University of Sussex

Forecasting Vegetation Condition in Pastoral Communities for Disaster Prevention

Edward Efui Salakpi

Candidate Number:191046

Supervisors: Seb Oliver, Pedram Rowhani

Submitted for the degree of Doctor of Philosophy

University of Sussex

Declaration

I hereby declare that this PhD thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree. This thesis is written in the papers based style, as approved by the University.

CHAPTER 3: Barrett, A. B., Duivenvoorden, S., Salakpi, E. E., Muthoka, J. M., Mwangi, J., Oliver, S., & Rowhani, P. (2020). **Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya**. Remote Sensing of Environment, 248, 111886. My contributions to this paper included, acquiring and preprocessing the MODIS data, testing various gap filling methods on MODIS, writing text in sections 3.2, 3.3.1.2, 3.5, 3.8, and 3.9.3, figures 3.1, 3.2, 3.6 and 3.9. I also contributed to the development of the whole project through discussions and commented on the whole paper.

CHAPTER 4: Salakpi, E. E., Hurley, P. D., Muthoka, J. M., Barrett, A. B., Bowell, A., Oliver, S., & Rowhani, P. (2021). **Forecasting Vegetation Condition with a Bayesian Auto-regressive Distributed Lags (BARDL) Model**. Natural Hazards and Earth System Sciences Discussions, 1-31 (In preprint). I was the lead author for this paper and led to all aspects including data preprocessing, modelling and writing. The other authors helped with access to data, providing guidance, comments and feedback.

CHAPTER 5: Salakpi, E. E., Hurley, P. D., Muthoka, J. M., Barrett, A. B., Bowell, A., Oliver, S., & Rowhani, P. (2021). **A Dynamic Hierarchical Bayesian Approach for Forecasting Vegetation Conditions**. Natural Hazards and Earth System Sciences (In preprint). I was the lead author for this paper and led to all aspects of the research including data preprocessing, modelling and writing. The other authors helped with access to data, providing guidance, comments and feedback.

Signature:

Edward Efui Salakpi

Project Summary

Drought is a slow occurring natural hazard that is known to be very complex. On average, drought events affect the livelihood of approximately 55 million people worldwide, with a large proportion from Africa. Addressing the challenges associated with drought requires understanding the drought categories and the factors that influence their occurrence. These categories include Meteorological drought, Hydrological drought, Agricultural drought, and Socio-Economic drought. Of all these, agricultural drought stands out due to its direct impact on people's livelihoods. This category of drought can adversely affect wildlife habitats, agriculture production, food security and the economy of the affected country or region. Governments and policymakers have explored early warning strategies that speculate the onset of drought and its severity. When in place, these strategies will help meet the United Nations' Sustainable Development Goals (SGDs) relating to food security. A major challenge with these strategies was that they were expensive to maintain, and drought forecasts were mainly based on expert judgement. Addressing these challenges requires cost-effective approaches that use easy to access satellite Earth observation data with machine learning methods to forecast drought.

Recent advances in high-performance computing and storage have enabled the development and implementation of robust early warning systems via machine learning. This PhD research aims to develop agricultural drought forecast models using satellite-based Vegetation Condition Index (VCI) and other agricultural drought indicators like precipitation and soil moisture.

Data sampled from Landsat and MODIS satellite images were used to develop a Gaussian Process model to forecast VCI. An Auto-regression modelling method was also used in this study for comparative analysis. The forecast models were very skilful for forecasting VCI for 2 to 6 weeks lead time. To extend the forecast range of VCI beyond 6 weeks, we used information from additional hydro-climatic factors within a Bayesian Auto-Regressive Distributed Lags (BARDL) model. The BARDL approach improved the forecast range by approximately two weeks. Finally, a Hierarchical Bayesian Model (HBM) was used to model agricultural drought in regions with diverse land cover types and agro-ecological zones. Forecasts from the HBM were more accurate than the BARDL approach, with an approximately one-week improvement in the forecast range.

Acknowledgements

First of all, I am eternally grateful to God for seeing me through my PhD journey.

I would also like to express my gratitude to the following people: My supervisors, **Prof. Seb Oliver**, for his exceptional ideas, training opportunities, encouragement, guidance, and making the Department of Physics and Astronomy conducive for research and studying. **Dr. Pedram Rowhani**, my supervisor, for his constant push and constructive criticism, which helped shape my research. And also linking me to people in the Department of Geography for training on acquiring and processing remote sensing data.

Dr. Peter D. Hurley, for his immense support with understanding Bayesian modelling and analysis as well as his friendship and advice.

I would also like to acknowledge Dr. Steven Duivenvoorden, Dr. Raphael Shirley and Dr. Adam Barrett for their support during the early stages of my research. A huge thanks goes to James M. Muthoka for his support and guidance with acquiring, processing, and analysing satellite Earth Observation data.

I would also want to thank Professor Eric Danquah and my former colleagues at WACCI * for their support and motivation to pursue a PhD in Data Science, and especially to Dr. John Eleblu for his direction and encouragement to apply for the DARA PhD scholarship.

Family and friends: My special gratitude goes to my wife, Dr. Naa Ayeley Salakpi, my son Elon Elinam Salakpi and my mother-in-law for their daily encouragement and prayers, especially on days when I am overwhelmed and down. To my sister Thelma Hammond, her husband Wilson Hammond and the children for providing me a home in the UK. Also, to Moses Palmer and Sampson Addo for their immense support prayer. Finally, to my parents, who have supported me in my education and career choice through their guidance, encouragement and prayers.

Special acknowledgment

Funding for PhD research and thesis was provided by the Newton Fund's Development in Africa with Radio Astronomy (DARA) Big Data project ST/R001898/1.

*West Africa Centre for Crop Improvement, University of Ghana

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Drought is a slow occurring natural hazard that is known to be very complex (Wilhite et al., 2007; Mishra et al., 2010). On average, drought events affects the livelihood of over 55 million people around the world, with a large proportion being people from Africa (Vatter, 2019). Addressing the challenges associated with drought requires understanding the drought categories and the factors that influence their occurrence. These categories include *meteorological drought* caused by precipitation (rainfall) deficit; *hydrological drought* which occurs as a result of a shortage of water in streams, lakes and groundwater; *agricultural drought* caused by a deficit in water (soil moisture) available to plants; and *socio-economic drought* caused by the scarcity of goods as a result of drought conditions (Heim, 2002). Of all these, agricultural drought stands out due to its direct impact on people's livelihoods.

Agricultural drought is the most complex in all the drought categories; its impact extends and varies over wide areas in affected regions (Boken et al., 2005). This category of drought, if not well managed, can adversely affect wildlife habitats, agriculture production, food security and the economy of the affected country or region (UNDRR, 2021; Chiang et al., 2021). Agricultural drought also has the potential of causing global food insecurity crises, especially when major food-producing countries are hit (D. Maxwell et al., 2012).

In Africa, the devastating effects of agricultural droughts are mostly seen in the arid and semi-arid lands (ASALs), especially in the eastern region of Africa. During the 2010-2011 drought in the Horn of Africa, over 200,000 lives were lost, and a total of 1.3 billion dollars was spent on post-disaster relief (UNDRR, 2021). The challenges associated with agricultural drought have also been linked to the cause of social unrest in most developing countries leading to a socio-economic drought situation (Kelley et al., 2015; Martínez-Fernández et al., 2016; Mishra et al., 2010).

Following the frequent occurrence of drought events, many countries started setting up national drought strategies to monitor and reduce the negative impact of drought hazards. However, to ensure these strategies meet the required standards, the World Meteorological Organization (WMO) advised governments to consider drought early warning strategies that provide timely forecasts on the onset and severity of drought events (Hayes et al., 2011). In addition, other global organisations and initiatives like the United Nations Office for Disaster Risk Reduction (UNDRR) and Paris Agreement recognise that implementing robust Earning Warning Systems (EWS) can accelerate the realisation of the United Nations' Sustainable Development Goals (SDG) related to food security (SDG 1,& 2) and extreme climate events (SDG13) (UNFCCC, 2015; UNDRR, 2021).

Through the efforts of organisations like the United Nations Development Programme (UNDP) and the United States Agency for International Development (USAID), many EWS have been developed and deployed in drought-prone regions. These EWS mostly fall under two categories (Funk et al., 2019):

- EWS for monitoring and forecasting weather and climatic indicators, e.g. Global Drought Information Systems* and
- EWS for monitoring vegetation health and impact of drought on food security, e.g. Famine Early Warning Systems Network (FEWS NET)†.

The FEWS NET is one of the popular systems used for monitoring and anticipating the impact of drought on food security in the east and southern Africa. It uses a combination of household livelihood data and biophysical indicators on climate and vegetation health to provide evidence-based guidance for disaster relief efforts (Funk et al., 2019). In east Africa, specifically Kenya, the National Drought Management Authority (NDMA) is the government agency mandated to spearhead drought risk management and establish measures for early action (NDMA, 2021). Like the FEWS NET, the NDMA also has an EWS that assimilates socio-economic data like Mid Upper Arm Circumference (MUAC), livestock/food commodity prices, pasture conditions in pastoral communities, and remotely sensed drought indicators to determine drought condition. The information from these indicators is used in monthly bulletins to aid anticipatory action and drought preparedness.

Although the EWS that focuses on agricultural drought has been useful, it has some downsides to its implementation in developing countries. First of all, financial constraints usually hinder the sustainability of such systems, especially in the area of data collection and information dissemination (Braimoh et al., 2018). Secondly, the systems rely primarily on weather station data; however, these stations do not cover all areas in developing countries. Thus, information relevant for drought early action is not available for all crop farmers and pastoralists (Masinde, 2014). Key amongst these challenges is that the forecast information required for

*www.drought.gov/gdm

†https://fews.net/

drought preparedness are in most cases based on expert judgement as stated by Funk et al., 2019. Fixing these challenges require an EWS that is cost-efficient in terms of implementation and management. The biophysical data used for such a system should also be timely, cover a wide area and be easy to acquire. Fortunately, advances in satellite Earth observation (EO) provide such data and enable the acquisition and derivation of many agro-climatic and biophysical indicators for measuring and monitoring drought events. Some of these indices include the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), Soil Moisture, Land Surface Temperature (LST) and Vegetation Condition Index (VCI). These indicators have been extensively used to study vegetation dynamics and to monitor agricultural drought (Dutta et al., 2013; Bolton et al., 2013; Shwetha et al., 2016; Ren et al., 2018). Information from these indicators coupled with advanced statistical methods can be used to develop machine learning models for monitoring and foresting agricultural drought. Machine learning algorithms make it possible to learn patterns from historical data for future predictions (Bishop, 2006). With the advances in high-performance computing and storage, developing and implementing machine learning models with high throughput EO data presents a more cost-effective approach to managing droughts.

Having identified the challenges with some existing EWS such as financial constraints, data availability and the fact that forecasts are based on expert judgements, we sought to answer the following questions:

- can we develop a cost-effective drought forecast model using a combination of data from satellite-based agricultural drought indicators and advanced machine learning methods?
- using the additional factors like precipitation and soil moisture levels that influence vegetation condition, can we develop a forecast model for long-term drought forecasts?
- can we study and simultaneously forecast drought and its impact in regions with different agro-ecologies or different land covers?

To address the first question, we developed a forecast model to forecast VCI, an agricultural drought indicator, using Gaussian Processes, a non-parametric model ideal for univariate time series modelling. An Auto-regression modelling method was also used for comparative analysis. Data used for this model were sampled from the Landsat [*] and the Moderate Resolution Imaging Spectroradiometer (MODIS) [†] satellite EO images. Details of this work are outlined in chapter 3.

Extending the forecast range of VCI as stated in the second research question required additional factors like precipitation and soil moisture via a multivariate modelling approach. An

[*] https://www.usgs.gov/core-science-systems/nli/landsat

[†] https://terra.nasa.gov/about/terra-instruments/modis

Auto-Regressive Distributed Lags (ARDL) model implemented within a Bayesian probabilistic modelling framework was used to address this. A detailed account of the methodology and results are in chapter 4

A Hierarchical Bayesian Modelling (HBM) approach was used to model agricultural drought for the different land cover maps and agro-ecological zones. The HBM method enable the incorporation of spatial variation into the forecast model. The chapter 5 of this thesis describes the methods and results from this work.

The research studies outlined in this thesis were conducted with EO data from Kenya. See chapter 2 for more details.

The following sections of this chapter will focus on a review of research studies on agricultural drought and the role of EO data and its drought indicators. This will be followed by a review of machine learning and its role in effective drought monitoring and forecasting. Finally, an overview of the various methods used in this thesis are outlined, along with their pros and cons.

1.2 Agricultural Drought Monitoring and Forecasting

Research on drought and its impact have been mostly directed towards meteorological and hydrological drought. The focus on these categories of drought is mainly because, unlike agricultural drought events, their onset is instantaneously visible (Boken et al., 2005). Secondly, vegetation in grasslands and shrublands, including food crops in communities that practice rain-fed agriculture, rely on an extensively studied water cycle driven by changes in the atmosphere and the oceans. These gave drought researchers a solid background for studying these categories of droughts. Over time, the focus shifted to agricultural drought because of its economic importance and the availability of several agricultural drought indices derived from temperature, soil moisture and vegetation health indicators like the NDVI (Mishra et al., 2011). Acquiring data for agricultural drought indicators at a global scale became possible due to advances in satellite Earth observation and high-performance computing systems for processing and storing remotely sensed Earth observation datasets.

1.2.1 The Role of Earth Observation and Remote Sensing

Satellite Earth observation datasets are images of the Earth captured by remotely sensing the energy dissipated from the Earth's surface. Depending on the source of energy, remote sensing can either be passive or active. Remote sensing that depends on Earth's energy source or the energy emanating from the Earth due to the sun's radiation is termed passive remote sensing. On the other hand, remote sensing measurements that are a result of an energy source from the instruments own radiation is termed active remote sensing (Richards, 2013). The instruments used for passive remote sensing are categorised as optical, while active instruments that work with microwaves signal are referred to as Synthetic Aperture Radar (SAR) (Richards, 2013; Woodhouse, 2017). The imaging instruments capture images in multiple spectral bands or channels with wavelengths ranging from visible (Red, Green, Blue) to infrared (Richards, 2013; Jensen, 2014). The number of specific wavelengths an imaging sensor can measure at a given time is referred to as its spectral resolution (Jensen, 2014). The pixel values, also known as the reflectance value of the image, represent the level of brightness of the signal received by the imaging instrument. The reflectance values from the various bands provide the data required for analysing changes in land cover and land use on the Earth's surface. The spatial resolution of the image refers to the physical scale of a single-pixel (pixel dimension). The value used to describe the spatial resolution is the length of one side of the pixel and is equivalent to the actual dimensions on the Earth's surface (i.e. a 500m spatial resolution implies 500m by 500m area on the ground) (Richards, 2013). The satellites carrying these instruments orbit the Earth at regular intervals, referred to as the repeat interval or temporal resolution. The temporal resolution range from daily to fortnightly and enables the monitoring of changes in an area of interest (Richards, 2013).

The Moderate Resolution Imaging Spectroradiometer (MODIS) is an example of such EO products. The Terra and Aqua instruments onboard the MODIS captures images in 36 spectral bands between 405nm and 14385nm. It has a temporal resolution of 1 to 2 days with spatial resolutions of 250m 500m, and 1000m (Schaaf et al., 2015). Other examples include United States Geological Survey (USGS) Landsat[*] and European Space Agency's (ESA) Sentinel [†] products. The launch of these satellites for remote sensing enabled the derivation of biophysical indicators for monitoring of the Earth's activities including agricultural drought research and to understand how humans and other living organisms interact with their environment (Ma et al., 2015).

1.2.2 Role of Drought Indices

When it comes to monitoring drought, choosing the appropriate indicator is very vital. Each of the categories of droughts has many indices, and one specific index does not provide all the required information when monitoring or measuring drought severity (Hao et al., 2017). For

[*]https://landsat.gsfc.nasa.gov/

[†]https://sentinels.copernicus.eu/web/sentinel/home

instance, the NDVI, despite extensive use (Pettorelli et al., 2005; Sruthi et al., 2015; Klisch et al., 2016; Nanzad et al., 2019) for studying vegetation health and drought, has some downsides. It is known to be sensitive to leaf chlorophyll and tends to saturate in regions with dense leaf covers. To address this challenge, the Enhanced Vegetation Index (EVI) was developed to cut down on the saturation and improve vegetation signal in regions with high plant biomass (Bolton et al., 2013; Huete et al., 2002). A significant factor that affects agricultural drought is soil moisture, which indicates water stress in plants. The Normalized Difference Water Index (NDWI) was proposed by Gao, 1996 to help monitor the vegetation stress resulting from moisture stress. In arid regions with sparse vegetation, monitoring vegetation conditions can be very challenging due to the background effect of soil. To address this effect, the Soil-Adjusted Vegetation Index (SAVI) was proposed by (A.R Huete, 1988). The potential of SAVI was recently explored by Ren et al., 2018 for above ground biomass estimation in deserts steppes in Magnolia. When it comes to agricultural drought monitoring and forecasting, the Vegetation Condition Index (VCI), Temperature Condition Index (TCI) and Vegetation Health Index (VHI) (Kogan, 1995) are the popular indices used. VCI is usually derived from NDVI and, in some cases, from SAVI as done by Bowell et al., 2021. TCI is derived from the LST and VHI from the combination of VCI and TCI. All the indices mentioned above were derived from the combination of spectral bands, and the equation for deriving them are outlined in table 1.1. The analysis of these indicators can be based on a Geographic object-based image analysis (GEOBIA) (Chen et al., 2018) approach, which involves the use of sections or groups of homogeneous pixels seen as objects (Chen et al., 2018). The analysis can also be pixel-based, where the reflectance values are extracted and analysed as a time series(Phiri et al., 2017). The forecast models in chapter 3, chapter 4 and chapter 5 were all pixel-based. The forecast models used in these chapters were also based on VCI. The choice VCI was primarily because the complex nature of agricultural drought required an indicator that adequately reflects the impact of hydro-climatic and biophysical factors like rainfall, temperature and soil moisture level (Vicente-Serrano et al., 2012; Yihdego et al., 2019). Details on VCI and how it is derived can be found in chapter 3 and chapter 4.

Table 1.1: Table of some remote sensing derived vegetation indices. NIR: Reflectance in Near Infra-Red Spectral Band, SWIR: Reflectance Short-Wave Infra-Red Spectral Band, RED: Reflectance Red Spectral Band, BLUE: Reflectance Blue Spectral Band L_1, L_2: leaf canopy background adjustment, C_1, C_2: coefficients of the atmospheric aerosol resistance ($L_1 = 1$,$L_2 = 0.5$ $C_1 = 6$, $C_2 = 7.5$)

Index	**Equation**
Normalized Difference Vegetation Index (NDVI)	$\frac{(NIR-RED)}{(NIR+RED)}$
Enhanced Vegetation Index (EVI)	$\frac{(NIR-RED)}{(NIR+C_1\times RED-C_2\ )\times BLUE+L_1)}$
Soil Adjusted Vegetation Index (SAVI)	$\frac{1+L_2(NIR-RED)]}{(NIR+RED+L_2\ )}$
Normalized Difference Water Index (NDWI)	$\frac{(NIR-SWIR)}{(NIR+SWIR)}$
Vegetation Condition Index (VCI)	$\frac{(NDVI-NDVI_{(\text{min})})}{(NDVI_{(\text{max})}+NDVI_{(\text{min})})} \times 100$
Temperature Condition Index (TCI)	$\frac{(LST_{(\text{max})})-LST}{(LST_{(\text{max})}+LST_{(\text{min})})} \times 100$
Vegetation Health Index (VHI)	$\alpha.VCI + (1-\alpha)TCI$

1.2.3 Role of Machine Learning

Recent advances in computational power and access to sophisticated data modelling algorithms have enabled the use of machine learning methods. These machine learning approaches use statistical methods to establish relationships or learn patterns in data. The discovered patterns and relationships are then used to characterise the data or make inferences about the future (Holloway et al., 2018; Bishop, 2006). Machine learning methods can generally be grouped into Supervised and Unsupervised learning. Supervised learning, in the context of remote sensing, involves the use of labelled pixels or pixels with known values ranges as target variables to train an algorithm to make predictions (Li et al., 2014; A. E. Maxwell et al., 2018). Depending on the type of desired output or target variables, supervised learning can be described as a classification task where target labels as discrete categories or as a regression task where the outputs are continuous values (Bishop, 2006). Popular supervised learning methods include the Support Vector Machine (SVM), used by Srivastava et al., 2012 in their paper for land use and land cover classification and Artificial Neural Networks (ANN), proposed by the (Marj et al., 2011) for forecasting drought using NDVI. Unsupervised learning, on the other hand, does not require having labelled pixels; instead, the algorithms used for this method can group data based purely on their characteristics and inherent relationships (Phiri et al., 2017). An example of such unsupervised learning methods is the K-Means clustering algorithm (Hartigan et al., 1979).

Machine learning has recently gained much traction; as a result, research works that focus on drought forecasting resort to this approach. Jalili et al., 2014, used the time series data of

NDVI, NDVI-Anomaly, TCI and VCI as inputs to forecast SPI values using Neural Networks (NN). The overall aim was to build an EWS to reduce the risk of severe disasters associated with drought. They used neural networks architectures based on Multilayer Perceptron (MLP) and Radial-Basis Function (RBF) as well as an SVM model. The predictions from the MLP network achieved 90% accuracy for monthly forecasts. They also concluded that TCI and VCI were better inputs compared to NDVIs. This observation was attributed to the fact that VCI and TCI were based on long-term data records, thus are suitable for comparing heterogeneous regions. The strength of this approach was that SPI could be predicted globally using EO data; however, it was less useful for people in pastoral communities who require timely information on vegetation condition. Which was one of the challenges addressed in this PhD research. Nay et al., 2018, trained a Gradient Boost Machine (GBM) model to predict vegetation health using EVI as an indicator. The aim was to develop a software application to determine future values of EVI based on lagged time series of spectral bands, EVI, NDVI, LST, LAI and photosynthetically active radiation (PAR) values. They noted from their results that the mean-squared difference reduced by 40% to 50% when the model was trained on lagged spectral bands and lagged EVI instead of using EVI only. Predictions over agricultural areas had a correlation of 0.75 and above with the observed EVI values. The GBM model was good at predicting high values of EVI and not the lower values that indicated vegetation stress. Another downside to this model was that it was developed for a one-step forecast only and not for multi-step forecasts as demonstrated in this thesis. Adede et al., 2019, used three-month lags of precipitation, SPI, and VCI to forecast future VCI3M (VCI from the 3-month rolling average) over three months using an ANN model. Before fitting the ANN, a General Additive Model (Hastie et al., 2017) was used to select the optimal input combinations. A one month VCI3M forecast from their ANN model had an R^2 score * ranging from 0.71 to 0.83. However, this score dropped significantly for forecasts beyond one month, indicating that the model was not robust for long-term forecasts, which we addressed with the second and third paper in this thesis. Aside from their high prediction accuracy, the models used cited studies are usually subject to high uncertainty. Thus, there is a need to have models that can produce probabilistic interpretations as part of its skill (AghaKouchak, 2014; Yan et al., 2017).

1.2.4 Bayesian Methods

Besides the AR method, which served as a base model in chapter 3, all the methods used in this thesis were implemented within a Bayesian probabilistic framework. Bayesian models are based

*Coefficient of Determination: The proportion of variability in the observed data that the model can explain on Bayes' theorem and allow prior knowledge about that parameter space. Parameter inference with this approach is made via Markov Chain Monte Carlo (MCMC) algorithms (Neal, 1993). Parameters and predicted values generated by the models are probability distribution functions (PDF), which provides a straightforward way to quantify uncertainties and make probabilistic interpretations (Martin, 2018; McElreath, 2018). These methods have been extensively used in Medical Research, Engineering, Physics and Astronomy. However, for drought research, it is mostly used in meteorological and hydrological drought studies. Wang et al., 2009 for instance, used a Bayesian model to monitor and forecast seasonal streamflow indices at multiple sites in the southeastern region of Australia.

Figure 1.1 shows various methods used in this thesis, including their advantages and disadvantages. From the flow chart, it can be seen that the method used in chapter 4 and chapter 5 builds on the shortfalls of the preceding chapter.

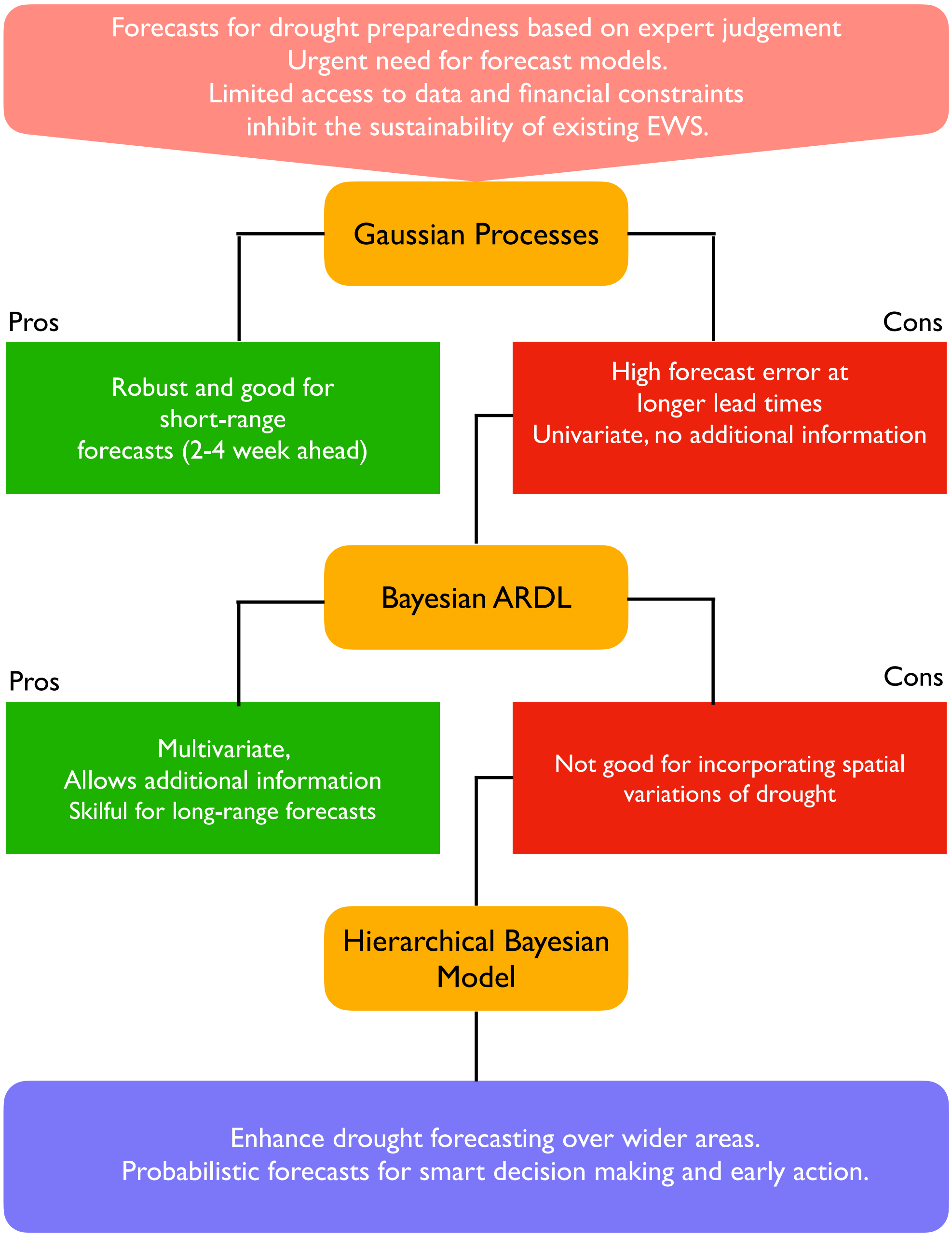

Figure 1.1: A flow chart depicting the pros and cons of the forecast models used in this thesis.

References

Adede, Chrisgone, Robert Oboko, Peter Waiganjo Wagacha and Clement Atzberger (May 2019). "A Mixed Model Approach to Vegetation Condition Prediction Using Artificial Neural Networks (ANN): Case of Kenya's Operational Drought Monitoring". In: *Remote Sensing* 11.9, p. 1099. ISSN: 2072-4292. DOI: 10.3390/rs11091099. URL: https://www.mdpi.com/2072-4292/11/9/1099.

AghaKouchak, A. (2014). "A baseline probabilistic drought forecasting framework using standardized soil moisture index: Application to the 2012 United States drought". In: *Hydrology and Earth System Sciences* 18.7, pp. 2485–2492. ISSN: 16077938. DOI: 10.5194/hess-18-2485-2014.

Bishop, Christopher M (2006). "Pattern recognition". In: *Machine learning* 128.9.

Boken, Vijendra K., Arthur P. Cracknell and Ronald L. Heathcote (May 2005). *Monitoring and Predicting Agricultural Drought*. Oxford University Press. ISBN: 9780195162349. DOI: 10.1093/oso/9780195162349.001.0001. URL: https://oxford.universitypressscholarship.com/view/10.1093/oso/9780195162349.001.0001/isbn-9780195162349.

Bolton, Douglas K. and Mark A. Friedl (2013). "Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics". In: *Agricultural and Forest Meteorology* 173, pp. 74–84. ISSN: 01681923. DOI: 10.1016/j.agrformet.2013.01.007. URL: http://dx.doi.org/10.1016/j.agrformet.2013.01.007.

Bowell, Andrew, Edward E. Salakpi, Kiswendsida Guigma, James M. Muthoka, John Mwangi and Pedram Rowhani (2021). "Validating commonly used drought indicators in Kenya". In: *Environmental Research Letters*. URL: http://iopscience.iop.org/article/10.1088/1748-9326/ac16a2.

Braimoh, Ademola, Bernard Manyena, Grace Obuya and Francis Muraya (2018). *Assessment of Food Security Early Warning Systems for East and Southern Africa*. Tech. rep. January. DOI: 10.1596/29269.

Chen, Gang, Qihao Weng, Geoffrey J. Hay and Yinan He (2018). "Geographic Object-based Image Analysis (GEOBIA): Emerging trends and future opportunities". In: *GIScience Remote Sensing* 55.2, p. 15481603.2018.1426092. ISSN: 1548-1603. DOI: 10.1080/15481603.2018.1426092. URL: https://www.tandfonline.com/doi/full/10.1080/15481603.2018.1426092.

Chiang, Felicia, Omid Mazdiyasni and Amir AghaKouchak (May 2021). "Evidence of anthropogenic impacts on global drought frequency, duration, and intensity". In: *Nature Commu-*

nications 2021 12:1 12.1, pp. 1–10. ISSN: 2041-1723. DOI: 10.1038/s41467-021-22314-w. URL: https://www.nature.com/articles/s41467-021-22314-w.

Dutta, Dipanwita, Arnab Kundu and N. R. Patel (2013). "Predicting agricultural drought in eastern Rajasthan of India using NDVI and standardized precipitation index". In: *Geocarto International* 28.3, pp. 192–209. ISSN: 10106049. DOI: 10.1080/10106049.2012.679975.

Funk, Chris et al. (June 2019). "Recognizing the Famine Early Warning Systems Network: Over 30 Years of Drought Early Warning Science Advances and Partnerships Promoting Global Food Security". In: *Bulletin of the American Meteorological Society* 100.6, pp. 1011–1027. ISSN: 0003-0007. DOI: 10.1175/BAMS-D-17-0233.1. URL: https://journals.ametsoc.org/view/journals/bams/100/6/bams-d-17-0233.1.xml.

Gao, Bo-cai (Dec. 1996). "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space". In: *Remote Sensing of Environment* 58.3, pp. 257–266. ISSN: 00344257. DOI: 10.1016/S0034-4257(96)00067-3. arXiv: bhmic00033. URL: https://linkinghub.elsevier.com/retrieve/pii/S0034425796000673.

Hao, Zengchao, Xing Yuan, Youlong Xia, Fanghua Hao and Vijay P. Singh (2017). "An overview of drought monitoring and prediction systems at regional and global scales". In: *Bulletin of the American Meteorological Society* September, BAMS–D–15–00149.1. ISSN: 00030007. DOI: 10.1175/BAMS-D-15-00149.1. URL: http://journals.ametsoc.org/doi/10.1175/BAMS-D-15-00149.1.

Hartigan, John A and Manchek A Wong (1979). "Algorithm AS 136: A k-means clustering algorithm". In: *Journal of the royal statistical society. series c (applied statistics)* 28.1, pp. 100–108.

Hastie, Trevor J and Robert J Tibshirani (2017). *Generalized additive models*. Routledge.

Hayes, Michael, Mark Svoboda, Nicole Wall and Melissa Widhalm (2011). "The lincoln declaration on drought indices: Universal meteorological drought index recommended". In: *Bulletin of the American Meteorological Society* 92.4, pp. 485–488. ISSN: 00030007. DOI: 10.1175/2010BAMS3103.1.

Heim, Richard (2002). "A Review of Twentieth- Century Drought Indices Used in the United States". In: August, pp. 1149–1165.

Holloway, Jacinta and Kerrie Mengersen (2018). "Statistical machine learning methods and remote sensing for sustainable development goals: A review". In: *Remote Sensing* 10.9, p. 1365.

Huete, A, K Didan, T Miura, E.P Rodriguez, X Gao and L.G Ferreira (Nov. 2002). "Overview of the radiometric and biophysical performance of the MODIS vegetation indices". In: *Remote Sensing of Environment* 83.1-2, pp. 195–213. ISSN: 0034-4257. DOI: 10.1016/S0034-

4257(02)00096-2. URL: https://www.sciencedirect.com/science/article/pii/S0034425702000962?via%3Dihub.

Huete, A.R (Aug. 1988). "A soil-adjusted vegetation index (SAVI)". In: *Remote Sensing of Environment* 25.3, pp. 295–309. ISSN: 0034-4257. DOI: 10.1016/0034-4257(88)90106-X. URL: https://www.sciencedirect.com/science/article/pii/003442578890106X?via%5C%3Dihub.

Jalili, Mahdi, Joobin Gharibshah, Seyed Morsal Ghavami, Mohammadreza Beheshtifar and Reza Farshi (2014). "Nationwide prediction of drought conditions in Iran based on remote sensing data". In: *IEEE Transactions on Computers* 63.1, pp. 90–101. ISSN: 00189340. DOI: 10.1109/TC.2013.118.

Jensen, John R. (2014). *Remote sensing of the environment : an earth resource perspective*. Pearson. ISBN: 1292034939.

Kelley, Colin P, Shahrzad Mohtadi, Mark A Cane, Richard Seager and Yochanan Kushnir (Mar. 2015). "Climate change in the Fertile Crescent and implications of the recent Syrian drought." In: *Proceedings of the National Academy of Sciences of the United States of America* 112.11, pp. 3241–6. ISSN: 1091-6490. DOI: 10.1073/pnas.1421533112. URL: http://www.ncbi.nlm.nih.gov/pubmed/25733898%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4371967.

Klisch, Anja and Clement Atzberger (2016). "Operational drought monitoring in Kenya using MODIS NDVI time series". In: *Remote Sensing* 8.4. ISSN: 20724292. DOI: 10.3390/rs8040267.

Kogan, F. N. (1995). "Application of vegetation index and brightness temperature for drought detection". In: *Advances in Space Research* 15.11, pp. 91–100. ISSN: 02731177. DOI: 10.1016/0273-1177(95)00079-T.

Li, Miao, Shuying Zang, Bing Zhang, Shanshan Li and Changshan Wu (2014). "A review of remote sensing image classification techniques: The role of Spatio-contextual information". In: *European Journal of Remote Sensing* 47.1, pp. 389–411. ISSN: 22797254. DOI: 10.5721/EuJRS20144723.

Ma, Yan, Haiping Wu, Lizhe Wang, Bormin Huang, Rajiv Ranjan, Albert Zomaya and Wei Jie (2015). "Remote sensing big data computing: Challenges and opportunities". In: *Future Generation Computer Systems* 51, pp. 47–60. ISSN: 0167739X. DOI: 10.1016/j.future.2014.10.029. URL: http://dx.doi.org/10.1016/j.future.2014.10.029.

Marj, Ahmad Fatehi and Allard M. J. Meijerink (2011). "Agricultural drought forecasting using satellite images, climate indices and artificial neural network". In: *International Journal of*

Remote Sensing 32.24, pp. 9707–9719. ISSN: 0143-1161. DOI: 10.1080/01431161.2011.575896. URL: https://www.tandfonline.com/doi/full/10.1080/01431161.2011.575896.

Martin, O. (2018). *Bayesian Analysis with Python: Introduction to statistical modeling and probabilistic programming using PyMC3 and ArviZ, 2nd Edition*. Packt Publishing. ISBN: 9781789349665. URL: https://books.google.co.uk/books?id=1Z2BDwAAQBAJ.

Martínez-Fernández, J., A. González-Zamora, N. Sánchez, A. Gumuzzio and C.M. Herrero-Jiménez (May 2016). "Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index". In: *Remote Sensing of Environment* 177, pp. 277–286. ISSN: 0034-4257. DOI: 10.1016/J.RSE.2016.02.064. URL: https://www.sciencedirect.com/science/article/pii/S0034425716300931#bb0225.

Masinde, Muthoni (2014). "An Effective Drought Early Warning System for Sub-Saharan Africa: Integrating Modern and Indigenous Approaches". In: DOI: 10.1145/2664591.2664629. URL: http://dx.doi.org/10.1145/2664591.2664629.

Maxwell, Aaron E, Timothy A Warner, Fang Fang, Aaron E Maxwell, Timothy A Warner, Fang Fang Implementation, Aaron E Maxwell and Timothy A Warner (2018). "Implementation of machine-learning classification in remote sensing : an applied review sensing : an applied review". In: *International Journal of Remote Sensing* 39.9, pp. 2784–2817. ISSN: 0143-1161. DOI: 10.1080/01431161.2018.1433343. URL: https://doi.org/10.1080/01431161.2018.1433343.

Maxwell, Daniel and Merry Fitzpatrick (2012). "The 2011 Somalia famine: Context, causes, and complications". In: *Global Food Security* 1.1. Special Issue on the Somalia Famine of 2011-2012, pp. 5–12. ISSN: 2211-9124. DOI: https://doi.org/10.1016/j.gfs.2012.07.002. URL: https://www.sciencedirect.com/science/article/pii/S221191241200003X.

McElreath, Richard (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

Mishra, Ashok K. and Vijay P. Singh (Sept. 2010). "A review of drought concepts". In: *Journal of Hydrology* 391.1-2, pp. 202–216. ISSN: 0022-1694. DOI: 10.1016/J.JHYDROL.2010.07.012. URL: https://www.sciencedirect.com/science/article/pii/S0022169410004257.

— (2011). "Drought modeling - A review". In: *Journal of Hydrology* 403.1-2, pp. 157–175. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2011.03.049. arXiv: arXiv:1011.1669v3. URL: http://dx.doi.org/10.1016/j.jhydrol.2011.03.049.

Nanzad, Lkhagvadorj, Jiahua Zhang, Battsetseg Tuvdendorj, Mohsen Nabil, Sha Zhang and Yun Bai (May 2019). "NDVI anomaly for drought monitoring and its correlation with climate factors over Mongolia from 2000 to 2016". In: *Journal of Arid Environments* 164, pp. 69–77. ISSN: 0140-1963. DOI: 10.1016/J.JARIDENV.2019.01.019. URL: https://www.sciencedirect.com/science/article/pii/S0140196318302659.

Nay, John, Emily Burchfield and Jonathan Gilligan (2018). "A machine-learning approach to forecasting remotely sensed vegetation health". In: *International Journal of Remote Sensing* 39.6, pp. 1800–1816. ISSN: 0143-1161. DOI: 10.1080/01431161.2017.1410296. URL: https://www.tandfonline.com/doi/full/10.1080/01431161.2017.1410296.

NDMA (2021). *National Drought Management Authority*. URL: https://www.ndma.go.ke/ (visited on 11/05/2021).

Neal, Radford M (1993). *Probabilistic inference using Markov chain Monte Carlo methods*. Department of Computer Science, University of Toronto Toronto, Ontario, Canada.

Pettorelli, Nathalie, Jon Olav Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J Tucker and Nils Chr Stenseth (2005). "Using the satellite-derived NDVI to assess ecological responses to environmental change". In: *Trends in ecology & evolution* 20.9, pp. 503–510.

Phiri, Darius and Justin Morgenroth (2017). "Developments in Landsat land cover classification methods: A review". In: *Remote Sensing* 9.9. ISSN: 20724292. DOI: 10.3390/rs9090967.

Ren, Hongrui, Guangsheng Zhou and Feng Zhang (May 2018). "Using negative soil adjustment factor in soil-adjusted vegetation index (SAVI) for aboveground living biomass estimation in arid grasslands". In: *Remote Sensing of Environment* 209, pp. 439–445. ISSN: 0034-4257. DOI: 10.1016/J.RSE.2018.02.068. URL: https://www.sciencedirect.com/science/article/pii/S0034425718300804.

Richards, John A. (2013). *Remote sensing digital image analysis : an introduction*. Springer. ISBN: 3642300618. URL: https://books.google.co.uk/books/about/Remote_Sensing_Digital_Image_Analysis.html?id=CBn-x1j_Aj8C&source=kp_book_description&redir_esc=y.

Schaaf, Crystal and Z. Wang (2015). *MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006 [Data set]*. DOI: 10.5067/MODIS/MCD43A4.006. URL: https://lpdaac.usgs.gov/dataset%7B%5C_%7Ddiscovery/modis/modis%7B%5C_%7Dproducts%7B%5C_%7Dtable/mcd43a4%7B%5C_%7Dv006.

Shwetha, H. R. and D. Nagesh Kumar (2016). "Prediction of high spatio-temporal resolution land surface temperature under cloudy conditions using microwave vegetation index and ANN". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 117, pp. 40–55. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2016.03.011. URL: http://dx.doi.org/10.1016/j.isprsjprs.2016.03.011.

Srivastava, Prashant K., Dawei Han, Miguel A. Rico-Ramirez, Michaela Bray and Tanvir Islam (2012). "Selection of classification techniques for land use/land cover change investigation". In: *Advances in Space Research* 50.9, pp. 1250–1265. ISSN: 02731177. DOI: 10.1016/j.asr.2012.06.032. URL: http://dx.doi.org/10.1016/j.asr.2012.06.032.

Sruthi, S. and M.A. Mohammed Aslam (2015). "Agricultural Drought Analysis Using the NDVI and Land Surface Temperature Data; a Case Study of Raichur District". In: *Aquatic Procedia* 4.Icwrcoe, pp. 1258–1264. ISSN: 2214241X. DOI: 10.1016/j.aqpro.2015.02.164. URL: http://linkinghub.elsevier.com/retrieve/pii/S2214241X15001650.

UNDRR (2021). *GAR Special Report on Drought 2021 — UNDRR*. Tech. rep. URL: https://www.undrr.org/publication/gar-special-report-drought-2021.

UNFCCC (2015). *ADOPTION OF THE PARIS AGREEMENT - Paris Agreement text English*. Tech. rep.

Vatter, Juliane (2019). *DROUGHT RISK The Global Thirst for Water in the Era of Climate Crisis*. Tech. rep. World Wildlife Fund (WWF) Germany. URL: www.studioazola.com.

Vicente-Serrano, Sergio M. et al. (2012). "Challenges for drought mitigation in Africa: The potential use of geospatial data and drought information systems". In: *Applied Geography* 34, pp. 471–486. ISSN: 01436228. DOI: 10.1016/j.apgeog.2012.02.001. URL: http://dx.doi.org/10.1016/j.apgeog.2012.02.001.

Wang, Q. J., D. E. Robertson and F. H. S. Chiew (2009). "A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites". In: *Water Resources Research* 45.5. ISSN: 00431397. DOI: 10.1029/2008WR007355. URL: http://doi.wiley.com/10.1029/2008WR007355.

Wilhite, Donald A, Mark D Svoboda, Michael J Hayes, D A Wilhite, M D Svoboda and · M J Hayes, givenun=0 (2007). "Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness *". In: *Water Resour Manage* 21, pp. 763–774. DOI: 10.1007/s11269-006-9076-5. URL: http://drought.unl.edu.

Woodhouse, Iain H (2017). *Introduction to microwave remote sensing*. CRC press.

Yan, Hongxiang, Hamid Moradkhani and Mahkameh Zarekarizi (2017). "A probabilistic drought forecasting framework: A combined dynamical and statistical approach". In: *Journal of Hydrology* 548, pp. 291–304. ISSN: 00221694. DOI: 10.1016/j.jhydrol.2017.03.004. URL: http://dx.doi.org/10.1016/j.jhydrol.2017.03.004.

Yihdego, Yohannes, Babak Vaheddoost and Radwan A. Al-Weshah (2019). "Drought indices and indicators revisited". In: *Arabian Journal of Geosciences* 12.3, p. 69. DOI: 10.1007/s12517-019-4237-z. URL: https://doi.org/10.1007/s12517-019-4237-z.

Chapter 2

Study Area and Data

2.1 Study Area

All the drought modelling and forecasting work done in this thesis was conducted with data over the ASAL regions of Kenya, East Africa. The population of Kenya as of 2015 was 46 million, of which 74% live in the rural areas (FAO, 2017). The economic activities in this area are mostly agro-pastoralism and wildlife conservation (Gebremeskel et al., 2019; Vatter, 2019). Both livestock and wildlife in Kenya rely mostly on pastures and grasslands as their primary source of fodder (Sibanda et al., 2017). The annual mean rainfall in the ASALs ranges from 20mm to 200mm, with temperatures ranging from 10^oC to 40^oC (Ayugi et al., 2016). Unfortunately, the extreme climatic variations in the eastern African region expose the country to frequent agricultural drought events. These drought events adversely affect the country's economy, the livelihood of pastoralist communities and wildlife ecology. The impact of drought on the economy and livelihood in Kenya is mainly linked to losses associated with drought-induced livestock diseases due to poor pasture and grassland conditions. In addition, people, especially children, suffer from moderate to severe malnutrition, given that meat and milk form a large part of the food consumed in Kenya (ReliefWeb, 2021). In 2017, for instance, the government of Kenya declared a national drought that affected approximately 2.6 million people. To reduce the adverse effect of such drought events in the future, cost-effective Early Warning Systems developed with Satellite Earth observation (EO) products are way forward.

2.2 Data

[h] The data from the following EO products were used for drought modelling and forecasting.

More details on the data description, acquisition and preprocessing can in found in chapters 3, 4, and 5.

Table 2.1: A Table of the Satellite Earth Observation Products used in this thesis

Data	**Source (Producer)**	**Spatial Resolution**	**Temporal Resolution**	**Acquisition Period**
Surface Reflectances	USGS Landsat 7 & 8	30m	16-Day	2000-2019
Surface Reflectances	NASA MODIS (MCD43A4 v006)	500m	Daily	2000-2019
Precipitation (mm)	Climate Hazards Group InfraRed Precipitation (CHIRPS)	5km	Daily	2001-2019
Soil Moisture (m^3/m^{-3})	European Space Agency Climate Change Initiative (CCI)	30km	Daily	2001-2018

References

Ayugi, Brian Odhiambo, Wang Wen and Daisy Chepkemoi (2016). "Analysis of Spatial and Temporal Patterns of Rainfall Variations over Kenya". In: 6.11. ISSN: 2225-0948. URL: www.iiste.org.

FAO (2017). *FAO Africa Sustainable Livestock 2050: A Country Brief - Republic of Kenya.* Tech. rep. URL: http://www.fao.org/3/a-i7348e.pdf.

Gebremeskel, Gebremedhin, Qiuhong Tang, Siao Sun, Zhongwei Huang, Xuejun Zhang and Xingcai Liu (June 2019). *Droughts in East Africa: Causes, impacts and resilience.* DOI: 10.1016/j.earscirev.2019.04.015.

ReliefWeb (2021). *Kenya: Drought - 2014-2021 — ReliefWeb.* URL: https://reliefweb.int/disaster/dr-2014-000131-ken (visited on 07/02/2019).

Sibanda, Mbulisi, Onisimo Mutanga, Mathieu Rouget and Lalit Kumar (2017). "Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives". In: *Remote Sensing* 9.1. ISSN: 20724292. DOI: 10.3390/rs9010055.

Vatter, Juliane (2019). *DROUGHT RISK The Global Thirst for Water in the Era of Climate Crisis.* Tech. rep. World Wildlife Fund (WWF) Germany. URL: www.studioazola.com.

Chapter 3

Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya

Adam B. Barrett[a,b], Steven Duivenvoorden[a,c], Edward E. Salakpi[a], James M. Muthoka[d], John Mwangi[e], Seb Oliver[a,c] and Pedram Rowhani[d]*

[a] The Data Intensive Science Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

[b] Sackler Centre for Consciousness Science, Department of Informatics, University of Sussex, Brighton BN1 9QJ, UK

[c] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

[d] School of Global Studies, Department of Geography, University of Sussex, Brighton, BN1 9QJ, UK

[e] The National Drought Management Authority (NDMA), Lonrho House, Nairobi, Kenya

*Corresponding author: P.Rowhani@sussex.ac.uk

Abstract

Droughts are a recurring hazard in sub-Saharan Africa, that can wreak huge socioeconomic costs. Acting early based on alerts provided by early warning systems (EWS) can potentially provide substantial mitigation, reducing the financial and human cost. However, existing EWS tend only to monitor current, rather than forecast future, environmental and socioeconomic indicators of drought, and hence are not always sufficiently timely to be effective in practice. Here we present a novel method for forecasting satellite-based indicators of vegetation condition. Specifically, we focused on the 3-month Vegetation Condition Index (VCI3M) over pastoral livelihood zones in Kenya, which is the indicator used by the Kenyan National Drought Management Authority (NDMA). Using data from MODIS and Landsat, we apply linear autoregression and Gaussian process modeling methods and demonstrate high forecasting skill several weeks ahead. As a bench mark we predicted the drought alert marker used by NDMA (VCI3M<35). Both of our models were able to predict this alert marker four weeks ahead with a hit rate of around 89% and a false alarm rate of around 4%, or 81% and 6% respectively six weeks ahead. The methods developed here can thus identify a deteriorating vegetation condition well and sufficiently in advance to help disaster risk managers act early to support vulnerable communities and limit the impact of a drought hazard.

keywords:*Drought; Forecasting; Early Warning Systems; Disaster Risk Reduction; Landsat; MODIS*

3.1 Introduction

Droughts are a major threat globally as they can cause substantial damage to society, especially in regions that depend on rain-fed agriculture. They particularly impact food security by significantly reducing agricultural production (Lesk et al., 2016) and raising food prices (Nelson et al., 2014; Brown et al., 2015), which often leads to increased levels of malnutrition, migration, disease, and other health concerns (Piguet et al., 2011; Stanke et al., 2013). The majority of droughts occur in sub-Saharan Africa (EM-DAT, 2019) where many communities rely on predictable rainfall patterns for their livelihood.

In East Africa, the main economic activity in the arid and semi-arid lands (ASAL) is subsistence rain-fed agriculture, as well as livestock farming using pastures and grasslands as the main source of fodder. As a result, the pastoral and agro-pastoral communities who live in these drylands are particularly vulnerable to drought (Nyong et al., 2007; Orindi et al., 2007), especially since their existing coping strategies have been compromised by population growth and land use change in recent years (Galvin et al., 2001). Governments and donor agencies in the region have thus developed several tools and early warning systems (EWS) to mitigate the impact of droughts on pastoralists.

Most EWS tend to monitor current key biophysical and socio-economic factors to assess the possible exposure of vulnerable people to specific hazards. However, once the impacts are visible, it may be too late to mitigate the consequences (Felix Kogan et al., 2013). Hence there is growing interest in moving toward a proactive humanitarian approach to disasters by developing preparedness actions based on climate forecasts (Coughlan de Perez et al., 2015; Lopez et al., 2018; Wilkinson et al., 2018). Additionally, it is estimated that being better prepared before a drought hits significantly reduces the costs and losses from these disasters (Venton et al., 2012). Hence, EWS now increasingly include expert knowledge and qualitative assessments of seasonal climate forecasts to assess the future development of food security, and define actions to mitigate possible losses (Coughlan de Perez et al., 2015; Tozier de la Poterie et al., 2015). However for drought conditions, a meteorological drought does not always lead to negative agricultural outputs (Bhuiyan et al., 2006). There is thus a growing interest to include forecasts of the impacts of these hazards (WMO, 2015; Sai et al., 2018; Sutanto et al., 2019).

In Kenya, following several periods of intense drought, the government established the National Drought Management Authority (NDMA) in 2016, to set up and operate a drought EWS, as well as to establish drought preparedness strategies and contingency plans. The NDMA provides monthly bulletins assessing food security in the 23 ASAL regions using current biophysical (e.g., rainfall, vegetation condition) and socio-economic (production, access, and utilisation) factors. One key biophysical indicator used by the NDMA drought phase classification is based on the Vegetation Condition Index(VCI) (F.N. Kogan, 1995; Klisch et al., 2016; Rulinda et al., 2011; Rojas et al., 2011).

The VCI, which expresses the Normalized Difference Vegetation Index (NDVI) in terms of where it currently lies within its expected range for the given pixel, is one of a number of satellite-based indicators that have been developed to detect and monitor drought (Zargar et al., 2011). While there is little agreement between VCI and precipitation-based meteorological drought indicators (Quiring et al., 2010; Bhuiyan et al., 2006), it is strongly linked to agricultural production and widely used to identify drought onset, intensity, duration, and impact (Jiao et al., 2016). The NDMA uses the 3-month averaged VCI (VCI3M) in its operational EWS (Klisch et al., 2016). Once the VCI3M goes below a threshold of 35, the NDMA triggers a rapid food security assessment and has access to the National Drought Contingency Fund in order to implement its preparedness strategies and contingency plans.

The main goal of this paper is to explore machine-learning techniques to forecast the vegetation indices that are commonly used in the pastoral areas of Kenya to monitor droughts. In order to provide useful information to drought risk managers, we aim to identify the right balance between forecast lead time and uncertainty. To this end, we evaluated the performance of our approaches up to ten weeks ahead.

Based on NDMA's experience, we particularly focused on the pastoral livelihood zones as the VCI3M is more reliable in identifying drought condition for grazing and browsing in the more arid regions of the country. Several studies have developed statistical and machine-learning approaches (Udelhoven et al., 2009; Meroni et al., 2014; Zambrano et al., 2018; Vrieling et al., 2016) to predict end-of-season crop, forage and biomass production. Recently, (Matere et al., 2019) developed a decision support tool based on a mechanistic model to estimate 6-monthly forecasts of forage condition. Here, we specifically focus on Gaussian Process (GP) modelling (Rasmussen et al., 2006), and linear autoregressive (AR) modelling (Hamilton, 1994) to forecast NDVI and VCI3M, which are derived from both Landsat (every 16 days at 30 m resolution) and the MODerate resolution Imaging Spectroradiometer (MODIS - daily data at 500 m resolution). GP modelling uses kernel-based non-parametric Bayesian inference on the structure of correlations between observations, and is widely applied to classification, interpolation, change detection and forecasting problems (Brahim-Belhouari et al., 2001; Chandola et al., 2011; Camps-Valls et al., 2016; Upreti et al., 2019). Linear AR is the regression of future observations on past observations, assuming a linear dependence. Previously it has been performed on monthly (i.e. temporally more sparse) NDVI data, see for example (Asoka et al., 2015) and (Papagiannopoulou et al., 2017), with mixed results in terms of forecasting potential (R^2-scores between 0 and 0.4 at a lead time of one month).

3.2 Study area

In Kenya, the livestock sector accounts for 13% of the national GDP and 43% of its agricultural GDP. Livestock farming mainly occurs in the ASAL which cover about 80% of the country (UNDP, 2013; FAO, 2014). In these regions, the pastoral communities rely on pastures and grasslands as the main source of fodder (Behnke et al., 2011). Thus, providing information on pasture productivity to these communities is key in times of drought.

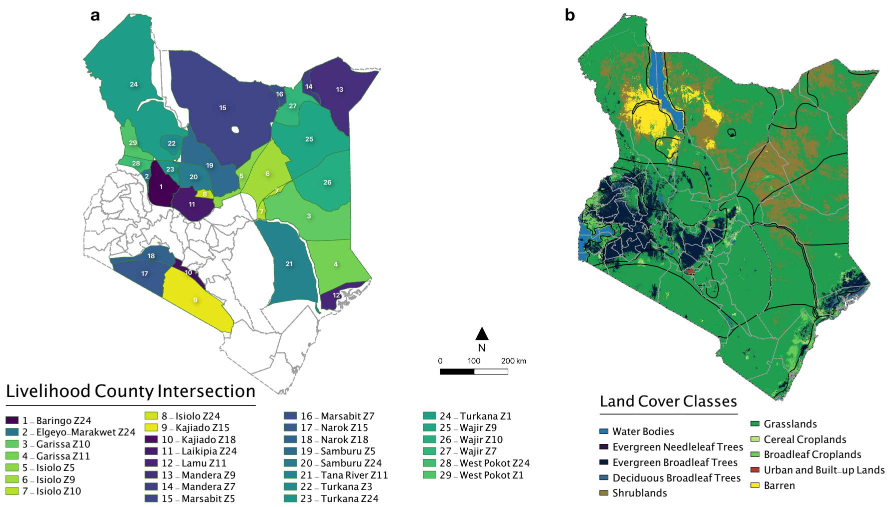

Figure 3.1: Maps of Kenya showing (a) Livelihood Zones and County intersections (Regions of Interest (ROI)) from which pixels were sampled for analysis, and (b) land-cover classification (according to the MODIS MCD12Q1 data). Analyses were performed for 29 regions, defined by pastoral livelihood zone and county intersections. A map showing the livelihood zones can be found in Fig. 3.15 in the Supplementary Material.

For the ASAL regions, the NDMA reports every month the VCI3M value at county level as well as over the different livelihood zones within the county. This study focused on the 10 (agro)-pastoral livelihood zones (see Fig. 3.15), which cross 15 counties. The names of the 29 livelihood zone county intersections can be found in Fig. 3.1; these are our regions-of-interest, which we refer to simply as 'regions'.

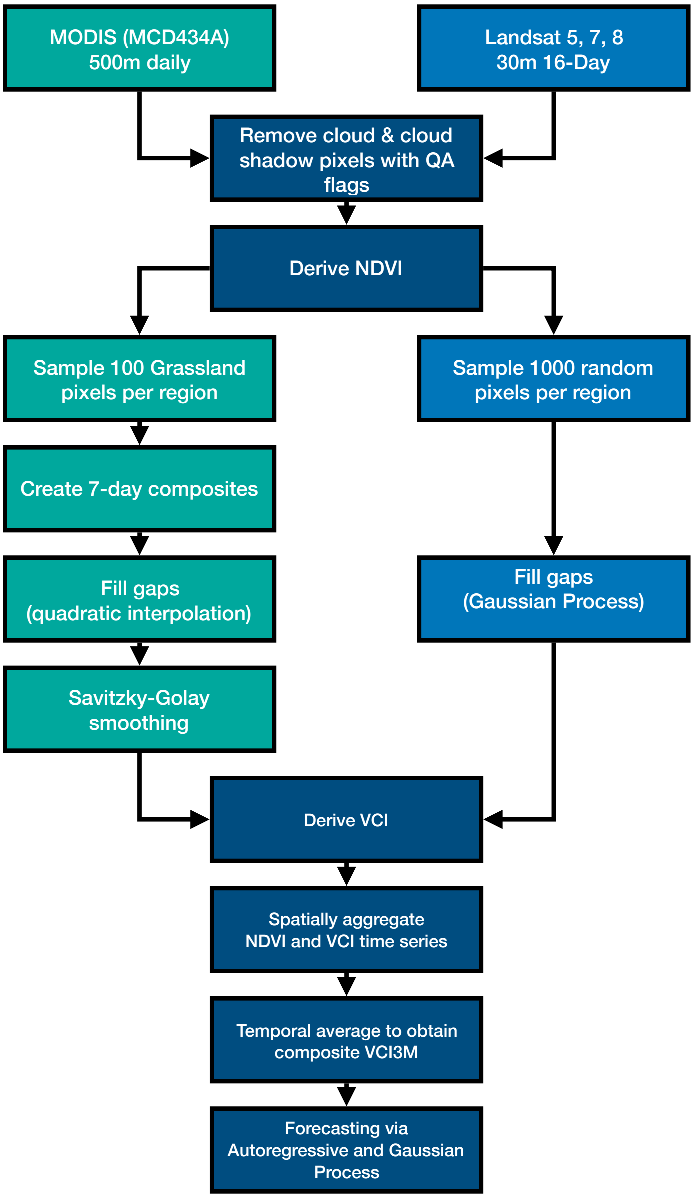

Figure 3.2: A flow chart of the data processing and analysis.

3.3 Methods

This research is based on two satellite-based Earth observation datasets, Landsat and MODIS. Description and justification of data selection, and a comparison between the two datasets can be found in 3.9. It should be noted that the analysis is based on a random subsample of the pixels within each of the 29 regions (Fig. 3.1). A summary of the entire work from data preparation to forecasting drought can be seen in Fig. 3.2.

3.3.1 Data preprocessing

3.3.1.1 Landsat

Temporal gridding and gap-filling on the Landsat data was done using Gaussian Process (GP) regression. For a given pixel, the GP regression took raw data as input, fit a temporal correlation structure to the data, and used this to output a time series of expected NDVI values, with observations provided every Saturday over the studied time period; see Figure 3.3 for an illustration and 3.10.1 for details. Two versions of GP gap-filling were carried out, which we refer to as forecast mode and non-forecast mode. For the non-forecast mode, the full time series from the given pixel were used to train the GP. The non-forecasting mode was used as the "ground truth" to test forecasts against. The forecast mode, by contrast, only used data up to a certain date, whichever date a forecast was being attempted from - since when doing forecasting with a near real-time data stream, one does not have access to future data.

3.3.1.2 MODIS

Weekly NDVI composites were obtained for each pixel by taking the mean (after cloud masking) of all available data over a 7-day time period. Gaps in the weekly time series were then filled using quadratic interpolation. Gaps longer than 6 weeks were left unfilled, see 3.10.2 for details. The gap-filled time series were then smoothed using the Savitzky-Golay method (Savitzky et al., 1964) to filter high-frequency measurement noise. The smoothing involved fitting, for each pixel, a polynomial to a window centred on the observation, and then replacing that observation with the output of the polynomial fit. The polynomial order was set to 2 (i.e. quadratic function) and the window length to 7 weeks. (Note that the combined interpolation and smoothing procedure does two rounds of quadratic interpolation where there are gaps, but that these are distinct: the interpolation fills a gap of up to 6 weeks with one quadratic function, while the smoothing modifies only one observation per fitted quadratic function.)

3.3.2 Indices

On both datasets, VCI time series were constructed from the NDVI time series according to the formula:

$$\mathrm{VCI_i} = 100 \times \frac{\mathrm{NDVI_i} - \mathrm{NDVI_{min,i}}}{\mathrm{NDVI_{max,i}} - \mathrm{NDVI_{min,i}}}, \tag{3.1}$$

where $\mathrm{NDVI_{min,i}}$ and $\mathrm{NDVI_{max,i}}$ are the minimum and maximum observed values for the NDVI of the pixel for the week of the year at time point i. The data within each region were aggregated taking the mean of the sampled pixels at each time point. Thus forecasting was applied on a single time series for each region. Finally, VCI3M was calculated as the mean VCI across the 12 weeks leading up to the given time point. Additionally, aggregate time series of NDVI anomaly were constructed (i.e., seasonal mean-subtracted NDVI, sometimes referred to as absolute anomaly; results for forecasting this can be found in 3.11).

With Landsat data, the mean, maximum and minimum value for the NDVI in (3.1) was computed using the non-forecast mode GP interpolated time series. Then forecast mode and non-forecast mode versions of each index were created. With the MODIS data, since large gaps were unfilled, whenever there were fewer than 25 individual pixel observations from a particular region at a given time, it was decided that there should be no datum in the aggregate VCI time series for that region (i.e., there should be a gap in the time series). Additionally, if the current aggregate NDVI observation was not present, a gap was placed in the VCI3M time series. Else, the mean was taken over all present observations from the most recent 12 weeks.

3.3.3 Forecasting

Machine-learning techniques offer a data-driven, empirical route to forecasting. Many different data inputs could be used to forecast these vegetation indices (e.g. precipitation and precipitation forecasts). However, perhaps the most simple is to use the past history of the indices themselves. This has the practical benefits of readily available data over large areas. Additionally, this approach will also take advantage from the fact that these indices are subject to plant growth and climate cycles giving periodic behaviour on large temporal scales that can be empirically modelled, while external perturbations, such as water availability, have persistent impact providing correlations on short temporal scales. Forecasts of NDVI anomaly and VCI3M were made using two separate methods, respectively based on Gaussian Process modelling (GP) and linear autoregressive (AR) modelling.

GP forecasting was performed by fitting a GP to the forecast mode aggregate time series for the index in question, and then using the GP to extrapolate. For details on GP modelling, see 3.10.1. The key step involved fitting a temporal correlation structure to the time series, i.e. a kernel $k(t, t')$ that describes the covariance between the index at any two times t and t'. The kernel with the highest evidence was the Radial Basis Function (RBF):

$$k_{\mathrm{RBF}}(t, t') = \sigma^2_{\mathrm{RBF}} \exp\Big(-0.5 \frac{|t - t'|^2}{l^2_{\mathrm{RBF}}} \Big) , \tag{3.2}$$

where σ^2_{RBF} and l_{RBF} are the signal variance and the length scale, respectively, and the modelling was carried out with the best fit version of this.

AR forecasting was performed with the following model-fitting and extrapolation method. For forecasting n weeks ahead, the following model was fit:

$$X_{t+n} = \sum_{i=0}^{p-1} a_i X_{t-i} + \epsilon_t , \tag{3.3}$$

where X is the index in question, subscripts denote the date (week), a_i are model coefficients, ϵ_t are the residuals (i.e. the errors), and p is called the model order. (This model assumes zero mean, so for VCI3M, the mean was removed prior to fitting the model, and then added back again after using the model to forecast the deviation from the mean.) Fitting the model to a segment of data involved finding the model coefficients that gave the minimum sum-square error, i.e. led to residuals with the minimum variance. To make a forecast, the model was fit using the most recent T consecutive observations (where T is called the training segment length), and then used to predict the observation n weeks after the most recent observation. This forecasting method was carried out along the entire available time series, fitting a distinct model to each segment of length T. The model order was set to $p = 3$ and the training segment length was set to $T = 200$, since forecast skill plateaued at these values.

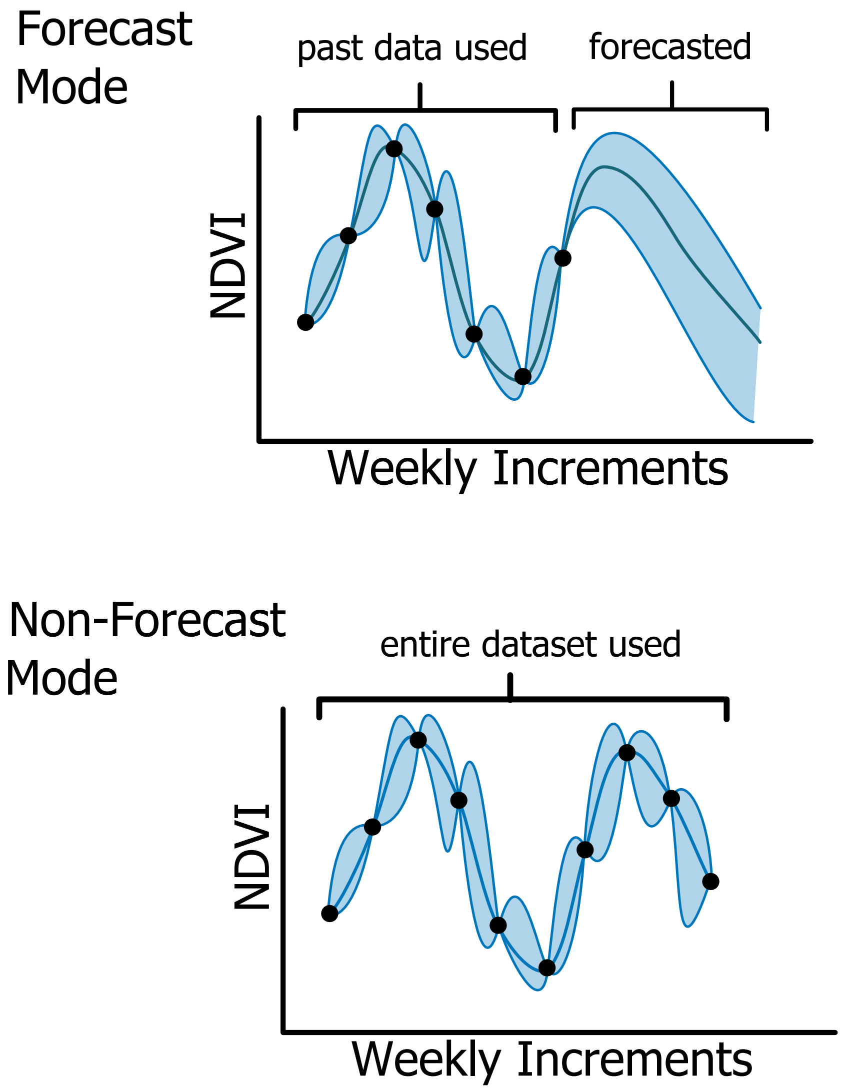

Figure 3.3: Illustration of the GP approach used for the Landsat data. In “forecast mode”, the correlations in the data up to a given date furnish a GP model, which can then be used for forecasting. In the “non-forecast mode”, the entire time series is used to train the GP, and provide a ground truth for the forecast.

3.3.4 Forecast assessment

Several metrics were used to assess the performance of the forecast methods tested on the data. In addition to RMSE, the R^2-score and the percentage of standard deviation remaining, S, were used. These are given by:

$$R^2\text{-score} = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}, \tag{3.4}$$

$$S = 100 \times \frac{\sqrt{\sum_i (y_i - f_i)^2}}{\sqrt{\sum_i (y_i - \bar{y})^2}}, \tag{3.5}$$

where the y_i are the true data, and the f_i are the forecasts. Note that $S \equiv 100 \times \sqrt{1 - R^2\text{-score}}$. To test for bias, we performed a linear regression of the actual index on the forecast index, and extracted the slope and intercept. Finally, receiver operating characteristic (ROC) curves were constructed for forecast-based drought-alert detection.

These performance indicators were also used to assess the sensitivity of our methods in space (comparing the results by region) and in time (to account for seasonality). Additionally, the forecast methods were evaluated for various drought categories (Klisch et al., 2016), compared against a persistence forecast (i.e. forecast obtained by taking the most recent observation to be the forecast value), and the impact of data gaps on forecast performance was analysed.

Forecasts on the MODIS data were assessed from January 1st 2004 onward, which was approximately the earliest date for which there were sufficient prior data for the AR method to be applied. Forecasts on the Landsat data were assessed only from January 1st 2014 onward, since the GP gap-filling method required training on data up to this date.

3.4 Results

3.4.1 Forecast value accuracy

The GP and AR forecasting methods were applied, on each of the two datasets, to regional aggregate VCI3M time series. We focus on performance results of GP forecasting on Landsat data and AR forecasting on MODIS data since these two combinations of data and forecasting method performed the best (as measured by R^2-score). We looked at lead times of up to ten weeks (see Figures 3.11 and 3.12). However, due to increasing uncertainty, the results provided here focus on two to six weeks forecasts of VCI3M.

Contour plots of forecast against actual data for two, four and six week forecasts are shown in Fig. 3.4. Table 3.1 shows the R^2-scores, RMSE, slope and intercept from each of these plots,

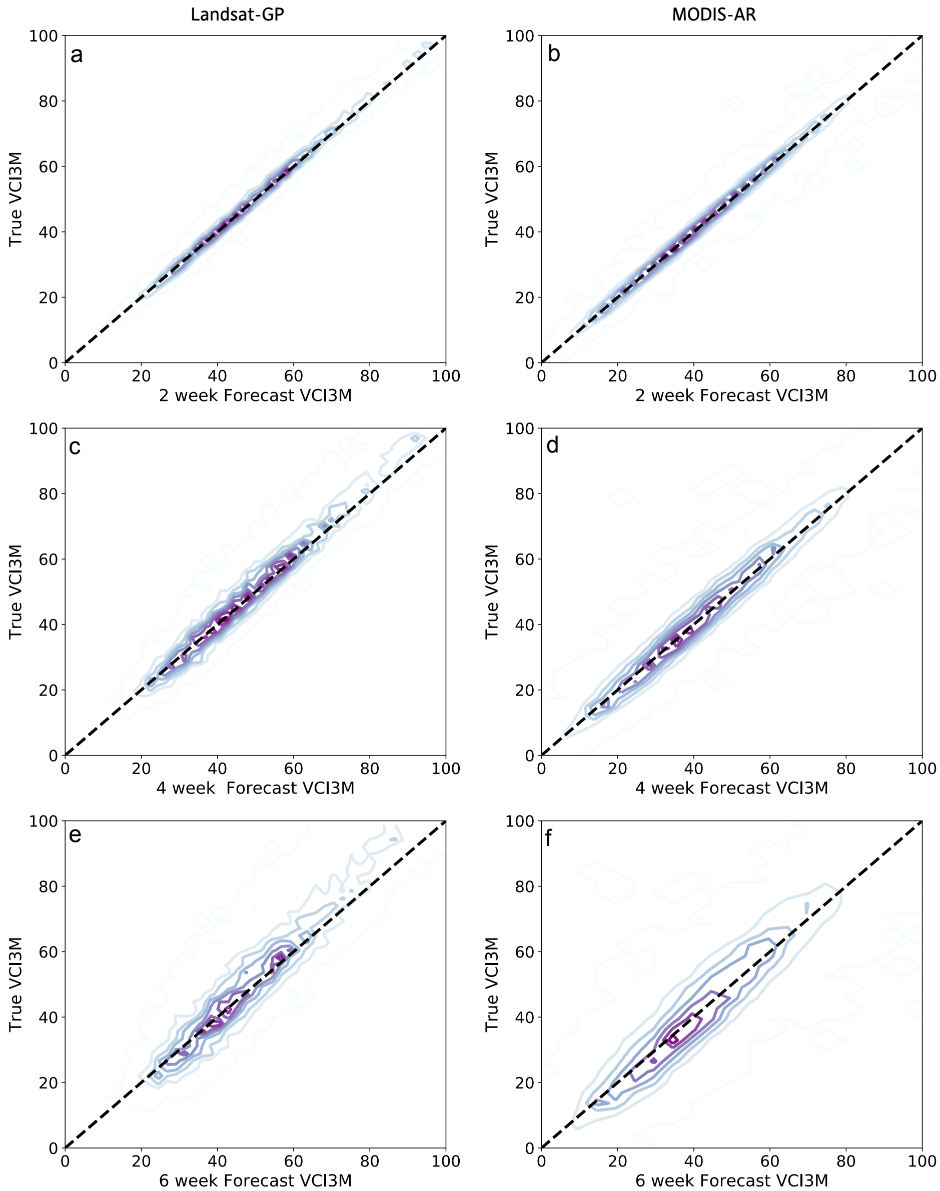

Figure 3.4: Contour plots of VCI3M against two, four and six weeks VCI3M forecasts. (a,c,e) show forecast performance for the GP method on Landsat data, and (b,d,f) show forecast performance for the AR method on MODIS data (across the 19 regions for which a 4 week forecast was possible more than 50% of the time, see main text for details).

Table 3.1: Performance statistics of VCI3M forecasts with lead times of 2, 4 and 6 weeks. Data for slope and intercept show ordinary least squares estimates ± standard error.

	Landsat GP			**MODIS AR**		
	2	**4**	**6**	**2**	**4**	**6**
		weeks			**weeks**	
R^2-score	0.99	0.95	0.87	0.99	0.96	0.88
RMSE	1.8	4.2	6.8	1.8	4.3	7.0
slope	1.0±0.0	1.1±0.0	1.1±0.0	1.0±0.0	1.0±0.0	1.0±0.0
intercept	-1±0	-2±0	-3±0	0±0	0±0	0±0

and demonstrates that there is substantial forecast skill from each method at each lead time (R^2-scores are substantial), and that the forecasts are unbiased (slopes are all approximately 1, and intercepts approximately 0). Corresponding results for NDVI anomaly time series can be found in the Supplementary Material, in Fig. 3.10 and Table 3.8. The much higher R^2-scores for VCI3M compared to NDVI anomaly is due to the fact that VCI3M is the 12 week composite of weekly VCI observations, and thus there is much greater correlation between VCI3M at time t and VCI3M at time $t+(2, 4, 6)$ weeks due to overlap in the composited weeks. Notwithstanding this, each forecasting method performed substantially better than a persistence forecast of VCI3M (i.e. taking the most recent observation to be the forecast value). For example, the AR method on the MODIS data achieved an RMSE of approximately half that of the persistence forecast for a lead time of 4 weeks, see Fig. 3.13 in the Supplementary Material.

For both methods, the forecast time series sometimes lag behind the true time series, since changes not foreseen by the models are incorporated only once they are observed, see Fig. 3.5 for examples. The two methods sometimes make different types of error. The GP method is more likely to predict a value that is closer to the long-term mean than the true value. This is because *a priori* to taking into account the most recent observations, the model assumes the forecast observation will be equal to the long term mean. By contrast, the AR method is more likely to predict a continuation of the recent trend. This is because the model assumes a continuation of the recent frequency profile, so if a faster-than-average trend is seen in either direction, the trend will be predicted to continue (Hamilton, 1994).

Due to the presence of non-interpolated gaps in the MODIS time series, there were weeks when a forecast assessment was not carried out on these data, see Table 3.5 for details. For 15 of the regions, a 4 week forecast could be made on more than 90 percent of weeks; however, for some regions, a forecast could rarely be made, see Fig. 3.16 in the Supplementary Material.

Additional checks were included to test the sensitivity of the methods to drought severity

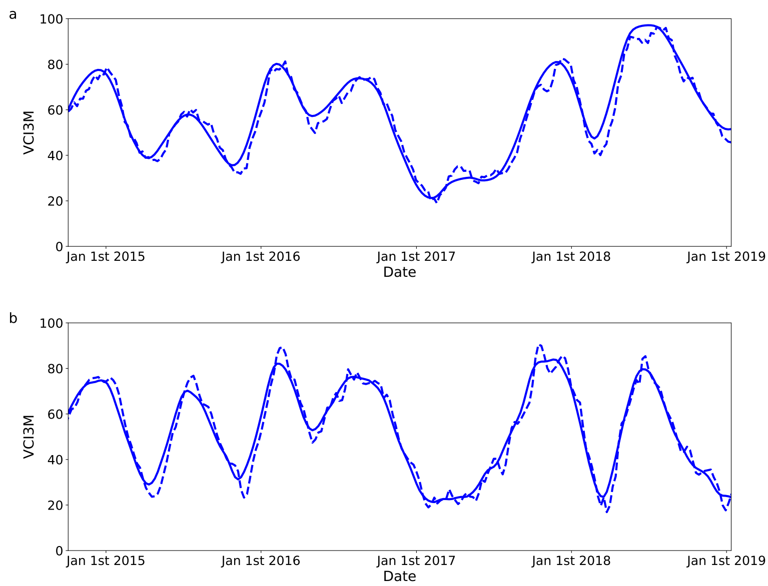

Figure 3.5: Sample aggregate VCI3M

time series from the intersection of Baringo county and livelihood zone 24 from (a) Landsat and (b) MODIS (solid lines). Dotted lines show forecasts at a lead time of 4 weeks, as given by the GP method on the Landsat data, and the AR method on the MODIS data.

and seasonality. The methods, when computed separately for each of the five categories on the NDMA drought scale (Klisch et al., 2016), perform better in terms of absolute RMSE when there was a state of drought than when the vegetation condition was normal or wet (Table 3.2). This could be explained by the fact that when conditions are relatively normal, the subsequent conditions could go in various directions. However, when there is an extreme drought, it is likely that it will persist (because vegetation cover is already below-normal). Note though that the relative RMSE as a proportion of the VCI3M value appears to be similar during drought than during normal vegetation condition. Neither method exhibited large seasonal differences, see Fig. 3.6. However, for the GP method applied to Landsat data at a 6-week lead time, RMSE is high at the start of the season but drops noticeably during the course of the season, suggesting that it is harder to predict how a season will start but easier to forecast how it will proceed once started.

The fact that seasonal differences in RMSE are not substantial provides reassurance that the forecast accuracy estimates are not inflated by the gap-filling during preprocessing. If this were the case, there would be a sustained drop in RMSE during the more overcast months of the year (March to May and October to December). While the GP forecasts on the Landsat data were computed from time series on which no future data were used for the interpolations (see 3.3.1), for the MODIS data interpolations did make use of future data. Therefore, to obtain

Table 3.2: RMSE in VCI3M forecast, for the true vegetation condition belonging to the different categories of drought, at lead times of 2, 4 and 6 weeks. Drought categories are defined by the VCI3M index: wet by VCI3M>50; normal by 35<VCI3M<50; moderate drought by 20<VCI3M<35; severe drought by 10<VCI3M<20; and extreme drought by VCI3M<10. (The extreme drought criterion was not met in any of the Landsat data.)

Drought category	**Landsat GP**			**MODIS AR**		
	2	**4**	**6**	**2**	**4**	**6**
		weeks			**weeks**	
Wet	2.2	5.3	9.0	2.2	4.8	7.5
Normal	1.7	3.4	5.0	1.6	4.0	6.5
Moderate drought	1.5	3.2	5.0	1.5	3.7	5.7
Severe drought,	1.1	2.5	5.5	1.4	3.3	5.4
Extreme drought				1.1	2.9	4.8

further reassurance that performance estimates of AR forecasting on the MODIS data were not inflated, a plot was made of RMSE at 4 weeks lead time against percentage of pixels from which a good observation was obtained on the date of the forecast, see Fig. 3.14 in the Supplementary Material. There was no apparent correlation (Pearson coefficient was 0.01), and hence it was concluded that the gap-filling was not leading to inflated forecast performance.

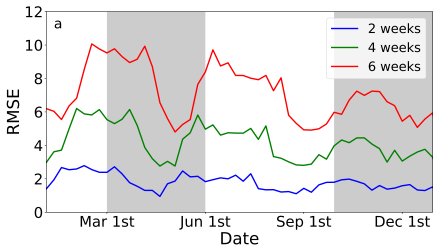

Figure 3.6: RMSE of VCI3M forecast for each week of the year. (a) GP forecasting on Landsat data. (b) AR forecasting on MODIS data. Grey shading indicate the rainy seasons, March-May and October-December.

3.4.2 Drought event forecast: ROC curves

To assess the usefulness of the AR and GP methods for drought forecasting, we tested their ability to detect specific drought events, as defined by the NDMA's alert threshold VCI3M<35 (Klisch et al., 2016). ROC curves were plotted for the detection of VCI3M<35 at lead times of two, four and six weeks, see Fig. 3.7(a, b). These curves show the probability of predicting a state of drought (VCI3M<35) when there will be a state of drought, i.e. hit rate, against the probability of predicting drought when there will not be drought, i.e. false alarm rate, for varying binarisation thresholds on the forecast. We further tested the ability to detect the onset of drought, see Fig. 3.7(c, d). For this, the hit rate was defined as the probability of predicting a transition from the normal condition (VCI3M>35) to the drought condition (VCI3M<35), given that this transition occurs; and the false alarm ratio was defined as the probability that a prediction of this transition is incorrect, given that this transition has been predicted. These curves give an indication that both methods have high skill at forecasting droughts, as measured by the VCI3M, as far as six weeks ahead. The AR forecast appears to do better than the GP forecast at predicting transitions from normal conditions to drought, which can be explained by the tendency of the AR forecast to predict a continuation of the recent trend, while the GP forecast is more likely to predict a reversion to the long term mean value.

The ROC curve performance is not highly dependent on the region (see Table 3.3). Even for the wetter Eastern regions, for which observations are sparser due to cloud cover, the hit and false alarm rates only differ by 1 to 2 percentage points compared with those computed across all regions. Further, ROC curves for predicting the NDMA drought categories of severe (10<VCI3M<20) or extreme (VCI<10) drought look similar to those for detecting VCI3M<35, see Fig. 3.17 in the Supplementary Material.

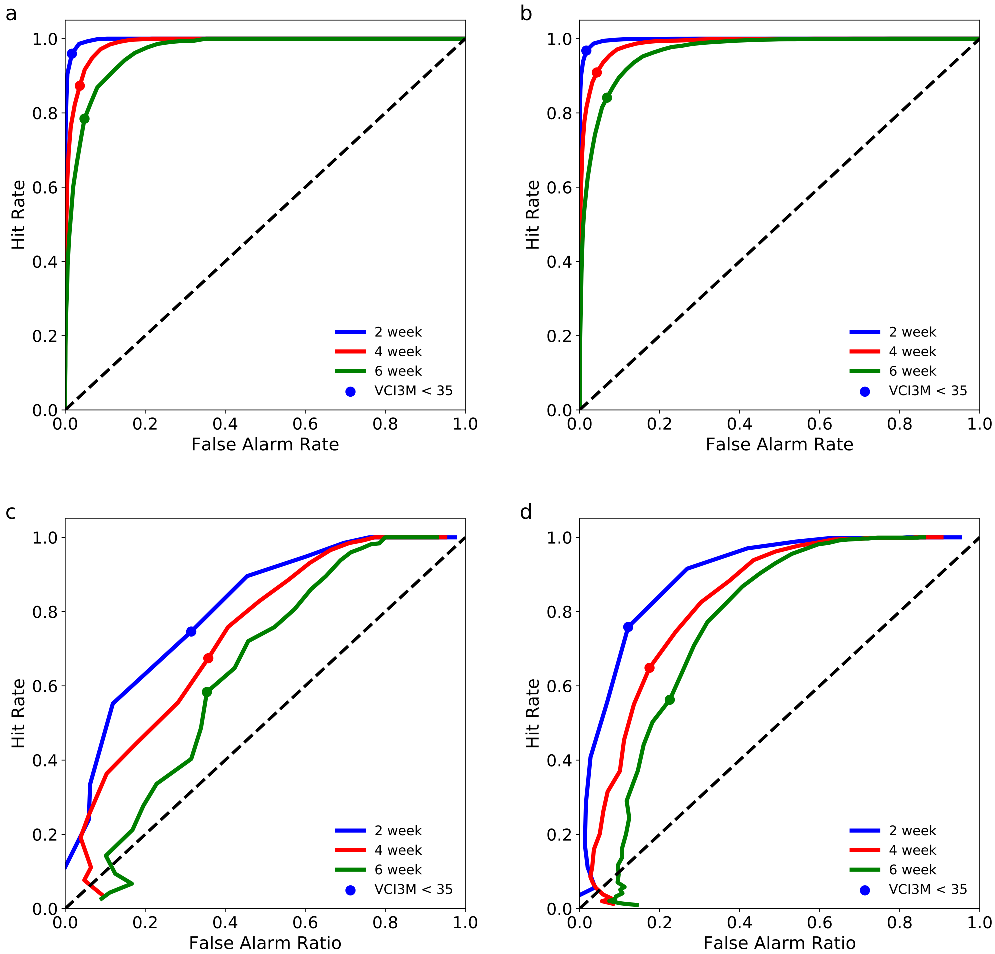

Figure 3.7: (a) ROC curve for drought detection (VCI3M $<$ 35) for lead times of 2, 4 and 6 weeks using the GP method on Landsat data. (b) ROC curve for drought detection using the AR method on MODIS data. (c, d) Respectively for the GP method on Landsat data and the AR method on the MODIS data, hit rate versus false alarm ratio for forecasting a transition to drought (VCI3M$<$ 35) given that the vegetation condition is normal (VCI3M$>$ 35) on the date of the forecast. The curves are plotted from applying different thresholds to convert the continuous forecast into a binary forecast of drought or no drought, see text for details. The shaded circles show the point obtained from forecasting drought when the predicted VCI3M$<$35.

Table 3.3: False alarm rate and hit rate (respectively, expressed in percent) for different regions in Kenya and at different lead times. This is based on forecasting drought if the predicted VCI3M is less than 35 (different performances could be obtained with different warning thresholds (see Figure 3.7. Regions are composed of the following zones: North – Z1,3 and 5; East – Z7, 9, 10 and 11 and South – (Z15 and 18))

Regions	**Landsat GP**						**MODIS AR**					
	2		**4**		**6**		**2**		**4**		**6**	
	weeks						**weeks**					
All	2	96	4	87	5	78	2	97	4	91	7	84
Z24	2	99	4	91	5	82	2	98	5	94	8	88
North	1	97	2	88	3	76	2	98	6	93	11	87
East	3	94	5	85	6	77	3	97	6	91	10	85
South	1	96	3	88	4	77	2	98	6	94	11	90

3.5 Discussion

Droughts are complex and hence inherently difficult to define and measure (Mishra et al., 2010). A large number of satellite-based indicators have been developed to identify meteorological, hydrological, and agricultural droughts (Zargar et al., 2011; AghaKouchak et al., 2015) with each performing well in space and time to a certain degree (Zhang et al., 2017). This paper uses two machine-learning methods to provide short-term forecasts of the 3-month VCI (VCI3M), which is used by Kenya's National Drought Management Authority (NDMA) in their drought Early Warning System (EWS). We have investigated the skill and robustness of our forecasts in a number of ways. Both of our methods showed high sensitivity and specificity for prediction of drought conditions (VCI3M<35) as well as the onset of drought, at lead times of 2, 4 and 6 weeks (see Fig. 3.7). They also perform better than a persistence forecast (a factor of two in RMSE for VCI3M, and R^2 improvement of 0.12). Compared to a similar study that used a Artificial Neural Network model to predict future VCI for four Kenyan counties (Adede et al., 2019), our forecasts provide higher skill, as they showed R^2-scores of 0.78 for a 1-month VCI3M forecast compared to 0.95 and 0.94 for our methods. Moreover, our two methods provide robust results with either dataset (i.e., MODIS and Landsat), and are not impacted by the preprocessing steps. Finally, the methods present a high skill in forecasting drought irrespective of the region, the drought category, and the season.

A very important strength of our methods is the high level of skill. It is instructive to understand the origin of this skill; particularly as it may be surprising since the methods are rather simple and do not include other variables (e.g., rainfall, precipitation). Part of the explanation is that, by using the indicators themselves to determine the forecast, we do track all the factors that impact the vegetation (e.g., disease, soil memory and land-use change) and not just meteorological factors. Additionally, the natural growth cycles of vegetation and their response to environmental factors introduce temporal correlations (persistence) in the indices which can be exploited in short-range forecasting. Furthermore, the VCI3M metric used in this study, and by the NDMA, is additionally smoothed over three months. This smoothing adds temporal dependency, which in turn increases the measured skill. It is important to recognise that this last improvement in skill comes from the inclusion of current data, so it is not a pure forecast and skill metrics should not be directly compared with skills of e.g. VCI forecasts. Nevertheless, the high apparent skill is extremely valuable to disaster risk managers who need to make decisions based on uncertain information (Wilkinson et al., 2018).

There are an increasing number of forecasting studies and methodologies for pastures that focus on several indices, with different lead times, and with varying skill (Matere et al., 2019; Adede et al., 2019; Papagiannopoulou et al., 2017; Meroni et al., 2014). Often, new forecast information developed by scientists to help the development and humanitarian sectors enhance disaster preparedness and response goes unused due to a “usability gap” between knowledge producers and users (Lemos, Kirchhoff et al., 2012). In our study, we aim to bridge this gap by focusing on VCI3M, a drought indicator that is currently used by the NDMA to classify drought severity in the arid and semi-arid regions of Kenya. Additionally, decision makers need reliable forecasts to develop robust anticipatory actions in order to mitigate the impacts of drought with limited financial resources (Wilkinson et al., 2018). Our methods provide skillful VCI3M forecasts with detailed information of hit rates and false-alarm ratios, which are often used to define anticipatory actions (see the Red Cross/Red Crescent Forecast-based Financing (FbF) manual for more information*). Finally, the methodology is also rather simple and easy to implement as it only relies on one data input derived from satellites, which are available globally. The methods can thus be applied everywhere, providing there is sufficient capacity and calibration data. Such co-production strategies allow us to bridge the usability gap (Dilling et al., 2011; Lemos, Arnott et al., 2018) and provides confidence that our forecasting methods may be used.

We have also concentrated on methods that produce accurate short-term forecasts, rather than less-certain, but longer-range forecasts. We can speculate that while the latter might

*http://fbf.drk.de/fileadmin/Content/Manual_FbF/01_Manual/01_Manual_For_Forecast-Based-Financing.pdf have greater value, the former might be more readily adopted in the monthly county bulletins released by the NDMA. Indeed, the forecasts developed here could, for example, help establish a new drought phase classification ('Early Alert') which, along with adequate preparedness actions developed by the disaster risk managers, would minimise the risk of a worsening drought condition. Anticipatory drought management strategies based on this 'Early Alert' could for example focus on livestock vaccination programmes, livestock movement monitoring, or the repair of strategic water sources which enhance the resilience of these communities before a drought hits.

Forecasts alone do not necessarily lead to good anticipatory actions. Whilst acting ahead of disasters is on average more financially effective than responding to an event (Venton et al., 2012), traditionally the humanitarian agencies tend to respond to disasters as financial resources are only available during or after an event. Additionally, due to the uncertainty in the models, anticipatory actions based on such forecasts do raise the risk of "acting in vain", which may have substantial negative impact on the humanitarian sector in the short term (Lopez et al., 2018). These agencies thus need access to adequate financial resources, e.g. FbF (Coughlan de Perez et al., 2015) to fund anticipatory actions based on skill-assessed forecast in order to factor in the possible negative consequences of acting in vain. For the forecast methods developed in this study, the chance of acting in vain will be low due to the high level of skill, which will ultimately lower the barriers to uptake.

3.6 Caveats and Future Work

As discussed above, our methods are already sufficiently skillful that they are usable as they stand. However, we have identified some minor limitations and relevant improvements to enhance the functionality, skill, lead-time and impact of our forecasts.

Our analysis has been based on relatively small samples of the available pixels, aggregated, spatially, at the level of the pastoral livelihood zone and county intersections. This limits the localisation specificity of our predictions. Additionally, our methodology using Landsat merges data from various land covers which may reduce accuracy. The processing of all pixels can be achieved within reasonable computational constraints and will allow us to aggregate over specific regions of interest. For example, one could perform the analysis for specific land covers within a county, or for individual grazing units, which would provide greater accuracy and additional functionality.

Our forecasts are currently unavailable, or are less accurate, in periods during or following cloud-cover gaps. More subtly, our validation will have favoured dry season observations, which are less affected by cloud cover, and this will have an impact on the validation of the forecast performance. However, as we found little variation in performance throughout the seasons we do not think these are significant problems. The impact of cloud cover will be reduced when all pixels are processed and aggregated.

As with any machine-learning method, the forecast and its estimated skill are only appropriate for the types of vegetation and environments for which it has been calibrated. The quality that we have obtained in different regions of Kenya gives us some confidence that the skills will not be substantially different. Nevertheless, the calibration, validation and skill assessment of the forecast will be an essential element of a practical and general tool. Future work should also explore how long a temporal baseline is required for good calibration. This will, in principle, allow a truly global forecasting tool.

The error estimates we currently provide in our forecasts are derived from the global validation of our performance. They should thus be correct on average but the errors will be overestimated in some situations while, correspondingly, underestimated in others. Future development can provide error estimates that are tailored for the specific conditions and data availability.

The indicators we have chosen are well motivated through their use in the existing EWS operated by NDMA. Identifying the most appropriate and useful indicators of such hazards is the subject of much debate and investigation (see e.g., (Zargar et al., 2011; AghaKouchak et al., 2015; Zhang et al., 2017)). But they may not be the most suitable to quantify the relevant socio-economic impacts of droughts (Sutanto et al., 2019). Subject to data availability, similar machine-learning approaches could be applied to more direct socio-economic indicators tracking food insecurity, such as malnutrition, food prices, or livestock condition.

Perhaps the most-significant limitation of our methods is that they are only appropriate for relatively short lead-times. Although a 4-week lead time can be useful, most contingency plans and drought preparedness policies are developed over seasonal timescales. It is thus key to extend this lead-time. While current observations of precipitation and temperature had little impact, including other observed climate variables (e.g., ENSO, sea surface temperature, (Funk et al., 2008)) or seasonal climate forecasts may enhance skill and lead times.

Future research will also be required on the effectiveness of the practical implementation of forecasts in EWS (Wilkinson et al., 2018). Clearly-defined triggers (e.g., threshold values based on forecasts, which may vary in time and space) will need to be defined and assessed and optimised against suitable performance metrics. Similarly, effective anticipatory actions need to be defined by the decision makers in relation to these triggers. Adequate policy and institutional arrangements will be needed to allow the various actors to engage and interact with a long-term perspective on risk management. This in turn, requires financial systems that can be accessed based on such forecasts to be able to act across various timescales before the disaster occurs (i.e. Forecast-based-Finance).

3.7 Conclusion

In conclusion we have developed two new forecasting methods which exploit the inherent temporal correlation in vegetation indices to provide highly skillful, short-range forecasts of VCI. The choice of input data, output indicators, simplicity of implementation, and demonstrated skill argues that these methods will be useful for drought early warning systems. We have identified ways this can be improved, but there is clear evidence here that our statistical persistence model provides strong skill over useful lead times. This can be an important contribution to anticipatory drought risk management in Kenya

Authors responsibilities

A.B.B., S.D. and E.S. are lead authors as they contributed equally to the paper and the order of the three names is alphabetical. A.B.B was responsible for developing and running the AR method. S.D. was responsible for developing and running the GP method, and for the accumulation and processing of the Landsat data. E.S. was responsible for the MODIS data accumulation and preprocessing. JMw, SO and PR developed the initial idea and provided feedback throughout. All authors wrote, reviewed and edited the final manuscript.

Acknowledgements

This research was funded by the Science and Technology Facilities Council (STFC) through the following projects: "AstroCast: Applying Astronomy Data Analysis to enhance disaster forecasting" – grant number ST/R004811/1; and "STFC Official Development Assistance (ODA) Institutional Award" attached to the same grant; and by the Science for Humanitarian Emergencies and Resilience (SHEAR) con- sortium project "Towards Forecast-based Preparedness Action" (ForPAc, www.forpac.org), grant numbers NE/P000673/1, funded by the UK Natural Environment Research Council (NERC), the Economic and Social Research Council (ESRC) and the UK Department for International Development (DfID). EES acknowledges support from the UK Newton Fund as part of the Development in Africa with Radio Astronomy (DARA) Big Data project delivered via STFC under grant number ST/R001898/1 and JMM was funded by the NERC-SHEAR Studentship Cohort grant number NE/R007799/1. This project was initiated through pump-priming funding from the University of Sussex's "Sussex Research" thematic programme and carried out as part of the interdisciplinary Data Intensive Science Centre at the University of Sussex (DISCUS). We acknowledge early contributions to that pilot work from Peter Hurley, Philip Rooney, Martin Jung, and Jörn Scharlemann. Martin Todd and Alexander Antonarakis provided useful feedback which improved the manuscript.

3.8 Supplementary Material

3.9 Data selection and comparison of datasets

3.9.1 Landsat

Landsat-5, 7 and 8 (Roy et al., 2014) red and near infrared (NIR) surface reflectances and quality assessment (QA) data over the 10 pastoral livelihood zones of Kenya, from January 1st, 2000 to February 1st, 2019, were obtained using the United States Geological Survey (USGS) EarthExplorer. Specifically, data from 1 000 pixels within each region were drawn from the Level-1 Precision Terrain (L1TP) processed dataset, which has well-characterized radiometry and is inter-calibrated across the different Landsat sensors. The spatial resolution of these data is 30m and the repeat interval is 16 days. Using the QA data, observations classified as clear from clouds or cloud shadows were kept. Pixels with fewer than half of the observations over the full time period were discarded (and replaced with an alternative random selection, with a few exceptions, see Fig. 3.16). The surface reflectances were combined to obtain NDVI.

3.9.2 MODIS

NDVI data were also gathered from the surface reflectances obtained from the daily, 500-meter resolution MODIS Terra/Aqua Nadir BRDF-Adjusted Reflectance product MCD43A4,v006 (Schaaf et al., 2015). Data from February 22nd, 2000 up to February 1st, 2019 were acquired via the NASA Land Processes Distributed Active Archive Center. QA maps files with binary quality flags were used to remove poor quality data resulting from cloud or unreliable BRDF corrections. Data were drawn from 100 pixels within each region, out of those that had been identified as grassland by the MODIS land cover classification maps (MCD12Q1,v006).

3.9.3 Comparison of the two datasets

The key differences between the two datasets are the spatial and temporal resolutions, see Table 3.4. The Landsat data have higher spatial resolution, whilst the MODIS data have higher temporal resolution. Since forecasting was being attempted at the level of large scale regions (livelihood zone and county intersections), and at a weekly temporal resolution, the expectation was that the MODIS data would have advantages, assuming individual Landsat and MODIS observations have similar signal-to-noise ratios. The processed MODIS time series with weekly observations have less measurement noise because they are composites of 7 daily observations (that themselves are 16-day composites of measurements taken every 1-2 days), whereas the

Table 3.4: Table comparing Landsat and MODIS products

Feature	Landsat	MODIS
Spatial Resolution	High resolution at 30 m	Medium resolution ranging from 250 m to 1 km
Temporal Resolution	16-day sampling (8-day when both Landsat-7 and 8 are used	Daily sampling monitoring dynamic variables
Quality	Cloud coverage at 30 m	Cloud coverage at 500 m

processed Landsat time series are derived from more temporally sparse data (up to 3 different Landsat missions, each yielding one observation every 16 days). Landsat data would have advantages in different applications where forecasts on smaller spatial scales are required. The Landsat data also has the advantage that the quality flags and cloud masks are defined on smaller scales.

The differences between the MODIS and Landsat datasets produced slightly different 'true' aggregate time series on which to assess the interpolation and forecasting methods. In addition to the different temporal resolution of the observations supplying the final time series, the MODIS data were aggregated across 100 random grassland pixels from each region, whereas the 1 000 Landsat pixels analysed were randomly distributed over the whole of each region. In choosing how many pixels to analyse per region, there is a trade-off between using a larger number of pixels for higher accuracy, and a smaller number of pixels for lower computational cost. Fewer MODIS pixels were used than Landsat pixels since they correspond to larger spatial regions. Both these choices of number of pixels should be sufficient for high accuracy of results, since for Landsat data the R^2-score comparing the average of all pixels from a region with the average of 100 or 1 000 random pixels was 0.990 and 0.9993 respectively. The MODIS grassland classification was not available at Landsat resolution, thus unambiguous classification of the smaller Landsat pixels was not possible. This is unlikely to have made much difference to pixel selection, given that the pastoral livelihood zones are mostly grasslands (Fig. 3.15).

3.10 Further details on preprocessing

3.10.1 Gaussian process modelling

A Gaussian Process is a probabilistic model defined as a collection of random variables for which any finite subset has a joint Gaussian distribution (Rasmussen et al., 2006). Formally, for the present application of interpolation or extrapolation of a time series, with observation at time t

denoted by X_t, the model is

$$X_t \sim \mathcal{N}\left[Y(t), \sigma_r^2\right] , \tag{3.6}$$

$$Y(t) \sim \mathcal{GP}\left[m(t), k(t, t')\right] . \tag{3.7}$$

Here $Y(t)$ is the true value of the observed index, and the measurement noise is σ_r, so that an observation X_t is a normal random variable with mean $Y(t)$ and standard deviation σ_r. The true values $Y(t)$ are also normally distributed, with the mean at time t given by the mean function $m(t)$, and the covariance between values at times t and t' given by the kernel function $k(t, t')$. To carry out interpolation or extrapolation from a time series, existing data are used to fit the mean, m, kernel, k, and measurement noise σ_r, and then expected values are produced for the desired times, based on the obtained fit.

For gap-filling on individual Landsat pixel NDVI time series, the model was determined as follows, using the `Pyro` programming package for `Python`. The mean, $m(t)$, was assumed to be constant, and the mean of the whole time series. To determine the kernel, Compositional Kernel Search (Duvenaud et al., 2013) was used. Specifically, a search through all the following kernels, and products and sums of pairs of them was carried out: Linear, Radial Basis Function (RBF), Periodic (with period p set to one year), Rational Quadratic, and Matern. The highest marginal likelihood was achieved by Radial Basis Function (RBF) plus Periodic ($k_{\text{RBF}} + k_{\text{P}}$), so this combination was selected as the kernel:

$$k_{\text{RBF}}(t, t') = \sigma_{\text{RBF}}^2 \exp\left(-0.5 \frac{|t - t'|^2}{l_{\text{RBF}}^2} \right), \tag{3.8}$$

$$k_{\text{P}}(t, t') = \sigma_{\text{P}}^2 \exp\left(-2 \frac{\sin^2(\pi |t - t'| / p)}{l_{\text{P}}^2} \right) . \tag{3.9}$$

There were thus 5 parameters to fit for each time series $(\sigma_r, \sigma_{\text{RBF}}, l_{\text{RBF}}, \sigma_{\text{P}}, l_{\text{P}})$. These were learned using Stochastic Variational Inference (Hoffman et al., 2013).

For the forecasting on the aggregated NDVI anomaly and VCI3M, a pure Radial Basis Function kernel was used, since for these anomaly indices, the periodic component is not present, see Section 3.3 in the main manuscript.

3.10.2 Gap-filling for MODIS

Interpolation of gaps in the raw MODIS time series was not carried out when the length of the gap was longer than a certain maximum, $L_{\max}$. In choosing $L_{\max}$, a trade off between quality and quantity of remaining observations had to be made: a small $L_{\max}$ would lead to fewer

Table 3.5: Comparison of outcomes for different choices of maximum allowed interpolation length $L_{\max}$ on the MODIS data. R^2-score of 4 week AR forecast and the percentage of the time that it was possible to make a forecast, for $L_{\max} = 4$, 6, and 8 weeks. Numbers show the median across all regions.

$L_{\max}$ **(weeks)**	R^2**-score**	**Forecasts attempted (%)**
4	0.60	84
6	0.58	93
8	0.63	98

forecasts being attempted, but interpolations closer to the ground truth, while a large $L_{\max}$ would lead to more forecasts being attempted, but with these forecasts being assessed against interpolations that are potentially far from the ground truth. The choice $L_{\max} = 6$ weeks was made, after exploring a range of values and finding R^2-score to be not sensitive to the precise choice within the range between 4 and 8 weeks, see Table 3.5. Note that interpolation on the Landsat data was carried out for all gaps, since the GP interpolation method makes use of the entire time series, and interpolated values within a long interpolation take values close to the seasonal mean.

Due to the presence of non-interpolated gaps in the MODIS time series, there were weeks when a forecast assessment was not carried out on these data. The criteria for being able to do AR forecasting on these data were: (i) the three most recent weekly aggregated observations had to be present, since these are required for making a prediction; (ii) there had to be an aggregated observation present for the week being forecast, so the quality of the prediction could be assessed.*

3.10.3 Comparison of other possible gap-filling methods

Various gap-filling methods have been used to deal with missing values resulting from the presence of clouds and atmospheric aerosols. These methods are based on either spatial information, temporal information or some combination of both spatial and temporal information (Kandasamy et al., 2013; Weiss et al., 2014). Temporal interpolation was chosen given that spatial interpolation methods suffer from the fact that there are frequently clouds over Kenya that cover large groups of neighbouring pixels (although a possible alternative, not considered here, would be to make use of other pixels that historically behave similarly in time (Cao et al., 2018).

*GP forecasting was still possible when (i) failed, but was also not carried out in that case, since performance would have been worse than usual in this case.

The performance of the temporal gap-filling methods employed, compared with alternative temporal gap-filling methods, was tested by removing observations, applying the method, and then comparing the interpolated observations with the removed observations. GP interpolation and linear, quadratic and cubic polynomial interpolation methods were tested, on both the Landsat and MODIS datasets. R^2-scores were obtained for using the interpolated values to predict the 'true' values for the missing observations.

For the Landsat data, one randomly chosen observation between 1/1/2014 and 1/2/2019 was removed from each of 2000 randomly selected individual pixel time series. From the MODIS data, 2000 random individual pixel NDVI time series (1/1/2014 to 1/2/2019) were chosen. 20 randomly selected NDVI values were dropped from each of the time series and the various gap-filling methods were used to interpolate the dropped values. The results for Landsat are shown in Table 3.6, and for MODIS in Table 3.7. Note that with these methods, the random samples are more likely to come from periods when there are not many gaps. It is an assumption that the results are valid across all periods.

Table 3.6: Comparison of GP method with commonly used interpolation methods as candidates for gap-filling on Landsat data. At the pixel level a random observation was removed, and then interpolated with each of the listed methods.

Method	**R^2-score**
GP	0.67
Linear	0.53
Quadratic	-0.07
Cubic	-1.92
Last value	0.34
Mean value	0.0

For the Landsat data, the GP method achieved the highest R^2-score, thus showing its utility, and justifying our choosing it. The R^2-score of 0.67, achieved by the GP method, is close to the R^2-score of 0.76 which is obtained from using one Landsat observation to predict a second Landsat observation from the same 16-day observation period (of which there were instances in the data). Fig. 3.8 shows a contour plot of the true versus interpolated NDVI observations using this method. This plot shows that the method doesn't introduce any biases- the slope and intercept are approximately 1 and 0 respectively.

For the MODIS data, GP, linear interpolation and quadratic interpolation all performed similarly well. Quadratic interpolation had the highest R^2-score, hence this method was chosen

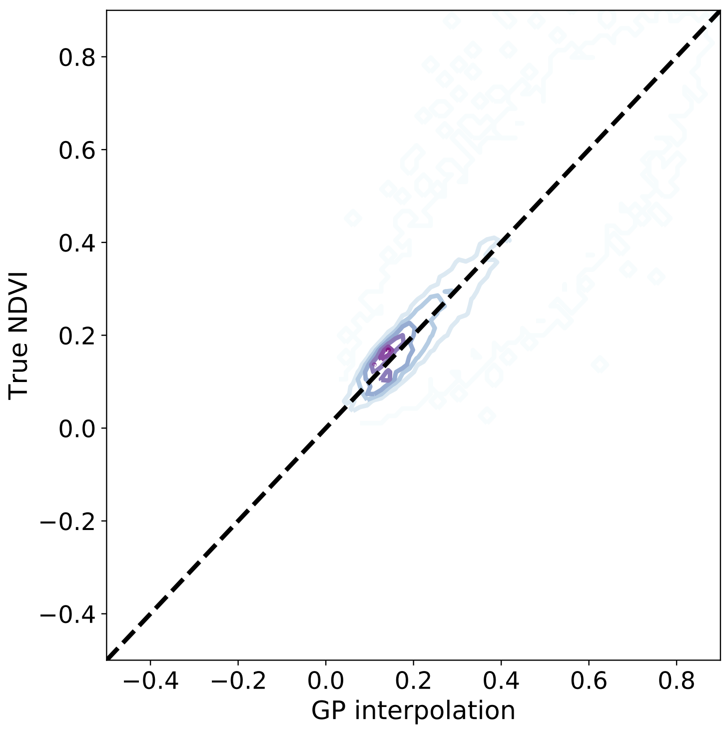

Figure 3.8: Contour plot of Landsat observed and predicted NDVI values from the GP interpolation.

for gap-filling on the MODIS data. The higher interpolation R^2-scores for MODIS, compared to Landsat, imply that the MODIS data is less noisy than the Landsat data. Assuming that observations from MODIS and Landsat have similar signal-to-noise ratio, this can be explained by the higher temporal resolution of MODIS, and the compositing of multiple observations for the weekly gridded MODIS data. Fig. 3.9 shows a contour plot of the true versus interpolated NDVI observations using the quadratic interpolation method. This again demonstrates that the interpolation doesn't introduce biases- the slope and intercept are approximately 1 and 0 respectively.

Table 3.7: Comparison of interpolation methods as candidates for gap-filling on MODIS data.

Method	**R^2-score**
GP	0.92
Linear	0.93
Quadratic	0.94
Cubic	0.92
Last value	0.70
Mean value	-0.02

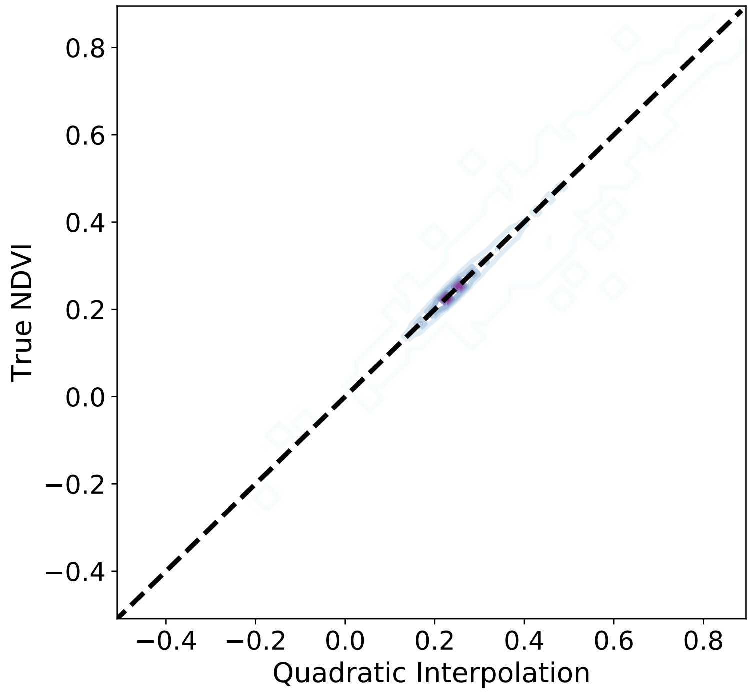

Figure 3.9: Contour plot of MODIS observed and predicted NDVI values from 2000 pixels for gap-filling by quadratic interpolation.

3.11 Further forecast results

Fig. 3.10 shows contour plots of forecast against actual NDVI anomaly data for the two methods, and Table 3.8 shows the R^2-scores, RMSE, slope and intercept from each of these plots. Figs. 3.11 and 3.12 plot the forecast performance of the two methods in terms of percentage of standard deviation remaining S, for lead times of 1 to 10 weeks. For NDVI anomaly, for both methods, S approaches the baseline of 100 as the lead time approaches 10 weeks, while for VCI3M, some forecast skill is still apparent at a lead time of 10 weeks. Fig. 3.13 compares the performance of the AR VCI3M forecast with that of the persistence VCI3M forecast, on the MODIS data; the persistence forecast being simply the most recent observation. The AR forecast performs substantially better than the persistence forecast, for example, achieving a RMSE of approximately half that of the persistence forecast for a lead time of 4 weeks. The GP VCI3M forecast on the Landsat data achieves a similar improvement on the persistence forecast. Fig. 3.14 shows, for the MODIS/AR method, the average RMSE of a 4 week forecast against the percentage of pixels from which there was a clear observation during the week the forecast was made. Fig. 3.16 shows forecast performance region by region. Fig. 3.17 shows alternative ROC curves for drought prediction using the AR method on the MODIS data, based on different thresholds for defining drought.

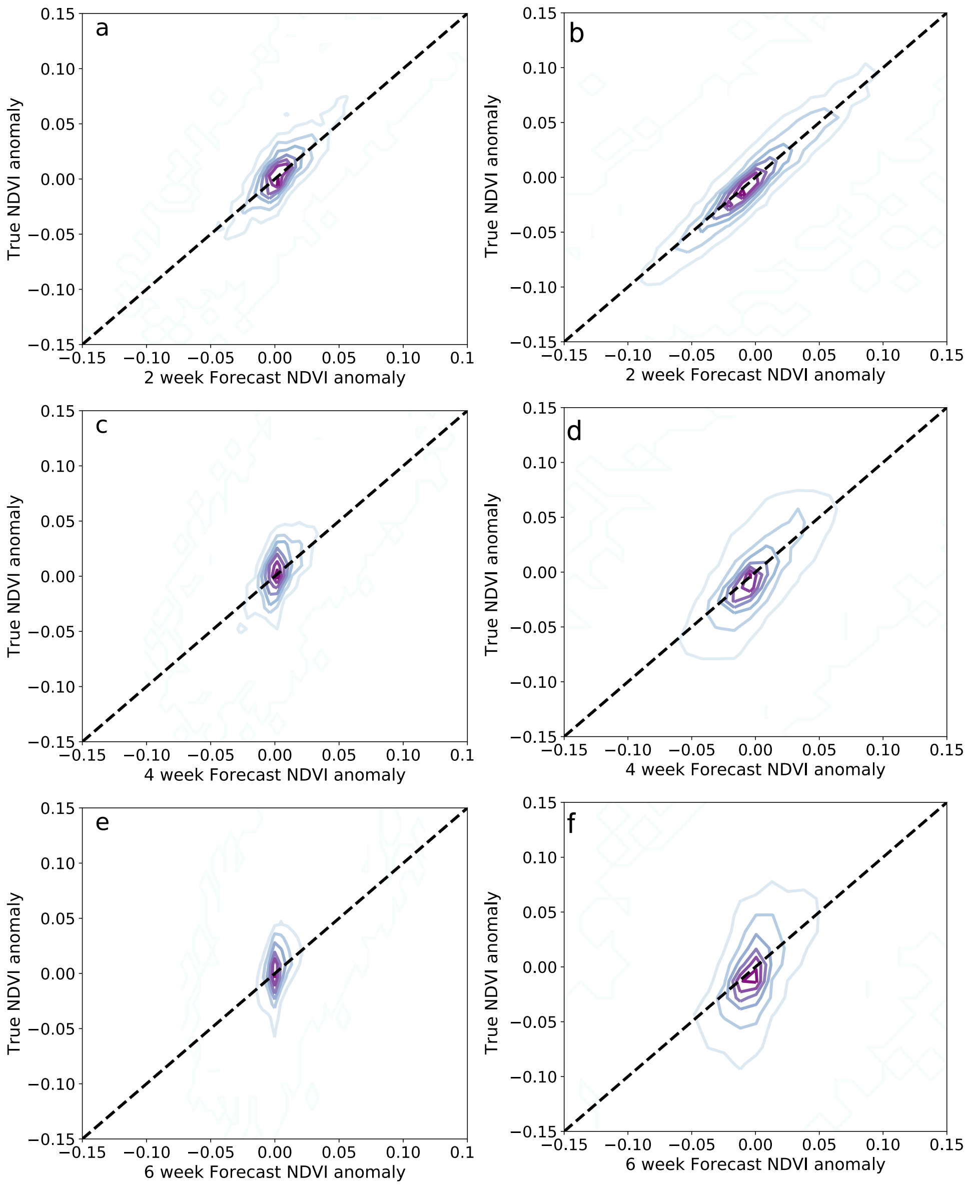

Figure 3.10: Contour plots of NDVI anomaly against two, four and six weeks NDVI anomaly forecasts. (a,c,e) show forecast performance for the GP method on Landsat data, and (b,d,f) show forecast performance for the AR method on MODIS data (across the 19 regions for which a 4 week forecast was possible more than 50% of the time, see main text for details).

Table 3.8: Performance statistics of NDVI anomaly forecasts with lead times of 2, 4 and 6 weeks. Data for slope and intercept show ordinary least squares estimates ± standard error.

	Landsat GP			**MODIS AR**		
	2	**4**	**6**	**2**	**4**	**6**
		weeks			**weeks**	
R^2-score	0.69	0.46	0.27	0.85	0.55	0.33
RMSE	0.029	0.039	0.045	0.025	0.043	0.053
slope	1.1±0.0	1.2±0.0	1.4±0.0	1.0±0.0	1.0±0.0	1.0±0.0
intercept	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0

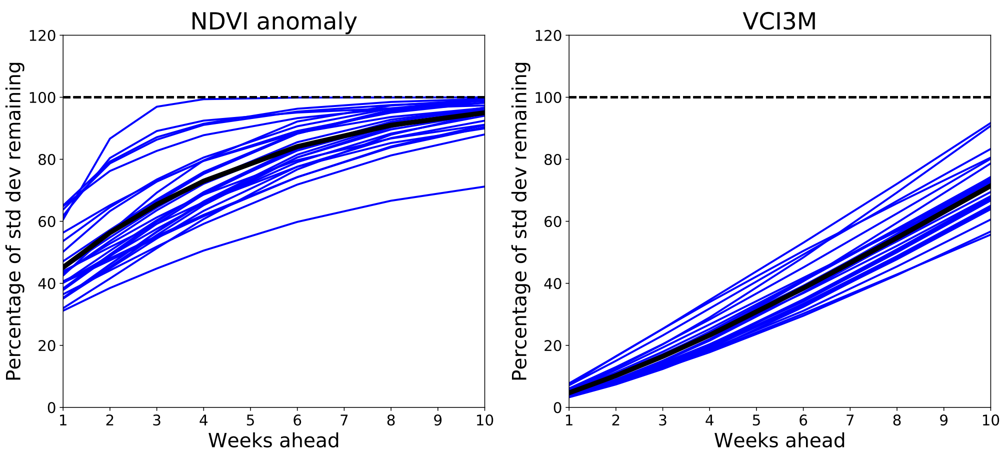

Figure 3.11: Forecast performance with a lead time of 1 to 10 weeks using the GP method on the Landsat data, as given by percentage standard deviation remaining S, for (Left) NDVI anomaly, and (Right) VCI3M. The blue lines show results for the individual regions, and the black line shows the median across all regions.

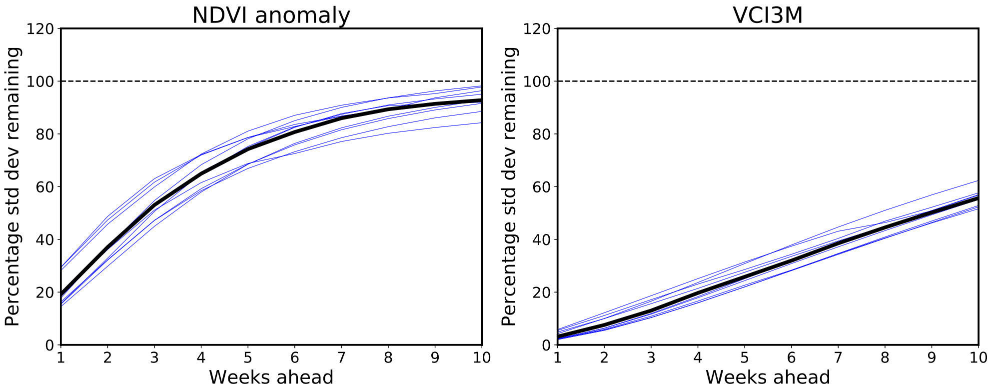

Figure 3.12: Forecast performance with a lead time of 1 to 10 weeks using the AR method on the MODIS data, as given by percentage standard deviation remaining, for (Left) NDVI anomaly, and (Right) VCI3M. The blue lines show results for the individual regions for which a forecast is possible more than 50% of the time, and the black line shows the median across all 19 of these regions.

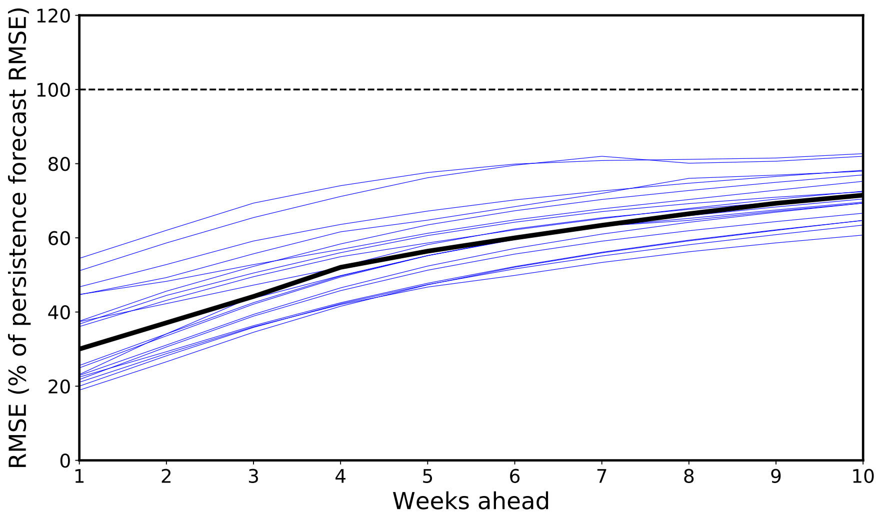

Figure 3.13: Comparison of AR forecast with persistence forecast on the MODIS data. For lead times of 1 to 10 weeks, the RMSE of the AR forecast as a percentage of the RMSE of the persistence forecast. The blue lines show results for the individual regions for which a 4 week forecast is possible more than 50% of the time, and the black line shows the median across these regions.

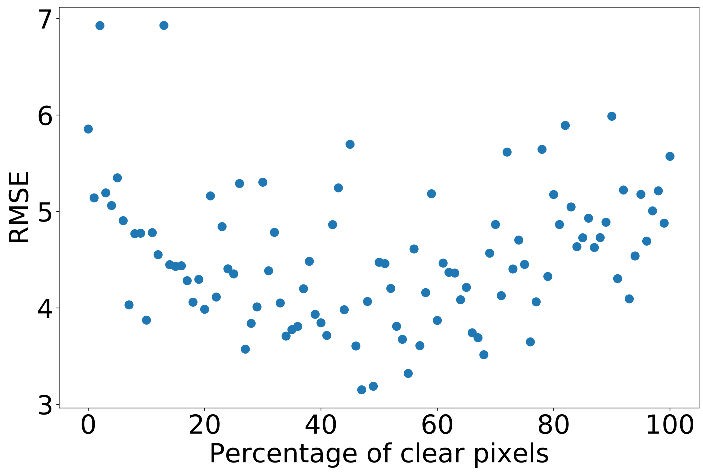

Figure 3.14: RMSE of 4 week forecast against percentage of clear pixels at most recent observation, for the AR method on the MODIS data. Plotted points are RMSE for each integer percentage of clear pixels. The Pearson correlation here is 0.01.

3.11.1 Effect of including observations from other regions in the AR model

For the MODIS data, we tested to see whether we could improve the prediction of VCI3M by incorporating the past of VCI3M from a distinct region in the AR model, i.e. Granger causality analysis was performed. Taking X as the VCI3M of the region to be forecast, as in equation (3.3), and Y to be the VCI3M from another region, the extended model was fit:

$$X_{t+n} = \sum_{i=0}^{p-1} a_i X_{t-i} + \sum_{i=0}^{q-1} b_i Y_{t-i} + \epsilon'_t \,, \tag{3.10}$$

and Granger causality measured as the percentage reduction in RMSE obtained when the extended model is used instead of the previous (reduced) model (3.3).

We tested for Granger causality of VCI3M from each region to each other region (within the set of regions for which predictions could be made more than 50% of the time). That is, for each pair of distinct regions, i and j, the 3 most recent observations from region j were added to the AR forecast model for region i, and the RMSE was compared with that obtained without including observations from region j. There was not strong Granger causality of VCI3M between most regions. For only a few combinations was there a reduction in RMSE of more than 5%, see Fig. 3.18. Nevertheless, these results suggest that, to create the optimal linear regression based forecasting method, data from all regions should be used. Future work will explore how best to extract any useful information from regions other than the one being forecast.

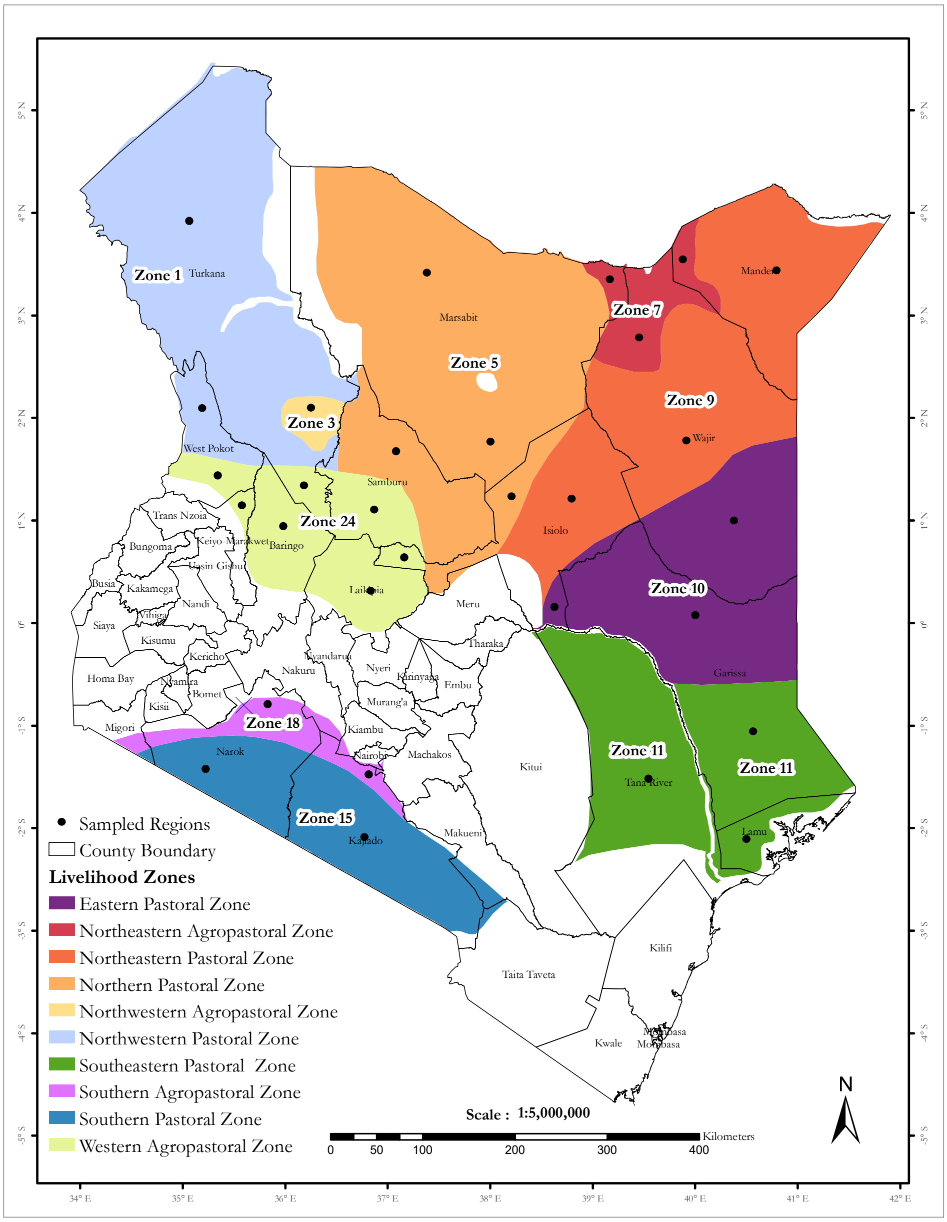

Figure 3.15: Map of Kenya showing the livelihood zones from which pixels were sampled.

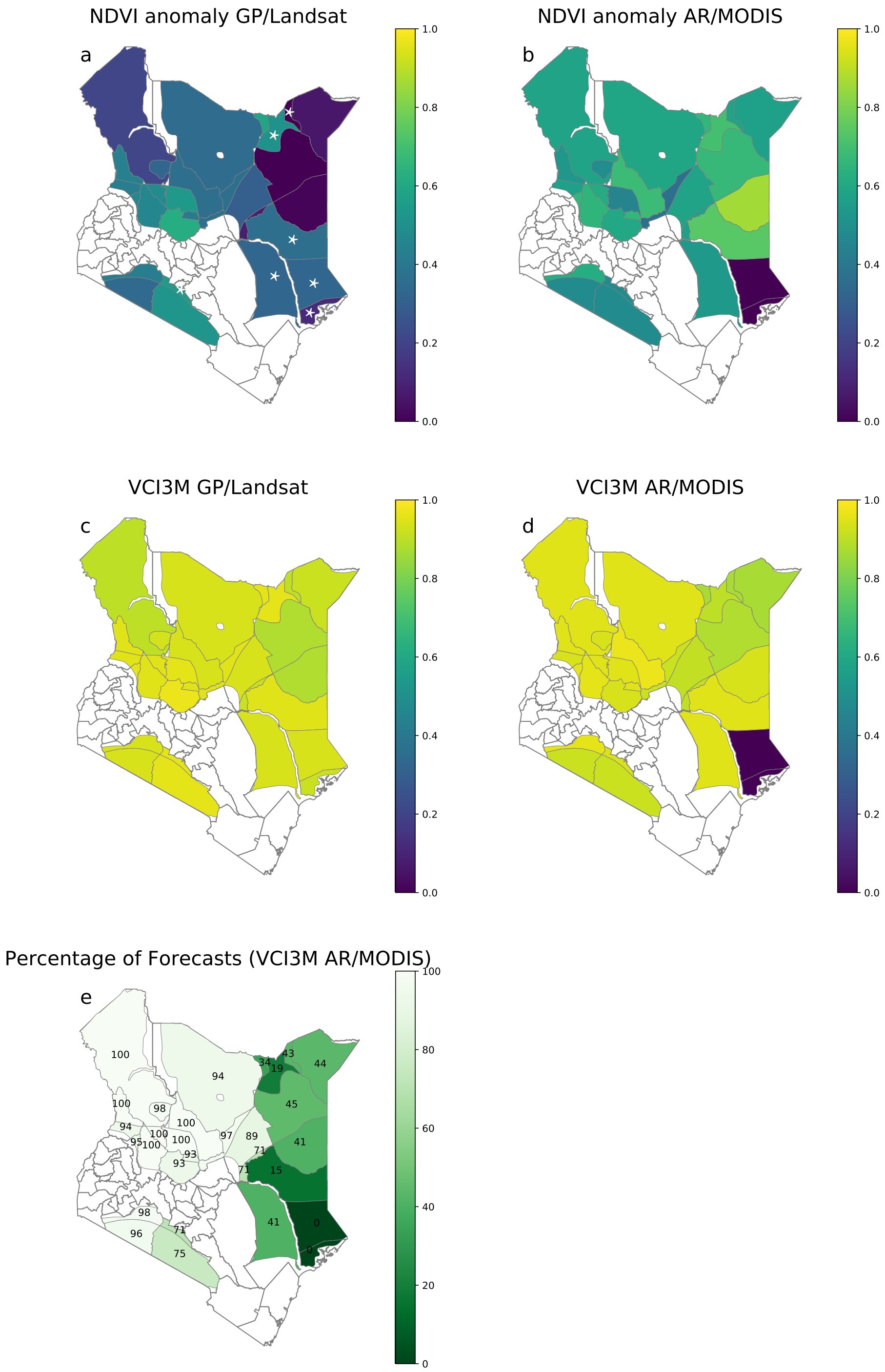

Figure 3.16: Maps of NDVI anomaly and VCI3M 4 week forecast performance region-by-region for: (a) NDVI anomaly with GP method on Landsat data; (b) NDVI anomaly with AR method on MODIS data; (c) VCI3M with GP method on Landsat data; (d) VCI3M with AR method on MODIS data. In (a), asterisks indicate regions where selected pixels had a minimum of 180, rather than 250, clean observations. (e) shows the percentages of weeks that the AR method provided a 4 week VCI3M forecast.

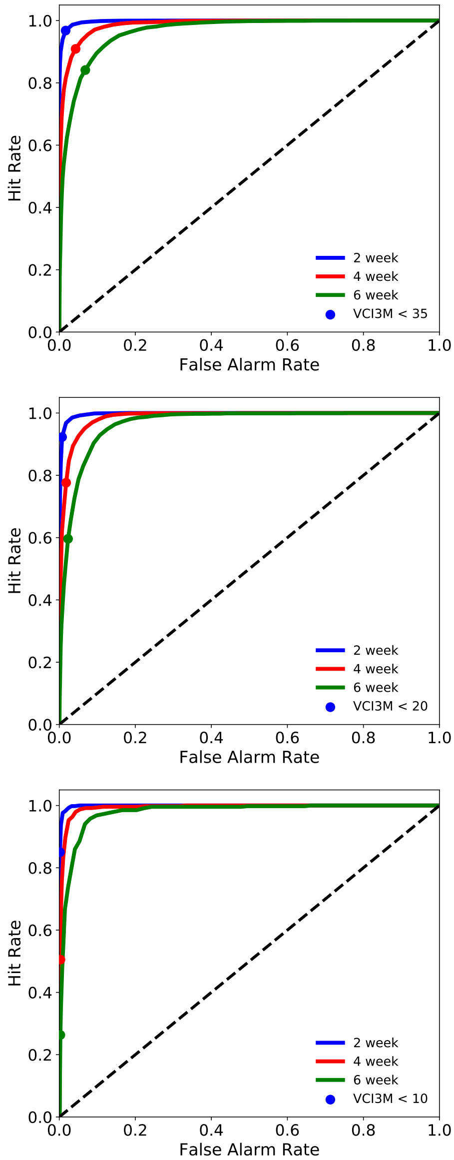

Figure 3.17: ROC curves for predicting drought with drought defined at various NDMA thresholds. Possible hit rates against possible false alarm rates for the AR method on the MODIS data for the detection of: (Top) Any drought, VCI3M<35, (Middle) Severe or extreme drought VCI3M<20, (Bottom) Extreme drought

VCI3M<10.

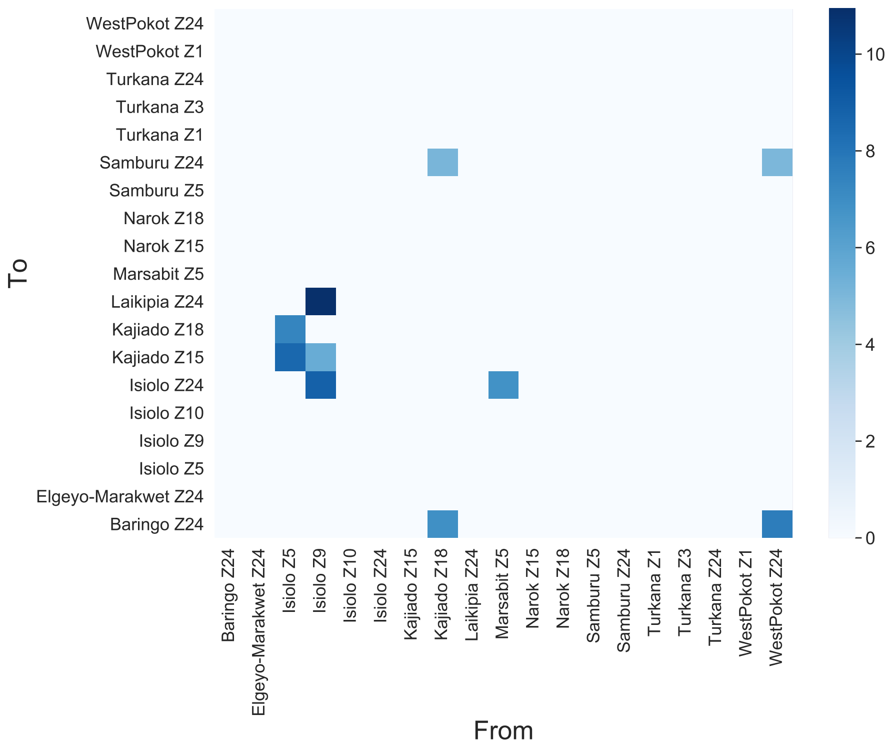

Figure 3.18: Granger causality of VCI3M

from each region to each other region, computed on the MODIS data, measured as percentage reduction in RMSE when observations from region ‘From’ are added to the AR model for forecasting region ‘To’ at a lead time of 4 weeks. Only substantial Granger causalities are shown, i.e. those with percentage reduction in RMSE of more than 5%.

References

Adede, Chrisgone, Robert Oboko, Peter Waiganjo Wagacha and Clement Atzberger (May 2019). "A Mixed Model Approach to Vegetation Condition Prediction Using Artificial Neural Networks (ANN): Case of Kenya's Operational Drought Monitoring". In: *Remote Sensing* 11.9, p. 1099. ISSN: 2072-4292. DOI: 10.3390/rs11091099. URL: https://www.mdpi.com/2072-4292/11/9/1099.

AghaKouchak, A, A Farahmand, FS Melton, J Teixeira, MC Anderson, Brian D Wardlow and CR Hain (2015). "Remote sensing of drought: Progress, challenges and opportunities". In: *Reviews of Geophysics* 53.2, pp. 452–480.

Asoka, Akarsh and Vimal Mishra (2015). "Prediction of vegetation anomalies to improve food security and water management in India". In: *Geophysical Research Letters* 42.13, pp. 5290–5298. DOI: 10.1002/2015GL063991.

Behnke, Roy and David Muthami (2011). *The Contribution of Livestock to the Kenyan Economy.* Tech. rep. URL: https://cgspace.cgiar.org/bitstream/handle/10568/24972/IGAD%7B%5C_%7DLPI%7B%5C_%7DWP%7B%5C_%7D03-11.pdf?sequence=1%7B%5C&%7DisAllowed=y.

Bhuiyan, C., R.P. Singh and F.N. Kogan (2006). "Monitoring drought dynamics in the Aravalli region (India) using different indices based on ground and remote sensing data". In: *International Journal of Applied Earth Observation and Geoinformation* 8.4, pp. 289–302. ISSN: 0303-2434. URL: http://www.sciencedirect.com/science/article/pii/S030324340600016X.

Brahim-Belhouari, S. and J. M. Vesin (Aug. 2001). "Bayesian learning using Gaussian process for time series prediction". In: *Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing (Cat. No.01TH8563)*, pp. 433–436. DOI: 10.1109/SSP.2001.955315.

Brown, Molly E. and Varun Kshirsagar (2015). "Weather and international price shocks on food prices in the developing world". In: *Global Environmental Change* 35, pp. 31–40. ISSN: 0959-3780. DOI: https://doi.org/10.1016/j.gloenvcha.2015.08.003. URL: http://www.sciencedirect.com/science/article/pii/S0959378015300248.

Camps-Valls, G., J. Verrelst, J. Munoz-Mari, V. Laparra, F. Mateo-Jimenez and J. Gomez-Dans (2016). "A Survey on Gaussian Processes for Earth-Observation Data Analysis: A Comprehensive Investigation". In: *IEEE Geoscience and Remote Sensing Magazine* 4.2, pp. 58–78. ISSN: 2168-6831. DOI: 10.1109/MGRS.2015.2510084.

Cao, Ruyin, Yang Chen, Miaogen Shen, Jin Chen, Ji Zhou, Cong Wang and Wei Yang (2018). "A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky-Golay filter". In: *Remote Sensing of Environment* 217, pp. 244–257. ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2018.08.022. URL: http://www.sciencedirect.com/science/article/pii/S0034425718303985.

Chandola, Varun and Ranga Raju Vatsavai (2011). "A scalable Gaussian process analysis algorithm for biomass monitoring". In: *Statistical Analysis and Data Mining: The ASA Data Science Journal* 4.4, pp. 430–445. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/sam.10129.

Coughlan de Perez, E., B. van den Hurk, M. K. van Aalst, B. Jongman, T. Klose and P. Suarez (Apr. 2015). "Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts". In: *Natural Hazards and Earth System Sciences* 15.4, pp. 895–904. ISSN: 1684-9981. DOI: 10.5194/nhess-15-895-2015. URL: https://www.nat-hazards-earth-syst-sci.net/15/895/2015/.

EM-DAT (2019). *The International Disaster Database*. Université catholique de Louvain, Belgium. URL: www.emdat.be.

Dilling, Lisa and Maria Carmen Lemos (2011). "Creating usable science: Opportunities and constraints for climate knowledge use and their implications for science policy". In: *Global environmental change* 21.2, pp. 680–689.

Duvenaud, David, James Lloyd, Roger Grosse, Joshua Tenenbaum and Ghahramani Zoubin (2013). "Structure Discovery in Nonparametric Regression through Compositional Kernel Search". In: *Proceedings of the 30th International Conference on Machine Learning*. Ed. by Sanjoy Dasgupta and David McAllester. Vol. 28. Proceedings of Machine Learning Research 3. Atlanta, Georgia, USA: PMLR, pp. 1166–1174. URL: http://proceedings.mlr.press/v28/duvenaud13.html.

FAO (2014). *Food and Agriculture Organization(FAO) Country programming framework for Kenya (2014-2017)*. Tech. rep. URL: www.fao.org/publications.

Funk, Chris, Michael D Dettinger, Joel C Michaelsen, James P Verdin, Molly E Brown, Mathew Barlow and Andrew Hoell (Aug. 2008). "Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development". In: *Proceedings of the National Academy of Sciences* 105.32, 11081 LP –11086. DOI: 10.1073/pnas.0708196105. URL: http://www.pnas.org/content/105/32/11081.abstract.

Galvin, KA, RB Boone, NM Smith and SJ Lynn (2001). "Impacts of climate variability on East African pastoralists: linking social science and remote sensing". In: *Climate Research* 19, pp. 161–172. ISSN: 0936-577X. DOI: 10.3354/cr019161. URL: http://www.int-res.com/abstracts/cr/v19/n2/p161-172/.

Hamilton, J.D. (1994). *Time series analysis.* Princeton, NJ: Princeton University Press.

Hoffman, Matthew D., David M. Blei, Chong Wang and John Paisley (2013). "Stochastic Variational Inference". In: *J. Mach. Learn. Res.* 14.1, pp. 1303–1347. ISSN: 1532-4435.

Jiao, Wenzhe, Lifu Zhang, Qing Chang, Dongjie Fu, Yi Cen and Qingxi Tong (2016). "Evaluating an Enhanced Vegetation Condition Index (VCI) Based on VIUPD for Drought Monitoring in the Continental United States". In: *Remote Sensing* 8.3. ISSN: 2072-4292. DOI: 10.3390/rs8030224. URL: https://www.mdpi.com/2072-4292/8/3/224.

Kandasamy, S., F. Baret, A. Verger, P. Neveux and M. Weiss (2013). "A comparison of methods for smoothing and gap filling time series of remote sensing observations - application to MODIS LAI products". In: *Biogeosciences* 10.6, pp. 4055–4071. DOI: 10.5194/bg-10-4055-2013. URL: https://www.biogeosciences.net/10/4055/2013/.

Klisch, Anja and Clement Atzberger (2016). "Operational Drought Monitoring in Kenya Using MODIS NDVI Time Series". In: *Remote Sensing* 8.4. ISSN: 2072-4292. DOI: 10.3390/rs8040267. URL: http://www.mdpi.com/2072-4292/8/4/267.

Kogan, F.N. (1995). "Application of vegetation index and brightness temperature for drought detection". In: *Advances in Space Research* 15.11. Natural Hazards: Monitoring and Assessment Using Remote Sensing Technique", pp. 91–100. ISSN: 0273-1177. URL: http://www.sciencedirect.com/science/article/pii/027311779500079T.

Kogan, Felix, Tatiana Adamenko and Wei Guo (2013). "Global and regional drought dynamics in the climate warming era". In: *Remote Sensing Letters* 4.4, pp. 364–372. DOI: 10.1080/2150704X.2012.736033. eprint: https://doi.org/10.1080/2150704X.2012.736033. URL: https://doi.org/10.1080/2150704X.2012.736033.

Lemos, Maria Carmen, James C Arnott et al. (2018). "To co-produce or not to co-produce". In: *Nature Sustainability* 1.12, p. 722.

Lemos, Maria Carmen, Christine J Kirchhoff and Vijay Ramprasad (2012). "Narrowing the climate information usability gap". In: *Nature climate change* 2.11, p. 789.

Lesk, Corey, Pedram Rowhani and Navin Ramankutty (Jan. 2016). "Influence of extreme weather disasters on global crop production". In: *Nature* 529.7584, pp. 84–87. ISSN: 14764687. DOI: 10.1038/nature16467.

Lopez, Ana, Erin Coughlan de Perez, Juan Bazo, Pablo Suarez, Bart van den Hurk and Marteen van Aalst (2018). "Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings". In: *Weather and Climate Extremes.* ISSN: 22120947. DOI: 10.1016/j.wace.2018.03.006.

Matere, Joseph, Piers Simpkin, Jay Angerer, Emmanuella Olesambu, Selvaraju Ramasamy and Folorunso Fasina (2019). "Predictive Livestock Early Warning System (PLEWS): Monitoring forage condition and implications for animal production in Kenya". In: *Weather and Climate Extremes*, p. 100209.

Meroni, M, D Fasbender, F Kayitakire, G Pini, F Rembold, F Urbano and MM Verstraete (2014). "Early detection of biomass production deficit hot-spots in semi-arid environment using FAPAR time series and a probabilistic approach". In: *Remote Sensing of Environment* 142, pp. 57–68.

Mishra, Ashok K. and Vijay P. Singh (2010). "A review of drought concepts". In: *Journal of Hydrology* 391.1, pp. 202–216. ISSN: 0022-1694. DOI: https://doi.org/10.1016/j.jhydrol.2010.07.012. URL: http://www.sciencedirect.com/science/article/pii/S0022169410004257.

Nelson, Gerald C. et al. (2014). "Climate change effects on agriculture: Economic responses to biophysical shocks". In: *Proceedings of the National Academy of Sciences* 111.9, pp. 3274–3279. ISSN: 0027-8424. DOI: 10.1073/pnas.1222465110. eprint: https://www.pnas.org/content/111/9/3274.full.pdf. URL: https://www.pnas.org/content/111/9/3274.

Nyong, A., F. Adesina and B. Osman Elasha (June 2007). "The value of indigenous knowledge in climate change mitigation and adaptation strategies in the African Sahel". In: *Mitigation and Adaptation Strategies for Global Change* 12.5, pp. 787–797. ISSN: "1573-1596". DOI: "10.1007/s11027-007-9099-0". URL: https://doi.org/10.1007/s11027-007-9099-0.

Orindi, Victor A, Anthony Nyong and Mario Herrero (2007). *Pastoral Livelihood Adaptation to Drought and Institutional Interventions in Kenya.* Tech. rep.

Papagiannopoulou, Christina, Diego Gonzalez Miralles, Stijn Decubber, Matthias Demuzere, Niko Verhoest, Wouter A Dorigo and Willem Waegeman (2017). "A non-linear Granger-causality framework to investigate climate-vegetation dynamics". eng. In: *GEOSCIENTIFIC MODEL DEVELOPMENT* 10.5, pp. 1945–1960. ISSN: 1991-9603. URL: http://dx.doi.org/10.5194/gmd-10-1945-2017.

Piguet, Etienne, Antoine Pécoud and Paul de Guchteneire (June 2011). "Migration and Climate Change: An Overview". In: *Refugee Survey Quarterly* 30.3, pp. 1–23. ISSN: 1020-4067. DOI: 10.1093/rsq/hdr006. eprint: http://oup.prod.sis.lan/rsq/article-pdf/30/3/1/4460951/hdr006.pdf. URL: https://doi.org/10.1093/rsq/hdr006.

Quiring, Steven M and Srinivasan Ganesh (2010). "Evaluating the utility of the Vegetation Condition Index (VCI) for monitoring meteorological drought in Texas". In: *Agricultural and Forest Meteorology* 150.3, pp. 330–339.

Rasmussen, C. E. and C. K. I. Williams (2006). “Gaussian Processes for Machine Learning”. In.

Rojas, O., A. Vrieling and F. Rembold (2011). “Assessing drought probability for agricultural areas in Africa with coarse resolution remote sensing imagery”. In: *Remote Sensing of Environment* 115.2, pp. 343–352. ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2010.09.006. URL: http://www.sciencedirect.com/science/article/pii/S0034425710002798.

Roy, David et al. (Mar. 2014). “Landsat-8: Science and Product Vision for Terrestrial Global Change Research”. In: *Remote Sensing of Environment* 2014, pp. 154–172. DOI: 10.1016/j.rse.2014.02.001.

Rulinda, Coco M., Wietske Bijker and Alfred Stein (2011). “The chlorophyll variability in Meteosat derived NDVI in a context of drought monitoring”. In: *Procedia Environmental Sciences”* 3. 1st Conference on Spatial Statistics 2011 – Mapping Global Change, pp. 32–37. ISSN: 1878-0296. DOI: https://doi.org/10.1016/j.proenv.2011.02.007. URL: http://www.sciencedirect.com/science/article/pii/S1878029611000089.

Sai, F., L. Cumiskey, A. Weerts, B. Bhattacharya and R. Haque Khan (2018). “Towards impact-based flood forecasting and warning in Bangladesh: a case study at the local level in Sirajganj district”. In: *Natural Hazards and Earth System Sciences Discussions* 2018, pp. 1–20. DOI: 10.5194/nhess-2018-26. URL: https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2018-26/.

Savitzky, Abraham and Marcel JE Golay (1964). “Smoothing and differentiation of data by simplified least squares procedures.” In: *Analytical chemistry* 36.8, pp. 1627–1639.

Schaaf, Crystal and Z. Wang (2015). *MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006 [Data set]*. DOI: 10.5067/MODIS/MCD43A4.006. URL: https://lpdaac.usgs.gov/dataset%7B%5C_%7Ddiscovery/modis/modis%7B%5C_%7Dproducts%7B%5C_%7Dtable/mcd43a4%7B%5C_%7Dv006.

Stanke, Carla, Marko Kerac, Christel Prudhomme, Jolyon Medlock and Virginia Murray (2013). “Health effects of drought: a systematic review of the evidence”. In: *PLoS currents* 5. URL: https://www.ncbi.nlm.nih.gov/pubmed/23787891.

Sutanto, Samuel J, Melati van der Weert, Niko Wanders, Veit Blauhut and Henny A J Van Lanen (2019). “Moving from drought hazard to impact forecasts”. In: *Nature Communications* 10.1, p. 4945. ISSN: 2041-1723. DOI: 10.1038/s41467-019-12840-z. URL: https://doi.org/10.1038/s41467-019-12840-z.

Tozier de la Poterie, Arielle and Marie-Ange Baudoin (June 2015). “From Yokohama to Sendai: Approaches to Participation in International Disaster Risk Reduction Frameworks”. In: *In-*

ternational Journal of Disaster Risk Science* 6.2, pp. 128–139. ISSN: 2192-6395. DOI: 10.1007/s13753-015-0053-6. URL: https://doi.org/10.1007/s13753-015-0053-6.

Udelhoven, T., M. Stellmes, G. del Barrio and J. Hill (2009). "Assessment of rainfall and NDVI anomalies in Spain (1989 – 1999) using distributed lag models". In: *International Journal of Remote Sensing* 30.8, pp. 1961–1976. DOI: 10.1080/01431160802546829. eprint: https://doi.org/10.1080/01431160802546829. URL: https://doi.org/10.1080/01431160802546829.

UNDP (2013). *Kenya – adaptation to climate change in arid lands (KACCAL)-Kenya case study*. Tech. rep. URL: http://gefonline.org/projectDetailsSQL.cfm?projID=3792..

Upreti, Deepak, Wenjiang Huang, Weiping Kong, Simone Pascucci, Stefano Pignatti, Xianfeng Zhou, Huichun Ye and Raffaele Casa (2019). "A Comparison of Hybrid Machine Learning Algorithms for the Retrieval of Wheat Biophysical Variables from Sentinel-2". In: *Remote Sensing* 11.5. ISSN: 2072-4292. DOI: 10.3390/rs11050481. URL: http://www.mdpi.com/2072-4292/11/5/481.

Venton, C, C Fitzgibbon, T Shitarek, L Coulter and O Dooley (2012). "The economics of early response and disaster resilience: lessons from Kenya and Ethiopia". In: *Independent report*.

Vrieling, Anton, Michele Meroni, Andrew G Mude, Sommarat Chantarat, Caroline C Ummenhofer and Kees CAJM de Bie (2016). "Early assessment of seasonal forage availability for mitigating the impact of drought on East African pastoralists". In: *Remote sensing of environment* 174, pp. 44–55.

Weiss, Daniel J., Peter M. Atkinson, Samir Bhatt, Bonnie Mappin, Simon I. Hay and Peter W. Gething (Dec. 2014). "An effective approach for gap-filling continental scale remotely sensed time-series". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 98, pp. 106–118. ISSN: 0924-2716. DOI: 10.1016/J.ISPRSJPRS.2014.10.001. URL: https://www.sciencedirect.com/science/article/pii/S0924271614002512.

Wilkinson, Emily, Lena Weingärtner, Richard Choularton, Meghan Bailey, Martin Todd, Dominic Kniveton and Courtenay Cabot Venton (2018). *Forecasting hazards, averting disasters: Implementing forecast-based early action at scale*. Tech. rep. Overseas Development Institute (ODI).

WMO (2015). "WMO Guidelines on Multi-Hazard Impact-Based Forecast and Warning Services". In.

Zambrano, Francisco, Anton Vrieling, Andy Nelson, Michele Meroni and Tsegaye Tadesse (2018). "Prediction of drought-induced reduction of agricultural productivity in Chile from MODIS, rainfall estimates, and climate oscillation indices". In: *Remote Sensing of Environment* 219, pp. 15–30. ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2018.10.006. URL: http://www.sciencedirect.com/science/article/pii/S0034425718304541.

Zargar, Amin, Rehan Sadiq, Bahman Naser and Faisal I Khan (2011). "A review of drought indices". In: *Environmental Reviews* 19.NA, pp. 333–349.

Zhang, Lifu, Wenzhe Jiao, Hongming Zhang, Changping Huang and Qingxi Tong (2017). "Studying drought phenomena in the Continental United States in 2011 and 2012 using various drought indices". In: *Remote Sensing of Environment* 190, pp. 96–106. ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2016.12.010. URL: http://www.sciencedirect.com/science/article/pii/S0034425716304813.

Chapter 4

Forecasting Vegetation Condition with a Bayesian Auto-regressive Distributed Lags(B-ARDL) Model

Edward E. Salakpi[a]*, Pete Hurley[a,b], James M. Muthoka[c], Adam B. Barrett[a,b], Andrew Bowell[a,b], Seb Oliver[a,b] and Pedram Rowhani[c]

[a] The Data Intensive Science Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

[b] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

[c] School of Global Studies, Department of Geography, University of Sussex, Brighton, BN1 9QJ, UK

*Corresponding author: e.salakpi@sussex.ac.uk

Abstract

Droughts form a large part of climate/weather-related disasters reported globally. In Africa, pastoralists living in the Arid and Semi-Arid Lands (ASALs) are the worse affected. Prolonged dry spells that cause vegetation stress in these regions have resulted in the loss of income and livelihoods. To curb this, global initiatives like the Paris Agreement and the United Nations recognised the need to establish Early Warning Systems (EWS) to save lives and livelihoods. Existing EWS use a combination of Satellite Earth Observation (EO) based biophysical indicators like the Vegetation Condition Index (VCI) and socio-economic factors to measure and monitor droughts. Most of these EWS rely on expert knowledge in estimating upcoming drought conditions without using forecast models. Recent research has shown that the use of robust algorithms like Auto-Regression, Gaussian Processes and Artificial Neural Networks can provide very skilled models for forecasting vegetation condition at short to medium range lead times. However, to enable preparedness for early action, forecasts with a longer lead time are needed. The objective of this research work is to develop models that forecast vegetation conditions at longer lead times on the premise that vegetation condition is controlled by factors like precipitation and soil moisture. To achieve this, we used a Bayesian Auto-Regressive Distributed Lag (BARDL) modelling approach which enabled us to factor in lagged information from Precipitation and Soil moisture levels into our VCI forecast model. The results showed a $\sim$ 2-week gain in the forecast

range compared to the univariate Auto-Regression model used as a baseline. The R^2 scores for the Bayesian ARDL model were 0.94, 0.85 and 0.74, compared to the Auto-Regression model's R^2 of 0.88, 0.77 and 0.65 for 6, 8 and 10 weeks lead time respectively.

Keywords: *Drought; Bayesian Models; Forecasting; Early Warning Systems; Disaster Risk Reduction*

4.1 Introduction

Drought events are amongst the most prevalent natural disasters reported globally and affect some 55 million people annually (Deleersnyder, 2018). In Africa, the devastating effects of droughts are mostly seen in the Arid and Semi-Arid Lands (ASALs), where people's lives and livelihoods mostly depend on agro-pastoral activities (Gebremeskel et al., 2019). Pastoralism in these regions contributes immensely to food security and local economies (Vatter, 2019). However, the ASALs grass- and shrublands, which serve as the main source of fodder for the livestock are among the first to be hit by low rains and extreme temperature (FAO, 2018). These dry spells, when prolonged, adversely impact the food markets, income, and eventually leads to the loss of livelihoods (FAO, 2018). As a consequence, several drought early warning systems (EWS) have been developed to avert and minimise the impacts of these hazards.

Global initiatives, such as the 2015 Paris Agreement and the United Nation's Sustainable Development Goals (SDGs) recognise the importance of establishing robust EWS to save lives and livelihoods (UNFCCC, 2015). Existing EWS combine data on biophysical indicators that measure hazard risk with a series of socio-economic factors to account for vulnerability and exposure for early action. Satellite Earth Observation (EO) rainfall estimates and vegetation health are some of the datasets commonly used to monitor these drought conditions. The USAID's * Famine Early Warning Systems Network (FEWS NET) utilises household livelihood information, rainfall estimates and the Normalized Difference Vegetation Index (NDVI) to monitor drought and its impact on food security (FEWSNET, 2019). In Kenya, the National Drought Management Authority (NDMA) monitors EO based biophysical indicators in combination with forage, livestock conditions and socio-economic data to monitor and anticipate future drought scenarios for early finance and early action (Klisch et al., 2016; FAO, 2017).

Recent research has highlighted robust methods for forecasting biophysical indicators used to measure vegetation condition. AghaKouchak, 2014 harnessed the persistence property in soil moisture with the ensemble streamflow prediction (ESP) to provide skilful forecasts of the standardized soil moisture index for up to two months ahead. Barrett et al., 2020 forecasted

*United States Agency for International Developmen (USAID) the Vegetation Condition Index (VCI) with Auto-Regression (AR) and Gaussian Process (GP) models using historical values of the same indicator. Both models performed well for lead times up to 6 weeks. Adede et al., 2019 used a multivariate approach that considered the effects of exogenous variables on VCI. The model was based on an Artificial Neural Network (ANN) and provided precise forecasts for one month lead time. While these models showed good accuracies for short-range forecasts, forecasts with longer lead times beyond six weeks will provide disaster risk managers ample time to prepare and implement relief measures.

This paper aims to build on existing forecast initiatives and develop models that accurately forecast VCI at longer lead times. More specifically, our approach will include the interaction between the lagged information from indicators and variables like precipitation, soil moisture, and vegetation condition in an Auto-regressive distributed lag (ARDL) model (Gujarati, 2003; Pesaran et al., 1999). ARDL models are useful in situations where variable Y_t at a time t is influenced by other variables X_t at time t and the same variables at previous time steps X_{t-i}.

Parameter estimation with ARDL models has traditionally been carried out with a maximum likelihood approach which produces point estimates and often results in over-fitting leading to imprecise predictions (Martin, 2018). To address this, the ARDL model used in this work was implemented within a Bayesian framework which allows the incorporation of prior knowledge of the model parameters. This approach generates a posterior probability distribution for the model parameters which enables more accurate quantification of prediction uncertainties and allows for more robust risk analysis (Lambert, 2018).

4.2 Study Area and Data

4.2.1 Study Area

This research was conducted in selected counties in the ASAL regions (see figure 5.2) of Kenya where the predominant activities are pastoralism and wildlife conservation. The farmers in these regions rely heavily on pastures and grasslands as the main source of feed for their animals (Sibanda et al., 2017). However, the erratic weather patterns in the eastern African region makes Kenya prone to frequent drought events which poses a threat to the country's food security and economy as a whole (Gebremeskel et al., 2019). During the 2008-2011 droughts the Kenyan economy lost a total of 21.1 billion USD (Cabot Venton et al., 2012; Cenacchi, 2014). Hence the need to develop drought EWS with the ability to provide timely medium to long-term for drought preparedness.

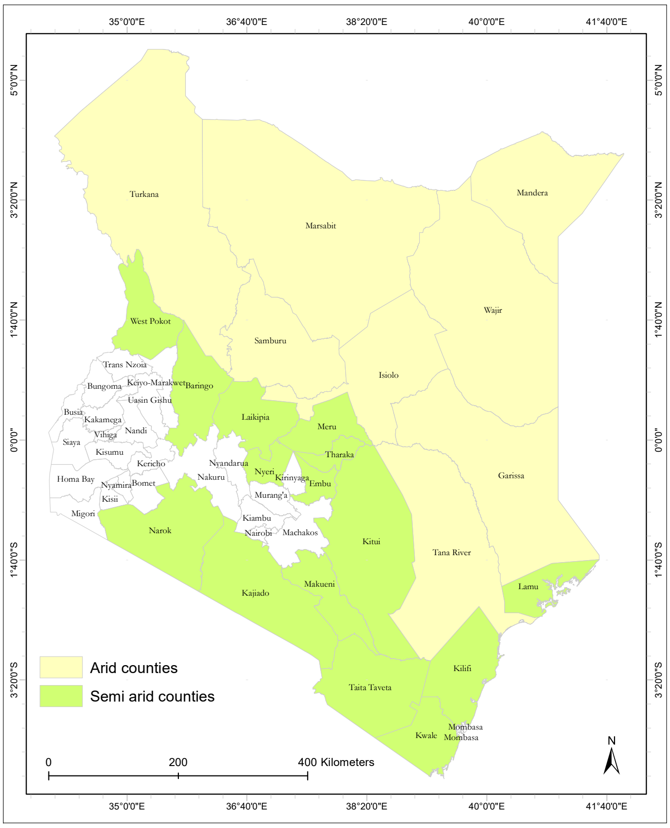

Figure 4.1: A map of Kenya showing the arid and semi-arid counties where the research was focused.

4.2.2 Data

Developing a highly skilled model required adequate historical data on drought indicators and biophysical factors acquired over a long period. Table 4.1 shows details of the satellite earth observation data used for this work.

Table 4.1: Summary of the datasets for the forecast model

Data	**Source (Producer)**	**Spatial Resolution**	**Temporal Resolution**	**Acquisition Period**	**Unit of Measure**
Precipitation	Climate Hazards Group InfraRed Precipitation (CHIRPS)	5km	Daily	2001-2018	mm
Soil Moisture	European Space Agency's Climate Change Initiative (CCI)	30km	Daily	2001-2018	m^3/m^{-3}
Surface Reflectances	NASA MODIS (MCD43A4 v006)	500m	Daily	2001-2018	N/A

4.2.2.1 Precipitation (Rainfall Estimates)

The precipitation data were acquired from the Climate Hazards Group InfraRed Precipitation (CHIRPS) project (Funk et al., 2015). The CHIRPS data comprise a combination of weather station data and rainfall estimates captured via satellite remote sensing using the Cold Cloud Duration (CCD) (Milford et al., 1990) approach. The approach is used to estimates rainfall by using remotely sensed information on the period of time a cloud remains at a given temperature. The dataset is available as daily 5km resolution images.

4.2.2.2 Soil Moisture

The daily 30km resolution soil moisture products by the European Space Agency's Climate Change Initiative (ESA-CCI) was used for this work. The data is produced from an algorithm that takes in back-scatter information from multiple active and passive Synthetic Aperture Radar (SAR) satellites. The values generated represent soil moisture at a soil depth of 10cm. The ESA-CCI Soil moisture products are available as passive, active or a combination of both. For this work, the combined version of the data is used (Gruber et al., 2019; Dorigo et al., 2017; Yang et al., 2017).

4.2.2.3 Surface Reflectance

The bidirectional reflectance distribution function (BRDF) corrected MODIS product, MCD43A4 Version 6, (Schaaf et al., 2015) was used to compute the NDVI and VCI. The product is available as daily 500m resolution images captured in 7 bands ranging from visible to infrared. Information on the vegetation health is derived from the Red and Near-Infra Red(NIR) bands via equation (4.1).

$$\text{NDVI} = \frac{NIR - Red}{NIR + Red} \tag{4.1}$$

4.3 Methods

4.3.1 Data pre-processing

The datasets were acquired from January 1, 2001, to December 31, 2018, to correspond with the availability of soil moisture data at the time of research. Apart from the precipitation, clouded and low-quality pixels from poor atmospheric and radiometric correction were removed using the quality flags from the Quality Assurance (QA) maps that came with the surface reflectance and soil moisture products. Pixels representing grasslands and shrublands areas within our regions of interest were retrieved with the European Space Agency (ESA)'s 2016 Sentinel 2 Land Use and Land Cover (LULC) map * **Ramoino2018**. For the coastal semi-arid counties like Lamu and Kwale we could not extract enough soil moisture data so no results were shown for these counties.

To measure the drought condition at a period in time, the minimum and maximum NDVI values for a chosen baseline time interval and the NDVI value for that period are used to compute the Vegetation Condition Index (VCI) via equation (4.2) (Kogan, 1995). VCI values range from 0-100, with values below 35 depicting a moderate to severe drought condition (Klisch et al., 2016).

$$\text{VCI}_i = 100 \times \frac{\text{NDVI}_i - \text{NDVI}_{min,i}}{NDVI_{\max,i} - NDVI_{\min,i}}, \tag{4.2}$$

where $\mathrm{N}DVI_{\min,i}$ and $\mathrm{N}DVI_{\max,i}$ are the long-term minimum and long-term maximum NDVI values of a pixel at i^{th} week of the year.

Temporal gaps created by the removal of poor quality pixels were filled with the Radial Basis Function (RBF) interpolation method (Rippa, 1999). This approach was used to avoid interpolated values for periods with longer gaps from going over the valid ranges. Noise resulting from faulty instruments were reduced with the Whittaker smoother (Eilers, 2003), which filters noise via a penalised least-squares. Since our target variable was computed from the long-term minimum and maximum NDVI, the additional variables were also converted to anomalies by

*Visit this link (http://2016africalandcover20m.esrin.esa.int/) to learn more subtracting their long-term means to produce soil moisture anomaly and precipitation anomaly. The persistence within individual variables was enhanced by computing with three months (12 weeks) rolling averages to derive three-month VCI (VCI3M), three-month precipitation (P3M) and three-month soil moisture (SM3M). Finally, the precipitation and soil moisture data were standardised to eliminate any associated units of measurements and avoid the dominance of certain variables. This was done by subtracting their mean and dividing it by the standard deviation.

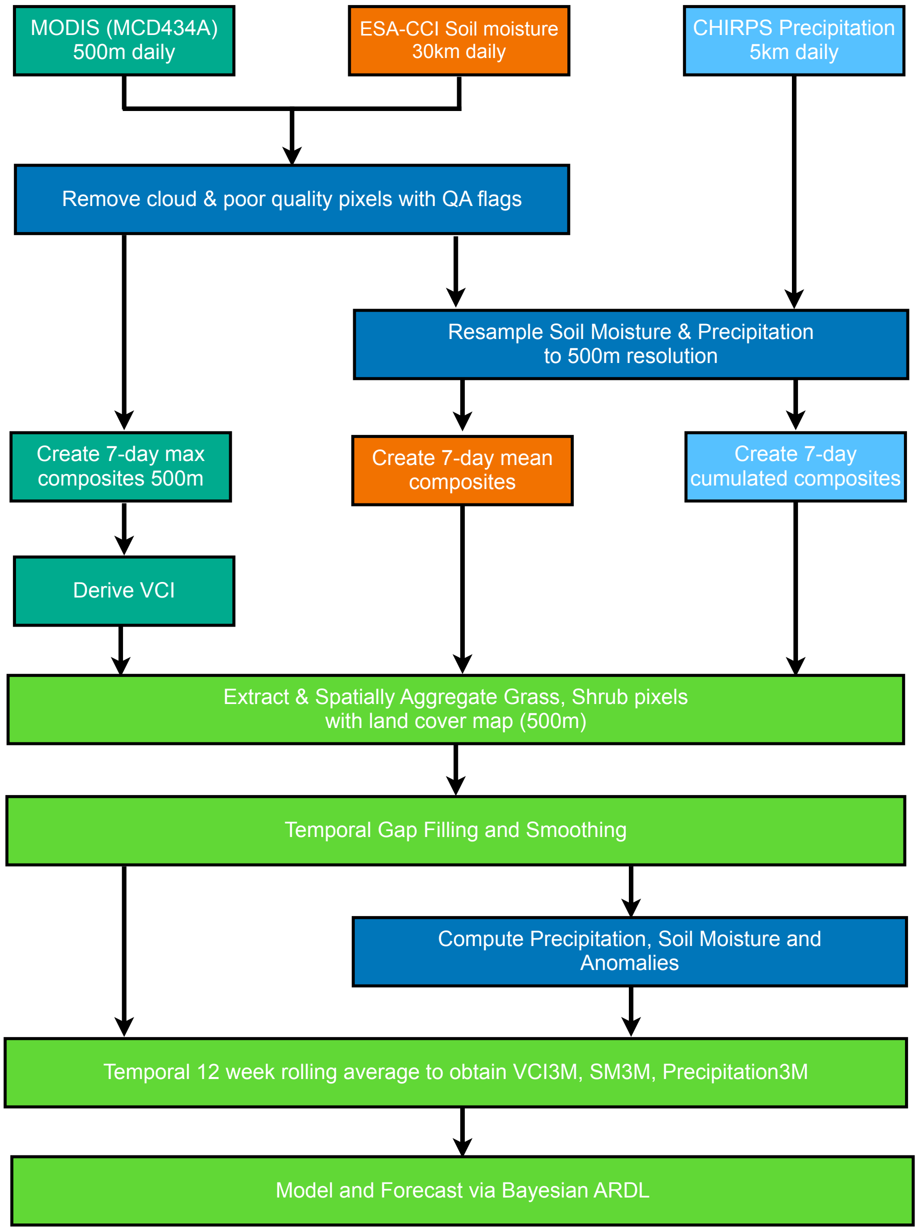

Figure 4.2: A flow chart showing data prepossessing and modelling

4.3.2 Drought Model and Forecasting

The Auto-Regressive Distributed Lag (ARDL) modelling approach used for this work is a generalised form of Auto Regression (AR) method mainly used for multivariate time series analysis. The method enables the variable of interest (dependent variable) to be modelled as a function of its lags and that of additional explanatory variables (independent variable) (Gujarati, 2003). An ARDL(p,q), consists of p, which represents the number of lags of the independent variable and q, which is the auto-regressive part of the model, represents the number of lags of the dependent variable. This approach has been extensively used in the field of economics and modelling the effect of climate and environmental variables on vegetation (Lei Ji et al., 2004; Ji et al., 2005).

For this study, however, parameter estimation for the ARDL was implemented within a Bayesian framework instead of using maximum likelihood methods based on Ordinary Least Squares (OLS). The Bayesian framework enables the incorporation of domain knowledge about the parameters through the use of informative priors. The model parameters, with this approach, are inferred using the Markov Chain Monte Carlo (MCMC) (Neal, 1993) sampling algorithm. The sampling process generates posterior probability distribution of the model parameters. As a consequence, we get a full probability distribution of forecast values for all lead time, which makes it easy to quantify forecast uncertainty for making informed decisions (Martin, 2018; Lambert, 2018).

The MCMC is a well-established sampling algorithm used for parameter inference in Bayesian models. However, Asaad et al., 2019, outlined some of its limitations and recommended the use of the Hamiltonian Monte Carlo (HMC) (Hoffman et al., 2014), an improved variant of the traditional MCMC algorithm which is based on Hamiltonian dynamics and converges faster to a global minimum for models with high dimensional parameter space. (Robert et al., 2018). Parameter inference for this work was done with the No-U-Turn Sampler (NUTS)(Hoffman et al., 2014) version of HMC implemented with PyMC3 (Salvatier et al., 2016) Python package.

The Bayesian ARDL model used for forecasting VCI3M with lagged P3M, and S3M is defined as:

$$D_{t+n} = \alpha_0 + \sum_{i=0}^{q} \beta_d D_{t-q} + \sum_{i=0}^{p} \theta_p P_{t-p} + \sum_{i=0}^{p} \delta_s S_{t-p} + \epsilon_{t-p} \tag{4.3}$$

where D_{t+n} is the drought index at n lead time, D_{t-q} are the lags (0, to q) of drought indicator (Dependent variable). P_{t-p}, S_{t-p} represent the lags 0, to p, for precipitation, and soil moisture respectively. α_0 is a constant representing the intercept and β_d, θ_p, and δ_s are the regression coefficients of the input variables with ϵ_{t-p} being the error term which is assumed to be Gaussian.

Equation (4.3) can re-written as:

$$D_{t+n} = \alpha + \sum_{i=0}^{i} \beta_i X_{t-i} + \epsilon_{t-i} \tag{4.4}$$

where n is the lead time, β_i are the model parameters and X_{t-i} represent the lagged input variables in equation 4.3.

The Bayesian approach makes explicit the prior beliefs about model parameters, which are then updated given some new data via the likelihood function, to give the posterior probability distribution.

Parameter inference with the Bayesian framework is based on Bayes' theorem via the equation below:

$$P(\theta|X_t) = \frac{P(X_t|\theta).P(\theta)}{P(X_t)} \tag{4.5}$$

where X_t represents $D_{t-q}, P_{t-p}, S_{t-p}$, $P(\theta|X_t)$ is the posterior or the probability of our model parameters given our data X_t, $P(X_t|\theta)$ is the likelihood or the probability of the data given the parameters, $P(\theta)$ is our prior belief about the parameters. $P(X_t)$, known as the evidence, is a normalisation term that represents the probability of the data. The term is difficult compute and usually ignored (Lambert, 2018; McElreath, 2016). Thus the equation (4.5) for Bayes' theorem is re-written as:

$$P(\theta|X_t) \propto P(X_t|\theta).P(\theta) \tag{4.6}$$

To put the ARDL model (equation 4.4) in the context of equation 4.6, the likelihood function $P(X_t|\theta)$ is written as:

$$P(X_t|\alpha, \beta_i, \sigma) \sim N(\alpha + \sum_{i=0}^{i} \beta_i X_{t-i}, \sigma_{t-i}) \tag{4.7}$$

Equation 4.6 is practically intractable due the complex integrals required (Lambert, 2018) thus, the need to use the HMC algorithm (Hoffman et al., 2014) for sampling model parameters.

The prior for the model's regression coefficients are assumed to be Gaussian $P(\theta) = N(\mu, \sigma)$ with μ set to 0 to allow inferred parameters to have both positive and negative values and a weakly informative σ of 0.5 as a regularization prior. This was done to avoid the approximation of unreasonable parameters (Martin, 2018).

4.3.3 Selecting optimal lags and forecasting

A full grid search was done with various combinations of p and q values for dependent and independent values to select the optimal p and q for the BARDL model. The Akaike Information Criterion (AIC) (Akaike, 1998) equation (4.8) and the R^2 (Equation 5.8) metric were used as the score criteria to choose the optimal lags. AIC enables model selection by determining the model that best fits the data. The model with lowest AIC value is preferred. Whereas the R^2 score explains how much variation in the observed data could be explained by the model. Valid R^2 scores range between 0&1 where models with scores close to 1 are considered more accurate. The search was done on lag values ranging from 1 to 16 weeks. The optimal lag values varied for different counties. However, across all counties, low AIC and high R^2 scores were obtained when all input variables were set to a lag of 6 weeks. The AIC scores are derived as follows:

$$\text{AIC} = 2K + n\log(\frac{RSS}{n}) \tag{4.8}$$

where the RSS is the residual sum of squares error, n is the number of data points and K is the number of estimated parameters.

The R^2 scores are derived as follows:

$$R^2\text{-score} = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}, \tag{4.9}$$

where the y_i are the observed data, $\bar{y}$ mean of the observed data and the f_i are the forecasts.

Forecasting with the BARDL was done using the direct multi-step forecast approach, where separate models are fitted for n step ahead forecasts (Ben Taieb, Sorjamaa et al., 2010; Ben Taieb and Rob J. Hyndman, 2014). To fit the model for n steps ahead, the data was restructured to offset values of the dependent (D_{t+n}), n weeks from lag 0 X_{t-0} for all input variables. A rolling window cross-validation approach (Rob J Hyndman et al., 2018) was used for model training and forecasting. With this approach, the data is divided into chunks of 500 data points, for each chunk, 400 data points are used to train the model and remaining 100 data points held-out for prediction. The observed values from held-out were then used to evaluate the model skill.

4.3.4 Forecast skill assessment

The performance of the models was assessed by measuring the *accuracy*, i.e. how well the forecasts agree with the observations and the *precision*, i.e. the quoted uncertainty and the accuracy of that uncertainty.

The model *accuracy* was evaluated with the R^2 (equation 5.8) and Root Mean Squared Error (RMSE) (equation 4.10). The RMSE measures the mean deviation between the observed and forecast values.

$$\text{R}MSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - f_i)^2}{n}} \tag{4.10}$$

where the y_i are the observed data, f_i are the forecasts and n the total number of data points.

The *precision*, was quantified with the Prediction Interval Coverage Probability (PICP) and the Mean Prediction Interval Width (MPIW) (Pang et al., 2018). The MPIW measures the average width between the upper $(u(D_i))$ and lower bound $l(D_i)$ of a proportion of forecast distribution (n weeks ahead) defined by a chosen prediction interval (e.g. 95%).

$$\mathrm{M}PIW_{t+n} = \frac{1}{N}\sum_{i=1}^{m} |u(D_i) - l(D_i)|. \tag{4.11}$$

The PICP shows the percentage of time the observed variable lies within the credible interval of the forecast distribution and is derived as follows:

$$PICP_{t+n} = \frac{1}{N}\sum_{i=1}^{m} c_i \tag{4.12}$$

where N is the number of predicted samples and c_i is either 0, or 1. If the observed drought target variable falls within the upper and lower bound of the forecast distribution (n weeks ahead) then $c_i = 1$; else $c_i = 0$ if otherwise.

The goal is to minimize the MPIW while maintaining a high PICP value. A high PICP value (0.90 to 0.99) indicates that the observed values lie within the forecast distribution and a low MPIW value indicates a more precise forecast (Su et al., 2018). For the AR model, the confidence interval used to derive its PICP and MPIW was computed with the forecast RMSE and z-score of 1.96 representing the 95% confidence level of a standard normal distribution. This was done to permit its comparison to the output of the full BARDL model.

The contribution of the individual lagged inputs in the ARDL model were also measured by computing their percentage relative importance via the Relative Weight Analysis method (Tonidandel et al., 2011). With this approach, the inputs variables are initially transformed into orthogonal variables. Through an iterative process, each orthogonal variable is added to a linear regression model and the change in R^2 score for each iteration is measured and expressed as a percentage of the total R^2 score.

The Receiver operating characteristic (ROC) curve (Fawcett, 2006) (see chapter: chapter 3, figure: 3.7) was also plotted to see how well the model forecasts drought events at a given threshold.

The forecast distribution from our BARDL model enabled the computation of forecast probabilities.The forecast probability of a drought event was computed from the full forecast distribution from our posterior at a drought threshold of VCI<35. The model's skill at accurately forecasting these probabilities was assessed by plotting and analysing a Reliability Diagram and Sharpness. The reliability diagrams were plotted by using same threshold of VCI3M<35 to initially convert the held-out observed VCI3M data at a given lead time into binary events where 0 indicates a 'No Drought' and 1 indicate a 'Drought' event. The forecast probabilities and observed binaries were binned into standard intervals and plotted as a joint distribution of forecast probabilities and the relative frequency of the true observed drought event where observed binaries were equal to 1. The sharpness plots, on the other hand, are frequencies of drought occurrences in each probability bin.

The Reliability Diagram shows how well forecast probabilities for a given drought event agreed with its corresponding observed event while the Sharpness shows the frequency of a forecasted drought event (WWRP, 2009; Wilks, 2006).

4.4 Results

4.4.1 Forecast accuracy

AR modelling approach had proved to be skilful for short-range (2 to 6 weeks lead time) VCI3M forecasts (Barrett et al., 2020). However, the goal of this study was to extend the forecast range beyond 6 weeks while maintaining high accuracy by using the BARDL model and considering the effect of exogenous factors like precipitation and soil moisture. The results shown in this section are for 6 to 12 weeks lead time for the BARDL models and with the AR modelling as a comparative baseline.

The contour plots in Figure 4.3, shows a joint distribution of the observed VCI3M and forecasted VCI3M at 6, 8, 10 & 12 weeks for both AR and the BARDL models. The contour lines represent the bin of tho joint histogram and for each plot, the correlation (r), RMSE and R^2 were computed. Overall, the results from the BARDL model showed a roughly 2-week gain in the performance metrics. For instance, R^2 score for the AR model at 6 weeks is equivalent to R^2 score at 8 weeks lead time for the BARDL models. This pattern can be seen across all forecast ranges for the RMSE as well.

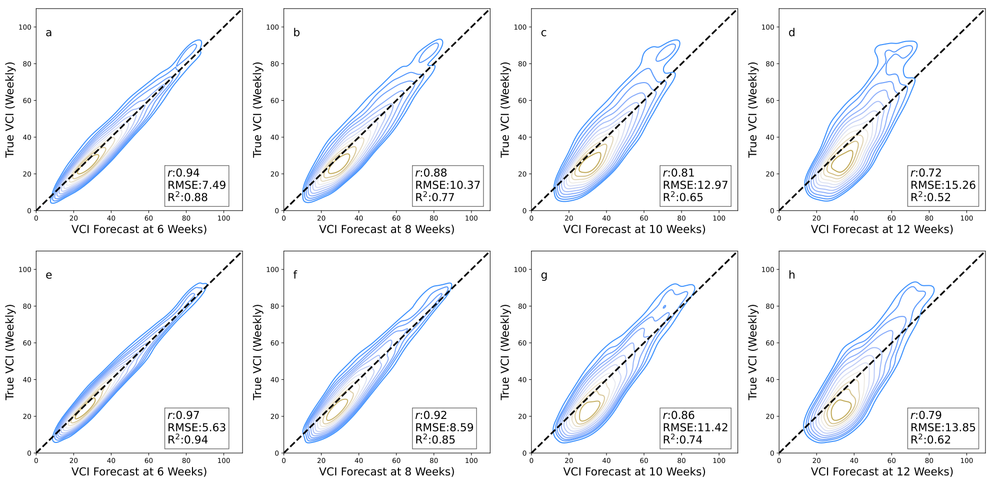

Figure 4.3: Contour plots showing VCI3M forecast against True VCI3M. Plots (a,b,c,d) shows the results from the AR method with VCI3M only, (e,f,g,h) shows the overall results for BARDL modelled with lags of VCI3M plus lags of Precipitation (P3M) and Soil Moisture (S3M) Anomalies for 6, 8, 10 and 12 weeks lead time for all counties

The performance metrics for the BARDL model in comparison to the AR model are shown in figure 4.4. This shows a significant improvement in performance at the same lead-time and, as a consequence, similar performance in the BARDL models is seen 2 weeks ahead of the AR models.

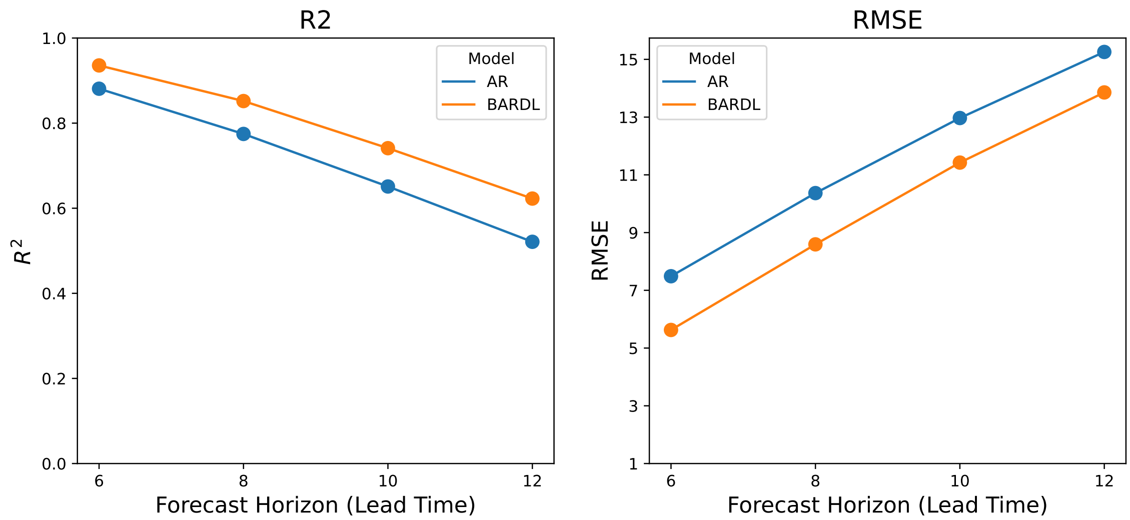

Figure 4.4: Performance metrics used to measure model accuracy as a function of forecast lead time. R^2(Left), RMSE (Right).

Table 4.2 shows the R^2 scores for 6 to 12 weeks forecasts for AR and BARDL models at the county level for arid and semi-arid regions. Just as observed in the contour plots, significant improvements are seen from 6 to 10 weeks lead time across all counties. In an arid county like Mandera, the R^2 improved from 0.84, 0.72 and 0.58 using AR to 0.93, 0.84 and 0.73 using BARDL for 6, 8 and 10 weeks lead times respectively. Kitui in the semi-arid region also showed an improvement in R^2 score from 0.84, 0.71 and 0.57 to 0.91, 0.81 and 0.67 for weeks 6, 8 and 10 respectively. Overall the BARDL method demonstrated better results compared to the AR across all counties.

4.4.2 Uncertainty Analysis (PICP and MPIW)

The PICP and MPIW for a 95% forecast confidence interval were computed for each lead time for both the AR and BARDL models. In figure 4.5, the time series plots show that the observed VCI3M values lie within the 95% forecast interval between 94–96% of the time across all lead times for both the BARDL and AR models, indicating that the errors bounds are very good for both. However, lower values of MPIW demonstrate that BARDL provided a more precise forecast. Appendix A, tabulates PICP and MPIW for 6 to 12 weeks forecasts for the AR and BARDL models for all counties (Table 4.3).

Table 4.2: R^2 scores (6 to 12 weeks lead times) for AR modelled with lags of VCI3M only, BARDL modelled with lags of VCI3M with Precipitation (P3M) and Soil Moisture (SM3M) for arid and semi-arid counties.

	County	**AR**				**BARDL**			
		6	**8**	**10**	**12**	**6**	**8**	**10**	**12**
Arid Counties	**Garissa**	0.88	0.78	0.66	0.54	0.92	0.83	0.71	0.60
	Isiolo	0.89	0.79	0.67	0.55	0.95	0.88	0.77	0.66
	Mandera	0.86	0.74	0.60	0.46	0.93	0.84	0.73	0.63
	Marsabit	0.91	0.81	0.69	0.54	0.96	0.90	0.80	0.68
	Samburu	0.88	0.75	0.59	0.43	0.95	0.87	0.75	0.62
	Tana-River	0.85	0.75	0.64	0.53	0.92	0.83	0.72	0.62
	Turkana	0.90	0.79	0.65	0.50	0.96	0.89	0.79	0.65
	Wajir	0.82	0.69	0.55	0.42	0.91	0.82	0.71	0.61
	Mean	**0.87**	**0.76**	**0.63**	**0.50**	**0.94**	**0.86**	**0.75**	**0.63**
	Std. Dev.	*0.03*	*0.04*	*0.04*	*0.05*	*0.02*	*0.03*	*0.03*	*0.03*

		6	**8**	**10**	**12**	**6**	**8**	**10**	**12**
Semi-Arid Counties	**Baringo**	0.92	0.83	0.70	0.56	0.95	0.86	0.74	0.60
	Kajiado	0.90	0.80	0.69	0.57	0.96	0.90	0.81	0.71
	Kilifi	0.84	0.72	0.60	0.48	0.88	0.76	0.62	0.49
	Kitui	0.84	0.70	0.56	0.43	0.92	0.81	0.68	0.53
	Laikipia	0.93	0.85	0.73	0.59	0.97	0.91	0.81	0.67
	Makueni	0.84	0.72	0.59	0.46	0.93	0.83	0.71	0.59
	Meru	0.83	0.67	0.49	0.33	0.92	0.81	0.67	0.52
	Narok	0.85	0.74	0.60	0.45	0.92	0.81	0.67	0.50
	Nyeri	0.90	0.81	0.68	0.54	0.93	0.85	0.73	0.60
	Taita-Taveta	0.86	0.74	0.60	0.47	0.92	0.81	0.69	0.59
	Tharaka-Nithi	0.81	0.64	0.45	0.28	0.83	0.63	0.39	0.17
	West-Pokot	0.91	0.82	0.69	0.54	0.95	0.86	0.72	0.57
	Mean	**0.87**	**0.75**	**0.62**	**0.48**	**0.92**	**0.82**	**0.69**	**0.54**
	Std. Dev.	*0.04*	*0.06*	*0.08*	*0.09*	*0.04*	*0.07*	*0.10*	*0.13*

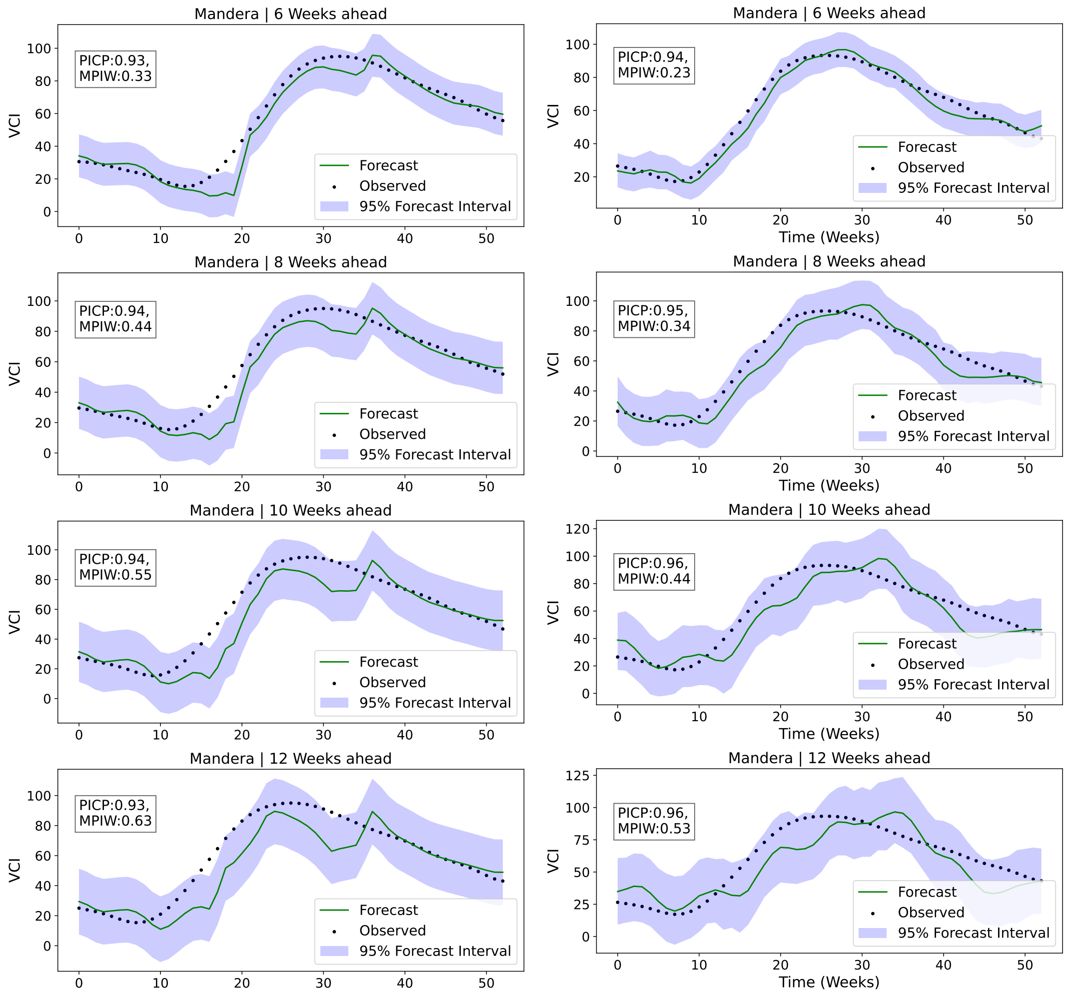

Figure 4.5: Time series plot showing uncertainty for 6, 8, 10, 12 weeks lead time for Mandera county. Plots on the left side are from the AR model and plots to the right are BARDL. The PICP and MPIW for the other counties can found in Appendix A

4.4.3 Drought Events ROC Curve

The Receiver Operating Characteristic (ROC) curve (figure 4.6) illustrates how well the model can discriminate drought events. Drought events are forecasted when the predicted VCI3M drops below a threshold and are deemed correct if the observed VCI3M is below 35 (moderate to severe drought) (Klisch et al., 2016). The ROC show the probability of a forecasted event being true (True Positive Rate (TPR)) against the chance of that predicted event being false alarm (False Positive Rate (FPR) as the threshold is varied. The Area Under the Curve (AUC), quantifies the ability of the forecast model to distinguish between drought events (Wilks, 2006; Bradley, 1997). The ROC curve and AUC metric for the BARDL model also demonstrated an improvement over the AR model. The points plotted on the curve represent the TPR and FPR where VCI3M<35. This indicates that when the AR model (Dotted curve), forecasts a drought condition (i.e VCI3M<35) for 6 weeks ahead, the probability of it being true is 86% with a FPR of 9%. Whereas a forecast by the BARDL model (Solid curve) at the same 6 weeks had a TRP of 89% and a FPR of 7%. The improvements with the BARDL model were mainly seen in the TPRs (6 to 10 week lead time) for the BARDL model while the FPR remained almost the same. The improvements seen in the ROC curves in figure 4.6 are however not reflective of the distinct improvement seen in figure 4.4. The observed difference was because whereas the R^2 and RMSE are comparing the explained variations and deviation between the observed and forecast VCI3M, the ROC is mainly assessing the skill of both models at predicting drought occurrence at the VCI3M<35 threshold. The difference indicates that although there is a general improvement in forecast accuracy, the tendency for both models to forecast VCI3M below the drought threshold do not differ vastly especially from 10 to 12 weeks ahead.

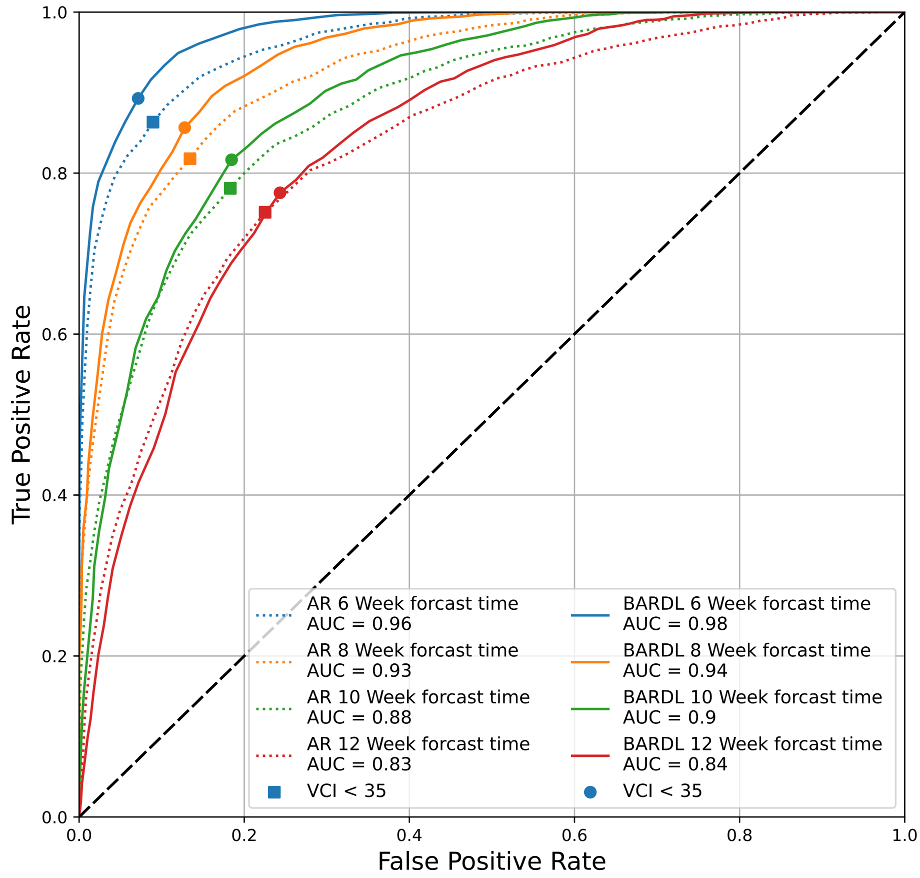

Figure 4.6: ROC Curve showing True Positive Rate (TPR), False Positive Rates(FPR) and AUC for 6,8,10,12 weeks for both AR (Dotted line) and BARDL (Solid line) forecasts. The VCI3M $<$ 35 threshold is plotted as points on the lines.

4.4.4 Forecast Reliability

Using the Bayesian approach also enabled the computation of forecast probabilities for a given drought event (No Drought Condition – VCI3M>35 or Drought Condition – VCI3M<35). To assess the skill for forecasting drought probabilities, we used the reliability diagrams in figure 4.7. The plot shows a joint distribution between the forecast probabilities in bins and the frequencies of the observed drought events that fall in those bins. For each lead time, the sharpness histogram which shows the frequency at which an event is forecasted are also plotted (WWRP, 2009). The reliability of a perfect model would follow the line $y = x$ which has been represented by a dashed line in figure 4.7. The closer a model is to this dashed line, the more reliable it is. Figure 4.7 shows the reliability for drought events (VCI3M<35) in arid counties, the forecast skill assessment of our BARDL model indicates that when we forecast a 'Drought' condition with a probability between 80% to 100% for 6 week lead time, it corresponds with the observed drought events about 88% to 99% of the time. In terms of the model's sharpness, it can be seen that most of the drought events forecasted by the BARDL model have a probability between 90% to 100%. The peak at the 0% to 10% bin of the sharpness plot shows the frequency of 'No Drought' forecasts in the arid counties. This indicates the likelihood of the model missing some drought events especially from 8 weeks lead time and beyond.

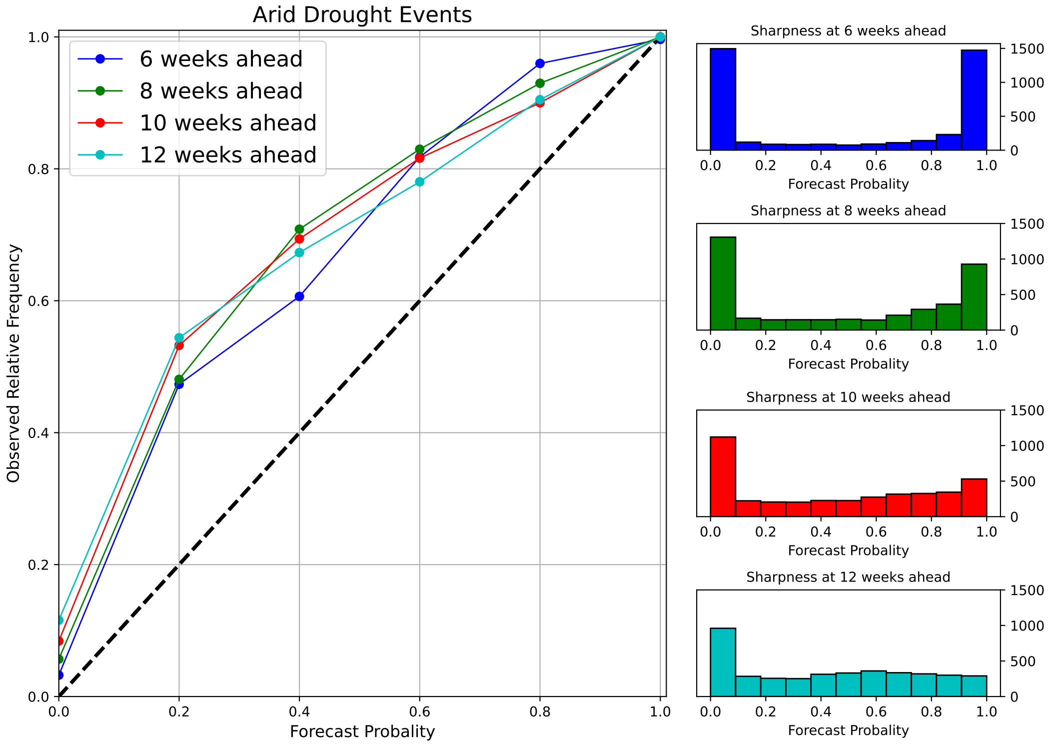

Figure 4.7: Reliability diagram showing forecast probability and their corresponding observed frequencies for 6, 8, 10, 12 weeks lead time together with their corresponding sharpness plots for drought events (VCI3M< 35) in the arid and semi-arid counties

4.4.5 Relative Importance

Figure 4.8 shows the cumulative percentage relative importance for the lags of VCI3M, P3M anomaly and SM3M anomaly. The lags of VCI3M contributes the most for shorter lead time and decreases longer lead times. The precipitation anomaly also contributes significantly to future VCI3M and its relative importance increases with increasing forecast lead times. The relative importance of soil moisture, although it varies less across various lead times, also contributes significantly. Detailed plots of the relative importance for individual lag contribution for each arid and semi-arid county in figure 4.9 (Appendix B). A critical look at these plots also showed that VCI3M responds better to precipitation anomaly in most arid counties like Turkana and Wajir compared to semi-arid counties like Kitui and West-Pokot.

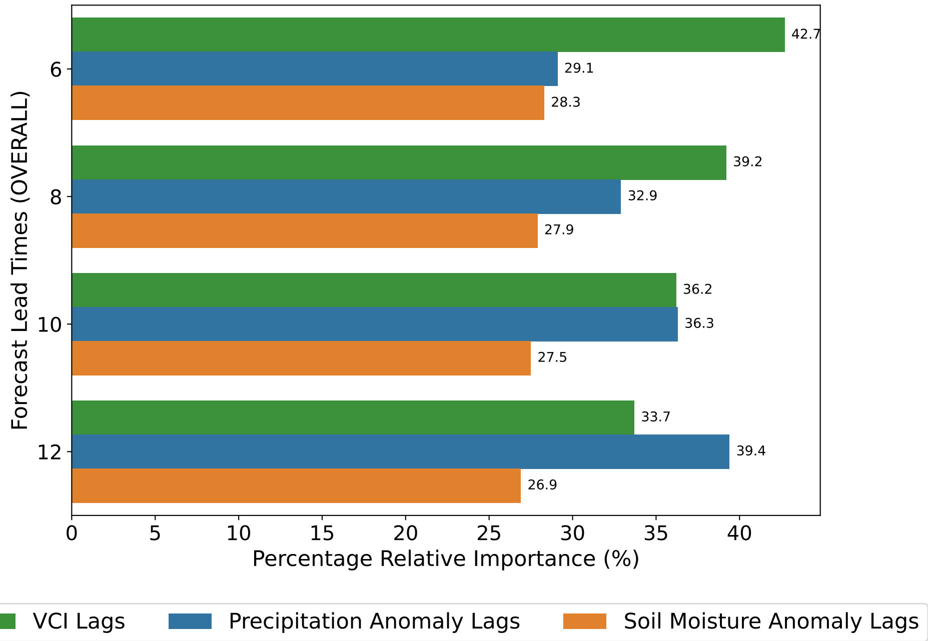

Figure 4.8: Bar plots showing the cumulative (All lags) relative importance of additional variables to the VCI3M forecast for all counties

4.5 Discussion

In this paper we increased the range of VCI3M forecasts, using additional lagged information from P3M and S3M anomalies. The VCI3M used here was derived from the 12-week rolling mean of VCI, as used by Kenya's National Drought Management Authority (NDMA) for monitoring and reporting agricultural droughts occurrences. The soil moisture data especially, though retrieved via a combination of remote sensing and a soil moisture model (Gruber et al., 2019), has proved useful for drought monitoring and forecast. The extensive model skill assessments done here shows that our Bayesian ARDL approach not only performs better compared to results from previous studies (Barrett et al., 2020) but also, the BARDL model, by design, provides additional uncertainty information for better decision making.

Our BARDL which incorporates the precipitation and soil moisture exhibited a 2-week gain in forecast range with overall R^2 scores of 0.94, 0.85 and 0.74 at 6, 8 and 10 weeks lead time respectively. Our forecasts were mostly driven by the variables at lag 0. However, the collective contribution of the additional lags substantially improved the forecast ranges. Finally, the skill assessment based on forecast probabilities indicated a good separation between No-Drought and Drought events.

The results from the model evaluation revealed a strong persistence within soil moisture and VCI3M, a property that enables future values to be inferred from their past values (AghaKouchak, 2014). Despite this inherent persistence in the VCI3M, it still required the information from additional biophysical factors to improve its forecast range as seen in figure 4.9 and the overall performance of the BARDL model. Another interesting observation from figure 4.8 also showed that VCI3M responded very slowly to short term moisture anomalies (Quiring et al., 2010; Vicente-Serrano, 2007). From a spatial perspective, both models (AR & BARDL) gave a higher forecast R^2 score in the arid areas compared to the semi-arid areas. This was more significant for the BARDL model.

Further evaluation of forecasts based on Kenya's long rain (March, April, May (MAM)) and short rain (October, November, December (OND)) seasons (Camberlin et al., 1997) also showed even better R^2 score for longer range forecast in MAM season compared to the OND for the BARDL model (see figures 4.11 and 4.12). This indicates that although VCI3M responds slowly to short term moisture levels, the impact of precipitation and soil moisture on vegetation condition is very important. The R^2 scores for the AR model in the MAM season however dropped significantly compared to the OND season. A possible reason for this observation, especially during the MAM season, could be attributed to the absence of information from the moisture levels (precipitation and soil moisture) in the AR model. The relative importance (figure 4.10) of the lagged exogenous factors for different seasons also confirms the reliance of future VCI3M on precipitation anomalies. The contribution of the lagged soil moisture anomalies during the OND seasons also increased compared to the MAM season. This also indicates that during the short rain season, vegetation condition is also controlled by soil moisture. When it comes to forecasting drought events, a much higher frequency is seen during the OND season (see figure 4.13. This is expected since there are fewer rains in the OND seasons.

Aside from the significant improvements in the forecast range and precision, the strength of our model hinges on the fact that we implemented it in a Bayesian context. Using the Bayesian approach yields a full probability distribution of forecasted VCI3M values which gave us the power to easily gain insight into the uncertainty of forecasted VCI3M values (Lambert, 2018). It also allowed the computation of probabilistic forecast of specific drought events (e.g. VCI3M falling in a particular range) (Wilks, 2006). For our target end-users and stakeholders like the NDMA, using the Bayesian model proposed in this paper as part of their EWS will enable them to confidently report on drought events. Also, policymakers and administrators of disaster relief organisations based on the forecast-based finance initiatives (Coughlan de Perez et al., 2015), can make better decisions and prioritise which drought alarms to act on. This will help with the efficient management of funds.

Although we have shown that we can extend forecast ranges with the added variables, the limitations to work include, the availability of soil moisture data. The ESA CCI Soil Moisture products used in this paper are released annually and are also a year behind. Thus they cannot currently be used for producing real-time forecasts. Another limitation was the use of the 2016 ESA Sentinel 2 land cover map for sampling grassland and shrub pixels across an 18-year period. Even though the land cover product accurately depicted areas with grassland and shrubs, pixel values from regions with significant land cover changes over time may affect the results.

4.6 Conclusion and Future Work

In this study we have made two key developments, these include primarily, the improvement in the forecast range of VCI3M using lagged information from precipitation (P3M) and soil moisture (S3M) by approximately 2 weeks compared to previous works. Secondly, modelling within the Bayesian framework also gave the added advantage of easily assessing model uncertainty and forecast probability of a drought event.

The forecast-based finance initiatives aimed at monitoring agricultural drought indicators and their impact on livelihoods should consider Bayesian approaches to enable better decision making. We would also recommend that soil moisture data be made available sooner and promptly to enable near real-time forecasting of vegetation condition via our proposed method.

The disparity in model performance between arid and semi-arid regions points to the fact that the differences in climate and vegetation land use and land cover (LULC) should also be considered when developing such forecast models. A natural expansion of our BARDL model would be to simultaneously explore and model for spatial variations like LULC in a county or any region of interest via a hierarchical modelling approach. Doing this will give us the advantage of pooling information between spatial variations, whilst still allowing flexibility between them.

Author Contribution

E.E.S. lead author, data preprocessing, modelling & running BARDL method; J.M.M. data acquisition, preprocessing, cartography and feedback; A.B.B. code for AR method; A.B. code for smoothing time series data; S.O., P.R., & P.H. conceptualised the initial idea and provided supervision and feedback; The final manuscript was edited and reviewed by all authors.

Competing Interests

All authors of the paper declare no known competing interests (financial, personal relationships) that could have influenced this study.

Acknowledgements

The work was funded by the UK Newton Fund's Development in Africa with Radio Astronomy (DARA) Big Data project delivered via STFC with grant number ST/R001898/1

References

Adede, Chrisgone, Robert Oboko, Peter Wagacha and Clement Atzberger (2019). "A mixed model approach to drought prediction using artificial neural networks: Case of an operational drought monitoring environment." In: *arXiv Learning* Figure 1, pp. 1–18. DOI: arXiv:1901.04927v1. URL: http://arxiv.org/abs/1901.04927.

AghaKouchak, A. (2014). "A baseline probabilistic drought forecasting framework using standardized soil moisture index: Application to the 2012 United States drought". In: *Hydrology and Earth System Sciences* 18.7, pp. 2485–2492. ISSN: 16077938. DOI: 10.5194/hess-18-2485-2014.

Akaike, Hirotogu (1998). "Information theory and an extension of the maximum likelihood principle". In: *Selected papers of hirotugu akaike*. Springer, pp. 199–213.

Asaad, Al-ahmadgaid B and Joselito C Magadia (2019). "Stochastic Gradient Hamiltonian Monte Carlo on Bayesian Time Series Modeling". In: *14th National Convention on Statistics Crowne*.

Barrett, Adam B, Steven Duivenvoorden, Edward E Salakpi, James M Muthoka, John Mwangi, Seb Oliver and Pedram Rowhani (2020). "Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya". In: *Remote Sensing of Environment* 248, p. 111886.

Ben Taieb, Souhaib and Rob J. Hyndman (2014). "Recursive and direct multi-step forecasting: the best of both worlds". In: *International Journal of Forecasting* September.

Ben Taieb, Souhaib, Antti Sorjamaa and Gianluca Bontempi (June 2010). "Multiple-output modeling for multi-step-ahead time series forecasting". In: *Neurocomputing* 73.10-12, pp. 1950–1957. ISSN: 09252312. DOI: 10.1016/j.neucom.2009.11.030.

Bradley, Andrew P. (1997). "The use of the area under the ROC curve in the evaluation of machine learning algorithms". In: *Pattern Recognition* 30.7, pp. 1145–1159. ISSN: 0031-3203. DOI: https://doi.org/10.1016/S0031-3203(96)00142-2. URL: https://www.sciencedirect.com/science/article/pii/S0031320396001422.

Cabot Venton, Courtenay, Catherine Fitzgibbon, Tenna Shitarek, Lorraine Coulter and Olivia Dooley (2012). *Economics of Resilience Final Report The Economics of Early Response and Disaster Resilience: Lessons from Kenya and Ethiopia*. Tech. rep.

Camberlin, P. and J. G. Wairoto (1997). "Intraseasonal wind anomalies related to wet and dry spells during the "long"and "short"rainy seasons in Kenya". In: *Theoretical and Applied Climatology* 58.1, pp. 57–69. DOI: 10.1007/BF00867432. URL: https://doi.org/10.1007/BF00867432.

Cenacchi, Nicola (2014). "Drought Risk Reduction in Agriculture: A Review of Adaptive Strategies in East Africa and the Indo-Gangetic Plain of South Asia". In: *Browser Download This Paper* September.

Coughlan de Perez, E., B. van den Hurk, M. K. van Aalst, B. Jongman, T. Klose and P. Suarez (2015). "Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts". In: *Natural Hazards and Earth System Sciences* 15.4, pp. 895–904. DOI: 10.5194/nhess-15-895-2015. URL: https://nhess.copernicus.org/articles/15/895/2015/.

Deleersnyder, Roxanna (2018). *Pastoralism in East Africa: challenges and solutions — Glo.be.* URL: https://www.glo-be.be/index.php/en/articles/pastoralism-east-africa-challenges-and-solutions (visited on 09/05/2021).

Dorigo, Wouter et al. (2017). "ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions". In: *Remote Sensing of Environment* 203, pp. 185–215. ISSN: 00344257. DOI: 10.1016/j.rse.2017.07.001. URL: https://doi.org/10.1016/j.rse.2017.07.001.

Eilers, Paul H. C. (2003). "A Perfect Smoother". In: *Analytical Chemistry* 75.14. PMID: 14570219, pp. 3631–3636. DOI: 10.1021/ac034173t. eprint: https://doi.org/10.1021/ac034173t. URL: https://doi.org/10.1021/ac034173t.

FAO (2017). *Easing the impact of drought in Kenya : FAO in Emergencies*. URL: http://www.fao.org/emergencies/resources/photos/photo-detail/en/c/1053828/ (visited on 02/02/2021).

— (2018). *Pastoralism in Africa's drylands Reducing risks, addressing vulnerability and enhancing resilience*. Tech. rep. URL: http://www.fao.org/3/ca1312en/CA1312EN.pdf.

Fawcett, Tom (2006). "An introduction to ROC analysis". In: *Pattern recognition letters* 27.8, pp. 861–874.

FEWSNET (2019). *Famine Early Warning Systems Network. FEWS NET*. URL: https://fews.net/ (visited on 09/05/2021).

Funk, Chris et al. (Dec. 2015). "The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes". In: *Scientific Data* 2.1, pp. 1–21. ISSN: 20524463. DOI: 10.1038/sdata.2015.66. URL: https://www.nature.com/articles/sdata201566.

Gebremeskel, Gebremedhin, Qiuhong Tang, Siao Sun, Zhongwei Huang, Xuejun Zhang and Xingcai Liu (June 2019). *Droughts in East Africa: Causes, impacts and resilience.* DOI: 10.1016/j.earscirev.2019.04.015.

Gruber, Alexander, Tracy Scanlon, Robin Van Der Schalie, Wolfgang Wagner and Wouter Dorigo (2019). "Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology". In: *Earth System Science Data* 11.2, pp. 717–739. ISSN: 18663516. DOI: 10.5194/essd-11-717-2019.

Gujarati, D.N. (2003). *Basic Econometrics.* Economic series. McGraw Hill. ISBN: 9780072335422. URL: https://books.google.co.uk/books?id=byu7AAAAIAAJ.

Hoffman, Matthew D. and Andrew Gelman (2014). "The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo". In: *Journal of Machine Learning Research* 15.2008, pp. 1593–1623. ISSN: 15337928. arXiv: 1111.4246.

Hyndman, Rob J and George Athanasopoulos (2018). *Forecasting: principles and practice.* OTexts.

Ji, Lei and Albert J. Peters (2005). "Lag and seasonality considerations in evaluating AVHRR NDVI response to precipitation". In: *Photogrammetric Engineering and Remote Sensing* 71.9, pp. 1053–1061. ISSN: 00991112. DOI: 10.14358/PERS.71.9.1053.

Klisch, Anja and Clement Atzberger (2016). "Operational drought monitoring in Kenya using MODIS NDVI time series". In: *Remote Sensing* 8.4. ISSN: 20724292. DOI: 10.3390/rs8040267.

Kogan, F. N. (1995). "Application of vegetation index and brightness temperature for drought detection". In: *Advances in Space Research* 15.11, pp. 91–100. ISSN: 02731177. DOI: 10.1016/0273-1177(95)00079-T.

Lambert, B. (2018). *A Student's Guide to Bayesian Statistics.* SAGE Publications. ISBN: 9781526418289. URL: https://books.google.co.uk/books?id=ZvBUDwAAQBAJ.

Lei Ji and A. J. Peters (2004). "Forecasting vegetation greenness with satellite and climate data". In: *IEEE Geoscience and Remote Sensing Letters* 1.1, pp. 3–6. DOI: 10.1109/LGRS.2003.821264.

Martin, O. (2018). *Bayesian Analysis with Python: Introduction to statistical modeling and probabilistic programming using PyMC3 and ArviZ, 2nd Edition.* Packt Publishing. ISBN: 9781789349665. URL: https://books.google.co.uk/books?id=1Z2BDwAAQBAJ.

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan.* Chapman & Hall/CRC Texts in Statistical Science. CRC Press. ISBN: 9781482253481. URL: https://books.google.co.uk/books?id=1yhFDwAAQBAJ.

Milford, J. R. and G. Dugdale (1990). "Monitoring of rainfall in relation to the control of migrant pests". In: *Philosophical Transactions - Royal Society of London, B* 328.1251, pp. 689–704. ISSN: 0080-4622. DOI: 10.1098/rstb.1990.0137.

Neal, Radford M (1993). *Probabilistic inference using Markov chain Monte Carlo methods.* Department of Computer Science, University of Toronto Toronto, Ontario, Canada.

Pang, Jingyue, Datong Liu, Yu Peng and Xiyuan Peng (2018). "Optimize the coverage probability of prediction interval for anomaly detection of sensor-based monitoring series". In: *Sensors (Switzerland)* 18.4. ISSN: 14248220. DOI: 10.3390/s18040967.

Pesaran, M. Hashem and Yongcheol Shin (1999). "An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis". In: *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium.* Ed. by SteinarEditor Strøm. Econometric Society Monographs. Cambridge University Press, pp. 371–413. DOI: 10.1017/CCOL521633230.011.

Quiring, Steven M. and Srinivasan Ganesh (2010). "Evaluating the utility of the Vegetation Condition Index (VCI) for monitoring meteorological drought in Texas". In: *Agricultural and Forest Meteorology* 150.3, pp. 330–339. ISSN: 0168-1923. DOI: https://doi.org/10.1016/j.agrformet.2009.11.015. URL: https://www.sciencedirect.com/science/article/pii/S0168192309002809.

Rippa, Shmuel (1999). "An algorithm for selecting a good value for the parameter c in radial basis function interpolation". In: *Advances in Computational Mathematics* 11.2, pp. 193–210. DOI: 10.1023/A:1018975909870. URL: https://doi.org/10.1023/A:1018975909870.

Robert, Christian P., Víctor Elvira, Nick Tawn and Changye Wu (Sept. 2018). "Accelerating MCMC algorithms". In: *Wiley Interdisciplinary Reviews: Computational Statistics* 10.5, e1435. ISSN: 1939-0068. DOI: 10.1002/WICS.1435. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/wics.1435%20https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.1435%20https://onlinelibrary.wiley.com/doi/10.1002/wics.1435.

Salvatier, John, Thomas V Wiecki and Christopher Fonnesbeck (2016). "Probabilistic programming in Python using PyMC3". In: *PeerJ Computer Science* 2, e55.

Schaaf, Crystal and Z. Wang (2015). *MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006 [Data set].* DOI: 10.5067/MODIS/MCD43A4.

006. URL: https://lpdaac.usgs.gov/dataset%7B%5C_%7Ddiscovery/modis/modis%7B%5C_%7Dproducts%7B%5C_%7Dtable/mcd43a4%7B%5C_%7Dv006.

Sibanda, Mbulisi, Onisimo Mutanga, Mathieu Rouget and Lalit Kumar (2017). "Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives". In: *Remote Sensing* 9.1. ISSN: 20724292. DOI: 10.3390/rs9010055.

Su, Dongqi, Ying Yin Ting and Jason Ansel (2018). "Tight Prediction Intervals Using Expanded Interval Minimization". In: arXiv: 1806.11222. URL: http://arxiv.org/abs/1806.11222.

Tonidandel, Scott and James M LeBreton (2011). "Relative importance analysis: A useful supplement to regression analysis". In: *Journal of Business and Psychology* 26.1, pp. 1–9.

UNFCCC (2015). *ADOPTION OF THE PARIS AGREEMENT - Paris Agreement text English.* Tech. rep.

Vatter, Juliane (2019). *DROUGHT RISK The Global Thirst for Water in the Era of Climate Crisis.* Tech. rep. World Wildlife Fund (WWF) Germany. URL: www.studioazola.com.

Vicente-Serrano, Sergio M. (2007). "Evaluating the impact of drought using remote sensing in a Mediterranean, Semi-arid Region". In: *Natural Hazards* 40.1, pp. 173–208. ISSN: 0921030X. DOI: 10.1007/s11069-006-0009-7.

Wilks, DS (2006). "Statistical methods in the atmospheric sciences". In.

WWRP (2009). *World Weather Research Programme (WWRP), Forecast Verification - Methods and FAQ.* URL: https://www.cawcr.gov.au/projects/verification/verif%5C_web%5C_page.html (visited on 22/06/2021).

Yang, Yongke, Pengfeng Xiao, Xuezhi Feng and Haixing Li (2017). "Accuracy assessment of seven global land cover datasets over China". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 125.April 2018, pp. 156–173. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2017.01.016. URL: http://dx.doi.org/10.1016/j.isprsjprs.2017.01.016.

4.7 Appendix

4.8 A table showing the PICP and MPIW (in brackets) estimates for the arid and semi-arid counties

Table 4.3: The PICP and MPIW (in parenthesis) estimates for the all arid and semi-arid counties.

	County	AR Model				BARDL Model			
		6	**8**	**10**	**12**	**6**	**8**	**10**	**12**
Arid Counties	Garissa	0.93 (0.33)	0.94 (0.46)	0.94 (0.57)	0.93 (0.67)	0.86 (0.21)	0.84 (0.31)	0.83 (0.39)	0.81 (0.46)
	Isiolo	0.93 (0.29)	0.93 (0.4)	0.93 (0.51)	0.93 (0.6)	0.94 (0.18)	0.92 (0.27)	0.9 (0.36)	0.89 (0.43)
	Mandera	0.93 (0.33)	0.94 (0.44)	0.94 (0.55)	0.93 (0.63)	0.94 (0.23)	0.95 (0.34)	0.96 (0.44)	0.96 (0.53)
	Marsabit	0.92 (0.25)	0.91 (0.36)	0.93 (0.46)	0.94 (0.56)	0.93 (0.15)	0.9 (0.23)	0.88 (0.32)	0.88 (0.38)
	Samburu	0.95 (0.26)	0.94 (0.37)	0.95 (0.47)	0.95 (0.56)	0.95 (0.17)	0.97 (0.27)	0.95 (0.37)	0.94 (0.44)
	Tana-River	0.94 (0.32)	0.93 (0.43)	0.94 (0.51)	0.94 (0.58)	0.87 (0.2)	0.86 (0.28)	0.85 (0.35)	0.85 (0.41)
	Turkana	0.95 (0.24)	0.95 (0.34)	0.95 (0.43)	0.95 (0.52)	0.92 (0.14)	0.92 (0.23)	0.93 (0.33)	0.94 (0.41)
	Wajir	0.94 (0.37)	0.94 (0.49)	0.94 (0.59)	0.95 (0.67)	0.9 (0.22)	0.89 (0.32)	0.9 (0.4)	0.9 (0.48)
	Mean	**0.94 (0.3)**	**0.94 (0.41)**	**0.94 (0.51)**	**0.94 (0.6)**	**0.91 (0.19)**	**0.91 (0.28)**	**0.9 (0.37)**	**0.9 (0.44)**

		6	**8**	**10**	**12**	**6**	**8**	**10**	**12**
Semi-Arid Counties	Baringo	0.95 (0.29)	0.96 (0.42)	0.95 (0.54)	0.95 (0.65)	0.95 (0.22)	0.94 (0.36)	0.94 (0.49)	0.95 (0.61)
	Kajiado	0.93 (0.3)	0.93 (0.42)	0.93 (0.53)	0.93 (0.63)	0.94 (0.18)	0.93 (0.29)	0.94 (0.4)	0.93 (0.48)
	Kilifi	0.94 (0.23)	0.94 (0.31)	0.95 (0.36)	0.94 (0.41)	0.88 (0.2)	0.89 (0.28)	0.89 (0.36)	0.9 (0.42)
	Kitui	0.93 (0.34)	0.95 (0.47)	0.94 (0.57)	0.94 (0.64)	0.9 (0.21)	0.89 (0.31)	0.88 (0.4)	0.89 (0.47)
	Laikipia	0.94 (0.24)	0.95 (0.35)	0.96 (0.46)	0.96 (0.56)	0.96 (0.17)	0.95 (0.28)	0.94 (0.4)	0.93 (0.5)
	Makueni	0.94 (0.34)	0.94 (0.46)	0.93 (0.56)	0.94 (0.64)	0.93 (0.22)	0.91 (0.32)	0.88 (0.4)	0.89 (0.47)
	Meru	0.95 (0.3)	0.95 (0.43)	0.95 (0.54)	0.95 (0.62)	0.93 (0.2)	0.93 (0.31)	0.92 (0.4)	0.91 (0.47)
	Narok	0.95 (0.27)	0.95 (0.37)	0.94 (0.45)	0.94 (0.53)	0.95 (0.19)	0.95 (0.29)	0.93 (0.39)	0.92 (0.48)
	Nyeri	0.94 (0.23)	0.95 (0.32)	0.96 (0.41)	0.95 (0.49)	0.91 (0.18)	0.89 (0.27)	0.88 (0.35)	0.89 (0.43)
	Taita-Taveta	0.92 (0.32)	0.92 (0.44)	0.92 (0.55)	0.93 (0.63)	0.85 (0.2)	0.84 (0.29)	0.84 (0.38)	0.85 (0.44)
	Tharaka-Nithi	0.94 (0.26)	0.94 (0.37)	0.95 (0.45)	0.94 (0.52)	0.92 (0.21)	0.91 (0.3)	0.9 (0.38)	0.9 (0.45)
	West-Pokot	0.96 (0.25)	0.96 (0.36)	0.95 (0.47)	0.95 (0.56)	0.95 (0.19)	0.94 (0.32)	0.93 (0.44)	0.95 (0.54)
	Mean	**0.94 (0.28)**	**0.94 (0.39)**	**0.94 (0.49)**	**0.94 (0.57)**	**0.92 (0.2)**	**0.91 (0.3)**	**0.91 (0.4)**	**0.91 (0.48)**

4.9 Relative Importance plots for each county

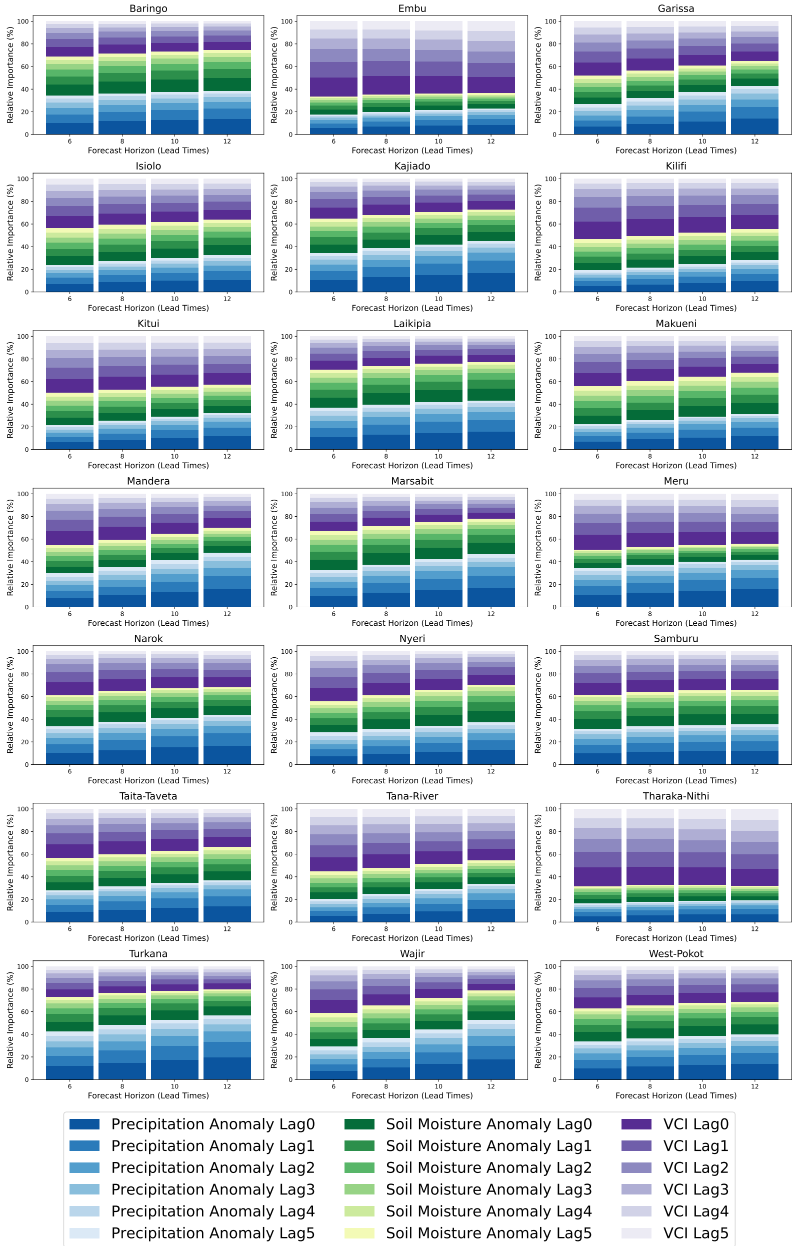

Figure 4.9: Relative Importance for each exogenous factors for each lag (0-5) variable per county.

4.10 Relative Importance plots for MAM and OND seasons

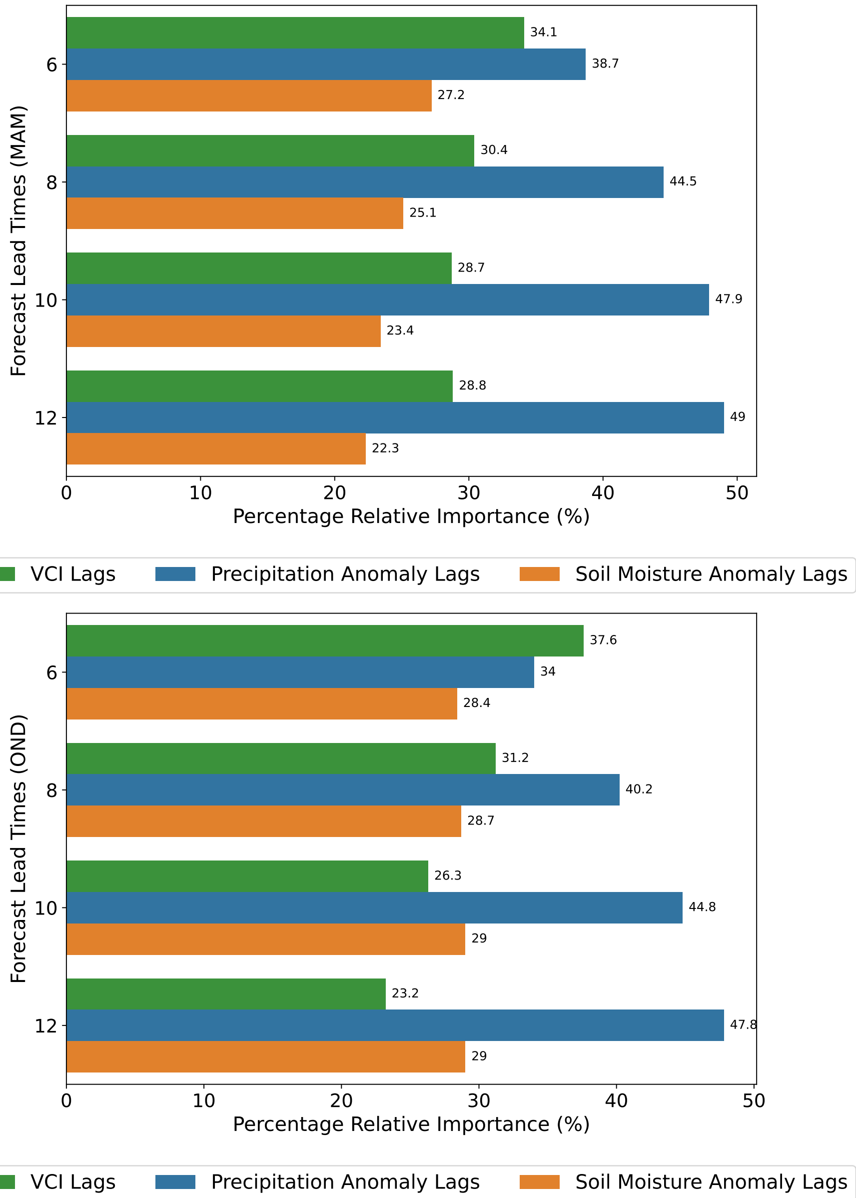

Figure 4.10: Cumulative lag relative importance plots for counties for the MAM and OND Seasons.

4.11 Contour plots showing forecast performance for MAM and OND seasons

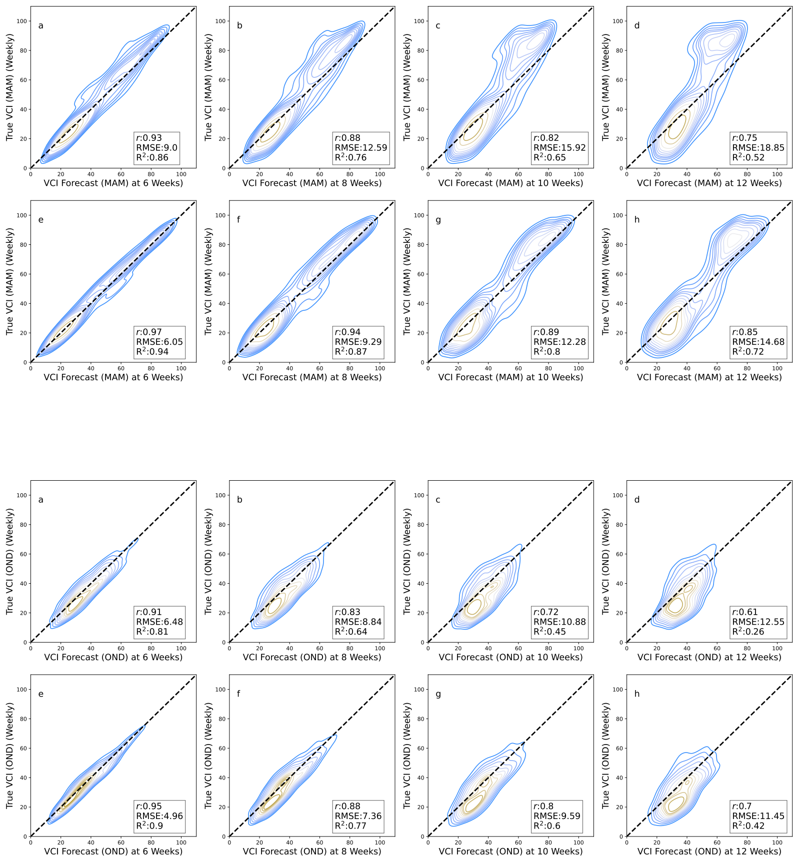

Figure 4.11: Contour plots showing VCI3M forecast against True VCI3M for MAM and OND Seasons. Plots (a,b,c,d) shows the results from the AR method with VCI3M only, (e,f,g,h) shows the overall results for BARDL modelled with lags of VCI3M plus lags of Precipitation (P3M) and Soil Moisture (S3M) Anomalies for 6, 8, 10 and 12 weeks lead time for all counties

4.12 Forecast performance metrics for MAM and OND seasons

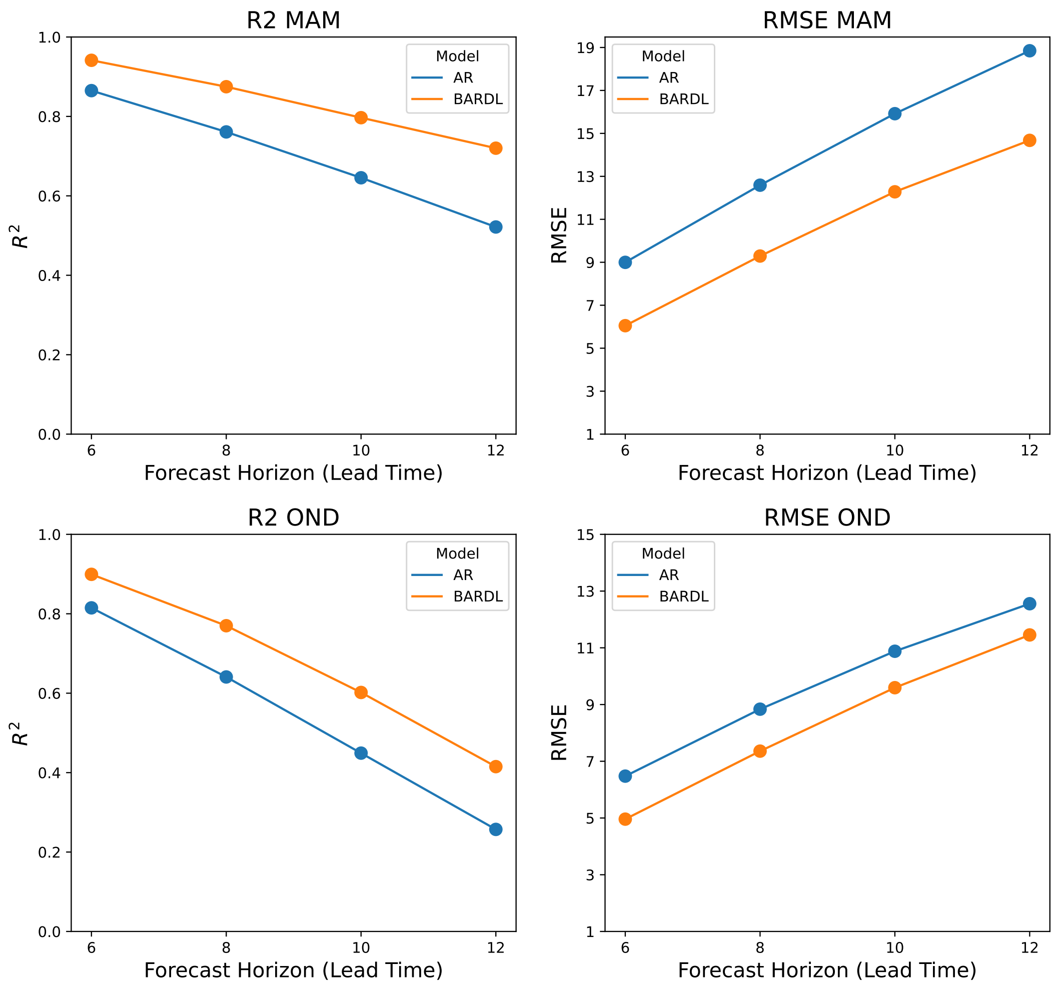

Figure 4.12: Performance metrics used to measure model accuracy as a function of forecast lead time for MAM and OND Season.

4.13 Forecast reliability for MAM and OND seasons

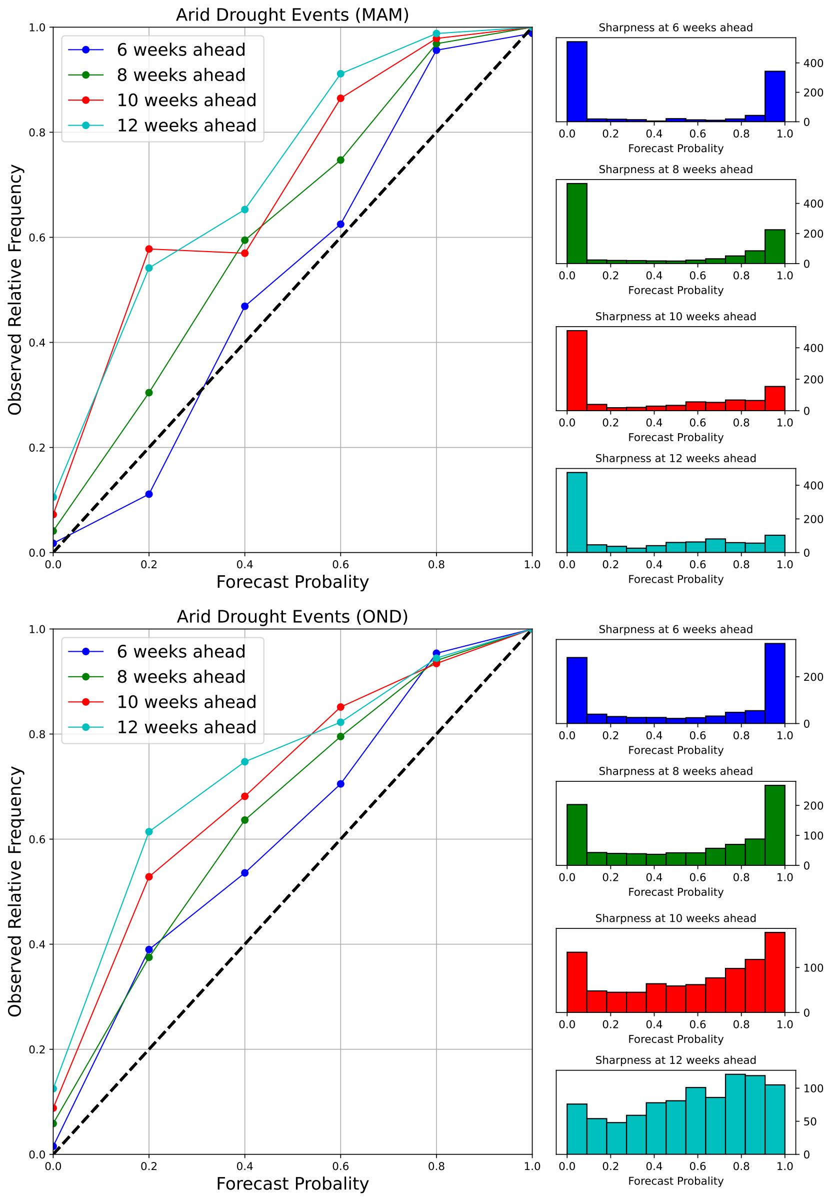

Figure 4.13: Reliability diagram showing forecast probability and their corresponding observed frequencies for 6, 8, 10, 12 weeks lead time together with their corresponding sharpness plots for drought events (VCI3M< 35) MAM and OND

Chapter 5

A Dynamic Hierarchical Bayesian Approach for Forecasting Vegetation Conditions

Edward E. Salakpi[a], Pete Hurley[a,b], James M. Muthoka[c], Andrew Bowell[a,b] Seb Oliver[a,b] and Pedram Rowhani[c]

[a] The Data Intensive Science Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

[b] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK

[c] School of Global Studies, Department of Geography, University of Sussex, Brighton, BN1 9QJ, UK

*Corresponding author: e.salakpi@sussex.ac.uk

Abstract

Agricultural drought, which occurs due to a significant reduction in the moisture required for vegetation growth, is the most complex amongst all drought categories. The onset of agriculture drought is slow and can occur over vast areas with varying spatial effects, differing in areas with a particular vegetation land cover or specific agro-ecological sub-regions. These spatial variations imply that monitoring and forecasting agricultural drought require complex models that consider the spatial variations in a given region of interest. Hierarchical Bayesian Models are suited for modelling such complex systems. Using partially pooled data with sub-groups that characterise spatial differences, these models can capture the sub-group variation while allowing flexibility and information sharing between these sub-groups. This paper's objective was to improve the accuracy and precision of agricultural drought forecast in spatially diverse regions with a Hierarchical Bayesian Model. Results showed that the Hierarchical Bayesian Model was better at capturing the variability for the different agro-ecological zones and vegetation land covers compared to a regular Bayesian Auto-Regression Distributed Lags model. The forecasted vegetation condition and associated drought probabilities were more accurate and precise with the Hierarchical Bayesian Model at 4 to 10 weeks lead times. Forecasts from the hierarchical model exhibited higher hit rates with a low probability of false alarms for drought events in semi-arid and arid zones. The Hierarchical Bayesian Model also showed good transferable forecast

skills over counties not included in the training data.

keywords: Drought; Hierarchical Bayesian Models; Forecasting; Early Warning Systems; Disaster Risk Reduction; MODIS

5.1 Introduction

Drought is a naturally occurring phenomenon that affects the food security of approximately 55 million people annually and can severely impact a country's economy (Deleersnyder, 2018; Nicolai-Shaw et al., 2017). Drought, in most cases, is associated with below-average precipitation and is referred to as meteorological drought. Prolonged meteorological drought event mainly leads to a significant reduction in the amount of soil moisture required for vegetation growth, thus resulting in an agricultural drought (Heim, 2002; Boken et al., 2005). Hence, agricultural drought events are considered a physical manifestation of meteorological drought (Boken et al., 2005). Agricultural drought, which is the focus of this paper, is the most complex amongst the drought categories (Boken et al., 2005). Its onset can be slow and can occur in vast areas with varying spatial impact (Boken et al., 2005). For instance, the impact of drought may differ within a given region depending on whether they are dominated by trees, grasslands or croplands. Spatial differences in drought impact can also arise due to the varied agro-ecological sub-regions within an affected area. These differences indicate that Early Warning Systems (EWS) for agricultural drought will require very complex models.

Drought EWS have been recognised by global initiatives like the United Nations Sustainable Development Goals (SDG) for effective drought monitoring and hazard preparedness (IISD, 2018). As such, international agencies like United Nations Development Programme (UNDP) and the United States Agency for International Development (USAID) * mandated to monitor drought hazards have developed and deployed several EWS. These systems assist drought management officials and people living in drought-prone communities to prepare for hazardous events (UN, 2018). The Famine Early Warning Systems Network (FEWS NET) † is an example of such EWS. The system, developed by the USAID, utilises household data together with agro-climatic indicators and vegetation health to monitor drought and its impact (FEWSNET, 2021). However, drought forecast for anticipatory action via the FEWS NET platform is mainly based on expert judgement (Funk et al., 2019) rather than the use of advanced statistical methods or machine learning models.

Recent advances in computational power and processing hardware have enabled researchers

*https://usaid.gov/

†https://fews.net/

to develop and deploy machine learning models (Bishop, 2006) such as Support Vector Machines (Shao et al., 2012) and various neural network architectures (Da Silva et al., 2017). Machine learning models enables the construction of predictive or prescriptive models using advanced statistical methods to capture hidden patterns in data (Bishop, 2006). In the field of drought research, most of the data used within machine learning models come from satellite Earth Observation (EO) images. These datasets are available over long temporal periods, cover vast areas and are easy to access. Therefore, they provide a cost-effective way of developing models for monitoring and forecasting drought events over vast regions. Examples of such EO datasets include precipitation, soil moisture levels, Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Vegetation Condition Index (VCI) (Kogan, 1995) all derived from remotely sensed EO data. Nay et al., 2018, for instance, used Gradient Boosting Machine to forecast EVI with lagged spectral bands from the Moderate Resolution Imaging Spectroradiometer (MODIS) EO data. Tian et al., 2019 worked on forecasting dryland vegetation condition using NDVI via an Eco-hydrological model driven by surface water extent also derived from MODIS images. Others include Barrett et al., 2020 and Adede et al., 2019 who applied Gaussian Processes and Artificial Neural Networks respectively in their research to develop robust models for short to medium-term forecasts of vegetation conditions. All the models used in the cited works were mainly implemented by aggregating data over similar land cover types and Agro-Ecological Zones (AEZ). The differences in the AEZs or land covers within the region were not considered.

In a previous study (Salakpi et al., 2021), we used a Bayesian regression method to model the relationship between biophysical drivers and their effect on forecasting vegetation conditions. The approach was based on the classical 'No-pooling' method (See figure 5.1), where we fitted separate regression models to data extracted from their respective regions. Pixels representing the biophysical indicators and vegetation conditions were sampled for different land cover and aggregated over the regions of interest. The modelling approach also treated the effects of climate and other biophysical factors on vegetation conditions independently for each region. The models were very skilful for medium to long term forecasts, but forecasts over regions with extensive cloud cover suffered due to the lack of data.

Although known to vary over the different regions, the effects of biophysical indicators on vegetation also show some similarities across the different regions (Sergio M. Vicente-Serrano, 2007). Data for such analysis can be pooled over all the regions of interest and analysed via the 'Complete-pooling' modelling approach to capture these similarities. This approach allows information sharing between the regions of interest, which is an advantage over the 'No-pooling' approach (Gelman and Hill, 2006). However, the 'Complete-pooling' method is not very useful when the pooled data has sub-groupings, e.g., a pooled soil moisture data from different regions with varied land cover types. In such a case, a more advanced approach would be to combine the strengths of both the 'No-pooling' and 'Complete-pooling' methods into a model known as a 'Partial-pooled' model or 'Hierarchical model' (Gelman and Hill, 2006; Gelman, Carlin et al., 2013). The hierarchical approach, which we demonstrate in this paper, enables flexibility between the sub-groups while treating them independently at the same time (Gelman and Hill, 2006). A Hierarchical Model, when implemented within a Bayesian framework, is referred to as a Hierarchical Bayesian Model (HBM) (Gelman, Carlin et al., 2013). HBMs have in recent times been recognised as a powerful approach for modelling and analysing very complex data. They have been extensively used for research in fields like Astrophysics, Neuroscience and Genetics (Sánchez et al., 2019; George et al., 2005; Storz et al., 2002). Although not commonly used in the study of vegetation dynamics and drought monitoring, Senf et al., 2017 used an HBM to study the spatial and temporal variation in broad-leaved forests phenology using Landsat data.

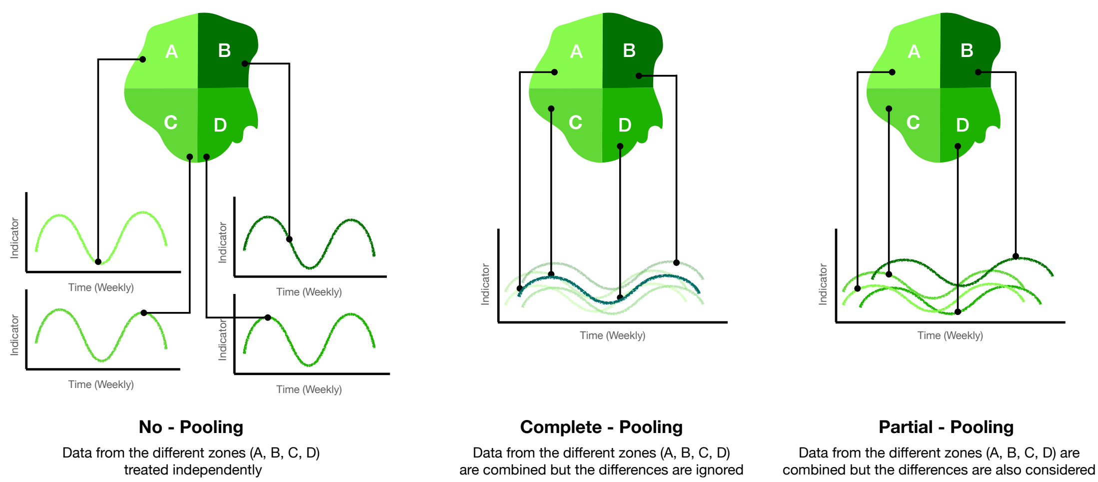

Figure 5.1: Figure illustrating the concept of 'No-Pooling', 'Complete-Pooling' and 'Partial-Pooling' of the data.

The HBM is an extension of the regular Bayesian regression where model parameters differ based on the variations within a given dataset (Gelman, Carlin et al., 2013; Gelman and Hill, 2006). Thus, this paper sought to test the concept of forecasting VCI, an EO based agricultural drought indicator, with an HBM and answer the following question. *'Can we improve forecast accuracy and precision by separately learning parameters for the effects of lagged precipitation* *and soil moisture on vegetation conditions in each AEZ or over varied land cover types?.*

Another advantage of using the HBM is its transferability (Senf et al., 2017). Transfer learning in this context refers to the process where models trained on a given dataset can be re-used to make predictions on different but related data that was not part of the training set (Z. Yang et al., 2017). The partially pooled data used in HBMs makes it suitable for transfer learning primarily because the training data are pooled from multiple regions, and the sub-groupings within the data are the same for the non-training sample data (Rosenstein et al., 2005).

Our objectives for this proof-of-concept are to:

- improve the forecast accuracy and precision of Bayesian Auto-Regression Distributed Lags (BARDL) model with a Hierarchical Bayesian Model in regions with diverse AEZs, and land covers.

- test the transfer learning property of hierarchical model that enables pre-trained models to be used on similar data from a different location without the need to retrain the model (Y. Yang et al., 2017).

5.2 Study Area and Data

5.2.0.1 Study Area

To test our concept of forecasting vegetation condition with HBM, we sampled data from some selected counties in Kenya (Baringo, Kitui, Marsabit, Narok, Tana-River, Turkana), shown in figure 5.2 with red boundary lines. The selected counties have diverse land use and land covers (LULC), ranging from crops to evergreen forests. These counties also have varied AEZs with rainfall and temperature patterns ranging from moderate to extreme. During the short and long rainfall seasons, annual mean precipitation range from 20mm to 200mm. Temperature across these counties also range from as low as $10^{o}C$ to $40^{o}C$ (Ayugi et al., 2016). The main economic activity in these counties is agriculture, predominantly agro-pastoral practices (Gebremeskel et al., 2019; Vatter, 2019). However, extreme climatic variations make this region prone to prolonged drought events, and the impact of these dry spells vary over the various land covers within the AEZs.

We selected only six counties because the algorithm used for parameter sampling by the HBM can be very time-consuming when the input data is more than 10,000 records. The sampling time is also mainly due to the complex nature HBMs.

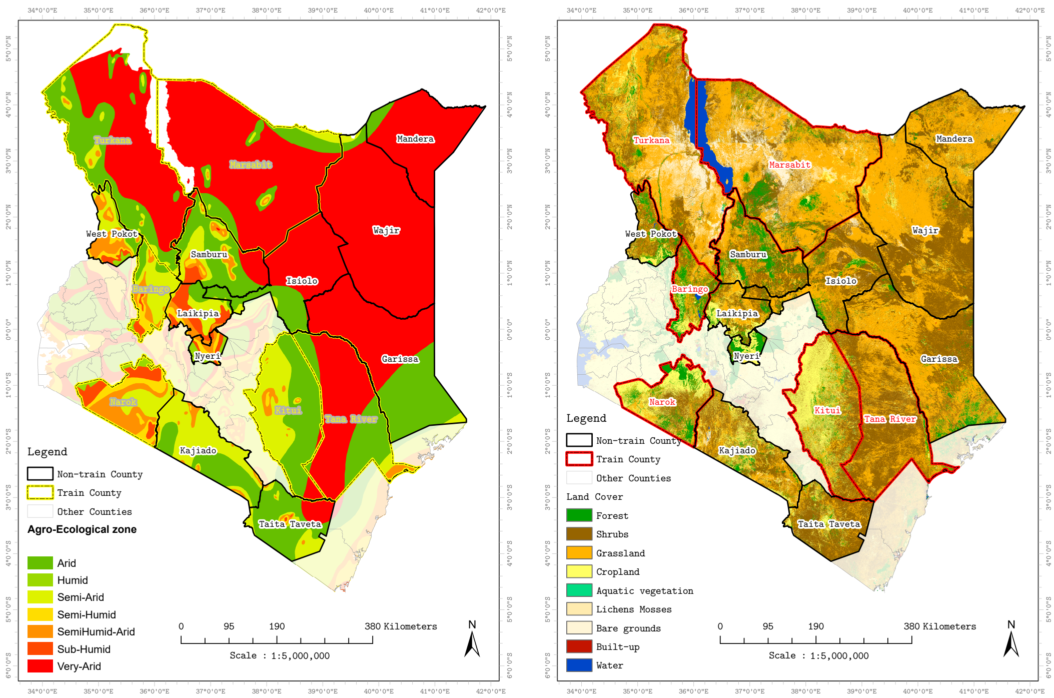

Figure 5.2: Maps of Kenya showing Agro-Ecological Zones (AEZ) and Land Cover maps for the counting from which pixels were sampled. Kenya AEZ boundary maps credit: IGAD Climate Prediction and Application Centre (ICPAC). Land Cover map credit: European Space Agency (ESA), Climate Change Initiative (CCI)

5.2.1 Data

Aside the AEZ boundary shapefiles, all the dataset used in this proof-of-concept is the same data used and described in a previous study here: (Salakpi et al., 2021).

5.2.2 Agro-Ecological Zones & Vegetation Land Covers

Two HBMs were developed in this study, one based on AEZs and the other on land covers. AEZs are geographical areas characterised by similar climatic patterns and soil moisture levels suitable for agriculture and vegetation growth. These zones were created by the Food and Agriculture Organization (FAO) in collaboration with International Institute for Applied Systems Analysis (IIASA) and are based on a framework that utilises a series of models with climate and land use information to quantify and map out the regions (Fischer et al., 2000). The zones are categorised as Humid, Semi-Humid, Arid, Semi-Arid and Very Arid. These AEZs, from their definition, exhibit distinct climate properties; thus, a modelling approach that can separately learn parameters for the effects of precipitation and soil moisture on vegetation conditions based on the difference AEZs can give a more accurate VCI forecast.

The AEZs in our study area include: Source: Sombroek et al., 1982

Table 5.1: Table describing the Agro-Ecological Zone, vegetation type and annual rainfall levels.

Zone Classification	**Vegetation Type**	**Average Annual Zone Rainfall (mm)**
Humid	Moist Forest	1100-2700
Sub-Humid	Moist and Dry Forest	1000-1600
Semi-Humid	Dry Forest and Moist Woodlands	800-1400
Semi-Humid to Arid	Dry Woodland and Bush lands	600-1100
Arid	Bush, Grass and Shrublands	450-900
Semi-Arid	Bush, Grass and Shrublands	300-500
Very-Arid	Desert, Sparse grass and shrub	150-350

Most drought-prone ROIs are made of diverse vegetation covers; these include Tree Covers (Forests), Grasslands, Shrubs and Croplands. The impact of drought on these land cover types varies both spatially and temporally. Thus, a drought forecast model should consider the varying effects of the biophysical factors on the various land covers. Using an HBM framework will enable us to achieve this. Data corresponding to the various vegetation land covers was extracted with the Sentinel 2, 2016, Land Use and Land Cover (LULC) map *.

*Visit this link (http://2016africalandcover20m.esrin.esa.int/) to learn more about the European Space Agency (ESA), Climate Change Initiative (CCI) Sentinel 2 Land Cover Map

5.3 Methodology

5.3.1 Data Pre-Processing

A major challenge with using optical EO images is cloud cover and cloud shadows. In addition, pixel reflectance values sometimes fall outside the meaningful range due to errors during the atmospheric and radiometric correction process. These clouded and poor-quality pixels were filtered out with the quality assurance maps that come with the EO products. Weekly averages of VCI, precipitation and soil moisture data corresponding to the vegetation land covers of interest were extracted from the selected counties using the European Space Agency (ESA) 2016 Sentinel 2 Land Use and Land Cover (LULC) map. Same data within the various AEZs were also extracted using the AEZ shapefiles produced by IGAD Climate Prediction and Application Centre (ICPAC) *. The temporal gaps, left by the removal of clouded pixels, were filled using the Radial Basis Function (BBF) interpolation method, which ensures values obtained through interpolation over wide intervals do not go beyond the valid ranges (Rippa, 1999). The noise resulting from optical instrument failures and gap-filling processes were reduced with a penalised least-squares method (Whittaker smoother) (Eilers, 2003; Klisch et al., 2016). A three-month (12 weeks) rolling average was applied to the VCI to make it VCI3M primarily because our stakeholders use it for their EWS. Applying the rolling averages enhanced the persistence in the VCI. Three-month Precipitation (P3M) and Soil moisture (SM3M) were also computed for consistency. Finally, to avoid the influence of strong seasonal cycles on the forecast values and make data stationary, the VCI3M, P3M and SM3M data were converted to anomalies by subtracting their seasonal means before fitting to the HBM. After forecasting, the subtracted seasonal means for the VCI3M (for each AEZ and land cover) were added back. All the variables were also standardised by subtracting the mean and divided by the standard deviation to make the variable unitless and avoid the dominance of certain variables over others.

5.3.2 Forecast Model

The HBM implemented in this work was done via an Auto-Regressive Distributed Lag (ARDL) model (Gujarati, 2003). The ARDL*(p,q)* is a time series regression method used for multivariate time series analysis where the variable of interest (dependent variable) is modelled with its lags and that of additional explanatory variables (independent variable) (Gujarati, 2003). The p represents the number of lags for the independent variable used in the model, and the q is the auto-regressive part of the model, representing the lags of the dependent variable passed to the

*http://geoportal.icpac.net/layers/geonode%3Aken_aczones

ARDL model. Within the HBM framework, a Bayesian probabilistic approach is used to infer model parameters instead of the Maximum likelihood approach. The data Y for the model is partially pooled as Y_{ij} where i is the index of the variable (e.g. precipitation), and j are the indices of the sub-groups (e.g. AEZs) within the data. This data structure enables parameter inference at both the global θ_i and sub-group levels θ_j at the same time as shown in figure 5.3. Using the Bayesian framework also allows us to incorporate informative priors into the parameter estimation process. Furthermore, we obtain a full posterior probability distribution for both the parameters and forecast values, instead of just point estimates, which enables gives a straightforward way to quantify forecast uncertainties (R. Ravines et al., 2006; Asaad et al., 2019).

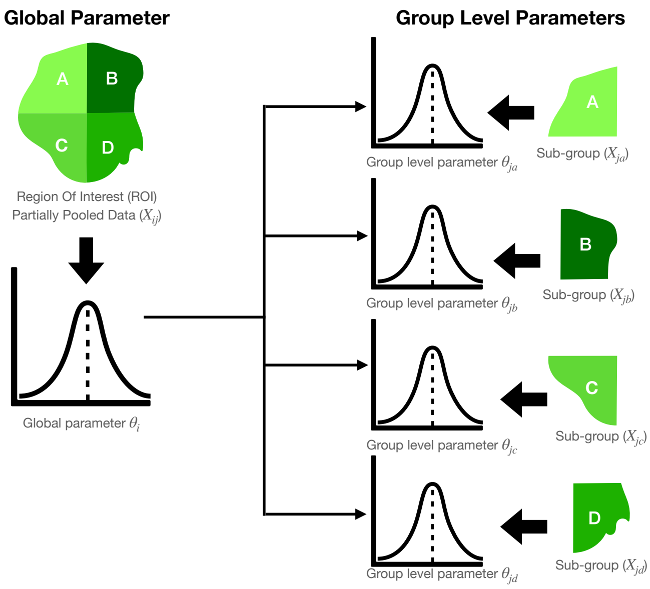

Figure 5.3: An illustration of the parameter structure of the Hierarchical Bayesian model based on partially pooled data (Y_{ij}). The global parameter (θ_i) represents the average posterior parameter distribution over an entire region of interest, while the group level parameters $\theta_{j(abcd)}$ are the individual posterior parameter distributions inferred from the sub group data (Y_{jabc}) within the region of interest.

The Bayesian framework used for the parameter inference is based on Bayes' theorem in equation 5.1:

$$P(\theta|X_t) = \frac{P(X_t|\theta).P(\theta)}{P(X_t)} \tag{5.1}$$

where X_t represents the input data of the ARDL model, $P(\theta|X_t))$ represents the probability of our model parameters given our data X_t also known as the posterior, $P(X_t|\theta)$ represents the probability of the data given the parameters referred to as the likelihood and $P(\theta)$ represents the prior belief about the parameters. $P(X_t)$ is the probability of data or evidence. The evidence is a normalisation term and usually ignored, making the posterior proportional to the likelihood and prior as seen in equation 5.2 (Lambert, 2018; R. McElreath, 2016).

$$P(\theta|X_t) \propto P(X_t|\theta).P(\theta) \tag{5.2}$$

It is important to note that working with the Bayes' framework allows us to explicitly define our prior beliefs about model parameters. These priors are then updated with the likelihood function to generate the posterior probability distribution when informed by observed data.

The HBM will enable us to fit the ARDL model by simultaneously inferring global parameters (Nodes A and B in figure 5.4) across the partially-pooled data as well as their sub-group level variations (Node G in figure 5.4) (Gelman and Hill, 2006). The sub-group levels, in this case, refers to the different LULC or AEZs within our data. The varying effect of the sub-groups was incorporated into our HBM as categorical variables (Node K in figure 5.4).

The HBM was based on an ARDL*(p=6,q=6)*, where the lagged of P3M, SM3M and VCI3M were all set to lags of 6 weeks. The nature of the input variables suggests a high likelihood for our model parameters to have a strong correlation. We addressed this by modelling our group-level parameters as a multivariate normal distribution using a Cholesky matrix decomposition as hyper-priors (prior of a prior distribution) (Nodes C, D and E in figure 5.4) (Richard McElreath, 2018). The Cholesky factorisation was used to transform the multivariate distribution to increase the efficiency of parameter sampling (Stan Development Team, 2018). However, since the HBM group-level parameters are modelled as conditional probabilities of the global parameters, the group level parameter tends not to separate well from the global mean. When this happens, the model does not converge, resulting in less precise forecasts. We handled this by introducing an offset factor (Node F in figure 5.4) to make the model non-centred (Betancourt et al., 2013). The global parameters were set to follow a normal distribution and centred on 0 to enable parameter values to take on positive and negative values. Due to the hierarchical structure of the model parameters, global prior distribution usually serves as hyper-priors for the group-level parameters.

Parameter approximation for the HBM was sampled with Hamiltonian Monte Carlo (HMC) algorithm (Hoffman et al., 2014), an improved version of the classic Markov Chain Monte Carlo (MCMC) based on the notion of Hamiltonian dynamic. For this research, however, the No-U-Turn Sampler (NUTS) (Hoffman et al., 2014) version of HMC was used.

Below (figure 5.4) is a Directed Acyclic Graph (DAG) representing of the HBM used for this study.

From the HBM Directed Acyclic Graph (DAG) in figure 5.4:

- Node A is the global (Mean) regression intercept or (α_i) parameter assumed to be Gaussian;
- Node B global (Mean) regression coefficients for each of the lagged input variables (precipitation and soil moisture) or (β_i) parameters for the 18 lagged variables (6 lags each for VCI3M, P3M, SM3M);
- Node C represent Cholesky covariance matrix used as hyper-priors for the group level α_j and β_j parameters ;
- Nodes D and E are the Cholesky standard deviation and correlation from the matrix decomposition, respectively;
- Node F represent offset distribution (Gaussian) hyper-prior to make the model non-centred;
- Node G is the prior group-level parameters for α_j and β_j parameters for each vegetation AEZ within our selected counties (i.e. Five AEZs (β_j) within each of the 18 (β_i) parameters plus 1 (α_i));
- Node H represents the error term in the HBM regression;
- Node I is the likelihood function (equation 5.5) of the HBM regression and is based on ARDL*(p=6,q=6)* shown in equations 5.3 & 5.4;
- Node J is our lagged inputs datasets;
- Node K is the categorical values that maps the input data to their respective AEZs;
- Node L is the observed VCI3M values at an i lead time.

The Hierarchical BARDL model in this study was defined as:

$$D_{t+n} = \alpha_{j[i]} + \sum_{i=0}^{q} \beta_{j[d]} D_{t-q} + \sum_{i=0}^{p} \theta_{j[p]} P_{t-p} + \sum_{i=0}^{p} \delta_{j[s]} S_{t-p} + \epsilon_{t-p} \tag{5.3}$$

where D_{t+n} is the VCI3M at n weeks ahead, D_{t-q} represent the data for lags 0 to q of VCI3M (Dependent variable). P_{t-p}, S_{t-p} are the lags 0, to p, P3M, and SM3M respectively. $\alpha_{j[i]}$ are the global (i) and group level (j) regression intercept, $\beta_{j[d]}$, $\theta_{j[p]}$, and $\delta_{j[s]}$ are the regression coefficients for the lagged P3M, and SM3M input variables at the global (i) and group level (j). ϵ_{t-p} is the regression error term. Equation (5.3) can be simplified as 5.4 and re-written as a Bayesian likelihood function $P(X_t|\theta)$ in equation 5.5:

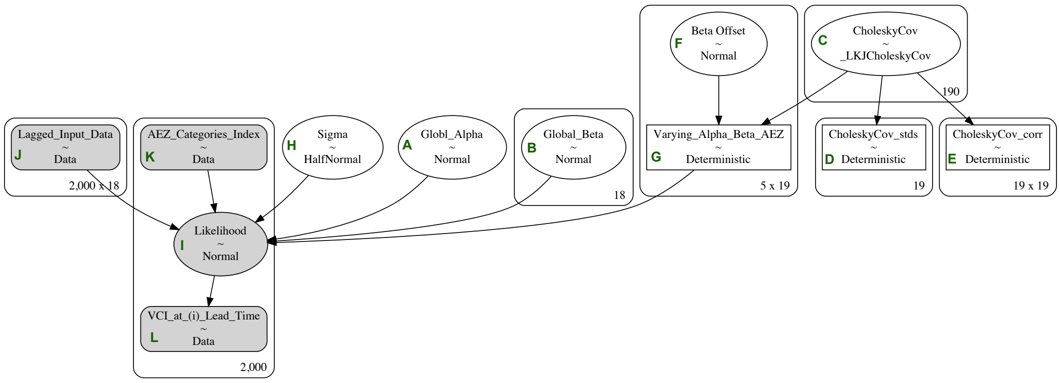

Figure 5.4: A Directed Acyclic Graph (DAG) schema representing the Hierarchical model based on varying Agro-Ecological Zones.

$$D_{t+n} = \alpha_{j[i]} + \sum_{i=0}^{i} \beta_{j[i]} X_{t-i} + \epsilon_{t-i} \tag{5.4}$$

where n is the lead time, $\beta_{j[i]}$ are the global and group level model parameters and X_{t-i} represent the lagged input variables in equation 5.3.

$$P(X_t|\alpha_{j[i]}, \beta_{j[i]}, \sigma) \sim N(\alpha_{j[i]} + \sum_{i=0}^{i} \beta_{j[i]} X_{t-i}, \sigma_{t-i}) \tag{5.5}$$

were $\alpha_{j[i]} \sim N(\mu_{\alpha_i}, \sigma^2_{\alpha_i}), \beta_{j[i]} \sim N\mu_{\beta_i}, \sigma^2_{\beta_i})$

and

$\sigma_{t-i} \sim HalfN(0, \sigma^2_d)$.

5.3.3 Forecasting and Model Evaluation

The forecast method used in this work was the direct multi-step forecast approach as described by Ben Taieb, Sorjamaa et al., 2010 and Ben Taieb and Hyndman, 2014.

$$D_{t+n} = \sum_{i=0}^{i} \nu_i X_{t-i} + \epsilon_{t-i} \tag{5.6}$$

where ν_i are the model parameters and X_{t-i} are the lagged inputs.

With this approach, separate models are fitted for every n lead time forecast. Meaning, for each n step forecast ahead (D_{t-n}), the observed VCI3M for the training dataset is shifted by n weeks ahead from lag0 X_{t-0} for all input variables.

After the parameter estimation via HMC sampling, the held-out dataset is passed to the fitted model (without the target variable) to produce forecast values for n weeks ahead. The held out observed values and mean values of our forecast distributions were used to compute the coefficient of determination (R^2) (Equation 5.8) and Root Mean Squared Error (RMSE), (Equation 5.7) metrics for forecast evaluation. The R^2 score quantifies the variation in the observed data that the model could explain, while the RMSE measures the average difference between the observed and forecast values.

$$RMSE = \sqrt{\frac{\sum_{i=n}^{n}(y_i - f_i)^2}{n}} \tag{5.7}$$

where the y_i are the observed data, f_i are the forecasts and n the total number of data points.

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}, \tag{5.8}$$

where the y_i are the observed data, and the f_i are the forecasts.

The forecast uncertainties were analysed with the Mean Prediction Interval Width (MPIW) and the Prediction Interval Coverage Probability (PICP) (Pang et al., 2018). The PICP computes the percentage of time the observed variable falls within a chosen prediction interval. The MPIW measures the mean distance between the upper (u) and lower (l) bound for a chosen prediction interval.

The MPIW was derived as follows:

$$MPIW_{t+n} = \frac{1}{N}\sum_{i=1}^{n} |u(D_i) - l(D_i)| \tag{5.9}$$

where $u(D_i)$ and $l(D_i)$ are the absolute upper and lower bounds values of the forecast distribution.

The PICP was derived as follows:

$$PICP_{t+n} = \frac{1}{N}\sum_{i=1}^{n} c_i \tag{5.10}$$

where N is the number of forecast samples, c_i is either 0 if the observed drought indicator at D_{t+n} value falls outside the prediction interval, and c_i is 1 if the observed value at is within the upper and lower bound of the forecast distribution.

Other forecast verification metrics used in this paper are the Receiver Operating Characteristic (ROC) curve (Wilks, 2006) curve and forecast probability Reliability Diagrams and Sharpness plots (Wilks, 2006; Jolliffe et al., 2012).

The ROC curve tells us the likelihood of a forecast being true (True Positive Rate (TPR)) for the given drought threshold and the probability of the forecast event being false (False Alarm

Rate (FAR)). In addition, the Area Under the Curve (AUC) was also computed to determine the propensity of our model to separate drought events for the set threshold (Bradley, 1997).

The Reliability Diagram allows us to assess the accuracy of the forecast probability predicted by our model. The probability of a drought event is computed using the full posterior distribution of our forecasts at a given drought threshold. The same threshold is used to convert observed data into binary events where 0 indicates a 'No Drought' and 1 indicate a 'Drought' event. The forecast probabilities and observed binaries are binned into probability intervals and used to plot the forecast reliability diagrams. The reliability of the forecast is assessed by the number of times an observed event agrees with a given forecast probability (Wilks, 2006). The sharpness plots, on the other hand, tells the frequency with which a drought event is predicted within a probability bin (WWRP, 2009).

5.4 Results

Our dynamic HBM for forecasting VCI3M were tested on datasets based on their AEZs and vegetation land covers. Two models were developed, A BARDL model based on a 'No-pooling' approach as a base model and an HBM based on the 'Partial-pooling' approach. The BARDL model was used to forecast VCI3M for the different AEZs, referred to as 'BARDL-AEZ', and different land covers, referred to as 'BARDL-LC'. The HBM, which was modelled with partially pooled AEZ data, is referred to as 'HBM-AEZ' and the model-based partially pooled land covers data will be referred to as 'HBM-LC'. The results shown in this section are a comparison of BARDL-AEZ to HBM-AEZ and BARDL-LC to HBM-LC.

The aim of modelling with HBM was to capture information on the variations within our partially pooled AEZ and land cover data. Figure 5.5 shows the time series plots of the observed and forecasted VCI3M from the BARDL and HBM at 4 weeks lead time for the different AEZs. From the plots, it is clear that the temporal variation differs for the various AEZs. The forecast values from the HBM-AEZ (figure 5.5 to the right) were more accurate than the BARDL-AEZ.

5.4.1 Model Performance for AEZ Based Models

The AEZ based models were used to forecast VCI3M for the Humid, Semi-Humid, Semi-Arid, Arid and Very-Arid zones. The R^2 scores and RMSE showed in figure 5.6 is for the Semi-Arid, Arid and Very-Arid zones since they were of most interest. The results for humid zones can be seen in figure 5.14. Both R^2 scores and RMSE in figure 5.6 (A & B) showed that the HBM-AEZ model performed better than the BARDL-AEZ model at all the lead times across all the AEZs. The R^2 scores were very high for forecasts in the very-arid zones, with HBM-AEZ having 0.97,

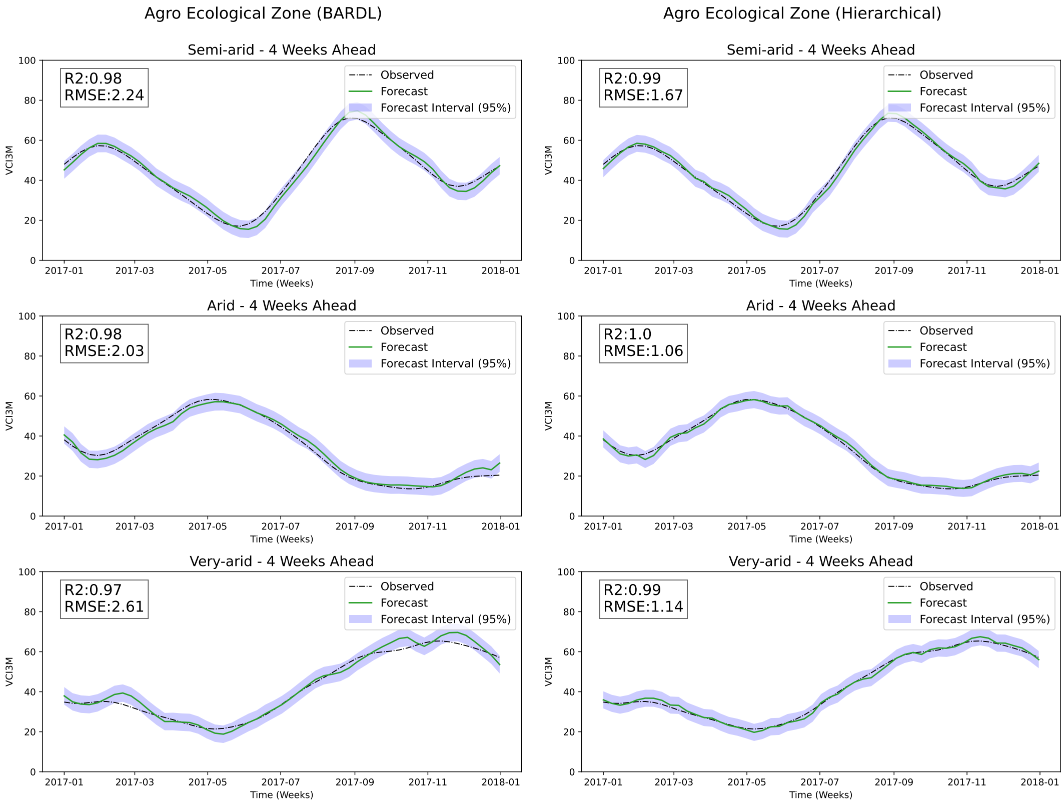

Figure 5.5: Time series Plots showing observed and forecast VCI3M at 4 week ahead in the semi-arid, arid and very-arid zones for the BARDL model (*left*) and HBM (*right*). The R^2 and RMSE metrics show that forecasts by the HBM are more accurate and have lower errors.

0.90, and 0.79 compared to 0.93, 0.86, and 0.75 for the BARDL-AEZ at 6, 8 and 10 weeks lead time, respectively. These scores indicate that the HBM was better at capturing the variability within the observed data than the BARDL model. In terms of the forecast errors (RMSE), the HBM-AEZ also performed better than the BARDL-AEZ model, with lower RMSE scores across the lead times.

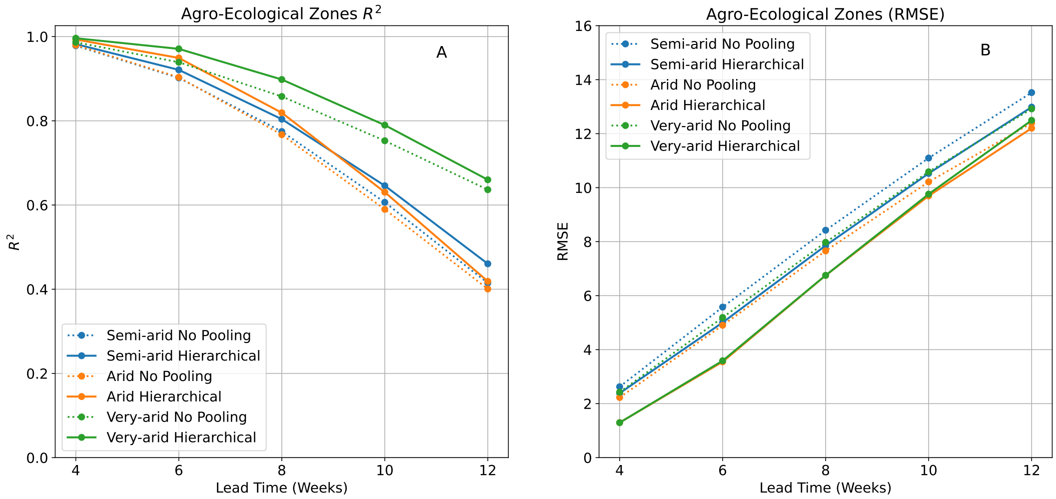

Figure 5.6: Plots showing R^2 Score (*left*) and RMSE (*right*) for BARDL-AEZ (Dotted) and HBM-AEZ (Solid) the VCI3M forecast over the different Agro-Ecological Zones

5.4.2 Model Performance for Land Cover Based Models

Figure 5.7 shows the performance metrics for the VCI3M forecast for the vegetation land covers. Overall, the HBM-LC performed better than the BARDL-LC except for the forest covers. (Where both models had almost identical R^2 scores across all lead times). The HBM-LC also performed well up to 10 weeks ahead for cropland with R^2 scores of 0.70 compared to 0.66 for the BARDL model. The R^2 score for forecasts over shrublands and grasslands remained between 0.90 and 0.70 up to 8 weeks ahead for the HBM-LC. The forecast errors from the RMSE plot (figure 5.7 B), showed a slightly different pattern. The forecast errors for all the land covers except for forest covers were lower for the HBM-LC. There was, however, no difference in R^2 and RMSE over forest cover, probably because the group-level effects did not differ significantly from the global effects.

5.4.3 Uncertainty Analysis

The forecast uncertainty of both forecasts models was analysed using the PICP and MPIW. The desired PICP value usually ranges between 0.90 to 0.99 Pang et al., 2018. The PICP indicates,

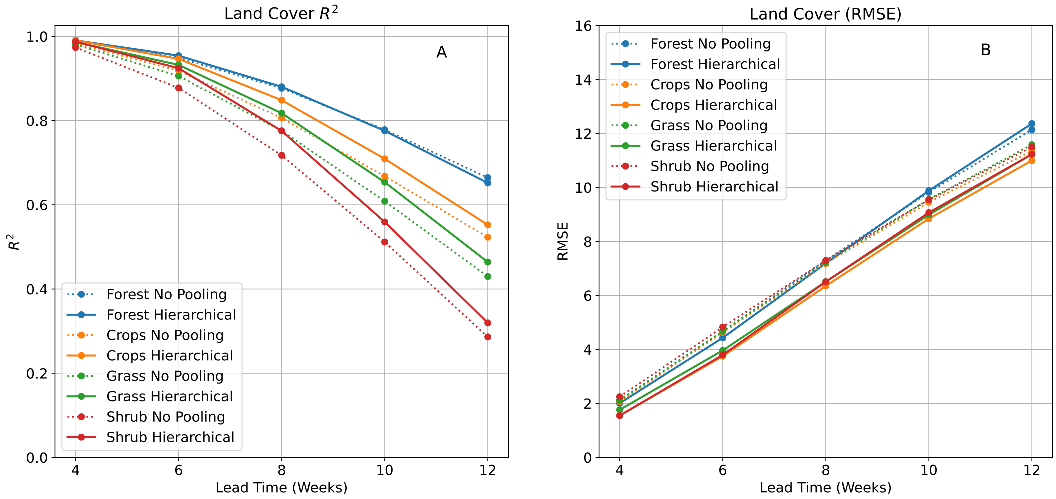

Figure 5.7: Plots showing R^2 (*left*) score and RMSE (*right*) for BARDL-LC (Dotted) and HBM-LC (Solid) the VCI3M forecast over the different vegetation land cover types.

the number of times observed values fall within our forecast interval. On the other hand, the MPIW values show forecast precision and are expected to remain very low. Figure 5.8 shows the time series plots of forecast and observed VCI3M for the arid zone in Baringo county. Each plot shows the 95% prediction interval along with the PICP and MPIW for 4 to 10 weeks lead time. The PICP values for both models indicate that observed values for all the lead times fall within a 95% credible interval of our forecast distributions over 90% of the time. The high PICP seen for the BARDL model from 8-Weeks was due to the wider forecast interval (error bars). A closer look at the MPIW values indicates that the HBM-AEZ forecasts are more precise than BARDL-AEZ, indicating that forecasts from HBM-AEZ have reduced uncertainties. A similar trend was seen for forecasts across all land covers. Overall, the MPIW metrics reiterate that forecasts by the HBM have lower errors than the BARDL. In addition, the partially pooled parameters also mean errors from the HBM is a better representation of the actual forecast error. Thus, even though PICP from 10 weeks ahead seems high for the BARDL model, they do not reflect the truth.

The mean PICP and MPIW for both AEZs and land covers over the selected counties are in tables 5.2 and 5.3.

5.4.4 Predicting Drought Event (ROC Curves)

Although our models produce accurate VCI3M values at the various lead times, our target users are also interested in whether or not a drought event alarm will be triggered at a defined threshold. Therefore, the skill of the forecast models at predicting drought events was assessed

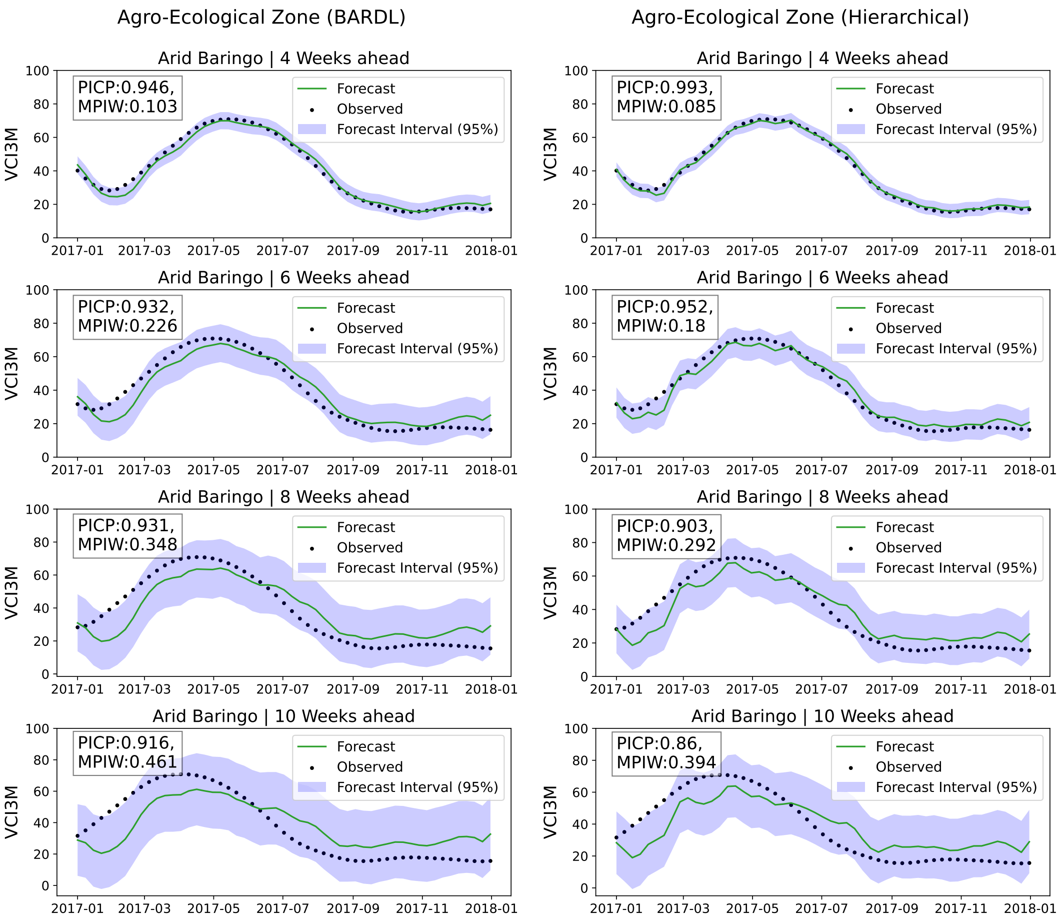

Figure 5.8: Plots showing forecast for Arid zones for 4 and 10 weeks lead times and their uncertainties (PICP & MPIW).

with the ROC curve with a threshold of VCI3M < 35%. VCI3M values below this threshold depict moderate to severe drought conditions (Klisch et al., 2016).

The ROC plots in figure (5.9) shows the TPR (Hit rate) and FAR (False Alarm) for the three arid zones. The dot on the curves indicates the VCI3M<35 threshold. Apart from the very-arid zones (fig 5.9 C), significant differences were seen between TPR and FAR for drought events predicted by the HBM-AEZ compared to the BARDL-AEZ (fig 5.9 A & B) at all lead times. The Hit rates for the HBM-AEZ were higher than the BARDL-AEZ and were mostly above 80% for drought events from 4 to 10 weeks ahead in the arid areas (fig 5.9 B) with false alarm rates between 1% to 18%. Drought events in semi-arid zones also had hit rates above 80% up until 8 weeks (fig 5.9 B). Both models performed very well at detecting moderate to severe drought events in the very-arid zones, as seen in figure (5.9 C), which was mostly because of the frequent occurrence of drought events in the very-arid zones.

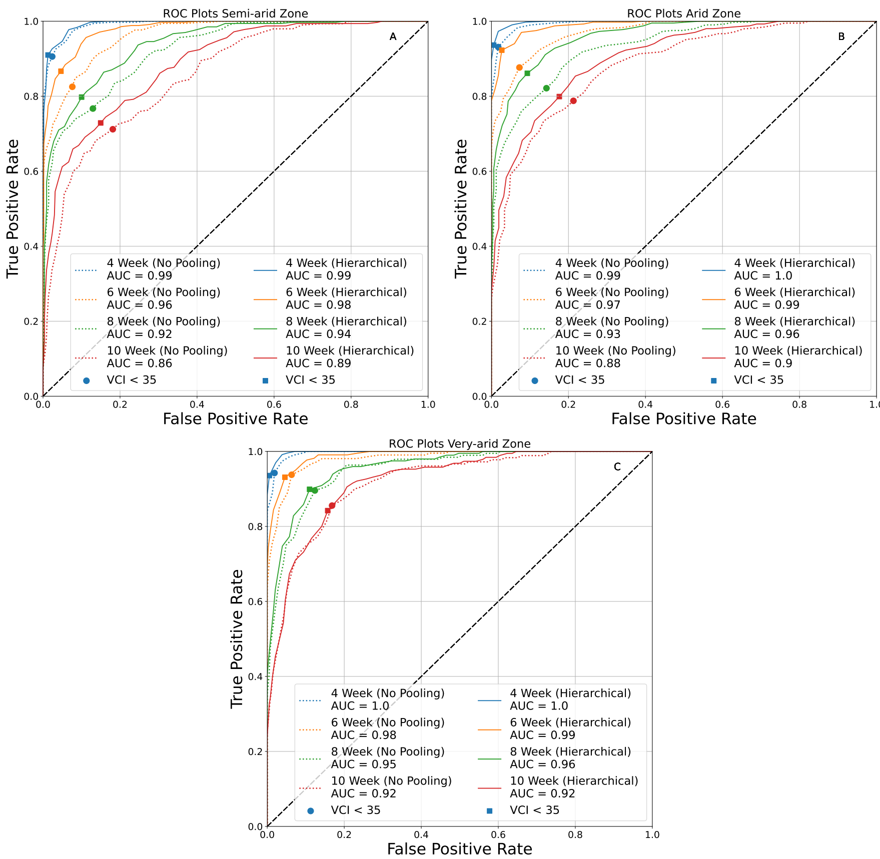

Figure 5.9: ROC plots generally showing higher Hit Rates for HBM in Semi Arid, Arid and Very Arid Zones

Figure 5.10 shows the ROC plots for the croplands, grasslands and shrubs for the BARDL-LC compared to HBM-LC. Overall, drought events predicted by the HBM-LC also had higher hit rates with lower false alarm rates than the BARDL-LC model. The hit rates for drought events over croplands remained above 80% up to 10 weeks ahead, with false alarm rates ranging between 1% to 16%. The TPR for grasslands and shrubs dropped quickly after 6 weeks. The TPRs were generally above 60% for all land covers at all lead times.

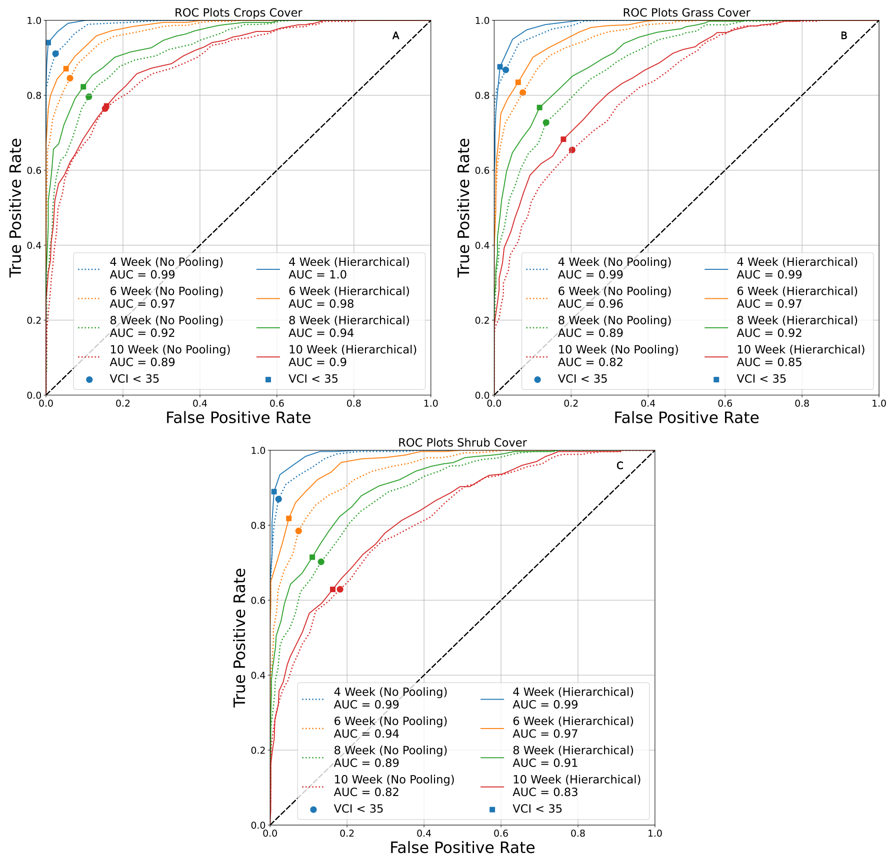

Figure 5.10: ROC Plots for Crops, Grass and Shrub Land Covers

5.4.5 Forecast Reliability

The reliability plots in figure 5.11 is a joint distribution of the binned forecast probabilities and relative frequency of the actual observed drought event (observed binaries = 1) for their respective probability bins. In a perfect system, the joint plots should lie on the diagonal line. The plots also show a histogram that depicts the model's sharpness. A perfect sharpness plot should have peaks at the extreme ends of the histogram. A peak close to the 0% probability bin indicates the frequency at which the model predicted a 'No Drought' event. Whereas a peak close to the 100% probability bin means otherwise. It is essential to state here that a forecast system is said to exhibit little or no sharpness when a sharpness peak is close to the long-term mean or climatology (Jolliffe et al., 2012).

The reliability diagrams for both BARDL-AEZ and HBM-AEZ (figure 5.11) showed some differences but were not very significant. The proximity of the reliability curves to diagonal, especially for the arid zones (figure 5.11, plots (A & C)) indicates the forecast probabilities from both models can be trusted for early warning and early action. From plots (A), we see that when the BARDL-AEZ model predicts drought event with a probability ranging between 80% to 100% at 4 to 6 weeks ahead, the forecast probability agrees with the observed frequency 90% to 99% of the time, which can also be seen in plots (C) for the HBM-AEZ model. For the very-arid zones forecast probabilities between 60% to about 80% (figure 5.11, plot(B & D)) corresponded with very high observed relative frequencies above 80%, a situation referred to as *'under forecasting'*. Under forecasting describes the situation where forecast probabilities do not adequately reflect observed events (Wilks, 2006). However, a closer look shows some subtle improvements with the HBM-AEZ, with a slight difference in the under forecasting effect from 4 to 6 week lead times. Regarding the sharpness of the models, a higher frequency of drought events was seen in the higher forecast probability bin for the HBM-AEZ 5.11, plot(C & D)) compared to the BARDL-AEZ 5.11, plot(A & B)) especially from 6 to 12 weeks in the arid zone. The reliability diagrams for croplands and grasslands for both BARDL-LC and HBM-LC models also showed similar patterns. Please see figure 5.15 in Appendix C.

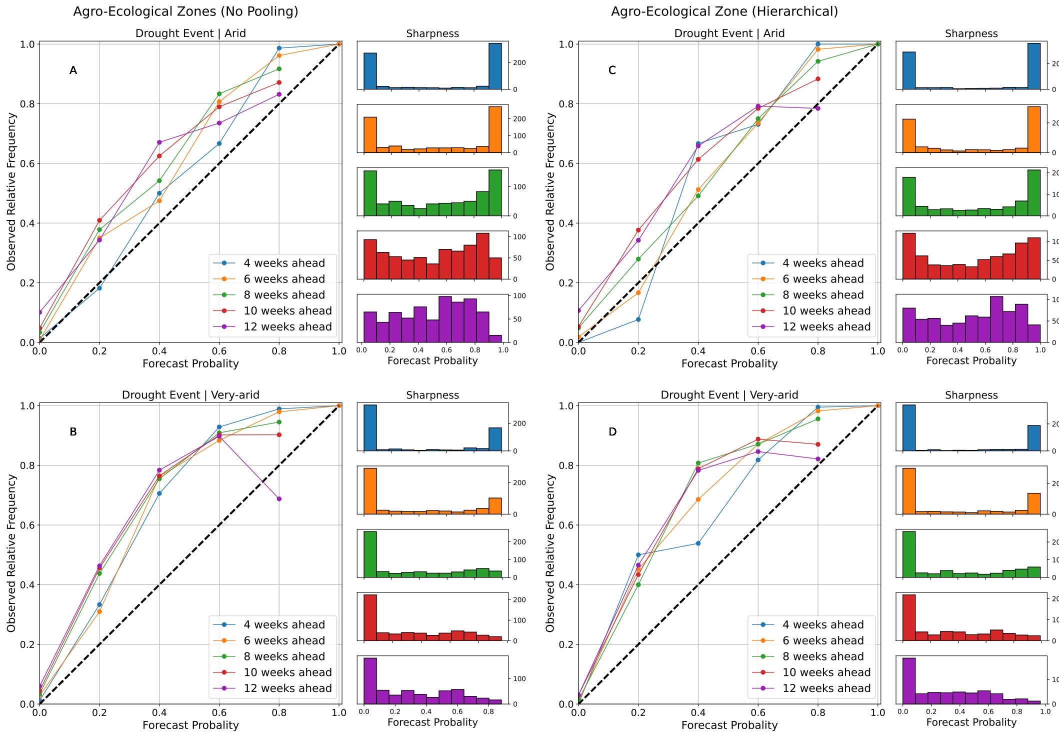

Figure 5.11: Reliability and sharpness plots showing a joint distribution of forecast probabilities and observed frequencies for various Arid and Very-Arid Agro-Ecological Zones for the different lead times

The skill of the models at predicting the onset and end of a drought period can be seen in figure 5.12. The figure shows a time series plot of observed and forecasted VCI3M at a 4-weeks lead time in a very-arid zone within Turkana county in Kenya for 2017. The plot also shows the forecast probability as a dot on vertical lines depicting the onset and end of a drought period. We can see from figure 5.12 (A) that at the start of a drought period where the observed VCI3M dropped below the threshold (VCI3M<35) line, the forecasted probability for the drought event predicted by the BARDL-AEZ was 9.4%. The low probability was because the forecasted VCI3M value model was higher than the observed value and threshold. However, the likelihood of a drought onset predicted by the HBM-AEZ in figure 5.12 (B) was 73%, prompting a trigger for early action. Towards the end of the drought period, the BARDL-AEZ model gave a high drought probability even though the drought duration had ended. Although these differences are not seen in all cases at the onset and end of a drought period, the few occurrences in some regions of interest emphasise that HBMs provide a better approach to forecasting VCI3M over a diverse region.

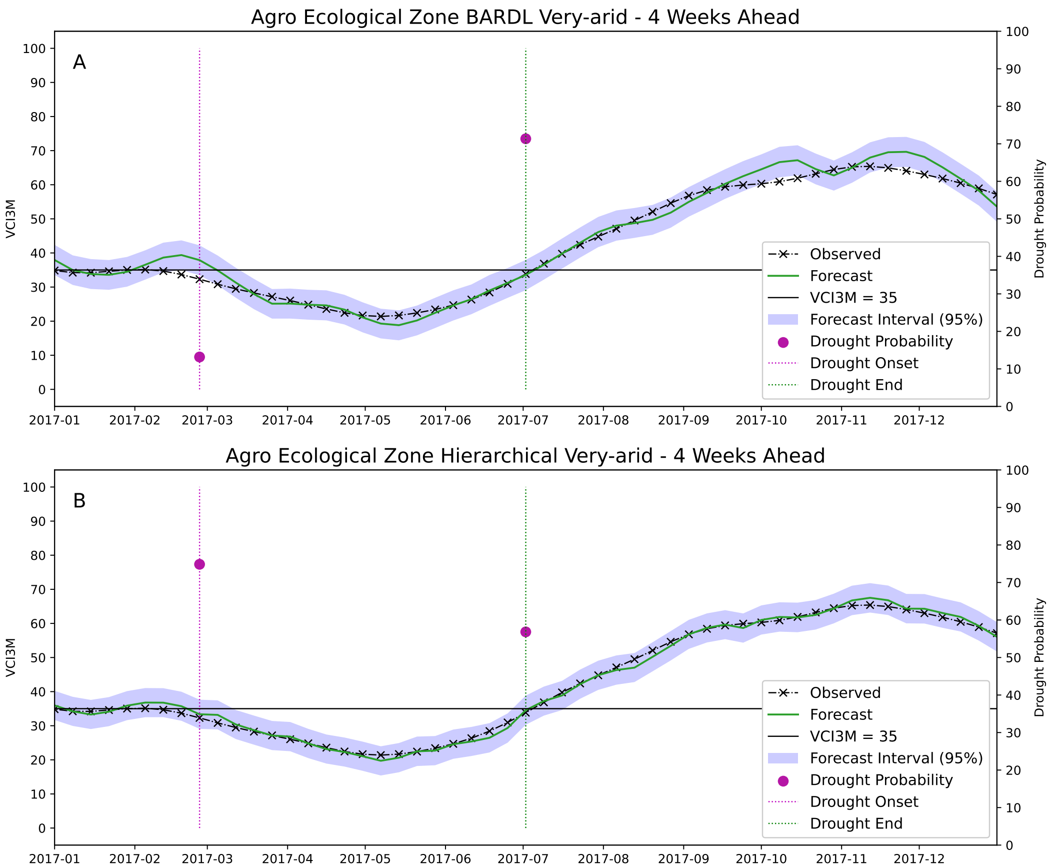

Figure 5.12: A time series plot showing the observed and forecasted VCI3M for the period of 2017. Forecast probabilities are indicated as points on the horizontal lines marking the onset and end of a drought periods

5.4.6 Test Transfer Learning

Although the data used for training and developing forecast models are usually sampled to represent a given area of interest, the goal in most cases is to have models that can scale up to produce forecasts over more expansive areas. The second objective of this study was to test the transfer learning capability of HBMs over other regions. The partially pooled data used for hierarchical parameter approximations were sampled from 6 counties. The trained models for the different lead times were then used to forecast VCI3M for the AEZs, and land covers over ten additional counties (shown with black boundaries in figure5.2), which were not part of the training sets. The comparison of their R^2 and RMSE metrics in figure 5.13 proved that both HBMs were able to forecast VCI3M over the non-trained counties accurately. For the AEZs, some significant differences were seen between the trained and non-trained counties in the semi-arid zones in terms of explained variances (R^2 score) (figure 5.13A). The case was different for forecast error in the same zone as seen in figure 5.13B. A significant difference was also seen for the forecast error over the very-arid area but not for the explained variations. The observation indicates that although the HBM model was able to capture the variations in the observed data there were instances in the very-zones where forecast values deviated significantly from the observed values. Performance over the different land covers, however, remained very close, especially for the RMSE (figure 5.13D) despite the gap seen for grassland in the R^2 score plots(figure 5.13C). These differences can be linked to the fact that although some non-trained counties may have similar AEZs or land covers, their climatic and vegetation phonology cycles are not similar. Aside from these observed differences, the HBMs could generalise and accurately forecast when given new unseen data.

5.5 Discussion

In this paper, we sought to improve the forecast accuracy of VCI3M over vast areas with varying AEZs, and land covers using an HBM. Compared to the non-hierarchical BARDL model, the HBM presented a more realistic approach for forecasting VCI3M in regions with different AEZs or land covers. The evaluation of the HBM based on R^2 metrics indicated that forecasts over the very-arid zones and forest cover areas showed higher accuracies at longer lead times. The high accuracy observed for the very-arid zones could be a consequence of the significant contribution from the lagged soil moisture to future VCI3M in addition to precipitation as seen in figures 5.16 and 5.17. For the forest areas, the observation could be because some dense forests show slight variation during drought periods.

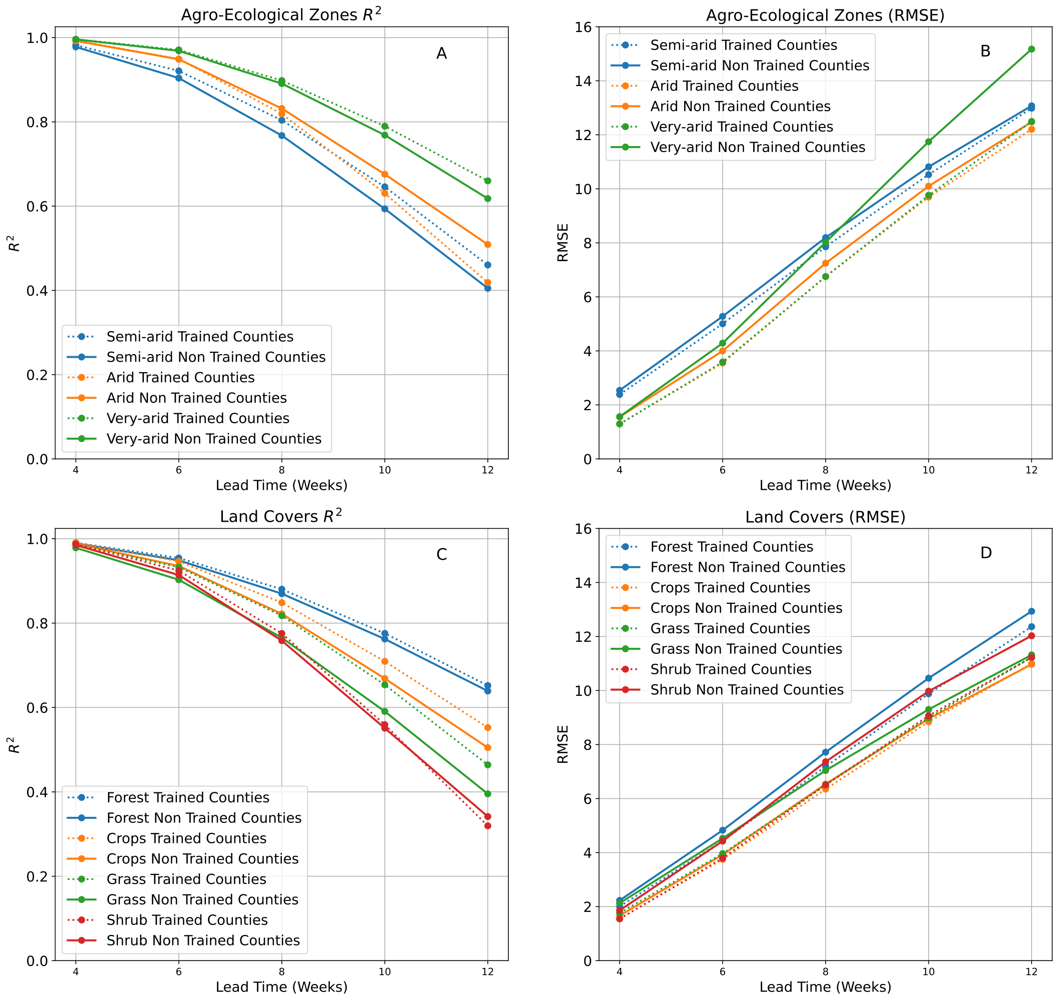

Figure 5.13: Plot showing R^2 score and RMSE for forecasts over counties not included in the training data used HBM (solid line) versus the counties included in the training data (dotted lines)

The strong relationship between lagged soil moisture VCIM over forest areas could be due to the frequent precipitation and high soil moisture retention in areas. On the other hand, the low contribution of soil moisture to forecasts in croplands, grasslands, and shrubs could be attributed to the low soil moisture levels over grass and shrub areas (James et al., 2003; Tyagi et al., 2013). For croplands, the low contribution of soil moisture could be due to several factors, including high temperature and soil type. However, in the very-arid areas, the high relative importance of soil moisture could be due to the rapid response of vegetation to sudden increases in soil moisture, especially after long periods of dryness.

Overall, results from the various skill assessments showed that forecasts with HBM were more precise with a low probability of false alarms rate for drought events than the BARDL model. The HBM was also able to effectively identify drought events in counties with diverse AEZs and some land covers.

Relating the overall forecast skill assessments from this work to previous works, the HBM showed an approximately one week increase in the forecast range compared to the results from the BARDL method used in Salakpi et al., 2021. On average, the HBM also exhibited an approximately 2-weeks increase in forecast range, compared to the auto-regression method used in Salakpi et al., 2021 and Barrett et al., 2020. Furthermore, using the HBM also enabled the simultaneous forecast of VCI3M for different AEZs and land covers which we could not do with the methods used in (Salakpi et al., 2021) and (Barrett et al., 2020). Finally, despite the improvement seen with the HBM, the BARDL models also proved to be useful at predicting drought events at the set threshold as demonstrated by (Salakpi et al., 2021).

Aside from the improvement in the forecast range, the HBM also had some added strengths. First of all, the hierarchical nature of the model parameters (see figure 5.3) enabled the incorporation of the varying (AEZs or land covers) effects of climate and other biophysical factors on vegetation conditions. Thus, modelling within the HBM framework made it possible to learn the within-sample parameters in addition to the global parameters and accurately forecast VCI3M values specific to the AEZs and land covers. Secondly, modelling within a Bayesian context means the model outputs probability distributions instead of point values. These distributions present a direct approach to quantifying forecast uncertainties. The probability distribution of forecasts also made it possible to derive forecast probabilities, which allowed us to quantify the likelihood of drought events in different locations. Finally, the HBM also made it possible to transfer trained models to similar datasets that were not part of the initial training data. Transferring the model also means even though the HBM model was calibrated on the data from Kenya, it can be scaled up to generate forecasts for wider regions without the need to re-calibrate.

The threat of agricultural drought to food security and global economies has pushed agencies like the USAID and FAO to develop early warning systems that continually monitor drought events. However, agricultural drought over vast and diverse ASAL regions poses a challenge to effective monitoring Boken et al., 2005; Sergio M Vicente-Serrano, 2006. Policy and decision-makers at these agencies, including Kenya's National Drought Management Authority (NDMA), our primary stakeholder, can incorporate the HBM demonstrated in this paper into their existing early warning systems to enhance their efforts. Aside from accounting for the different AEZs or land covers, the forecasted drought probabilities from the HBM will also enable intelligent decision making for drought relief agencies that practice the Forecast based Financing (FbF) (Coughlan de Perez et al., 2015) for drought early action.

The methods used in this paper also had a few limitations. A fundamental limitation was the timely availability of the ESA CCI Soil Moisture data. A setback that can affect the prospects of producing real-time forecasts. Parameter inference via HMC sampler also takes a long time to complete partly due to the complex nature of the HBM and the number of data points involved. However, this was not considered a significant limitation as it only occurs during the model training phase. Once the model converges, and sampling completes, the posterior predictive sampling or forecasting VCI3M takes seconds.

5.6 Conclusion and Future Work

In this paper, we presented a proof-of-concept that HBM can factor spatial differences into drought forecast. Using this approach also allowed us to understand the vegetation dynamics in Agro-climatic areas and regions with diverse vegetation covers. For instance, we saw an approximately one week gain in forecast range for vegetation conditions in very-arid as well as forests (Tree cover) and cropping areas. Furthermore, we have shown that soil moisture contributes more when forecasting VCI3M over very-arid areas and forest covers. We also showed that HBM trained with data in one area could be transferred to other similar datasets in other regions. Future research work should consider more complex HBMs that takes into account variations for different land cover types within the various Agro-Ecological zones and the seasonal differences.

Author Contribution

EES lead author, data preprocessing, modelling (Coding) & running BARDL and HBM methods; JMM data acquisition, preprocessing, cartography and feedback; AB code for smoothing time series data; SO, PR, & PH conceptualised the initial idea and provided supervision and feedback; The final manuscript was edited and reviewed by all authors.

Competing Interests

All authors of the paper declare no known competing interests (financial, personal relationships) that could have influenced this study.

Acknowledgements

The work was funded by the UK Newton Fund's Development in Africa with Radio Astronomy (DARA) Big Data project delivered via STFC with grant number ST/R001898/1 and by the Science for Humanitarian Emergencies and Resilience (SHEAR) consortium project 'Towards Forecast-based Preparedness Action' (ForPAc, www.forpac.org), Grant Number NE/P000673/1, funded by the UK Natural Environment Research Council (NERC), the Economic and Social Research Council (ESRC), and the UK Department for International Development (DfID).

Link to Data and Code repository

https://github.com/edd3x/Hierarchical-Bayesian-ARDL.git

References

Adede, Chrisgone, Robert Oboko, Peter Waiganjo Wagacha and Clement Atzberger (May 2019). "A Mixed Model Approach to Vegetation Condition Prediction Using Artificial Neural Networks (ANN): Case of Kenya's Operational Drought Monitoring". In: *Remote Sensing* 11.9, p. 1099. ISSN: 2072-4292. DOI: 10.3390/rs11091099. URL: https://www.mdpi.com/2072-4292/11/9/1099.

Asaad, Al-ahmadgaid B and Joselito C Magadia (2019). "Stochastic Gradient Hamiltonian Monte Carlo on Bayesian Time Series Modeling". In: *14th National Convention on Statistics Crowne*.

Ayugi, Brian Odhiambo, Wang Wen and Daisy Chepkemoi (2016). "Analysis of Spatial and Temporal Patterns of Rainfall Variations over Kenya". In: 6.11. ISSN: 2225-0948. URL: www.iiste.org.

Barrett, Adam B, Steven Duivenvoorden, Edward E Salakpi, James M Muthoka, John Mwangi, Seb Oliver and Pedram Rowhani (2020). "Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya". In: *Remote Sensing of Environment* 248, p. 111886.

Ben Taieb, Souhaib and Rob J. Hyndman (2014). "Recursive and direct multi-step forecasting: the best of both worlds". In: *International Journal of Forecasting* September.

Ben Taieb, Souhaib, Antti Sorjamaa and Gianluca Bontempi (June 2010). "Multiple-output modeling for multi-step-ahead time series forecasting". In: *Neurocomputing* 73.10-12, pp. 1950–1957. ISSN: 09252312. DOI: 10.1016/j.neucom.2009.11.030.

Betancourt, M. J. and Mark Girolami (Dec. 2013). "Hamiltonian Monte Carlo for Hierarchical Models". In: *Current Trends in Bayesian Methodology with Applications*, pp. 79–101. arXiv: 1312.0906. URL: http://arxiv.org/abs/1312.0906.

Bishop, Christopher M (2006). "Pattern recognition". In: *Machine learning* 128.9.

Boken, Vijendra K., Arthur P. Cracknell and Ronald L. Heathcote (May 2005). *Monitoring and Predicting Agricultural Drought*. Oxford University Press. ISBN: 9780195162349. DOI: 10.1093/oso/9780195162349.001.0001. URL: https://oxford.universitypressscholarship.com/view/10.1093/oso/9780195162349.001.0001/isbn-9780195162349.

Bradley, Andrew P. (1997). "The use of the area under the ROC curve in the evaluation of machine learning algorithms". In: *Pattern Recognition* 30.7, pp. 1145–1159. ISSN: 0031-3203. DOI: https://doi.org/10.1016/S0031-3203(96)00142-2. URL: https://www.sciencedirect.com/science/article/pii/S0031320396001422.

Coughlan de Perez, E., B. van den Hurk, M. K. van Aalst, B. Jongman, T. Klose and P. Suarez (2015). "Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts". In: *Natural Hazards and Earth System Sciences* 15.4, pp. 895–904. DOI: 10.5194/nhess-15-895-2015. URL: https://nhess.copernicus.org/articles/15/895/2015/.

Da Silva, Ivan Nunes, Danilo Hernane Spatti, Rogerio Andrade Flauzino, Luisa Helena Bartocci Liboni and Silas Franco dos Reis Alves (2017). "Artificial neural network architectures and training processes". In: *Artificial neural networks*. Springer, pp. 21–28.

Deleersnyder, Roxanna (2018). *Pastoralism in East Africa: challenges and solutions — Glo.be.* URL: https://www.glo-be.be/index.php/en/articles/pastoralism-east-africa-challenges-and-solutions (visited on 09/05/2021).

Eilers, Paul H. C. (2003). "A Perfect Smoother". In: *Analytical Chemistry* 75.14. PMID: 14570219, pp. 3631–3636. DOI: 10.1021/ac034173t. eprint: https://doi.org/10.1021/ac034173t. URL: https://doi.org/10.1021/ac034173t.

FEWSNET (2021). *About Us — Famine Early Warning Systems Network.* URL: https://fews.net/about-us (visited on 09/05/2021).

Fischer, G., H.T. van Velthuizen and F.O. Nachtergaele (2000). "Global Agro-Ecological Zones Assessment: Methodology and Results". In.

Funk, Chris et al. (June 2019). "Recognizing the Famine Early Warning Systems Network: Over 30 Years of Drought Early Warning Science Advances and Partnerships Promoting Global Food Security". In: *Bulletin of the American Meteorological Society* 100.6, pp. 1011–1027. ISSN: 0003-0007. DOI: 10.1175/BAMS-D-17-0233.1. URL: https://journals.ametsoc.org/view/journals/bams/100/6/bams-d-17-0233.1.xml.

Gebremeskel, Gebremedhin, Qiuhong Tang, Siao Sun, Zhongwei Huang, Xuejun Zhang and Xingcai Liu (June 2019). *Droughts in East Africa: Causes, impacts and resilience.* DOI: 10.1016/j.earscirev.2019.04.015.

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari and Donald B Rubin (2013). *Bayesian data analysis.* CRC press.

Gelman, Andrew and Jennifer Hill (2006). *Data analysis using regression and multilevel/hierarchical models.* Cambridge university press.

George, Dileep and Jeff Hawkins (2005). "A hierarchical Bayesian model of invariant pattern recognition in the visual cortex". In: *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* Vol. 3. IEEE, pp. 1812–1817.

Gujarati, D.N. (2003). *Basic Econometrics.* Economic series. McGraw Hill. ISBN: 9780072335422. URL: https://books.google.co.uk/books?id=byu7AAAAIAAJ.

Heim, Richard (2002). "A Review of Twentieth- Century Drought Indices Used in the United States". In: August, pp. 1149–1165.

Hoffman, Matthew D. and Andrew Gelman (Jan. 2014). "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo". In: *J. Mach. Learn. Res.* 15.1, pp. 1593–1623. ISSN: 1532-4435.

IISD (2018). *WMO Checklist Helps Enhance Early Warning Systems — News — SDG Knowledge Hub — IISD*. URL: https://sdg.iisd.org/news/wmo-checklist-helps-enhance-early-warning-systems/ (visited on 03/09/2021).

James, Sarah E., Meelis Pärtel, Scott D. Wilson and Duane A. Peltzer (Apr. 2003). "Temporal heterogeneity of soil moisture in grassland and forest". In: *Journal of Ecology* 91.2, pp. 234–239. ISSN: 1365-2745. DOI: 10.1046/J.1365-2745.2003.00758.X. URL: https://besjournals.onlinelibrary.wiley.com/doi/full/10.1046/j.1365-2745.2003.00758.x%20https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2745.2003.00758.x%20https://besjournals.onlinelibrary.wiley.com/doi/10.1046/j.1365-2745.2003.00758.x.

Jolliffe, Ian T and David B Stephenson (2012). *Forecast verification: a practitioner's guide in atmospheric science*. John Wiley & Sons.

Klisch, Anja and Clement Atzberger (2016). "Operational drought monitoring in Kenya using MODIS NDVI time series". In: *Remote Sensing* 8.4. ISSN: 20724292. DOI: 10.3390/rs8040267.

Kogan, F. N. (1995). "Application of vegetation index and brightness temperature for drought detection". In: *Advances in Space Research* 15.11, pp. 91–100. ISSN: 02731177. DOI: 10.1016/0273-1177(95)00079-T.

Lambert, B. (2018). *A Student's Guide to Bayesian Statistics*. SAGE Publications. ISBN: 9781526418289. URL: https://books.google.co.uk/books?id=ZvBUDwAAQBAJ.

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press. ISBN: 9781482253481. URL: https://books.google.co.uk/books?id=1yhFDwAAQBAJ.

McElreath, Richard (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

Nay, John, Emily Burchfield and Jonathan Gilligan (2018). "A machine-learning approach to forecasting remotely sensed vegetation health". In: *International Journal of Remote Sensing* 39.6, pp. 1800–1816. ISSN: 0143-1161. DOI: 10.1080/01431161.2017.1410296. URL: https://www.tandfonline.com/doi/full/10.1080/01431161.2017.1410296.

Nicolai-Shaw, Nadine, Jakob Zscheischler, Martin Hirschi, Lukas Gudmundsson and Sonia I. Seneviratne (2017). "A drought event composite analysis using satellite remote-sensing based soil moisture". In: *Remote Sensing of Environment* 203, pp. 216–225. ISSN: 00344257. DOI: 10.1016/j.rse.2017.06.014. URL: https://doi.org/10.1016/j.rse.2017.06.014.

Pang, Jingyue, Datong Liu, Yu Peng and Xiyuan Peng (2018). "Optimize the coverage probability of prediction interval for anomaly detection of sensor-based monitoring series". In: *Sensors (Switzerland)* 18.4. ISSN: 14248220. DOI: 10.3390/s18040967.

R. Ravines, Romy, Alexandra M. Schmidt and Helio S. Migon (Mar. 2006). "Revisiting distributed lag models through a Bayesian perspective". In: *Applied Stochastic Models in Business and Industry* 22.2, pp. 193–210. ISSN: 1524-1904. DOI: 10.1002/asmb.628. URL: http://doi.wiley.com/10.1002/asmb.628.

Rippa, Shmuel (1999). "An algorithm for selecting a good value for the parameter c in radial basis function interpolation". In: *Advances in Computational Mathematics* 11.2, pp. 193–210. DOI: 10.1023/A:1018975909870. URL: https://doi.org/10.1023/A:1018975909870.

Rosenstein, Michael T, Zvika Marx, Leslie Pack Kaelbling and Thomas G Dietterich (2005). "To transfer or not to transfer". In: *NIPS 2005 workshop on transfer learning*. Vol. 898, pp. 1–4.

Salakpi, E. E., P. D. Hurley, J. M. Muthoka, A. B. Barrett, A. Bowell, S. Oliver and P. Rowhani (2021). "Forecasting Vegetation Condition with a Bayesian Auto-regressive Distributed Lags (BARDL) Model". In: *Natural Hazards and Earth System Sciences Discussions* 2021, pp. 1–31. DOI: 10.5194/nhess-2021-223. URL: https://nhess.copernicus.org/preprints/nhess-2021-223/.

Sánchez, Carles and Gary M Bernstein (2019). "Redshift inference from the combination of galaxy colours and clustering in a hierarchical Bayesian model". In: *Monthly Notices of the Royal Astronomical Society* 483.2, pp. 2801–2813.

Senf, Cornelius, Dirk Pflugmacher, Marco Heurich and Tobias Krueger (June 2017). "A Bayesian hierarchical model for estimating spatial and temporal variation in vegetation phenology from Landsat time series". In: *Remote Sensing of Environment* 194, pp. 155–160. ISSN: 00344257. DOI: 10.1016/j.rse.2017.03.020.

Shao, Yang and Ross S. Lunetta (2012). "Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 70, pp. 78–87. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2012.04.001. arXiv: arXiv:1011.1669v3. URL: http://dx.doi.org/10.1016/j.isprsjprs.2012.04.001.

Sombroek, Wim G, HMH Braun, BJA Van der Pouw et al. (1982). *Exploratory soil map and agro-climatic zone map of Kenya, 1980. Scale 1: 1,000,000.* Kenya Soil Survey.

Stan Development Team (2018). *The Stan Core Library*. Version 2.18.0. URL: http://mc-stan.org/%203.

Storz, Jay F and Mark A Beaumont (2002). "Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model". In: *Evolution* 56.1, pp. 154–166.

Tian, Siyuan, Albert I. J. M. Van Dijk, Paul Tregoning and Luigi J. Renzullo (Dec. 2019). "Forecasting dryland vegetation condition months in advance through satellite data assimilation". In: *Nature Communications* 10.1, p. 469. ISSN: 2041-1723. DOI: 10.1038/s41467-019-08403-x. URL: http://www.nature.com/articles/s41467-019-08403-x.

Tyagi, J. V., Nuzhat Qazi, S. P. Rai and M. P. Singh (2013). "Analysis of soil moisture variation by forest cover structure in lower western Himalayas, India". In: *Journal of Forestry Research* 24.2, pp. 317–324. ISSN: 1007662X. DOI: 10.1007/s11676-013-0355-8.

UN (2018). *Early Warning Systems — United Nations*. URL: https://www.un.org/en/climatechange/climate-solutions/early-warning-systems (visited on 03/09/2021).

Vatter, Juliane (2019). *DROUGHT RISK The Global Thirst for Water in the Era of Climate Crisis*. Tech. rep. World Wildlife Fund (WWF) Germany. URL: www.studioazola.com.

Vicente-Serrano, Sergio M (Feb. 2006). "Differences in Spatial Patterns of Drought on Different Time Scales: An Analysis of the Iberian Peninsula". In: *Water Resources Management* 20.1, pp. 37–60. ISSN: 0920-4741. DOI: 10.1007/s11269-006-2974-8. URL: http://link.springer.com/10.1007/s11269-006-2974-8.

— (2007). "Evaluating the impact of drought using remote sensing in a Mediterranean, Semi-arid Region". In: *Natural Hazards* 40.1, pp. 173–208. ISSN: 0921030X. DOI: 10.1007/s11069-006-0009-7.

Wilks, DS (2006). "Statistical methods in the atmospheric sciences". In.

WWRP (2009). *World Weather Research Programme (WWRP), Forecast Verification - Methods and FAQ*. URL: https://www.cawcr.gov.au/projects/verification/verif%5C_web%5C_page.html (visited on 22/06/2021).

Yang, Yongke, Pengfeng Xiao, Xuezhi Feng and Haixing Li (2017). "Accuracy assessment of seven global land cover datasets over China". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 125.April 2018, pp. 156–173. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2017.01.016. URL: http://dx.doi.org/10.1016/j.isprsjprs.2017.01.016.

Yang, Zhilin, Ruslan Salakhutdinov and William W Cohen (2017). "Transfer learning for sequence tagging with hierarchical recurrent networks". In: *arXiv preprint arXiv:1703.06345*.

5.7 Appendix

5.8 Forecast Metrics Semi-Humid and Humid Zones

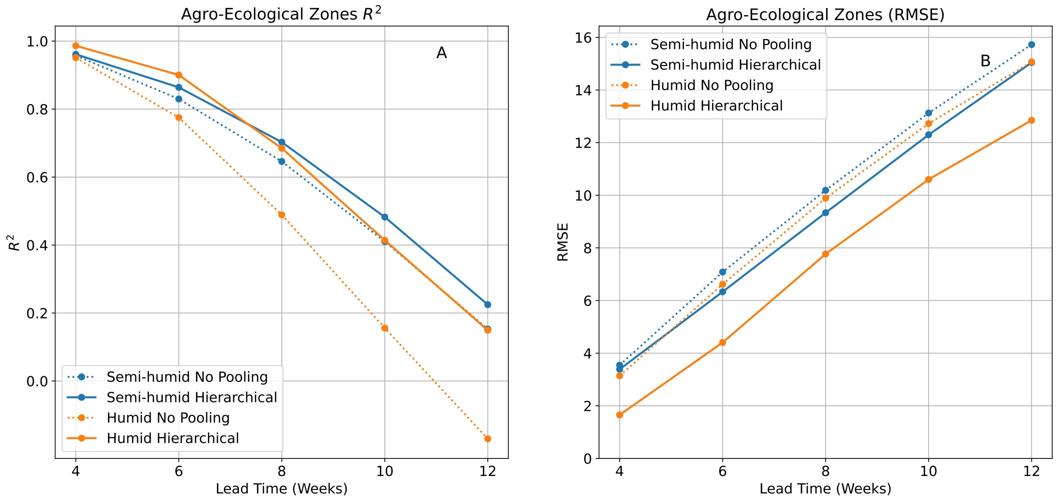

Figure 5.14: Plots showing R^2 Score and RMSE for BARDL-AEZ (Dotted) and HBM-AEZ (Solid) the VCI3M forecast over the different humid zones

5.9 PICP and MPIW for Land Covers and Agro-Ecological Zones

Table 5.2: Table showing a PICP and MPIW (In Parenthesis) for the various Agro-Ecological Zones

Models	AEZ	4	6	8	10	12
BARDL	Humid	0.88 (0.09)	0.88 (0.21)	0.9 (0.32)	0.91 (0.42)	0.91 (0.52)
	Semi-Humid	0.87 (0.1)	0.88 (0.22)	0.88 (0.33)	0.89 (0.43)	0.9 (0.52)
	Semi-Arid	0.97 (0.1)	0.96 (0.22)	0.94 (0.32)	0.95 (0.42)	0.95 (0.5)
	Arid	0.98 (0.11)	0.98 (0.23)	0.97 (0.33)	0.97 (0.43)	0.96 (0.51)
	Very-Arid	0.96 (0.11)	0.95 (0.23)	0.94 (0.33)	0.94 (0.42)	0.94 (0.5)
Hierarchical	Humid	0.97 (0.09)	0.95 (0.18)	0.94 (0.29)	0.94 (0.4)	0.93 (0.48)
	Semi-Humid	0.81 (0.09)	0.84 (0.18)	0.88 (0.29)	0.88 (0.39)	0.88 (0.48)
	Semi-Arid	0.94 (0.09)	0.94 (0.18)	0.95 (0.29)	0.95 (0.39)	0.95 (0.48)
	Arid	1.0 (0.09)	0.98 (0.18)	0.96 (0.29)	0.95 (0.39)	0.94 (0.48)
	Very-Arid	1.0 (0.09)	0.97 (0.18)	0.94 (0.29)	0.93 (0.39)	0.93 (0.48)

Table 5.3: Table showing a PICP and MPIW (In Parenthesis) for the various vegetation land covers

Model	Land Covers	4	6	8	10	12
BARDL	Forest	0.97 (0.09)	0.96 (0.19)	0.95 (0.29)	0.95 (0.38)	0.94 (0.46)
	Crops	0.97 (0.09)	0.95 (0.19)	0.94 (0.29)	0.95 (0.38)	0.95 (0.46)
	Grass	0.97 (0.09)	0.96 (0.19)	0.96 (0.29)	0.97 (0.38)	0.96 (0.46)
	Shrub	0.96 (0.09)	0.96 (0.19)	0.96 (0.29)	0.96 (0.38)	0.96 (0.46)
Hierarchical	Forest	0.94 (0.08)	0.93 (0.17)	0.94 (0.27)	0.94 (0.37)	0.94 (0.46)
	Crops	0.99 (0.08)	0.97 (0.17)	0.96 (0.27)	0.95 (0.37)	0.95 (0.46)
	Grass	0.98 (0.08)	0.97 (0.17)	0.96 (0.27)	0.96 (0.37)	0.95 (0.46)
	Shrub	0.98 (0.08)	0.98 (0.17)	0.98 (0.27)	0.97 (0.37)	0.96 (0.46)

5.10 Reliability Diagram for Crop and Grass Covers

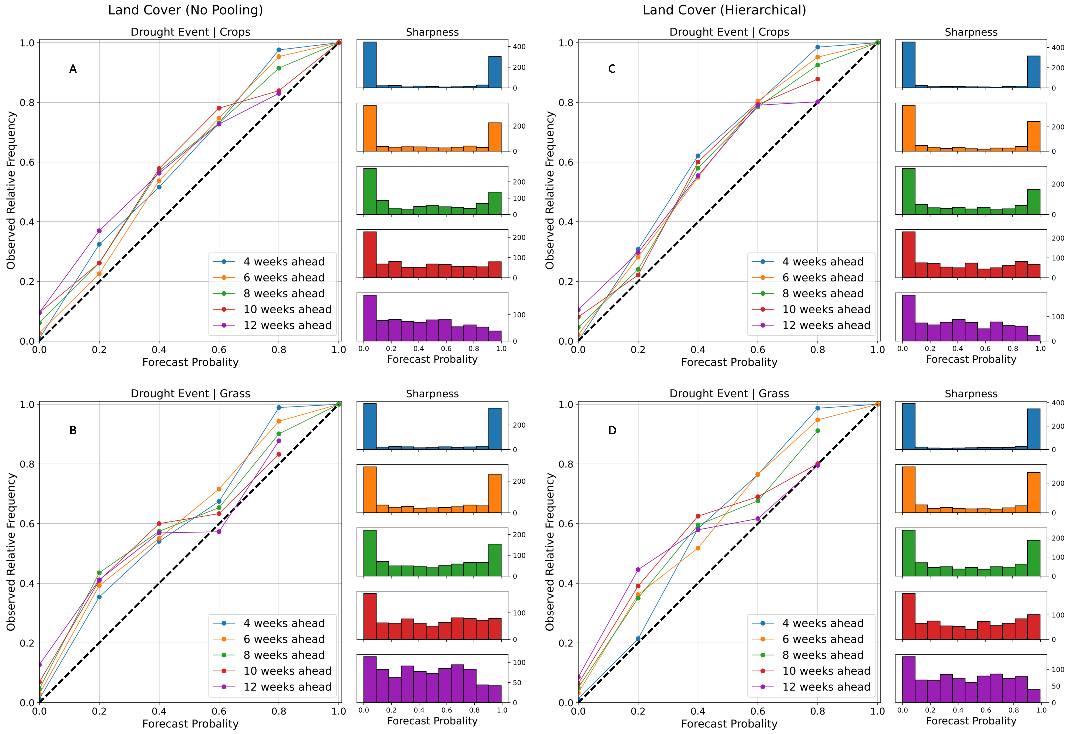

Figure 5.15: Reliability and sharpness plots showing a joint distribution of forecast probabilities and observed frequencies for various Agro-Ecological Zones and Land Cover for different lead times

5.11 Percentage Relative Importance

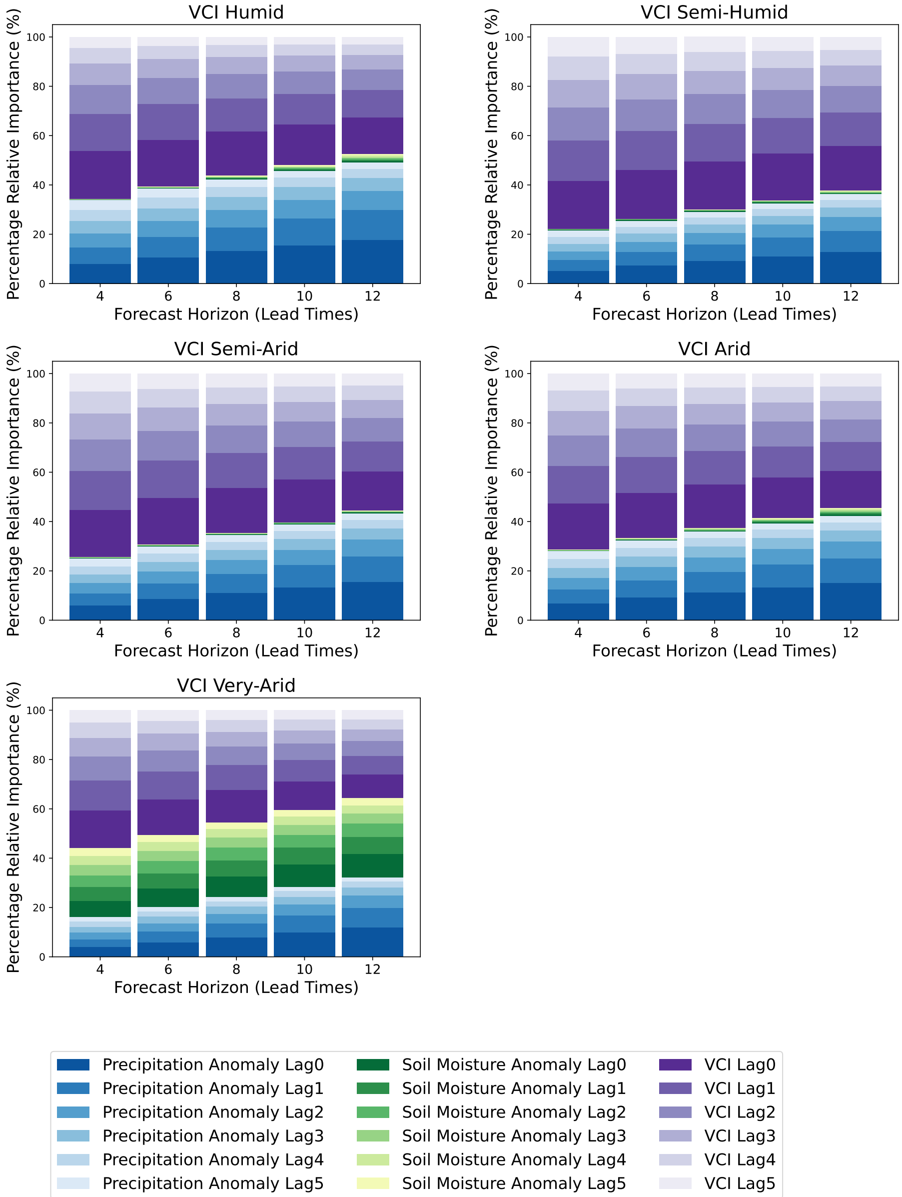

Figure 5.16: Plots showing the relative importance of the lagged input variables (VCI3M, P3M, SM3M) and VCI3M at 4 to 12 lead times the different Agro-Ecological zones

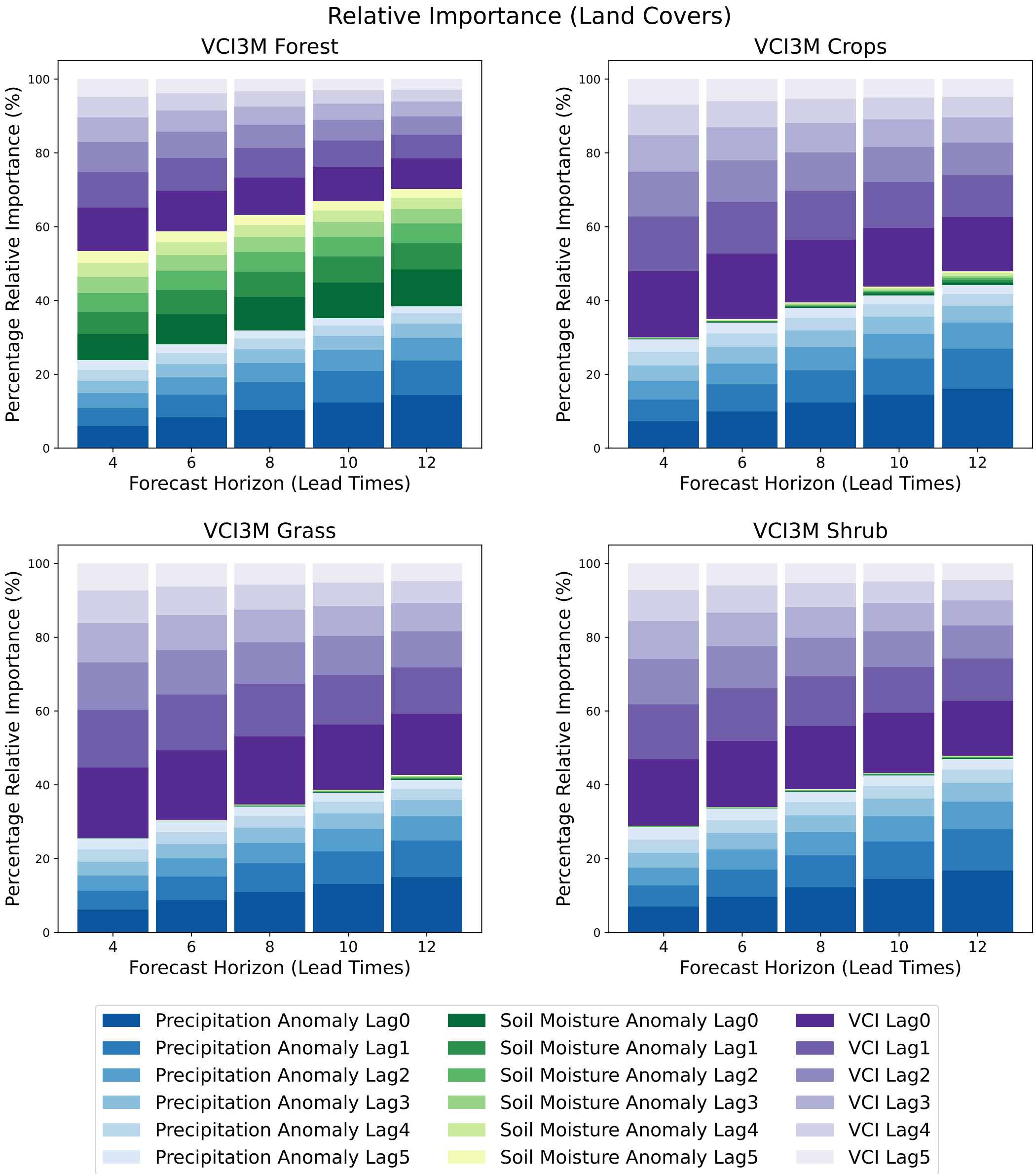

Figure 5.17: Plots showing the relative importance of the lagged input variables (VCI3M, P3M, SM3M) and VCI3M at 4 to 12 lead times the different vegetation land covers

Chapter 6

Discussion

In this PhD thesis, we explored various probabilistic machine learning approaches for forecasting agricultural drought for the benefit of pastoralist communities in Kenya. The overall goal was to develop Early Warning Systems (EWS) that enable the early response to severe drought conditions that have, over the last decade, negatively affected the country's economy and livelihoods. To this end, we focused on agricultural drought, given its direct socio-economic impact. The indicator used to measure and analyse drought severity in this context was the Vegetation Condition Index (VCI) (Kogan, 1995). The VCI, which is based on the NDVI, allowed us to study and understand the vegetation dynamics and forecast the onset of drought using linear Auto-Regression (AR), Gaussian Processes (GP), Bayesian Auto-Regressive Distributed Lags (BARDL) and A Dynamics Hierarchical Bayesian Model (HBM).

In chapter 3 the GP model takes advantage of the temporal correlation of a VCI3M time series. It captures the covariance between VCI3M at any two consecutive time points and generates forecasts via extrapolation. The R^2-score showed the model was very skilful at forecasting VCI3M at 2 to 6 weeks lead time. The GP also proved to be efficient at predicting drought events (i.e. VCI3M<35) with very low false alarm rates across all the lead times. However, GP performed very poorly when anticipating drought onset or transitions between VCI3M>35 and VCI3M<35. An interesting observation was its tendency to forecast the long term mean of VCI3M instead of following the trend. Also, the AR method modelled VCI3M as linear regression with a 3-week lag order showed very accurate forecasts. The AR method performed slightly better than GP, especially regarding the hit rate for VCI3M<35 events. The AR also exhibited a good skill for determining the onset of vegetation stress.

The skills exhibited by both models were independent of the livelihood zones within our regions of interest. However, both models relied heavily on the intrinsic temporal correlation of VCI3M only, which led to very accurate short-term forecasts. The accuracy of the GP model, coupled with its lower false alarm rate, made it of most interest to our stakeholders in Kenya. One of such stakeholders is the National Drought Management Authority (NDMA), to which we proposed adding a new drought classification, 'Early Alert', to their monthly report for drought preparedness (Barrett et al., 2020). The NDMA operates a forecast-based financing (FbF) mechanism where drought-prone communities receive anticipatory cash support when drought indicator drops below a threshold (UNDRR, 2021; Coughlan de Perez et al., 2015). The FbF approach has proven to be a cost-effective approach to address challenges associated with the impact of agricultural drought (Guimarães Nobre, 2019; UNDRR, 2021). However, FbF for early action requires ample time for drought preparedness which the short-term forecast could not address, calling for the need research into long-term forecast models.

Forecasting VCI3M at longer ranges required additional information from other factors that drive agriculture drought. The objective of the paper in chapter 4 was to develop a Bayesian Auto-Regressive Distributed Lag (BARDL) models that forecast longer-range VCI3M with information from additional factors like precipitation and soil moisture. Results from the BARDL was compared to the AR method used in chapter 3. The BARDL showed significant improvement over the AR model with an approximately two week gain in lead time. Compared to the AR method, the BARDL also demonstrated a better propensity toward predicting drought events (VCI3M<35) at longer lead times. The uncertainty on the forecasts from the Bayesian models was also lower. The added advantage of the BARDL approach was using forecast distribution to determine the forecast probabilities of drought events at a (VCI3M<35) threshold. Thus, when we forecast a VCI3M for a given time ahead, we can also tell the probability of the average forecast distribution falling below the threshold.

The frequency with which the BARDL predicted drought events was higher during the short rain season of Kenya. Further analysis also showed that the forecasting long-term VCI3M is mainly driven by lagged precipitation. The contribution of the lagged precipitation, however, increased with longer lead times. This observation confirmed that although precipitation affects VCI3M, its response to the changing moisture levels is slow (Quiring et al., 2010). A closer inspection of the model performance metrics at the county level revealed some spatial variations in model performance. For example, counties in the arid regions had better accuracies compared to the semi-arid regions. Recognising that the effects of agricultural drought may vary depending on the Agro-Ecological zones or vegetation land cover (Boken et al., 2005) led us to think about how to incorporate such differences into our drought models.

In chapter 5 of this thesis, we demonstrated the use of the Hierarchical Bayesian Model (HBM) to concurrently forecast VCI3M within different Agro-Ecological Zones and over areas with varying land covers. The HBM was developed from partially-pooled data of VCI3M, precipitation and soil moisture for the different Agro-Ecological Zones and land covers. The results from the HBM, compared to the BARDL, gave more precise forecasts. The results also showed that the forecast from the HBM had approximately a 1-week lead time over the BARDL model. This improvement could also be seen with predicting drought events at VCI3M<35 for arid and semi-arid zones and grasslands and croplands. Forecast metrics over forests were, however same for both models. The observation over forest areas was probably due to minor temporal variations in vegetation condition over forest covers (Los et al., 2019) and the fact that the within-group effects of precipitation and soil moisture on VCI3M in forest areas are not too different from the global effect. The results also showed that the HBM was better at predicting the onset of drought events with a higher success rate than the BARDL model, which is essential for early drought action and preparedness. We also showed that forecast models developed via the HBM are easily transferable to regions with similar data. This transferability means models calibrated in the smaller region can be scaled up to forecast drought over more extensive areas without retraining the HBM model. The transferability is also an advantage because training HBMs are computationally expensive and time-consuming.

Throughout this thesis, we have demonstrated that Bayesian models can be employed for effective drought monitoring and forecasting models mainly because they inherently enable uncertainty analysis and probabilistic interpretation. Although these models are based on complex mathematical algorithms, they are not difficult to implement because all the required software tools are available and easy to access. Furthermore, once the models are well calibrated, the cost associated with setting up the data pipelines and deploying these models are low.

In partnership with Kenya's National Drought Monitoring Agency (NDMA), our main stakeholder for this project, the forecast model developed with Gaussian Processes (GP) has been deployed on servers at The Regional Centre for Mapping of Resources for Development (RCMRD) in Nairobi, Kenya, as part of their web-based Rangeland Decision Support Tool [*]. The deployment of the data pipelines and the VCI forecast model was facilitated by the direct and real-time access to the MODIS EO satellite data provided by the RCMRD and SERVIR-Eastern and Southern Africa Project [†] (SERVIR-ESA). The direct access enables the real-time download of daily MODIS data whenever the satellite passes over the east and southern region of Africa. The pre-processed data over Kenya is feed into the GP data pipeline which produces VCI forecasts. The decision support tool is currently used by organisation like the NDMA, Kenya Rangeland Trust and Kenya Red Cross for monitoring and forecasting short term vegetation condition for timely drought anticipation and early action.

Discussions are also currently going with the stakeholders to test and deploy the other models based on the BARDL and the HBM. The proposed way forward is to replace the current soil moisture data with a near real-time version and re-write the codes for the data pipelines to produce VCI3M forecast at the pixel level. Doing this will allow drought monitoring at both the county and ward levels.

[*] Link to the tool: http://tools.rcmrd.org/

[†] https://servir.rcmrd.org/

Both the BARDL and HBM methods used in this thesis have their strengths, thus, a preferred method or a combination of these methods can also be implemented in the trigger phase of the Early Action protocols (EAP) (IFRC, 2021) used by drought monitoring agencies that practice FbF. Using our models to trigger early warnings will enhance the anticipatory efforts and build confidence in governments and the donor community to invest more in the FbF initiative. A well-implemented EAP eventually translates into building community resilience and the improvement of people's livelihoods, especially agro-pastoralists living in drought-prone communities.

References

Barrett, Adam B, Steven Duivenvoorden, Edward E Salakpi, James M Muthoka, John Mwangi, Seb Oliver and Pedram Rowhani (2020). "Forecasting vegetation condition for drought early warning systems in pastoral communities in Kenya". In: *Remote Sensing of Environment* 248, p. 111886.

Boken, Vijendra K., Arthur P. Cracknell and Ronald L. Heathcote (May 2005). *Monitoring and Predicting Agricultural Drought*. Oxford University Press. ISBN: 9780195162349. DOI: 10.1093/oso/9780195162349.001.0001. URL: https://oxford.universitypressscholarship.com/view/10.1093/oso/9780195162349.001.0001/isbn-9780195162349.

Coughlan de Perez, E., B. van den Hurk, M. K. van Aalst, B. Jongman, T. Klose and P. Suarez (Apr. 2015). "Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts". In: *Natural Hazards and Earth System Sciences* 15.4, pp. 895–904. ISSN: 1684-9981. DOI: 10.5194/nhess-15-895-2015. URL: https://www.nat-hazards-earth-syst-sci.net/15/895/2015/.

Guimarães Nobre, G. (2019). "Floods, droughts and climate variability: From early warning to early action". English. PhD thesis. Vrije Universiteit Amsterdam. ISBN: 9789402816488.

IFRC (2021). "FbF Practitioners Manual". In: *International Federation of Red Cross and Red Crescent*. URL: https://manual.forecast-based-financing.org/.

Kogan, F.N. (1995). "Application of vegetation index and brightness temperature for drought detection". In: *Advances in Space Research* 15.11. Natural Hazards: Monitoring and Assessment Using Remote Sensing Technique", pp. 91–100. ISSN: 0273-1177. URL: http://www.sciencedirect.com/science/article/pii/027311779500079T.

Los, Sietse O., F. Alayne Street-Perrott, Neil J. Loader, Cynthia A. Froyd, Aida Cuní-Sanchez and Robert A. Marchant (2019). "Sensitivity of a tropical montane cloud forest to climate change, present, past and future: Mt. Marsabit, N. Kenya". In: *Quaternary Science Reviews*

218, pp. 34–48. ISSN: 0277-3791. DOI: https://doi.org/10.1016/j.quascirev.2019.06.016. URL: https://www.sciencedirect.com/science/article/pii/S0277379119300733.

Quiring, Steven M. and Srinivasan Ganesh (Mar. 2010). "Evaluating the utility of the Vegetation Condition Index (VCI) for monitoring meteorological drought in Texas". In: *Agricultural and Forest Meteorology* 150.3, pp. 330–339. ISSN: 01681923. DOI: 10.1016/j.agrformet.2009.11.015.

UNDRR (2021). *GAR Special Report on Drought 2021 — UNDRR*. Tech. rep. URL: https://www.undrr.org/publication/gar-special-report-drought-2021.

Chapter 7

Conclusion and Recommendation

In this thesis, we learnt about agricultural drought and its economic importance to people in agro-pastoral communities. We also saw the role of satellite Earth observation data and machine learning for effective drought forecasting to enhance early action and resilience in drought-prone communities.

The overall objective of this thesis was to use remotely sensed Earth observation data to develop models that forecast VCI3M, a popular satellite-derived agricultural drought indicator. The models exhibited good forecast skills, especially when additional factors like precipitation and soil moisture were considered. Using more complex models like the Hierarchical Bayesian Model, we further improved forecast precision by accounting for the spatial variability in our regions of interest. Working within the Bayesian framework enabled the quantification of forecast uncertainties and also derive forecast probabilities on the severity of drought events.

To further improve our models in the thesis, future work should consider a combination of Bayesian frameworks and deep learning methods for forecasting vegetation conditions and predicting drought impact.