University of Reading

Department of Meteorology

Using machine learning to predict the intensification and propagation of East African storms

Sean Kelley

Primary Supervisor: Dr. Eliza Karlowska

Co-supervisors: Dr. Kieran Hunt, Prof. Andy Turner

A report submitted in partial fulfilment of the requirements of
the University of Reading for the degree of
Master of Science in *Climate Change and Artificial Intelligence (AI)*

August 13, 2025

Declaration

I, Sean Kelley, of the Department of Meteorology, University of Reading, confirm that this is my own work and figures, tables, equations, code snippets, artworks, and illustrations in this report are original and have not been taken from any other person's work, except where the works of others have been explicitly acknowledged, quoted, and referenced. I understand that if failing to do so will be considered a case of plagiarism. Plagiarism is a form of academic misconduct and will be penalised accordingly.

I give consent to a copy of my report being shared with future students as an exemplar.

I give consent for my work to be made available more widely to members of UoR and public with interest in teaching, learning and research.

Sean Kelley
August 13, 2025

Abstract

Keywords: a maximum of five keywords/keyphrase separated by commas

Word count: 574

Report code: https://github.com/seangtkelley/uor-msc-dissertation-xai-african-storms

Acknowledgements

Contents

List of Figures

List of Tables

Nomenclature

c Speed of light in a vacuum

h Planck constant

Glossary

El Niño-Southern Oscillation A shift in position of sea surface pressure anomalies between each side of the tropical Pacific Ocean with a period of 2-5 years. ix, 3

Indian Ocean Dipole An irregular oscillation of sea surface temperatures in the Indian Ocean. ix, 3

Intertropical Convergence Zone A belt near the equator where the northeasterly and southeasterly trade winds converge. ix, 2

Machine Learning Subset of artificial intelligence that enables systems to learn from data and improve their performance over time without being explicitly programmed. ix, 3

Madden-Julian Oscillation Main component of tropical intraseasonal variability via a coupling of circulation and convection that travels slowly eastward over the Indian and Pacific Oceans. ix, 2

Mesoscale Convective System A group of thunderstorms organised into a single cloud system that lasts several hours, often resulting in extreme rainfall, flash flooding and hail. ix, 1, 2

Synoptic Scale Meteorological phenomena that occur at a horizontal scale of 200 kilometres and above, often associated with large-scale weather systems. 2

Teleconnection Climate patterns related to each other at large distances, typically thousands of kilometres apart, often influencing weather patterns across regions. 2

Tropical Easterly Jet A high-altitude, easterly wind current stretching over the tropics from South Asia to Africa which is most prominent during the Asian monsoon. ix, 3

Acronyms

ENSO El Niño-Southern Oscillation. 3

IOD Indian Ocean Dipole. 3

ITCZ Intertropical Convergence Zone. 2

MCS Mesoscale Convective System. 1–3

MJO Madden-Julian Oscillation. 2

ML Machine Learning. 3

TEJ Tropical Easterly Jet. 3

Chapter 1

Introduction

Gebrechorkos et al. (2019) Mesoscale Convective System (MCS)

1.1 Research Objectives

This section outlines the primary objectives of the research, which include:

- To investigate the factors contributing to the intensification and propagation of East African storms.
- To develop a machine learning model capable of predicting storm behavior based on historical data.
- To evaluate the performance of the proposed model against existing forecasting methods.

Chapter 2

Background

2.1 Mesoscale Convective Systems (MCSs)

Mesoscale Convective Systems (MCSs) constitute a critical component of regional weather forecasting and climatology due to their significant size, duration, and impact. Officially, MCSs are defined as a complex of thunderstorms which become organised on a scale larger than any of the individual thunderstorms (NWS, 2025). Consequently, these storm systems often last for several hours and cover areas of tens of thousands of square kilometres. In addition, they often produce severe weather phenomena, including flooding, strong winds, and hail (Houze, 2014). Unlike many Synoptic Scale systems, MCSs are not usually associated with a well-defined centre of circulation and instead are characterised by their multi-scale organisation, typically incorporating a variety of convective cells and larger-scale features, such as squall lines or mesoscale convective complexes (AMS, 2024; NWS, 2025). These systems are prevalent throughout the world and thus are key to understanding regional climatology. For example, in the United States, MCSs are a primary driver of warm-season precipitation over the Great Plains (Haberlie and Ashley, 2019). Similarly, the Sahel region of Africa is heavily affected by these storms, producing some of the strongest MCSs globally due to it being a climatic transition zone with strong seasonal cycles (Zipser et al., 2006).

2.1.1 MCSs in the Horn of Africa

The Horn of Africa is a region with complex topography and large-scale climatic variability that affects MCS development. From the west, the Sahel fades into the Ethiopian Highlands and later the East African Rift Valley, characterised by high topographic relief and complex orography. In the east, the Ethiopian Highlands transition into the low-lying coastal plains of Somalia. Unlike most other countries at this latitude, most of Somalia is arid or semi-arid, with the exception of its border region with Kenya (Beck et al., 2023). This contrast in geography is reflected in storm development and precipitation patterns. The mountains of Ethiopia dominate local convective processes (Negash et al., 2024) while the low-lying areas on the south-eastern coast of the region are not nearly as conducive to MCS development and thus are much more susceptible to storm patterns over the Indian Ocean and the Gulf of Aden (Camberlin et al., 2024). The combination of land surface temperature and soil moisture also impact the storm development and intensification. Notably, multiple studies over distinct regions have demonstrated that strong soil moisture gradients can intensify convection (Barton et al., 2021; Klein and Taylor, 2020; Taylor et al., 2017). These processes are likewise relevant to the Horn of Africa, especially in transitional climates bridging arid and humid zones. Large-scale teleconnections also play a major role in governing MCS activity in this region. The Madden-Julian Oscillation (MJO) is a dominant intra-seasonal factor which modulates rainfall in the tropics and in East Africa, its active phases coincide with increased convection and extreme rainfall events (Camberlin et al., 2019; Ochieng et al., 2023; Pohl and Camberlin, 2006). Quite uniquely for the tropics, the Intertropical Convergence Zone (ITCZ) passes over the region twice per year leading to two distinct rainy seasons, one short and one long (Palmer et al., 2023; Tefera et al., 2025). The Tropical Easterly Jet (TEJ) is another key feature enhancing vertical wind shear which can both facilitate or hinder MCS development (Farnsworth et al., 2011; Vashisht et al., 2021). Additionally, the El Niño-Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) have also been shown to modulate rainfall patterns in the region, with coupled regional climate models able to reproduce these patterns at various timescales (Dubache et al., 2019; Endris et al., 2019; Vashisht et al., 2021). Overall, despite this complex array of factors, MCSs have been shown to account for over 60% of extreme rainfall in Ethiopia and Somalia (Hill et al., 2023).

MAYBE NEED A PARAGRAPH ABOUT LIMITED RESOURCES AND DATA AVAILABILITY IN THE REGION

2.2 Data-driven Scientific Discovery

The physical sciences have traditionally favoured numerical methods for complex system modelling, most often relying on the resultant equations from first principles and physical laws as the basis for simulation. Since most of these systems are too complex to be solved analytically, numerical methods are used to approximate solutions within certain constraints limiting grid resolution, stability, and subgrid parametrisation [SOURCE FOR NWP and maybe physics too]. The primary benefit of such techniques is that they are inherently based on well-established physical laws, which can provide a degree of interpretability and trustworthiness.

I DON"T LIKE THIS SETNENCE: However, these methods often struggle with the non-linear and chaotic nature of many physical systems, particularly in meteorology where small changes can lead to vastly different outcomes [SOURCE FOR CHAOS IN METEOROLOGY, Lorenz maybe].

Data-driven methods have historically been avoided due the lack of data and computational infeasibility, even when such methods were known and their potential well-understood [SOURCE FOR EARLY NEURAL NETS]. However, the advent of large datasets and ever-increasing computational power as hardware manufacturers kept pace with Moore's Law [glossary entry] has led to a resurgence of interest in data-driven methods [SOURCE FOR MOORES LAW]. Particular credit can be attributed to the surprise success of deep learning in computer vision and natural language processing, which has led to a broader acceptance of Machine Learning (ML) techniques across many scientific disciplines [SOURCE FOR ALEXNET and something NLP].

PARAGRAPH ABOUT ML METHODS GENERICALLY IN SCIENCE HERE

2.2.1 Applications in Meteorology

The field of meteorology has been no exception to the increasing popularity of data-driven methods, with a growing body of literature exploring the application of ML techniques to various meteorological problems. The use of ML in meteorology is particularly promising due to the non-linear nature of atmospheric processes and hard-to-model subgrid processes, which have long been a source of instability in numerical models [SOURCE FOR COMPLEXITY OF ATMOSPHERIC PROCESSES].

PIML FOR LARGE SCALE MODELS came about to address ml models not obeying the laws of physics, e.g. the biug NVIDIA one, PhyDL-NWP https://arxiv.org/abs/2505.14555

Modern nowcasting methods, such as physics-conditional deep generative models (e.g., Nowcast-Net), unify physical evolution with statistical learning to achieve short-term, high-resolution forecasts of extreme precipitation by leveraging radar data and ensemble neural networks[NowcastNet SOURCE https://www.nature.com/articles/s41586-023-06184-4].

DOwnscaleing: PhyDL-NWP as well https://arxiv.org/abs/2505.14555

subgrid parameterization: gotta be something

2.3 Explainable and Interpretable ML (XAI)

Techniques such as LIME, SHAP, and Layerwise Relevance Propagation (LRP) have been widely adopted to interpret complex ML models...

2.3.1 SHapley Additive exPlanations (SHAP)

SPECIFIC EXPLANATION OF SHAP AS ThIS IS ThE BACKGROUND

2.4 Literature Review

2.4.1 XAI in Meteorology

For example, LIME has helped identify crucial features in random forest models for seasonal precipitation prediction, with explanations aligning closely to known meteorological phenomena like ENSO impacts[https://www.sciencedirect.com/science/article/abs/pii/S1352231024004722].

2.4.2 XAI for Convective Systems

Case Studies in Convective Systems:

Tropical Convection and MCSs: LRP has revealed the importance of large-scale vertical velocity and wind shear in predicting convective area and organization in the tropics, with moisture and thermodynamic factors primarily affecting the convective area, and horizontal wind fields impacting system organization[https://www.sciencedirect.com/science/article/abs/pii/S1352231024004722].

Paleoclimate reconstruction of Indian monsoon [https://cp.copernicus.org/articles/21/1/2025/].

Relevant Example: Indian Monsoon Lows, Hunt and Turner (2023) demonstrated the utility of interpretable gradient-boosted decision-tree ensembles in uncovering new dynamical relationships governing Indian monsoon low-pressure systems (Hunt and Turner, 2024).

Limitations include for Hunt Turner 2023: - only applied to indian monsoon. would be interesting to apply to other regions and storm types - less focus on any prediction capability of models, more just focused on finding novel meteorological relationships

relevant and motivate the work

Chapter 3

Methodology

Chapter 4

Results

4.1 Summary

Chapter 5

Discussion and Analysis

5.1 Significance of the findings

In this chapter, you should also try to discuss the significance of the results and key findings, in order to enhance the reader's understanding of the investigated problem

5.2 Limitations

Discuss the key limitations and potential implications or improvements of the findings.

5.3 Summary

Chapter 6

Conclusion

References

AMS, 2024: cyclonic scale - glossary of meteorology. URL https://glossary.ametsoc.org/wiki/Cyclonic_scale.

Barton, E. J., C. M. Taylor, C. Klein, P. P. Harris, and X. Meng, 2021: Observed soil moisture impact on strong convection over mountainous tibetan plateau. *Journal of Hydrometeorology*, **22**, 561–572, https://doi.org/10.1175/JHM-D-20-0129.1.

Beck, H. E., and Coauthors, 2023: High-resolution (1 km) köppen-geiger maps for 1901–2099 based on constrained cmip6 projections. *Scientific Data 2023 10:1*, **10**, 1–16, https://doi.org/10.1038/s41597-023-02549-6.

Camberlin, P., O. A. Dabar, B. Pohl, M. M. Waberi, K. Hoarau, and O. Planchon, 2024: Contribution of western arabian sea tropical cyclones to rainfall in the horn of africa and southern arabian peninsula. *Journal of Geophysical Research: Atmospheres*, **129**, e2024JD041 109, https://doi.org/10.1029/2024JD041109.

Camberlin, P., W. Gitau, G. Kiladis, E. Bosire, and B. Pohl, 2019: Intraseasonal to interannual modulation of diurnal precipitation distribution over eastern africa. *Journal of Geophysical Research: Atmospheres*, **124**, 11 863–11 886, https://doi.org/10.1029/2019JD031167.

Dubache, G., B. A. Ogwang, V. Ongoma, and A. R. M. T. Islam, 2019: The effect of indian ocean on ethiopian seasonal rainfall. *Meteorology and Atmospheric Physics*, **131**, 1753–1761, https://doi.org/10.1007/S00703-019-00667-8/METRICS.

Endris, H. S., C. Lennard, B. Hewitson, A. Dosio, G. Nikulin, and G. A. Artan, 2019: Future changes in rainfall associated with enso, iod and changes in the mean state over eastern africa. *Climate Dynamics*, **52**, 2029–2053, https://doi.org/10.1007/S00382-018-4239-7/FIGURES/12.

Farnsworth, A., E. White, C. J. Williams, E. Black, and D. R. Kniveton, 2011: Understanding the large scale driving mechanisms of rainfall variability over central africa. *Advances in Global Change Research*, **43**, 101–122, https://doi.org/10.1007/978-90-481-3842-5_5.

Gebrechorkos, S. H., C. Bernhofer, and S. Hülsmann, 2019: Impacts of projected change in climate on water balance in basins of east africa. *The Science of the total environment*, **682**, 160–170, https://doi.org/10.1016/J.SCITOTENV.2019.05.053.

Haberlie, A. M., and W. S. Ashley, 2019: A radar-based climatology of mesoscale convective systems in the united states. *Journal of Climate*, **32**, 1591–1606, https://doi.org/10.1175/JCLI-D-18-0559.1.

Hill, P. G., T. H. Stein, and C. Cafaro, 2023: Convective systems and rainfall in east africa. *Quarterly Journal of the Royal Meteorological Society*, **149**, 2943–2961, https://doi.org/10.1002/QJ.4540.

Houze, R. A., 2014: Mesoscale convective systems. *International Geophysics*, **104**, 237–286, https://doi.org/10.1016/B978-0-12-374266-7.00009-3.

Hunt, K. M., and A. G. Turner, 2024: Using interpretable gradient-boosted decision-tree ensembles to uncover novel dynamical relationships governing monsoon low-pressure systems. *Quarterly Journal of the Royal Meteorological Society*, **150**, 1–24, https://doi.org/10.1002/QJ.4582.

Klein, C., and C. M. Taylor, 2020: Dry soils can intensify mesoscale convective systems. *Proceedings of the National Academy of Sciences of the United States of America*, **117**, 21 132–21 137, https://doi.org/10.1073/PNAS.2007998117/SUPPL_FILE/PNAS.2007998117.SAPP.PDF.

Negash, E., B. V. Schaeybroeck, P. Termonia, M. V. Ginderachter, K. V. Weverberg, and J. Nyssen, 2024: Topoclimate and diurnal cycle of summer rain over the ethiopian highlands in a convection-permitting simulation. *International Journal of Climatology*, **44**, 406–427, https://doi.org/10.1002/JOC.8334.

NWS, N., 2025: Noaa's national weather service - glossary. URL https://forecast.weather.gov/glossary.php.

Ochieng, P. O., I. Nyandega, B. Wambua, and V. Ongoma, 2023: Linkages between madden–julian oscillation and drought events over kenya. *Meteorology and Atmospheric Physics*, **135**, 1–23, https://doi.org/10.1007/S00703-022-00948-9/METRICS.

Palmer, P. I., and Coauthors, 2023: Drivers and impacts of eastern african rainfall variability. *Nature Reviews Earth & Environment 2023 4:4*, **4**, 254–270, https://doi.org/10.1038/s43017-023-00397-x.

Pohl, B., and P. Camberlin, 2006: Influence of the madden–julian oscillation on east african rainfall: Ii. march–may season extremes and interannual variability. *Quarterly Journal of the Royal Meteorological Society*, **132**, 2541–2558, https://doi.org/10.1256/QJ.05.223.

Taylor, C. M., and Coauthors, 2017: Frequency of extreme sahelian storms tripled since 1982 in satellite observations. *Nature 2017 544:7651*, **544**, 475–478, https://doi.org/10.1038/nature22069.

Tefera, A. K., G. Liguori, W. Cabos, and A. Navarra, 2025: Seasonal forecasting of east african short rains. *Scientific Reports 2025 15:1*, **15**, 1–10, https://doi.org/10.1038/s41598-025-86564-0.

Vashisht, A., B. Zaitchik, and A. Gnanadesikan, 2021: Enso teleconnection to eastern african summer rainfall in global climate models: Role of the tropical easterly jet. *Journal of Climate*, **34**, 293–312, https://doi.org/10.1175/JCLI-D-20-0222.1.

Zipser, E. J., D. J. Cecil, C. Liu, S. W. Nesbitt, and D. P. Yorty, 2006: Where are the most intense thunderstorms on earth? *Bulletin of the American Meteorological Society*, **87**, 1057–1072, https://doi.org/10.1175/BAMS-87-8-1057.

Appendix A

An Appendix Chapter