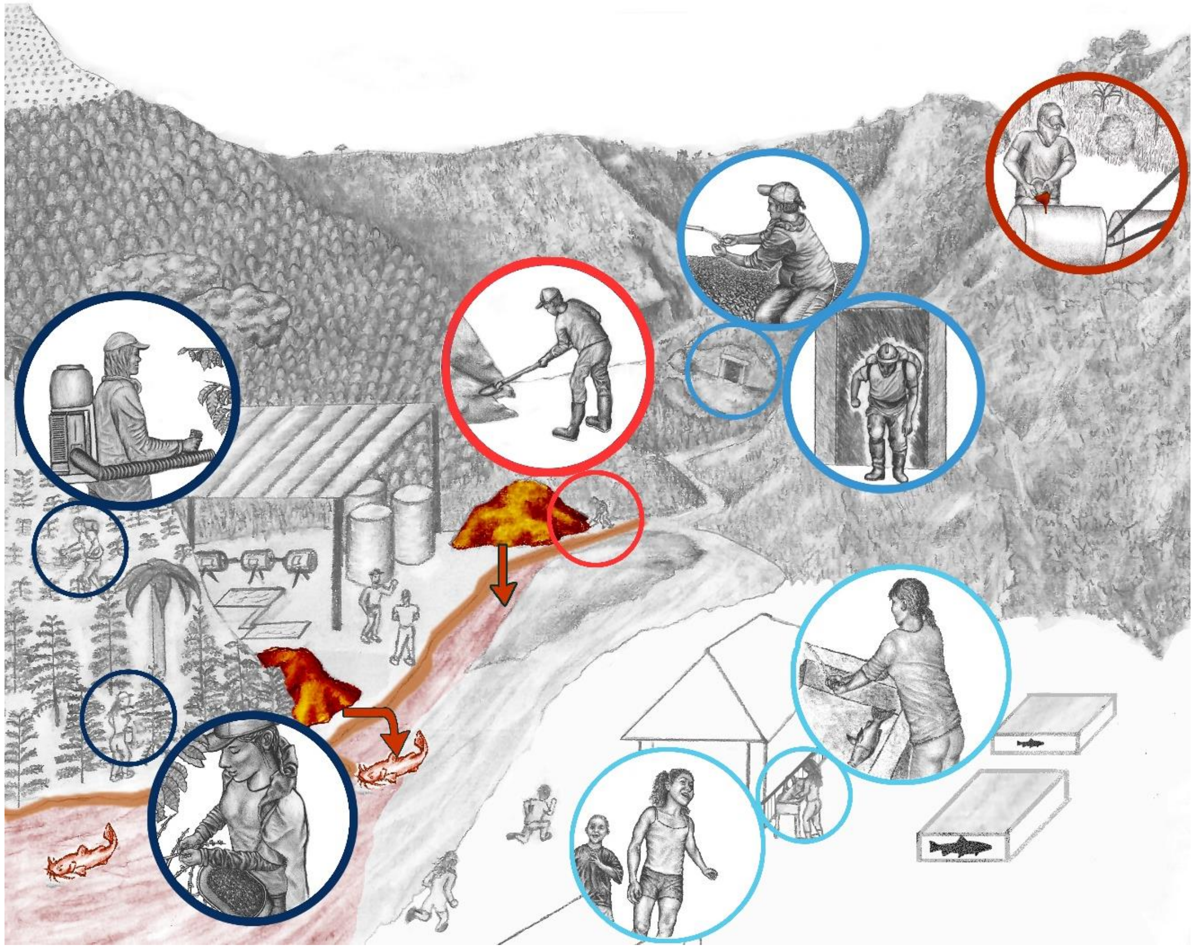

Figure 3: Visual of Community Interconnectedness in Andes (Adapted from O'Brien 2021, used with permission). Due to River networks in the area being a common topic of discussion and a central mechanism for mercury contamination, it is highlighted in red along with old tailings piles. The diversity of stakeholder groups is represented by varying shades of blue, except for two groups: processing plant workers and covert processors, shown in red due to their direct contact with mercury.

Clearly, the two uses that hold the potential impact water quality most are coffee and gold mining industries. Without proper pretreatment, gold mining waste pollutes waterways and has caused coffee farmers to avoid it as a source of potential supplemental irrigation for their agriculture operations, which is particularly important during times of low flows tied to drought. Gold mining operations also affect local fisheries and community members by contributing to the bioaccumulation of mercury in both fish and humans. Coffee farming operations impact the environment by depleting oxygen with polluted return flows. This has potential impact on both the gold mining industry, local fishery, and community members. Improper disposal of coffee waste can change essential water quality characteristics, which could have an impact on the quality of gold able to be collected by miners, reduce the amount of healthy fish populations at the local fishery, and make consumers sick. Without intervention, these water quality issues will cause environmental and health risks to the global community (Esdaile and Chalker, 2018).

To reduce long-term water availability risk, the Andes community should prepare a water policy to ensure risk is properly mitigated. There is not currently any type of water rights policy implemented within the community, but the policy should consider coffee farming, gold mining, aquaculture, and individual community member operations in its development. To properly govern and manage water in a developing economy, it is important that policy is founded upon proper modeling and data collection. Seasonal forecasts answer this call and have the potential to be very useful for agricultural activities and serve as a good starting point for early warning and response planning within a community (Johansen et al., 2017; Hellmuth, 2007). Statistical and probabilistic forecast methods have been used to provide deterministic and categorical forecast information that can be calibrated to suit decision-maker needs. The challenge is to incorporate such forecast information, with its explicit uncertainties, into decision-making (Hellmuth, 2007). And even skillful forecasts must be tailored such that their delivery suits decision-makers. Very often, forecasts are translated from a raw value, e.g., streamflow, to an economic output using sectoral models.

The Andes community presents an excellent opportunity to test the development of a forecast that could promote the long-term sustainability of water resources through climate change. Each of the communities within Andes has unique requirements with respect to water availability and quality. The sources of pollution pose a great threat in developing countries and should be remediated efficiently and quickly. According to local sources, there is a notable conflict between gold miners, coffee farmers, local fisheries, and community members when it comes to water use and allocation. Coffee growers are highly concerned about water quality issues and face growing concerns with respect to how changing climate could affect water availability.

Water availability could be impacted in two distinct ways due to climate uncertainty: scarcity or surplus. The discussion up to this point has focused on water scarcity conflicts in the area of study. However, in cases of extreme water surplus, this research could be used to develop a forecast with appropriate leads to aid in landslide mitigation planning for local stakeholders.

Research Question

The aforementioned literature and local context illustrate the need for seasonal management of water resources, where water availability and quality, and competition have the potential to stress supply, and therefore economic activity. This research focuses on one research questions as a means of understanding the unique climate uncertainty conditions in Andes, Antioquia, Colombia, which may be applied more broadly—geographically—to water resource availability and competitive use problems:

1. Can an accurate statistical forecast model be built to aid in local water sustainability discussions?

Methodology Overview

The methodology of this research consists of two major efforts: the preliminary analysis of data, and the prediction of streamflow in Andes, Antioquia, Colombia using a statistical forecasting modeling approach. The model component uses multiple linear regression (MLR) and principal component regression (PCR) methodologies to create suites forecasts. The skill of deterministic and ensemble forecasts are assessed using a variety of standard forecast metrics including coefficient of variation, ranked probability skill score (RPSS), and mean absolute percent error (MAPE).

Thesis Organization

The organization of this research is broken into five major efforts: literature review, the gathering and analysis of data, formation and methodology of forecasting models, results, and discussion. Chapter 2 is the literature review section which covers the topics of climate uncertainty, water scarcity, statistical forecast modeling, and the common drivers of hydrology in Colombia. Chapter 3 consists of an overview of the data gathered. Chapter 4 discusses how the forecast model was set up and executed. Chapter 5 highlights the overall results from the four different model variations evaluated, and then focuses on the more detailed results for the best-determined model. Chapter 6 focuses on the discussion of how the completed forecast model can help members of the Andes community mitigate their risk when it comes to water availability, whether that be due to extreme drought and water allocation issues or during high flood periods where landslide risk increases dramatically. Chapter 7 summarizes the main findings, the significance of the research, and future research to be completed.

II. Literature Review

Chapter Overview

To determine the appropriate data and model for predicting seasonal streamflow it is necessary to investigate the topics and methods that reveal the best approaches. As such, climate uncertainty, water scarcity, forecast modeling, and other research being completed within the area of study are reviewed. This section of the thesis is essential to consolidate what is already known, address any knowledge gaps, and determine how this research can contribute to the further understanding of forecast modeling. The conversation on climate uncertainty, water scarcity, and other research completed in the case study location is brief but essential in gaining required background knowledge. The focus of this section is statistical forecast modeling with an emphasis on the moisture transport process in South America.

Climate Uncertainty and Water Scarcity in Colombia

Studies have found that climate uncertainty will spur an increase in global extreme climate events, which have the potential to impact both water demand and supply (Jimenez Cisneros et al., 2014; Allan and Soden, 2008). These impacts, such as persistent droughts and/or flooding events, have presented considerable challenges to decision-makers involved in regional water planning and management. These challenges include the execution of water reduction measures for agriculture, manufacturing, and residential sectors while simultaneously maintaining fiscal stability and public relations throughout regions (Zimmerman et al., 2016).

Colombia has previously been considered a nation with high water wealth; however, climate uncertainty is expected to bring new water scarcity conflicts throughout the entire region over the next 50 years (Vargas-Pineda, 2020). Many studies throughout this region focus on the water footprint of important crops, water supply available, perceptions of water scarcity, and planning and water management practices (Vargas-Pineda, 2020; Naranjo-Merino, 2018; Murtinho et al., 2013). Changes in water variability and scarcity will greatly impact the coffee industry, the primary crop throughout Colombia, due to its high sensitivity to climatic extremes such as droughts, torrential rains, and landslides (Frank et al., 2011). In Colombia, both water quantity and quality are strongly influenced by climate variability and potential long-term seasonal changes are very likely to develop (Felipe et al., 2013). The perception of water scarcity throughout the region has a significant impact on the implementation of adaption strategies amongst communities and government organizations; with limited community knowledge of the approaching water scarcity being of great concern (Murtinho et al., 2013).

Research has shown that it is very possible for the climate to be affected in both the long-term availability and the short-term variability of water resources throughout many regions (Olmstead 2010). These potential regional impacts could include increased frequency and magnitude of droughts, heatwaves, and floods. Climate uncertainty may also impact essential water resources by influencing long-term changes in precipitation, temperature, humidity, wind intensity, duration of accumulated snowpack, vegetation, soil moisture, and runoff (Olmstead 2010, Solomon et al. 2007, Keating 2021, Delorit et al. 2017). The Bolivian High, eastern and northern winds that carry moisture towards the Andes Mountains from the Amazon, has also been shown to impact precipitation in South America (Vizy and Cook, 2007).

Until recently, there were no set of actions, policies, or management measures unique to the problem of adapting water resource management to global warming (Stakhiv, 1998). In 2019, a book was published that outlines potential modeling frameworks for considering uncertainty in decision making processes (Marchau et al., 2019). These approaches have made it possible for deep uncertainty to be included in practical decision-making processes. Other studies have promoted the principles of water resources planning and river basin management using skillful season-ahead streamflow forecasts (Delorit et al., 2017; Block, 2009). These studies suggest, with a reasonable degree of confidence, that vulnerability within communities due to climate change and water scarcity can be minimized or stabilized in most cases with proper planning and community involvement.

Forecast Modeling

Forecast modeling predominately falls into two categories: dynamical or statistical; however, it can also be used in a hybrid sense (Block et al., 2009; Sabzi, 2017). The dynamical approach is a physically-based model that seeks to simulate physical processes such as runoff or infiltration to produce streamflow predictions (Souza Filho and Lall, 2003; Keating, 2021). The second technique, statistical modeling, utilizes predictor variables to directly estimate the dependent variable through numerical techniques such as multiple linear regression. The hybrid approach is becoming more popular in practical applications, but recent studies have found there are numerous instances where the results have been greatly distorted during implementation (Zhu et al., 2018). Overall, these three forecasting techniques have been widely used in the literature to estimate diverse objectives such as streamflow quantity, earthquake probability, global solar radiation, fish population rates, and even lake-level fluctuations (Block & Rajagopalan, 2009; Sabzi, 2017).

Statistical forecast modeling only provides value if the forecast lead allows time for informed decision-making to occur. Due to the ability to investigate lagged relationships between variables of interest, statistical prediction models have been found to generally outperform dynamical models when it comes to predicting streamflow in many regions, particularly where teleconnective strength between global variables, e.g., El Niño Southern Oscillation (ENSO), and streamflow is greatest (Zimmerman et al., 2016, Delorit et al., 2017, Keating, 2021). Statistical prediction models typically capitalize on identifying patterns and pattern changes from climate-related anomalies at large spatial and temporal scales. They also rely entirely on the relevance and availability of historical data (Chambers et al., 1971). Due to this, there are unique instances where statistical prediction models may not be fully capable of incorporating a complete physical understanding of the climate uncertainty in an area with limited data supply (Block, 2009). However, most models are still able to facilitate in decision-making processes and assist in the development of local policy (Johansen, 2017; Rod et al., 2020). In short, the simplicity of statistical models, and the ease with which outputs are translation to decision-ready outputs, make them attractive to both modelers and decision makers.

There have been a handful of studies which have properly leveraged appropriate lead times to ensure skillful streamflow forecasts, at the seasonal scale. These studies are subsequently applied to sectoral models aid in the development of water allocation policy, energy management, irrigation plans, municipality water management strategy, and environmental services (Cai, 2008; Delorit et al., 2017; Alexander et al., 2021). There are many potential early actions that can be taken when forecast lead times are properly aligned to a decision maker's decision points, and are sufficiently skillful. Though, it is generally accepted that forecast lead and skill are inversely related. Longer lead times allow for a greater range of early actions to be taken but must be balanced to ensure forecast uncertainty remains within acceptable ranges (Bazo et al., 2019).

Although the improvements in seasonal climate forecast skill and advocacy for integration into risk reduction strategies may be well documented, demonstrated forecast use in localized water allocation and policy strategies is extremely limited. There is also a scarcity of literature on how forecast modeling can be used to predict the quality of water for a community's local stream or watershed. One recently published study details water quality forecasts for the purposes of maintaining the proper quality of life in the Great Barrier Reef, Queensland, Australia (Khan et al., 2020). This research used the Short-term Water Information and Forecasting Tools (SWIFT), a statistical-based forecasting tool, designed for operational streamflow forecasting and scientific research in Australia (Hapuarachchi et al., 2017; Kabir et al., 2018).

The techniques used to create the water quality model for this study could be reproduced to develop water quality models at other locations; however, since the method presented in this study is computationally intense a supercomputing platform is an essential requirement (Khan et al., 2020). A previous study, which was completed in 2005, shows that simpler techniques can be used to develop a forecast for streamflow quality but might give up some accuracy in the prediction of pH and some other essential water quality indicators (Kurunc et al., 2005). This research focuses on evaluating different forecasting techniques such as the autoregressive integrated moving average (ARIMA) statistical forecasting technique to determine river water quality. Even though these studies generally apply forecasting techniques to address the growing concern of potable water available to communities, neither of these studies address the importance of community involvement in successful water policy implementation or evaluate how changes in forecasted water availability impact overall community stability due to their dependence on proper water quality and quantity.

Other Relevant Research within Andes, Antioquia, Colombia.

Researchers have taken a keen interest in Andes, Antioquia, Colombia due to a slew of potential areas of environmental concern, water conflict, and economic activities that occur in the area. Most of these matters deal with extreme land and water pollution from ASGM practices, but there have also been a few areas of study that focus on understanding how local agriculture has impacted the region (Schwartz, 2021; O'Brien, 2020). Building resilient communities through the implementation of sociotechnical solutions, focused-working groups, the formation of diverse research teams, and community involvement has also been evaluated by many academics (Smits et al., 2020; Schwartz, 2021; O'Brien, 2020; O'Brien, 2021).

Even though this research has been crucial in the development of safe ASGM practices and ASGM climate impact understanding, there has been little to no progress made to address the growing concern of climate uncertainty. In interviews conducted during physical immersion within the Andes community, numerous members expressed concern for how climate uncertainty could impact their essential economic operations and water availability in the future (Schwartz, 2021). Andes' localized conditions and community interconnectedness make this an ideal location for continued research on the implementation of forecasting modeling, and how technical solutions can be successfully implemented at the community level.

III. Data

Chapter Overview

In this section, the processes for gathering required data and performing preliminary analysis on these data are discussed. The data gathered for this study were obtained from a combination of both global and localized databases. While stakeholders might prefer the use of locally collected data, continuous data that accounts for regional and global forcings should be prioritized when creating a statistical forecasting model. This is particularly true given the known teleconnective strength between the region of interest and global circulations. The gathering and preliminary analyses performed on the streamflow data will be discussed first, followed by each predictor variable used in the modeling. Each of these variables will be listed, explained, and justified; with any transformations or modifications to the data discussed.

Gathering Required Streamflow Data

Andes, Antioquia, Colombia is a rural, largely underdeveloped area, and as such, obtaining consistent data is difficult. A streamflow gauge (Station Code: 26197010) was discovered for the San Juan River approximately 3-km downstream of Andes, Antioquia, Colombia. This streamflow gauge started recording daily mean values in January 1972 until it stopped capturing readings at the end of 2015. There are not currently any other streamflow gauges in the area that have been reported by the Colombian government. This gauge is considered to be reflective of flows near Andes given consumptive extractions from the San Juan River are minimal; most uses are either run-of-river or provide return flows.

The streamflow data gathered were obtained as mean daily values from January 1st, 1972, to December 31st, 2015, from Colombia's national hydrology database known as de Datos de Hidrología y Meteorología (DHIME) (Fig. 4). These data were then aggregated to the monthly time scale, in order to match other input data resolution. After completing the aggregation to the monthly scale, it was discovered that 21 of the 528 months of data were missing. These missing months were from the years 1999, 2000, 2011, and 2015. To fill in these missing data, synthetic data were created in accordance with the standard practice of using each month's respective mean from the remaining years. The streamflow data for the years 1972-1981 were not included in the development of the forecasting model due to ENSO data only being available from 1982-present.

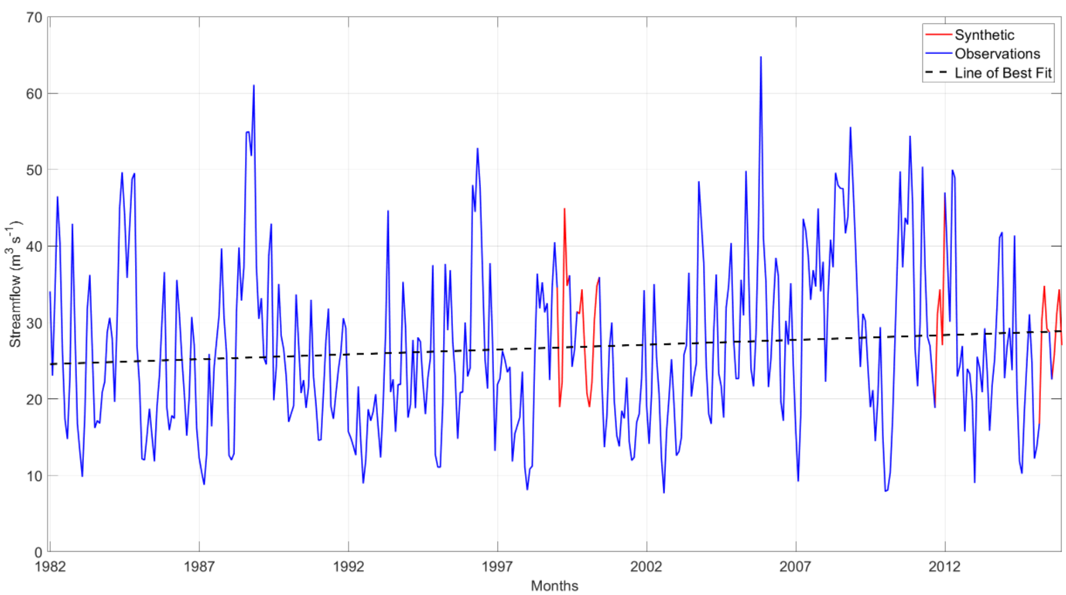

Figure 4: Time-series of the Streamflow in the San Juan River.

Preliminary Analyses on Streamflow Data

The preliminary analyses performed on the streamflow data include time-series analysis, descriptive statistics, Student's t-test, and the Mann-Kendall test. Each of these tests and analyses performed helps identify the basic nature of the data gathered and highlights any potential discontinuities. Here the aforementioned tests are performed both on the monthly data (n = 408) and for the season of interest (September-December, n = 34).

When analyzing the entire dataset (n = 408), the Mann-Kendall test reveals a slight positive increase in monthly flows for the period 1982-2015, with 95% confidence (Fig. 4). Using the Student's t-statistic and the Mann-Kendall's test revealed that the median, minimum, and maximum flows for the first ten years of data (1982-1992) and the last ten years of data (2005-2015) are different at the 95% confidence interval (Table 1); with minimum and maximum values becoming more extreme over time.

Table 1: Minimum, Median, and Maximum for Streamflow Dataset ($\boldsymbol{\alpha = 0.95}$)

Value	First Ten Years (1982-1992)	Last Ten Years (2005-2015)
Minimum	8.77 $m^3\ s^{-1}$	7.92 $m^3\ s^{-1}$
Median	22.99 $m^3\ s^{-1}$	29.24 $m^3\ s^{-1}$
Maximum	61.1 $m^3\ s^{-1}$	64.8 $m^3\ s^{-1}$

Next, descriptive analysis on the streamflow dataset was completed to gain a sense of how the data are distributed, to determine the appropriate modeling techniques (i.e., linear transformations of nonlinear data enable linear modeling approaches). By performing a goodness of fit test on the streamflow data, it was determined that these data are not normal or log-normal but are best fit with a gamma distribution, which is a distribution type capable of mimicking most nonlinear distributions (Fig. 5).

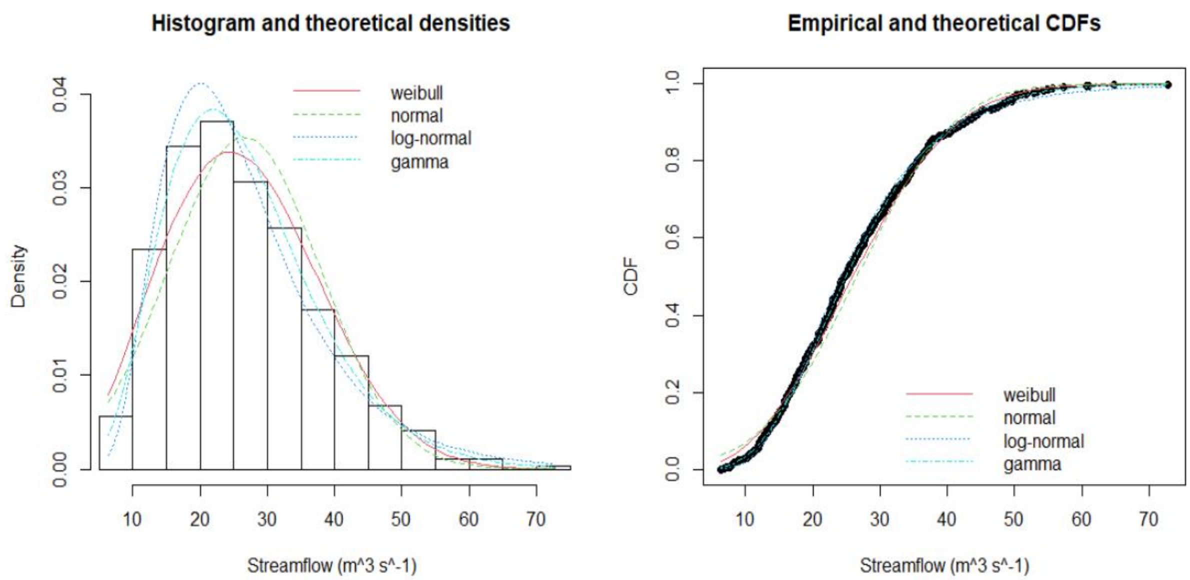

Figure 5: Descriptive Statistics on Streamflow Data (1982-2015)

To aid in the basic understanding of the dependent variable, the tests ran on the entire streamflow dataset were re-accomplished for the September-December (SOND) season of interest. The data for streamflow SOND exhibits the same slight positive increase over the 1982-2015 period and varies significantly on an interannual basis (Fig. 6). Instances of extreme drought are extremely rare during September – December, while high flow events will become more likely due to increased precipitation during this time.

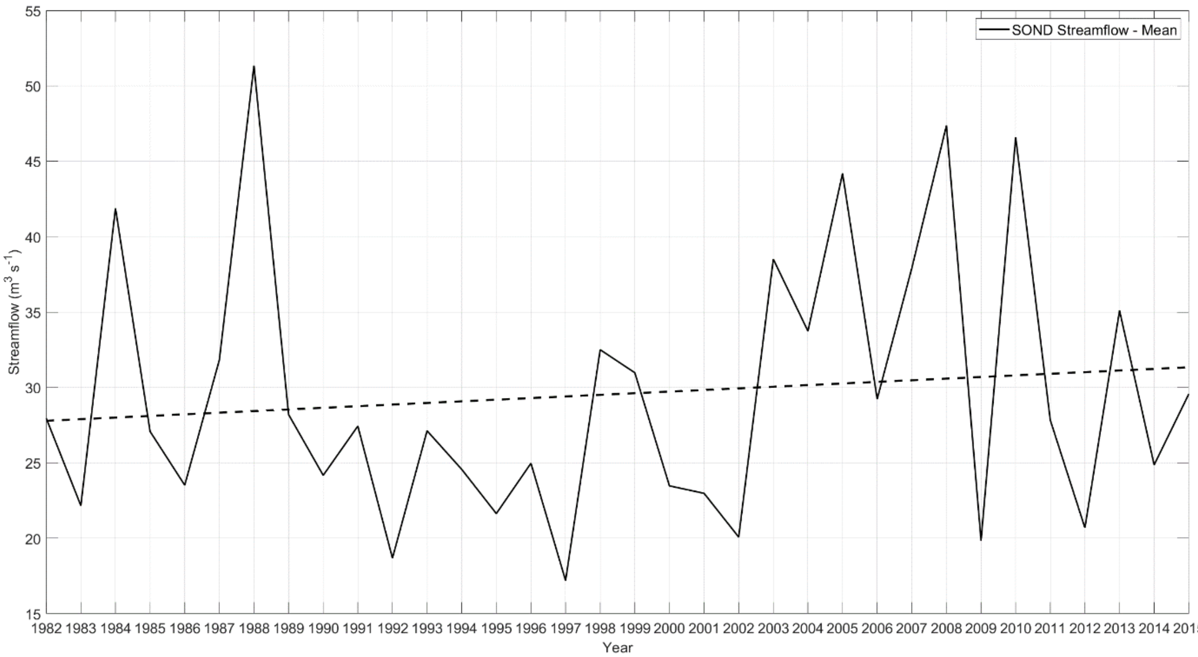

Figure 6: Time-Series for SOND Streamflow

The mean of the first half of data (1982-1997) and the last half of the data (1998-2015) are different upon comparison with 95% confidence. This was conducted using Student's t-test on each of the time periods, which was found to be significant at the 95% confidence level. To highlight the difference in values between these two periods, the specific values for the minimum, median, and maximum are presented in Table 2.

Table 2: Minimum, Median, and Maximum for SOND Dataset ($\boldsymbol{\alpha = 0.95}$)

Value	First Half (1982-1997)	Second Half (1998-2015)
Minimum	17.20 m^3 s^{-1}	19.84 m^3 s^{-1}
Median	27.08 m^3 s^{-1}	29.56 m^3 s^{-1}
Maximum	51.33 m^3 s^{-1}	47.37 m^3 s^{-1}

To test if the peaks and troughs in streamflow are becoming more extreme over time, the Mann-Kendall's test was performed on the 75th and 25th percentiles of the SOND time series. No trend was found, which suggests that peak and low flow years are not increasingly variable.

Lastly, a descriptive analysis of the data was completed to gain a sense of how it is distributed. By performing a goodness of fit test on the streamflow data, it was determined that these data are not normal but are best fit with the log-normal distribution (Fig. 7).

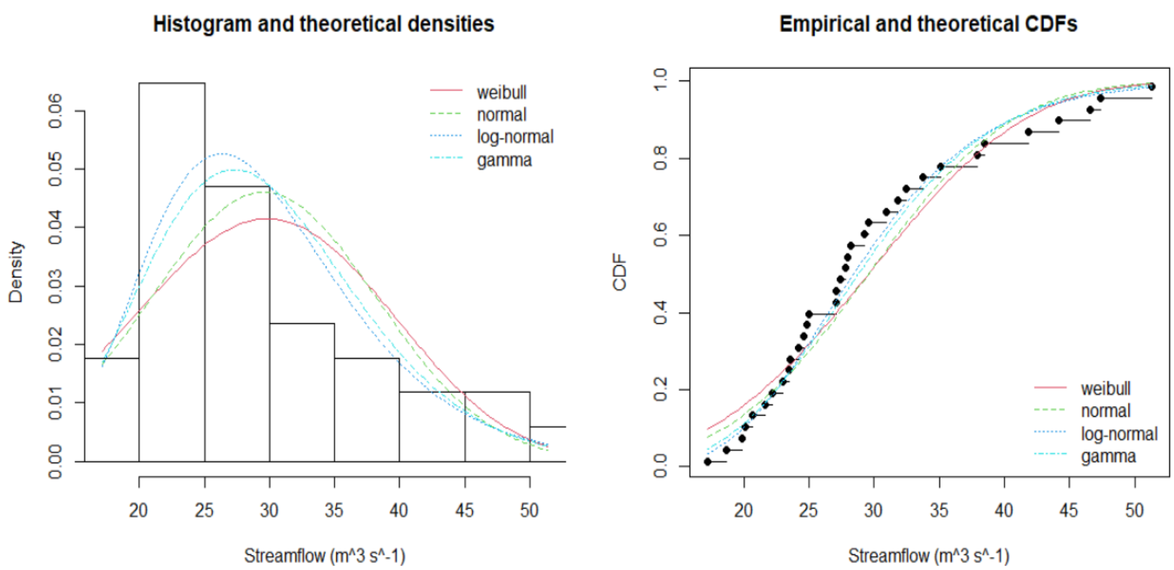

Figure 7: Descriptive Statistics on SOND Streamflow (1982-2015)

Variables of Interest: Local and Global Predictors

A mixture of global and local predictor variables that have historically been shown to have an influence on the streamflow are evaluated in this section. Global predictors, e.g., sea surface temperature (SSTs) and sea level pressure (SLP), are generally tied to general circulations like El Niño Southern Oscillation (ENSO). The SSTs evaluated here consist of the 3.4 region of the ENSO anomaly (120W-180W, 5S-5N). The SLP evaluated in this research emanated from the Southern Oscillation Index (SOI), which is a pressure differential between Darwin, Australia, and Tahiti. The local predictors include precipitation, soil moisture, and temperature. The following sections list, explain, and justify any transformations or modifications made to the collected predictor variable datasets.

Precipitation

Reliable, station-based precipitation data were not available at the local level or from any of the Colombian databases. To remedy this, globally gridded precipitation (mm) data were obtained from the NOAA Physical Sciences Division database at the 1.0 x 1.0˚ grid resolution. These data were collected at the hourly time step for the period 1979 – present. These data were obtained from the latitude range 283.7˚E – 284.2˚E and longitude 5.5˚N – 6.3˚N, which is the coordinate range closest to Andes, for which data are available. There were no missing values in this dataset. These data were aggregated to the monthly scale for the period 1982-2015.

Soil Moisture

Soil moisture (mm) data were obtained at the 0.5 x 0.5˚ grid resolution from the NOAA Physical Sciences Division database. These data were collected at the daily time step for the period 1948 – present, from the latitude range 248˚E – 284.5˚E and longitude 5.5˚N – 6˚N. There were no missing values in this dataset. These data were aggregated to the monthly scale for the appropriate 1982-2015 period. No other modifications were made to this dataset.

The El Niño–Southern Oscillation (ENSO)

ENSO involves changes in sea surface temperatures in the Pacific, and is related atmospheric circulation patterns that largely drive global climate. The ENSO index is characterized by a 3-month running mean anomaly of Sea Surface Temperatures (SST). When SSTs are higher than the long-term mean, it is referred to as El Niño, and when they are below the long-term mean, it is referred to as a La Niña. The two phases of ENSO are marked by different 'strengths' measured as the magnitude of the departure from the long-term mean. The ability to monitor and forecast ENSO events has improved significantly since the 1990s (Balmaseda et al., 1995; Trenberth, 2002; Block et al., 2009; Delorit et al., 2017; Chen, 2020).

ENSO displays a spatial pattern in the Equatorial Pacific, typically extending from 160˚E, east to the coast of South America, and between 5˚N and 5˚S. Because the intensity and behavior of ENSO vary across this wide area, researchers have proposed regions for which indices are calculated: Niño 1 + 2, Niño 3, Niño 3.4, and Niño 4 (Fig. 8).

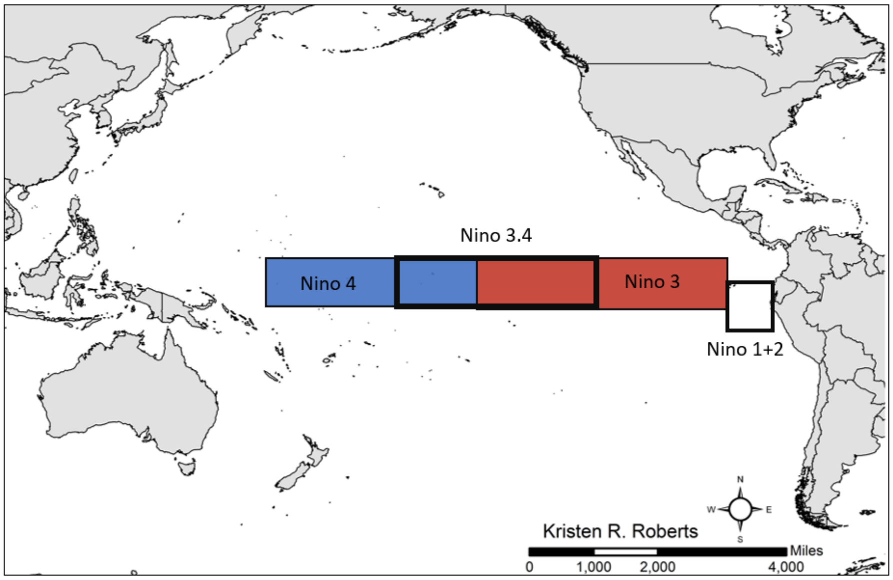

Figure 8: Niño Regions

Research has commonly shown the El Niño and La Niña phases have a periodicity of 2-8 years (Neelin et al., 1998; Bombardi et al., 2014; Delorit et al., 2017). The ENSO data used in this research were gathered from the National Oceanic and Atmospheric Administration's website in anomaly values. ENSO is available from 1982 – the present. In this research, the 3.4 index was chosen because it displays the highest correlation with streamflow values for the SOND season of interest (Fig. 11. A). This finding is consistent with other South American-based studies (Delorit et al., 2017).

The Niño 3.4 region is related to local streamflow and precipitation data (Fig 9). Visually, it is apparent that the highest streamflow and precipitation values occur during a negative SST (La Niña) phase. During the La Niña phases, both precipitation and streamflow appear elevated, e.g., 1988, 1996, 2008, and 2011. The opposite is also true; a strong El Niño most often results in severe low flows and precipitation, e.g., 1982, 1997, and 2010. Extreme high and low streamflow values also seem to be related to the strength of the Niño 3.4 phase.

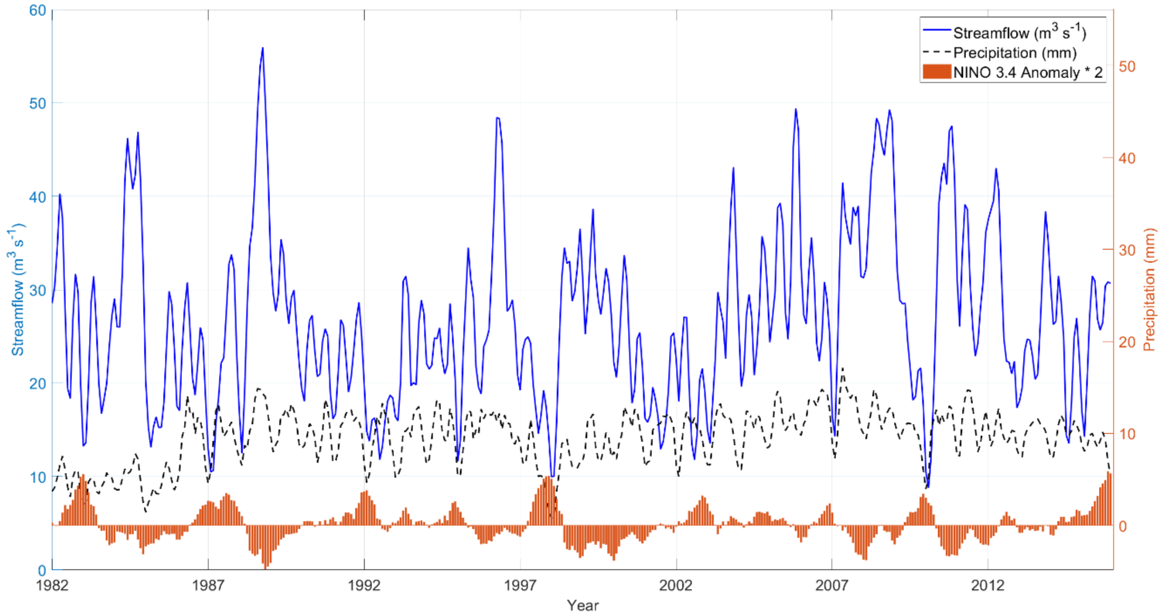

Figure 9: Total annual precipitation (black dashed), streamflow (blue solid), and Niño 3.4 sea surface temperature anomalies (orange bars).

Southern Oscillation Index (SOI)

The Southern Oscillation Index (SOI) is a standardized index, based on the observed sea level pressure (SLP) differences between Tahiti and Darwin, Australia. The SOI is one measure of the large-scale fluctuations in air pressure occurring between the western and eastern tropical Pacific (i.e., the state of the Southern Oscillation) during El Niño and La Niña episodes. In general, the three-month moving mean time series of the SOI correspond very well with changes in ocean temperatures across the eastern tropical Pacific (Yan et al., 2011). The negative phase of the SOI represents below-normal air pressure at Tahiti, while above-normal air pressure at Darwin. Prolonged periods of negative SOI values coincide with abnormally warm ocean waters across the eastern tropical Pacific typical of El Niño episodes. The methodology used to calculate SOI is available below.

$$SOI = \frac{(Standardized\ Tahiti - Standardized\ Darwin)}{Monthly\ Standard\ Deviation} \qquad \text{(Eqn. 1)}$$

The SOI data were gathered from NOAA's website in their standardized anomaly values. SOI is available from January 1951 – present in both their anomaly and standardized forms. The only modification to the data was curtailing it to the proper time frame 1982-2015. ENSO and SOI have a strong negative correlation (Fig 10).

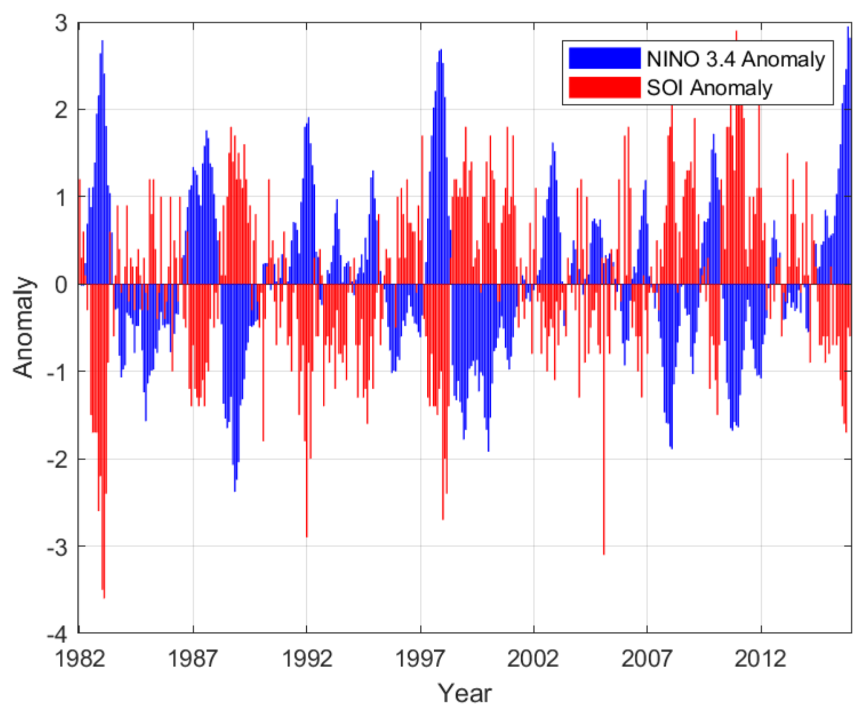

Figure 10: Relationship between ENSO and SOI

Correlation Plots Between DV and Predictor Variables: ENSO, SOI, and Precipitation

NOAA's Linear Correlations in Atmospheric Seasonal/Monthly Averages mapping tool was used to evaluate which region of SST, precipitation, and SLP were most highly correlated with SOND streamflow (Fig. 11). Plot A shows streamflow has a strong negative correlation with SSTs in the equatorial pacific, which matches previous findings (Fig. 9). Plots B and C show streamflow has high positive correlations with both precipitation and SOI.

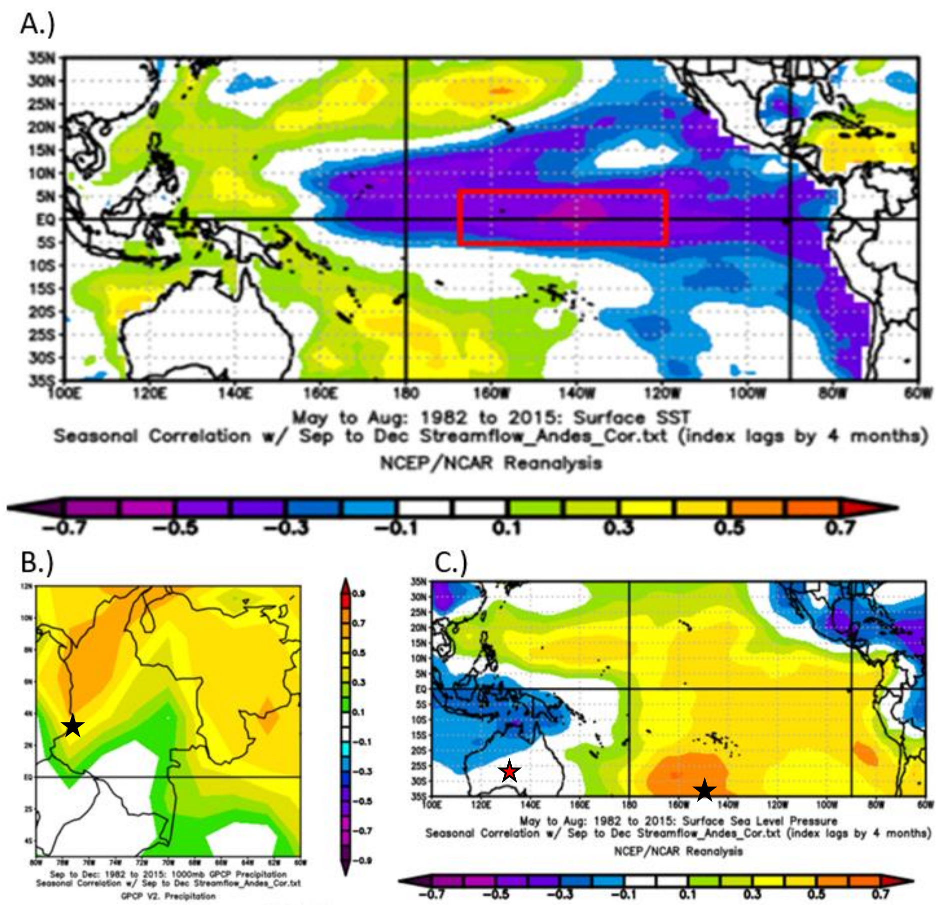

Figure 11: (A) Correlation of MJJA SST with SOND streamflow (1982-2015). Note box represents Niño 3.4; (B) Correlation of Precipitation with SOND streamflow (1982-2015); (C) Correlation of MJJA SLP with SOND streamflow (1982-2015).

This section provided the necessary information on the variables used in the development of the forecast models (i.e., streamflow, precipitation, Niño 3.4 anomaly, SOI, and soil moisture). These variables were assembled using both local and globalized datasets for the 1982-2015 period, with the smallest possible number of modifications made to the datasets. To gain a better understanding of streamflow in the San Juan River, basic statistical tests such as time-series analysis, the Mann-Kendall test, the goodness of fit test, and Student's t-test were used to provide insight into the data's behavior and authenticity. These tests resulted in the identification of basic trends and relationships, such as increasing variability in extreme streamflow values or positive or negative correlations between the dependent and independent variables. Streamflow has both positive and negative correlations with the predictor variables. To start, all Niño regions Niño 1 + 2, Niño 3, Niño 3.4, and Niño 4 were evaluated using both statistical tests and correlation mapping tools to determine which region was most highly related with streamflow in Andes. These tests resulted in the use of Niño 3.4 region as a predictor variable, while the other regions were not analyzed further. Streamflow is positively correlated with precipitation, soil moisture, and the southern oscillation index, while it is negatively correlated with Niño 3.4 anomaly.

IV. Forecasting Model Methodology

Chapter Overview

Here, the statistical forecast model configuration and modes are discussed. This includes the initial model setup utilizing techniques including simple multiple linear regression (MLR), principal component regression (PCR), and PCR with cross-validation (PCA+CV). Four independent variable configurations are tested for each modeling technique and will be referred to as *predictor modes.* The section concludes with a presentation of validation and performance metrics used to evaluate statistical and probabilistic skill.

Forecasting Model Framework and Steps

The model used to evaluate these data is based primarily on techniques developed by a student conducting research in South America, though never before for locations to the east of the cordillera of the Andes Mountains (Delorit et al. 2017; Zimmerman et al., 2016; Keating, 2021). The modeling processes are divided into three different steps: modeling, post-analysis/performance metric production, and evaluation. The modeling portion of the methodology is focused on the creation of MLR, PCA, and PCA+CV forecasts. The post-analysis portion focuses on the production of forecast skill metrics that are used to determine which, if any, models possess skill and enables a discussion of the tradeoffs of differently parametrized models, i.e., independent variable combinations. The evaluation process allows the four created *predictor modes* to be ran through the models, and various lead times to be tested in order to explore tradeoffs between forecast issue date and skill, which ultimately targets stakeholder's value of information.

Forecast Modeling Techniques: Multiple Linear Regression, Principal Component Analysis, and Cross-validated Principal Component Regression Models

Three different forecasting models were created to analyze the predictability of streamflow in Andes: multiple linear regression, principal component regression-based forecasting model, and a principal component model with cross-validation. Multiple linear regression forecast modeling is a statistical technique that uses several explanatory variables (precipitation, temperature, ENSO anomaly, SOI anomaly, etc.) to predict the outcome of a response variable, i.e., streamflow. This technique is commonly used to create models in many different fields and has become very popular due to its simplicity (Equation 2),

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \epsilon \qquad \text{(Eqn. 2)}$$

Where, i is the number of observations, y_i= dependent variable, β_0 = y-intercept, x_i = predictor variables, β_n = slope coefficient for each predictor variable, and ϵ = model error.

This type of deterministic model is simple and can be improved by using principal components in place of independent variables. This process results in a set of principal components representing the variance in the predictors. It is important to clarify that principal component analysis (PCA) is not a regression; it is a technique used to create new independent variables from the original set. PCA is commonly applied in forecasting and hindcasting to ensure a fair forecast is made by reducing both model dimensionality, when principal component retention rulesets are applied, and multicollinearity (Delorit et al., 2017).

Reduction of model dimensionality is possible when then number of input variables is reduced. PCA naturally produces an independent variable set that is ordered by variance in the underlying input dataset. That is, the first PC explains the most variability, the second explains less variability than the first PC, but more than the third PC, and so on. Because of this, rulesets are typically used to determine the number of PCs to retain. By retaining fewer PCs in the PCR, than the total number produced by PCA, dimensionality—complexity—of the underlying model is reduced, without significant loss of skill. One such rule is Joffille's rule, which is used as a PC retention and dimensionality reduction technique in this research. This method commonly retains any principal component whose Eigenvalue was greater or equal to 0.7.

The last model created, PCA+CV, was a leave-one-out crossvalidated hindcast model across the dataset. This technique produces a bias-reduced, deterministic prediction; however, since this form of cross-validation removes the streamflow value for the time-step being predicted, the deterministic skill of the model, e.g., r^2 can be significantly lower than non-crossvalidated alternatives. Utilizing Jofille's rule in the crossvalidated model results in the PCA+CV model being an unbiased and conservative estimate of streamflow.

Variable Combinations to Allow Adequate Lead Time and Evaluate Scenarios

To evaluate how streamflow is influenced during the season of interest (SOND), up to 35 independent variables are made available from the underlying predictor set: Niño 3.4, SOI, precipitation, soil moisture, and the difference of ENSO and SOI, by adjusting temporal combinations. These 35 variables were created using the one-month aggregated mean values across May, June, July, and August in different combinations, e.g., May, June, July, August, MJJA, MJJ, and MJ, for the years 1982-2015. The purpose of the combinations is to test the time-skill tradeoffs for different forecast leads against the predictability of streamflow. Using these combinations allows the forecasting lead-time to be extended or adjusted up to three months if required. For example, if information up to June 30^{th} is included in the forecast being issued for SOND, this forecast has an issue date equal to the day after the last day of information is collected (July 1^{st}) and has a lead equivalent to the number of days or months between the issue date and the first day of the season of interest (2 months).

Next, these 35 input variables are iteratively used as inputs to the multiple linear regression (MLR) and principal component regression (PCR) and cross-validated principal component regression (PCR+CV) models for both normal and log-normal streamflow. Due to the distribution of streamflow typically being log-normal, a transformation of the dependent variable may be necessary to improve model skill. However, sometimes this transformation is not always necessary or particularly beneficial from a forecast skill perspective. Each forecasting model, e.g., MLR, PCR, and

PCR+CV was run in four modes: *basic predictors*, *expanded predictors*, *top predictors* from each category, and *ENSO Only* for both normal and log-normal streamflow. In total 24 unique model runs were completed, with the goal of balancing input complexity and deterministic skill. Each mode is explained below.

1. *Basic Predictors* is comprised of twenty-eight independent variables. These twenty-eight variables consist of seven ENSO 3.4 variables, seven SOI index variables, seven precipitation variables, and seven soil moisture variables (Table 3). This model is the baseline model that was used to test the forecasting model code and determine if ENSO, SOI, precipitation, and soil moisture were good predictors of streamflow in Andes.
2. *Expanded Predictors* (Appendix, Table 10) consists of twenty-nine independent variables. These twenty-nine variables consist of seven ENSO 3.4 variables, seven SOI index variables, seven precipitation variables, the most highly correlated soil moisture variable, and seven new variables that represent the difference between ENSO and SOI. This new variable, known as the Difference, was inserted into the model due to ENSO and SOI being highly correlated with streamflow, but in the opposite directions (Fig 10).
3. *Top Predictors* (Appendix, Table 11) includes the top predictors from each of the variable categories: MJJ Niño 3.4, August SOI, August Precipitation, and May Soil Moisture. The newly created Difference variable was not included in this model because it was discovered to not be a significant contributor to streamflow

prediction. The variables used in this model were determined by viewing which singular variable from each category (precipitation, soil moisture, ENSO, and SOI) had the highest correlation with streamflow in the basic mode (Fig. 14). This combination was evaluated to see how forecasts would perform with limited input information.

4. *ENSO Only* consists of the seven ENSO 3.4 combinations i.e., May ENSO, June ENSO, July ENSO, August ENSO, MJJA ENSO, MJJ ENSO, and MJ ENSO (Table 9, Chapter 6). This fourth model was created to evaluate how easily streamflow could be predicted if community members only knew information about the current ENSO cycle.

When creating the above variable combinations, it was mathematically necessary to have at least two more observations (years of streamflow) than independent variables. This was due to linear modeling constraints outlined in previous research (Chambers et al. 1972). Due to this finding, independent variable combinations were kept below thirty-two variables for the MLR, PCA, and PCA+CV models. The input variables for the *basic predictor* mode are shown in Table 3, while the other three variable combinations are showcased in the Appendix. This table showcases how each variable for the basic predictor mode was combined and gives a basic description of each created variable. Each of the variables were made using either local or globally gridded data mean values for the time-period 1982-2015 as shown in the variable description. Lead-times are described for each variable to illustrate the relationship between each variable combination and the associated forecast lead.

Table 3: Basic Predictor Mode Variable Combinations

Variable	Variable Type	Variable Title	Lead Time	Variable Description
1	ENSO	Niño 3.4_May	3-Month Lead	Mean value, 1982-2015
2		Niño 3.4 _Jun	2-Month Lead	
3		Niño 3.4 _Jul	1-Month Lead	
4		Niño 3.4 _Aug	0-Month Lead	
5		Niño 3.4 _MJJA	0-Month Lead	
6		Niño 3.4 _MJJ	1-Month Lead	
7		Niño 3.4 _MJ	2-Month Lead	
8	SOI	SOI_May	3-Month Lead	
9		SOI _Jun	2-Month Lead	
10		SOI _Jul	1-Month Lead	
11		SOI _Aug	0-Month Lead	
12		SOI _MJJA	0-Month Lead	
13		SOI _MJJ	1-Month Lead	
14		SOI _MJ	2-Month Lead	
15	Precipitation	Precip_May	3-Month Lead	Globally gridded, Mean value, 1982-2015
16		Precip _Jun	2-Month Lead	
17		Precip _Jul	1-Month Lead	
18		Precip _Aug	0-Month Lead	
19		Precip _MJJA	0-Month Lead	
20		Precip _MJJ	1-Month Lead	
21		Precip _MJ	2-Month Lead	
22	Soil Moisture	SM_May	3-Month Lead	
23		SM_Jun	2-Month Lead	
24		SM _Jul	1-Month Lead	
25		SM _Aug	0-Month Lead	
26		SM _MJJA	0-Month Lead	
27		SM _MJJ	1-Month Lead	
28		SM _MJ	2-Month Lead	

Validation Metrics

To evaluate the skill of each model run, and to provide a basis for exploration of model-to-model tradeoffs, the mean absolute percent error (MAPE), and ranked probability skill score (RPSS) are calculated. MAPE is the most widely used forecasting uncertainty statistic and has been used in multiple instances to analyze forecast skill (Delorit et al., 2017; Keating, 2021; Zimmerman et al., 2013). When utilizing MAPE, a score above 50 percent signifies inaccurate forecasting, a score between 20 and 50 percent signifies a prediction model of "reasonable" quality, a score between 10 and 20 percent represents a good forecasting model, and a score that falls below 10, the forecast model is said to be of "excellent" quality,

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_t-\hat{y}_t}{y_t}\right| \times 100 \qquad \text{(Eqn. 3)}$$

where n is the number of observations; y_t is mean streamflow observation at time *t*; and $\hat{y}_t$ = predicted streamflow value at *t*.

RPSS is a metric used to account for probabilistic skill in the model outputs. Ensembles are generated by drawing, with replacement, from a normal distribution of the errors between the observations and the deterministic model values. The number of random errors generated is equal to the product of the number of ensembles desired and the number of predicted values in the deterministic model output. For example, if 100 ensembles are desired for a 50-year forecast, the total number of errors is 5,000. The random errors are added to the replicated deterministic model outputs. Following the example above, the 50-year deterministic forecast output is replicated 100 times, such that the 5,000 random error terms may be added to each forecast value. These ensembles are then used to calculate ranked probability skill score (RPSS).

$$RPSS = \frac{\overline{RPS} - \overline{RPS}_{reference}}{0 - \overline{RPS}_{reference}} = 1 - \frac{\overline{RPS}}{\overline{RPS}_{reference}} \qquad \text{(Eqn. 4)}$$

Where: $\overline{RPS}$ = ranked probability score

For RPSS, a score greater than zero constitutes a forecast outperforming the climatological odds. In this calculation, a reference to climatology is first established by separating the distribution of observed data into categories based on the characteristics of the distribution. This becomes the standard against which the prediction ensembles are tested. Climatology and the prediction model are scored based on the number of ensemble predictions that fall in the same category as the observed data, for the predefined categories.

Any time the forecast and the observed value occur in the same bin the RPSS increases. In cases where the forecast is not in agreement with observations, the RPSS falls proportional to the degree of inaccuracy. For example, if the forecast is for above-normal streamflow, and the observation is below normal, the RPSS score is reduced more than if the forecast were for near-normal. If both forecasts are wrong; the first scenario misses by two categories, and is penalized accordingly.

The following figure shows an example of what 1000 random ensemble generations using the PCA+CV method looks like in comparison to the observed streamflow data.

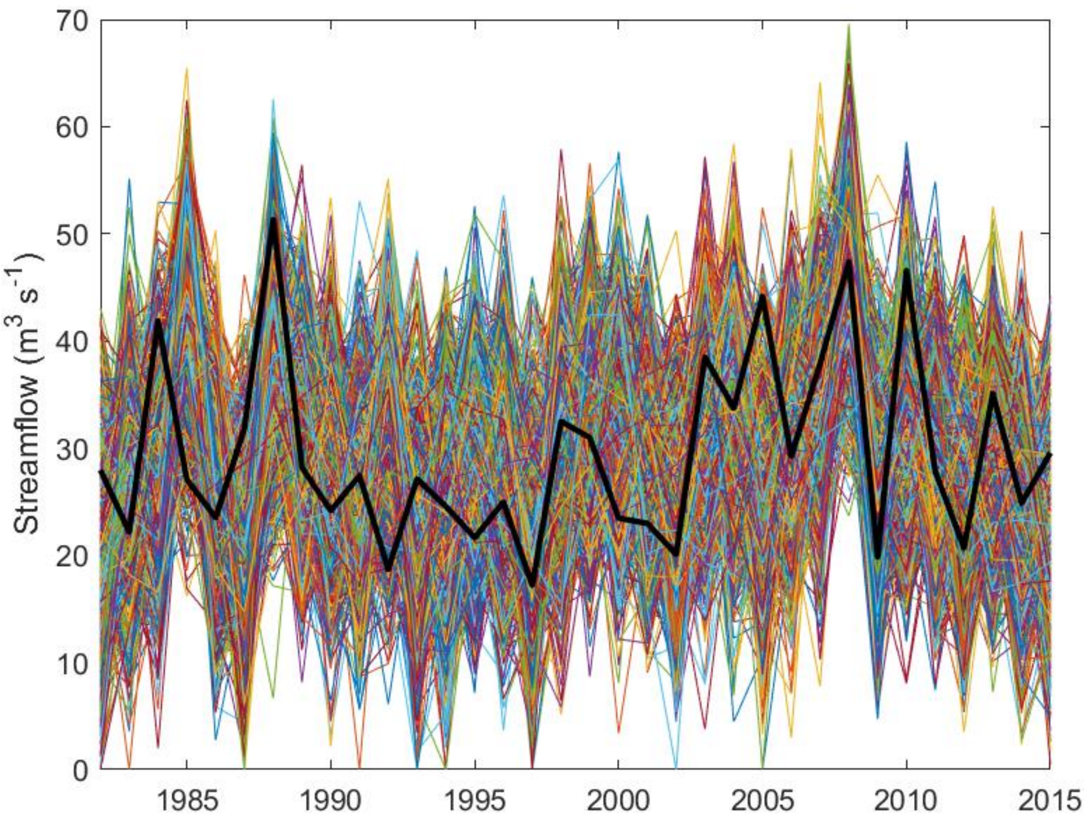

Figure 12: 1000 Generated Ensembles for Basic Predictor Mode, PCA+CV Model

V. Results

Chapter Overview

In this section the results are organized to ensure a certain straightforwardness. First, all four combinations of independent variables discussed in the variable setup section (i.e., *predictor modes*) were analyzed to determine which combination of variables best forecasts streamflow in Andes. After the best configuration was determined, the forecasting results for that configuration are discussed in depth. These results showcase the independent variable correlations with streamflow, a correlation heat map against the other independent variables, comparisons of MLR, PCR, and PCR+CV, and deterministic and probabilistic skill metrics of the best-performing model.

Competitive Model Comparison and Best Model Results

Each of the four variable combinations were analyzed using both normal and log-normal streamflow. However, due to the similarity in skill metrics for normal and log-normal flow inputs, the log-normal translation is not discussed, because the assumption of normality removes transformations. The four combinations of model modes yielded similar results, though the *basic predictors* mode produced the best skill scores (Table 4). The *extended predictors* mode produces a lower adjusted R^2 than the simplified *basic predictors*, which suggests that additional information, in the form of increasing the number of independent variables, adds false skill. The *top predictors* mode, which retains just those variables that correlate highest with streamflow has the lowest ratio between R^2 and adjusted R^2, which is to be expected, though overall, it is not as deterministically, or probabilistically skillful as the *basic predictors*. Finally, the *ENSO only* mode produced skillful results, and is valuable as a limited-data model, i.e., SSTs explain roughly 60% of SOND streamflow variability. Still, the most skillful mode remains *basic predictors*, and as such, the remainder of the results focus on its performance.

Table 4: Model Comparison between Multiple Variable Combinations

		Basic Predictors	Extended Predictors	Top Predictors	ENSO Only
Deterministic Skill Metrics	Linear Prediction: R^2	**0.959**	0.803	0.446	0.367
	Linear Prediction: adjusted R^2	**0.772**	0.276	0.370	0.197
	Linear Prediction: *p-value*	**0.0243**	0.261	0.0014	0.0732
	All PC: R^2	**0.959**	0.961	0.446	0.367
	Cross-Validation: R^2	**0.240**	0.003	0.289	0.083
	Cross-Validation – Joffille's: R^2	**0.228**	0.120	0.222	0.135
Probabilistic Skill Metrics	*MAPE* (50th percentile forecast)	**21.87%**	23.24%	21.98%	24.52%
	RPSS [5,24,5]	**21.14%**	7.55%	8.67%	14.38%

The results of the PCA suggest that there are two major modes of variability in the underlying independent variable set, which is observed in the Scree plot (Fig. 13). For the *basic predictors* mode, 5 PCs were retained using Joffille's rule, which represents 91% of variability explained and an 82% reduction in dimensionality. Back-correlating the independent variables with the PCs revealed that May, June, and August Niño 3.4

variable combinations were identified as the predictors that explains the most variability in streamflow through the PCA. That is PCs 1, 2, and 3 are most correlated with the aforementioned variables.

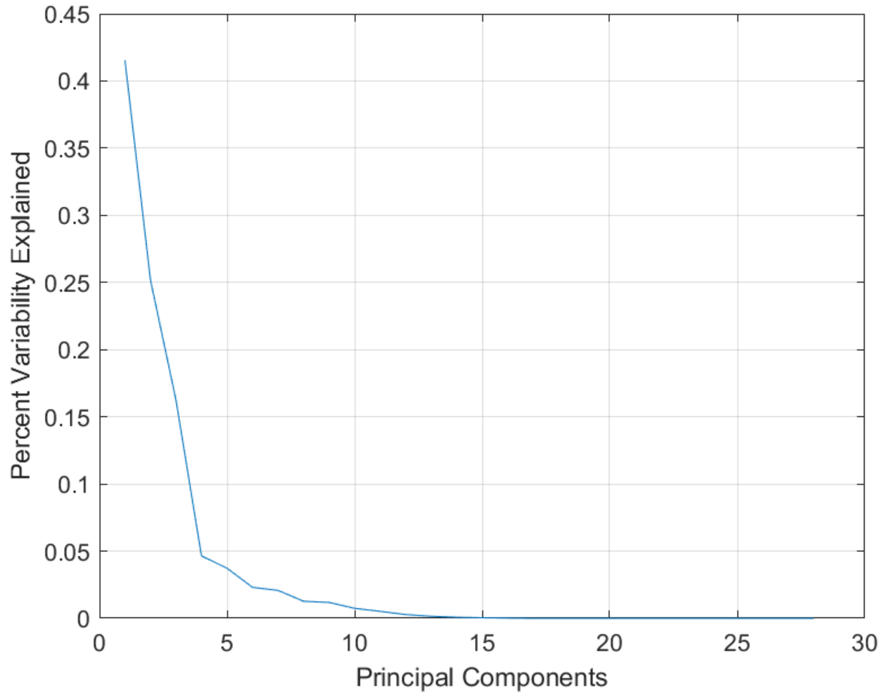

Figure 13: Scree plot to show the variability explained by principal components

To better understand the *basic predictor* mode, a cross-correlation map of its 28 independent variable combinations (Table 3) against streamflow, reveals that Niño 3.4 and soil moisture are negatively related to streamflow, while SOI and precipitation are positively correlated (Fig. 14). These correlations range between -0.5 to 0.5, with Niño 3.4 and SOI having the strongest correlations. This suggests that global forcings are more highly linked to seasonal trends in streamflow than local variables.

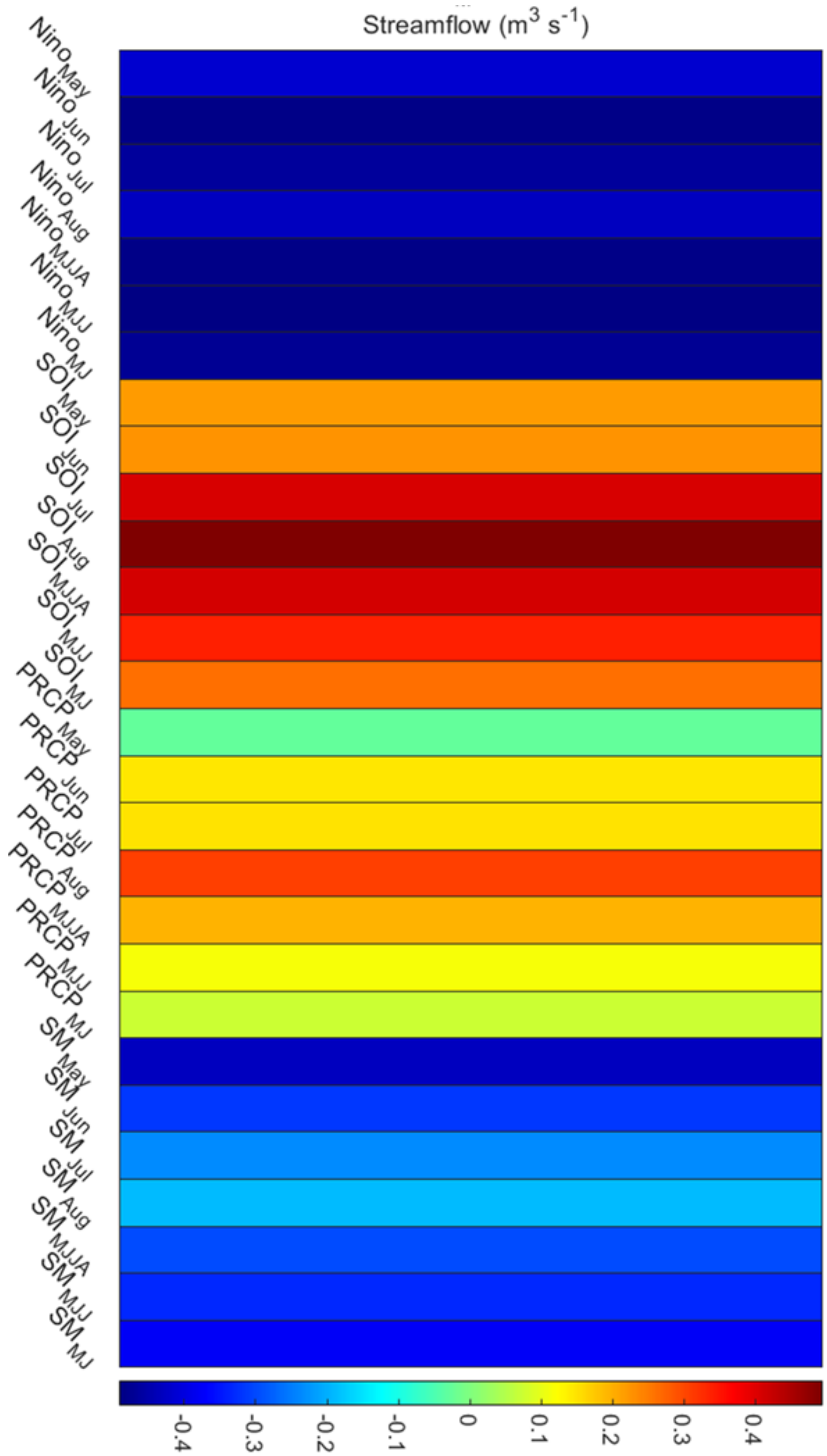

Figure 14: Cross-Correlation Heatmap between DV and IVs

Deterministic, temporal comparisons of the MLR, PCR, and PCR+CV models to observed streamflow for the *basic* and top *predictor modes* illustrate the tradeoff in model skill and model fairness (Figs. 15-17).

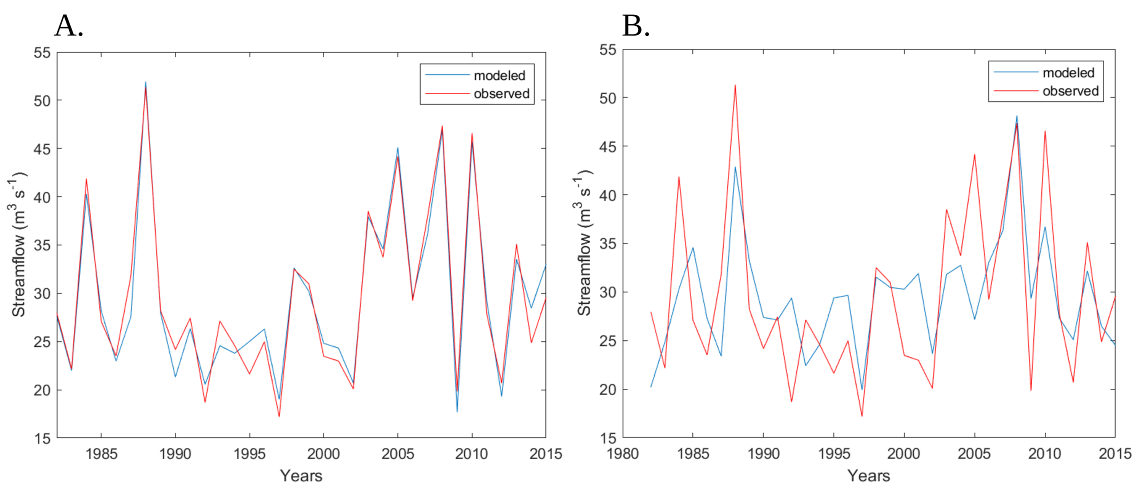

Figure 15: Model Performance between All PC's for (A) *Basic Predictors* and (B) *Top Predictor* Modes

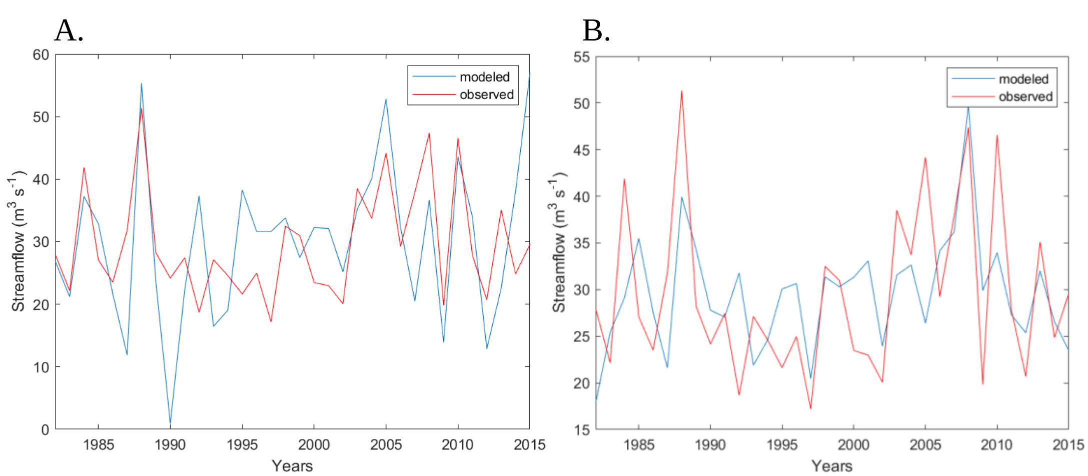

Figure 16: All directly as Predictors, drop-one-year Cross-Validation Model Performance between (A) *Basic Predictors* and (B) *Top Predictors* Modes

Even though the top predictor variable does not perform as well in MLR and PCA, it still performs extremely well in the PCA+CV modeling without a ruleset (Table 4, Fig. 16).

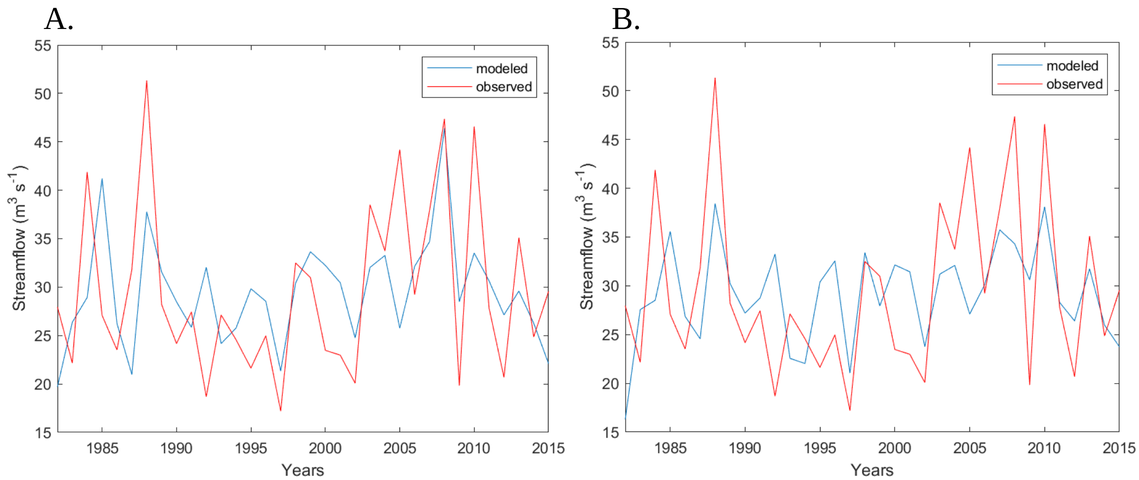

Figure 17: Joffille's Rule, drop-one-year Cross Validation Model Performance between Modes. (A) *Basic Predictors*, Rule-based n= 5 PC(S) as Predictors (B) *Top Predictors*, Rule-based n= 3 PC(S) as Predictors

The *basic predictors* mode does a better job of predicting the observed streamflow extremes during cross-validation, and with the Joffille's PC-retention ruleset, than any of the other variable combinations analyzed. This metric also performs the best when utilizing, the MAPE and RPSS model validation metrics. To establish forecast skill using MAPE at various forecast percentiles and RPSS, 1000 ensembles were created to be able to establish errors as discussed in the forecasting methodology (Fig. 12).

MAPEs for all modes are of "reasonable" quality and are relatively similar (Table 5). Forecast percentile scores are derived from the ensembles and presented as a measure of forecast pessimism (25th percentile) and optimism (75th percentile). Generally, optimistic forecast ensembles perform better than pessimistic ensembles, which is a reflection of the fact that low-flow conditions in Andes are not synonymous with drought conditions, which are extremely rare in Andes.

Table 5: MAPE Results for All Variable Combinations

Description	**Basic Predictors**	**Expanded Predictors**	**Top Predictors**	**ENSO Only**
Deterministic	21.87	23.24	21.98	24.52
25th percentile forecast	30.78	34.56	31.68	31.16
50th percentile forecast	21.63	22.91	21.54	24.61
75th percentile forecast	23.34	24.29	22.68	24.92

The accuracy of this created ensemble forecast is shown in a boxplot format to establish how well the PCR+CV values compare to the original observed streamflow values. The PCR+CV model using *basic predictor* mode performs well as shown in the boxplot (Fig. 18). Nineteen of the thirty-four years of predictions are within the interquartile range. None of the predictions fall outside the 5th and 95th percentile, i.e., not considered outliers in the forecast, though 15 predictions do exist outside of the interquartile range.

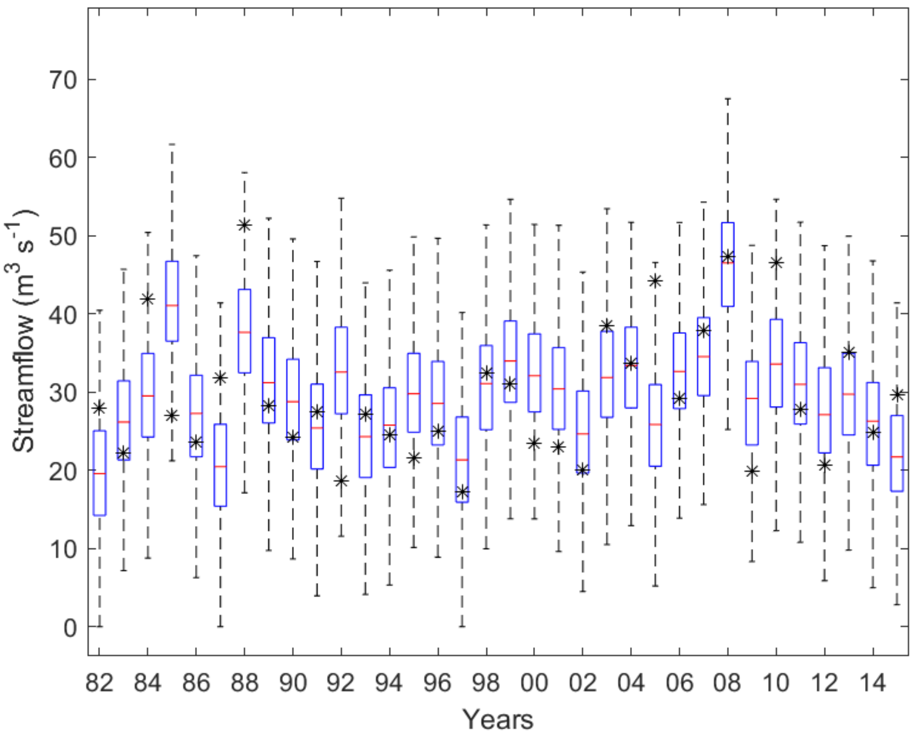

Figure 18: Cross-validated September forecast of Streamflow (m^3s^{-1})

The same information, as presented in a boxplot can be presented probabilistically (Fig. 19) and as a function of the deterministic forecast according to categorical binning of observed streamflow data (Table 6)**.** Using flow-regime breakpoints to establish categories (i.e., above normal, 6; near normal, 18; or below normal, 10), forecast certainty can be evaluated for the ensemble predictions. Given these 'breaks' a flow below 24 m^3s^{-1} constitutes a below normal flow, while a normal flow exceeds this value but is below 38 m^3s^{-1}, and any flow above 38 m^3s^{-1}, is considered above normal. When the observation falls within the strongest prediction category, a "hit" occurs. For the PCR+CV forecast, the number of years where the probabilistic ensemble majority correctly corresponds with the observed category is 20 of 34 predicted years (59%).

While this outcome is not impressive, changing the metric just slightly, to a case where a forecast is only issued when the total ensemble proportion in the largest category is greater than 50%, i.e., more than half of the ensembles occur in a single category, the forecast accuracy improves to 70%. The model issues a forecast in just 10 of 34 years and declines to forecast in the remaining 24 years, which are mostly years with normally observed flows.

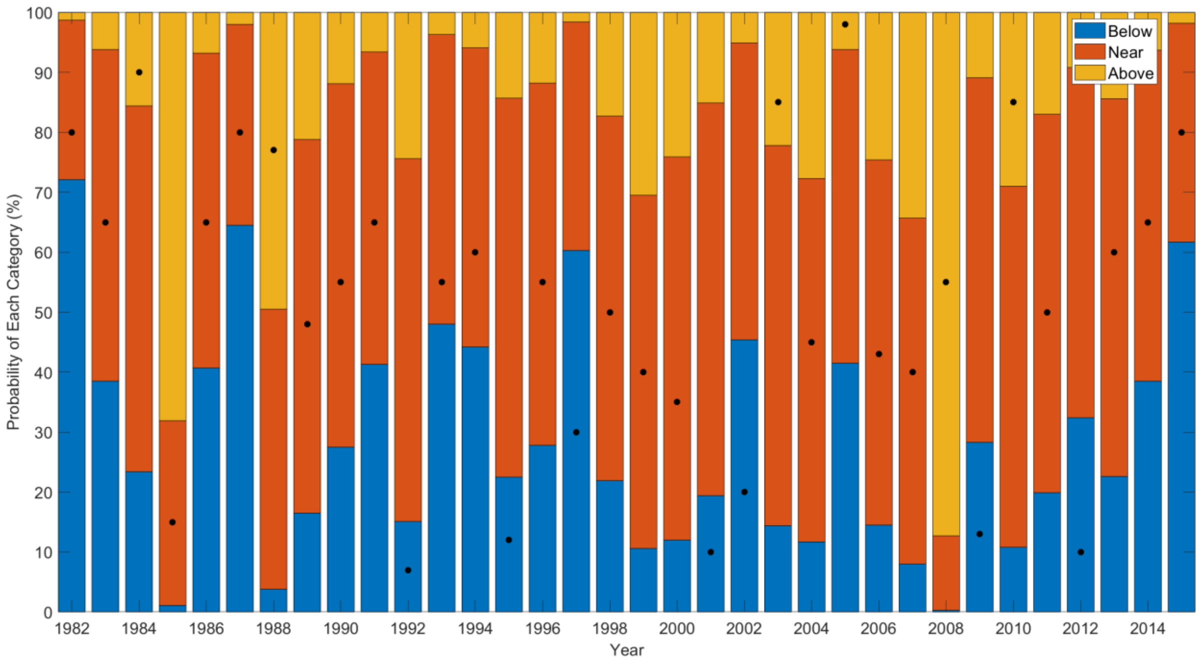

Figure 19: Categorical Forecast Strength for PCA+CV Basic Predictors Mode

Table 6: Contingency Table for Above Normal, Near Normal, or Below Normal

		Forecast		
		Below	Normal	Above
Observed	Below	1	6	0
	Normal	3	17	1
	Above	0	4	2

In Andes during the September-December season of interest, above and below normal streamflow values, i.e., floods are of primary concern to stakeholders, as are droughts, however rare. To gain an understanding of how the forecast model compares to climatology when predicting extremes, the high and low thresholds are set, such that the forecast ensembles can be tested for extreme flow prediction skill. Thresholds, create categories of flow conditions, that are commonly referred to as “bins”. Several bin constructs were explored, across two (above- and below-normal flow) and three bin methods (above-, near-, and below-normal flow) (Table 7 and Table 8).

Table 7: RPSS - Two Bin (*n* = 34)

Climatology	Below Normal	Above Normal	RPSS Basic	RPSS Extended	RPSS Top	RPSS Enso Only
Uniform	17	17	33.25%	22.91%	14.42%	15.91%
2-bin (high flow focus)	29	5	84.28%	66.55%	73.27%	74.24%
2-bin (low flow focus)	5	29	18.24%	1.97%	10.26%	21.33%

Table 8: RPSS – Three Bin (*n* = 34)

Climatology	Below Normal	Normal	Above Normal	RPSS Basic	RPSS Extended	RPSS Top	RPSS Enso Only
15 %	5	24	5	21.14%	7.55%	8.67%	14.38%
25%	9	16	9	-4.53%	-19.32%	-5.50%	-13.56%
Even	11	12	11	-1.82%	-13.03	-0.77%	-9.58%
3-bin histogram	10	18	6	15.23%	5.50%	12.99%	10.99%

When predicting above or below streamflow values the cross-validated models are better than climatology in all instances (Table 7). The basic predictor mode greatly outperforms all the other variable combinations and predicts above or below normal streamflow values better than climatology. This forecasting model has the capability to predict below above-normal streamflow values better than below-normal. This is advantageous considering low flow events during this season are rare, and again not drought-like.

The 3-bin combinations do not provide positive results in all instances; however, this analysis shows that this model is still helpful for predicting extremes (Table 8). Understanding how the model performs when predicting extreme streamflow events, such as droughts and flooding, is vital to creating a water allocation policy that mitigates climate uncertainty for the appropriate stakeholders. For example, in a 15% tail scenario (5 below, 24 near, 5 above), the cross-validated forecasting model performs better than climatology across all four predictor modes. This model does not perform better than climatology when looking at even bins or 25% of the tails, which is attributable to the fact that the streamflow observations are not uniformly distributed. Lastly, using the probability density function of the observed streamflow to categorize flow regimes (10 below, 18 near, 6 above), which is likely the most statistically accurate of the three-bin scenarios, all modes are skillful. This suggests that if streamflow observations follow the same general trend, i.e., more above-normal flows than below-normal, the model holds the potential to provide value, at least categorically.

There is a trade-off between model skill and forecast lead time (Fig. 20). This concept is very important to remember when providing stakeholders forecasts with adequate lead times. Shorter lead times will provide forecast models of higher skill, while longer lead times will produce forecasts of less skill. Until mitigation policy is made, longer lead times are preferred for the Andes community, especially since the results for the 15% tail scenario (5 below, 24 near, 5 above) perform better than climatology in all instances. As stakeholders feel comfortable with reducing the forecast lead time, models of higher skill can be created.

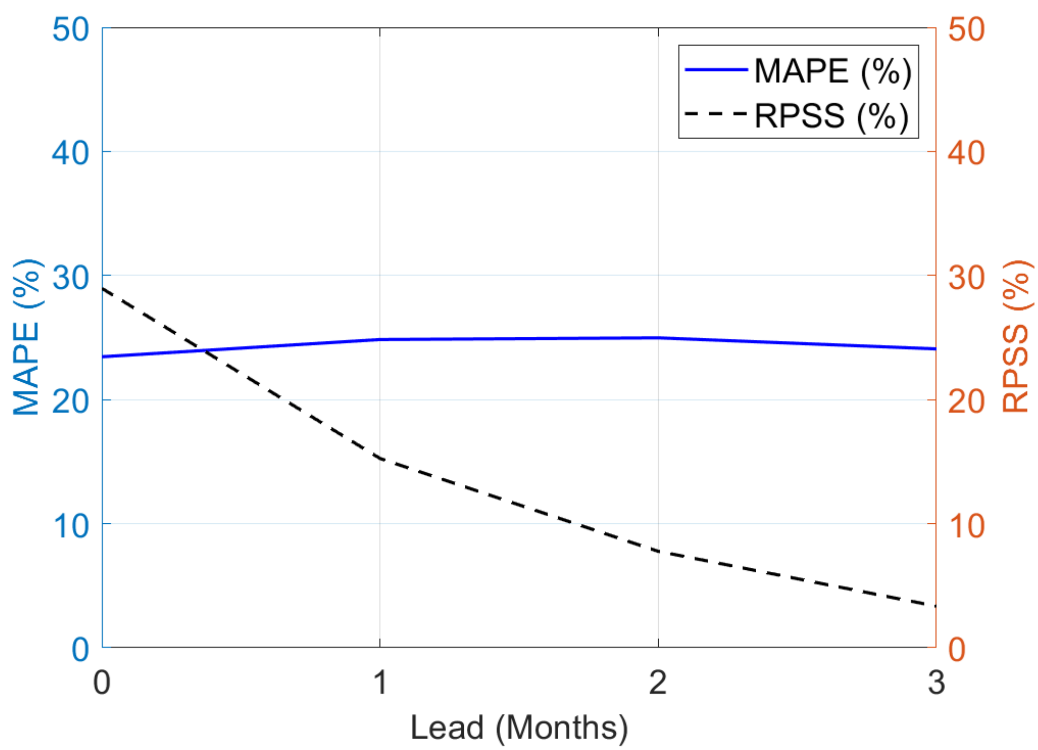

Figure 20: Model Skill vs. Lead Time for 15% tail scenario (5 below, 24 near, 5 above).

VI. Discussion

Chapter Overview

This chapter discusses the implications of the results in the context of Andes, analyzes the importance of implementing forecasting techniques to help the community address the vulnerabilities highlighted in the background, and reviews the research question initially proposed at the beginning of this work.

Review of Research Questions Initially Proposed

The research question analyzed in this work was written to address water resource availability and competitive use problems.

1. *Can an accurate forecast model be built to aid in local water sustainability discussions*?

Accurate statistical forecast models can be built to aid in the local water sustainability discussion for communities like Andes. The model built for the Andes community resulted in a model of reasonable quality according to the MAPE score. This may be able to be improved if localized soil moisture and precipitation data were made available in the area. However, this research shows that for small communities like Andes where reliable data is not available forecasts that perform better than climatology can still be produced using globally gridded data to aid in sustainability discussions.

Climate Uncertainty and Water Resources in Andes

The Andes community is susceptible to hydrologic extremes. In the case of drought, water availability for coffee farming, gold mining operations, fisheries, and population consumption would become extremely controversial and limited due to the number of intersectoral demands; and, in the more likely case of high flows certain sectors could be negatively impacted by landslides; rapid, river geomorphology, or inundations. Independent of extremes, the relationships between intersectoral water users will become increasingly important as climate shifts. Coffee farmers and the ASGM community serve as good examples since these are the two most practiced economic drivers in the area. Ideally, both communities desire enough water to continue their current operations and provide for their families. However, changing flow regimes will likely require collaboration.

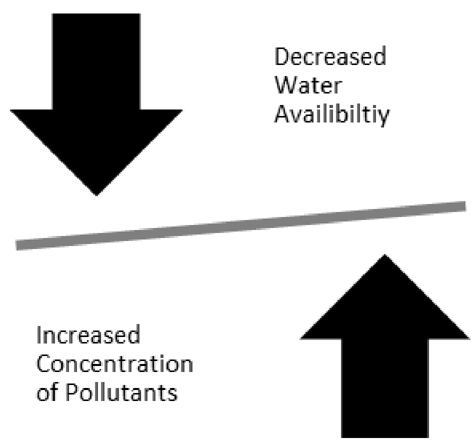

Figure 21: Relationship Between Water Availability and Quality

When water becomes scarce, its quality will also plummet as ASGM and coffee farming operations remain constant. These water quality issues do not only impact local members of the community but also millions of people downstream (Bedoya, 2009). For the ASGM community, it is vital for them to understand how their current operations impact water quality in the area and how they can help ensure the quality of their local water resources does not plummet due to the future intensification of drought.

Coffee farmers are at the highest risk when it comes to climate uncertainty and any slight changes in local weather patterns. This should be very concerning to stakeholders and researchers in the area since this industry currently makes up over 90 percent of the local economy. As previously discussed, coffee farmers are very dependent on ideal temperature and precipitation conditions to produce good quality crops; with any changes in local weather phenomena greatly reducing their crop yield. There are documented instances where local coffee farmers have expressed concern with noticeable changes in local weather patterns, mentioning longer periods of drought and intensified storm events. These changes might explain the decreased quality and yield coffee farmers in the local area are currently experiencing. As periods of drought intensify, irrigation could be used to meet crop-water demand, but only if sufficient surface water supply and quality are provided.

Supplemental irrigation from the San Juan River is not currently an option for coffee farmers due to rumored ASGM gold extraction and disposal techniques concerning the use of heavy metals such as mercury and cyanide. Coffee farming disposal techniques also threaten the quality of water in the San Juan, but further research needs to be done in this area to truly understand these impacts. If made available, using the San Juan River as a reliable source of irrigation for coffee farms during periods of extreme drought would greatly reduce the risk coffee farmers face from climate uncertainty. Focusing efforts on this singular industry and ensuring coffee farming remains stable through all the changes brought forth by a changing climate is one way to ensure the entire community of Andes continues to thrive into the future. Increased periods of drought are not the only concern for community members, but so is the increased frequency and intensity of storm events.

Instances of increased precipitation could cause decreased coffee yield and make the community susceptible to a higher rate of landslides. When analyzing the precipitation dataset using statistical analysis, the mean of the first ten years of precipitation data and the last ten years differ significantly. These results showcase there has been an increase in precipitation throughout Andes since 1979. If proper water management practices are not developed to allow adequate water runoff, this increased precipitation could greatly impact the coffee farming industry through overwatering coffee fields. Developing water retention systems during periods of high-intensity precipitation events could also lead to a safe alternative for a potential source of crop irrigation during periods of extreme drought.

Current literature has noted that the intensification of storms has been shown to increase the probability of landslide occurrence throughout mountainous zones, especially in communities with poor water management practices (Guidicini, 1977; Haque, 2019; Kirschbaum, 2020; Ward, 2020). Landslides are a serious concern for people of the Andes community and have the potential to impact all members of the community in a tremendous way. Further research can be done on the relationship between climate uncertainty and landslide mitigation in this area.

The Importance of Implementing Forecasting Techniques to Help Address Climate Uncertainty in Small Communities.

The skillful forecasts produced here hold the potential to inform actions of the Andes community, as they look to adapt to changing climate. Appropriately balancing lead time with forecast uncertainty can be a very difficult task for decision makers and modelers. As a forecast's lead is increased, i.e., the expanding gap between the forecast issue date and the forecast period, the skill of the forecast generally falls. Balancing lead and skill is a key consideration; longer lead times enable decision time, while on the other hand better information allows for more accurate preparations to be made.

If forecasting techniques, such as the one applied throughout this work, are to support the management of essential resources (i.e., water management) in a community, public involvement is crucial. There is a need for researchers to engage multiple levels of stakeholders in the area to facilitate awareness and education on water availability issues brought forth by climate uncertainty. Stakeholder dialog and participation can help ensure members are receiving the information they desire and shape the services provided to the community. However, discussing these types of climate issues presents a unique challenge, especially when trying to communicate forecasts and statistical predictions to community members in different countries.

There is an opportunity to link this type of forecasting modeling with ethnographic data from members of the Andes community. This linkage is essential to see if/why people of the area need this type of modeling, and how it could be integrated for community use. In October 2021, research colleagues conducted 31 interviews with over 40 local stakeholders including academics, geologists, engineers, ASG miners, farmers people participating in both farming and ASGM, workers at ore processing plants, aquaculture producers, apiarists, aqueduct managers, leaders of community organizations, employees of coffee cooperatives, environmental NGOs, historians, librarians, panela producers, and more. The information gained through these interviews will be vital to the successful implementation of forecast modeling into a community to directly aid in landslide mitigation and/or water allocation policy.

Models that are less complex have a better chance of being accepted by community members. For this reason, an ENSO Only model was developed for streamflow prediction in Andes. This simplified model mode explains nearly 60% of seasonal streamflow and outperforms climatology according to the RPSS metric. As such, a forecast of this type could be used to provide a simple categorical forecast. While the same forecast might not possess a sufficient deterministic skill to evoke uptake, the categorical accuracy could spur preparations, e.g., a forecast for above-normal streamflow could drive community preparations for flooding or landslide mitigation actions.

ENSO Only Forecasting Model

Having an understanding the current ENSO 3.4 phase and strength can reduce fear of climate uncertainty in community members throughout Andes. Based on PCA, water availability in Andes is mostly influenced by the ENSO 3.4 global phenomena. To build the MLR, PCA, and PCA+CV the seven unique variable combinations were used (Table 9). If all of these variables are kept in the model, there would be a zero-month lead due to the use of August data as a predictor. However, of these seven variables, Niño 3.4_May explained over 85% of the variability in streamflow for the SOND season of interest. This means that if any member of the community is concerned with the amount of water available during SOND, understanding the May ENSO cycle and intensity of that cycle can give decision-makers a strong indication of water availability in SOND.

Table 9: ENSO Only Predictor Variable Combinations

Variable	**Variable Type**	**Variable Description**	**Lead Time**	**This is Random Section Now**
1	ENSO	Niño 3.4_May	3-Month Lead	Mean value, 1982-2015
2		Niño 3.4 _Jun	2-Month Lead	
3		Niño 3.4 _Jul	1-Month Lead	
4		Niño 3.4 _Aug	0-Month Lead	
5		Niño 3.4 _MJJA	0-Month Lead	
6		Niño 3.4 _MJJ	1-Month Lead	
7		Niño 3.4 _MJ	2-Month Lead	

Using only ENSO 3.4 information, the PCA+CV model created is better at predicting above or below normal streamflow conditions in comparison to climatology.

The model can also explain streamflow during SOND with reasonable quality, i.e., MAPE = 24.52 (Table 5). Streamflow and precipitation amounts experienced in the Andes community are related to the current ESNO cycle and strength (Fig. 9); with extremes being associated with increased ENSO phase intensity.

Due to this relationship with streamflow in Andes, it is important to understand how the ENSO 3.4 variable could be impacted by climate change. Positive ENSO anomaly is negatively related to precipitation and streamflow, meaning, during a positive phase of ENSO 3.4 (El Niño) there is less precipitation and streamflow available throughout the community in Andes. Research has recently been completed to analyze how climate change will impact the variability of sea surface temperature in the Niño 3.4 region (Fig. 22). This research has shown that sea surface temperatures are expected to increase in variability throughout the Niño 3.4 region through the year 2100. For Andes, this projection is not positive, especially since this increase in Niño 3.4 variability will directly impact the strength of extremes in the Andes region (Fig. 9). During positive Niño 3.4 events, precipitation and water supply in the area are going to decrease, while the opposite is true of negative events. This finding makes advocacy for landslide mitigation and water policies including distribution, storage, and quality of water of extreme importance for future community readiness in Andes.

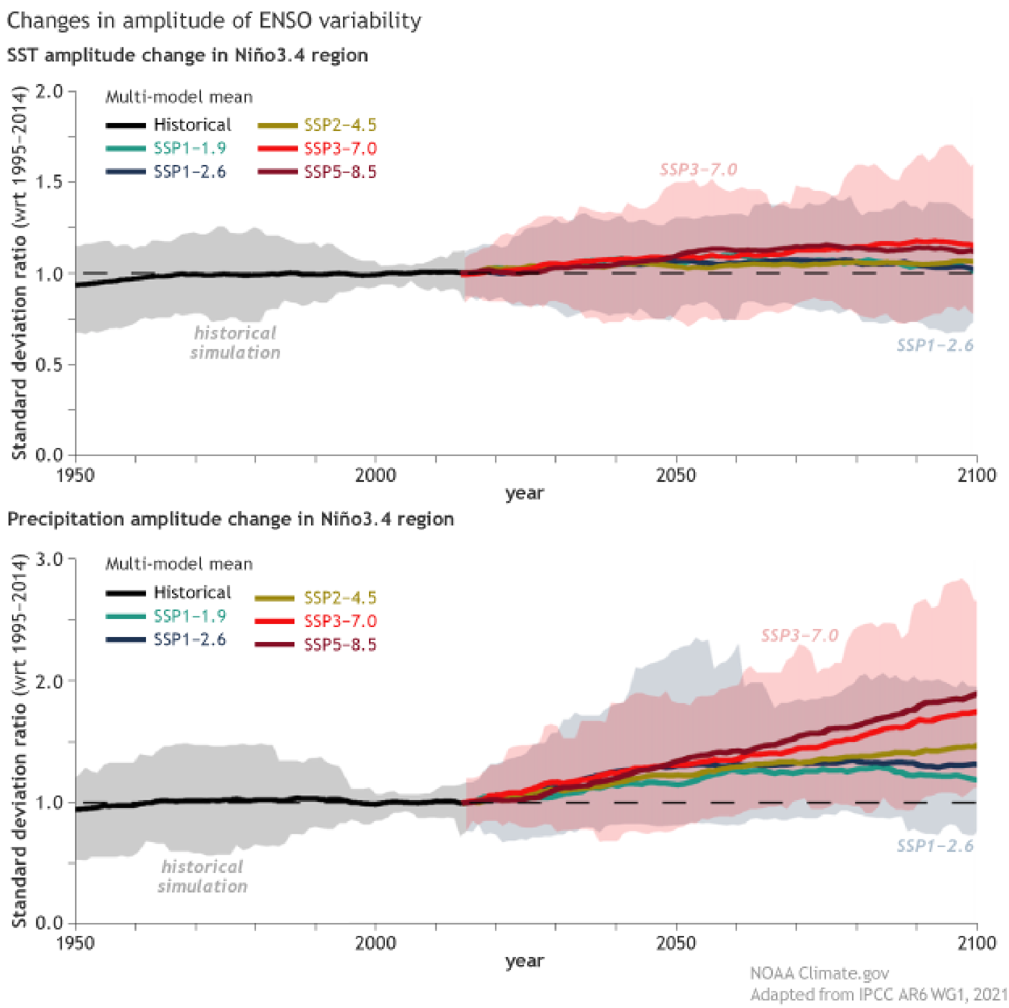

Figure 22: Niño 3.4 and Precipitation Predictions. Adopted from: NOAA Climate.gov.

Even though water distribution policy and instances of extreme drought may be a concern in the late twentieth century, landslide risk poses a high threat right now. The statistical analysis, and MLR, PCA, and PCA+CV models suggest intensification is likely to make high flows, i.e., landslides, more of a current issue in Andes. While higher flows might alleviate supply and some quality concerns, it will require different collaborations to ensure economic and social productivity.

VII. Conclusion

Chapter Overview

This chapter reviews the findings, discusses the significance, and provides insight into areas of future research.

Review of Findings

Forecasting can be a very helpful tool to facilitate conversations about local water policy and conservation efforts. Data have shown that the magnitude and frequency of extreme events have a direct impact on the amount of water available in the local community of Andes. To deconflict areas of future conflict when it comes to water resources, hindcast forecast modeling can be accurately used to grasp an understanding of how local water availability has been impacted in the past. Based on these models, forward-looking forecasts can be created to aid in water resource management adjusted for climate change. Understanding the current phase and strength of the Niño 3.4 cycle can greatly reduce uncertainty in local precipitation and streamflow patterns.

The biggest challenge when using forecasts to understand climate uncertainty is engaging local stakeholders. Coffee farmers make up the largest percentage of the local economy and are the most climate-vulnerable group. So far, much of the local discourse and research has been focused on the ASGM community, its environmental impacts, and potential remediation techniques. These issues are important; however, due to the economic risk coffee farming destabilization brings to the community, these concerns of climate uncertainty and water policy need to be addressed promptly. The best solution going forward is to engage all stakeholders, facilitate awareness and education on water resource management, and to support open dialog across many levels throughout the community such that all members can help shape local water management and sustain their ways of life into the future.

Significance of Research

This study was a win in both the testing of forecasting techniques and in providing modeling tools to aid the Andes community in reducing their current vulnerability to climate uncertainty when it comes to water resources. This research helps the individual community members and other researchers working in Andes by locating viable data. Before this research was completed, other researchers working in the area were unaware of any reliable streamflow, precipitation, or soil moisture data. This analysis shows that the use of globally gridded data is a reliable option of use for other researchers working in this area. This research can also give members of the Andes who are concerned with how climate change is going to impact their current operations a little more security if they know the current ENSO cycle phase and strength. This information alone is a strong predictor of how streamflow and precipitation in the area will be influenced.

Future Research

This research is an advancement for researchers working in Andes, Antioquia, Colombia on environmental pollution and remediation projects. Without data or a basic understanding of how water will be impacted due to climate uncertainty, remediation of pollutants in the San Juan River or water allocation policies amongst intersectoral communities are near impossible. Even though this research provides a good starting point, it can still be improved. This model is built on hindcast data, but this process could be improved by making forward-looking forecasts to see how streamflow will be impacted in the future due to a changing climate. Also, if the community works to set up more reliable methods of data collection, the forecasting models completed in this analysis could be re-accomplished to see if using localized precipitation and soil moisture data makes a difference in the forecasting quality and prediction capability. Another option for future research in relation to the forecast modeling completed in this research is linking the outputs with sectoral models for the area. Lastly, research on how to communicate this information and important findings to members of the community in the right way can be completed.

Appendix

Table 10: Variable Description for Expanded Predictor Mode

Variable	Variable Type	Variable Title	Variable Description
1	ENSO	Niño 3.4_May	Mean value, 1982-2015
2		Niño 3.4 _Jun	
3		Niño 3.4 _Jul	
4		Niño 3.4 _Aug	
5		Niño 3.4 _MJJA	
6		Niño 3.4 _MJJ	
7		Niño 3.4 _MJ	
8	SOI	SOI_May	
9		SOI _Jun	
10		SOI _Jul	
11		SOI _Aug	
12		SOI _MJJA	
13		SOI _MJJ	
14		SOI _MJ	
15	Precipitation	Precip_May	Globally gridded, Mean value, 1982-2015
16		Precip _Jun	
17		Precip _Jul	
18		Precip _Aug	
19		Precip _MJJA	
20		Precip _MJJ	
21		Precip _MJ	
22	Soil Moisture	SM_May	
23	Difference: ENSO - SOI	Dif_May	Difference of Mean Values
24		Dif_Jun	
25		Dif_Jul	
26		Dif_Aug	
27		Dif_MJJA	
28		Dif_MJJ	
29		Dif_MJ	

Table 11: Variable Description for Top Predictor Mode

Variable	Variable Type	Variable Title	Variable Description
1	ENSO	Niño_MJJ	Mean value, 1982-2015
2	SOI	SOI_Aug	Mean value, 1982-2015
3	Precipitation	PRCP_Aug	Globally gridded, Mean value, 1982-2015
4	Soil Moisture	SM_May	Globally gridded, Mean value, 1982-2015

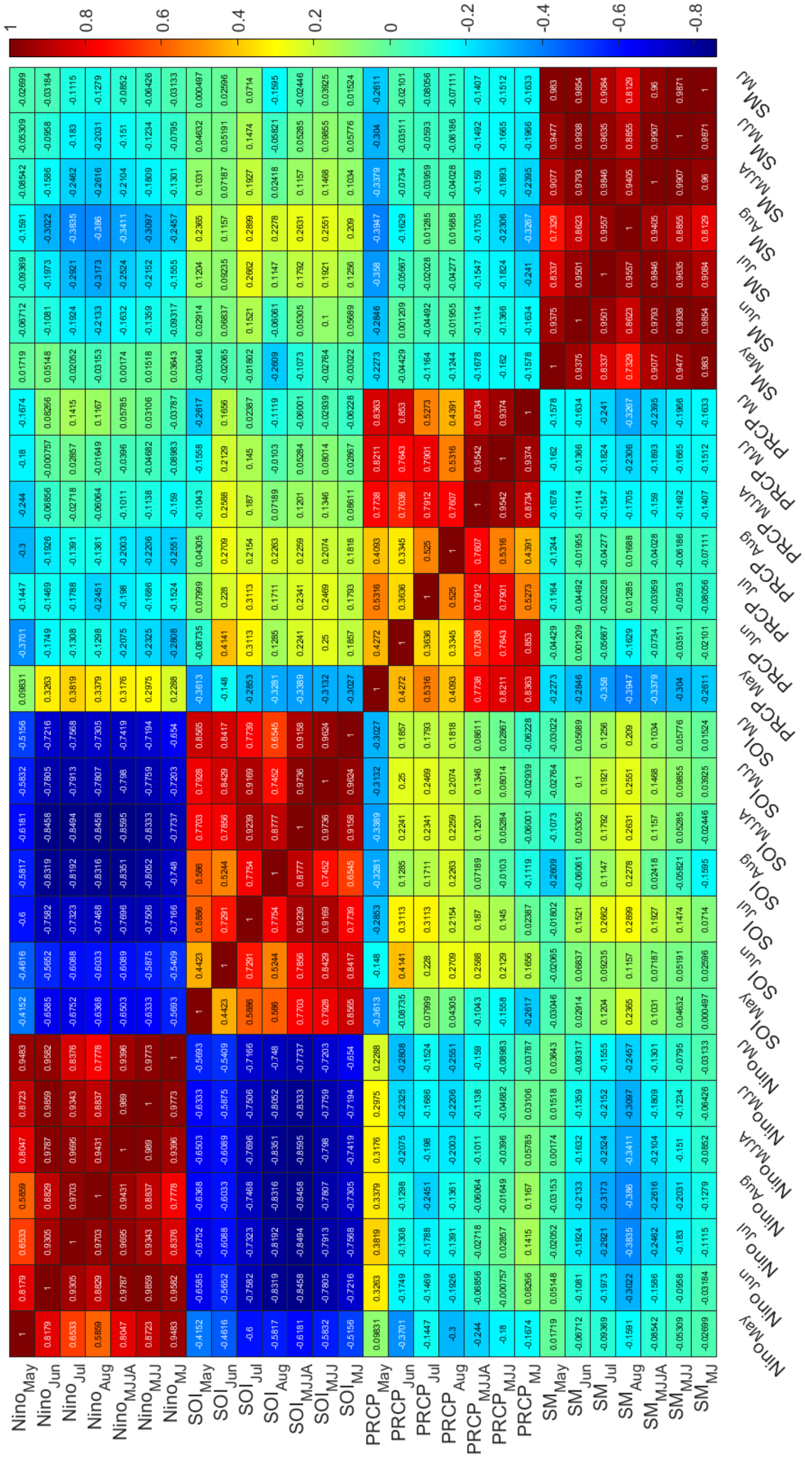

Figure 23: Heatmap Showing Correlation Between Variables for Basic Predictor Mode

Bibliography

Aksoy, H. (2000). Use of gamma distribution in hydrological analysis. *Turkish Journal of Engineering and Environmental Sciences*, *24*(6), 419-428.

Alexander, S., Yang, G., Addisu, G., & Block, P. (2021). Forecast-informed reservoir operations to guide hydropower and agriculture allocations in the Blue Nile basin, Ethiopia. *International Journal of Water Resources Development*, *37*(2), 208-233.

Allan, R. P., and B. J. Soden (2008), Atmospheric warming and the amplification of precipitation extremes, Science, 321(5895), 1481–1484.

Allison, E. H., Perry, A. L., Badjeck, M. C., Neil Adger, W., Brown, K., Conway, D., ... & Dulvy, N. K. (2009). Vulnerability of national economies to the impacts of climate change on fisheries. *Fish and fisheries*, *10*(2), 173-196.

Balmaseda, M. A., Davey, M. K., & Anderson, D. L. (1995). Decadal and seasonal dependence of ENSO prediction skill. *Journal of Climate*, *8*(11), 2705-2715.

Bazo, J., Singh, R., Destrooper, M. and Coughlan de Perez, E.: Pilot experiences in using seamless forecasts for early action: The "ready-set-go!" approach in the Red Cross, in Sub-seasonal to Seasonal Prediction, pp. 387–398, Elsevier., 2019.

Bedoya, J., & Martinez, E. (2009). Air quality in the Aburrá Valley Antioquia-Colombia. Dyna, 76(158), 7-15.

Beyene, A., Kassahun, Y., Addis, T., Assefa, F., Amsalu, A., Legesse, W., ... & Triest, L. (2012). The impact of traditional coffee processing on river water quality in Ethiopia and the urgency of adopting sound environmental practices. *Environmental monitoring and assessment*, *184*(11), 7053-7063.

Block, P. and Rajagopalan, B.: Statistical – dynamical approach for streamflow modeling at Malakal, Sudan, on the White Nile River, J. Hydrol. Eng., 185–196, doi:10.1061/(ASCE)1084-0699(2009)14:2(185), 2009.

Block, P. J., Souza Filho, F. A., Sun, L., & Kwon, H. H. (2009). A streamflow forecasting framework using multiple climate and hydrological models 1. *JAWRA Journal of the American Water Resources Association*, *45*(4), 828-843.

Bombardi, R. J., Carvalho, L. M., Jones, C., & Reboita, M. S. (2014). Precipitation over eastern South America and the South Atlantic Sea surface temperature during neutral ENSO periods. *Climate Dynamics*, *42*(5-6), 1553-1568.

Bosselmann, A. S., Dons, K., Oberthur, T., Olsen, C. S., Ræbild, A., & Usma, H. (2009). The influence of shade trees on coffee quality in small holder coffee agroforestry systems in Southern Colombia. *Agriculture, ecosystems & environment*, *129*(1-3), 253-260.

Cai, X. (2008). Water stress, water transfer and social equity in Northern China—Implications for policy reforms. *Journal of Environmental Management*, *87*(1), 14-25.

Célleri, R., Buytaert, W. O. U. T. E. R., De Bièvre, B. E. R. T., Tobón, C., Crespo, P., Molina, J., & Feyen, J. (2010). Understanding the hydrology of tropical Andean ecosystems through an Andean Network of Basins. *IAHSAISH Publication*, *336*, 209-212.Chambers, J. C., Mullick, S. K., & Smith, D. D. (1971). *How to choose the right forecasting technique*. Harvard University, Graduate School of Business Administration.

Chen, H. C., Tseng, Y. H., Hu, Z. Z., & Ding, R. (2020). Enhancing the ENSO predictability beyond the spring barrier. *Scientific reports*, *10*(1), 1-12.

Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., & Palmer, R. N. (2004). The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic change*, *62*(1), 337-363.

Cutter, Susan L. & Sahar Derakhshan (2020) Temporal and spatial change in disaster resilience in US counties, 2010–2015, Environmental Hazards, 19:1, 10-29, DOI: 10.1080/17477891.2018.1511405

Davis AP, Gole TW, Baena S, Moat J (2012) The impact of climate change on indigenous Arabica coffee (Coffea arabica): predicting future trends and identifying priorities. PLoS ONE 7:e47981. https://doi.org/10.1371/journal.pone.0047981

DaMatta FM et al (2016) Sustained enhancement of photosynthesis in coffee trees grown under free-air CO2 enrichment conditions: disentangling the contributions of stomatal, mesophyll, and biochemical limitations. J Exp Bot 67:341–352. https://doi.org/10.1093/jxb/erv463

Delorit, J., E. C. G. Ortuya, and P. Block. (2017). "Evaluation of model-based seasonal streamflow and water allocation forecasts for the Elqui Valley, Chile," Hydrology and Earth System Sciences; Katlenburg-Lindau, vol. 21, no. 9, pp. 4711–4725, doi: http://dx.doi.org.afit.idm.oclc.org/10.5194/hess-21-4711-2017

Didier, M., Broccardo, M., Esposito, S., & Stojadinovic, B., (2018) A compositional demand/supply framework to quantify the resilience of civil infrastructure systems (Re-CoDeS), Sustainable and Resilient Infrastructure, 3:2, 86-102, DOI: 10.1080/23789689.2017.1364560

Ficke, A. D., Myrick, C. A., & Hansen, L. J. (2007). Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries, 17*(4), 581-613.

Fischer, G., Shah, M., van Velthuizen, H., 2002. Climate Change and Agricultural Vulnerability. International Institute of Applied Systems Analysis, Vienna.

Fowler, H. J., & Ekström, M. (2009). Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *29*(3), 385-416.

Frank, E., Eakin, H., & López-Carr, D. (2011). Social identity, perception and motivation in adaptation to climate risk in the coffee sector of Chiapas, Mexico. *Global environmental change*, *21*(1), 66-76.

Gay, C., Estrada, F., Conde, C., Eakin, H., & Villers, L. (2006). Potential impacts of climate change on agriculture: a case of study of coffee production in Veracruz, Mexico. *Climatic Change*, *79*(3), 259-288.

GOA. (2020). Climate Resilience. *United States Government Accountability Office*. GAO-21-46.

Goto, Y. B., & Fukunaga, E. T. (1956). Coffee: harvesting and processing for top quality coffee.

Gottesfeld, P., Tirima, S., Anka, S.M., Fotso, A., Nota, M.M., 2019. Reducing lead and silica dust exposures in small-scale mining in Northern Nigeria. Ann. Work Expo. Health 63, 1–8. https://doi.org/10.1093/annweh/wxy095.Guidicini, G., & Iwasa, O. Y. (1977). Tentative correlation between rainfall and landslides in a humid tropical environment. *Bulletin of the International Association of Engineering Geology-Bulletin de l'Association Internationale de Géologie de l'Ingénieur*, *16*(1), 13-20.

Guidicini, G., & Iwasa, O. Y. (1977). Tentative correlation between rainfall and landslides in a humid tropical environment. *Bulletin of the International Association of Engineering Geology-Bulletin de l'Association Internationale de Géologie de l'Ingénieur*, *16*(1), 13-20.

Hamlet, A. F., Huppert, D., & Lettenmaier, D. P. (2002). Economic value of long-lead streamflow forecasts for Columbia River hydropower. *Journal of Water Resources Planning and Management*, *128*(2), 91-101.

Hapuarachchi, H. A. P., Kabir, A., Zhang, X. S., Kent, D., Bari, M. A., Tuteja, N. K., ... & Ahmad, Z. (1815). Performance evaluation of the national 7-day water forecast service. In *Proceedings of 22nd International Congress on Modelling and Simulation* (Vol. 1821).

Harriman, L.: Cyclone Phailin in India: Early warning and timely actions saved lives, Environ. Dev., 9, 93–100, doi:10.1016/j.envdev.2013.12.001, 2014.

Harrod, C., Ramírez, A., Valbo-Jørgensen, J., & Funge-Smith, S. (2019). How climate change impacts inland fisheries. *Impacts of climate change on fisheries and aquaculture*, 375.

Haque, U., Da Silva, P. F., Devoli, G., Pilz, J., Zhao, B., Khaloua, A., ... & Glass, G. E. (2019). The human cost of global warming: Deadly landslides and their triggers (1995–2014). *Science of the Total Environment*, *682*, 673-684.

Hellmuth, M. E., Moorhead, A., Thomas, M. C., & Williams, J. (2007). Climate risk management in Africa: Learning from practice.

Hitz, S., & Smith, J. (2004). Estimating global impacts from climate change. *Global Environmental Change*, *14*(3), 201-218.

Informacion ´ del Municipio [WWW Document], 2018. Alcaldia Andes (accessed 20 May 2020). http://www.andes-antioquia.gov.co/MiMunicipio/Paginas/Informacion-de l-Municipio.aspx

Jayakumar M, Rajavel M, Surendran U, Gopinath G, Ramamoorthy K (2017) Impact of climate variability on coffee yield in India—with a micro-level case study using long-term coffee yield data of humid tropical Kerala. Clim Chang 145:335–349. https://doi.org/10.1007/s10584-017-2101-2

Jimenez Cisneros, B. E., T. Oki, N. W. Arnell, G. Benito, J. G. Cogley, P. Doll, T. Jiang, and S. S. Mwakalila (2014), Freshwater resources, in € Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change, edited by C. B. Field et al., pp. 229–269, Cambridge Univ. Press, Cambridge, U. K.

Johansen, C., Horney, J., & Tien, I. (2017). Metrics for evaluating and improving community resilience. Journal of Infrastructure Systems, 23(2), 04016032.

Joshi, H. C., & Sukumaran, P. K. (1991). Water pollution investigation in the River Tungabhadra near Harikar, Karnataka. *Indian Journal of Animal Sciences*, *61*(2), 230-234.

Kabir, A., Hasan, M., Hapuarachchi, P., Zhang, S., Liyanage, J., Gamage, N., Laugesen, R., Plastow, K., MacDonald, A., Bari, M., Tuteja, N., Robertson, D., Shrestha, D., Bennett, J.C., 2018. Evaluation of multi-model rainfall forecasts for the national 7-day ensemble streamflow forecasting service, 2018. In: Hydrology and Water Resources Symposium (HWRS 2018): Water and Communities. Engineers Australia, Melbourne, pp. 393–406. In: Hydrology and Water Resources Symposium (HWRS 2018): Water and Communities. Melbourne: Engineers Australia, 2018: 393- 406.

Keating, C., Lee, D., Bazo, J., & Block, P. (2021). Leveraging multi-model season-ahead streamflow forecasts to trigger advanced flood preparedness in Peru. *Natural Hazards and Earth System Sciences*, *21*(7), 2215-2231.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., & Senior, C. A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, *4*(7), 570-576.

Khan, U., Cook, F. J., Laugesen, R., Hasan, M. M., Plastow, K., Amirthanathan, G. E., ... & Tuteja, N. K. (2020). Development of catchment water quality models within a realtime status and forecast system for the Great Barrier Reef. *Environmental Modelling & Software*, *132*, 104790.

Kirschbaum, D., Kapnick, S. B., Stanley, T., & Pascale, S. (2020). Changes in extreme precipitation and landslides over High Mountain Asia. *Geophysical Research Letters*, *47*(4), e2019GL085347.

Kurunç, A., Yürekli, K., & Cevik, O. (2005). Performance of two stochastic approaches for forecasting water quality and streamflow data from Yeşilırmak River, Turkey. *Environmental Modelling & Software*, *20*(9), 1195-1200.

Lin BB (2007) Agroforestry management as an adaptive strategy against potential microclimate extremes in coffee agriculture. Agric For Meteorol 144:85–94. https://doi.org/10.1016/j.agrformet.2006.12.009

Mason, S. J., & Stephenson, D. B. (2008). How do we know whether seasonal climate forecasts are any good?. In *Seasonal Climate: Forecasting and Managing Risk* (pp. 259-289). Springer, Dordrecht.

Mortensen, E., Wu, S., Notaro, M., Vavrus, S., Montgomery, R., De Piérola, J., ... & Block, P. (2018). Regression-based season-ahead drought prediction for southern Peru conditioned on large-scale climate variables. *Hydrology and earth system sciences*, *22*(1), 287-303.

Murtinho, F., Tague, C., De Bièvre, B., Eakin, H., & Lopez-Carr, D. (2013). Water scarcity in the Andes: a comparison of local perceptions and observed climate, land use and socioeconomic changes. *Human ecology*, *41*(5), 667-681.

Naranjo-Merino, C. A., Ortíz-Rodriguez, O. O., & Villamizar-G, R. A. (2018). Assessing green and blue water footprints in the supply chain of cocoa production: a case study in the northeast of Colombia. *Sustainability*, *10*(1), 38.

Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F. F., Wakata, Y., Yamagata, T., & Zebiak, S. E. (1998). ENSO theory. *Journal of Geophysical Research: Oceans*, *103*(C7), 14261-14290.

O'Brien, R. M., Phelan, T. J., Smith, N. M., & Smits, K. M. (2021). Remediation in developing countries: A review of previously implemented projects and analysis of stakeholder participation efforts. *Critical Reviews in Environmental Science and Technology*, *51*(12), 1259-1280.

O'Connor, R. E., Bard, R. J., & Fisher, A. (1999). Risk perceptions, general environmental beliefs, and willingness to address climate change. *Risk analysis*, *19*(3), 461-471.

Olmstead, S. M. (2010). The economics of managing scarce water resources. *Review of Environmental Economics and policy*, *4*(2), 179-198.

Olmstead, S. M. (2014). Climate change adaptation and water resource management: A review of the literature. *Energy Economics*, *46*, 500-509.

Pham, Y., Reardon-Smith, K., Mushtaq, S., & Cockfield, G. (2019). The impact of climate change and variability on coffee production: a systematic review. *Climatic Change*, *156*(4), 609-630.

Rajaee, M., Obiri, S., Green, A., Long, R., Cobbina, S., Nartey, V., Buck, D., Antwi, E., Basu, N., 2015. Integrated assessment of artisanal and small-scale gold mining in Ghana—part 2: natural sciences review. Int. J. Environ. Res. Public Health 12, 8971–9011. https://doi.org/10.3390/ijerph120808971.Regonda, S. K., Rajagopalan, B., & Clark, M. (2006). A new method to produce categorical streamflow forecasts. *Water resources research*, *42*(9).

Rød, B., Lange, D., Theocharidou, M., & Pursiainen, C. (2020). From risk management to resilience management in critical infrastructure. Journal of Management in Engineering, 36(4), 04020039.

Rosenzweig, C., Hillel, D., 1998. Climate Change and the Global Harvest: Potential Impacts of the Greenhouse Effect on Agriculture. Oxford University Press, New York

Sabzi, H. Z., King, J. P., & Abudu, S. (2017). Developing an intelligent expert system for streamflow prediction, integrated in a dynamic decision support system for managing multiple reservoirs: A case study. *Expert systems with applications*, *83*, 145-163.

Sadoff, C., and M. Muller (2009), Water Management, Water Security and Climate Change Adaptation: Early Impacts and Essential Responses, Global Water Partnership, Stockholm.

Salaam-Blyther, T. (2012). Global access to clean drinking water and sanitation: US and international programs.

Schwartz, M., Smits, K., Smith, J., Phelan, T., & Baena, O. J. R. (2021). Incorporating positive deviance into comprehensive remediation projects: A case study from artisanal and small-scale gold mining in the municipality of Andes, Colombia. *Environmental Science & Policy*, *123*, 142-150.

Smith, J.B., Schellnhuber, H.J., Mirza, M.Q., Fankhauser, S., Leemans, R., Erda, L., Ogallo, L., Pittock, B., Richels, R., Rosenzweig, C., Safriel, U., Tol, R.S.J., Weyant, J., Yohe, G., 2001. Vulnerability to climate change and reasons for concern: a synthesis. In: McCarthy, J., Canziana, O., Leary, N., Dokken, D., White, K. (Eds.), Climate Change 2001: Impacts, Adaptation, and Vulnerability. Cambridge University Press, New York, pp. 913–967.

Smits, K. M., O'Brien, R. M., Smith, N., & Phelan, T. J. (2020, December). Integrating scientific and local knowledge to enhance remediation efficacy using community-informed conceptual site models. In *AGU Fall Meeting Abstracts* (Vol. 2020, pp. H031-0010).

Solomon, S., Manning, M., Marquis, M., & Qin, D. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC* (Vol. 4). Cambridge university press.

Souza Filho, F.A. and U. Lall, 2003. Seasonal to Interannual Ensemble Streamflow Forecasts for Ceara, Brazil: Applications of a Multivariate, Semiparametric Algorithm. Water Resources Research 39(11):1307, doi: 10.1029/2002WR001373.

Stakhiv, E. Z. (1998). Policy implications of climate change impacts on water resources management. *Water Policy, 1*(2), 159-175.

Tavares PD, Giarolla A, Chou SC, Silva AJD, Lyra AD (2018) Climate change impact on the potential yield of Arabica coffee in Southeast Brazil. Reg Environ Chang 18:873–883. https://doi.org/10.1007/s10113-017-1236-z

TCI (2016) A brewing storm: the climate change risks to coffee. The Climate Institute. http://www.limateinstitute.org.au/coffee.html. Accessed 05/07/2019

Tirima, S., Bartrem, C., von Lindern, I., von Braun, M., Lind, D., Anka, S.M., Abdullahi, A., 2016. Environmental remediation to address childhood lead poisoning epidemic due to artisanal gold mining in Zamfara, Nigeria. Environ. Health Perspect. 124, 1471–1478. https://doi.org/10.1289/ehp.1510145.

Trenberth, K. E., Caron, J. M., Stepaniak, D. P., & Worley, S. (2002). Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures. *Journal of Geophysical Research: Atmospheres*, *107*(D8), AAC-5.

UN Environment, 2019. Global Mercury Assessment 2018. Geneva, Switzerland. https://www.unep.org/resources/publication/global-mercury-assessment-2018 (accessed 12 Dec 2021).

Vargas-Pineda, O. I., Trujillo-González, J. M., & Torres-Mora, M. A. (2020). Supply–demand of water resource of a basin with high anthropic pressure: case study Quenane-Quenanito Basin in Colombia. *Air, Soil and Water Research*, *13*, 1178622120917725.

Vizy, E. K. and Cook, K. H. (2007). Relationship between Amazon and high Andes rainfall, J. Geophys. Res.-Atmos., 112, D07107, https://doi.org/10.1029/2006JD007980.

Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., Ruiter, M. C. D., Duncan, M. J., ... & Winsemius, H. C. (2020). Natural hazard risk assessments at the global scale. *Natural Hazards and Earth System Sciences*, *20*(4), 1069-1096.

Wilkinson, E., Weingärtner, L., Choularton, R., Bailey, M., Todd, M., Kniveton, D. and Venton, C. C.: Forecasting hazards, averting disasters. Implementing forecast-based early action at scale, London., 2018.

Yan, H., Sun, L., Wang, Y., Huang, W., Qiu, S., & Yang, C. (2011). A record of the Southern Oscillation Index for the past 2,000 years from precipitation proxies. *Nature Geoscience*, *4*(9), 611-614.

Zhu, J., Wu, P., Chen, H., Zhou, L., & Tao, Z. (2018). A hybrid forecasting approach to air quality time series based on endpoint condition and combined forecasting model. *International journal of environmental research and public health*, *15*(9), 1941.

Zimmerman, B. G., Vimont, D. J., & Block, P. J. (2016). Utilizing the state of ENSO as a means for season-ahead predictor selection. *Water resources research*, *52*(5), 3761-3774.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 074-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
24-03-2022	Master's Thesis	September 2020 – March 2022

TITLE AND SUBTITLE	
The Impacts of Climate Uncertainty on Streamflow in Andes, Antioquia, Colombia.	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Roberts, Kristen R., 2nd Lt, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENY) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865	AFIT-ENV-MS-21-M-252
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFCEC
Air Force Civil Engineering Center 2261 Hughes Ave, Ste.155 JBSA Lackland, TX 78236-9853	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Natural hazards, such as hurricanes, wildfires, floods, and droughts impact human systems that rely on predictable patterns in the natural elements with which they interact. These events threaten communities everywhere, and humanity continually seeks to adapt. Skillful prediction of the impacts of climate change on linked, human-natural systems, like surface water resources, can help ensure physical risks within vulnerable communities are mitigated, resource sustainability is maximized, and intersectoral markets continue to contribute to socioeconomic stability. Due to water resources being a primary conduit through which climate uncertainty impacts people, economies, and ecosystems, its study is worthy of investigation; particularly, where those resources are uncertain and demanded by a variety of competitive users. This work evaluates a season-ahead statistical prediction model of growing season streamflow (September – December) in Andes, Antioquia, Colombia, against a suite of global and local predictor variables: precipitation, soil moisture, Niño 3.4 sea-surface temperature anomaly, and Southern Oscillation Index anomaly. Skillful results, which are defined as streamflow forecasts that outperform a specified climatological baseline, are produced for the models when analyzing extreme streamflow events (r^2 = 0.77, mean absolute percentage error = 21.87, ranked probability skill score = 0.21). Even a lean model, consisting of just Niño 3.4 as a predictor, produces skillful results (r^2 = 0.37, mean absolute percentage error = 21.98, ranked probability skill score = 0.087). Viewed cumulatively, these results suggest streamflow predictions and forecasts can identify the role of global and local climate on communities, inform how and when changes should be implemented, and justify stakeholder decisions.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Lt Col Justin D. Delorit, AFIT/ENV
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	95	(937) 255-3636, ext. 4826 (justin.delorit@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18