



**UNIVERSITY OF LEEDS**

---

# Severe Weather over East Africa

---

by

BETHANY JANE WOODHAMS

Submitted in accordance with the requirements for the degree of Doctor of  
Philosophy

The University of Leeds  
School of Earth and Environment

September 2019



# Declaration of authorship

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to work of others.

The publication Woodhams et al., 2018, *What Is the Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall over Tropical East Africa?*, *Monthly Weather Review* 146(9), p2757–2780, doi: <https://doi.org/10.1175/MWR-D-17-0396.1>, jointly authored with Cathryn E. Birch, John H. Marsham, Caroline L. Bain, Nigel M. Roberts and Douglas F. A. Boyd, is included as chapter 2 of this thesis. The text was solely written by the candidate, with comments from co-authors. The candidate performed all data analysis and produced all figures. Birch, Marsham, Bain and Roberts contributed to method refinement and interpretation of results. Boyd was responsible for running the operational model.

The publication Woodhams et al., 2019: *Identifying Key Controls on Storm Formation over the Lake Victoria Basin.*, *Monthly Weather Review*, 147(9), p3365–3390, doi: <https://doi.org/10.1175/MWR-D-19-0069.1>, jointly authored with Cathryn E. Birch, John H. Marsham, Todd P. Lane, Caroline L. Bain and Stuart Webster, is included as chapter 3 of this thesis. The text was solely written by the candidate, with comments from co-authors. The candidate performed all data analysis and produced all figures. Birch, Marsham, Lane and Bain contributed to interpretation of results. The candidate ran the model simulations, with guidance from Webster.

The fourth chapter Woodhams et al., *Aircraft observations of the lake-land breeze circulation over Lake Victoria from the HyVic pilot flight campaign*, jointly authored with John H. Marsham, Cathryn E. Birch, Caroline L. Bain, Andrew J. Hartley, Paul Barrett, Jennifer K. Fletcher, Stuart Webster and Uganda National Meteorological Authority, has been prepared for submission to *Quarterly Journal of the Royal Mete-*

*orological Society*. The text was solely written by the candidate, with comments from co-authors. The candidate performed all data analysis and produced all figures. The candidate, Hartley, Fletcher and Barrett were mission scientists on the HyVic flights. All co-authors contributed to interpretation of results. Webster ran the operational and additional research simulations. The Facility for Airborne and Atmospheric Measurement (FAAM) team distributed the observational data.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2019 The University of Leeds and Bethany Jane Woodhams

# Acknowledgements

Firstly, I would like to thank my supervisor, Cathryn Birch, for her amazing support at every step, and her encouragement to pursue every opportunity which came my way. Thank you also to John Marsham, whose supervision meetings I always left, full of exciting and new ideas. Many thanks to Caroline Bain, for always inspiring more ‘blue-sky thinking’. I am also grateful to everyone in the Dynamics research group at Leeds, for creating such a supportive and friendly research atmosphere.

This PhD was funded by the NERC SPHERES DTP (grant NE/L002574/1). Many thanks to all those involved in the DTP, especially Nigel. NERC also funded many training courses which I attended, including the Atmospheric Measurement Summer School on the Isle of Arran, and the FAAM training course.

Additional funding was provided through a CASE award with the Met Office. Many thanks to all those at the Met Office (especially Caroline Bain, Stu Webster, Andy Hartley, Paul Barrett and Chris Tubbs) for making it feel like my second home. In particular, the UM could not have been conquered without the generous help of Stu, who would often answer my emails at ridiculous hours.

I am grateful to the Menzies Centre at King’s College London, from which I received the Australian Bicentennial Award, and to the Royal Meteorological Society, from which I received the Rupert Ford Award, in order to fund a three month visit to the University of Melbourne. Thank you to Todd Lane, who supervised me during my visit, and to all the wonderful people I met in Melbourne. Thank you also to LeedsforLife and the WMO, for travel assistance to attend the 2018 GEWEX conference in Canmore, Canada.

The HyVic pilot flight campaign could not have taken place without the tireless work of all those at FAAM. A special thank you to Mo Smith, Doug Anderson, Hannah Price, Dave Sproson, David Simpson and Barney Black for all their help. Thank you to the HyVic dream team (Jennifer Fletcher, Andy Hartley, Paul Barrett and Rosalind

West) for all their help and support during the campaign, and for the many Tuskers shared. Thank you also to Festus Luboyera and Solomon Mangeni at Uganda National Meteorological Authority (UNMA), without whom HyVic may never have made it off the ground.

Thank you also to those at Kenya Meteorological Department (KMD), Tanzania Meteorological Agency (TMA), and UNMA, who took the time to speak with me and share their insight and knowledge during a visit to East Africa in December 2016.

A special thanks to Paddy, Josh, Louise, Sam and Sagar, who have put up with me as a housemate during this time. In particular, thank you to Paddy and Josh for the epic mah-jong nights, and fun times in Woodside Avenue. Thank you to the East Africa Posse (Dean and Declan) and UM Ugali for fun, but helpful, discussions over the years. A big shout-out to Office 2.10 in Fairbairn (Anya, Ben and Dean) - we saw some things, but came out stronger on the other side. To Jenny, Sarah, Sam, and all those who I have met during my PhD, thank you for your friendship. A huge thank you to the one and only Craig Poku, for all the ridiculous conversations, baking and science we have shared.

Thank you to my family, who have supported me from the beginning, and helped me believe that I can do anything I set my mind to. Although it is highly unlikely that they will ever read this thesis, I hope they know how much I appreciate them!

To Chetan: you have supported me throughout this process, and built me up when I was feeling down. Thank you for listening to me complain, sharing my excitement when I found out something ‘cool’, and bringing a general smattering of gooviness to the proceedings.

Acknowledgement must be given to NBC’s *Friends* and FOX’s *New Girl*, for keeping me sane in these last few months. I cannot finish without mentioning the 1996 cinematic masterpiece *Twister*, without which I would not be who I am today.

Finally, thank you again to everyone involved in my PhD, for making it such an enjoyable time in my life. Even in the last few days of preparing this thesis, I have not stopped smiling as I have looked back on the amazing opportunities that I have been lucky enough to experience, and all the wonderful people I have met. I count myself very lucky that I have spent the last four years being able to do what I love.

# Abstract

Convective storms in East Africa have severe socio-economic consequences for the region. Lake Victoria is a hotspot for severe weather, which contributes to the estimated 5,000 fatalities on the lake every year, and its reputation as the most dangerous water body in the world. Although East Africa presents its own unique challenges, understanding and predicting tropical convection remains one of the grand challenges of meteorology.

Between 2012–2019, the UK Met Office ran a convection-permitting (CP) model over East Africa, aimed at improving forecasts of severe weather, in particular over Lake Victoria. In this thesis, an assessment of the added-value of the CP forecast, relative to the Met Office operational global model, which was convection parametrised, is performed. This is the first evaluation of a CP model in the tropics conducted over an extended period of time (2 years). The evaluation demonstrates the enhanced skill of the CP model for the prediction of sub-daily rainfall, but large uncertainties remain in the predicted location of precipitation, as well as issues with the distribution of rainfall intensities.

Poor skill in forecasting over East Africa can also be attributed to an incomplete understanding of the processes responsible for severe weather, which is addressed with modelling and observational studies in this thesis. CP simulations with 1.5 km horizontal grid-spacing highlight the importance of the lake–land breeze circulation, and its interaction with the large-scale prevailing winds, in the formation and propagation of storms. The influence of convection over land on nocturnal storm formation over the lake is also demonstrated. First results from the HyVic pilot flight campaign, using the FAAM aircraft to sample the lake–land breeze circulation over Lake Victoria in unprecedented detail, are presented, and highlight the need for very high-resolution modelling to capture the observed small-scale features of the circulation.





# Contents

<b>Declaration of authorship</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Contents</b>	<b>ix</b>
<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xix</b>
<b>Abbreviations</b>	<b>xxi</b>
<b>1. Introduction</b>	<b>1</b>
1. Motivation . . . . .	1
2. Background . . . . .	4
2.a. Challenge of forecasting convection . . . . .	4
2.b. Convection-permitting models . . . . .	6
2.c. Lake–land breeze circulations . . . . .	8
3. East Africa . . . . .	13
3.a. Climatology of East Africa . . . . .	13

3.b.	Seasonal and climate variability . . . . .	17
3.b.1.	Variability of the Long and Short Rains . . . . .	17
3.b.2.	MJO . . . . .	17
3.b.3.	ENSO and IOD . . . . .	18
3.b.4.	Seasonal prediction . . . . .	19
3.c.	Climate predictions for East Africa . . . . .	20
3.d.	Lake Victoria . . . . .	21
3.d.1.	The effect of Lake Victoria on the diurnal cycle of precipitation . . . . .	21
3.d.2.	Factors influencing regional climate over the Lake Victoria basin . . . . .	24
3.d.3.	Feedbacks between convection over land and lake . . . . .	33
3.d.4.	Discussion and summary of Lake Victoria literature . . . . .	33
4.	Methods . . . . .	34
4.a.	East Africa CP Model verification . . . . .	34
4.a.1.	MetUM East Africa CP model . . . . .	34
4.a.2.	Verification Methods . . . . .	37
4.b.	Case study analysis . . . . .	40
4.c.	Aircraft observations - HyVic pilot flight campaign . . . . .	41
5.	Thesis aims and structure . . . . .	42
	References . . . . .	45

**2. What Is the Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall over Tropical East Africa? 63**

Abstract . . . . .	65
--------------------	----

1. Introduction . . . . .	65
---------------------------	----

2.	Methods . . . . .	67
2.a.	Models . . . . .	67
2.b.	Observations . . . . .	68
2.c.	Verification methods . . . . .	69
2.c.1.	Objective analysis . . . . .	69
2.c.2.	Fractions Skill Score . . . . .	70
3.	Results . . . . .	71
3.a.	Forecast example . . . . .	71
3.b.	Diurnal cycle . . . . .	72
3.c.	Precipitation rates . . . . .	73
3.d.	Lake Victoria objective analysis . . . . .	73
3.e.	FSS for 24-h accumulation rates . . . . .	76
3.e.1.	Localized FSS . . . . .	76
3.e.2.	LFSS seasonal variability . . . . .	76
3.f.	FSS for 3-hourly precipitation rates . . . . .	77
4.	Discussion and conclusions . . . . .	81
	Appendix: FSS Computational Details and Choice of Rainfall Threshold . . . . .	84
	References . . . . .	85
	Erratum . . . . .	89

**3. Identifying Key Controls on Storm Formation over the Lake Victoria Basin . . . . . 91**

	Abstract . . . . .	93
1.	Introduction . . . . .	93
2.	Methods . . . . .	96
2.a.	Model . . . . .	96

2.b.	Observations . . . . .	96
3.	Results . . . . .	97
3.a.	Case study introduction . . . . .	97
3.a.1.	Dry period (9–11 July 2015) . . . . .	97
3.a.2.	Long rains (LR) storm (5–8 May 2015) . . . . .	98
3.a.3.	Dry season (DS) storm (27–30 July 2016) . . . . .	100
3.b.	Case study: Dry period . . . . .	100
3.c.	Case study: Long rains storm . . . . .	103
3.d.	Case study: Dry season storm . . . . .	107
3.e.	Case study comparison . . . . .	109
4.	Summary . . . . .	112
5.	Discussion and conclusions . . . . .	114
	References . . . . .	116
	Erratum . . . . .	119
<b>4.</b>	<b>Aircraft observations of the lake–land breeze circulation over Lake Victoria from the HyVic pilot flight campaign</b>	<b>121</b>
1.	Introduction . . . . .	122
2.	Methods . . . . .	124
2.a.	Flights . . . . .	124
2.a.1.	C130: Evening flight . . . . .	126
2.a.2.	C131: Morning flight . . . . .	126
2.b.	Aircraft data . . . . .	127
2.c.	MetUM simulations . . . . .	128
3.	Results . . . . .	129
3.a.	General conditions . . . . .	129

3.b.	Lake breeze . . . . .	129
3.c.	Land breeze . . . . .	136
4.	Discussion and conclusions . . . . .	142
	References . . . . .	146
<b>5.</b>	<b>Conclusions</b>	<b>151</b>
1.	Overview . . . . .	151
2.	Wider impact of the work . . . . .	155
3.	Limitations of the work . . . . .	156
4.	Recommendations for future work . . . . .	158
4.a.	Evaluation of convection-permitting models . . . . .	158
4.b.	Controls on storms over the Lake Victoria basin . . . . .	159
4.c.	Idealised modelling of land breezes . . . . .	160
4.d.	Lake surface temperatures . . . . .	161
4.e.	Extended field campaign . . . . .	161
5.	Summary . . . . .	162
	References . . . . .	164



# List of Figures

1.1. Orography of East Africa. . . . .	2
1.2. Diagram showing a sea breeze circulation cell. . . . .	9
1.3. Annual cycle of rainfall and wind over East Africa. . . . .	14
1.4. Rainfall onset and seasons over East Africa. . . . .	16
1.5. Map showing the time of maximum precipitation over East Africa. . . . .	22
1.6. Time-longitude plot of diurnal rainfall occurrence over Lake Victoria. . . . .	24
1.7. Vertical cross-sections over Lake Victoria to show the effect of the lake during the daytime. . . . .	31
1.8. Vertical cross-sections over Lake Victoria to show the effect of the lake during the nighttime. . . . .	32
1.9. Frequency distributions of precipitation and wind from observations, East Africa CP model, and global MetUM. . . . .	35
2.1. A map showing orography over the domain spanned by the operational Met Office CP model in East Africa in 2017. . . . .	66
2.2. Examples of forecasts from the global and CP models and corresponding IR and GPM observations. . . . .	71
2.3. Mean diurnal cycle of precipitation over the land and Lake Victoria from the global and CP models and observations. . . . .	72
2.4. Precipitation rates as a function of percentile for the global and CP models and observations. . . . .	74
2.5. Hit rate, false alarm ratio and ROC area under curve for the global and CP models for a range of storm size and intensities. . . . .	75

2.6.	Mean Fractions Skill Score (FSS) as a function of neighborhood size for 24-h rainfall accumulations, shown for the global and CP models and a 24-h persistence forecast. . . . .	76
2.7.	Maps of Localised Fractions Skill Score (LFSS) for the global and CP models for 24-h rainfall accumulations. . . . .	77
2.8.	Maps of perturbation in LFSS for different seasons computed for the global model. . . . .	78
2.9.	Maps of perturbation in LFSS for different seasons computed for the CP model. . . . .	79
2.10.	Mean FSS as a function of neighborhood size, computed at 3-h intervals for the global and CP models, and a 24-h persistence forecast. . . . .	80
2.11.	Mean FSS as a function of forecast lead time for the global and CP models, and a 24-h persistence forecast. . . . .	81
2.12.	Probability distribution of observed precipitation rate and mean FSS as a function of observed precipitation rate for the global and CP models. . . . .	82
2.13.	(A1) Mean FSS as a function of percentile threshold. . . . .	86
3.1.	A map showing the CP model domain and the transect used in the analysis. Orography over the domain and along the transect is also shown. . . . .	94
3.2.	Observed and simulated brightness temperature and precipitation during the dry period case study. . . . .	98
3.3.	Observed and simulated brightness temperature and precipitation during the long rains storm case study. . . . .	99
3.4.	Observed and simulated brightness temperature and precipitation during the dry season storm case study. . . . .	101
3.5.	Mean simulated 10 m wind and 1.5 m specific humidity during each case study. . . . .	102
3.6.	Simulated 10 m wind and 1.5 m specific humidity during the dry period case study. . . . .	103
3.7.	Cross-sections of various simulated variables during the dry period case study. . . . .	104
3.8.	CAPE and CIN at 1800 and 0600 LT for each case study . . . . .	105



3.9. Cross-sections of various simulated variables during the long rains storm case study. . . . .	106
3.10. Cross-sections of various simulated variables during the dry season storm case study. . . . .	108
3.11. Mean profiles of specific humidity and wind during each case study, with monthly means from ERA-Interim for comparison. . . . .	109
3.12. Hovmöller of along-transect 10 m wind, vertical motion at 1000 m AGL, and OLR during each case study. . . . .	110
3.13. Hovmöller of 1.5m virtual potential temperature and specific humidity, and OLR during each case study. . . . .	111
3.14. Schematics showing winds and moisture during the dry period case study.	114
3.15. Schematics showing winds, moisture and precipitation during the storm case studies. . . . .	115
4.1. Flight tracks for the HyVic pilot flight campaign. . . . .	125
4.2. Model domains for MetUM runs over HyVic period. . . . .	128
4.3. General conditions during HyVic flights. . . . .	130
4.4. Lake breeze observed during flight C130. . . . .	131
4.5. Profiles of the boundary layer during flights C130 and C131. . . . .	134
4.6. Land breeze observed during flight C131. . . . .	137
4.7. Virtual potential temperature and along-transect winds measured by sondes during morning flight. . . . .	138
4.8. Along-transect winds, specific humidity and vertical velocity along aircraft runs over lake during the morning flight. . . . .	140



# List of Tables

3.1. Simulation details of the case studies. . . . .	97
3.2. Summary of key meteorological variables and other important information for each case study. . . . .	113
4.1. A summary of the flights performed as part of the HyVic pilot field campaign. . . . .	124
4.2. Aircraft observations used in this paper. . . . .	127



# Abbreviations

**AGL** above ground level

**AMMA** African Monsoon Multidisciplinary Analysis

**ARPEGE** Action de Recherche Petite Echelle Grande Echelle

**ATD** arrival time difference

**AUC** area under curve

**AVAPS** Airborne Vertical Atmospheric Profiling System

**AWS** automatic weather station

**BL** boundary layer

**CAPE** convective available potential energy

**CLM** Community Land Model

**CMIP** Coupled Model Intercomparison Project

**COSMO** Consortium for Small-Scale Modelling

**CP** convection-permitting

**DYMECS** Dynamical and Microphysical Evolution of Convective Storms

**EAC** East African Community

**ECMWF** European Centre for Medium-Range Weather Forecasts

**ENDGAME** Even Newer Dynamics for General Atmospheric Modelling of the Environment

**ENSO** El Niño Southern Oscillation

**FAAM** Facility for Airborne Atmospheric Measurements

**FCFA** Future Climate for Africa

**FLake** Freshwater lake

**FSS** Fractions Skill Score

**GCM** General Circulation Model

**GCRF** Global Challenges Research Fund

**GFS** Global Forecast System

**GHACOF** Greater Horn of Africa Climate Outlook Forum

**GIN** GPS Inertial Navigation

**GLOBE** Global Land One-kilometer Base Elevation

**GloSea5** Global Seasonal Forecast System 5

**GPM** Global Precipitation Measurement

**GTS** Global Telecommunications System

**HPE** hydrostatic primitive equations

**IAS** indicated air speed

**IFS** Integrated Forecasting System

**IMPALA** Improving Model Processes for African Climate

**IOD** Indian Ocean Dipole

**IR** infrared

**ITCZ** Intertropical Convergence Zone

**KMD** Kenya Meteorological Department

**LES** Large Eddy Simulation

**LFSS** Localised Fractions Skill Score

**LST** lake surface temperature

**LT** Local Time

**MCS** mesoscale convective system

**MetUM** Met Office Unified Model

**MJO** Madden–Julian oscillation

**MSG** Meteosat Second Generation

**MSL** above mean sea level

**MSS** Meteorological Service Singapore

**NCAR** National Centre for Atmospheric Research

**ND** New Dynamics

**NERC** Natural Environment Research Council

**NOAA** National Oceanic and Atmospheric Administration

**NWP** numerical weather prediction

**OLR** outgoing longwave radiation

**POM** Princeton Ocean Model

**RA1T** Regional Atmosphere 1 for the Tropics

**RAMS** Regional Atmospheric Modelling System

**RCM** Regional Climate Model

**RCOF** Regional Climate Outlook Forum

**ROC** receiver operating characteristic

**SAR** synthetic aperture radar

**SAWS** South African Weather Service

**SEVIRI** Spinning Enhanced Visible and Infrared Imager

**SHL** Saharan Heat Low

**SST** sea surface temperature

**SWFDP** Severe Weather Forecasting Demonstration Project

**SWIFT** Science for Weather Information and Forecasting Techniques

**TMA** Tanzania Meteorological Agency

**TRMM** Tropical Rainfall Measuring Mission

**UAV** unmanned aerial vehicle

**UNECA** United Nations Economic Commission for Africa

**UNMA** Uganda National Meteorological Authority

**VCP** Voluntary Cooperation Programme

**IEWS** Lake Victoria Intense Early Warning System

**WMO** World Meteorological Organization

**WRF** Weather Research and Forecasting

**WVSS-II** Water Vapour Sensing System II



# Chapter 1.

## Introduction

### 1. Motivation

The complex geography of East Africa distinguishes the region as an important and unique place to study the local meteorology. Bounded by the Indian Ocean to the east, punctuated by the steep slopes of the Great Rift Valley and Ethiopian Highlands, and scattered with vast inland water bodies, East Africa (Fig. 1.1) is subject to a variety of large-scale and mesoscale forcings, which determine its weather. Of particular significance is Lake Victoria, the world's largest tropical lake, which is situated in a depression of the Great Rift Valley. Convection is enhanced over the lake by nocturnal land breezes, posing a danger to many users of the lake.

Consisting of many developing nations, East Africa has a high vulnerability to severe weather. The East African Community (EAC)<sup>1</sup> recognises that floods, landslides and drought are among the main reasons for displacement of East Africans from their homes (EAC Secretariat 2014). The EAC Deputy Secretary General in charge of Finance and Administration, Hon. Jesca Eriyo, states that poverty increases the vulnerability to natural hazards of many East Africans. This vulnerability is attributed to a lack of knowledge, as well as a lack of technology for adaption and preparedness.

Vulnerability is further increased by the high population density in the Lake Victoria basin, where urban areas are some of the fastest-growing in the world (Seto et al. 2012). It is estimated that 30 million people live on the shores of Lake Victoria, of which  $\sim 3.5$  million (including 200,000 fishermen) are dependent on the lake for their livelihoods (Semazzi 2011). The lake is also an important trade and transport link between its

---

<sup>1</sup>The EAC is a regional intergovernmental organisation with 6 Partner States: the Republics of Burundi, Kenya, Rwanda, South Sudan and Uganda, and the United Republic of Tanzania

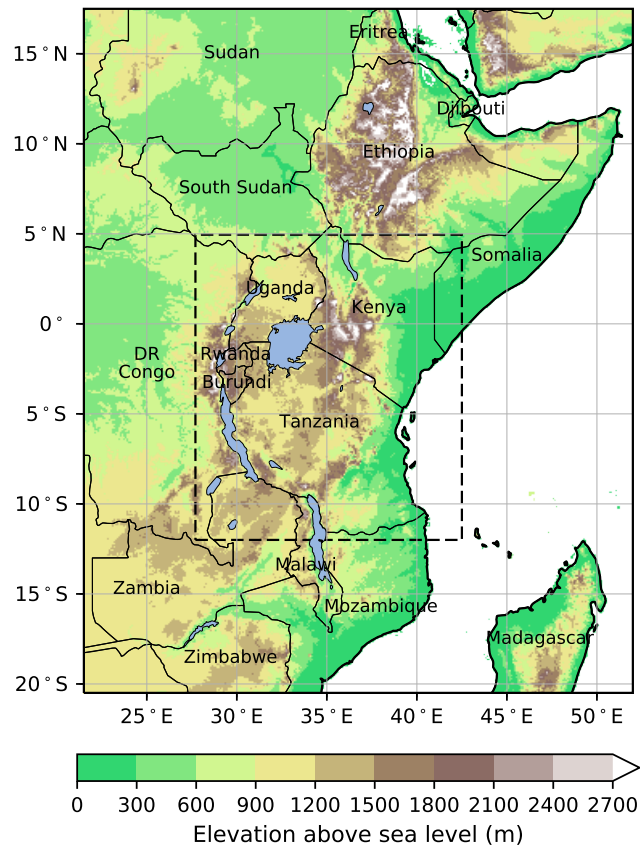


Figure 1.1.: A map showing the orography over East Africa. Elevation data are from the Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model (Hastings and Dunbar 1999).

neighbouring countries. In addition to supporting its catchment area, Lake Victoria is a major source of the Nile, thereby impacting over 300 million inhabitants of the Nile basin (Semazzi 2011). The lake and its basin are integral to food production, with much of the surrounding land used for agriculture and the lake itself supporting fisheries. The lake also supplies hydroelectric power and water for domestic and industrial use (Anyah et al. 2006). The United Nations Economic Commission for Africa (UNECA) has identified Eastern Africa to have the fastest-growing economy on the continent, with an annual average growth rate of 6.7% between 2013 and 2017, more than double the African average. However, agriculture forms a large part of the GDP and productivity can fluctuate rapidly as a result of natural disasters, in particular flooding and droughts (UNECA 2019).

In addition to effects on food security and energy, severe weather has a more direct and immediate impact on fishermen and other users of the lake; boating accidents, many associated with severe weather and dangerous water currents, are thought to cause 5,000 fatalities on the lake every year (Semazzi 2011; Cannon et al. 2014). High winds associated with severe storms drive large waves on the lake, which can capsiz the small wooden canoes used by many fishermen. Although warnings are issued to fishermen, some are ignored due to a lack of trust in forecasts following many false alarms (Muwembe 2012). Many fishermen would rather take the risk of getting caught in a storm than miss a day's work.

The Lake Victoria basin also has a heightened vulnerability to the effects of climate change. Based on a high-resolution climate projection, Thiery et al. (2016) suggest that, under scenario RCP8.5, the increase of precipitation extremes over the lake itself will be approximately twice as large compared to the surroundings by 2100. The same model predicts a decrease in the annual mean precipitation over the lake of 6%. However, there is much uncertainty surrounding climate projections in this region (Onyutha et al. 2016).

An added hazard is the fluctuating water levels of Lake Victoria, which have varied over  $\sim 3$  m during the last 65 years (Vanderkelen et al. 2018). This fluctuation is mainly attributed to changes in precipitation, which is the largest and most variable term in the water budget of Lake Victoria (Yin and Nicholson 1998; Vanderkelen et al. 2018).

Compared to West Africa, the meteorology of East Africa is less well understood. Many modelling and field studies have recently taken place in West Africa (e.g. African Monsoon Multidisciplinary Analysis (AMMA), Cascade, Fennec, DACCIWA) where there is now improved understanding of the long-lived mesoscale convective systems (MCSs) which propagate westward across the region. Differences in the properties of MCSs between East and West Africa have not been well-studied, but it is anticipated that storms are more localised and shorter-lived in East Africa, with initiation and propagation controlled by sea breezes, lake and land breezes, and orography.

Being able to accurately predict severe convective events is of vital importance in East Africa in order to protect lives and livelihoods. However, forecasting convection remains one of the grand challenges of atmospheric science. Models struggle to correctly simulate convection due to the small scales involved, as well as an incomplete theoretical understanding of convection. Simulation of convection is a particular problem for coarse global models with parametrised convection. In the last decade, the development of higher-resolution models, and improved supercomputing power to run such models, has

provided a step-change in the ability of models to forecast convection (Clark et al. 2016). This development offers the opportunity for significant improvements to forecasts in East Africa.

Although accurate forecasts in East Africa are limited by forecast model performance, an incomplete understanding of the meteorology of the region also contributes to poor forecast skill. Both of these factors will be addressed in this thesis.

## 2. Background

### 2.a. Challenge of forecasting convection

Houze (1997) describes atmospheric convection as ‘the overturning of the atmosphere that is required to neutralize the vertical distribution of moist static energy’. Without convection, the large-scale Hadley and Walker circulations would be insufficient to redistribute the excess moist energy at the surface in the tropics (Lafore et al. 2017). The majority of global precipitation, especially in the tropics, occurs as a result of convection (e.g. Houze et al. 2015; Lafore et al. 2017). By determining the distribution of cloud and water vapour in the tropics, convection also plays a key role in the Earth’s radiative budget (Lafore et al. 2017). Convection may be organised across many scales: from a single cumulus or cumulonimbus cloud; to MCSs; to global-scale phenomena such as the Madden–Julian oscillation (MJO) (Madden and Julian 1971, 1972), convectively coupled equatorial waves (Kiladis et al. 2009) and the Intertropical Convergence Zone (ITCZ) (Moncrieff et al. 2012). Despite its importance and prevalence, forecasting tropical convection remains a huge challenge for numerical weather prediction (NWP), in part due to the wide range of scales with which it interacts.

The idea of using numerical models to predict weather originated from the pioneering work of Richardson (1922). NWP models have now been in use operationally for over 50 years (Clark et al. 2016). Bauer et al. (2015) go as far as to suggest that the computational problem of NWP is ‘comparable to the simulation of the human brain and of the evolution of the early Universe’. The first numerical models were used to predict synoptic scales, or larger, at which hydrostatic balance holds (gravitational force is balanced by the pressure-gradient force in the vertical). These models were based on the hydrostatic primitive equations (HPEs), consisting of the equation of motion (in 3D, where the vertical component is replaced by the hydrostatic equation), continuity equation, equation of state and the First Law of Thermodynamics (Haltiner and Williams

1980). These equations were generally used as the dynamical core of models between 1960 and 1990 (Arakawa 2002a). However, the hydrostatic approximation only holds where the vertical scale is much smaller than the horizontal scale (Daley 1988). When attempting to model convection, for which updrafts on the scale of hundreds of metres can be important, hydrostatic balance is no longer a reasonable assumption.

In the 1970s and 1980s, non-hydrostatic models were developed to simulate single convective clouds and mesoscale phenomena, such as sea breezes and mountain and valley flows (see Clark et al. 2016 and references therein). The Met Office began to experiment with non-hydrostatic models, based on the model developed by Tapp and White (1976), but available computational power was insufficient to run such models operationally until the early 2000s. In 2003, the non-hydrostatic dynamical core New Dynamics (ND) (Davies et al. 2005) was incorporated into the operational Met Office Unified Model (MetUM).

Even with the hydrostatic balance assumption removed, models with large horizontal grid-spacing must parametrise small-scale phenomena such as convection. Arakawa (2002b) defines cumulus parametrisation as ‘the problem of formulation of the collective effects of cumulus convection in terms of large-scale variables’. In other words, although convection itself cannot be modelled, parametrisation attempts to predict the effects of convection that would occur as a result of the large-scale environment. The ability of the parametrisation to predict precipitation is key, but a parametrisation scheme must also adjust the atmosphere to feed back onto the large scale and interact with the boundary layer (Arakawa 2002b).

Parametrised models exhibit a wide variety of issues affecting the prediction of convective rainfall, unsurprising given the complex task to which they are assigned. In particular, convective activity tends to peak at midday or earlier in parametrised models, rather than in the late afternoon as observed (Dai et al. 1999; Yang and Slingo 2001; Bechtold et al. 2004). Guichard et al. (2004) show that this lag arises due to the inability of such models to simulate the different stages of convective growth (dry, shallow and deep) and their interactions, described in Chaboureau et al. (2004). Instead, the models tend to produce deep convection too early by responding instantaneously to the low-level temperature and convective available potential energy (CAPE). Guichard et al. (2004) suggest that a variety of parametrisations would be necessary to correctly represent each stage of convective growth. Chaboureau et al. (2004) propose that inadequacies in the parametrisation of the surface and boundary layer also contribute to the offset in timing, in addition to the parametrisation of convection.

Another shortcoming of parametrised models is that, due to their coarse resolution, they are unable to produce realistic storm structures. The distribution of rainfall rates in parametrised models is skewed; too much light rain, or ‘drizzle’, is predicted (over large areas), whereas localised, high rainfall intensities cannot be simulated (Sun et al. 2006; Dai 2006; Stephens et al. 2010). Coarse resolution also requires smoothing to be applied to orography, therefore mountain and valley flows cannot be accurately represented (Clark et al. 2016). Given the myriad of issues with convective parametrisations, improving parametrisations is an active area of research, for example within the ParaCon project (<https://www.metoffice.gov.uk/research/approach/collaboration/paracon>).

As computing power has increased in recent decades, the horizontal grid-spacing of models has correspondingly decreased. Finer structures can now be resolved and convection may be simulated explicitly (i.e. the need for parametrisations is removed). These high-resolution models are known as convection-permitting (CP), and are discussed in the next section. However, it is also worth noting that, despite ongoing model improvements and additional computing resources, the prediction of convection is fundamentally limited by the fact that the atmosphere is governed by non-linear, differential equations (Lorenz 1969). As a result, two forecasts from the same model which are initialised with slightly different initial conditions may diverge rapidly from one another. Analogously, uncertainty in the initial conditions of a model may lead to the divergence of a forecast from the truth. The growth rate of errors on convective scales are approximately ten times greater than for synoptic scales, in part because the perturbations depend on convective instabilities rather than baroclinic instabilities (Hohenegger and Schär 2007). Small-scale errors may then influence larger scales.

## 2.b. Convection-permitting models

Over the last few decades there has been a step-change in the modelling of convection, with the introduction of CP models in which convective structures are explicitly simulated, rather than being parametrised (Clark et al. 2016). Models with explicit convection have been shown to shift the peak convection to later in the day, more in line with observations (e.g. Lean et al. (2008) for the UK and Birch et al. (2014b) for West Africa). Whilst there are obvious benefits of an improved diurnal cycle, work by Marsham et al. (2013) and Birch et al. (2014b) as part of the Cascade project has shown that correctly modelling the diurnal cycle feeds back positively onto the large scale. Because convection occurs too early in parametrised models, radiative heating

at the surface is reduced and latent heating occurs too early in the day. Over West Africa, this results in an unrealistic strengthening of the pressure gradient between the Saharan Heat Low (SHL) and the high pressure in the Gulf of Guinea at night, affecting the northward advection of moisture associated with the monsoon. It is possible that similar effects occur elsewhere in the world, and so it is important that the diurnal cycle is well represented in models. Additional benefits of CP models are discussed in chapter 2, section 1.

Despite the step-change in forecasting capability, CP modelling still faces significant challenges. Several studies note that, although the representation of convection and its organisation is improved, CP models offer little to no improvement in forecast skill at the grid-scale (e.g. Done et al. 2004; Weisman et al. 2008; Weusthoff et al. 2010). Much of this issue is likely related to the high dependence of regional models on their global driving models (Guichard et al. 2010; Vié et al. 2011; Melhauser and Zhang 2012; Birch et al. 2013; Schumacher et al. 2013; Luo and Chen 2015). In many cases, poor driving model performance can be attributed to a lack of observations for data assimilation, especially over Africa.

Many national meteorological services around the world, including several in East Africa, now make use of CP models in their operational forecasts. A comprehensive list of these countries, and their respective models, is given in chapter two, section 1. However, a recent addition to this growing list is the South African Weather Service (SAWS), which runs CP MetUM with 4.4 and 1.5 km horizontal grid-spacing. Operational radar in the region (Terblanche et al. 2001) has allowed detailed verification of the models. The CP forecasts were shown to exhibit an improved diurnal cycle and more realistic distribution of rainfall rates compared to parametrised MetUM (Stein et al. 2019). Keat et al. (2019) also found good agreement between the distribution of storm sizes in the model and observations. However, both studies noted that deep cloud was often initiated several hours too early in the models, although this did not seem to affect the diurnal cycle of precipitation.

Many existing evaluation studies of CP models have been performed for short time periods, such as for case studies, or over just one season. Mittermaier et al. (2013) performed an assessment of CP MetUM in the UK over a 41 month period. However, given the different nature of convective organisation in the two regions, it is anticipated that model performance will greatly differ between the tropics and midlatitudes. It is important to study the behaviour of these models over longer time periods in order to understand seasonal dependence, and build more robust statistics about their

performance.

In addition to their contribution to NWP, CP models are increasingly being used in climate projections (Prein et al. 2015). The improved realism of rainfall in CP models allows more accurate simulation of extreme precipitation events. In particular, Kendon et al. (2014) used a CP MetUM climate configuration to show a projected increase in the highest-intensity convective summer rainfall over the UK. A parametrised configuration was unable to replicate this result. As part of the Future Climate for Africa (FCFA) Improving Model Processes for African Climate (IMPALA) project, a pan-Africa 4.4 km convection-permitting simulation (CP4) has been run for present day and  $\sim 2100$  under the RCP8.5 emissions scenario (Stratton et al. 2018). A parametrised 25 km simulation was also run alongside this. The simulations are currently being used to investigate future predictions of precipitation and heatwaves. A good understanding of the performance of CP models and their ability to represent various small-scale processes is vital to aid interpretation of future climate CP predictions.

### **2.c. Lake–land breeze circulations**

The differing heat capacities between land and water drive breezes between land and water bodies. These flows are commonly found along coastlines and are referred to as the sea breeze. However, such flows may also occur around inland bodies of water as a lake breeze. Since water bodies warm more slowly than land in response to shortwave solar radiation, air over water warms more slowly than air over land during the day. Figures 1a-d from Pielke (2015) (P15) show an idealised formation of a sea breeze. During the early morning, before daytime heating has begun, pressure surfaces above the water and land are flat (P15, Fig. 1a). Turbulent mixing in an unstable boundary layer, and the volumetric expansion of the air as it is heated, mix mass upward over the land (P15, Fig. 1b). The upward transport of air causes an increase in pressure a short distance above the surface, and a density gradient between the land and water at this height. In response to the density gradient, a flow of air from the land to water is induced. The loss of mass at this height over land leads to lower pressure at the surface over land. Onshore flow—the sea breeze—occurs as a result of the pressure gradient induced at the surface, with a maximum during the afternoon (P15, Fig. 1d). As the cooler air over the water flows onshore, the temperature gradient is advected onshore (Pielke 2015).



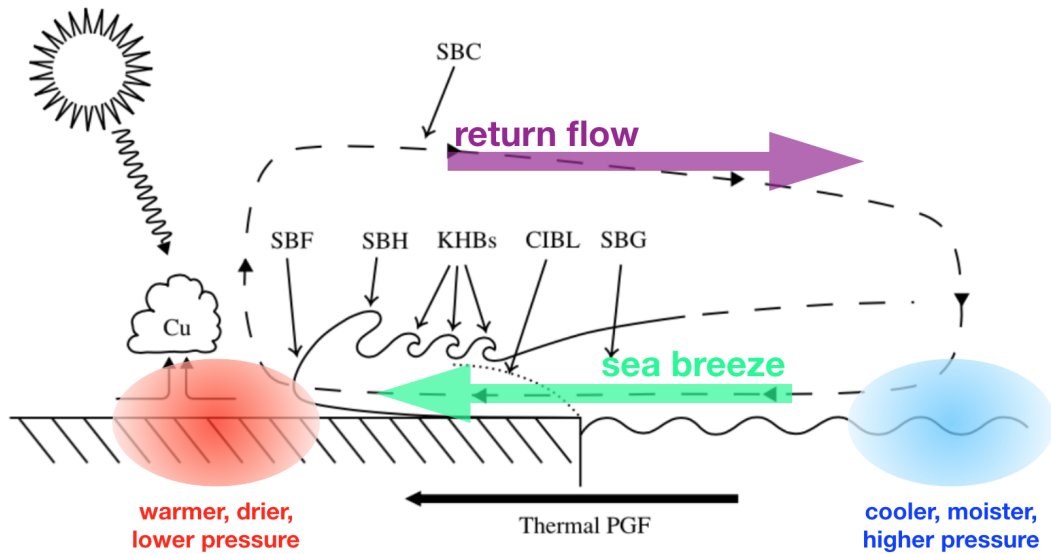


Figure 1.2.: Figure 1 from Miller et al. (2003) (adapted) showing the sea breeze system, including; sea breeze circulation (SBC), sea breeze gravity current (SBG), sea breeze front (SBF), sea breeze head (SBH), Kelvin-Helmholtz billows (KHBs), and the convective internal boundary layer (CIBL). [© 2003 by the American Geophysical Union. Published by John Wiley and Sons.]

Tijm and Van Delden (1999) argue that sound waves (which do not obey hydrostatic balance) are responsible for generating the pressure gradient. As the air over land warms and expands diabatically, sound waves are generated, propagating in all directions. Vertically propagating waves increase pressure throughout the atmospheric column over land. According to Tijm and Van Delden (1999), horizontally propagating sound waves are responsible for inducing the horizontal density gradient, by reducing surface pressure over land and increasing surface pressure over the sea.

After sunset, land loses heat by radiative cooling more quickly than water, such that air above the water is warmer than above the lake. The cooler air over land sinks, increasing the density near the surface (Pielke 2015). Because the pressure surface slightly above the land surface is lowered, an onshore wind is induced at this level, which transports mass from above the ocean to over the land (P15, Fig. 1f). Low pressure is induced above the water surface as a result of the loss of mass, creating a pressure gradient and offshore flow—the land breeze—at the surface (P15, Fig. 1g).

Although the description of sea and land breeze formation above implies that the sea and land breezes form in response to the formation of an upper-level flow, this upper-

level flow is often referred to as the ‘return flow’. The surface and upper flows form a closed circulation which is shown in Fig. 1.2 for a sea-breeze cell. The sea breeze flow occurs over a depth of  $\sim 500\text{--}1000$  m (Parker et al. 2017), and the return flow between  $\sim 1000\text{--}2000$  m above the sea breeze (e.g. Moroz 1967; Lyons 1972; Keen and Lyons 1978).

Wexler (1946) notes that sea, lake and land breezes are most common under clear sky conditions. In the tropics, such circulations may occur almost daily during dry seasons and are frequent, but tend to be weaker, during the wet seasons. In the midlatitudes, the circulations are most common in summer, but can still occur in spring and autumn. A study of sea breezes across the south coast of England by Simpson et al. (1977) showed that the sea breeze could penetrate over 100 km inland on some days, but remain offshore on others. These differences were attributed to differences in the strength of the density gradient between land and sea air. Complementing this, (Pielke and Segal 1986) showed that the strength of a sea or land breeze is governed by the strength of the temperature gradient and the depth of the temperature perturbation (Pielke and Segal 1986).

Synoptic-scale flows can also affect the characteristics of these breezes (Sills 1998 and references therein). Simulations by Estoque (1962) showed that sea breezes were weakened during onshore synoptic flow, whereas offshore flow strengthened the sea breeze by enhancing the temperature gradient. Further work by Arritt (1993) showed that light offshore winds produced the strongest sea-breeze circulation, whereas onshore winds great than  $3\text{ ms}^{-1}$  suppressed the circulation. These results from simulations are also supported by observational evidence in Simpson et al. (1977) and Simpson (1994).

Sea/lake and land breezes are often described as density currents (e.g. Simpson 1969; Simpson and Britter 1980). Density (or gravity) currents are horizontal flows which develop in response to a horizontal density gradient. In the atmosphere, density gradients are likely a result of temperature differences (Simpson 1999), although air masses with different molecular make-ups can also affect density. An important example is moist air; because the molecular weight of water is lower than the average molecular weight of air, an air mass with a greater mixing ratio will be less dense than a drier air mass at the same temperature and pressure. A useful way to identify density currents is with virtual potential temperature  $\theta_v$ , first introduced by Lilly (1968). Virtual potential temperature takes into account the temperature and moisture content of air and can

be used as a proxy for buoyancy. It is defined as:

$$\theta_v = \theta(1 + 0.61r - r_L), \quad (1.1)$$

where  $\theta$  is the potential temperature,  $r$  is the mixing ratio of water vapour, and  $r_L$  is the mixing ratio of liquid water. Low  $\theta_v$  indicates dense air, whereas high  $\theta_v$  indicates buoyant air.

In the case of sea and land breezes, cool (low  $\theta_v$ ) air of the ‘density current’ moves out over a relatively warmer surface (land or water respectively), which heats the air from below, and erodes the front of the density current (Parker et al. 2017). In the day, when the land warms rapidly and a convective boundary layer forms, intense vertical mixing dominates the dynamics, mixing air in the sea/lake breeze with air over land (Linden and Simpson 1986). Overall, these processes act to smooth the temperature gradient, whereas a ‘pure’ density current is defined by a sharp gradient across its front. However, when turbulence decreases (e.g. in the late afternoon/evening), a density current with a strong temperature, humidity and wind gradient across its leading edge can form at the sea/lake breeze front, and propagate much further inland (Linden and Simpson 1986; Parker et al. 2017).

Various observational studies have shown that sea breezes tend to be stronger than their land counterparts, even if the magnitude of the land–sea temperature contrast is the same during the day and night (Mak and Walsh 1976). Mak and Walsh (1976) used a simple theoretical model to attribute the differences to increased stability in the boundary layer overnight compared to the daytime. Stability inhibits vertical motion and limits the depth of the nocturnal boundary layer, weakening the whole land breeze circulation (Mak and Walsh 1976; Parker et al. 2017). Wexler (1946) notes that mountainous terrain adjacent to the shore enhances the strength and frequency of local winds, as the lake and land breezes combine with upslope and downslope winds respectively.

The front of a sea (or lake or land) breeze is generally associated with sharp gradients in temperature, moisture and wind (Miller et al. 2003). A distinguishing feature of a density current is that the speed of the flow behind the front exceeds the propagation speed of the front, leading to convergence at the leading edge (Parker et al. 2017). In order to conserve mass, vertical motion must occur, along with mixing with environmental air. Vertical motion can lead to the development of deep convection at the sea breeze front (e.g. Byers and Rodebush 1948; Pielke 1974). Fovell (2005) suggest that

the sea breeze induces widespread but gentle lifting ahead of it. This lifting moistens and destabilises the atmosphere, and forms a moist tongue (sometimes partly visible as a shelf cloud) ahead of the leading edge at the height of the sea-breeze cell lid. In the Fovell (2005) study, an additional trigger—in this case horizontal convective rolls—is required to initiate the deep convection. Convective initiation can also occur around irregular coastlines as a result of regions of convergence between sea breezes from differently oriented coastlines (e.g. Laird et al. 1995). In the region of subsidence over water, clear skies are likely (Pielke and Segal 1986).

Overnight, convection can occur at the land breeze front. Even though the strength of the land breeze is likely weaker than the sea breeze, high moisture content and instability, due to the water surface remaining warm, can result in vertical motion and cloud formation (Pielke and Segal 1986). Neumann and Mahrer (1975) and Physick (1976) performed theoretical studies of lake–land breeze circulations over circular lakes and showed strong divergence over the centre of the lake during the day. At night, the land breeze from all directions was shown to converge over the centre of the lake, and produce vertical motion. From observations of Lake Michigan, Moroz and Hewson (1966) suggested that the return flow over lakes is more pronounced than for sea-breeze circulations. These processes are discussed in more detail with specific regard to Lake Victoria in section 3.d and investigated in chapters 3 and 4.

Lake breeze circulations around the world have been studied via field campaigns and modelling. In particular, there is a wealth of research focused on Lake Michigan and other Great Lakes in North America (Lyons 1972; Lyons and Olsson 1972; Keen and Lyons 1978; Sills et al. 2011). Relative to Lake Victoria, the frequency of occurrence of lake-land breeze circulations is reduced in this region due to the strong influence of synoptic weather conditions. In general, discussion of land breezes is neglected in the literature, especially over lakes where two land breezes may oppose one another from opposite shores. However, the role of convergent land breezes in the formation of Great Lake snow storms has been shown by Passarelli and Braham (1981) and Ballentine (1982).

There are a number of studies on the effect of sea breezes and their collisions with other density currents. Goler and Reeder (2004) performed highly-idealised simulations to model the formation of the Morning Glory by the collision of opposing sea breezes over the Cape York Peninsula, in northern Australia. The simulation showed that a prevailing easterly wind deepens the east-coast sea breeze, causing it to override the west-coast sea breeze. Waves are excited and propagate on the west-coast sea breeze.

The Morning Glory is produced when the phase speed of the east-coast sea breeze and waves are similar, such that the waves are amplified. However, their model did not include any moist processes and the water temperature was kept fixed. To some extent, the exclusion of moist processes allows the experiment to be analogous to colliding land breezes during the early morning. However, it is anticipated that the situation would differ greatly with the inclusion of a moisture source below the collision/waves. Kingsmill and Crook (2003) observed the collision of gust fronts and sea breezes in Florida, showing that collisions could result in secondary density currents or bores (depending on a variety of factors) which could go on to initiate further convection. Birch and Reeder (2013) and Birch et al. (2014a) showed that wave-cloud lines in Australia and the Arabian Sea respectively were caused by the overriding of the sea breeze with accelerated synoptic flow from the land overnight.

This section has highlighted the effects of lake–land breeze circulations on temperatures, humidity, wind, cloudiness, and precipitation in the region of a lake. The specific impact of Lake Victoria on the regional climate of the Lake Victoria basin is discussed in detail in section 3.d.

### **3. East Africa**

#### **3.a. Climatology of East Africa**

East Africa lies at the intersection of two distinct monsoon systems; the northeast (NE) and southeast (SE) monsoons (Okoola 1999). Figure 1.3 shows the annual cycle of 850 hPa winds. The strongest winds occur during boreal summer, when onshore southeasterlies in the southern hemisphere swing to the east to become southwesterlies in the northern hemisphere, flowing out across the Arabian Sea. These southwesterlies constitute the Somali Jet (Findlater 1977). Aerial observations have shown that the core of the jet is situated at approximately 40°E, at altitudes of 1–2 km. This circulation pattern is strongest in July and August and begins to weaken in September. Flow is southerly over Kenya, situated on the equator. The southerly flow pattern does not extend far into Ethiopia, as it is diverted by the Ethiopian highlands.

During boreal autumn, the wind pattern begins to reverse and, by November, northeasterlies replace the southwesterlies in the northern hemisphere. These northeasterlies do not penetrate into the southern hemisphere until December, replacing the onshore southeasterlies in Tanzania. The location at which the wind reverses from easterly to

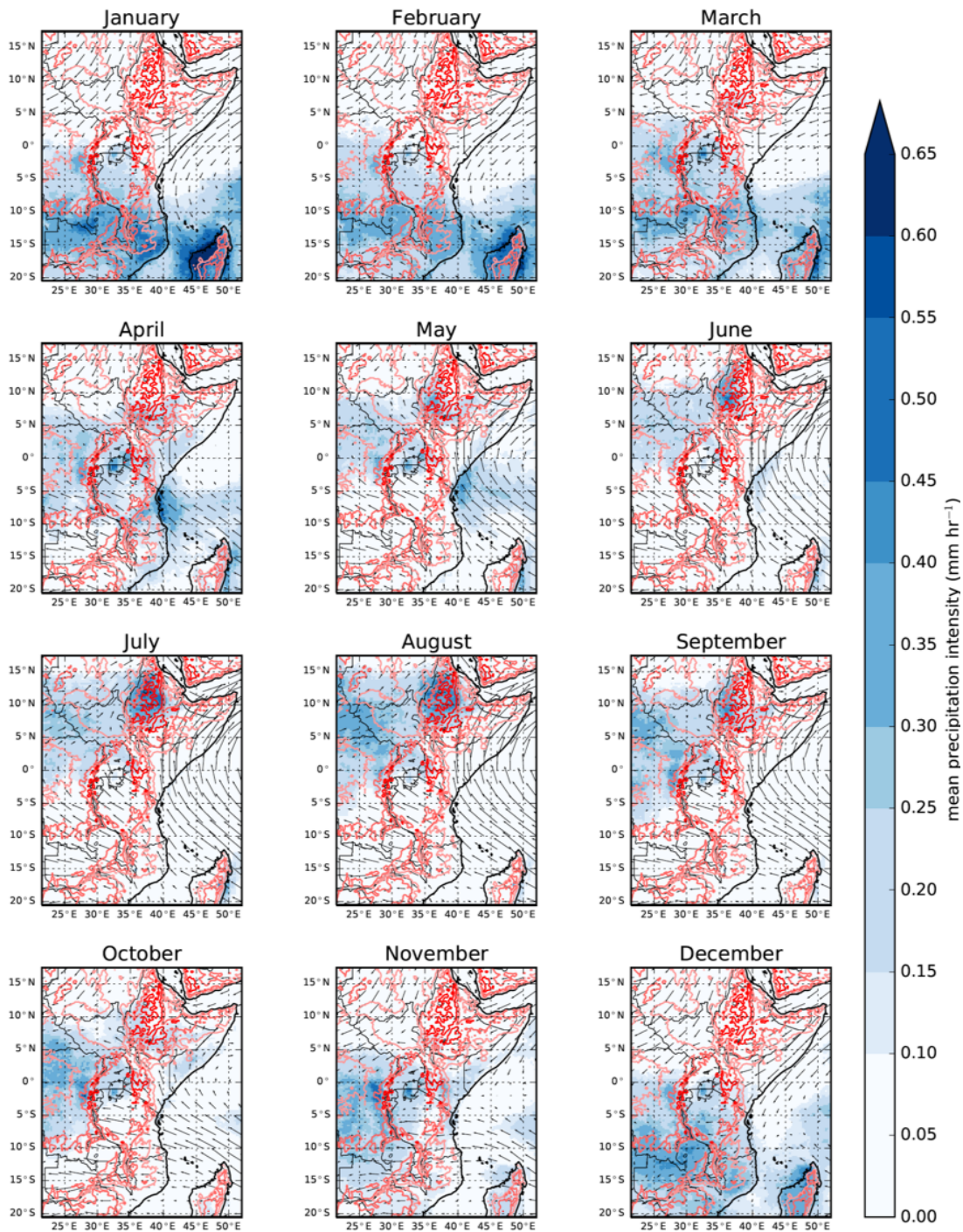


Figure 1.3.: Mean monthly rainfall (shading) between 1979–2015 from TRMM 3B43 (Liu et al. 2012) and 850 hPa wind (vectors) from ERA-Interim (Dee et al. 2011) between 1998–2014. Red contours show 500, 1000 and 2000 m orography.

westerly shifts south away from the equator, to about 6°S. These strong northeasterlies persist until February and begin to weaken in March. The winds are calmest in April, and onshore winds are once again southeasterly in Tanzania.

During the transitions to and from the NE monsoon in boreal winter and SE monsoon in boreal summer, convergence occurs over the region; this is the ITCZ, which is associated with intense convection and rainfall. The progression of the rainfall northward and southward throughout the year is shown in maps of the mean monthly rainfall (Fig. 1.3). As the ITCZ travels north and south over the course of a year, much of the region experiences a bimodal annual cycle of rainfall, with a wet season during both boreal spring and autumn (Okoola 1998). Average dates of onset and cessation (according to the method of Dunning et al. 2016) of these two seasons are shown in Figs. 1.4a-d. Mean annual cycles of precipitation for various stations over the region are shown in Fig. 1.4e, highlighting the bimodal annual cycle over central East Africa, as well as a unimodal cycle in the far north of the regions, centred on boreal summer.

In the regions experiencing bimodal rainfall (between black dashed lines in Figs. 1.4a-d), the two rainfall seasons are known as the ‘long rains’ (March–May, MAM) and the ‘short rains’ (October–December, OND). In a study of rainfall in Kenya and northern Tanzania between 1958 and 1987, Camberlin et al. (2009) actually found that the average duration of the ‘short rains’ exceeds that of the ‘long rains’ by around a week. However, the variability in the duration of the short rains is much greater, and the mean total amount of precipitation during the long rains exceeds that of the short rains by over 50%. In addition, the mean fraction of days on which it rains within the season is over 40% greater, and the intensity of rain per day is over 30% greater during the long rains. The wet seasons are not continuously wet, but are composed of alternating wet and dry periods.

Given that East Africa is located in the deep tropics, a similar distance from the equator as the South Asian, West African and Australian monsoons, it might be expected that the region should experience a ‘wet’ climate, with a high total annual rainfall amount and an annual precipitation cycle typical of a monsoonal climate. However, the climate in East Africa is actually semiarid/arid, with much of the region receiving less than 2 mm day<sup>-1</sup> (Yang et al. 2015).

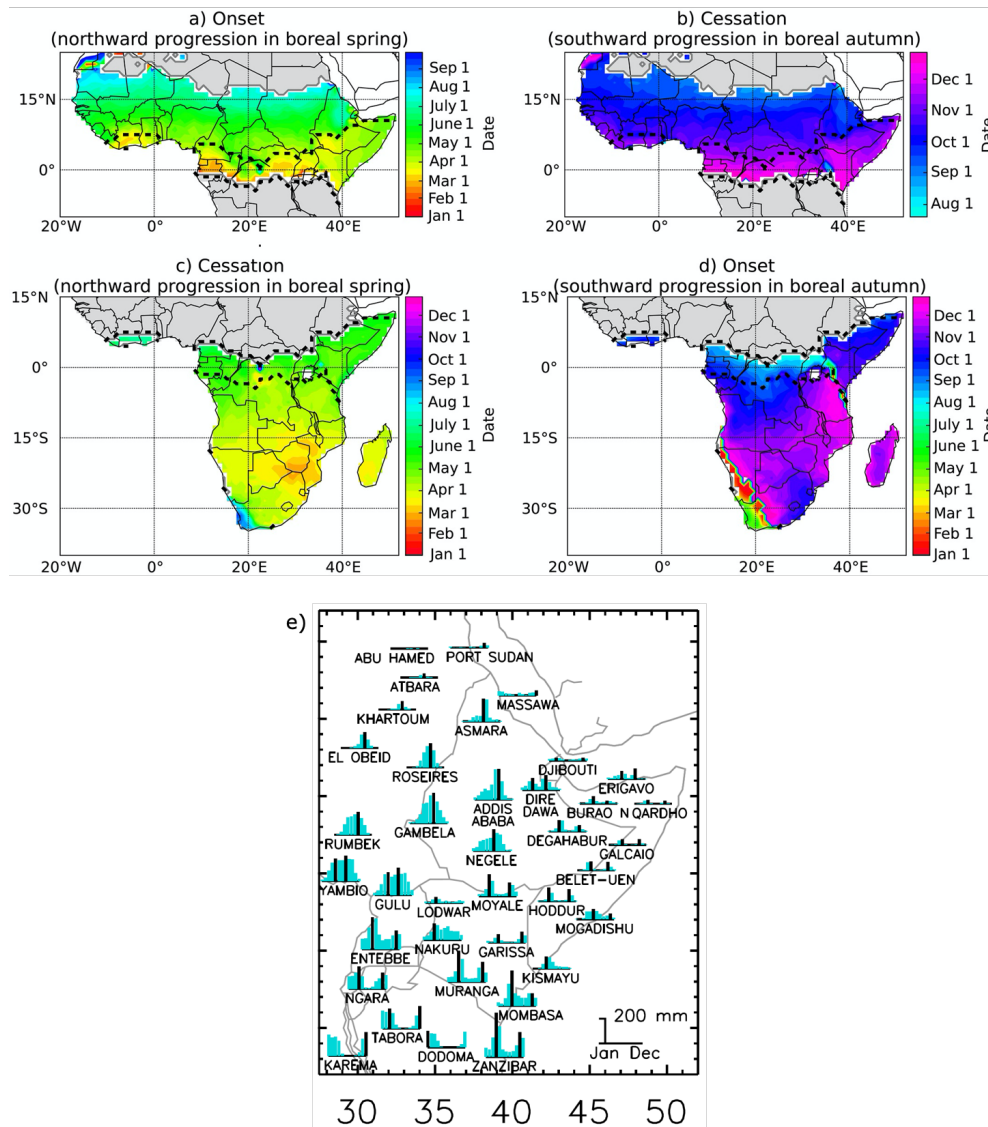


Figure 1.4.: (a-d) Figure 5 from Dunning et al. (2016). Southward and northward progression of onset and cessation across the annual/biannual boundaries. (a) The northward progression of onset in boreal spring from onset of the first/long rains into onset of the West African Monsoon. (b) The southward progression of cessation in boreal autumn from the cessation of the West African Monsoon into the end of the second/short rains. (d) The southward progression of onset in boreal autumn, from the onset of the second/short rains into the onset of the annual rains over southern Africa. (c) The northward progression of cessation in boreal spring from the end of the annual rains over southern Africa into the end of the first/long rains. These were computed using the GPCP data set over 1998–2013. Grey indicates regions not considered for these plots. [© 2016. Dunning et al. Under the terms of the Creative Commons Attribution License. Published by John Wiley and Sons.] (e) Taken from Fig. 3 from Nicholson (2017). Schematic of the seasonal cycle at select stations. Black vertical lines/lines at each station indicate the month or months of maximum rainfall at the indicated stations. [© 2017. American Geophysical Union. All Rights Reserved. Published by John Wiley and Sons.]



### **3.b. Seasonal and climate variability**

#### **3.b.1. Variability of the Long and Short Rains**

The vast majority of research concerning weather and climate in East Africa has investigated seasonal and climate timescales. Much research has been dedicated to the different characteristics of the long and short rains which exhibit variability in many different ways, including their onset and cessation dates, total precipitation amount, number of wet days and daily rainfall intensity. The variability of these factors is often interrelated; Camberlin et al. (2009) found that, during the short rains, the onset is related to the cessation, number of wet days and daily intensity. However, during the long rains these factors are unrelated, making the long rains less predictable than the short rains. Camberlin et al. (2009) also showed that, for both wet seasons, the onset date of the rains varies more than the cessation date.

The short rains are known to have greater interannual variability than the long rains (Ogallo 1989; Nicholson 1996) and most floods and droughts from the past 50 years have been associated with the short rains Hastenrath et al. (2011). Hastenrath (2007) and Hastenrath et al. (2007, 2010, 2011) show that a ‘Walker-like’ circulation exists over the Indian Ocean during the short rains. This consists of an ascending branch over Indonesia, a subsiding branch over East Africa, and surface westerly winds across the Indian Ocean. Easterly winds in the upper troposphere complete the circulation. When the westerly surface winds are strong, subsidence over East Africa is increased, and rainfall is suppressed. The reverse is true when the surface winds are weaker. This circulation does not exist during the long rains, since the upper-tropospheric winds are westerly at this time of year. Again, this decreases the predictability of the long rains.

#### **3.b.2. MJO**

The Madden-Julian Oscillation (MJO) is a 30-60 day oscillation in pressure, wind and convection over the Indian and West Pacific Oceans, first formalised by Madden and Julian (1971, 1972). It is the leading mode of intraseasonal variability in the tropics and is manifested as convective cloud clusters propagating eastward across the Equator. The MJO is described by its ‘phase’, which gives an indication of the location of its centre of active convection.

The influence of the MJO on East African rainfall, in particular on its intraseasonal

variability, was investigated by Pohl and Camberlin (2006a). They noted that, even within the wet seasons, alternating wet and dry spells with a period of  $\sim 20$ – $50$  days (note similar time period to that of the MJO) would occur. In addition, these wetter periods were found to coincide with anomalous easterly winds along the equator. By carrying out composite analysis of the atmospheric circulation during wet and dry periods, Pohl and Camberlin (2006a) found that a moderate to strong MJO event influences rainfall over the Highlands ( $\sim 36^\circ\text{E}$ ,  $1^\circ\text{S}$ , Fig. 1.1) and the coast in very different ways. In the highlands, rainfall is anomalously high during phases 2–4 of the MJO, when the active cluster of MJO convection is over the Indian Ocean. Using OLR observations, Pohl and Camberlin (2006a) attributed this rain to increased deep convection over the region, and the anomalous advection of moisture into the region from the Atlantic. Over the coast, rainfall is anomalously high during phases 7–8 of an MJO event, when the active cluster of MJO convection is over the central Pacific Ocean. During this time, surface easterly trade winds are anomalously strong due to the pressure gradient between the Indian and Atlantic Ocean. As a result, moisture advection from the Indian Ocean is increased. However, this period of enhanced precipitation differs from that over the highlands; OLR records show that the enhanced precipitation is not of deep convective origin, but is attributed to shallow convection and stratiform clouds.

Similar work was carried out by Omeny et al. (2008), who also found the east/west split in response to the MJO phase. Further work by Pohl and Camberlin (2006b) suggests that, between 1979 and 1995, fluctuations in the amplitude of the MJO explained 44% of the variance in the long rains. In particular, years with a large MJO amplitude during boreal spring recorded earlier onsets of the rains, and increased total precipitation during the season. More recent work by Hogan et al. (2015) emphasises that the influence of the MJO is most prominent during the wet seasons.

### **3.b.3. ENSO and IOD**

The El Niño Southern Oscillation (ENSO) also modulates rainfall in East Africa. Ogallo (1988) found that the short rains are stronger in El Niño years, although rainfall is decreased between July and September. Indeje et al. (2000) showed that different regions have different responses to the different phases of ENSO (i.e. positive or negative). Indeed, 8 different homogeneous rainfall regions were identified. Mutai and Ward (2000) suggest that the different responses relate to the varied nature of topography within the region. During El Niño years, Moron et al. (2013) showed that the duration of the long rains increases, but the intensity is reduced. The inverse is true during La Niña

years.

The Indian Ocean Dipole (IOD), first described by Saji et al. (1999), is a mode of the Indian Ocean, during which the west of the ocean is anomalously warm and the east is anomalously cold. This is the ‘positive’ phase of the IOD. Temperature anomalies are reversed in the ‘negative’ phase (Marchant et al. 2007). There is much discussion as to the relationship between ENSO and the IOD, and whether the IOD can be considered as independent of ENSO. A detailed review of arguments for and against this is given in Marchant et al. (2007). It should be noted that strong ENSO events have been shown to ‘set up’ IOD events (Hastenrath et al. 1993; Black et al. 2003). Black et al. (2003) found that it is only strong IOD events (lasting several months) which are shown to have an effect on the short rains over East Africa. Composites of wind over the region during strong positive IOD events show that equatorial westerlies are weakened, which affects the Walker-like circulation of the Indian Ocean described by Hastenrath (2007).

### **3.b.4. Seasonal prediction**

The teleconnections discussed above, as well as sea surface temperatures (SSTs), can be used in statistical seasonal predictions (e.g. Farmer 1988; Mutai et al. 1998; Nicholson 2014, 2015b). At present, seasonal forecasts are produced at the Greater Horn of Africa Climate Outlook Forum (GHACOF), one of several World Meteorological Organization (WMO) supported Regional Climate Outlook Forums (RCOFs) (Ogallo et al. 2008). GHACOF events are held prior to the long and short rains and summer rainfall season (June–September in the north of the region, Fig. 1.4). Forecasters from various meteorological centres in the region come together to produce a consensus forecast for the season, using a range of statistical and dynamical models, as well as local knowledge. Walker et al. (2019) compared the skill of the tercile forecasts produced by GHACOF to dynamical forecasts produced by the UK Met Office Global Seasonal Forecast System 5 (GloSea5), a coupled ocean-atmosphere model (MacLachlan et al. 2015). In general, GloSea5 outperformed the GHACOF forecasts; for the latter, the likelihood of near-normal rainfall was over-predicted, likely due to an under-confidence in predicting above or below normal rainfall. GloSea5 showed positive skill for the prediction of the short rains, but almost no skill during the long rains.

### 3.c. Climate predictions for East Africa

The effect of anthropogenic climate change on the future climate of East Africa was investigated by Shongwe et al. (2011) using Coupled Model Intercomparison Project (CMIP) General Circulation Models (GCMs). They found that CMIP (phase 3) models strongly suggest that mean precipitation rates will increase over the region, along with the intensity of the highest rainfall events. The models also predict less severe drought. However, it is noted that the timescales over which changes to rainfall patterns may occur vary widely between models. Projections by CMIP5 models also indicate a general increase in precipitation, although variability between models is high (e.g. Otieno and Anyah 2013; Kent et al. 2015). Studies by Vizu and Cook (2012) and Cook and Vizu (2013) suggest a reduction in the long rains, causing the number of dry days to increase by 5–10% by 2060. This is attributed to increased rainfall in the Congo Basin, which causes increased moisture advection out of East Africa. The study also expects the short rains to be reduced in eastern Ethiopia and Somalia. These results somewhat contradict other studies, although Cook and Vizu (2013) do predict an increase in extreme wet events over the Ethiopian Highlands and over Kenya and Tanzania during the short rains, more in line with other studies.

In contradiction to the predicted wetting trend, studies looking at recent changes in the climate have shown a drying trend in the long rains. Williams and Funk (2011) show a decrease in the long rains since 1983, and Nicholson (2015a) describes prolonged periods of drought (interrupted by several devastating floods) since 2008. The decline in rainfall totals during the long rains has been linked to SSTs, in particular an increase in the gradient in SSTs between Indonesia and the central Pacific (Lyon and DeWitt 2012; Liebmann et al. 2014; Yang et al. 2014).

The contradiction between the recent drying trend and future climate predictions has been termed the ‘Eastern Africa Paradox’. (Rowell et al. 2015) attempted to reconcile the two trends by testing a variety of hypotheses. Poor quality of rainfall data (leading to identification of an incorrect trend) and land use changes were largely ruled out. Natural variability was shown to be important, but unlikely to be the only driver. (Rowell et al. 2015) identified that the effect of changing anthropogenic aerosol emissions may have affected SSTs, leading to the observed changes in precipitation. It was also suggested, but not investigated, that some physical processes are inadequately represented in CMIP models, which may cast doubt on the future projections.

Wainwright et al. (2019) showed that the historical decrease in long rains can be at-

tributed to a later onset and earlier cessation, not a change in the daily rainfall intensity. The delay in onset was attributed to warming of the ocean to the south of Eastern Africa, whereas the earlier cessation was attributed to increased southerlies, caused by lower surface pressure over Arabia and a warmer north Arabian sea. It is suggested that the predicted increase in precipitation is therefore mainly due to an increase in daily mean precipitation, rather than a change in season length (Dunning et al. 2018; Wainwright et al. 2019). Wainwright et al. (2019) note that further work is required to elucidate to what extent the changes driving a shorter long rains season are due to natural variability, or changing SSTs due to anthropogenic aerosol and carbon emissions.

### **3.d. Lake Victoria**

Lake Victoria, the second largest freshwater lake in the world, lies on the equator at the intersection of Kenya, Uganda and Tanzania (Fig. 1.1). The lake has a surface area of almost 69,000 km<sup>2</sup>, and approximate dimensions of 350 km north to south and 250 km east to west. For reference, the lake is larger than the two nearby countries of Rwanda and Burundi combined. The lake is located just over 1,100 m above sea level and lies between the eastern and western ridges of the East Africa Rift System. It is therefore surrounded by mountainous terrain, in particular the Kenyan highlands to the east, where peaks exceed 4,000 m.

With an average depth of 45 m, and a maximum depth of only 92 m, Lake Victoria is one of the shallowest lakes in East Africa (Song et al. 2004). A key feature of the bathymetry is that the west and southwest of the lake are much shallower, especially compared to the northwest (Talling 1966). As a result, it may be expected that lake surface temperatures (LSTs) would exhibit a temperature gradient, with warmer water in the southwest and cooler water in the northwest. However, remotely sensed observations from the 1999–2008 ARC-Lake dataset (MacCallum and Merchant 2012) presented in Fig. 5 in Thiery et al. (2015) actually show the opposite.

#### **3.d.1. The effect of Lake Victoria on the diurnal cycle of precipitation**

Lake Victoria greatly influences the local circulation, and consequently the diurnal cycle of rainfall, over its basin. Flohn and Fraedrich (1966) first noted the diurnal circulation system, consisting of a divergent lake breeze during the afternoon and convergent land

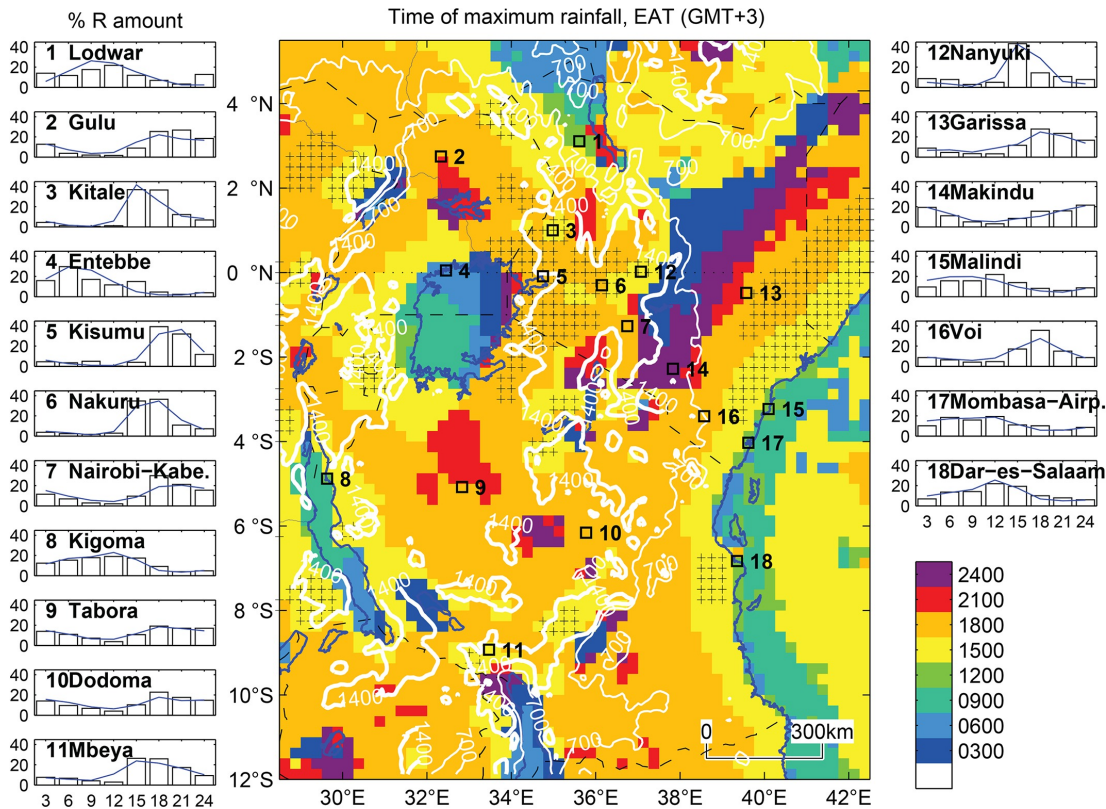


Figure 1.5.: Figure 4 from Camberlin et al. (2018). Three hour period during which maximum precipitation occurs, computed from TRMM data (1998–2014). 3 EAT stands for the period between 0130 and 0430 East African Time. White lines correspond to 700 and 1400 m contours. Pluses mark regions where over 35% of total precipitation was recorded in the wettest 3 hour period. Side panels: percentage of rainfall amounts in each 3-h period at stations located by squares, with bars showing TRMM data, and blue lines showing in situ rainfall records (6- to 26-year averages, depending on stations (Tomsett 1975)). [© 2017 Royal Meteorological Society. Published by John Wiley and Sons.]

breeze during the morning, the latter of which is responsible for around half of the annual precipitation over the lake (Flohn and Burkhardt 1985). In addition to convergence produced by the land breeze, convective activity over the lake is also enhanced by an unstable boundary layer as the atmosphere cools, but the lake surface temperature remains fairly constant (Flohn and Burkhardt 1985; Datta 1981).

Fig. 1.5, taken from Camberlin et al. (2018), shows a map of the time of day at which the maximum frequency of precipitation occurs (computed from the Tropical Rainfall Measuring Mission (TRMM) dataset). Over land, the maximum rainfall occurs at 1800 Local Time (LT), where  $LT = UTC + 3h$ . Over most of the lake, the rainfall maximum

occurs between 0600–0900 LT. Datta (1981) presented observations from weather stations close to the lake shore, showing that the lake breeze reaches maximum strength at 1500 LT, whereas the land breeze reaches maximum strength between 0600 and 0900 LT.

Easterly prevailing winds introduce an east–west asymmetry in the spatial pattern of wind and precipitation over the Lake Victoria basin (Datta 1981; Song et al. 2004). To the east of the lake, the rainfall maximum tends to occur between 1800 and 0100 LT, as the lake breeze converges with the prevailing easterlies. At this time, rainfall is suppressed over central and western parts of the lake by the westward advection of latent heat from convection in the east, which stabilises the atmosphere over the lake, and by subsidence related to the lake breeze return flow (between approximately 2000 and 5000 above mean sea level (MSL)) (Datta 1981; Ba and Nicholson 1998; Thiery et al. 2015). On the western shore, the rainfall maximum occurs during the early morning to midday, albeit with large variation throughout the year (Datta 1981; Ba and Nicholson 1998). Datta (1981) attributes this morning maximum to convergence between the prevailing easterlies and land breeze across the western shore. Mukabana and Pielke (1996) also suggest that a downslope drainage flow from the high grounds in central Africa may combine with the land breeze to decelerate the prevailing easterlies, and produce convergence in this region.

Maximum rainfall over the whole lake is coincident with the peak in land breeze strength (between 0600–0900 LT) (Datta 1981). Using an infrared (IR) satellite product, Ba and Nicholson (1998) noted that the peak in cold cloud frequency over the lake occurs around 0200 LT during the long rains and 0800 LT during the short rains. However, an asymmetry in the time of maximum rainfall exists across the lake, with a clear shift from 2100 LT over the very east of the lake to 0900 LT in the west (Fig. 1.5). Overnight, downslope flow from the Kenya Highlands, formed of prevailing easterlies accelerated by a katabatic component and land breeze, produce subsidence to the east of the lake, suppressing convection in this region (Mukabana and Pielke 1996).

Camberlin et al. (2018) used a mean diurnal cycle from satellite rainfall estimates to suggest that afternoon storms formed over the Kenya Highlands could propagate westward over Lake Victoria, where they would be rejuvenated overnight (Fig. 1.6). However, a mismatch in propagation speeds ( $4.1 \text{ m s}^{-1}$  and  $9.9 \text{ m s}^{-1}$  over the highlands and lake respectively) made this connection unclear.

Lake Victoria is not the only body of water to modify the timing of local precipitation. Figure 1.5 shows that early morning rainfall maximums also exist over Lakes

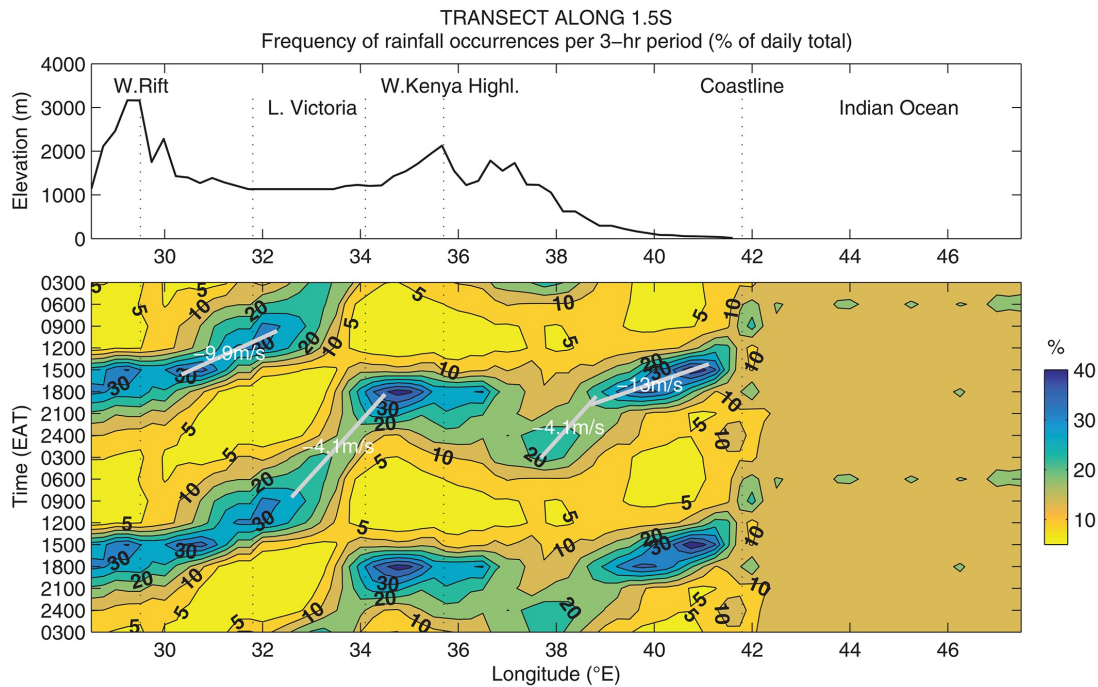


Figure 1.6.: Figure 8 from Camberlin et al. (2018). Time-longitude plot of diurnal rainfall occurrence along  $1.5^{\circ}\text{S}$ . (a) Elevation cross-section. (b) Percentage of local rainfall occurrence during each 3 hour period. For convenience, two consecutive diurnal cycles are plotted. Heavy grey lines show approximate propagation and corresponding speeds. [© 2017 Royal Meteorological Society. Published by John Wiley and Sons.]

Tanganyika and Malawi (to the south of Lake Victoria). Also of note is the clear signal of inland sea breeze propagation along the Somali coastline; over the ocean, maximum rainfall occurs around midday, but the time of maximum rainfall increases moving inland to 0300 LT at 400 km inland (Fig. 1.5).

### 3.d.2. Factors influencing regional climate over the Lake Victoria basin

As apparent from the previous section, the spatial and diurnal pattern of rainfall is not only a result of the lake-land breeze circulation. Local orography induces local circulation patterns, and the interaction of large-scale monsoonal winds with these mesoscale circulations is also important. Modelling studies, generally using Regional Climate Models (RCMs), have been performed to investigate the contribution of these various controls to the amount and spatial and diurnal patterns of rainfall, as well as other aspects of the regional climate. Although numerical models have associated errors, they can be used to perform sensitivity tests to isolate contributions from different



controls, which is not possible using observations. In addition, sensitivity tests can be performed to investigate the response of the regional climate to changes in these controls; in particular, the effect of varying the lake–land temperature contrasts by changes to the physical characteristics of the lake surface has been considered in various studies. Some studies have also included coupled lake models (both 1D and 3D) which provide insight into feedbacks between the lake and atmosphere. Results from these various modelling studies are presented below, including a critical discussion of their limitations.

Okeyo (1986) investigated the effect of Lake Victoria on convective activity over the Kenyan highlands. Using a mesoscale model, simulations spanning Kenya and Lake Victoria were performed, with and without the presence of the lake. Although convergence occurred over the highlands at 1800 LT in both cases, the presence of the lake set up a divergent lake breeze which merged with the upslope flow, and advected moist air from above the lake toward the highlands. Evaporation above the lake surface cooled the air by approximately 5 K, increasing the pressure gradient between the lake and highlands. The advection of the moist air added  $\sim 2 \text{ g kg}^{-1}$  of moisture over the highlands. Vertical velocities near the surface in the region of convergence showed little change between cases, suggesting little effect of the lake on near-surface convergence. However, vertical velocities were increased at upper levels, where latent heat release by the moister air invigorated the updraft. As a result, convection deepened and rainfall was enhanced. This experiment was performed with a coarse grid of only 40 km horizontal grid-spacing and heavy smoothing was applied to the terrain. In addition, the simulations were initialised with an atmosphere at rest at 0600 LT and most of the analysis was presented only 12 hours later at 1800 LT, likely not enough time to allow for model spin-up.

Mukabana and Pielke (1996) conducted a study to isolate the effects of the local (mesoscale) circulations and the large-scale winds. The study investigated a case study day from April 1985 (long rains). The modelling experiments were performed using the Regional Atmospheric Modelling System (RAMS), developed at Colorado State University (Pielke et al. 1992), with 20 km horizontal grid-spacing and 30 vertical levels. A control experiment was run, alongside an experiment in which the atmosphere was initialised from rest (to isolate mesoscale effects) and an experiment with land–water contrasts and orography removed (to isolate the effect of monsoonal winds). Advection of moisture from the Indian Ocean by the monsoonal winds was shown to contribute much more moisture to the region than the advection of moisture from water bodies by the mesoscale circulations (Mukabana and Pielke 1996). Note that Okeyo (1986)

did consider this latter effect to be important. Deep convection was shown to form in the regions of convergence between local circulations and large-scale winds, but its advection was shown to be controlled only by the large-scale winds. However, the results of this study should be considered with caution; not only was the model resolution coarse, but little time was given for model spin-up and smoothing was applied to the orography. Resultingly, large discrepancies existed between the observations and simulated precipitation in the control run. In addition, the authors suggest that the rainfall patterns seen in the model run with no orography could be a result of imbalances in the model caused by the removal of the orography.

Anyah et al. (2006) conducted experiments to further investigate the roles of orography and large-scale moisture advection. These experiments were performed using the National Centre for Atmospheric Research (NCAR) regional climate model RegCM3 (Pal et al. 2007) coupled to a 3D lake model based on the Princeton Ocean Model (POM) (Blumberg and Mellor 1987), but adapted for use over Lake Victoria (as described in Song et al. 2002, 2004 and Anyah 2005). Orography was shown to enhance the thermal contrast between the lake and land overnight, as katabatic flow cooled air near the shore by mixing it with cooler air from the mountains. As a result, a stronger land breeze formed, enhancing convergence and convection over the lake. The presence of steep orography to the east of the lake also affects daytime convection; simulations showed that strong anabatic winds entrain the lake breeze front further away from the lake. Correspondingly, rainfall occurs further east.

Sensitivity tests on the advection of moisture into the region across the four sides of a bounding box ( $5^{\circ}\text{N}$ – $7^{\circ}\text{S}$ ,  $25^{\circ}\text{E}$ – $41^{\circ}\text{E}$ ), showed that the amount of rainfall across the region exhibited a high dependency on advection through the eastern side of the box, but almost no dependency on advection through the remaining three sides (Anyah et al. 2006). This result demonstrates the importance of the transport of moisture from the Indian Ocean over the region by the large-scale winds, in agreement with Mukabana and Pielke (1996). Anyah et al. (2006) also highlighted the interaction of orography with the monsoonal winds. Simulations demonstrated that more rainfall would occur to the west of the lake if the steep orography to the east was reduced. Convergence over the mountains, between the lake breeze and prevailing winds (reinforced by anabatic flows), enhanced convection in this region. The convection was further fuelled by the moisture advected from the east, which was therefore unable to penetrate further westward. This blocking effect was also noted by Song et al. (2004).

Properties of the lake, especially its surface, can affect the lake–land breeze con-

trasts and feed back to the rainfall. Fraedrich (1972) used an idealised, cylindrically-symmetric, numerical model to simulate the nocturnal circulation over the lake. The strength of the circulation showed a strong dependence on the temperature difference between the lake and land. The study also suggested that the magnitude of latent heat release over Lake Victoria during the nocturnal circulation is comparable to that of a disturbance over the Gulf of Mexico. It should be noted that this study used a motionless atmosphere (in reality there are easterly prevailing winds) and assumed a hydrostatic atmosphere (likely not suitable for small-scale dynamics such as this circulation).

Anyah and Semazzi (2004) used NCAR's RegCM2 (Giorgi et al. 1993a,b) to investigate the effect of lake surface temperatures. Three four-month simulations were run across the 1988 short rains, including a control with a LST of 24°C and two with LSTs perturbed by  $\pm 1.5^\circ\text{C}$ . In all cases, the temperature was uniform across the whole lake and kept constant throughout the run. For a cooler lake surface, divergence over the lake increased. Subsequently, rainfall also decreased by around 50%. For a warmer lake surface, convergence increased over the region and some areas received an increase in rainfall of over 100%. In both cases, it should be noted that the rainfall anomaly pattern varied greatly between months. In the case with warmer LSTs, the rainfall anomaly over the lake itself was actually negative in November and December, but a strong positive anomaly occurred to the south west. It was hypothesised that, due to the more southerly position of the ITCZ at this time, the additional moisture (from increased evaporation from the warmer lake surface) was advected southward toward the ITCZ by the prevailing flow. Note that the model resolution in this study was very coarse (60 km horizontal grid-spacing). This study showed the effect of changing surface temperatures over a season, but shorter-term fluctuations may also have an effect.

In Anyah and Semazzi (2004), LSTs were kept constant and were uniform across the whole lake. This unrealistic LST distribution may have significantly affected the precipitation anomaly pattern. In addition, it is expected that a two-way feedback would exist between the lake and regional climate. To investigate this, Song et al. (2004) used a nested coupled model, consisting of the North Carolina State University (NCSU) version of the NCAR regional climate model (NCSU-RegCM2) and the 3D POM, run across the short rains season of 1988. In the model, the temperature of the surface water varied by up to 1°C between day and night (the greatest variation occurred near the shore and there was little change over the centre of lake) as the surface water rapidly warmed, and the mixed layer deepened through the late morning. Effects on

water temperature were limited to within  $\sim 30$  m of the surface. In earlier work, a coupled simulation was also performed using just a 1D lake model (see Song et al. 2002). By comparing the model runs using 1D and 3D lake models, the effect of lake hydrodynamics on the climate could be elucidated. Both model configurations exhibited a temperature dipole due to the bathymetry (shallower and warmer in the SE, deeper and cooler in the NW—note that this does not agree with observations in Thiery et al. (2015)). However, the southeasterly prevailing winds forced an anticlockwise gyre over the whole lake in the 3D coupled model, which advected and mixed the water and resulted in a warm tongue in the NE and cold tongue in the NW. The mixing of the warm water in the SE increased the mean LST by  $\sim 1^\circ\text{C}$ . However, despite the mean increase in evaporation, mean rainfall over the lake decreased in the 3D simulation. The decrease was attributed to a weakened temperature dipole across the lake, which altered the winds and reduced convergence and rainfall over the SW of the lake. The model used in this study had a coarse resolution of 20 km and only 9 vertical levels. The study also neglected effects on the larger-scale circulation and rainfall over the surrounding land.

In a later study, using an updated regional model (RegCM3), Anyah and Semazzi (2009) used TRMM observations to show that the spatial pattern of rainfall over the basin was improved when the model was coupled with POM, compared to a 1D lake model. This improvement was attributed to changes in the pattern of evaporation caused by the redistribution of heat within the lake by horizontal advection. This study also showed that the 3D structure of lake temperatures was, to some extent, a product of the equatorial location of the lake (i.e. the Coriolis parameters were important). This result suggests that some generalised results from studies of large lakes in the midlatitudes may not be applicable to Lake Victoria.

Anyah et al. (2006) also used RegCM3-POM to investigate the effect of replacing the lake with marsh (to represent water hyacinth). As a result of the differing albedo, roughness, and thermal capacity of water and marsh, the lake retained less heat during the day. Consequently, the temperature contrast between the lake and land at night was reduced, as was the convergence and rainfall. During the day, convection was increased to the east of the lake in the marsh simulation, which the authors attributed to an increased temperature gradient between the lake and land. This enhanced temperature dipole was generated by greater cooling of air above the lake, as more energy from solar heating went into evaporation, as opposed to being absorbed by the lake surface.

Sun et al. (2015) performed some similar experiments to those by Anyah and Semazzi

(2004), but using the Weather Research and Forecasting (WRF) model (Skamarock and Klemp 2008). The WRF model was run with a much smaller horizontal grid-spacing of 4 km. Experiments with uniform LSTs (as in Anyah and Semazzi 2004) showed that increasing the temperature increased the amount of rainfall (due to an increase in lake–land temperature contrast), but generally maintained a similar spatial pattern (two rainfall maxima, over NW and E of lake). However, at the highest LST (25°C), the rainfall pattern was disrupted and large amounts of rainfall occurred over almost the whole lake. Sun et al. (2015) suggest that this non-linear response occurred because, above a certain temperature (in this case between 24 and 25.5°C), the main control on convection is convective instability over the lake, rather than the temperature contrast between lake and land. Simulations were also performed to compare rainfall occurrence using uniform LSTs, and using the asymmetric LST pattern from Anyah (2005). It was shown that the asymmetry plays an important role in the location of rainfall. A complementary study by Sun et al. (2014) coupled WRF with POM to investigate the LST pattern produced and its effect on rainfall. The study was comparable to Song et al. (2004) and Anyah (2005), since all three were performed over a short rains season. However, there were differences between the LST patterns produced between Sun et al. (2015) and Song et al. (2004) and Anyah (2005). These differences may have arisen due to the different years of the studies or from the different resolutions and physics of the models used, the latter of which suggests the LST distribution exhibits a high sensitivity to the climate.

The UK Met Office used the Hadley Centre regional climate model (HadRM3P) to study the Lake Victoria basin. Williams et al. (2015) used an ensemble of HadRM3P simulations with a horizontal grid-spacing of 50 km, driven by five members of the Hadley Centre global perturbed-physics ensemble (QUMP). These simulations did not include an integrated lake model. The study demonstrated that the presence of the lake reduces the expected rate of temperature increase over land during the morning; Williams et al. (2015) suggest that the land breeze is responsible for the advection of air away from the warming land. In the afternoon, temperatures over land cool faster than expected, which Williams et al. (2015) attributed to cooler air advected onshore by the lake breeze. It was also suggested that cloud cover associated with the lake breeze could reduce temperatures.

Thierry et al. (2015) performed simulations using the regional Consortium for Small-Scale Modelling (COSMO) model in climate mode (COSMO-CLM, Rockel et al. 2008), with a horizontal grid-spacing of  $\sim 7$  km. This atmosphere model was fully coupled to the 1D Freshwater lake (FLake) model and the Community Land Model (CLM) version

3.5 (Davin and Seneviratne 2012). The modelling studies discussed so far were limited to time periods between one day, up to a season. The simulation in Thiery et al. (2015) was run for a 13 year period. A second simulation, with lakes removed, was also run.

For the first time, LSTs were compared to observations, which showed that LSTs increased from SW to NE in both the model (with lakes) and observations (Thiery et al. 2015). This result contradicts previous studies (e.g Song et al. 2002, 2004; Anyah 2005; Anyah and Semazzi 2009; Sun et al. 2014) and the expected temperature distribution, given the shallower water in the SW. However, no explanation was offered for this pattern. Given that the study in Thiery et al. (2015) spanned 13 years, whereas other studies (e.g. Song et al. 2002, 2004; Sun et al. 2014) spanned one short rains season (or just a few day of the season), differences may be accounted for if the LST pattern varies seasonally. However, the simulations in Anyah (2005) and Anyah and Semazzi (2009) were more idealised and not tied to a specific season. The differences may then be attributed to the different lake models (1D FLake in Thiery et al. 2015 and 3D POM in the other studies). Another explanation could be a long-term change in LST pattern, such that the earlier studies may still be correct. An explanation for the discrepancy is impossible to deduce without observations from individual years or seasons. However, the earlier studies using POM may need to be considered with care.

Thiery et al. (2015) clearly showed the cooling effect of Lake Victoria (and other lakes in the region) on the climate. Over the lakes themselves, the average cooling (compared to the run with no lakes) was  $-0.67$  K, but localised areas exceeded  $-2$  K over Lake Victoria. As a result of downwind advection, the cooling effect was also experienced over land in some regions close to the lakes. Over the lakes, rainfall amounts showed large increases (for example, an 85% increase over Lake Victoria), but there was little effect on rainfall totals over land.

Thiery et al. (2015) also presented vertical cross-sections from west to east across the Lake Victoria region (Figs. 1.7 and 1.8), which clearly show the effect of the lake on the stability and circulation of the atmosphere. Firstly, near-surface air over the lake region is much more moist throughout the diurnal cycle due to the presence of the lake (Figs. 1.7d and 1.8d). During the day, the presence of the lake cools the air within approximately 1.5 km of the surface (Fig. 1.7c), which induces a pressure increase near the surface (Fig. 1.7e). Subsequently, a lake breeze is generated to the west with a depth of  $\sim 1$  km, which advects moist lake air onshore (Fig. 1.7d). Updrafts occur around  $30.5^\circ\text{E}$ , with westerlies back toward the lake between  $\sim 3$ – $5$  km MSL, and subsidence of drier air over the centre of the lake forming a complete circulation cell (Fig. 1.7d).

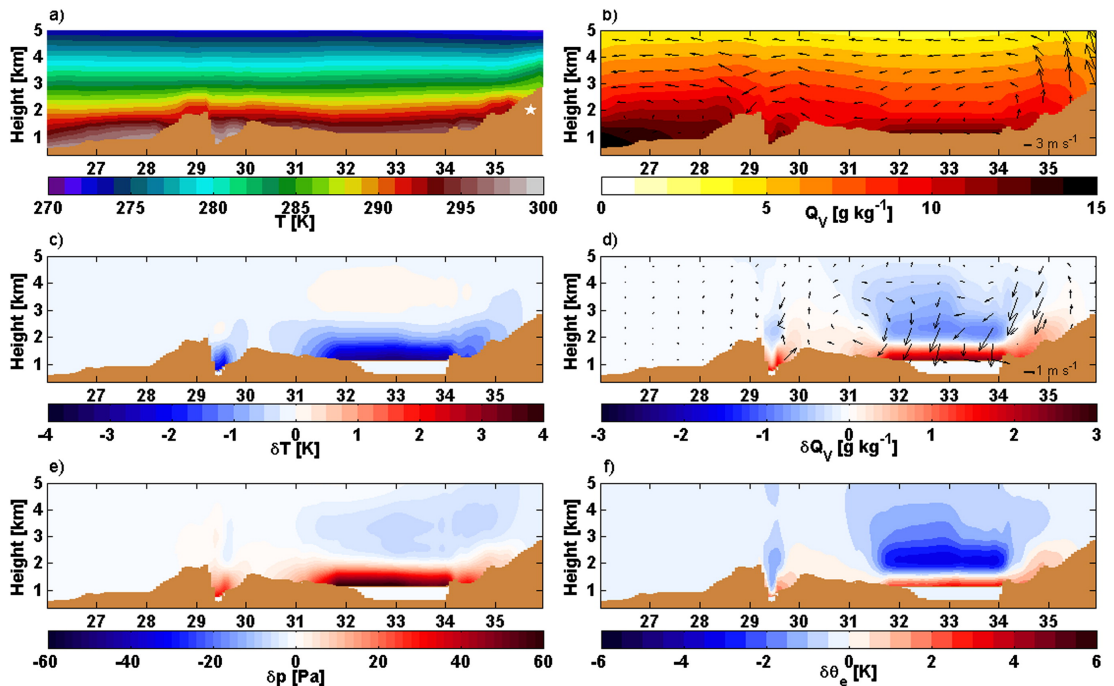


Figure 1.7.: Figure 13 from Thiery et al. (2015). Vertical cross sections along a transect from west to east across Lake Victoria for 1999–2008 for the model run with lakes (see text) from 0900–1800 UTC (daytime) of (a) temperature and (b) specific humidity, including longitudinal circulation (vectors), and for the mean change due to lake presence in (c) temperature, (d) specific humidity and circulation (vectors), (e) pressure and (f) equivalent potential temperature. [© Copyright 2015 AMS.]

The result of the subsidence is a more stable layer between 1.5 and 6 km (Fig. 1.7f), which suppresses convection over the lake. Thiery et al. (2015) argue that, to the east, the primary local circulation is controlled by orography rather than the lake, although the lake breeze does enhance the circulation. The overall effect is similar to the west, as moist lake air is advected over the land (Fig. 1.7d). However, Thiery et al. (2015) suggest that the strengthening of the anabatic flow by the land breeze induces near-surface divergence, and therefore subsidence, over land closest to the shore. As a result, the subsidence and stabilising effect over the eastern side of the lake is reduced relative to the west (Figs. 1.7d,f).

During the night and early morning, the presence of the lake increases the temperature in the boundary layer above Lake Victoria, such that a relative low pressure is induced (Figs. 1.8c,e). A land breeze forms over both the eastern and western shores, but is stronger from the east due to katabatic winds (Fig. 1.8d). Convergence between the land breezes over the centre of the lake forces ascent of the moist air. The stability of

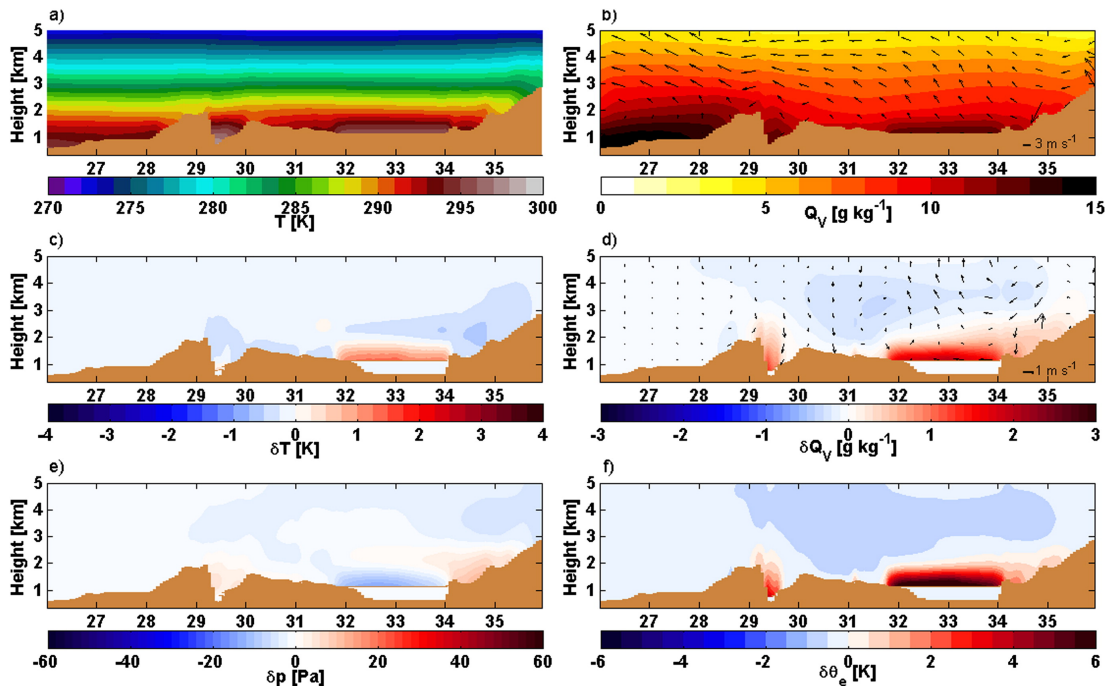


Figure 1.8.: Figure 14 from Thiery et al. (2015). As in fig. 1.7, but from 2100–0600 UTC (nighttime). [© Copyright 2015 AMS.]

the boundary layer over the lake is reduced by the higher temperatures and moisture (Figs. 1.8c,d,f) which, combined with the vertical motion, enhances convection over the lake.

By comparing the diurnal cycle in low-level wind convergence and  $\theta_v$  (buoyancy), Finney et al. (2019) used the present-day CP4 simulations (Stratton et al. 2018) to study factors responsible for mesoscale flow. Buoyancy differences between air over the lake and nearby surrounding land were linked to lake and land breezes, whereas buoyancy differences between air over the mountains and air at the same altitude over the lake were linked to anabatic and katabatic flows. As a result of these mesoscale flows, convergence occurred over the lake at night, whilst divergence occurred over the mountains. In the day, the reverse was true. An additional outcome of this study was that a stronger nocturnal land breeze was simulated in the CP configuration compared to a 25 km parametrised configuration. As a result, evaporation and moisture flux convergence were increased over the lake, leading to unrealistically high rainfall. However, the diurnal cycle over orography to the east of the lake showed improvement in the CP4 run relative to the parametrised run.



### 3.d.3. Feedbacks between convection over land and lake

The lake–land breeze and orographic circulations influence the timing and position of convection over both Lake Victoria and the surrounding land. However, studies have also shown that daytime convection over land may feed back onto convection over the lake. Thiery et al. (2016) used satellite observations to show that particularly intense overnight storms on the lake are preceded by intense evening storms over land. This link was attributed to increased moisture availability following storms over land and a modification of the lake–land breeze circulation, rather than explicit propagation of the land-based storms toward the lake. Daytime cold pools over land reduce the lake–land temperature contrast, decreasing the advection of moisture away from the lake surface. Overnight, persisting cold pools increase the lake–land contrast, enhancing over-lake convergence.

The results of Thiery et al. (2016) were used to propose a statistical early warning system, Lake Victoria Intense Early Warning System (VIEWS), to predict the top 1% of storms. The predictor can be driven using real-time overshooting top data derived from the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument with a detection algorithm developed by Bedka et al. (2010) and Bedka (2011). The total number of pixels with overshooting tops is aggregated through the day and fed into a logistic regression model, developed using 9 years' worth of data, to predict the likelihood of a severe (top 1%) storm over Lake Victoria. The scripts to run this model are freely available on Github under an MIT licence ([github.com/wthiery/VIEWS](https://github.com/wthiery/VIEWS)) and have little computational cost, therefore can be used by various lake users and management units. However, this product is only able to predict the 3–4 most intense storms per year. It is likely that less-intense storms also cause loss of life. In concluding, Thiery et al. (2017) note that the system could be adapted to take into account a distribution of storm strengths. Despite high skill, the warning time from the model is very short, so NWP remains important for forecasting storms at a more useful lead time.

### 3.d.4. Discussion and summary of Lake Victoria literature

Many of the modelling studies over Lake Victoria have been performed using coarse models with parametrised convection. Such studies would benefit from being repeated using high-resolution CP models. Such models would include better representation of orography and likely the diurnal cycle and lake–land breeze circulation. Most of the

studies were only performed for one season. Given the high interannual and seasonal variability in the region, it may not be appropriate to generalise some of the results. In addition, effects on the climate, or feedbacks on LSTs, were generally presented as means across the simulation period (from days to months to years), such that the effect and importance of day-to-day variability was neglected.

To summarise this section, the existing modelling studies have shown the importance of interactions between large-scale circulations and mesoscale circulations (driven by orography and lake–land temperature contrasts), in determining the amount, and spatial and diurnal patterns, of rainfall. The sensitivities of the climate to large-scale moisture advection, physical characteristics of the lake and the orography, demonstrate the requirement of an NWP model which can accurately represent all these processes, including their interactions and feedbacks.

## 4. Methods

The thesis matter was investigated in two main ways: firstly, by performing verification on a state-of-the-art CP forecast model for East Africa, in particular of its skill in the prediction of intense precipitation; and secondly, by investigating the underlying mechanisms behind severe weather in the region.

### 4.a. East Africa CP Model verification

#### 4.a.1. MetUM East Africa CP model

The UK Met Office ran a convection-permitting model over an East African domain from 2011 to early 2019. The configuration of this model is fully described in chapter 2, section 2.a. Output from the model was disseminated to forecasters in East Africa via the Voluntary Cooperation Programme (VCP) Africa Web Viewer, a password-protected online portal distributed by the Met Office. The model spanned 20.5°S to 17.5°N and 21.5°E to 52°E, as shown by the domain in Fig. 1.1. The model had a resolution of 4.4 km, allowing the convective parametrisation to be switched off, and the convection to be modelled explicitly.

The configuration of this model in spring 2012 was compared to the global model, as well as observations including brightness temperature from satellite, arrival time difference

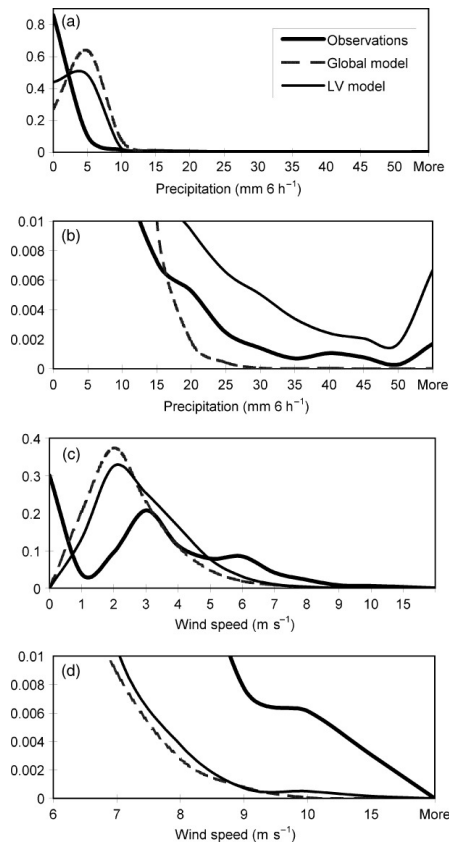


Figure 1.9.: Figure 3 from Chamberlain et al. (2014). Normalized frequency distribution of accumulated precipitation and wind speed for observations, the global model and the CP Lake Victoria model. (a) The whole distribution of precipitation values. (b) Zoom on less frequent precipitation events (<1%) to capture the changes in high levels of rainfall. (c) The whole distribution of wind speed values. (d) Zoom on less frequent wind events to capture the changes in high wind speeds. [© 2013 Crown copyright, the Met Office. *Meteorological Applications* © 2013 Royal Meteorological Society. Published by John Wiley and Sons.]

(ATD) lightning and wind and precipitation data from automatic weather stations (AWSs), for the two month period February–March 2012 by Chamberlain et al. (2014). In this older configuration, LSTs were prescribed from ARCLake (MacCallum and Merchant 2011) monthly night-time climatologies and the domain spanned a smaller area, centred on Lake Victoria. Overall, the global model produced too much light rainfall, too often. A similar problem occurred in the CP model, although to a lesser extent. A large difference between the two models was that the global model was unable to produce any of the very high intensity rainfall seen in observations from ground stations, whereas the CP model predicted too many intense rainfall events. Both models were extremely poor at predicting wind speeds. According to observations,

wind conditions were stagnant for 30% of the analysis period. When wind was recorded, the most common speed was  $3 \text{ ms}^{-1}$ , with a secondary peak occurring at  $6 \text{ ms}^{-1}$ , which Chamberlain et al. (2014) attributed to the local orography, nocturnal low-level jets, or lake breezes. In contrast, in both models, the probability of a given wind speed increased from zero at  $0 \text{ ms}^{-1}$  to a peak at  $2 \text{ ms}^{-1}$ , after which the probability decayed (Fig. 1.9). The incorrect probability distribution of wind speeds suggests that the model may not correctly represent local circulations, such as the lake–land breeze circulation, or circulations associated with orography.

The main aim of Chamberlain et al. (2014) was to assess the skill of the models in forecasting severe events in the region. Two case studies of particularly severe events were examined, in which the CP model was better able to simulate the high-intensity rainfall and wind than the global model. Forecasts by the two models were analysed by identifying severe events from cloud top temperature, surface gusts, wind shear between 250 and 700 hPa, accumulated precipitation, and winds at 850 hPa. Using brightness temperature and ATD lightning to identify storms in the observations, the CP model was shown to have more skill in predicting storm occurrence (increased hit rate) compared to the global model, especially at forecast lead times of T+6 h and T+12 h. However, it was also shown to over-predict severe events, leading to more ‘false alarms’. Both models were unable to predict storms at lead times greater than T+12 h. It is noted that the threshold above which a storm should be classified as severe, with the potential to cause disruption to boats on the lake, is not well understood and likely to depend on the quality and loading of the boat.

A study by Eagle et al. (2015) aimed to assess the model since an update in February 2015, when the ND dynamical core was replaced with the Even Newer Dynamics for General Atmospheric Modelling of the Environment (ENDGAME) dynamical core, along with some other physics changes which are detailed in their report. The CP model was run in both the old and new configurations for ten case studies between March and June 2014. Overall, both configurations produced similar results, suggesting that the upgrades had little effect on the simulation of convection.

Comparing simulated and observed brightness temperature, neither the old nor new CP configurations produced enough cloud. The cloud produced was often too shallow (brightness temperature too warm). Despite this, the models actually simulated too much convective rain. Moreover, the simulated storms tended to be too small. As a result, the simulated storms were often far too intense. It is suggested that this problem stemmed from the resolution being too coarse to properly resolve the updrafts.

In addition, where one large storm was observed, the models often represented this as several smaller storm fragments, possibly due to issues with sub-grid mixing or the model microphysics. The configurations did not simulate enough light rain, especially surrounding intense storm cores. These biases have also been found in models over the UK as part of the Dynamical and Microphysical Evolution of Convective Storms (DYMECS) project (Stein et al. 2015) and when running the UM in other tropical regions of the world. A particular issue, perhaps related to LSTs, was that early afternoon storms in the region tended to be too intense, whereas the night-time storms over the lake were too weak, or not forecast at all.

A 1.5 km configuration over a smaller domain, spanning 3.9°S to 3.9°N and 28.7°E to 37.3°E, was also run by Eagle et al. (2015). This configuration actually showed less skill than the 4.4 km configurations, since the fragmentation of larger storms was increased. However, Eagle et al. (2015) did anticipate that, once the fragmentation issue was resolved, moving to a higher resolution model would improve the forecasts.

The verification in both Chamberlain et al. (2014) and Eagle et al. (2015) were performed over short time periods. In this thesis, a longer-term assessment is performed, in order to build up more robust statistics about the model performance. In addition, the verification is performed over a larger model domain, although particular attention is paid to Lake Victoria. A wide variety of verification metrics are used to understand different aspects of the model performance. In particular, verification was performed to determine how the skill of the model varies across the domain and throughout the diurnal cycle. This information can provide more insight into reasons behind good/poor model performance.

#### **4.a.2. Verification Methods**

Forecasters in East Africa have had access to a wide range of global models (including global MetUM, Météo France Action de Recherche Petite Echelle Grande Echelle (ARPEGE), European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) and National Oceanic and Atmospheric Administration (NOAA) Global Forecast System (GFS)) for many years, via the Synergie system (Voidrot-Martinez and Berthou 1991) and other online portals. The issues and biases with global models are well documented (section 2.b) and familiar to forecasters, who take into account and compensate for these biases when producing a forecast. Given the relatively new addition of CP models, forecasters are less familiar with how to use

and interpret such models.

Given that the main differences between CP and parametrised models are their representation of convection, and therefore prediction of precipitation, the ability of the East Africa CP model to forecast rainfall is assessed in chapter 2. In particular, the model performance is compared to that of the operational global MetUM, to quantify the added-value of the CP model compared to the global model forecast.

There are many aspects of a precipitation forecast which can be assessed, including the timing, location, intensity and structure of the precipitation. To verify the location of storms, traditional verification metrics (which rely on point-to-point verification) are unsuitable. In a parametrised model, rainfall often occurs over (unrealistically) large areas. Consequently, there are often areas of overlap between the forecast rainfall and observed rainfall, even if the structure of the precipitation is incorrect. In a CP model, even if the structure of a storm is more realistic, a forecast storm which is offset from its observed location by more than its horizontal extent, will exhibit no overlap between the forecast and observations. Using point-to-point verification, the forecast is then penalised for ‘missing’ the observed storm, as well as producing a ‘false alarm’ at the location of the simulated storm. This is known as the ‘double-penalty’ problem (Roberts and Lean 2008; Mittermaier 2012).

Although it is important that storms are forecast in the correct location, it is unlikely that a model can predict the exact location of a storm. Forecasters using CP models should be aware of this and use judgement to forecast storms over a region, rather than in a specific location. To minimise the double-penalty problem during verification, metrics have been developed to compare the model and observations over larger spatial areas, instead of on a point-to-point basis. These are known as ‘fuzzy’ or ‘neighbourhood’ metrics.

One such metric is the Fractions Skill Score (FSS) developed by Roberts and Lean (2008). The FSS is based on the Brier score. However, instead of comparing actual rainfall amounts in the forecast and observations, the FSS compares the fraction of grid-points which exceed a threshold rainfall amount within a specified neighbourhood. The FSS is constructed from the ratio of this fractions Brier score to the fractions Brier score of a low-skill reference forecast and is fully described in chapter 2, section 2.c.2. The FSS produces one score over the whole model domain at every time step. This score may be averaged over many time steps or used to assess the change in skill with time. In chapter 2, the skill score is adapted to use multiple forecast times to produce a map of how skill varies in space, termed the Localised Fractions Skill Score (LFSS).

Knowledge of how the performance of the model varies within the region is useful to forecasters, and may be used to infer aspects of the model that work well. For example, if the skill is higher over mountainous regions, this suggests that the model captures orographic rainfall well. Similarly, if the scores are high over lakes, this implies that the model is capturing lake–land–atmosphere interactions well.

The observations used to evaluate the models are satellite precipitation estimates from the NASA Global Precipitation Measurement (GPM) mission (Huffman 2017; Huffman et al. 2018). These estimates come primarily from microwave observations, but are supplemented with other information, for example from IR channels and monthly ground gauge totals. Satellite products are useful because they are gridded, so the models and observations can be directly compared over a large region. However, the algorithms used to produce the product can introduce large errors. A full description of the GPM product, along with a discussion of its limitations, is given in chapter 2, section 2.b.

An important aspect of this thesis is ensuring that the findings will be of use to the regional forecasters. As such, visits to Kenya Meteorological Department (KMD), Tanzania Meteorological Agency (TMA) and Uganda National Meteorological Authority (UNMA) were carried out in December 2016, to meet with forecasters and gain a better understanding of the forecasting process and forecasting tools available. The opportunity was also used to discuss perceived strengths and weaknesses of the different models being used. One of the major hurdles faced by operational CP models is their correct integration into standard operating procedures at national meteorological services. Given that parametrised models have been used operationally for many decades, forecasters are more familiar with how to use them and how to compensate for their limitations. Time and training is required for forecasters to familiarise themselves with how to interpret a CP model and use it in conjunction with global models.

In particular, it was noted that rainfall forecasts issued by the centres tended to cover a whole 24 hour period, and not give much indication of when in the day the rainfall might occur. This practice likely originates from the inability of global models to simulate the diurnal cycle, therefore it is better to use a 24 hour accumulation. However, the East Africa CP model was also being used to predict 24 hour accumulations, which is not making optimal use of its capabilities. In chapter 2, the performance of the CP and global models are compared for the prediction of 24 hour accumulations, as well as 3 hour rainfall rates, to showcase differences between CP and parametrised models.

During these visits to East Africa, the daily WMO Severe Weather Forecasting Demonstration Project (SWFDP) teleconference between forecasters from various countries in

Eastern Africa was observed. At this meeting, regions expecting rainfall of greater than 50 mm in 24 hours, or winds exceeding 25 knots ( $\sim 12.8 \text{ ms}^{-1}$ ) are agreed upon and marked on a map. The SWFDP rainfall threshold was used to inform the threshold used in FSS computation in chapter 2. Ultimately, a value of  $\sim 30$  mm in 24 hours was chosen as a compromise between matching the SWFDP threshold, and having a high enough number of rainfall events to sample.

#### 4.b. Case study analysis

Studies of convection over the Lake Victoria basin have tended to investigate the mean diurnal cycle, computed from a large number of events (section 3.d). This approach has established that the lake–land breeze plays an important role in the formation of storms in the region. However, the presence of the lake–land breeze circulation does not mean that a storm will always form or behave in the same way. Variability in the timing and location of storm initiation, the storm lifetime, and its propagation, is masked by the mean diurnal cycle. The controls on the processes responsible for this variability are not well understood, given the scarcity of high quality observations in the region and limited modelling studies. In chapter 3, detailed case studies are used to investigate processes responsible for storm initiation, both on the local and large scales.

Given the lack of observations, the case studies were investigated using high-resolution CP MetUM simulations, with 1.5 km horizontal grid-spacing. Despite some concerns about model performance arising from chapter 2, specific cases where the operational East Africa model was known to perform well were chosen to be run at the higher resolution. In addition, three model initialisation times were used for each case study, such that the best match to observations (brightness temperature and precipitation from satellite products) could be chosen.

Three case studies were investigated, two of which involved a significant storm over Lake Victoria. In the first case, taken from the 2015 long rains season, a storm formed over land to the east of Lake Victoria during the evening, and propagated over the lake during the night. In the second case, taken from July 2016 (during a dry season), a storm formed over the lake itself during the early morning and showed little propagation. The first storm grew to be very large (spanning almost the whole lake), whereas the July storm was small but intense. Given the known importance of the lake–land breeze circulation, a third case in which no storm formed was run to investigate the



circulation in isolation, and effectively act as a ‘control’ run to compare and contrast with the two storm cases.

Although case studies offer opportunity for detailed analysis, a limited number of case studies cannot build a complete picture, and a full range of large-scale conditions cannot be sampled. However, the case studies presented in chapter 3 were chosen to sample two very different types of storm over the lake, providing contrasting conditions for comparison.

#### **4.c. Aircraft observations - HyVic pilot flight campaign**

In general, high-resolution, detailed in-situ observations are sorely lacking in East Africa. As a result, the model evaluation in chapter 2 depends solely on satellite retrieval estimates of rainfall, and chapter 3 relies on CP modelling to reproduce case studies in high-resolution detail. In part, the lack of observations is attributed to issues faced by many developing countries, including the expense of purchasing and maintaining instrumentation and the need for infrastructure to support the instruments and facilitate the distribution of data. Notably, radar and reliable radiosonde launches are difficult to sustain. Whilst some AWSs and rain gauges do exist in the region, data is not widely shared. In addition, the time resolution is often insufficient for studying features such as lake–land breezes and short-lived propagating storms. In addition, observations in the centre of Lake Victoria are particularly difficult to obtain.

A unique opportunity arose to observe the lake–land breeze circulation over Lake Victoria in unprecedented detail, using the Natural Environment Research Council (NERC) Facility for Airborne Atmospheric Measurements (FAAM) BAe-146 aircraft. Having found out that the aircraft would be based in Entebbe in late January 2019 as part of the MOYA field campaign (Tollefson 2019), the HyVic pilot flight campaign was designed as an add-on. The campaign consisted of just three flights (although this was later reduced to two), to act as a proof-of-concept for a future extended campaign, in addition to collecting an initial, novel observational dataset.

The FAAM aircraft is fitted with various instruments to measure temperature, pressure, humidity, winds (and other variables) along its track. The instruments used in the campaign are described in Table 4.2. Where multiple instruments measured the same variable, comparisons were made between datasets, and a suitable instrument was chosen following discussions with experienced users of the aircraft data. Note that an offset in the favoured wind probe was identified and the data is awaiting reprocessing by

FAAM. However, the possible errors in the data are unlikely to affect any conclusions, so have therefore been included in the thesis. In addition to measurements made on board the aircraft, dropsondes were deployed from the aircraft and were able to measure similar variables. Dropsondes are essentially the same instrument as radiosondes, but travel downward instead of upward. These allowed full profiles of the atmosphere to be sampled.

The maximum flight time of the aircraft is approximately 5 hours, but this was reduced to  $\sim 4.5$  hours given the altitude of Entebbe and the warm temperatures. For taking measurements along straight level runs, the aircraft cruises at  $\sim 100 \text{ ms}^{-1}$  indicated airspeed. Over water (including Lake Victoria), the aircraft can profile down to 50 ft ( $\sim 15 \text{ m}$ ) above ground level (AGL) and perform extended runs at 100 ft ( $\sim 30 \text{ m}$ ) AGL. Over land, the minimum safe altitude is 500 ft ( $\sim 150 \text{ m}$ ) AGL. These minimum altitudes only apply during daylight hours and when visibility is good. Otherwise, the aircraft is restricted to 1,000 ft above the highest obstacle within 30 nm of the aircraft track. These restrictions had to be taken into consideration during flight planning.

The timings of the two HyVic flights were chosen to coincide with mature lake and land breezes, but restrictions on low-level flying in the dark also had to be considered. The flight tracks consisted of flying a northwest to southeast transect at various heights from Entebbe, across the lake, and into Tanzania. Further justification of the flight plans is given in chapter 4, in the context of results from chapter 3.

## 5. Thesis aims and structure

This thesis aims to improve the prediction and understanding of severe weather over East Africa by: evaluation of NWP forecasts, in particular from an operational CP model; diagnosing the key factors responsible for convective initiation over the Lake Victoria basin; and using observations to build a more detailed understanding of the lake–land breeze circulation over Lake Victoria. The project also contributes toward identifying and understanding biases in CP models, and provides motivation for a more intensive field campaign in the region.

The main objectives of the PhD are:

- Investigate the skill of a convection-permitting forecast model for predicting intense precipitation over East Africa

- How does the skill compare to that of a global (parametrised) forecast model?
- How does the skill vary on sub-daily timescales, and with forecast lead time?
- How does skill vary within the model domain?
- Investigate key controls on storm formation over the Lake Victoria basin
  - What is the role of the lake–land breeze circulation and other local forcings?
  - How important are large-scale features (moisture availability/circulation) for storm formation?
  - What additional information about storm formation can be gained from case studies, compared to previous studies using the mean diurnal cycle?
- Investigate the lake–land breeze circulation over Lake Victoria using novel aircraft observations from the HyVic pilot flight campaign
  - What are the key characteristics of the circulation, e.g. wind speeds, moisture content, gradient across fronts?
  - How well is the lake–land breeze circulation represented in CP models when compared to observations?
  - How can this pilot campaign be used to inform a future, extended campaign?

Chapter 2 presents a comparison of the skill of the operational MetUM global and CP configurations over East Africa. This chapter aims to highlight key differences between the behaviours of the two models, in particular their ability to predict the timing and location of storms and the structure of precipitation features. This is the first time that a CP model in the tropics has been verified over an extended period of time (2 years). The aim is to provide regional forecasters with increased confidence in how and when to use the model. This chapter also provides feedback for model developers of where improvements can be made in the model.

High-resolution CP MetUM simulations of three case studies of the lake–land breeze circulation and storms over the Lake Victoria basin are analysed in chapter 3. The objective of this chapter is to identify key controls on the initiation of storms in the region, including the lake–land breeze circulation, orography, and large- and local-scale moisture and circulation. A more detailed understanding of storm triggers and

conditions favourable for convection could lead to improved forecasts of convection.

Chapter 4 presents observational data from the HyVic pilot flight campaign using the FAAM aircraft. Observations of the evening lake breeze and morning land breeze over the Lake Victoria basin are presented, including vertical profiles from dropsondes over the lake and land. This is the first time that such detailed observations have been recorded over Lake Victoria; the work aims to build a better understanding of the lake–land breeze circulation. The observations are also compared to high-resolution CP MetUM model runs to investigate the representation of the lake–land breeze circulation in models, including the impact of model grid-spacing.

The results of chapters 2–4 are summarised in chapter 5, alongside a discussion and synthesis of the results. Final conclusions and recommendations for future work are also made.

---

## References

- Anyah, R., and F. Semazzi, 2004: Simulation of the sensitivity of Lake Victoria basin climate to lake surface temperatures. *Theor. Appl. Climatol.*, **79** (1-2), 55–69.
- Anyah, R. O., 2005: Modeling the variability of the climate system over Lake Victoria Basin. Ph.D. thesis, North Carolina State University. Available at: <http://www.lib.ncsu.edu/resolver/1840.16/3820>.
- Anyah, R. O., and F. Semazzi, 2009: Idealized simulation of hydrodynamic characteristics of Lake Victoria that potentially modulate regional climate. *Int. J. Climatol.*, **29** (7), 971–981.
- Anyah, R. O., F. H. Semazzi, and L. Xie, 2006: Simulated physical mechanisms associated with climate variability over Lake Victoria Basin in East Africa. *Mon. Wea. Rev.*, **134** (12), 3588–3609.
- Arakawa, A., 2002a: A Personal Perspective on the Early Years of General Circulation Modeling at UCLA. *General Circulation Model Development - Past, Present and Future*, Academic Press, chap. 1, 1–65.
- Arakawa, A., 2002b: Future Development of General Circulation Models. *General Circulation Model Development - Past, Present and Future*, Academic Press, chap. 23, 721–780.
- Arritt, R. W., 1993: Effects of the large-scale flow on characteristic features of the sea breeze. *Journal of Applied Meteorology*, **32** (1), 116–125.
- Ba, M. B., and S. E. Nicholson, 1998: Analysis of convective activity and its relationship to the rainfall over the Rift Valley lakes of East Africa during 1983-90 using the Meteosat infrared channel. *J. Appl. Meteor.*, **37** (10), 1250–1264.
- Ballentine, R. J., 1982: Numerical simulation of land-breeze-induced snowbands along the western shore of Lake Michigan. *Mon. Wea. Rev.*, **110** (11), 1544–1553.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525** (7567), 47.
- Bechtold, P., J.-P. Chaboureaud, A. Beljaars, A. Betts, M. Köhler, M. Miller, and J.-L. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation

- over land in a global model. *Quart. J. Roy. Meteor. Soc.*, **130 (604)**, 3119–3137.
- Bedka, K., J. Brunner, R. Dworak, W. Feltz, J. Otkin, and T. Greenwald, 2010: Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients. *J. Appl. Meteor. Climatol.*, **49 (2)**, 181–202.
- Bedka, K. M., 2011: Overshooting cloud top detections using MSG SEVIRI Infrared brightness temperatures and their relationship to severe weather over Europe. *Atmos. Res.*, **99 (2)**, 175–189.
- Birch, C., D. Parker, A. O’Leary, J. Marsham, C. Taylor, P. Harris, and G. Lister, 2013: Impact of soil moisture and convectively generated waves on the initiation of a West African mesoscale convective system. *Quart. J. Roy. Meteor. Soc.*, **139 (676)**, 1712–1730.
- Birch, C., M. Reeder, and G. Berry, 2014a: Wave-cloud lines over the Arabian Sea. *J. Geophys. Res.: Atmospheres*, **119 (8)**, 4447–4457.
- Birch, C. E., D. Parker, J. Marsham, D. Copsey, and L. Garcia-Carreras, 2014b: A seamless assessment of the role of convection in the water cycle of the West African Monsoon. *J. Geophys. Res.: Atmospheres*, **119 (6)**, 2890–2912.
- Birch, C. E., and M. J. Reeder, 2013: Wave-cloud lines over northwest Australia. *Quart. J. Roy. Meteor. Soc.*, **139 (674)**, 1311–1326.
- Black, E., J. Slingo, and K. R. Sperber, 2003: An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Mon. Wea. Rev.*, **131 (1)**, 74–94.
- Blumberg, A. F., and G. L. Mellor, 1987: A description of a three-dimensional coastal ocean circulation model. *Three-dimensional coastal ocean models*, 1–16.
- Byers, H. R., and H. R. Rodebush, 1948: Causes of thunderstorms of the Florida peninsula. *J. Meteor.*, **5 (6)**, 275–280.
- Camberlin, P., W. Gitau, O. Planchon, V. Dubreuil, B. M. Funatsu, and N. Philippon, 2018: Major role of water bodies on diurnal precipitation regimes in Eastern Africa. *Int. J. Climatol.*, **38 (2)**, 613–629.
- Camberlin, P., V. Moron, R. Okoola, N. Philippon, and W. Gitau, 2009: Components of rainy seasons’ variability in Equatorial East Africa: onset, cessation, rainfall

- frequency and intensity. *Theor. Appl. Climatol.*, **98** (3-4), 237–249.
- Cannon, T., F. Krüger, G. Bankoff, and L. Schipper, 2014: Putting culture at the centre of risk reduction. *World Disasters Report 2014: focus on culture and risk*, T. Cannon, and L. Schipper, Eds., International Federation of Red Cross and Red Crescent Societies, 185–209.
- Chaboureau, J.-P., F. Guichard, J.-L. Redelsperger, and J.-P. Lafore, 2004: The role of stability and moisture in the diurnal cycle of convection over land. *Quart. J. Roy. Meteor. Soc.*, **130** (604), 3105–3117.
- Chamberlain, J., C. Bain, D. Boyd, K. McCourt, T. Butcher, and S. Palmer, 2014: Forecasting storms over Lake Victoria using a high resolution model. *Meteor. Appl.*, **21** (2), 419–430.
- Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: a step-change in rainfall forecasting. *Meteor. Appl.*, **23** (2), 165–181.
- Cook, K. H., and E. K. Vizy, 2013: Projected changes in East African rainy seasons. *J. Climate*, **26** (16), 5931–5948.
- Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate*, **19** (18), 4605–4630.
- Dai, A., F. Giorgi, and K. Trenberth, 1999: Observed and model simulated precipitation diurnal cycle over the contiguous United States. *J. Geophys. Res.*, **104**, 6377–6402.
- Daley, R., 1988: The normal modes of the spherical non-hydrostatic equations with applications to the filtering of acoustic modes. *Tellus A*, **40** (2), 96–106.
- Datta, R., 1981: Certain aspects of monsoonal precipitation dynamics over Lake Victoria. *Monsoon Dynamics*, J. Lighthill, and R. Pearce, Eds., Cambridge University Press, 333–349.
- Davies, T., M. J. Cullen, A. J. Malcolm, M. Mawson, A. Staniforth, A. White, and N. Wood, 2005: A new dynamical core for the Met Office’s global and regional modelling of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **131** (608), 1759–1782.
- Davin, E. L., and S. I. Seneviratne, 2012: Role of land surface processes and diffuse/direct radiation partitioning in simulating the European climate. *Biogeosci.*,

**9 (5)**, 1695–1707.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137 (656)**, 553–597.

Done, J., C. A. Davis, and M. Weisman, 2004: The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) model. *Atmos. Sci. Lett.*, **5 (6)**, 110–117.

Dunning, C. M., E. Black, and R. P. Allan, 2018: Later wet seasons with more intense rainfall over Africa under future climate change. *Journal of Climate*, **31 (23)**, 9719–9738.

Dunning, C. M., E. C. Black, and R. P. Allan, 2016: The onset and cessation of seasonal rainfall over Africa. *Journal of Geophysical Research: Atmospheres*, **121 (19)**, 11–405.

Eagle, C., H. Lean, and S. Webster, 2015: Met Office Technical Report: Convective scale NWP model configurations for the Lake Victoria region. Tech. rep., Met Office.

East African Community Secretariat, 2014: Floods and Conflicts the major cause of Disasters in East Africa. Available at: <https://www.eac.int/press-releases/144-environment-natural-resources/863-floods-and-conflicts-the-major-cause-of-disasters-in-east-africa>. Last accessed: 12 June 2019.

Estoque, M. A., 1962: The sea breeze as a function of the prevailing synoptic situation. *Journal of the Atmospheric Sciences*, **19 (3)**, 244–250.

Farmer, G., 1988: Seasonal forecasting of the Kenya coast Short Rains, 1901–84. *J. Climatol.*, **8 (5)**, 489–497.

Findlater, J., 1977: Observational aspects of the low-level cross-equatorial jet stream of the western Indian Ocean. *Pure Appl. Geophys.*, **115 (5-6)**, 1251–1262.

Finney, D. L., and Coauthors, 2019: Implications of improved representation of convection for the East Africa water budget using a convection-permitting model. *J. Climate*, **(2019)**.

Flohn, H., and T. Burkhardt, 1985: Nile runoff at Aswan and Lake Victoria: a case of



- a discontinuous climate time series. *Z. Gletscherk. Glazialgeol*, **21**, 125–130.
- Flohn, H., and K. Fraedrich, 1966: Tagesperiodische zirkulation und niederschlagsverteilung am Victoria-See (Ostafrika) (The daily periodic circulation and distribution of rainfall over Lake Victoria, in German). *Meteorologische Rundschau*, **19 (6)**, 157–165.
- Fovell, R. G., 2005: Convective initiation ahead of the sea-breeze front. *Monthly weather review*, **133 (1)**, 264–278.
- Fraedrich, K., 1972: A simple climatological model of the dynamics and energetics of the nocturnal circulation at Lake Victoria. *Quart. J. Roy. Meteor. Soc.*, **98 (416)**, 322–335.
- Giorgi, F., M. R. Marinucci, and G. T. Bates, 1993a: Development of a second-generation regional climate model (RegCM2). Part I: Boundary-layer and radiative transfer processes. *Mon. Wea. Rev.*, **121 (10)**, 2794–2813.
- Giorgi, F., M. R. Marinucci, G. T. Bates, and G. De Canio, 1993b: Development of a second-generation regional climate model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions. *Mon. Wea. Rev.*, **121 (10)**, 2814–2832.
- Goler, R. A., and M. J. Reeder, 2004: The Generation of the Morning Glory. *J. Atmos. Sci.*, **61 (12)**, 1360–1376.
- Guichard, F., and Coauthors, 2004: Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quart. J. Roy. Meteor. Soc.*, **130 (604)**, 3139–3172.
- Guichard, F., and Coauthors, 2010: An intercomparison of simulated rainfall and evapotranspiration associated with a mesoscale convective system over West Africa. *Wea. Forecasting*, **25 (1)**, 37–60.
- Haltiner, G. J., and R. T. Williams, 1980: *Numerical prediction and dynamic meteorology*. 2nd ed., John Wiley & Sons, Inc., 477 pp.
- Hastenrath, S., 2007: Circulation mechanisms of climate anomalies in East Africa and the equatorial Indian Ocean. *Dyn. Atmos. Oceans*, **43 (1)**, 25–35.
- Hastenrath, S., A. Nicklis, and L. Greischar, 1993: Atmospheric-hydrospheric mecha-

- nisms of climate anomalies in the western equatorial Indian Ocean. *J. Geophys. Res.: Oceans (1978–2012)*, **98 (C11)**, 20 219–20 235.
- Hastenrath, S., D. Polzin, and C. Mutai, 2007: Diagnosing the 2005 drought in equatorial East Africa. *J. Climate*, **20 (18)**, 4628–4637.
- Hastenrath, S., D. Polzin, and C. Mutai, 2010: Diagnosing the droughts and floods in equatorial East Africa during boreal autumn 2005–08. *J. Climate*, **23 (3)**, 813–817.
- Hastenrath, S., D. Polzin, and C. Mutai, 2011: Circulation mechanisms of Kenya rainfall anomalies. *J. Climate*, **24 (2)**, 404–412.
- Hastings, D. A., and P. K. Dunbar, 1999: Global Land One-kilometer Base Elevation (GLOBE) digital elevation model, documentation, volume 1.0. Key to Geophysical Records Documentation (KGRD), Doc. 34, National Oceanic and Atmospheric Administration, National Geophysical Data Center. Available at: <https://repository.library.noaa.gov/view/noaa/13424>.
- Hogan, E., A. Shelly, and P. Xavier, 2015: The observed and modelled influence of the Madden–Julian Oscillation on East African rainfall. *Meteor. Appl.*, **22 (3)**, 459–469.
- Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88 (11)**, 1783–1794.
- Houze, R. A., Jr, 1997: Stratiform precipitation in regions of convection: A meteorological paradox? *Bull. Amer. Meteor. Soc.*, **78 (10)**, 2179–2196.
- Houze, R. A., Jr, K. L. Rasmussen, M. D. Zuluaga, and S. R. Brodzik, 2015: The variable nature of convection in the tropics and subtropics: A legacy of 16 years of the Tropical Rainfall Measuring Mission satellite. *Rev. Geophys.*, **53 (3)**, 994–1021.
- Huffman, G., 2017: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V05, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Last accessed: 20 Dec 2017, DOI: 10.5067/GPM/IMERG/3B-HH/05.
- Huffman, G. J., and Coauthors, 2018: Algorithm Theoretical Basis Document (ATBD) Version 5.2. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Tech. rep., Available at: [https://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V5.2.pdf](https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2.pdf).

- Indeje, M., F. H. Semazzi, and L. J. Ogallo, 2000: ENSO signals in East African rainfall seasons. *Int. J. Climatol.*, **20** (1), 19–46.
- Keat, W. J., and Coauthors, 2019: Convective initiation and storm life cycles in convection-permitting simulations of the Met Office Unified Model over South Africa. *Quart. J. Roy. Meteor. Soc.*, **145**, 1323–1336.
- Keen, C. S., and W. A. Lyons, 1978: Lake/land breeze circulations on the western shore of Lake Michigan. *J. Appl. Meteor.*, **17** (12), 1843–1855.
- Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior, 2014: Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nat. Climate Change*.
- Kent, C., R. Chadwick, and D. P. Rowell, 2015: Understanding uncertainties in future projections of seasonal tropical precipitation. *Journal of Climate*, **28** (11), 4390–4413.
- Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47** (2).
- Kingsmill, D. E., and A. N. Crook, 2003: An observational study of atmospheric bore formation from colliding density currents. *Mon. Wea. Rev.*, **131** (12), 2985–3002.
- Lafore, J. P., and Coauthors, 2017: Deep Convection. *Meteorology of tropical West Africa: The forecasters' handbook*, D. J. Parker, and M. Diop-Kane, Eds., 1st ed., John Wiley & Sons, chap. 3, 90–129.
- Laird, N. F., D. A. Kristovich, R. M. Rauber, H. T. Ochs III, and L. J. Miller, 1995: The Cape Canaveral sea and river breezes: Kinematic structure and convective initiation. *Monthly weather review*, **123** (10), 2942–2956.
- Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.*, **136** (9), 3408–3424.
- Liebmann, B., and Coauthors, 2014: Understanding recent eastern Horn of Africa rainfall variability and change. *Journal of Climate*, **27** (23), 8630–8645.
- Lilly, D. K., 1968: Models of cloud-topped mixed layers under a strong inversion. *Quart.*

- J. Roy. Meteor. Soc.*, **94** (401), 292–309.
- Linden, P., and J. Simpson, 1986: Gravity-driven flows in a turbulent fluid. *J. Fluid Mech.*, **172**, 481–497.
- Liu, Z., D. Ostrenga, W. Teng, and S. Kempler, 2012: Tropical Rainfall Measuring Mission (TRMM) precipitation data and services for research and applications. *Bull. Amer. Meteor. Soc.*, **93** (9), 1317–1325.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26** (4), 636–646.
- Luo, Y., and Y. Chen, 2015: Investigation of the predictability and physical mechanisms of an extreme-rainfall-producing mesoscale convective system along the Meiyu front in East China: An ensemble approach. *J. Geophys. Res.: Atmospheres*, **120** (20), 10 593–10 618.
- Lyon, B., and D. G. DeWitt, 2012: A recent and abrupt decline in the east african long rains. *Geophysical Research Letters*, **39** (2).
- Lyons, W. A., 1972: The climatology and prediction of the Chicago lake breeze. *J. Appl. Meteor.*, **11** (8), 1259–1270.
- Lyons, W. A., and L. E. Olsson, 1972: Mesoscale air pollution transport in the Chicago lake breeze. *J. Air Pollut. Control Assoc.*, **22** (11), 876–881.
- MacCallum, S., and C. Merchant, 2011: ARC-lake: data product description. Tech. rep., University of Edinburgh. Available at: [http://www.geos.ed.ac.uk/arclake/ARCLake\\_DPD\\_v1.1.1](http://www.geos.ed.ac.uk/arclake/ARCLake_DPD_v1.1.1).
- MacCallum, S. N., and C. J. Merchant, 2012: Surface water temperature observations of large lakes by optimal estimation. *Canadian Journal of Remote Sensing*, **38** (1), 25–45.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141** (689), 1072–1084.
- Madden, R. A., and P. R. Julian, 1971: Detection of a 40-50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28** (5), 702–708.

- Madden, R. A., and P. R. Julian, 1972: Description of global-scale circulation cells in the tropics with a 40-50 day period. *J. Atmos. Sci.*, **29** (6), 1109–1123.
- Mak, M. K., and J. E. Walsh, 1976: On the relative intensities of sea and land breezes. *J. Atmos. Sci.*, **33** (2), 242–251.
- Marchant, R., C. Mumbi, S. Behera, and T. Yamagata, 2007: The Indian Ocean dipole—the unsung driver of climatic variability in East Africa. *African Journal of Ecology*, **45** (1), 4–16.
- Marsham, J. H., N. S. Dixon, L. Garcia-Carreras, G. Lister, D. J. Parker, P. Knippertz, and C. E. Birch, 2013: The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophys. Res. Lett.*, **40** (9), 1843–1849.
- Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69** (11), 3350–3371.
- Miller, S., B. Keim, R. Talbot, and H. Mao, 2003: Sea breeze: Structure, forecasting, and impacts. *Rev. Geophys.*, **41** (3).
- Mittermaier, M., 2012: Verification. Tech. rep., Met Office. Paper 17.6, MOSAC-17. Available at: [http://www.metoffice.gov.uk/media/pdf/f/3/MOSAC\\_17.6\\_Mittermaier.pdf](http://www.metoffice.gov.uk/media/pdf/f/3/MOSAC_17.6_Mittermaier.pdf).
- Mittermaier, M., N. Roberts, and S. A. Thompson, 2013: A long-term assessment of precipitation forecast skill using the Fractions Skill Score. *Meteor. Appl.*, **20** (2), 176–186.
- Moncrieff, M. W., D. E. Waliser, M. J. Miller, M. A. Shapiro, G. R. Asrar, and J. Caughey, 2012: Multiscale convective organization and the YOTC virtual global field campaign. *Bull. Amer. Meteor. Soc.*, **93** (8), 1171–1187.
- Moron, V., P. Camberlin, and A. W. Robertson, 2013: Extracting subseasonal scenarios: an alternative method to analyze seasonal predictability of regional-scale tropical rainfall. *J. Climate*, **26** (8), 2580–2600.
- Moroz, W. J., 1967: A lake breeze on the eastern shore of Lake Michigan: Observations and model. *J. Atmos. Sci.*, **24** (4), 337–355.
- Moroz, W. J., and E. Hewson, 1966: The mesoscale interaction of a lake breeze and low

- level outflow from a thunderstorm. *Journal of Applied Meteorology*, **5 (2)**, 148–155.
- Mukabana, J. R., and R. A. Pielke, 1996: Investigating the influence of synoptic-scale monsoonal winds and mesoscale circulations on diurnal weather patterns over Kenya using a mesoscale numerical model. *Mon. Wea. Rev.*, **124 (2)**, 224–244.
- Mutai, C., M. Ward, and A. Colman, 1998: Towards the prediction of the East Africa short rains based on sea-surface temperature–atmosphere coupling. *Int. J. Climatol.*, **18 (9)**, 975–997.
- Mutai, C. C., and M. N. Ward, 2000: East African rainfall and the tropical circulation/convection on intraseasonal to interannual timescales. *J. Climate*, **13 (22)**, 3915–3939.
- Muwembe, K. Y., 2012: Verification of mobile weather alert forecasts over Lake Victoria in Uganda. M.S. thesis, University of Reading.
- Neumann, J., and Y. Mahrer, 1975: A theoretical study of the lake and land breezes of circular lakes. *Mon. Wea. Rev.*, **103 (6)**, 474–485.
- Nicholson, S., 1996: A review of climate dynamics and climate variability in Eastern Africa. *The limnology, climatology and paleoclimatology of the East African lakes*, 25–56.
- Nicholson, S. E., 2014: The predictability of rainfall over the Greater Horn of Africa. Part I: Prediction of seasonal rainfall. *J. Hydrometeor.*, **15 (3)**, 1011–1027.
- Nicholson, S. E., 2015a: Long-term variability of the East African ‘short rains’ and its links to large-scale factors. *Int. J. Climatol.*
- Nicholson, S. E., 2015b: The predictability of rainfall over the Greater Horn of Africa. Part II: Prediction of monthly rainfall during the long rains. *J. Hydrometeor.*, **16 (5)**, 2001–2012.
- Nicholson, S. E., 2017: Climate and climatic variability of rainfall over eastern Africa. *Reviews of Geophysics*, **55 (3)**, 590–635.
- Ogallo, L., 1988: Relationships between seasonal rainfall in East Africa and the Southern Oscillation. *Int. J. Climatol.*, **8 (1)**, 31–43.
- Ogallo, L., 1989: The spatial and temporal patterns of the East African seasonal rainfall

- derived from principal component analysis. *Int. J. Climatol.*, **9** (2), 145–167.
- Ogallo, L., P. Bessemoulin, J.-P. Ceron, S. Mason, and S. J. Connor, 2008: Adapting to climate variability and change: the Climate Outlook Forum process. *Bull. World Meteor. Organ.*, **57** (2), 93–102.
- Okeyo, A. E., 1986: The impact of Lake Victoria on the convective activities over the Kenya Highlands. *J. Meteorolog. Soc. Japan Ser. II*, **64**, 689–695.
- Okoola, R., 1998: Spatial evolutions of the active convective patterns across the Equatorial Eastern Africa region during northern hemisphere spring season using Outgoing Longwave Radiation records. *Meteor. Atmos. Phys.*, **66** (1-2), 51–63.
- Okoola, R. E., 1999: A diagnostic study of the eastern Africa monsoon circulation during the Northern Hemisphere spring season. *Int. J. Climatol.*, **19** (2), 143–168.
- Omeny, P. A., L. Ogallo, R. Okoola, H. Hendon, and M. Wheeler, 2008: East African rainfall variability associated with the Madden-Julian Oscillation. *J. Kenya Meteor. Soc.*, **2** (2), 105–114.
- Onyutha, C., H. Tabari, A. Rutkowska, P. Nyeko-Ogiramoi, and P. Willems, 2016: Comparison of different statistical downscaling methods for climate change rainfall projections over the Lake Victoria basin considering CMIP3 and CMIP5. *J. Hydro-environ. Res.*, **12**, 31–45.
- Otieno, V. O., and R. O. Anyah, 2013: CMIP5 simulated climate conditions of the Greater Horn of Africa (GHA). Part II: projected climate. *Climate dynamics*, **41** (7-8), 2099–2113.
- Pal, J. S., and Coauthors, 2007: Regional climate modeling for the developing world: the ICTP RegCM3 and RegCNET. *Bull. Amer. Meteor. Soc.*, **88** (9), 1395–1409.
- Parker, D. J., and Coauthors, 2017: Local Weather. *Meteorology of tropical West Africa: The forecasters' handbook*, D. J. Parker, and M. Diop-Kane, Eds., 1st ed., John Wiley & Sons, chap. 4, 130–174.
- Passarelli, R. E., Jr, and R. R. Braham, Jr, 1981: The role of the winter land breeze in the formation of Great Lake snow storms. *Bull. Amer. Meteor. Soc.*, **62** (4), 482–491.
- Physick, W., 1976: A numerical model of the sea-breeze phenomenon over a lake or gulf. *J. Atmos. Sci.*, **33** (11), 2107–2135.

- Pielke, R., 2015: Mountain Meteorology — Land and Sea Breezes. *Encyclopedia of Atmospheric Sciences*, G. R. North, J. Pyle, and F. Zhang, Eds., 2nd ed., Academic Press, Oxford, 80–83, Retrieved from <https://app.knovel.com/hotlink/khtml/id:kt011A8Q72/encyclopedia-atmospheric/land-and-sea-breezes>.
- Pielke, R., and M. Segal, 1986: Mesoscale Circulations Forced by Differential Terrain Heating. *Mesoscale Meteorology and Forecasting*, P. Ray, Ed., Springer, 516–548.
- Pielke, R. A., 1974: A three-dimensional numerical model of the sea breezes over south Florida. *Mon. Wea. Rev.*, **102** (2), 115–139.
- Pielke, R. A., and Coauthors, 1992: A comprehensive meteorological modeling system—RAMS. *Meteor. Atmos. Phys.*, **49** (1-4), 69–91.
- Pohl, B., and P. Camberlin, 2006a: Influence of the Madden-Julian Oscillation on East African rainfall. Part I: intraseasonal variability and regional dependency. *Quart. J. Roy. Meteor. Soc.*, **132** (621), 2521–2539.
- Pohl, B., and P. Camberlin, 2006b: Influence of the Madden-Julian Oscillation on East African rainfall. Part II: March-May season extremes and interannual variability. *Quart. J. Roy. Meteor. Soc.*, **132** (621B), 2541–2558.
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53** (2), 323–361.
- Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136** (1), 78–97.
- Rockel, B., A. Will, and A. Hense, 2008: The regional climate model COSMO-CLM (CCLM). *Meteorologische Zeitschrift*, **17** (4), 347–348.
- Rowell, D. P., B. B. Booth, S. E. Nicholson, and P. Good, 2015: Reconciling past and future rainfall trends over east Africa. *J. Climate*, **28** (24), 9768–9788.
- Saji, N., B. N. Goswami, P. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401** (6751), 360–363.



- Schumacher, R. S., A. J. Clark, M. Xue, and F. Kong, 2013: Factors influencing the development and maintenance of nocturnal heavy-rain-producing convective systems in a storm-scale ensemble. *Mon. Wea. Rev.*, **141** (8), 2778–2801.
- Semazzi, F., 2011: Enhancing safety of navigation and efficient exploitation of natural resources over Lake Victoria and its basin by strengthening meteorological services on the lake. Tech. rep., North Carolina State University Climate Modeling Laboratory.
- Seto, K. C., B. Güneralp, and L. R. Hutyra, 2012: Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Natl. Acad. Sci.*, **109** (40), 16 083–16 088.
- Shongwe, M. E., G. J. van Oldenborgh, B. van den Hurk, and M. van Aalst, 2011: Projected changes in mean and extreme precipitation in Africa under global warming. Part II: East Africa. *J. Climate*, **24** (14), 3718–3733.
- Sills, D., J. Brook, I. Levy, P. Makar, J. Zhang, and P. Taylor, 2011: Lake breezes in the southern Great Lakes region and their influence during BAQS-Met 2007. *Atmos. Chem. Phys.*, **11** (15), 7955–7973.
- Sills, D. M. L., 1998: Lake and land breezes in southwestern Ontario: observations, analyses and numerical modeling. Ph.D. thesis, York University, Toronto. Available at: [http://www.yorku.ca/pat/research/dsills/papers/PHD/Sills\\_Dissertation\\_1998.pdf](http://www.yorku.ca/pat/research/dsills/papers/PHD/Sills_Dissertation_1998.pdf).
- Simpson, J., 1969: A comparison between laboratory and atmospheric density currents. *Quart. J. Roy. Meteor. Soc.*, **95** (406), 758–765.
- Simpson, J., and R. Britter, 1980: A laboratory model of an atmospheric mesofront. *Quart. J. Roy. Meteor. Soc.*, **106** (449), 485–500.
- Simpson, J., D. Mansfield, and J. Milford, 1977: Inland penetration of sea-breeze fronts. *Quart. J. Roy. Meteor. Soc.*, **103** (435), 47–76.
- Simpson, J. E., 1994: *Sea breeze and local winds*. Cambridge University Press.
- Simpson, J. E., 1999: *Gravity currents: In the environment and the laboratory*. 2nd ed., Cambridge University Press, 244 pp.
- Skamarock, W. C., and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J. Computational Phys.*,

**227 (7)**, 3465–3485.

Song, Y., F. H. Semazzi, and L. Xie, 2002: Development of a coupled regional climate simulation model for the Lake Victoria basin. *The East African Great Lakes: Limnology, Palaeolimnology and Biodiversity*, Springer, 153–186.

Song, Y., F. H. Semazzi, L. Xie, and L. J. Ogallo, 2004: A coupled regional climate model for the Lake Victoria basin of East Africa. *Int. J. Climatol.*, **24 (1)**, 57–75.

Stein, T., and Coauthors, 2019: An evaluation of clouds and precipitation in convection-permitting forecasts for South Africa. *Wea. Forecasting*, **34 (1)**, 233–254.

Stein, T. H., R. J. Hogan, P. A. Clark, C. E. Halliwell, K. E. Hanley, H. W. Lean, J. C. Nicol, and R. S. Plant, 2015: The DYMECS project: A statistical approach for the evaluation of convective storms in high-resolution NWP models. *Bull. Amer. Meteor. Soc.*, **(2015)**.

Stephens, G. L., and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.: Atmospheres (1984–2012)*, **115 (D24)**.

Stratton, R. A., and Coauthors, 2018: A Pan-African Convection-Permitting Regional Climate Simulation with the Met Office Unified Model: CP4-Africa. *J. Climate*, **31 (9)**, 3485–3508.

Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95 (3)**, 409–426.

Sun, X., L. Xie, F. Semazzi, and B. Liu, 2015: Effect of lake surface temperature on the spatial distribution and intensity of the precipitation over the Lake Victoria basin. *Mon. Wea. Rev.*, **143 (4)**, 1179–1192.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? *J. Climate*, **19 (6)**, 916–934.

Talling, J., 1966: The annual cycle of stratification and phytoplankton growth in Lake Victoria (East Africa). *International Review of Hydrobiology*, **51 (4)**, 545–621.

Tapp, M., and P. White, 1976: A non-hydrostatic mesoscale model. *Quart. J. Roy. Meteor. Soc.*, **102 (432)**, 277–296.

Terblanche, D., G. Pegram, and M. Mittermaier, 2001: The development of weather

- radar as a research and operational tool for hydrology in South Africa. *J. Hydrol.*, **241** (1-2), 3–25.
- Thiery, W., E. L. Davin, H.-J. Panitz, M. Demuzere, S. Lhermitte, and N. Van Lipzig, 2015: The impact of the African Great Lakes on the regional climate. *J. Climate*, **28** (10), 4061–4085.
- Thiery, W., E. L. Davin, S. I. Seneviratne, K. Bedka, S. Lhermitte, and N. P. van Lipzig, 2016: Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.*, **7**, 12786.
- Thiery, W., L. Gudmundsson, K. Bedka, F. H. Semazzi, S. Lhermitte, P. Willems, N. P. van Lipzig, and S. I. Seneviratne, 2017: Early warnings of hazardous thunderstorms over Lake Victoria. *Environ. Res. Lett.*, **12** (7), 074012.
- Tijm, A., and A. Van Delden, 1999: The role of sound waves in sea-breeze initiation. *Quart. J. Roy. Meteor. Soc.*, **125** (558), 1997–2018.
- Tollefson, J., 2019: Tropical Africa could be key to solving methane mystery. Available at: <https://www.nature.com/articles/d41586-019-00457-7>. Last accessed: 11 Sep 2019.
- Tomsett, J., 1975: The diurnal variation of precipitation in East Africa. Tech. rep., No. 25. EAMD.
- UN Economic Commission for Africa, 2019: Macroeconomic and Social Developments in Eastern Africa 2019. *Towards the Implementation of the African Continental Free Trade Area in Eastern Africa*. Executive Summary. Available at: [https://www.uneca.org/sites/default/files/uploaded-documents/SRO-EA/22ICE/Eng/sub-regional\\_profile\\_2019\\_executive\\_summary\\_eng.pdf](https://www.uneca.org/sites/default/files/uploaded-documents/SRO-EA/22ICE/Eng/sub-regional_profile_2019_executive_summary_eng.pdf). Last accessed: 12 June 2019.
- Vanderkelen, I., N. P. van Lipzig, and W. Thiery, 2018: Modelling the water balance of Lake Victoria (East Africa)—part 1: observational analysis. *Hydrol. Earth Syst. Sci.*, **22** (10), 5509–5525.
- Vié, B., O. Nuissier, and V. Ducrocq, 2011: Cloud-resolving ensemble simulations of Mediterranean heavy precipitating events: uncertainty on initial conditions and lateral boundary conditions. *Mon. Wea. Rev.*, **139** (2), 403–423.
- Vizy, E. K., and K. H. Cook, 2012: Mid-twenty-first-century changes in extreme events

- over northern and tropical Africa. *J. Climate*, **25** (17), 5748–5767.
- Voidrot-Martinez, M.-F., and C. Berthou, 1991: SYNERGIE: A workstation for operational weather monitoring and forecasting. *Third Workshop on Meteorological Operational Systems, 18-22 November 1991*, ECMWF, Shinfield Park, Reading, ECMWF, 241-245, URL <https://www.ecmwf.int/node/12965>.
- Wainwright, C. M., J. H. Marsham, R. J. Keane, D. P. Rowell, D. L. Finney, E. Black, and R. P. Allan, 2019: ‘Eastern African Paradox’ rainfall decline due to shorter not less intense Long Rains. *npj Climate and Atmospheric Science*, **2** (1), 1–9.
- Walker, D. P., C. E. Birch, J. H. Marsham, A. A. Scaife, R. J. Graham, and Z. T. Segele, 2019: Skill of dynamical and GHACOF consensus seasonal forecasts of East African rainfall. *Climate Dyn.*, 1–25.
- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23** (3), 407–437.
- Weusthoff, T., F. Ament, M. Arpagaus, and M. W. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Mon. Wea. Rev.*, **138** (9), 3418–3433.
- Wexler, R., 1946: Theory and observations of land and sea breezes. *Bull. Amer. Meteor. Soc.*, **27** (6), 272–287.
- Williams, A. P., and C. Funk, 2011: A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Climate Dyn.*, **37** (11-12), 2417–2435.
- Williams, K., J. Chamberlain, C. Buontempo, and C. Bain, 2015: Regional climate model performance in the Lake Victoria basin. *Climate Dyn.*, **44** (5-6), 1699–1713.
- Yang, G.-Y., and J. Slingo, 2001: The diurnal cycle in the tropics. *Mon. Wea. Rev.*, **129** (4), 784–801.
- Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2014: The East African long rains in observations and models. *Journal of Climate*, **27** (19), 7185–7202.
- Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2015: The Annual Cycle of East African Precipitation. *J. Climate*, **28** (6), 2385–2404.

Yin, X., and S. E. Nicholson, 1998: The water balance of Lake Victoria. *Hydrol. Sci. J.*, **43** (5), 789–811.



## Chapter 2.

# What Is the Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall over Tropical East Africa?

Published in *Monthly Weather Review* (2018)





## What Is the Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall over Tropical East Africa?

BETH J. WOODHAMS,<sup>a</sup> CATHRYN E. BIRCH,<sup>a,b</sup> JOHN H. MARSHAM,<sup>a,c</sup> CAROLINE L. BAIN,<sup>b</sup>  
NIGEL M. ROBERTS,<sup>d</sup> AND DOUGLAS F. A. BOYD<sup>b</sup>

<sup>a</sup> *School of Earth and Environment, University of Leeds, Leeds, United Kingdom*

<sup>b</sup> *Met Office, Exeter, United Kingdom*

<sup>c</sup> *National Centre for Atmospheric Science, University of Leeds, Leeds, United Kingdom*

<sup>d</sup> *MetOffice@Reading, Met Office, Reading, United Kingdom*

(Manuscript received 31 December 2017, in final form 18 June 2018)

### ABSTRACT


Forecasting convective rainfall in the tropics is a major challenge for numerical weather prediction. The use of convection-permitting (CP) forecast models in the tropics has lagged behind the midlatitudes, despite the great potential of such models in this region. In the scientific literature, there is very little evaluation of CP models in the tropics, especially over an extended time period. This paper evaluates the prediction of convective storms for a period of 2 years in the Met Office operational CP model over East Africa and the global operational forecast model. A novel localized form of the fractions skill score is introduced, which shows variation in model skill across the spatial domain. Overall, the CP model and the global model both outperform a 24-h persistence forecast. The CP model shows greater skill than the global model, in particular on subdaily time scales and for storms over land. Forecasts over Lake Victoria are also improved in the CP model, with an increase in hit rate of up to 20%. Contrary to studies in the midlatitudes, the skill of both models shows a large dependence on the time of day and comparatively little dependence on the forecast lead time within a 48-h forecast. Although these results provide more motivation for forecasters to use the CP model to produce subdaily forecasts with increased detail, there is a clear need for more in situ observations for data assimilation into the models and for verification. A move toward ensemble forecasting could have further benefits.

### 1. Introduction

Forecasting tropical convection remains a huge challenge for numerical weather prediction (NWP). In particular, there is an urgent need to improve forecasting of high-impact weather, such as heavy convective precipitation, in sub-Saharan Africa. Because of their coarse resolution, traditional global models use parameterization schemes to represent convection. Parameterized models are not designed to produce realistic storm structures; they almost ubiquitously produce too much light rain and are unable to capture the highest-intensity events (Sun et al. 2006; Dai 2006; Stephens et al. 2010). Such models are unable to simulate the diurnal cycle; the convective

maximum tends to occur at midday, rather than late afternoon, as observed (Yang and Slingo 2001; Bechtold et al. 2004). Marsham et al. (2013) and Birch et al. (2014b) showed that an incorrect diurnal cycle may introduce errors into the synoptic-scale flow. In addition, flows over complex topography and those within deep convective storms cannot be well represented at coarse resolutions (Clark et al. 2016).

Low skill in tropical precipitation prediction means that forecasters often use other information from models, such as the synoptic-scale circulation and stability measures, to determine favorable conditions for deep convection. Consequently, predictions are heavily reliant on forecasters' experience and knowledge of the model and meteorology of the region (Lafore et al. 2017).

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Beth Woodhams, eebjw@leeds.ac.uk



This article is licensed under a [Creative Commons Attribution 4.0 license](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>).

Convection-permitting (CP) models provide a step change in the representation of convective storms because they explicitly represent the storms themselves (Clark et al. 2016). CP models produce more realistic-looking precipitation fields and have an improved diurnal cycle, with the peak in convection shifted further toward the observed late-afternoon maximum (Lean et al. 2008; Done et al. 2004; Weisman et al. 2008; Weusthoff et al. 2010; Birch et al. 2014b; Prein et al. 2015). Marsham et al. (2013) and Garcia-Carreras et al. (2013) have shown that the finer horizontal grid spacing of CP models allows the simulation of cold pools, which can affect synoptic-scale fluxes in the models and trigger new convection. CP models also better capture the organization and propagation of convection (Weisman et al. 2008; White et al. 2018). The benefits of using CP models are even being utilized for future climate projections to provide more details on regional and local scales (Prein et al. 2015; Stratton et al. 2018).

However, increased realism of rainfall may not translate to increased forecast skill. CP models can produce too much rain, in particular at high intensities, while the proportion of low-intensity events is too small (Lean et al. 2008; Kendon et al. 2012; Marsham et al. 2013; Done et al. 2004). The horizontal grid spacing remains on a scale larger than most convective updrafts, meaning that cloud structures are still underresolved, causing errors in the multiscale interactions, upscale growth, and timing of storms (Clark et al. 2016; Lean et al. 2008). Lean et al. (2008) showed that CP models often perform poorly at the start of runs, since it takes time for the high-resolution detail to “spin up” from the initial fields provided by the driving model.

High-resolution simulations of mesoscale convective systems (MCSs) in various regions of the world have shown that the accuracy of the initial conditions (ICs) and boundary conditions provided by the driving model are key factors responsible for the correct prediction of initial development of a storm (Birch et al. 2013; Guichard et al. 2010; Melhauser and Zhang 2012; Schumacher et al. 2013; Luo and Chen 2015; Vié et al. 2011). Sensitivity to ICs is a particular concern in much of the tropics because of a lack of routine observations for data assimilation into the driving model. These studies also showed that the predictability of a storm may depend on the type of synoptic flow. High-resolution CP models are fundamentally limited by the fact that predictability times are reduced on cloud-resolving scales, compared to synoptic scales. Small-scale errors may grow more quickly and influence larger scales (Lorenz 1969; Hohenegger and Schär 2007).

In recent years, as computing power has increased, CP models have become more feasible and more commonly

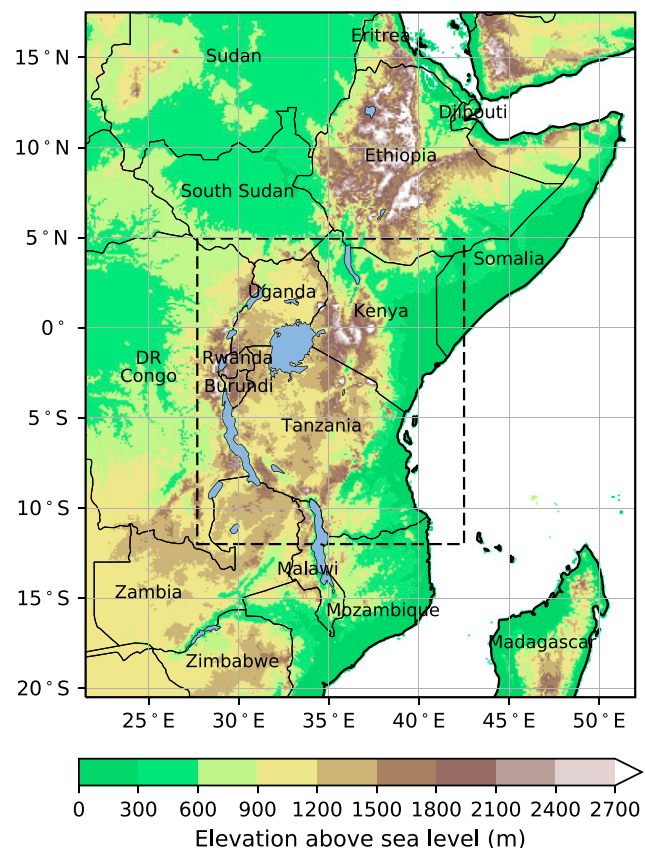


FIG. 1. A map showing the orography over the domain spanned by the current operational Met Office CP model in East Africa. Orography data are from the Global Self-Consistent, Hierarchical, High-Resolution Geography (GSHHG) database (Wessel and Smith 1996). The dashed box encloses the LV subdomain used in the analysis.

used in operational weather forecasts. They have been used at a country level by many national meteorological services for over a decade. The Met Office uses a CP version of the Met Office Unified Model (MetUM) over the United Kingdom and some tropical domains (Tang et al. 2013); the U.S. Weather Research and Forecasting (WRF) Model is CP and used by multiple agencies worldwide (Michalakes et al. 2001), as is the CP version of COSMO, created in Germany (Baldauf et al. 2011). The Application of Research to Operations at Mesoscale (AROME) model developed in France (Seity et al. 2011) and the Japan Meteorological Agency’s Non-hydrostatic Mesoscale Model over Japan (Saito et al. 2006) are also CP models. Many of these models are configured to run operationally beyond their country of origin. Although many CP forecast models are used across the tropics, published verification is limited, especially for an extended period.

Equatorial East Africa (Fig. 1) is a region that receives the majority of its rainfall via deep convective storms. The prediction of severe convection is acutely important in this region, which is at high risk of severe

flooding and drought. It is particularly important to provide accurate forecasts for the Lake Victoria (LV) basin. This densely populated region supports around 35 million people, including 200 000 fishermen. Nocturnal storms over Lake Victoria produce intense precipitation and high winds, which capsize boats. There are an estimated 5000 deaths on the lake each year, with many attributed to severe weather (World Bank DGF 2011). Forecasting convection in this environment with orography and land–lake circulations is a challenge. The localized nature of convection and a lack of observations, especially upper air, to assimilate into forecasts add to the difficulty.

In 2011, the Met Office began running an operational CP forecast model over Lake Victoria, with a horizontal grid spacing of 4.4 km, funded through the Met Office Voluntary Cooperation Programme (VCP) and intended to aid the forecast of severe weather events, in particular over the Lake Victoria basin (Chamberlain et al. 2014). The domain was extended to cover a larger region in February 2014, which currently remains operational. Output from the model is disseminated to operational meteorologists in East Africa (principally, Kenya, Uganda, Tanzania, Rwanda, and Burundi). Model output is available to view on the VCP Africa Web Viewer, a password-controlled site open to African forecasters, which also shows output from the global model and recent satellite imagery and arrival time difference (ATD) lightning.

Some verification and comparison with the global MetUM was performed on the smaller domain by Chamberlain et al. (2014) for spring 2012. Overall, and in agreement with studies of other CP models, both the global and the CP models produced too much light rainfall, especially the global model. The CP model predicted too many intense rainfall events, whereas the global model was unable to produce any of the highest-intensity rainfall rates observed. Objective analysis showed the CP model to have more skill in predicting when a storm would occur, compared to the global model. However, it was also shown to overpredict severe events, leading to more “false alarms.”

Away from NWP, Thiery et al. (2017) developed a prototype of a statistical storm predictor [Lake Victoria Intense Early Warning System (VIEWS)] that uses satellite observations to forecast nocturnal storms over the lake. The model is based on the strong correlation between intense storms over land during the afternoon and intense nocturnal storms over Lake Victoria found by Thiery et al. (2016). While this prototype can achieve high hit rates and low false alarm rates, it has a lead time of only a few hours. NWP models are vital for providing much earlier warnings, with statistical model and nowcasting techniques playing an important role in warning confidence and refinement closer to the event.

This paper investigates whether a CP model over East Africa provides additional skill for the forecasting of severe tropical rainfall, compared to a global model. The rainfall field in the model was chosen for verification following meetings with forecasters from East Africa. While forecasters use other model fields, such as surface pressure, relative humidity, and winds, to ensure that the model rainfall field is realistic, the rainfall field is the main tool used to produce forecasts. The most commonly used output is 24-h rainfall accumulations, despite the availability of accumulations at 3-hourly intervals. A particular aim of the paper is to determine whether the CP model provides any added value on subdaily time scales or at specific locations. The effect of forecast lead time and time of day is also investigated. Much of the analysis is performed using the fractions skill score (FSS) proposed by Roberts and Lean (2008), along with an examination of model biases. The analysis builds on that of Chamberlain et al. (2014) by performing more detailed verification on a larger model domain over a period of 2 years. The paper emphasizes the implications of the findings on the operational aspects of the model, including how it is run and how it may best be used by forecasters.

The model configurations, observational data, and verification methods are introduced in section 2. Section 3 presents the characteristics of precipitation and its diurnal cycle in the models, alongside verification of the model skill at different spatial scales and at different times of day. The implications of these results are discussed and conclusions drawn in section 4.

## 2. Methods

This study consists of a comparison of the forecast skill of the MetUM CP model and the MetUM global model for rainfall over East Africa, alongside an analysis of the performance of the models at different forecast lead times and different times of day. The period of study is the 2 years between 18 July 2014 and 17 July 2016. Analysis was performed over the whole model domain, as well as a subdomain centered on LV, shown within the dashed box in Fig. 1 and chosen to include all countries using the CP model.

### a. Models

Both the CP model over East Africa and the global model are subsets of the MetUM and were used operationally in 2017. The start date of the analysis period marks the upgrade of the global model to GA6.1 and the Even Newer Dynamics for General atmospheric modeling of the environment (ENDGame) dynamical core (Wood et al. 2014), with a reduced horizontal grid spacing of

approximately 17 km in the meridional direction by 25 km in the zonal direction (in the tropics). The model is initialized every 6 h, at 0000, 0600, 1200, and 1800 UTC, although only the 0000 and 1200 UTC initializations are considered to allow direct comparison with the CP model. Convection is parameterized in the global model using the mass-flux scheme introduced by Gregory and Rowntree (1990) and with subsequent enhancements.

The CP model has a horizontal grid spacing of 4.4 km in both directions and spans the domain from 20.5°S to 17.5°N and from 21.5° to 52°E (Fig. 1). The model grid consists of  $762 \times 950$  grid points and has 70 vertical levels up to a lid of 40 km. The model is run with a time step of 100 s. Convection is treated explicitly, although the model also uses a convective available potential energy (CAPE)-dependent closure scheme (Roberts 2003) used by the Met Office in all operational models with this configuration. This scheme adjusts the time scale over which instability is removed in accordance with the amount of CAPE that is present in order to restrict the parameterized mass flux. This allows explicit convection while trying to parameterize the effects of smaller clouds that the CP model cannot resolve. The parameters within the microphysics and subgrid mixing have not been “tuned” for the tropics, but are the same as those used in the 4-km U.K. (UK4) model (Eagle et al. 2015). The New Dynamics (ND) dynamical core (Davies et al. 2005) was replaced with ENDGame [along with some other physics changes detailed in Eagle et al. (2015)] approximately halfway through the study period. Eagle et al. (2015) ran the model in both configurations and found that they produced very similar results, suggesting that the model may be considered to have similar behavior throughout the whole analysis period.

The CP model is initialized twice per day, at 2100 and 0900 UTC, using 3-h forecasts from the global model (i.e., ICs are taken from 3 h into the global model 1800 and 0600 UTC runs, respectively). The models are allowed 3 h for spin up until the first diagnostics are outputted at 0000 and 1200 UTC, respectively. Forecasts are outputted up to a lead time of 48 h from the first diagnostics (Chamberlain et al. 2014). The two initializations are referred to as the 0000 and 1200 UTC initializations since these are the times of the first available forecasts. The model is forced by lateral boundary conditions from the global model, updated every 3 h. Smoothing is applied at the edges to remove discontinuities between the models.

In both models, lake surface temperatures (LSTs) are prescribed as the foundation water surface temperature (temperature below the diurnal warm layer) taken from daily Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA), available on a  $1/20^\circ$  ( $\sim 6$  km) grid

(Fiedler et al. 2014). Observations are obtained from in situ data received via the Global Telecommunication System (GTS; although no in situ observations existed over Lake Victoria during this study) and satellite sea surface temperature (SST) data from the Group for High-Resolution SST (GHRSSST). Only overnight observations and daytime observations for wind speeds greater than  $6 \text{ m s}^{-1}$  are used, such that the effect of the diurnal warm layer above the sea surface is not included. These observations are assimilated onto a background field of the analysis from the previous day, slightly relaxed toward the MacCallum and Merchant (2011) ARC-Lake nighttime climatology (Donlon et al. 2012). Using OSTIA, the temperature in the model is updated once per day to the predawn value. When observations are unavailable, the LST in OSTIA will relax toward the ARC-Lake climatology over a period of 30 days (Fiedler et al. 2014).

At the time of the study, the OSTIA system included no lake-specific processing; satellite retrievals were optimized for over oceans, meaning that the different wind and cloud regimes and the elevations and continental locations of some lakes may have introduced errors into the satellite retrievals (Fiedler et al. 2014). However, Fiedler et al. (2014) note that Lake Victoria performs very well, compared to other lakes across the globe, likely due to its large size, position on the equator, and relatively low elevation. This is important since Thiery et al. (2015) and Argent et al. (2015) showed that accurate LSTs are crucial for reproducing observed precipitation patterns over and around Lake Victoria. In addition, Lakes Malawi and Tanganyika (also in the model domain) are said to perform well. In JJA 2009, the lakes in this region were found to have an average of several hundred observations per day due to satellite overpasses (Fiedler et al. 2014). However, since JJA is a dry period, the number of observations during the wet seasons may be decreased as sampling issues due to cloud cover increase.

For some analysis purposes, a time series of model data was required. For both models, a set of four different time series was formed by stitching together data from either forecast lead times between  $T + 12$  and  $T + 33$  h or between  $T + 24$  and  $T + 45$  h, for both the 0000 and 1200 UTC initializations. In addition, four different 24-h accumulation periods were defined using the same lead time bounds and the two initialization times.

### b. Observations

Few in situ observations are available for the region during the analysis period. There were no working radars and very few weather stations, forcing a reliance on satellite-derived data for observations. Precipitation

intensity observations are sourced from the Global Precipitation Measurement (GPM) mission, in particular the IMERG Final Precipitation version 5 (V05) level 3 product at  $0.1^\circ$  (Huffman 2017; Huffman et al. 2018). The product is available at 30-min time steps, but only times matching the model output were used. GPM IMERG is the successor to the Tropical Rainfall Measuring Mission (TRMM) 3B42 V7 product (Liu et al. 2012; TRMM 2011). IMERG takes observations from a network of satellites in the GPM constellation and unifies them to create a gridded product. In particular, the GPM *Core Observatory* satellite hosts a dual-frequency precipitation radar (DPR) and a conical-scanning multichannel microwave imager (GMI) (Hou et al. 2014). These two instruments are used as a reference to intercalibrate the passive microwave (PMW) precipitation estimates from other satellites in the GPM constellation using the method developed for TRMM by Huffman et al. (2007) and including calibration against the Global Precipitation Climatology Centre (GPCC) gauge analysis by Schneider et al. (2008) (Huffman et al. 2018; Hou et al. 2014). Since the PMW sensors do not have complete coverage of the ground, the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center morphing technique with Kalman filter (CMORPH-KF) (Joyce et al. 2004; Joyce and Xie 2011) is applied, which estimates precipitation outside the sensed area by propagation of PMW estimates with motion vectors derived from geosynchronous IR satellite imagery. For even further coverage, the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Cloud Classification System (PERSIANN–CCS) uses IR retrievals, calibrated against PMW retrievals, to estimate rainfall (Hong et al. 2004; Sorooshian et al. 2000). For brevity, the GPM IMERG product is referred to as GPM for the remainder of the paper.

Many studies in different parts of the world, including many with complex topography, have shown that the GPM product outperforms its predecessor, the TRMM 3B42 V7 product, on various spatial and temporal scales (Tang et al. 2016a,b; Kim et al. 2017; Xu et al. 2017; Wang et al. 2017b; Sharifi et al. 2016; Prakash et al. 2018). In particular, GPM is better able to detect low-intensity rainfall due to four more high-frequency channels on the GMI instrument, compared to the corresponding instrument for TRMM. From studies over mainland China, Tang et al. (2016a) suggested that GPM required improvement in dry climates and high altitudes. This could be a cause for concern over the Horn of Africa and East African highlands. Kim et al. (2017) noted uncertainties with orographic convection, and Xu et al. (2017) found that GPM has issues for

orography greater than 4500 m. O and Kirstetter (2018) found an underestimation of the diurnal variation over mountains in the United States. Further studies have shown that GPM underestimates high-intensity precipitation events (Wang et al. 2017a; O et al. 2017).

For comparison with GPM, the bias-corrected V1.0 CMORPH precipitation intensity dataset on an 8-km grid (NCEP 2017) is also used. The CMORPH algorithm by Joyce et al. (2004) is also used to produce GPM as described above, but for CMORPH, V1.0, is used without the Joyce and Xie (2011) Kalman filter. As for GPM, CMORPH is available every 30 min, but only times matching the model output are used. The performance of CMORPH has been shown to be very variable, both spatially and temporally (Zeweldi and Gebremichael 2009; Habib et al. 2012; Haile et al. 2013). Studies have shown that CMORPH is often unable to capture the highest-intensity events and therefore overestimates the frequency of lower-intensity rain events (Kumar et al. 2016; Habib et al. 2012). However, over complex terrain in Mexico, Nesbitt et al. (2008) found an overestimation of precipitation within deep convective systems. Several studies over Ethiopia found a general underestimation of rainfall rates (Haile et al. 2013; Romilly and Gebremichael 2011; Hirpa et al. 2010). Romilly and Gebremichael (2011) note that over northwestern Ethiopia, where the ITCZ has a strong effect and the climate is humid, CMORPH tends to overestimate rainfall at low elevations but performs well at higher elevations. Contrastingly, in the northeast, Hirpa et al. (2010) found an underestimation at high elevations. This shows the variability in performance of CMORPH over even a small area.

All model and observational data were interpolated onto the same grid for analysis. The global model has the coarsest resolution ( $\sim 17 \text{ km} \times 25 \text{ km}$ ), but square grid boxes were required for some of the analysis; therefore, all data were interpolated onto a regular  $0.25^\circ$  grid.

### c. Verification methods

#### 1) OBJECTIVE ANALYSIS

Objective analysis was performed to compare the skill of the models in forecasting storms over Lake Victoria. A storm was identified if rainfall occurred with a minimum given intensity over a minimum given area. A range of size and intensity thresholds was sampled. Stormy 3-h periods were identified in both the models and observations to obtain the number of hits (storm in both the observations and forecast), false alarms (storm was forecast but did not occur), misses (storm occurred but was not forecast), and correct negatives (storm was not forecast and did not occur). These were then used to compute the hit rate [hits/(hits + misses)] and false

alarm ratio [false alarms/(hits + false alarms)]. The false alarm rate [false alarms/(correct negatives + false alarms)] was also computed.

Using a fixed size and intensity threshold to define a storm over the lake in the observations, the hit rate was plotted against the false alarm rate for a variety of forecast intensity thresholds (keeping the size threshold the same as in the observations) to produce a receiver operating characteristic (ROC) curve (Swets 1973; Mason 1982). A good forecast should maximize the hit rate and minimize the false alarm rate; hence, the ROC curve should lie in the upper-left half of the ROC diagram. The closer the curve lies to false alarm rate = 0 and hit rate = 1, the greater the skill. This skill was captured by computing the area under the curve (AUC), which is greater than 0.5 for a skillful forecast and equal to 1 for a perfect forecast (Mason and Graham 2002; Wilks 2011).

## 2) FRACTIONS SKILL SCORE

The ability of the models to forecast storms at the correct time and in the correct location was assessed. The small horizontal grid spacing of the CP model meant that traditional point-to-point verification methods were not appropriate because small-scale errors can be heavily penalized. If the forecast position of the storm is offset from the true position of the storm, point-to-point verification applies a “double penalty”: where the storm was forecast will record false alarms, and where the storm did occur will record misses, resulting in a low skill score (Roberts and Lean 2008; Mittermaier 2012). Forecasters using the model should be aware of this and use judgement to forecast storms over a region, rather than at a specific location.

For this reason, the FSS, developed by Roberts and Lean (2008), is used in this study. This metric aims to measure the variation in skill with spatial scale and hence the smallest scale at which a model has skill. The verification process is as follows: first, a rainfall rate threshold is chosen. At each grid box, the fraction of points that exceeds this threshold within a surrounding  $n \times n$  gridpoint “neighborhood” is recorded for both the model and observations for increasing values of  $n$ . The square of the difference between the fraction in the model  $M_{(n)}$  and the fraction in the observations  $O_{(n)}$  is averaged over all grid points at a given time to compute the mean squared error (MSE):

$$\text{MSE}_{(n)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [O_{(n)ij} - M_{(n)ij}]^2, \quad (1)$$

where  $N_x$  and  $N_y$  are the number of longitude and latitude points, respectively. Since the MSE has a high dependence on the frequency of the event, it is compared

against the MSE of a reference forecast with low-skill  $\text{MSE}_{(n)\text{ref}}$ , defined in Murphy and Epstein (1989) as

$$\text{MSE}_{(n)\text{ref}} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [O_{(n)ij}^2 + M_{(n)ij}^2], \quad (2)$$

to obtain the fractions skill score:

$$\text{FSS}_{(n)} = 1 - \frac{\text{MSE}_{(n)}}{\text{MSE}_{(n)\text{ref}}}. \quad (3)$$

A score of 1 implies that the forecast has a perfect match for the neighborhood size used, whereas a score of 0 implies that the forecast has no skill. The skill score that would be achieved, on average, by a random forecast with the same fraction of events  $f_0$  over the domain as the observations is given by  $f_0$ . A “target” or uniform forecast is defined as a forecast in which the model fraction  $M_{(n)}$  at each grid point is equal to  $f_0$ . The score for such a forecast is given by  $0.5 + (f_0/2)$  and can be approximated to 0.5 for comparisons when the frequencies are small (Roberts and Lean 2008; Roberts 2008). In fact, if the FSS is being used as a measure of spatial displacement, it is the value of 0.5 that should be used (Skok 2015; Skok and Roberts 2016). The spatial scale at which the model reaches this skill can be interpreted as the spatial scale on which the model has skill. More information regarding the details of the FSS calculation is given in the appendix.

The traditional FSS generates a skill score at each time step to describe the whole domain and allow variation in the model performance with time to be studied. However, FSS cannot be used to understand how the skill of the model varies across the spatial domain. An adapted version of the FSS, termed the localized fractions skill score (LFSS), was used to do this. The mean squared error between the fraction of grid points exceeding the threshold in the model and observations is found by taking the mean over time for a given neighborhood size, rather than over the spatial domain. In this version of the score, Eqs. (1) and (2) are therefore replaced by

$$\text{MSE}_{(n)} = \frac{1}{N_t} \sum_{k=1}^{N_t} [O_{(n)k} - M_{(n)k}]^2, \quad (4)$$

and

$$\text{MSE}_{(n)\text{ref}} = \frac{1}{N_t} \sum_{k=1}^{N_t} [O_{(n)k}^2 + M_{(n)k}^2], \quad (5)$$

respectively, where  $N_t$  is the number of time steps. Rather than a skill score per time, a skill score is instead obtained per grid point over a period of time for the

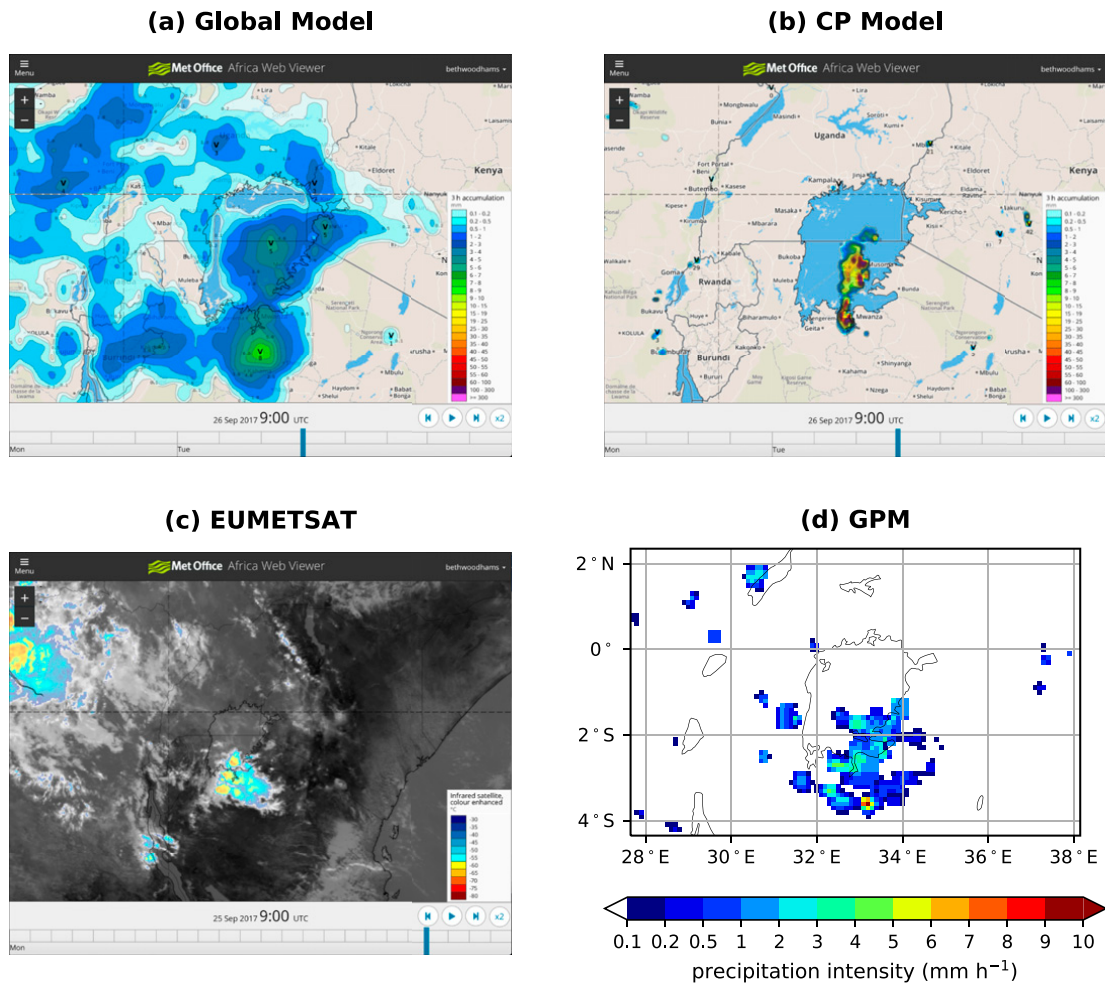


FIG. 2. The  $T + 33$  h 3-h precipitation accumulation forecasts from (a) the global model and (b) the CP model, and (c) IR observations from EUMETSAT for 0900 UTC 26 Sep 2017; screenshots taken from the Met Office VCP Africa Web Viewer. (d) Instantaneous precipitation intensity at 0900 UTC from GPM.

neighborhood size of interest. Because the neighborhood size is still in the spatial domain, the LFSS cannot have a “target” score to find the displacement of storms in time, which would require the size of the neighborhood to be in the time dimension. The LFSS is therefore still used to understand the spatial skill. While the FSS should be used to quantitatively find a neighborhood size on which the model has skill, the LFSS should be used more qualitatively to find regions that have greater or lesser skill relative to the whole domain.

### 3. Results

#### a. Forecast example

Figure 2 shows an example of a 3-h rainfall accumulation forecast over Lake Victoria from (Fig. 2a) the global model and (Fig. 2b) the CP model, as would be seen by a forecaster using the Met Office VCP Africa Web Viewer. This example is from 0900 UTC 26 September 2017

(1200 LT; LT = UTC + 3 h). This case was chosen as an exemplar case of a good forecast by the CP model, which exhibits many of the characteristics typical of the model performance for a significant storm event over the lake. Both models correctly predicted an event over the lake, with the CP model predicting a more realistic storm structure. Although the global model forecast did show some structure in the precipitation field, it did not produce an organized storm. Rather, it forecast rain over most of the lake. The CP model forecast a linear storm structure, with a region of heavy precipitation over the south of the lake and another region over land on the southern shore. IR observations available on the Africa Web Viewer in Fig. 2c show that the regions of rainfall forecast by the CP model coincided well with areas of low IR. Precipitation rates from GPM (Fig. 2d) show that the heaviest region of precipitation on the southern shore collocated fairly well with the area of highest accumulations in the CP model forecast, although it was slightly

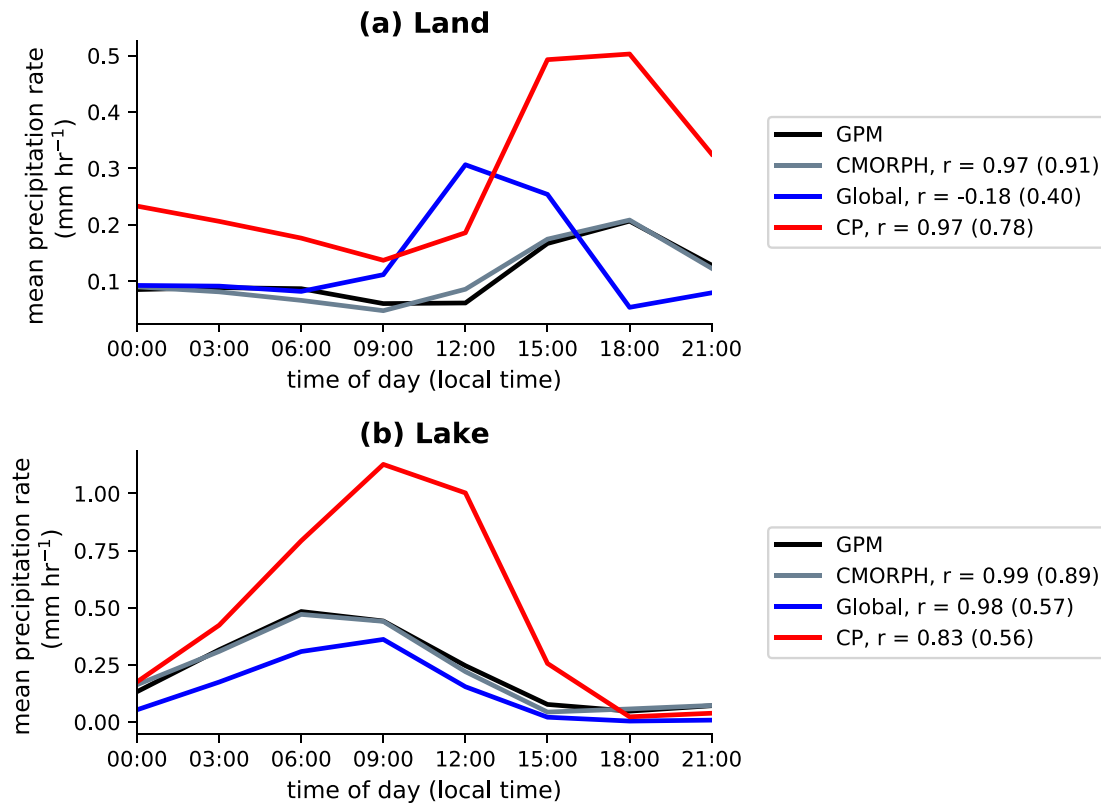


FIG. 3. Mean diurnal cycle of precipitation rates over (a) the land grid points within the LV subdomain (dashed black line in Fig. 1) and (b) the lake grid points of LV for GPM, CMORPH, and the global and CP models. The Pearson's  $r$  value for the correlation of the mean diurnal cycle with that of GPM is given in the legend. The correlation over the full 2-yr time series is also given in brackets. The four model time series (comprising different initialization and lead times as described in section 2a) are used for the global and CP model data, so the results reflect an average across these datasets. This is true in all figures unless stated otherwise.

farther south. The observed rainfall rates suggest that the forecast accumulation was too high and also too localized.

### b. Diurnal cycle

The mean diurnal cycles over (Fig. 3a) the land grid points within the LV subdomain and (Fig. 3b) Lake Victoria (lake grid points only) are shown for GPM observations (and CMORPH for comparison) and for the global and CP models. Over land, the observations show that the convective maximum occurs at 1800 LT. During the day, the land surrounding the lake heats up faster than the lake itself, causing divergence and hence subsidence over the lake (Chamberlain et al. 2014). Onshore winds trigger storms over the mountains to the east of the lake, which reach peak intensity late in the afternoon. However, the global model shows a mean rainfall maximum at 1200 LT, which is larger than the peak in the observations by a factor of 1.5. There is actually a weak negative correlation between the mean diurnal cycle in the model and observations (Pearson's  $r$  value of  $-0.18$ ). Comparing the full 2-yr time series from the global model to that of the observations, the

correlation coefficient is 0.40, showing an improved relationship compared to the mean. The erroneous midday maximum is in agreement with previous studies of global models (Yang and Slingo 2001; Bechtold et al. 2004). The peak in rainfall in the CP model occurs at 1800 LT, in agreement with the observations, although the mean rainfall rate at 1500 LT is nearly as great. The CP model overpredicts the maximum rainfall rate by a factor of 2.4, consistent with known biases in CP MetUM (Lean et al. 2008; Kendon et al. 2012), but improved in recent research configurations (Aranami et al. 2015; Zerroukat and Shipway 2017). Despite this, correlation coefficients of 0.97 and 0.78 are achieved by the CP model for the mean diurnal cycle and the full 2-yr time series, respectively. CMORPH exhibits a similar diurnal cycle to GPM over land, with a correlation coefficient of 0.97 for the mean diurnal cycle and 0.91 across the whole time series.

At night, the lake remains warmer than the land, such that convergence and, hence, convection occurs over the lake, and the observed precipitation peaks at 0600 LT (Song et al. 2004). The maximum rainfall occurs later, at



0900 LT, in the global and CP models. The global model slightly underpredicts the maximum mean rainfall rate over the lake, whereas the rate in the CP model is a factor of 2.3 greater than GPM. The correlation between the mean diurnal cycle and the models is greater for the global model (0.98) than the CP model (0.83), although both show strong correlations. The correlation coefficients between both models and the observations over the full time series are very similar: 0.57 and 0.56 for the global and CP model, respectively. The ability of the global model to capture the correct diurnal cycle over the lake, when it is unable to do so over land, suggests that the forcings that lead to storms over the lake are particularly strong and on a scale greater than the grid spacing of the global model. The consistently warm temperature of the lake relative to the land may allow storms to be triggered in the parameterization scheme, without the need for solar insolation, while the positioning of the rain is constrained by the location of the lake. CMORPH correlates very well with GPM, with a correlation coefficient of 0.99 for the mean diurnal cycle and 0.89 for the whole time series over the lake.

### c. Precipitation rates

The precipitation rates corresponding to different percentiles are shown in Fig. 4 for the observational data (GPM) and the two models over the full domain (solid line) and over the LV subdomain (dashed line). Precipitation rates for the additional observations from CMORPH are also shown. The corresponding rainfall rates were found using the data from all times and all grid points within the analysis period and corresponding domain. Most of the rainfall intensities within this time were  $0 \text{ mm h}^{-1}$ . This is reflected in Figs. 4a and 4b, which show the rainfall rates as a function of percentile, computed including times of no rain. Figure 4a demonstrates that while the models do predict a large proportion of dry events, they predict too many rainfall events, compared to observations. While 90.9% of all data points in the full domain have no rain in the observations, this is true of 75.8% and 81.1% of data points in the global and CP models, respectively. Similar results hold over the LV subdomain. Figure 4b shows that above the 99.9th percentile, the rain rates in the global model are 15% and 20% lower than the observations over the full domain and LV subdomain, respectively. However, the CP model produces far too much heavy rain, over a factor of 3.5 greater than the observations for the 99.9th percentile, and increasing with greater percentiles. However, studies such as Wang et al. (2017a) and O et al. (2017) suggest that GPM may underestimate high rainfall events, so the difference between the CP model and true rainfall amount may not be so large. Overall, there

is little difference between the full domain and LV subdomain.

After removing data points with no rain, Figs. 4c and 4d show how the rainfall rates are distributed. Figure 4c shows that of the rain that was produced by both models (and over both domains), the rain was too light for the majority of data points. Figure 4d shows that the global model is unable to predict the highest-intensity events, while the CP model has a disproportionate amount of unrealistically extreme events, with the precipitation rate around a factor of 3 too large above the 99.9th percentile.

The contribution of rain of different intensities to the mean rainfall intensity is shown in Fig. 4e for GPM and the two models. Again, this shows that the global model produces light precipitation too often, compared to GPM. For the very highest intensities, the global model matches the observations well. Despite being predicted relatively too often, the lightest-intensity precipitation contributes to the mean rainfall too infrequently in the CP model, because rain rates above  $5 \text{ mm h}^{-1}$  are produced far too frequently.

Comparing the two observational datasets, Fig. 4a shows that CMORPH is drier than GPM. GPM has higher extreme rainfall rates over the whole domain, but CMORPH has the highest rain rates over the LV subdomain (Figs. 4b,d). This difference could be related to the performance of the two observational datasets over the complex topography within the subdomain. Low rainfall rates contribute to the mean rainfall less in CMORPH than in GPM, but intermediate rainfall rates (between  $6$  and  $19 \text{ mm h}^{-1}$ ) contribute more (Fig. 4e). Overall, the differences between the characteristics of the observational datasets are smaller than the differences between the models and observations. Therefore, only results using GPM observations are presented in the remainder of the paper. Much of the following work was reproduced using CMORPH and yielded very similar results.

### d. Lake Victoria objective analysis

Because of the dangers faced by fishermen when a storm occurs over Lake Victoria, the ability of the models to forecast storms over the lake itself is extremely important. Objective analysis was performed to compute the hit rate, false alarm ratio, and ROC AUC for a range of minimum storm sizes and intensities over the lake. The intensity threshold was defined by a percentile threshold (calculated using only lake grid points) to avoid issues with the different distributions of rain rates between the models and observations.

Figures 5a and 5b show the hit rates achieved by the global and CP models, respectively, as a function of minimum size and intensity, and Fig. 5c shows the difference in

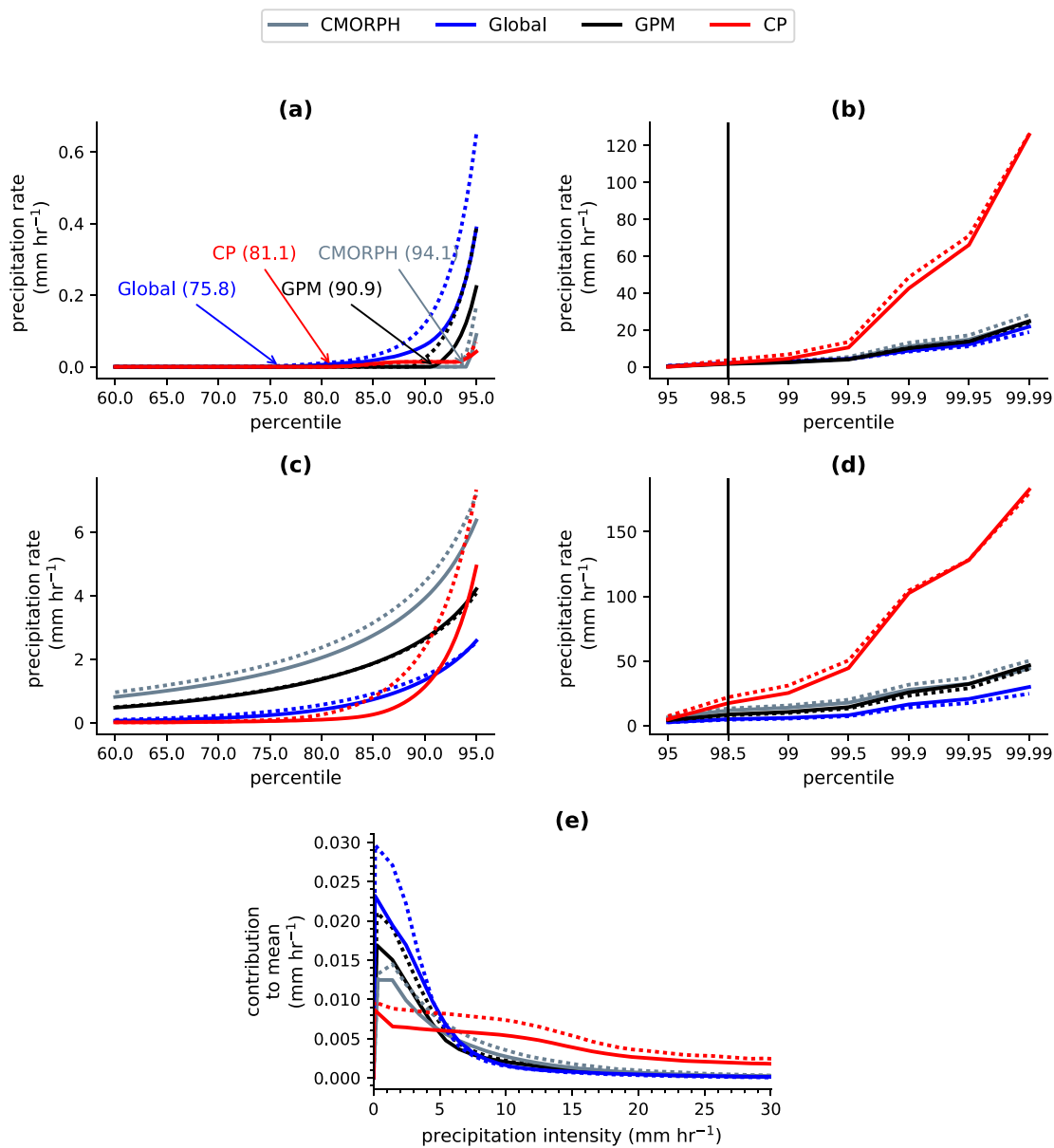


FIG. 4. (a)–(d) Precipitation rates corresponding to different percentile thresholds for GPM, CMORPH, and the global and CP models. (a),(b) The corresponding precipitation rates including points with precipitation rates of 0 mm h<sup>-1</sup>; (c),(d) neglect these points. (e) Contribution to the mean rainfall intensity, computed by binning rainfall rates into 1 mm h<sup>-1</sup> bins and multiplying the average rainfall rate within the bin by the probability of the rainfall rate lying within that bin. The area under the curve is equal to the mean rainfall. For all panels, the solid lines show results for the full domain, and the dashed lines show results for the LV subdomain.

skill between the two models. For all three metrics, results are only plotted when a storm of the specified size and intensity was detected in the observations for at least 2.5% of the time steps (corresponding to over 500 storm events) in order to have a large sample size for more robust statistics. The sizes and intensities where this criterion was not met are hatched. The hit rate for both the global and CP models decreases as the minimum size and intensity of the storm increases. Overall, the CP model has a higher hit rate than the global model, and for the lowest intensities, it

predicts up to 20% more storms correctly. For a minimum storm size of 20 grid points (approximately 12 500 km<sup>2</sup>, 1/5 of the size of Lake Victoria) and rainfall above the 92nd percentile, the CP model has a hit rate of 57%, compared to 33% in the global model.

Similar plots of the false alarm ratios are shown in Figs. 5d–f. For both models, the false alarm ratio increases with increasing size and intensity. The global model has a smaller false alarm ratio for the majority of storm sizes and intensities, especially for small, low-intensity storms.

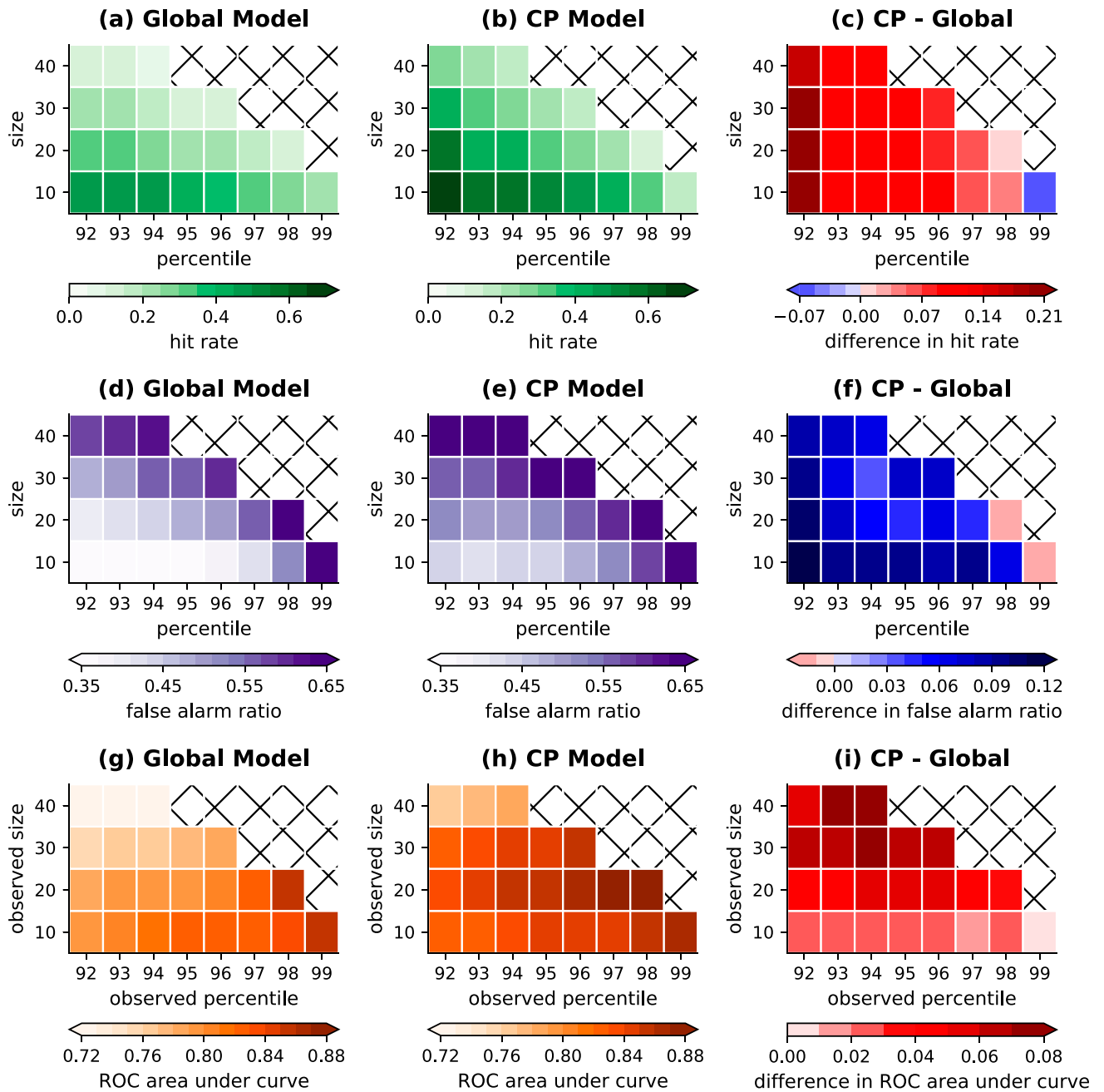


FIG. 5. (a)–(c) Hit rate, (d)–(f) false alarm ratio, and (g)–(i) ROC AUC for storms of a minimum intensity over a minimum size (number of grid points) over LV (lake grid points only) on 3-h intervals. The first column shows results for the global model, the second column for the CP model, and the third column is the difference between the CP and global models. In the difference plots, red shows where the CP model has increased skill relative to the global model. The hatched areas show where a storm of the given size or intensity was observed over less than 2.5% of the time steps.

For storms with a minimum size of 20 grid points and rainfall above the 92nd percentile, the global and CP models have false alarm rates of 51% and 41%, respectively.

The ROC AUC of both models generally decreases as the size of the observed storm increases (Figs. 5g,h). However, the AUC generally increases as the intensity of the observed storm increases. Figure 5i shows that for

all intensities and sizes, the CP model scores more highly than the global model, especially for large storms. The ROC AUC for storms with a minimum size of 20 grid points and rainfall above the 92nd percentile is 0.28 for the CP model and 0.22 for the global model. The greater skill of the CP model suggests that its greater hit rate outweighs the increased fraction of false alarms.

### e. FSS for 24-h accumulation rates

To verify the model over a long time period, the FSS was used. The rainfall threshold was chosen to be the 98.5th percentile for reasons discussed in the [appendix](#). Over the full domain (LV subdomain), the 98.5th percentile corresponded to 30.2 (32.0) mm day<sup>-1</sup> in the observations and, taking an average of the four combinations of forecast initializations and forecast lead time periods, 29.0 (27.6) mm day<sup>-1</sup> in the global model and 96.2 (109.3) mm day<sup>-1</sup> in the CP model.

Figures 6a and 6b show the mean FSS as a function of spatial scale for 24-h accumulation forecasts from the different models and a 24-h persistence forecast for the full domain and the LV subdomain, respectively. All three forecasts perform better than a random forecast over both domains, although not by much at the grid scale. Both models improve upon the persistence forecast at all spatial scales. Over the full domain, the two models show very similar skill, only reaching the target skill for spatial scales greater than almost 400 km. Over the LV subdomain, the CP model has increased skill, compared to the global model, reaching the target skill at a spatial scale of around 350 km, compared to just over 375 km for the global model.

#### 1) LOCALIZED FSS

Maps of LFSS, computed for 24-h accumulations at a spatial scale of 425 km ( $n = 17$ , because this is the nearest neighborhood size at which the models first become skillful over the whole domain, according to [Fig. 6](#)) and for a threshold of the 98.5th percentile, are shown in [Figs. 7a and 7b](#) for the global and CP models, respectively. The resulting spatial patterns are consistent across all spatial scales and for all percentile thresholds greater than the 95th percentile, although the magnitude of the score does change. Both models show enhanced skill over Lake Victoria and many of the regions of high orography, especially the Ethiopian highlands and the mountain ridges on either side of Lake Victoria. The global model shows particularly poor performance along the Somalian coastline, compared to the CP model. The green contour encloses regions where the mean rainfall is in the top 25% in both the observations and models. Although some of the regions of enhanced LFSS coincide with overlapping regions of the heaviest rainfall, this is not exclusive, suggesting that the model does not only perform well where rainfall is heaviest and most common. As shown in [Fig. 7c](#), the CP model generally shows greater skill over the land and lake than the global model, but the global model shows higher skill over the ocean. Since only a small proportion of the sea is included in the LV subdomain, this explains why the skill of the global

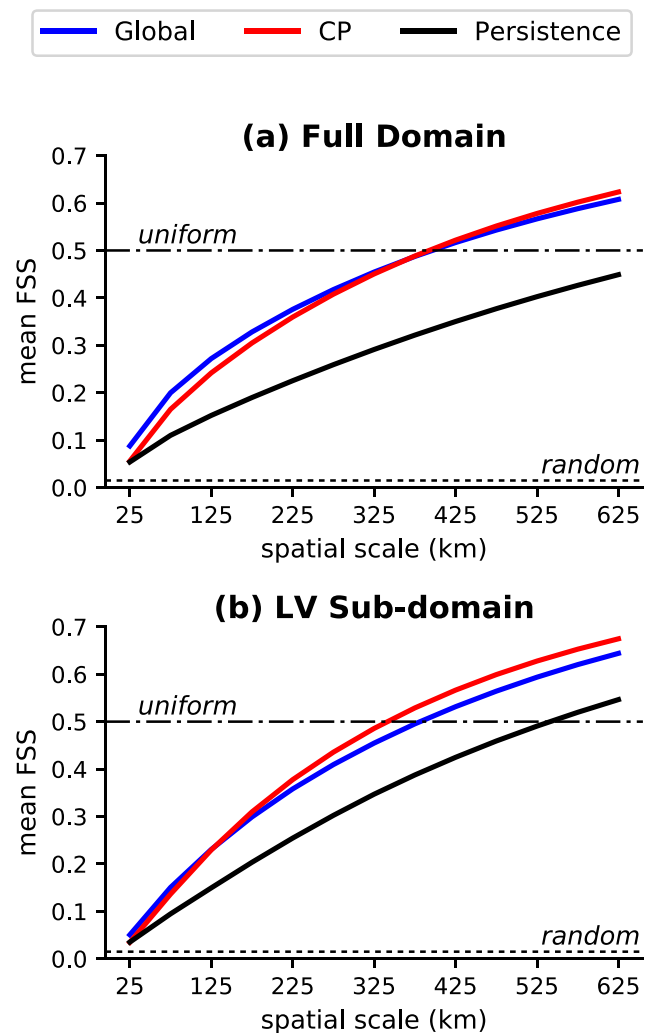


FIG. 6. Mean FSS as a function of neighborhood size for 24-h rainfall accumulations above the 98.5th percentile, shown for the global and CP models and a 24-h persistence forecast over (a) the full domain and (b) the LV subdomain.

model relative to the CP model drops over the LV subdomain ([Fig. 6](#)). Over land and on Lake Victoria are the most important places to forecast correctly, since these are where people live and work. The ubiquitously high scores over the ocean in the global model suggest that the LV subdomain should be used for further analysis to avoid contamination by ocean grid points. As such, for the remainder of the paper, only FSSs over the LV subdomain are presented.

#### 2) LFSS SEASONAL VARIABILITY

Figures 8 and 9 show how the LFSS, computed separately over the different seasons, differs from the LFSS computed over the whole time period for the global and CP models, respectively. Although quite noisy, the results are fairly similar for both models. This is likely because the annual cycle of convection is controlled by

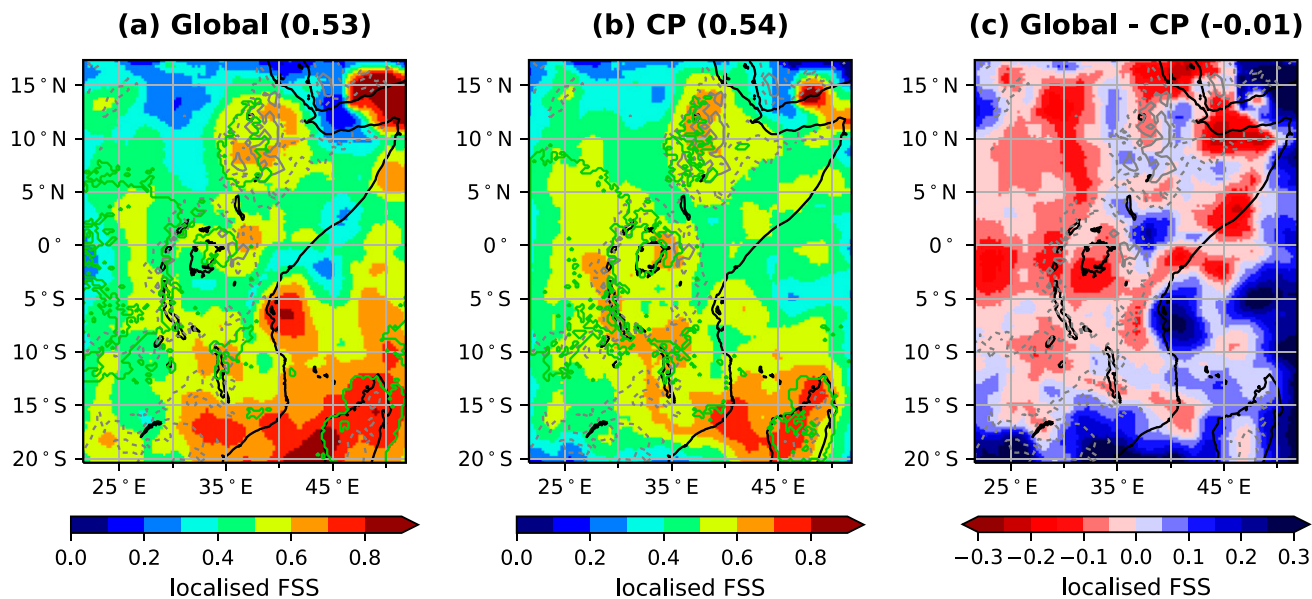


FIG. 7. Map of LFSS, computed for (a) the global and (b) the CP model, using a threshold of the 98.5th percentile and a neighborhood size of 425 km ( $n = 17$ ) for 24-h accumulations. (c) Map of the difference between the LFSS in the global model and CP model. The solid green contour encloses regions with mean rainfall in the top 25% in both the observations and the model. The gray dashed and solid contours mark orography of height 1 and 2 km, respectively. The numbers in brackets give the area average within the LV subdomain.

the intertropical convergence zone (ITCZ), the location of which should be similar in both models, since the global model fields are used to initiate the CP model.

During the rainy seasons (MAM and OND), when the ITCZ is located over equatorial East Africa, the variability in skill in both models is generally small around the Lake Victoria basin, with areas of both moderately increased and decreased skill. The skill over the ocean is decreased during the short rains (OND), but increased during the long rains (MAM). The dry seasons in the equatorial region correspond to when the ITCZ and main band of rainfall are to the north of the region in June–September (JJAS) and to the south in January–February (JF). For both models, there is broadly a large negative perturbation in skill over the opposite region of the domain to where the ITCZ sits. However, within the large areas of reduced skill, there are small areas of highly increased skill. Some of these are located in regions of high orography or over some of the small lakes or parts of the coastline. Within a contour enclosing the regions for which the top 25% of mean rainfall overlap in the model and observations, the mean LFSS perturbation is close to zero for all seasons and both models, showing that the model does not just perform well where rainfall is heaviest or most common.

*f. FSS for 3-hourly precipitation rates*

Figure 10 shows how the FSS for the global and CP models and a 24-h persistence forecast varies with spatial scale for the forecast rainfall rate at 3-h intervals

throughout the day. Similar to the analysis on 24-h accumulations, a threshold of the 98.5th percentile was chosen, corresponding to  $2.1 \text{ mm h}^{-1}$  in the observations and (taking an average of the four combinations of forecast initializations and forecast lead time periods)  $2.3 \text{ mm h}^{-1}$  in the global model and  $3.8 \text{ mm h}^{-1}$  in the CP model.

All three forecasts consistently beat the random forecast skill. The CP model outperforms the global model and persistence forecast at almost all spatial scales and at all times of day. The global model generally has greater skill than the persistence forecast, except at 1800 LT. Before 1200 LT, the target “uniform” skill is generally not achieved by any of the forecasts within spatial scales below 425 km. In particular, the skill of the CP model increases after 1200 LT and is very high at 1500 and 1800 LT, reaching the target skill at a spatial scale of approximately 275 km. These times of day correspond to the convective maximum over land.

Figure 11 shows how the skill of (Fig. 11a) the global model and (Fig. 11b) the CP model at different times of day varies with forecast lead time, using the 98.5th percentile as a threshold. Plots are only shown for a neighborhood size of 475 km ( $n = 19$ , the spatial scale at which the models have reached the target score for almost all times of day), but trends seen are broadly consistent for all spatial scales and percentiles greater than the 95th percentile. Each line corresponds to a different time of day, being formed of three points, since each time of day is forecast from three different initializations (the

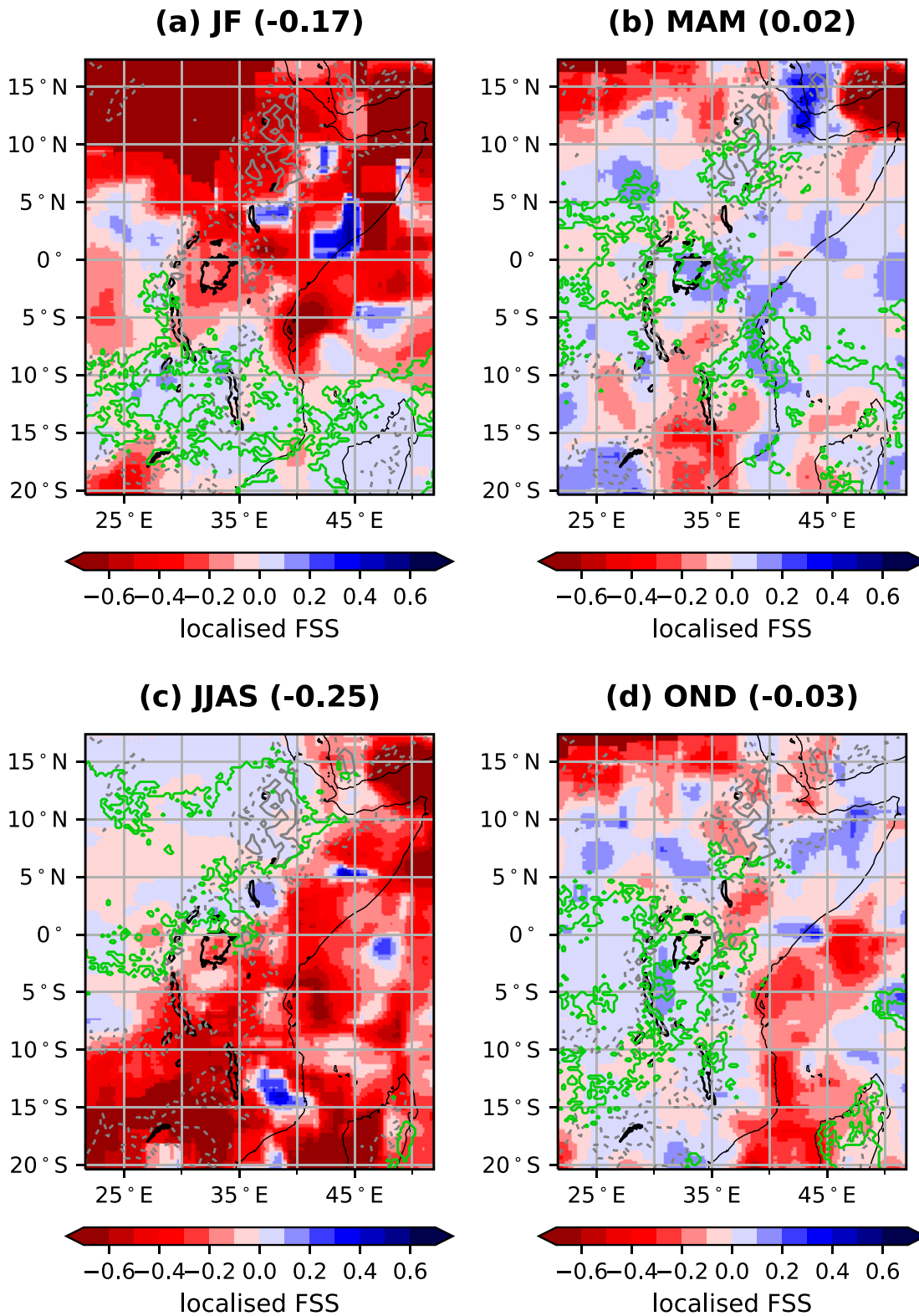


FIG. 8. Maps of perturbation in LFSS from Fig. 7a for different seasons, computed for the global model, using a threshold of the 98.5th percentile and a neighborhood size of 425 km ( $n = 17$ ) for 24-h accumulations. The solid green contour encloses regions with mean rainfall in the top 25% in both the observations and the model for the given season. The gray dashed and solid contours mark orography of height 1 and 2 km, respectively. The numbers in brackets give the area average within the LV subdomain.

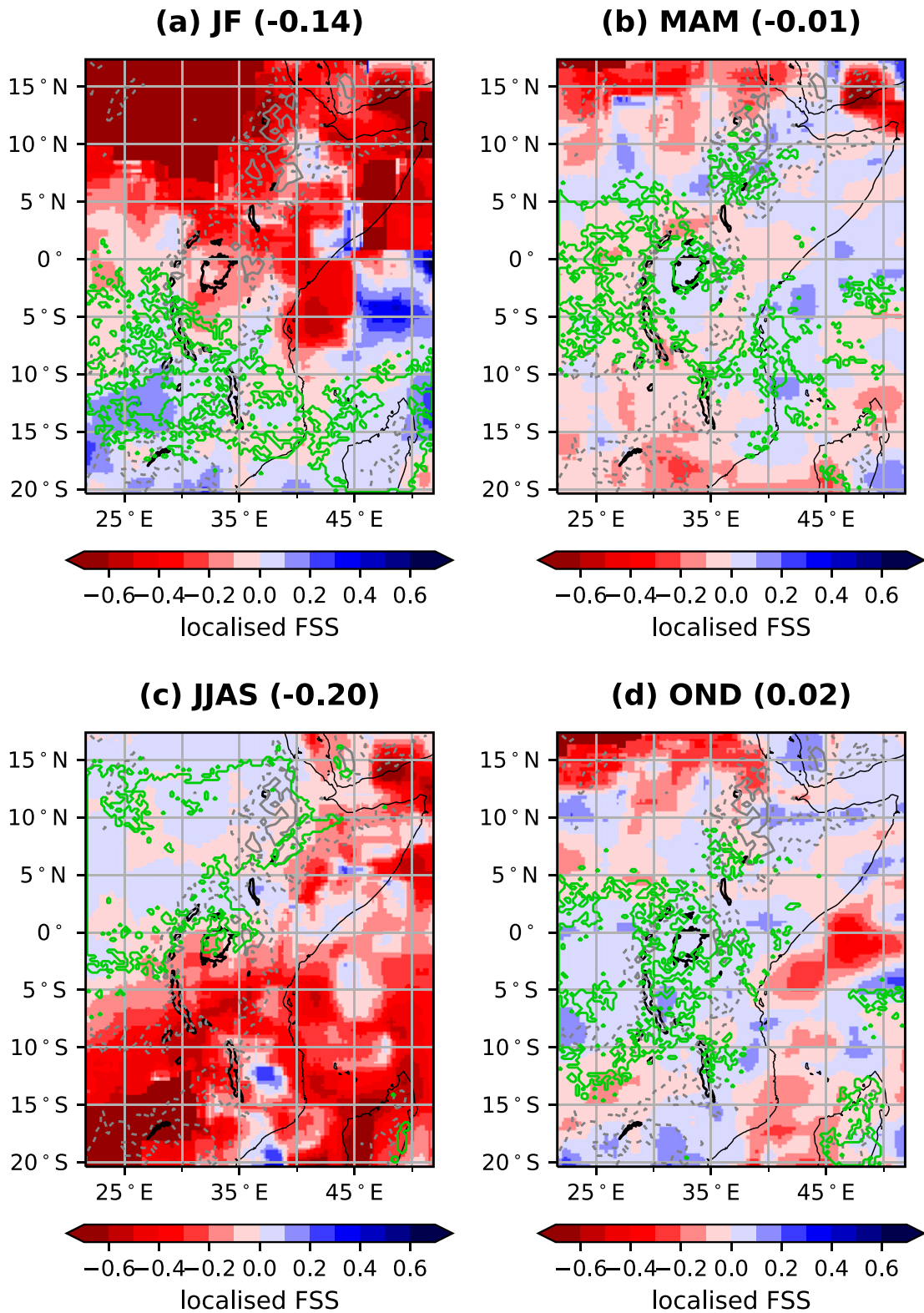


FIG. 9. As in Fig. 8, but computed for the CP model and therefore showing the perturbation in LFSS from Fig. 7b.

first 12 h of forecast are discarded due to spin up). Although there is small degradation in the skill of both models as forecast lead time increases, this is almost negligible, compared to the variation in skill between the

different times of day, especially in the CP model. These results contrast with those of Roberts (2008) for U.K. precipitation, who showed no dependence of the score on time of day. This is because rainfall in the midlatitudes is

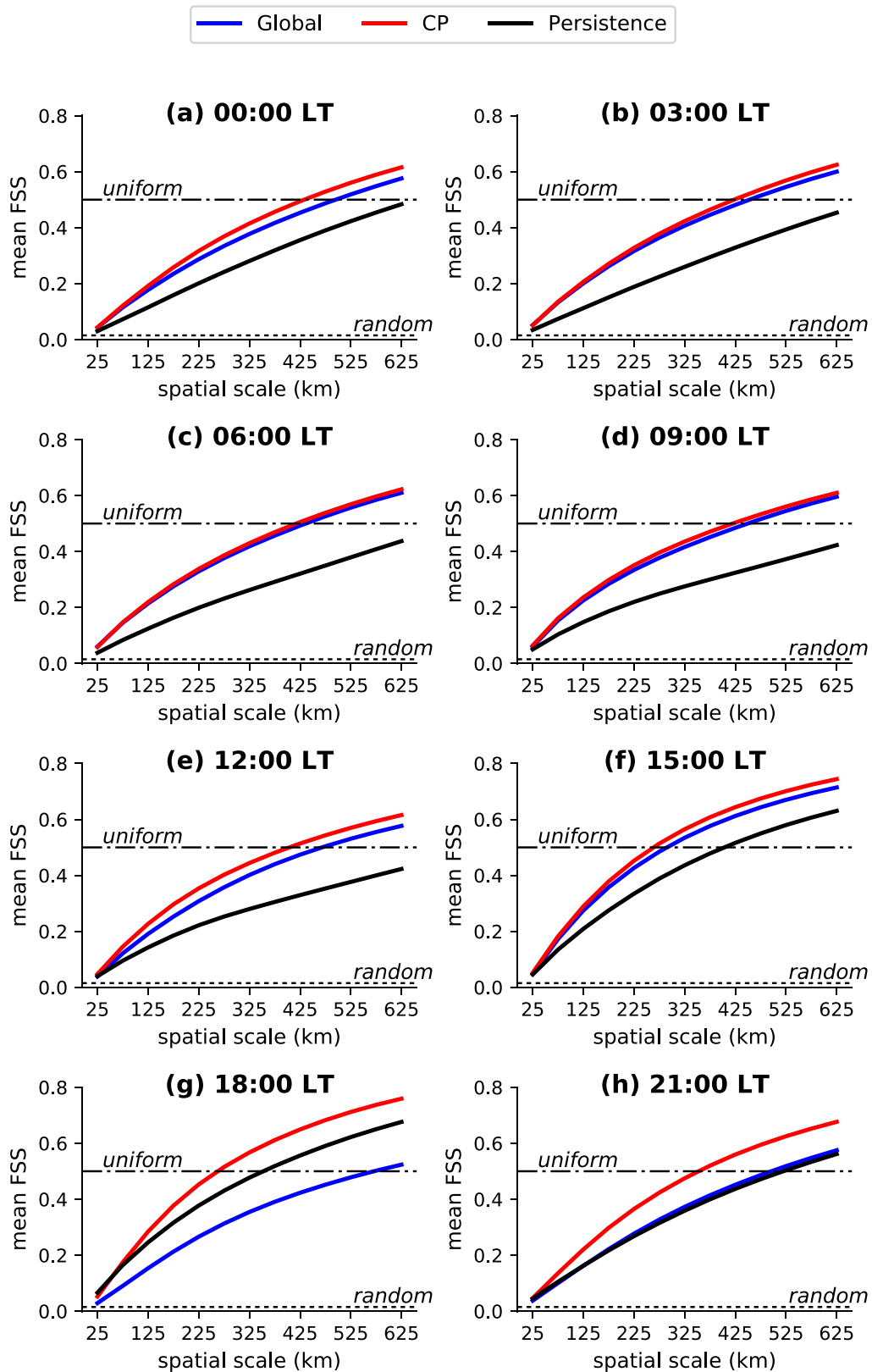


FIG. 10. Mean FSS over the LV subdomain as a function of neighborhood size for rainfall above the 98.5th percentile, computed at 3-h intervals (corresponding to the model diagnostic output times) for the global and CP models and a 24-h persistence forecast.



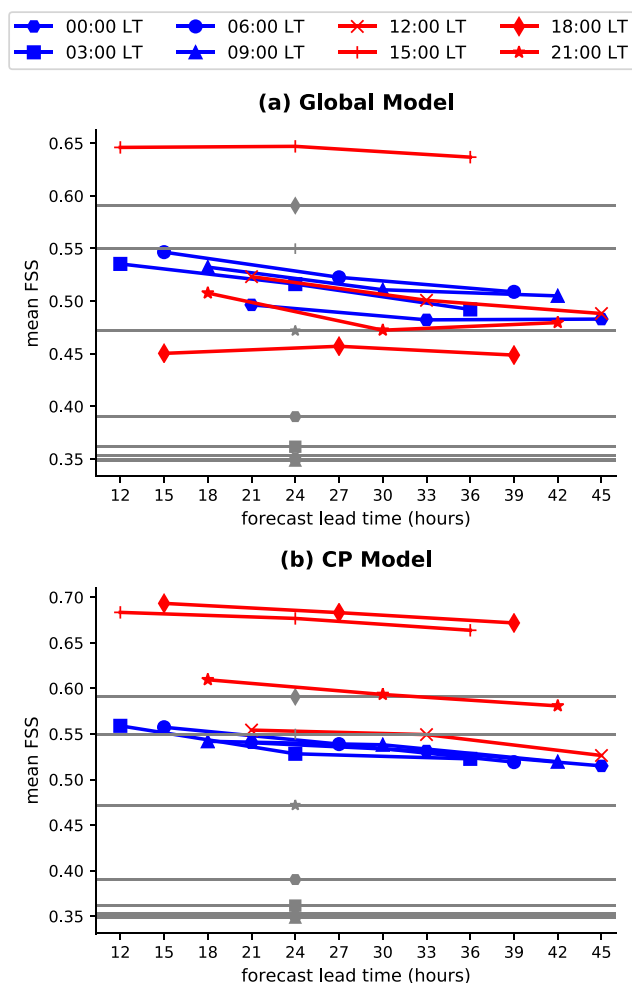


FIG. 11. Mean FSS as a function of forecast lead time, using the 98.5th percentile for (a) the global model and (b) the CP model over the LV subdomain. A neighborhood size of 475 km ( $n = 19$ ) is used, but results are consistent for all spatial scales. Each line corresponds to a different model diagnostic output time and is formed of three points (e.g., for 0000 LT, the first and last points are the  $T + 21$  and  $T + 45$  forecasts from 0000 UTC initializations and the middle point is the  $T + 33$  forecast from 1200 UTC initializations). The gray lines show the skill of a 24-h persistence forecast.

generally controlled by frontal systems, whereas convection in the tropics is strongly forced by the diurnal cycle.

Figure 11a shows that the skill of the global model is higher at 1500 LT by over 1/4, compared to other times of day. The CP model has increased skill at 1500 and 1800 LT by around 1/4 relative to most other times of day (Fig. 11b), and scores at 2100 LT are also elevated. These times correspond to the period when convection is at a maximum over land.

Figure 12a shows the probability distribution of the 98.5th percentile of precipitation rates in the GPM observations obtained from all time steps. Except between 1500 and 2100 LT, there is a high likelihood that the 98.5th percentile of the rainfall intensity is small.

Figures 12b and 12c show the mean FSS associated with a given 98.5th percentile of rainfall intensity for the global and CP models, respectively, for different times of day. Except at 1800 LT in the global model, the smallest intensities score the lowest FSS at all times of day in both models. The high skill for very low rainfall at 1800 LT in the global model is likely because the global model always forecasts such little rain at this time, so any time steps that occasionally do have little rainfall at 1800 LT will automatically score highly.

One reason for the increased skill in the global model at 1500 LT is that rainfall rates between 1 and 3.5 mm h<sup>-1</sup> are likely to have a greater FSS than at other times of day (Fig. 12b). Furthermore, the likelihood of the lowest rainfall rates (which generally have reduced scores) is decreased (Fig. 12a). Since this time lies between the midday convective maximum in the global model and the observed 1800 LT maximum, this increased skill is likely an artifact of storms over land nearing the end of their lives in the model coincident with growing storms in the observations.

Increased skill in the CP model between 1500 and 1800 LT occurs because the scores associated with intermediate rainfall rates have an increased mean FSS, compared to other times of day (Fig. 12c). In addition, the likelihood of the 98.5th percentile being below 0.5 mm h<sup>-1</sup> is greatly reduced. This is also true at 2100 LT, which explains the increased skill at this time of day. It is speculated that increased skill in the CP model coincident with the convective maximum over land is due to the model's ability to more accurately predict the location of storms at this time of day, not simply because there is more heavy rain and therefore a reduced number of low-scoring, low-intensity events.

#### 4. Discussion and conclusions

The precipitation forecast produced by a CP model, which was run operationally for East Africa by the Met Office in 2017, was verified over a 2-yr period from July 2014 to July 2016 using observations from GPM. Its performance was also compared to that of the operational global MetUM in order to understand the value added by a CP forecast. Verification was performed for the prediction of 24-h accumulations, as well as for rainfall rates at 3-hourly time intervals. Much of this verification used the fractions skill score, including a novel form of the skill score, termed the localized fractions skill score, to look at spatial variations in performance. Although extended assessments of CP models have been performed in the midlatitudes (Mittermaier et al. 2013), such an assessment has not been performed in the tropics before.

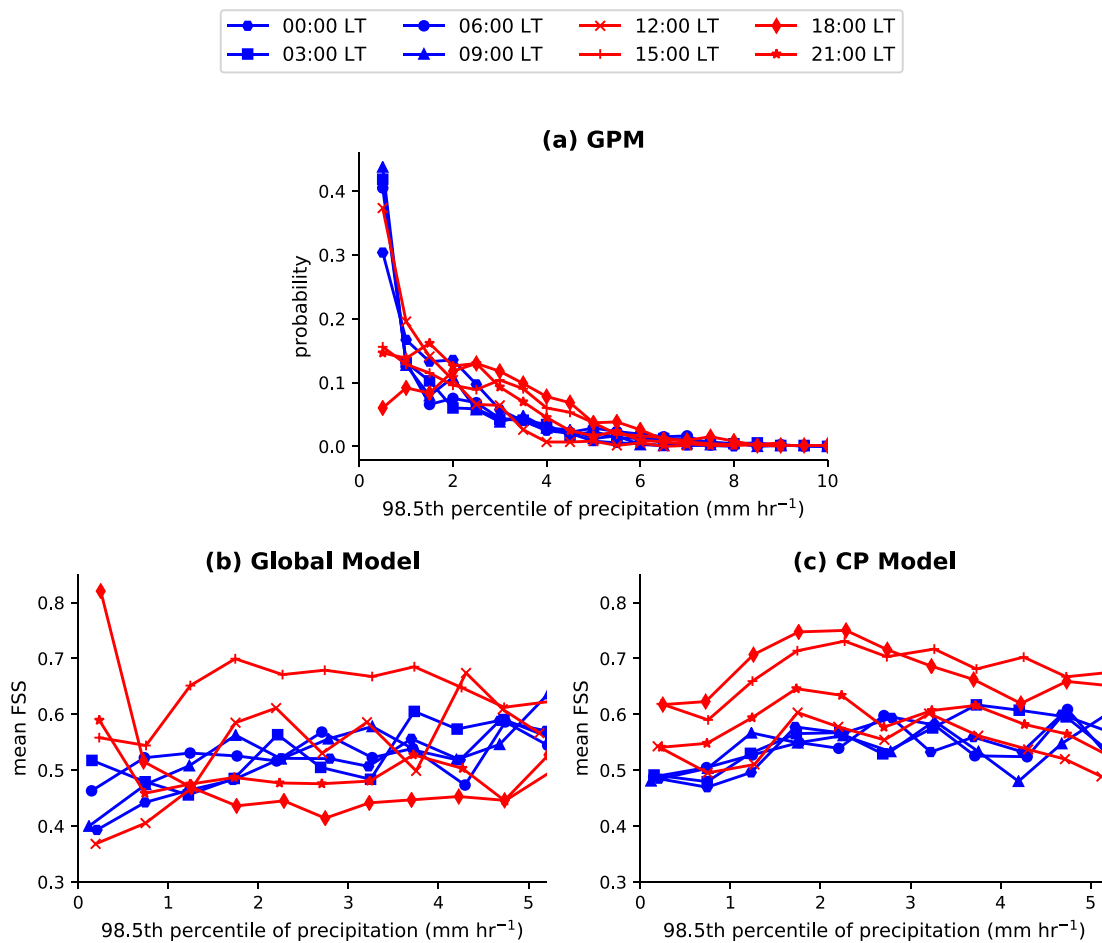


FIG. 12. (a) A probability distribution for the precipitation rate at the 98.5th percentile according to GPM observations. Mean FSS as a function of the 98.5th percentile of observed precipitation intensity across the LV subdomain for 3-h intervals (corresponding to the model diagnostic output times) for (b) the global model and (c) the CP model. The intensity is only plotted within an intensity band where there is a high-enough sample of data. A neighborhood size of 475 km ( $n = 19$ ) is used, with consistent results across all spatial scales.

The models show large biases in the distribution of rainfall rates, compared to the observations. In particular, both models are too wet, and the CP model has excessive rain rates. Therefore, percentile thresholds were chosen to define events in both the objective and FSS analysis instead of absolute intensities. Since there is uncertainty associated with the GPM observations, there is uncertainty about the “truth” against which the models are compared, but using a percentile threshold goes some way to reducing this problem.

The generally increased skill in the CP forecast shown over Lake Victoria is encouraging, since the primary reason for the existence of the model is to improve safety on the lake. Objective analysis shows that the CP model is better able to predict whether a storm will occur, compared to the global model. Despite a higher hit rate than the global model, the CP model does produce more false alarms, in agreement with Chamberlain et al.

(2014). Since the livelihoods of fishermen are dependent on their ability to go out on the lake, too many false alarms may lower trust in the forecast and thus reduce adherence to future warnings, even if issued correctly. Higher ROC AUC scores for the CP model suggest that the increased hit rate does outweigh the negative effect of increased false alarms, but the benefit of this does depend on the needs of the lake users.

Using the FSS, both the global and the CP model show greater skill than a 24-h persistence forecast, demonstrating that the models are valuable as forecasting tools. The CP model improves upon the forecast produced by the global model, especially on subdaily time scales; severe storms are more realistic, the diurnal cycle is improved, and this is reflected in generally elevated FSS for the CP model. Presently, forecasters in East Africa tend to use 24-h accumulation values produced by the models in their forecasts (Kenya Meteorological

Department; Tanzania Meteorological Agency; Uganda National Meteorological Authority, 2016, personal communications). Given that the CP model performs better on 3-hourly time intervals, the provision of forecasts with diurnal detail should be increased to provide more guidance to users of the lake as to exactly when it may be hazardous.

The CP model has the greatest benefit in the prediction of storms over land relative to the global model, but the global model performs better over the ocean. Birch et al. (2014a) show that convective initiation in explicit models usually corresponds to low-level convergence, whereas parameterized models tend to be unresponsive to convergence, and convection is generally triggered by instability. Regions of convergence are more likely over land than ocean due to the orography and influence of the lakes. Therefore, the increased scores over land may be due to the increased ability of the CP model to respond to convergence. This hypothesis is also supported by the fact that FSSs for the CP model peak at 1800 LT and are also elevated at 1500 and 2100 LT. These correspond to times of storm initiation over land, suggesting that the model is better able to forecast initiation, but is unable to capture storm propagation. The global model performs poorly along the Somalian coastline, suggesting that the model cannot capture the response of convection to the sea breeze (Birch et al. 2015), but the CP model shows improved performance here.

Despite the superior skill of the models—especially the CP model—relative to a persistence forecast, the level of skill is still fairly poor. Roberts and Lean (2008) suggest that an FSS of at least 0.5 should be the “target,” above which a model may be considered skillful. Figure 6 shows that neither the CP model nor the global model reaches this skill within a neighborhood size of over 350 km for 24-h accumulations. On subdaily time scales, the skillful target is reached by the CP model at a spatial scale of around 275–350 km, between 1500 and 2100 LT, but at other times of day, the target is not reached below approximately 425 km. This shows how heavier rain is easier to forecast, in the CP model in particular, when convection is most widespread. It should be noted that these results are only for the heaviest 1.5% of events, and for lower-percentile thresholds that represent lighter and more extensive rain, the spatial scale at which useful skill is reached is improved. Clark et al. (2016) note that the scales on which CP models have reasonable skill is large, compared to the grid spacing of the model, as well as the horizontal extent of convective rainfall events. Similar results are also shown by Mittermaier et al. (2013) for very high-intensity events in the UK4 model. Results such as these indicate that a CP model should not be interpreted deterministically; rather, an “ensemble” of the model should be run in order to obtain probabilistic forecasts (Clark et al. 2016). However, these are expensive and require greater computing power.

Another reason for such low skill, compared to CP models in the midlatitudes, is the lack of observations available for data assimilation in the tropics. Given the sensitivity of CP models to initial conditions (Birch et al. 2013; Guichard et al. 2010; Melhauser and Zhang 2012; Schumacher et al. 2013; Luo and Chen 2015; Vié et al. 2011), an increase in in situ observations, such as radiosondes and observations of lake surface temperature, has the potential to greatly increase the skill of the model.

Thiery et al. (2015), Argent et al. (2015), and Anyah and Semazzi (2004) showed the importance of accurate LSTs for the prediction of rainfall over Lake Victoria. Since LST observations are only available once per day, and the model contains no interactive lake model, the temperature of the lake will deviate from the truth throughout the duration of the simulation. Thiery et al. (2015) showed that a one-dimensional lake model—in this case, the Freshwater Lake (FLake) model (Mironov 2008; Mironov et al. 2010)—did add value to climate simulations using the COSMO model in climate mode. The FLake model has been used with MetUM (Rooney and Bornemann 2013) before; however, given that the diurnal cycle is fairly well simulated in the current model, the additional computational expense of running the lake model may outweigh the benefits. It would be interesting to see if an interactive lake model could shift the peak in precipitation over the lake slightly earlier to match the observations.

The forecast skill shows a much stronger dependence on the diurnal cycle than the forecast lead time, in contrast with similar studies performed in the midlatitudes. The results suggest that a good use of computer resources could be to initialize the model only once per day, instead of twice, and run out to a longer lead time. This would allow for more advance warnings to be given to users of the lake. However, further tests would need to be run to investigate model performance at lead times greater than 48 h. An immediate solution could be to consider forecasts from the multiple sequential initializations as an ensemble and maintain access to them on the VCP Africa Web Viewer, rather than older forecasts being rapidly replaced by newer ones.

The excessive rainfall rates for the highest-intensity events in the CP model are a cause for concern. This bias exists because semi-Lagrangian advection (as used in MetUM) does not conserve mass. Mass restoration schemes, such as Priestley (1993) and Zerroukat (2010), may easily be applied to global models, but they cannot directly be applied to limited-area models (LAMs) without knowledge of flux through the domain boundaries (Aranami et al. 2015). Aranami et al. (2015) developed a mass restoration scheme for LAMs, which reduces the excessive rainfall rates. This scheme is very

computationally expensive, but was followed by a less computationally intense scheme by [Zerroukat and Shipway \(2017\)](#), which requires no computation of lateral fluxes. This scheme is used in the latest CP MetUM configuration for the tropics, called RA1-T, which also includes further changes, such as tuning of the microphysical parameters to the tropics.

CP models are still fairly novel, and the way in which they must be used and interpreted is extremely different to global models. The uses and limitations of global models are well known, since these models have been used for many years. For a forecaster to feel comfortable using a CP model, time is required to adapt to the different interpretation and to learn its strengths and weaknesses through experience. In essence, a forecaster must learn to “trust” the model. This can be accelerated by more assessment of operational CP models and a greater understanding of when they provide increased skill over a global model forecast.

*Acknowledgments.* This work was supported by the NERC SPHERES DTP (Grant NE/L002574/1) and the HyCRISTAL project (Grant NE/M02038X/1). This work was also supported by U.K. Research and Innovation as part of the Global Challenges Research Fund, Grant NE/P021077/1 (GCRF African SWIFT). The GPM IMERG data were provided by the NASA Goddard Space Flight Center’s Precipitation Measurement Missions Science Team and Precipitation Processing System, which develop and compute the GPM IMERG as a contribution to GPM, including data archiving at the NASA GES DISC. The CMORPH data were provided by the NOAA Climate Prediction Center. The authors thank Humphrey Lean for his useful discussion and comments on the work, Stuart Webster for helpful discussions regarding CP MetUM configurations, and Peter Willets for help in processing CMORPH data. The authors would also like to thank forecasters at the Kenya Meteorological Department, Tanzania Meteorological Agency and Uganda National Meteorological Authority for their helpful discussions and Karen McCourt and Chris Tubbs for facilitating these meetings. The authors are grateful to the anonymous reviewers for improving the clarity and content of the paper.

## APPENDIX

### FSS Computational Details and Choice of Rainfall Threshold

As described in [section 2c\(2\)](#), the FSS relies on a rainfall threshold, above which an “event” is deemed to have occurred. However, [Fig. 4](#) shows that the rainfall

intensities produced by CP models and global models can be vastly different from one another and from observations. For many forecasting applications, the magnitude of the intensity may be bias corrected, and it is more important for the model to forecast whether or not a storm will occur. Therefore, in the computation of the FSS, the precipitation intensity corresponding to a given percentile is used instead of an absolute threshold. Consequently, the models and observations may have different minimum rainfall rates above which an event is defined. For the  $N$ th rainfall percentile, at each time step, the top  $(100 - N)\%$  of grid points (including dry grid points) in the model and observations are considered as “events.” If it is very dry at a time step in either the model or observations, such that there are not  $(100 - N)\%$  of grid points that receive rainfall, these fields are still included in the comparison. This means that the effective percentile threshold is increased for some time steps. The reason for doing this is to include comparisons in which the percentile threshold is just missed (not quite enough rain), but a meaningful comparison can still be made. If time steps with little rain were not included, the sample size would drop considerably because there are many times when the fractional rainfall coverage over the domain is very low.

An issue arises when either the model or observations is much drier than the other, and at least one has less than  $(100 - N)\%$  coverage, such that the ratio of the number of grid points being compared in the model and observations is not equal to one. This leads to a frequency bias, which can become very large when the ratio diverges from one. [Roberts and Lean \(2008\)](#) showed that a frequency bias can greatly reduce the FSS. Therefore, if the frequency bias is more than a factor of 3, the percentile is altered in the wetter dataset to keep the frequency bias within a factor of 3. Again, this means that the effective percentile threshold is increased (spatial coverage is reduced) for some times.

To ensure that only the highest-intensity events are considered, a further constraint is imposed, such that the rainfall must also be above the  $(100 - N)$ th percentile, computed using data from all time steps. This prevents the inclusion of light rainfall at time steps with light rain over large areas.

[Figure A1](#) shows how the mean FSS over the domain varies as a function of percentile threshold. A neighborhood size of 425 km ( $n = 17$ ) is used, but a similar pattern of decline in FSS as percentile threshold increases emerges at all spatial scales. At lower thresholds, the decline is gradual but becomes rapid at around the 95th percentile, because it is more difficult to get the correct positioning of more localized rain. In this study, a threshold of the 98.5th percentile is used to verify the performance of the models in predicting the most

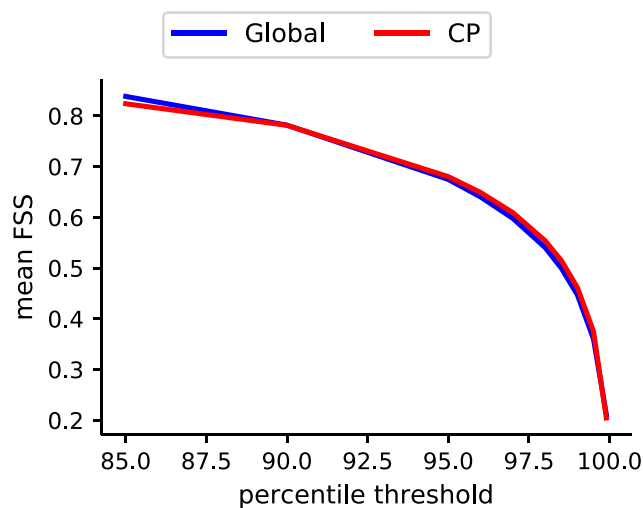


FIG. A1. Mean FSS as a function of percentile threshold for a neighborhood size of 425 km ( $n = 17$ ) for 24-h rainfall accumulations, shown for the global and CP models over the full domain.

extreme events. Although this percentile lies in the region of Fig. A1, in which there is a rapid decline in skill, a high percentile is necessary to focus on the more extreme events, while not so high that a meaningful measure of skill is lost. However, it should be noted that low skill scores are to be expected in comparison to similar studies over the United Kingdom by Roberts and Lean (2008), Roberts (2008), and Mittermaier (2012).

The time steps that do not have coverage of at least 1.5% and, therefore, have a higher effective percentile threshold will be at a disadvantage, given the rapid decline in average score for high-percentile thresholds shown in Fig. A1. To reduce this unfair effect on the FSS, only time steps with at least 0.25% coverage in both datasets are included.

#### REFERENCES

- Anyah, R., and F. Semazzi, 2004: Simulation of the sensitivity of Lake Victoria basin climate to lake surface temperatures. *Theor. Appl. Climatol.*, **79**, 55–69, <https://doi.org/10.1007/s00704-004-0057-4>.
- Aranami, K., T. Davies, and N. Wood, 2015: A mass restoration scheme for limited-area models with semi-Lagrangian advection. *Quart. J. Roy. Meteor. Soc.*, **141**, 1795–1803, <https://doi.org/10.1002/qj.2482>.
- Argent, R., X. Sun, F. Semazzi, L. Xie, and B. Liu, 2015: The development of a customization framework for the WRF Model over the Lake Victoria basin, eastern Africa on seasonal timescales. *Adv. Meteor.*, **2015**, 653473, <https://doi.org/10.1155/2015/653473>.
- Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905, <https://doi.org/10.1175/MWR-D-10-05013.1>.
- Bechtold, P., J.-P. Chaboureaud, A. Beljaars, A. Betts, M. Köhler, M. Miller, and J.-L. Redelsperger, 2004: The simulation of the diurnal cycle of convective precipitation over land in a global model. *Quart. J. Roy. Meteor. Soc.*, **130**, 3119–3137, <https://doi.org/10.1256/qj.03.103>.
- Birch, C. E., D. J. Parker, A. O’Leary, J. H. Marsham, C. M. Taylor, P. Harris, and G. Lister, 2013: Impact of soil moisture and convectively generated waves on the initiation of a West African mesoscale convective system. *Quart. J. Roy. Meteor. Soc.*, **139**, 1712–1730, <https://doi.org/10.1002/qj.2062>.
- , J. H. Marsham, D. J. Parker, and C. M. Taylor, 2014a: The scale dependence and structure of convergence fields preceding the initiation of deep convection. *Geophys. Res. Lett.*, **41**, 4769–4776, <https://doi.org/10.1002/2014GL060493>.
- , D. J. Parker, J. H. Marsham, D. Copsey, and L. Garcia-Carreras, 2014b: A seamless assessment of the role of convection in the water cycle of the West African Monsoon. *J. Geophys. Res. Atmos.*, **119**, 2890–2912, <https://doi.org/10.1002/2013JD020887>.
- , M. J. Roberts, L. Garcia-Carreras, D. Ackerley, M. J. Reeder, A. P. Lock, and R. Schiemann, 2015: Sea-breeze dynamics and convection initiation: The influence of convective parameterization in weather and climate model biases. *J. Climate*, **28**, 8093–8108, <https://doi.org/10.1175/JCLI-D-14-00850.1>.
- Chamberlain, J., C. Bain, D. Boyd, K. McCourt, T. Butcher, and S. Palmer, 2014: Forecasting storms over Lake Victoria using a high resolution model. *Meteor. Appl.*, **21**, 419–430, <https://doi.org/10.1002/met.1403>.
- Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteor. Appl.*, **23**, 165–181, <https://doi.org/10.1002/met.1538>.
- Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate*, **19**, 4605–4630, <https://doi.org/10.1175/JCLI3884.1>.
- Davies, T., M. J. Cullen, A. J. Malcolm, M. Mawson, A. Staniforth, A. White, and N. Wood, 2005: A new dynamical core for the Met Office’s global and regional modelling of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **131**, 1759–1782, <https://doi.org/10.1256/qj.04.101>.
- Done, J., C. A. Davis, and M. Weisman, 2004: The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) Model. *Atmos. Sci. Lett.*, **5**, 110–117, <https://doi.org/10.1002/asl.72>.
- Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.*, **116**, 140–158, <https://doi.org/10.1016/j.rse.2010.10.017>.
- Eagle, C., H. Lean, and S. Webster, 2015: Convective scale NWP model configurations for the Lake Victoria region. Met Office Tech. Rep., 27 pp.
- Fiedler, E. K., M. J. Martin, and J. Roberts-Jones, 2014: An operational analysis of lake surface water temperature. *Tellus*, **66**, 21247, <https://doi.org/10.3402/tellusa.v66.21247>.
- Garcia-Carreras, L., and Coauthors, 2013: The impact of convective cold pool outflows on model biases in the Sahara. *Geophys. Res. Lett.*, **40**, 1647–1652, <https://doi.org/10.1002/grl.50239>.
- Gregory, D., and P. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev.*, **118**, 1483–1506, [https://doi.org/10.1175/1520-0493\(1990\)118<1483:AMFCSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2).

- Guichard, F., and Coauthors, 2010: An intercomparison of simulated rainfall and evapotranspiration associated with a meso-scale convective system over West Africa. *Wea. Forecasting*, **25**, 37–60, <https://doi.org/10.1175/2009WAF2222250.1>.
- Habib, E., A. T. Haile, Y. Tian, and R. J. Joyce, 2012: Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates. *J. Hydrometeorol.*, **13**, 1784–1798, <https://doi.org/10.1175/JHM-D-12-017.1>.
- Haile, A. T., E. Habib, and T. Rientjes, 2013: Evaluation of the Climate Prediction Center (CPC) morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile River. *Hydrol. Processes*, **27**, 1829–1839, <https://doi.org/10.1002/hyp.9330>.
- Hirpa, F. A., M. Gebremichael, and T. Hopson, 2010: Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *J. Appl. Meteor. Climatol.*, **49**, 1044–1051, <https://doi.org/10.1175/2009JAMC2298.1>.
- Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88**, 1783–1794, <https://doi.org/10.1175/BAMS-88-11-1783>.
- Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteor. Climatol.*, **43**, 1834–1853, <https://doi.org/10.1175/JAM2173.1>.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Huffman, G. J., 2017: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree  $\times$  0.1 degree V05. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 20 December 2017, <https://doi.org/10.5067/GPM/IMERG/3B-HH/05>.
- , and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , and Coauthors, 2018: NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG). NASA Tech. Rep., 35 pp., [https://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V5.2.pdf](https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2.pdf).
- Joyce, R. J., and P. Xie, 2011: Kalman filter-based CMORPH. *J. Hydrometeorol.*, **12**, 1547–1563, <https://doi.org/10.1175/JHM-D-11-022.1>.
- , J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503, [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts, 2012: Realism of rainfall in a very high-resolution regional climate model. *J. Climate*, **25**, 5791–5806, <https://doi.org/10.1175/JCLI-D-11-00562.1>.
- Kim, K., J. Park, J. Baik, and M. Choi, 2017: Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.*, **187**, 95–105, <https://doi.org/10.1016/j.atmosres.2016.12.007>.
- Kumar, B., K. C. Patra, and V. Lakshmi, 2016: Daily rainfall statistics of TRMM and CMORPH: A case for trans-boundary Gandak River basin. *J. Earth Syst. Sci.*, **125**, 919–934, <https://doi.org/10.1007/s12040-016-0710-1>.
- Lafore, J. P., and Coauthors, 2017: West African synthetic analysis and forecast. *Meteorology of Tropical West Africa: The Forecasters' Handbook*, D. Parker and M. Diop-Kane, Eds., Wiley-Blackwell, 423–451.
- Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.*, **136**, 3408–3424, <https://doi.org/10.1175/2008MWR2332.1>.
- Liu, Z., D. Ostrenga, W. Teng, and S. Kempler, 2012: Tropical Rainfall Measuring Mission (TRMM) precipitation data and services for research and applications. *Bull. Amer. Meteor. Soc.*, **93**, 1317–1325, <https://doi.org/10.1175/BAMS-D-11-00152.1>.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646, [https://doi.org/10.1175/1520-0469\(1969\)26<636:APARBN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2).
- Luo, Y., and Y. Chen, 2015: Investigation of the predictability and physical mechanisms of an extreme-rainfall-producing meso-scale convective system along the Meiyu front in East China: An ensemble approach. *J. Geophys. Res. Atmos.*, **120**, 10 593–10 618, <https://doi.org/10.1002/2015JD023584>.
- MacCallum, S., and C. Merchant, 2011: ARC-lake: Data product description. University of Edinburgh Tech. Rep., [http://www.geos.ed.ac.uk/arclake/ARCLake\\\_DPD\\\_v1\\\_1\\\_1](http://www.geos.ed.ac.uk/arclake/ARCLake\_DPD\_v1\_1\_1).
- Marshall, J. H., N. S. Dixon, L. Garcia-Carreras, G. Lister, D. J. Parker, P. Knippertz, and C. E. Birch, 2013: The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophys. Res. Lett.*, **40**, 1843–1849, <https://doi.org/10.1002/grl.50347>.
- Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30** (4), 291–303.
- Mason, S. J., and N. E. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quart. J. Roy. Meteor. Soc.*, **128**, 2145–2166, <https://doi.org/10.1256/003590002320603584>.
- Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69**, 3350–3371, <https://doi.org/10.1175/JAS-D-11-0315.1>.
- Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, 2001: Development of a next-generation regional Weather Research and Forecast Model. *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology*, World Scientific, 269–276.
- Mironov, D. V., 2008: Parameterization of lakes in numerical weather prediction. Part 1: Description of a lake model. Tech. rep., Consortium for Small-Scale Modelling Tech. Rep. 11, 47 pp., <http://www.cosmo-model.org/content/model/documentation/techReports/docs/techReport11.pdf>.
- , E. Heise, E. Kourzeneva, B. Ritter, N. Schneider, and A. Terzhevik, 2010: Implementation of the lake parameterisation scheme Flake into the numerical weather prediction model COSMO. *Boreal Environ. Res.*, **15**, 218–230, <http://www.borenav.net/BER/pdfs/ber15/ber15-218.pdf>.
- Mittermaier, M., 2012: Verification. Met Office Tech. Rep. 17.6, MOSAC-17, Met Office, 10 pp., [https://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/f3/mosac\\_17.6\\_mittermaier.pdf](https://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/f3/mosac_17.6_mittermaier.pdf).
- , N. Roberts, and S. A. Thompson, 2013: A long-term assessment of precipitation forecast skill using the fractions skill score. *Meteor. Appl.*, **20**, 176–186, <https://doi.org/10.1002/met.296>.
- Murphy, A. H., and E. S. Epstein, 1989: Skill scores and correlation coefficients in model verification. *Mon. Wea. Rev.*, **117**, 572–582, [https://doi.org/10.1175/1520-0493\(1989\)117<0572:SSACCI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2).

- NCEP, 2017: CMORPH bias corrected V1.0. NCEP, accessed 24 November 2017, [http://ftp.cpc.ncep.noaa.gov/precip/CMORPH\\_V1.0/CRT/8km-30min/](http://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/CRT/8km-30min/).
- Nesbitt, S. W., D. J. Gochis, and T. J. Lang, 2008: The diurnal cycle of clouds and precipitation along the Sierra Madre Occidental observed during NAME-2004: Implications for warm season precipitation estimation in complex terrain. *J. Hydrometeorol.*, **9**, 728–743, <https://doi.org/10.1175/2008JHM939.1>.
- O, S., and P.-E. Kirstetter, 2018: Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. *Quart. J. Roy. Meteor. Soc.*, <https://doi.org/10.1002/qj.3218>, in press.
- , U. Foelsche, G. Kirchengast, J. Fuchsberger, J. Tan, and W. A. Petersen, 2017: Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.*, **21**, 6559–6572, <https://doi.org/10.5194/hess-21-6559-2017>.
- Prakash, S., A. K. Mitra, A. AghaKouchak, Z. Liu, H. Norouzi, and D. Pai, 2018: A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *J. Hydrol.*, **556**, 865–876, <https://doi.org/10.1016/j.jhydrol.2016.01.029>.
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, <https://doi.org/10.1002/2014RG000475>.
- Priestley, A., 1993: A quasi-conservative version of the semi-Lagrangian advection scheme. *Mon. Wea. Rev.*, **121**, 621–629, [https://doi.org/10.1175/1520-0493\(1993\)121<0621:AQCVOV>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<0621:AQCVOV>2.0.CO;2).
- Roberts, N. M., 2003: The impact of a change to the use of the convection scheme to high resolution simulations of convective events (stage 2 report from the storm scale numerical modelling project). Met Office Tech. Rep. 407, 30 pp.
- , 2008: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteor. Appl.*, **15**, 163–169, <https://doi.org/10.1002/met.57>.
- , and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Romilly, T. G., and M. Gebremichael, 2011: Evaluation of satellite rainfall estimates over Ethiopian river basins. *Hydrol. Earth Syst. Sci.*, **15**, 1505–1514, <https://doi.org/10.5194/hess-15-1505-2011>.
- Rooney, G. G., and F. J. Bornemann, 2013: The performance of Flake in the Met Office Unified Model. *Tellus*, **65A**, 21363, <https://doi.org/10.3402/tellusa.v65i0.21363>.
- Saito, K., and Coauthors, 2006: The operational JMA non-hydrostatic mesoscale model. *Mon. Wea. Rev.*, **134**, 1266–1298, <https://doi.org/10.1175/MWR3120.1>.
- Schneider, U., T. Fuchs, A. Meyer-Christoffer, and B. Rudolf, 2008: Global precipitation analysis products of the GPCP. Tech. rep., Deutscher Wetterdienst Tech. Rep., 12 pp., [http://www.mapcruzin.com/environmental-shapefile-maps/water/precipitation/GPCP\\_intro\\_products\\_2008.pdf](http://www.mapcruzin.com/environmental-shapefile-maps/water/precipitation/GPCP_intro_products_2008.pdf).
- Schumacher, R. S., A. J. Clark, M. Xue, and F. Kong, 2013: Factors influencing the development and maintenance of nocturnal heavy-rain-producing convective systems in a storm-scale ensemble. *Mon. Wea. Rev.*, **141**, 2778–2801, <https://doi.org/10.1175/MWR-D-12-00239.1>.
- Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson, 2011: The AROME-France convective-scale operational model. *Mon. Wea. Rev.*, **139**, 976–991, <https://doi.org/10.1175/2010MWR3425.1>.
- Sharifi, E., R. Steinacker, and B. Saghafian, 2016: Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sens.*, **8**, 135, <https://doi.org/10.3390/rs8020135>.
- Skok, G., 2015: Analysis of fraction skill score properties for a displaced rainband in a rectangular domain. *Meteor. Appl.*, **22**, 477–484, <https://doi.org/10.1002/met.1478>.
- , and N. Roberts, 2016: Analysis of fractions skill score properties for random precipitation fields and ECMWF forecasts. *Quart. J. Roy. Meteor. Soc.*, **142**, 2599–2610, <https://doi.org/10.1002/qj.2849>.
- Song, Y., F. H. Semazzi, L. Xie, and L. J. Ogallo, 2004: A coupled regional climate model for the Lake Victoria basin of East Africa. *Int. J. Climatol.*, **24**, 57–75, <https://doi.org/10.1002/joc.983>.
- Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, [https://doi.org/10.1175/1520-0477\(2000\)081<2035:EOPSSE>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2).
- Stephens, G. L., and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**, D24211, <https://doi.org/10.1029/2010JD014532>.
- Stratton, R. A., and Coauthors, 2018: A pan-African convection-permitting regional climate simulation with the Met Office Unified Model: CP4-Africa. *J. Climate*, **31**, 3485–3508, <https://doi.org/10.1175/JCLI-D-17-0503.1>.
- Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? *J. Climate*, **19**, 916–934, <https://doi.org/10.1175/JCLI3672.1>.
- Swets, J. A., 1973: The relative operating characteristic in psychology: A technique for isolating effects of response bias finds wide use in the study of perception and cognition. *Science*, **182**, 990–1000, <https://doi.org/10.1126/science.182.4116.990>.
- Tang, G., Y. Ma, D. Long, L. Zhong, and Y. Hong, 2016a: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over mainland China at multiple spatiotemporal scales. *J. Hydrol.*, **533**, 152–167, <https://doi.org/10.1016/j.jhydrol.2015.12.008>.
- , Z. Zeng, D. Long, X. Guo, B. Yong, W. Zhang, and Y. Hong, 2016b: Statistical and hydrological comparisons between TRMM and GPM level-3 products over a midlatitude basin: Is day-1 IMERG a good successor for TMPA 3B42V7? *J. Hydrometeorol.*, **17**, 121–137, <https://doi.org/10.1175/JHM-D-15-0059.1>.
- Tang, Y., H. W. Lean, and J. Bornemann, 2013: The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteor. Appl.*, **20**, 417–426, <https://doi.org/10.1002/met.1300>.
- Thiery, W., E. L. Davin, H.-J. Panitz, M. Demuzere, S. Lhermitte, and N. van Lipzig, 2015: The impact of the African Great Lakes on the regional climate. *J. Climate*, **28**, 4061–4085, <https://doi.org/10.1175/JCLI-D-14-00565.1>.
- , —, S. I. Seneviratne, K. Bedka, S. Lhermitte, and N. P. van Lipzig, 2016: Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.*, **7**, 12786, <https://doi.org/10.1038/ncomms12786>.
- , L. Gudmundsson, K. Bedka, F. H. Semazzi, S. Lhermitte, P. Willems, N. P. van Lipzig, and S. I. Seneviratne, 2017: Early warnings of hazardous thunderstorms over Lake Victoria. *Environ. Res. Lett.*, **12**, 074012, <https://doi.org/10.1088/1748-9326/aa7521>.

- TRMM, 2011: TRMM (TMPA) rainfall estimate L3 3 hour 0.25 degree  $\times$  0.25 degree V7. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 16 November 2016, <https://doi.org/10.5067/TRMM/TMPA/3H/7>.
- Vié, B., O. Nuissier, and V. Ducrocq, 2011: Cloud-resolving ensemble simulations of Mediterranean heavy precipitating events: Uncertainty on initial conditions and lateral boundary conditions. *Mon. Wea. Rev.*, **139**, 403–423, <https://doi.org/10.1175/2010MWR3487.1>.
- Wang, R., J. Chen, and X. Wang, 2017a: Comparison of IMERG Level-3 and TMPA 3B42V7 in estimating typhoon-related heavy rain. *Water*, **9**, 276, <https://doi.org/10.3390/w9040276>.
- Wang, Z., R. Zhong, C. Lai, and J. Chen, 2017b: Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.*, **196**, 151–163, <https://doi.org/10.1016/j.atmosres.2017.06.020>.
- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW Model. *Wea. Forecasting*, **23**, 407–437, <https://doi.org/10.1175/2007WAF2007005.1>.
- Wessel, P., and W. H. Smith, 1996: A global, self-consistent, hierarchical, high-resolution shoreline. *J. Geophys. Res.*, **101**, 8741–8743, <https://doi.org/10.1029/96JB00104>.
- Weusthoff, T., F. Ament, M. Arpagaus, and M. W. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Mon. Wea. Rev.*, **138**, 3418–3433, <https://doi.org/10.1175/2010MWR3380.1>.
- White, B., A. Buchanan, C. Birch, P. Stier, and K. Pearson, 2018: Quantifying the effects of horizontal grid length and parameterized convection on the degree of convective organization using a metric of the potential for convective interaction. *J. Atmos. Sci.*, **75**, 425–450, <https://doi.org/10.1175/JAS-D-16-0307.1>.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.
- Wood, N., and Coauthors, 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quart. J. Roy. Meteor. Soc.*, **140**, 1505–1520, <https://doi.org/10.1002/qj.2235>.
- World Bank DGF, 2011: Towards a regional framework for weather and climate services for food aid, food security, maritime transport safety contributing to disaster risk reduction in Lake Victoria region. World Bank Development Grant Facility Tech. Rep., 17 pp., [https://www.wmo.int/pages/prog/www/DPFS/Meetings/RAI-EA-SWFDP-RSMT\\_Nairobi2011/documents/DGF-Lake%20Victoria-scoping.pdf](https://www.wmo.int/pages/prog/www/DPFS/Meetings/RAI-EA-SWFDP-RSMT_Nairobi2011/documents/DGF-Lake%20Victoria-scoping.pdf).
- Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.*, **122**, 910–924, <https://doi.org/10.1002/2016JD025418>.
- Yang, G.-Y., and J. Slingo, 2001: The diurnal cycle in the tropics. *Mon. Wea. Rev.*, **129**, 784–801, [https://doi.org/10.1175/1520-0493\(2001\)129<0784:TDCITT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0784:TDCITT>2.0.CO;2).
- Zerroukat, M., 2010: A simple mass conserving semi-Lagrangian scheme for transport problems. *J. Comput. Phys.*, **229**, 9011–9019, <https://doi.org/10.1016/j.jcp.2010.08.017>.
- , and B. Shipway, 2017: ZLF (Zero Lateral Flux): A simple mass conservation method for semi-Lagrangian-based limited-area models. *Quart. J. Roy. Meteor. Soc.*, **143**, 2578–2584, <https://doi.org/10.1002/qj.3108>.
- Zeweldi, D. A., and M. Gebremichael, 2009: Evaluation of CMORPH precipitation products at fine space–time scales. *J. Hydrometeorol.*, **10**, 300–307, <https://doi.org/10.1175/2008JHM1041.1>.



## Erratum

Page numbers refer to the page numbers of the published paper (shown on the top outer corner).

- Note that WRF may be run with convective parametrisation or in CP mode (p2758, 2nd col).
- $N_t$  in equations 4–5 refers to the total number of forecast times that are included in the study and compared to observations. In this study, the analysis is performed over 2 years (including a leap day) with 8 forecast times per day, therefore  $N_t = 5848$  (p2762, 2nd col).
- Correct values for receiver operating characteristic (ROC) area under curve (AUC) are 0.79 and 0.83 for the global and CP models respectively (p2767).
- The transition as a CP model produces high-resolution detail from the smooth initial conditions from the coarser driving model is known as ‘spin-up’. The 3h spin-up period on p2760 refers to the time between the CP model being initialised from the global model and the time at which the first diagnostics are output (T+0). However, the effects of the model spin-up may last beyond 3h, which is the spin-up referred to on p2771.
- On p2771–2773, it is argued that FSS in Roberts and Lean (2008) show no dependence on time of day because ‘rainfall in the midlatitudes is generally controlled by frontal systems, whereas convection in the tropics is strongly forced by the diurnal cycle’. However, it should be noted that Roberts and Lean (2008) chose 10 convectively active days for their analysis. Although there may also have been frontal activity during the analysis period, this is not stated in their paper.



## Chapter 3.

# Identifying Key Controls on Storm Formation over the Lake Victoria Basin

Published in *Monthly Weather Review* (2019)



## Identifying Key Controls on Storm Formation over the Lake Victoria Basin

BETH J. WOODHAMS

*University of Leeds, Leeds, United Kingdom*

CATHRYN E. BIRCH

*University of Leeds, Leeds, and Met Office, Exeter, United Kingdom*

JOHN H. MARSHAM

*University of Leeds, and National Centre for Atmospheric Science, Leeds, United Kingdom*

TODD P. LANE

*University of Melbourne, and ARC Centre of Excellence for Climate Extremes, Melbourne, Australia*

CAROLINE L. BAIN AND STUART WEBSTER

*Met Office, Exeter, United Kingdom*

(Manuscript received 15 March 2019, in final form 20 June 2019)


### ABSTRACT

The Lake Victoria region in East Africa is a hot spot for intense convective storms that are responsible for the deaths of thousands of fishermen each year. The processes responsible for the initiation, development, and propagation of the storms are poorly understood and forecast skill is limited. Key processes for the life cycle of two storms are investigated using Met Office Unified Model convection-permitting simulations with 1.5 km horizontal grid spacing. The two cases are analyzed alongside a simulation of a period with no storms to assess the roles of the lake–land breeze, downslope mountain winds, prevailing large-scale winds, and moisture availability. While seasonal changes in large-scale moisture availability play a key role in storm development, the lake–land-breeze circulation is a major control on the initiation location, timing, and propagation of convection. In the dry season, opposing offshore winds form a bulge of moist air above the lake surface overnight that extends from the surface to ~1.5 km and may trigger storms in high CAPE/low CIN environments. Such a feature has not been explicitly observed or modeled in previous literature. Storms over land on the preceding day are shown to alter the local atmospheric moisture and circulation to promote storm formation over the lake. The variety of initiation processes and differing characteristics of just two storms analyzed here show that the mean diurnal cycle over Lake Victoria alone is inadequate to fully understand storm formation. Knowledge of daily changes in local-scale moisture variability and circulations are keys for skillful forecasts over the lake.

### 1. Introduction

Lake Victoria, in tropical East Africa (Fig. 1), is a hot spot for intense convective storms and lightning (Flohn and Fraedrich 1966; Virts et al. 2013; Albrecht

et al. 2016). It is estimated that 30 million people live on the shores of Lake Victoria, of which ~3.5 million, including 200 000 fishermen, are dependent on the lake for their livelihoods (Semazzi 2011). Boating accidents, mainly associated with severe weather and dangerous water currents, are thought to cause 5000 deaths on the

 Denotes content that is immediately available upon publication as open access.

*Corresponding author:* Beth J. Woodhams, b.j.woodhams@leeds.ac.uk



This article is licensed under a [Creative Commons Attribution 4.0 license](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>).

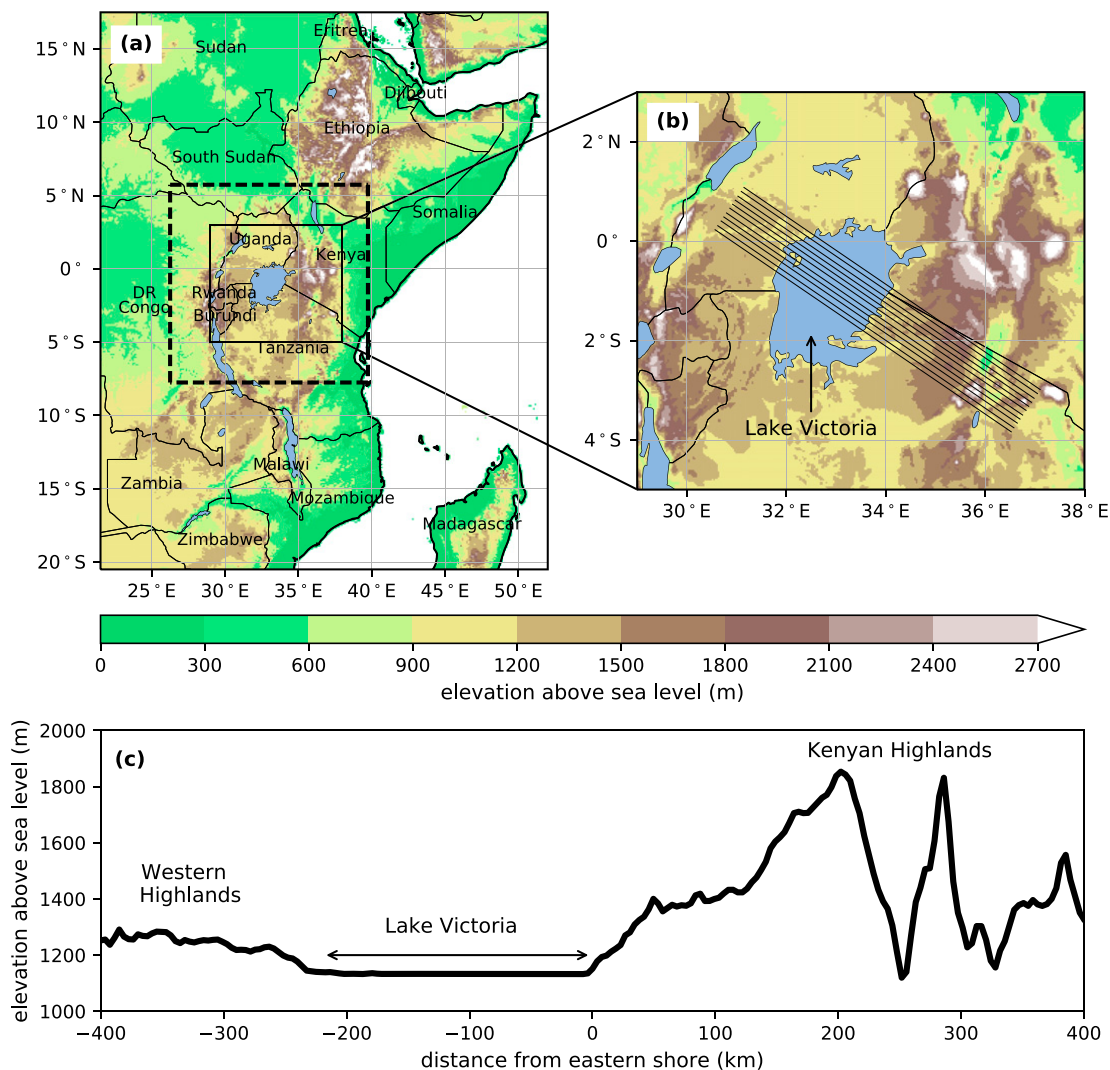


FIG. 1. (a) A map showing the orography over East Africa, covering the domain of the 4.4 km simulation. The dashed box encloses the 1.5 km domain. Elevation data are from the Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model (Hastings and Dunbar 1999). (b) An enlarged section of the 1.5 km domain. Elevation data show the 1.5 km simulation input orography. The black lines show transects parallel to the eastern shore. (c) The mean 1.5 km simulation orography along the transects plotted in (b).

lake every year (Semazzi 2011; Cannon et al. 2014). Despite recent improvements in weather forecast products over East Africa, including the addition of convection-permitting (CP) modeling (Chamberlain et al. 2014; Woodhams et al. 2018), forecasting storms over Lake Victoria remains a major challenge. A basic process-based understanding of the initiation, development, and propagation of individual storms within the Lake Victoria basin region is lacking, as is an understanding of the factors impacting the intensity of the storms, in particular the rainfall and near-surface winds.

Due to Lake Victoria's position on the equator, prevailing winds are from the east over most of the troposphere, with some seasonal variability due to the

movement of the ITCZ (Flohn and Fraedrich 1966; Nicholson 1996). The frequency of storms over the lake is enhanced as the ITCZ moves north over the lake during March–May (known locally as the “long rains”) and returns south between October and December (the “short rains”), although storms occur throughout the year. Winds over the lake have a northerly component between November and March and southerly component between April and October, as the position of the ITCZ changes. Modeling studies have shown that most of the available moisture over the region is advected from the Indian Ocean by the prevailing easterly monsoonal winds, although much of this is blocked by the high ridge of mountains to the east of the lake (Mukabana and Pielke 1996; Anyah et al. 2006).

A regional diurnal circulation system, driven by temperature and moisture differences between the lake and land, locally modifies the large-scale wind pattern across Lake Victoria (Flohn and Fraedrich 1966). This circulation consists of a divergent lake breeze during the afternoon and convergent land breeze overnight, which have a control on convection and rainfall over the lake and surrounding catchment. Flohn and Burkhardt (1985) suggested that this circulation occurs on average 175 days  $\text{yr}^{-1}$  and is responsible for approximately half of the annual precipitation over the lake. Using observations from stations close to the lake shore, Datta (1981) showed the lake breeze to reach its maximum strength at 1500 LT (LT = UTC + 3 h), with rainfall over land occurring in the late afternoon. The land breeze was shown to be strongest between 0600 and 0900 LT, when rainfall is most frequent over the lake. There is a dependence of the strength of the lake–land-breeze circulation on the lake–land temperature contrast (Fraedrich 1972). While the temperature of the lake remains almost constant at around 25°C, the temperature of the air over land can vary between 15° and 30°C over the diurnal cycle (Lumb 1970).

The interaction of the lake and land breezes with the easterly prevailing winds causes an asymmetry in the spatial pattern of winds and rainfall over the lake (Datta 1981; Song et al. 2004). During the evening, the relatively cool and moist lake breeze across the eastern shore converges with the warm and relatively dry air from over the Kenyan Highlands, which is advected toward the lake by the prevailing winds. This convergence maximizes mean rainfall along the east coast between 1800 and 0100 LT, when rainfall is at a minimum over central and western parts of the lake. This minimum occurs due to two secondary effects highlighted in observations by Datta (1981) and Ba and Nicholson (1998) and in model simulations by Thiery et al. (2015). The first is associated with the release of latent heat from deep convection along the eastern shore that is advected westward in the prevailing wind and acts to stabilize the atmosphere over the lake. The second is the compensating subsidence over the lake produced by the return flow of the lake breeze, which occurs between 2000 and 5000 MSL.

The mean rainfall maximum over central and western parts of the lake occurs around 0800 LT and is attributed to convergence between the prevailing easterlies and the land breeze from the western shore during the morning (Datta 1981). Since the atmosphere cools overnight, but the lake surface temperature remains fairly constant, a steep lapse rate in the BL destabilizes the atmosphere over the lake, such that it is conducive to deep convection (Flohn and Burkhardt 1985; Datta 1981). While the

above understanding was developed using a limited number of gauge stations, the diurnal cycle of the spatial rainfall pattern was later verified by several studies using satellite data (Ba and Nicholson 1998; Yin et al. 2000; Nicholson and Yin 2002; Camberlin et al. 2018).

In addition to lake and land breezes, topography has been shown to play an important role in the diurnal cycle over the Lake Victoria basin (Lumb 1970; Okeyo 1986; Mukabana and Piekle 1996; Anyah et al. 2006). The lake sits on a plateau in the Great African Rift System at approximately 1100 m, surrounded by high orography to the west, and the particularly significant slopes of the Kenyan Highlands to the east (Fig. 1c). It is suggested that the highlands may encourage daytime upslope (anabatic) winds, which induce the lake-breeze front (and associated convergence and rainfall) farther inland. At night, downslope (katabatic) winds are expected; as cooler air is brought to the lake shore, a more intense land breeze is generated, with greater convergence and convection over the lake. Asymmetry in the spatial pattern of rainfall over the lake is enhanced, since the steeper mountains to the east encourage stronger katabatic flow, pushing the area of convergence farther west (Anyah et al. 2006).

Other lake-breeze circulations around the world have been studied via field campaigns and modeling. In particular, there is a wealth of research focused on Lake Michigan and the other Great Lakes in North America (Lyons 1972; Lyons and Olsson 1972; Keen and Lyons 1978; Sills et al. 2011). Relative to Lake Victoria, the frequency of lake–land-breeze circulation development is reduced in this region due to the strong influence of midlatitude weather systems. Discussion of land breezes is generally neglected in the literature, especially the interaction of two land breezes from opposite shores. However, the role of land-breeze convergence in the formation of Great Lake snowstorms has been shown by Passarelli and Braham (1981), Ballentine (1982), and Hjelmfelt and Braham (1983).

While many studies have looked at the effect of Lake Victoria on the mean diurnal cycle of storm activity (Mukabana and Piekle 1996; Song et al. 2004; Anyah et al. 2006; Thiery et al. 2015; Camberlin et al. 2018), the dynamical processes involved have not been studied in detail or at the individual storm level. This is due in part to a lack of observations—especially over the lake itself and at a high time and spatial resolution—as well as the requirement of very high-resolution modeling to capture the details of storm dynamics. Studies have shown that high-resolution CP modeling is vital for reproducing observed phenomena associated with sea breezes (Birch and Reeder 2013; Birch et al. 2014, 2015) and the same is likely applicable to lake breezes. In particular,

CP models have improved representation of cold pools (Marshall et al. 2013). Woodhams et al. (2018) and Finney et al. (2019) showed that CP MetUM simulations have an improved diurnal cycle in rainfall over East Africa relative to parameterized models, although Finney et al. (2019) showed that rainfall intensities can be too high over the lake.

This paper uses CP simulations with the Met Office Unified Model (MetUM) at high horizontal resolution (1.5 km grid spacing) to perform a process-based study of the initiation and propagation of two storms over the Lake Victoria basin. The existing literature described here tends to focus on the mean or composited diurnal cycle of circulation or storms, but lacks any detailed analysis of individual storm cases. Such studies have shown that large-scale moisture fluxes into the basin have a strong control on storm formation and that the lake–land–breeze circulation and orography play important roles in determining the mean diurnal cycle of convection. However, it is hypothesized that these different factors and their interactions will differ from the mean on daily time scales. This hypothesis is investigated by comparing case studies of a wet season and dry season storm to a third case study of a dry period over the lake (i.e., a “control” case, which illustrates the underlying lake–land–breeze circulation in dry conditions). This paper identifies the key controls on storm formation and shows that the mean diurnal cycle is insufficient to understand the processes responsible for individual storms due to a diverse range of conditions and triggers for storm initiation. It is noted that the three cases presented here are insufficient to fully represent the variety of processes leading to the formation of storms over the Lake Victoria basin. However, the chosen cases highlight a number of key factors that should be investigated going forward.

The model configuration for the simulations is described in section 2 and the case studies are introduced in section 3a. The dry period case study is used to describe the lake–land–breeze circulation in section 3b and the storm case studies are analyzed in sections 3c and 3d. Section 3e compares and contrasts all three cases. Schematics of the three cases and a comparison table are presented in section 4. The results are discussed and conclusions are drawn in section 5.

## 2. Methods

### a. Model

The MetUM was used to perform the numerical simulations, with a configuration similar to that in Stratton et al. (2018), based on the Met Office UKV

regional model (Tang et al. 2013). The simulations consisted of a driving global model and two doubly one-way nested limited-area convection-permitting (CP) models, with horizontal grid spacing of 4.4 and 1.5 km, respectively. The driving global model and nested regional models were based on the Even Newer Dynamics for General atmospheric modeling (ENDGame) dynamical core (Wood et al. 2014). The driving model fields were taken from archived analysis from the operational MetUM global model [approximate horizontal grid spacing of 17 by 25 km in the tropics, with convection parameterized using the mass-flux scheme introduced by Gregory and Rowntree (1990) and with subsequent enhancements]. The 4.4 km nest was the same as that of the East Africa operational forecast model run by the Met Office in 2014 (Woodhams et al. 2018) and is shown in Fig. 1a. The 1.5 km nest (enclosed within the dashed line in Fig. 1a) was centered on Lake Victoria with approximate dimensions of 1500 km by 1500 km. The configuration of the regional models had 80 terrain-following vertical levels up to a lid of 38.5 km (lowest level  $\sim 1.5$  m) and included the Aranami et al. (2015) mass restoration scheme, which reduces the excessive rainfall rates typical of MetUM models of this configuration (Lean et al. 2008; Kendon et al. 2012). Note that this scheme was not present in the MetUM configuration described in Woodhams et al. (2018), but was present in the configuration in Finney et al. (2019). Time steps of 100 and 60 s were used for the 4.4 and 1.5 km simulations, respectively.

Lake surface temperatures (LSTs) were prescribed as the foundation water surface temperature (temperature below the diurnal warm layer) taken from daily Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA), available on a  $1/20^\circ$  ( $\sim 6$  km) grid (Fiedler et al. 2014). For this dataset, LST observations are obtained from in situ data received via the Global Telecommunication System (GTS, although no in situ observations existed over Lake Victoria during this study) and satellite sea surface temperature (SST) data from the Group for High Resolution SST (GHRSSST). A full discussion of this method and its limitations is given in Woodhams et al. (2018).

### b. Observations

The simulated precipitation and OLR were compared to observations to ensure that the models were producing storms similar to observed. Observations from the IMERG Final Precipitation version 5 (V05) level 3 product from the Global Precipitation Measurement (GPM) mission (Huffman 2017; Huffman et al. 2018) were used. This product is available at 30 min intervals on a regular grid of  $0.1^\circ$ . The product combines and



TABLE 1. Simulation details of the cases studies.

Case	Start date	Run length	Period of interest	Description
Dry period	0000 UTC 9 Jul 2015	72 h	1200 LT 10 Jul–1200 LT 11 Jul 2015	Three-day period with no significant rain over Lake Victoria
Long rains storm	1800 UTC 5 May 2015	72 h	1200 LT 6 May–1200 LT 7 May 2015	Storm forms over land on the evening of 6 May and propagates onto Lake Victoria overnight
Dry season storm	1800 UTC 27 Jul 2016	72 h	1800 LT 28 Jul–1200 LT 29 Jul 2016	Storm forms over Lake Victoria during the early morning of 29 Jul

intercalibrates passive microwave (PMW) precipitation estimates from satellites in the GPM constellation with observations from the Dual-frequency Precipitation Radar (DPR) and a conical-scanning multichannel microwave imager [GPM Microwave Imager (GMI)] on board the GPM *Core Observatory*, using the method developed for the Tropical Rainfall Measuring Mission (TRMM) by Huffman et al. (2007) (Hou et al. 2014). Further algorithms are applied to the satellite data to improve rainfall estimates and spatial coverage. These are detailed in Huffman et al. (2018) and summarized here: calibration against the Global Precipitation Climatology Centre (GPCC) gauge analysis by Schneider et al. (2008); application of the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center morphing technique with Kalman filter (CMORPH-KF) to estimate precipitation outside the sensed area by propagation of PMW estimates with motion vectors derived from geosynchronous IR satellite imagery (Joyce et al. 2004; Joyce and Xie 2011); and application of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Cloud Classification System (PERSIANN–CCS), using IR retrievals calibrated against PMW retrievals (Sorooshian et al. 2000; Hong et al. 2004). Although GPM has been shown to outperform its TRMM predecessor, the product is known to have issues over high orography (Tang et al. 2016; Kim et al. 2017; Xu et al. 2017; Sungmin and Kirstetter 2018) and can underestimate high-intensity events (Wang et al. 2017; Sungmin et al. 2017).

Brightness temperatures  $T_b$  computed from the 10.8  $\mu\text{m}$  IR channel from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument on board the Meteosat Second Generation satellite were also used (Schmetz et al. 2002; EUMETSAT 2012). These data are available at 15 min intervals with approximately 4 km grid spacing, but were interpolated onto a regular grid of  $0.1^\circ$ . For comparison with these observations, the model OLR was converted to a brightness temperature using the Stefan–Boltzmann relationship,  $\text{OLR} = \sigma T_b^4$ . It should be noted that the observed and modeled

brightness temperatures are not directly comparable since the observations only take into account wavelengths within the 10.8  $\mu\text{m}$  IR channel.

Data from the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim) (Dee et al. 2011) on a  $0.7^\circ$  grid with 37 pressure levels were also used to compute monthly mean profiles of winds and specific humidity. The profiles simulated by the MetUM show good agreement with the profiles from ERA-Interim reanalysis for the case study periods, allowing a direct comparison of the simulations to ERA-Interim.

### 3. Results

#### a. Case study introduction

Three case studies were analyzed for this study: one of a 3-day dry period taken from July 2015, one of a storm during the long rains season in May 2015, and one of a storm during the dry season in July 2016. Each case study was run out to 72 h, with output every 1 h. Details of the three case study simulations are summarized in Table 1. To choose the storm case studies, periods when storm events occurred in both GPM observations and the East Africa operational forecast model (UKMO 2017, unpublished data) were chosen, to increase the likelihood that the simulations would produce a storm of sufficient quality to analyze. Similarly, a 3-day dry period was identified in both the observations and model for the dry period case study.

##### 1) DRY PERIOD (9–11 JULY 2015)

The dry period case is taken from the middle of the 2015 JJAS dry season, during a period in which almost no rain fell over Lake Victoria. This case is used to consider the baseline lake–land–breeze circulation, unperturbed by strong convective activity. On all three days of this simulation, both the synoptic-scale and lake–land–breeze circulation patterns over Lake Victoria were remarkably similar (not shown), therefore only results from midday on 10 July to midday on 11 July are analyzed. During the afternoon of 10 July, some

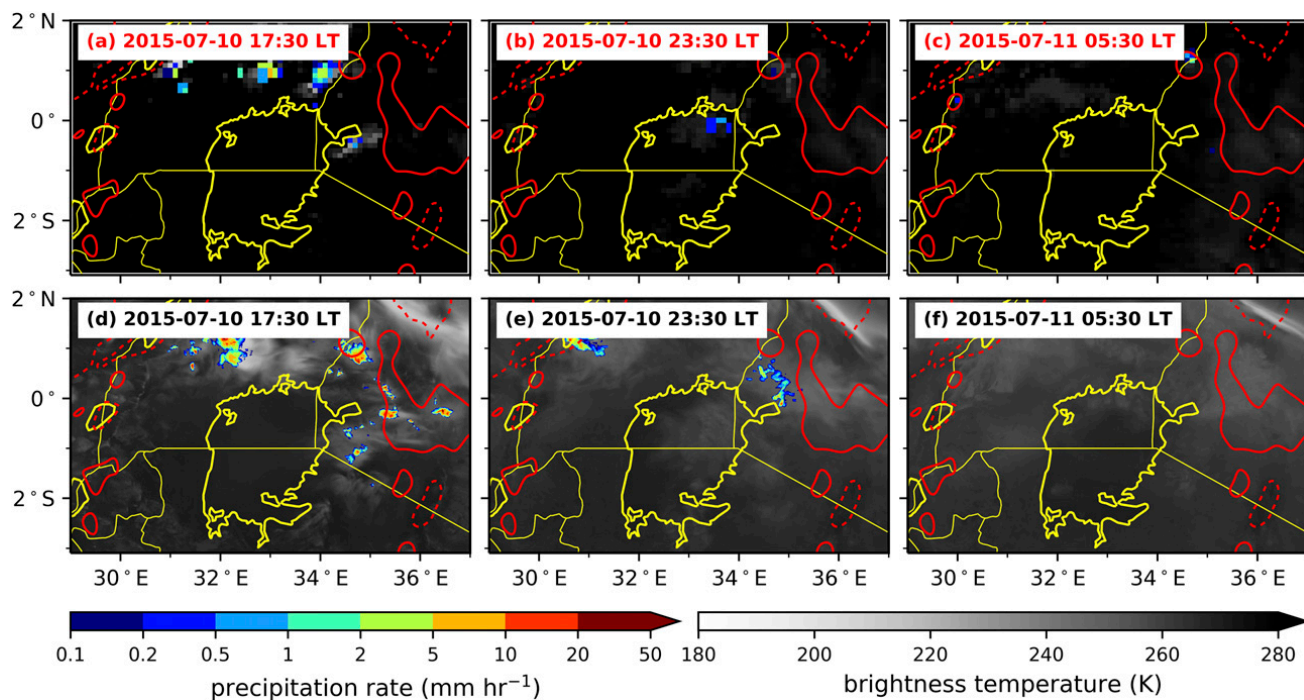


FIG. 2. (a)–(c) (red titles) Observed precipitation from GPM and observed  $10.8\ \mu\text{m}$  brightness temperature from SEVIRI and (d)–(f) (black titles) simulated precipitation (colors) and brightness temperature computed from simulated OLR (grayscale) between 10–11 Jul 2015 (dry period). Red contours give orography at 1 (dashed) and 2 km (solid).

scattered convection occurred to the north and northeast of the lake (Figs. 2a,d), and overnight there was some cloud (reduced brightness temperatures) over the lake in both the model and observations (Figs. 2b,e). Given the relatively warm cloud-top temperatures, the cloud was likely low, although observations show a small area of light rain did occur over the north of the lake, which was not captured by the model. The lake was clear in both the model and observations during the early morning (Figs. 2c,f). Given that the model remained dry over the lake and produced precipitation over land comparable to the observations, it is considered that this simulation is suitable for investigating the lake–land–breeze circulation in relatively dry conditions.

## 2) LONG RAINS (LR) STORM (5–8 MAY 2015)

A similar pattern of storm activity occurred on all three days of this case study, taken from the 2015 long rains season. The storm that initiated on the evening of 6 May is investigated, since this storm was the most significant of the 3-day period. In both the observations and simulation, convection initiated along the eastern shore of Lake Victoria and to the northeast of the lake during the late afternoon (Figs. 3a,j). The convection along the eastern shore was observed to dissipate into the evening (Figs. 3b,c), while the convection to the northeast organized and deepened, with its southwestern edge extending over the northeast of the lake

(Figs. 3d,e). In the model, the convection along the eastern shore persisted into the evening (Figs. 3k,l). The convection to the northeast also organized in the model, but did not extend toward the lake (Figs. 3k–n).

The core of the observed storm moved westward from the land to the lake between 2330 and 0130 LT (Figs. 3d,e). The storm grew and rainfall rates increased, especially over the lake, during the early morning (Fig. 3f). As the storm continued to propagate westward through the morning, it formed large areas with rainfall exceeding  $10\text{--}20\ \text{mm h}^{-1}$ , with much of the heaviest rainfall over the lake (Figs. 3g,h). In the simulation, the convection along the eastern shore propagated westward over the lake from 2130 LT (Figs. 3l–o). This convection organized and deepened as it propagated, forming several embedded convective cells with rainfall rates exceeding  $20\ \text{mm h}^{-1}$  along its leading edge. Similar to the observed storm, the greatest rainfall rates occurred over the northwest of the lake. However, the stratiform region of the simulated storm was situated over the lake to the south and east of the core, whereas it was observed over land to the north of the core.

Between 0730 and 0930 LT, the observed storm began to weaken, although widespread light rainfall persisted to the north of the storm (Figs. 3h,i). In the simulation, the storm began to dissipate once the leading edge reached the northwest shore, and some new convective initiation occurred to the north of the lake (Figs. 3q,r).

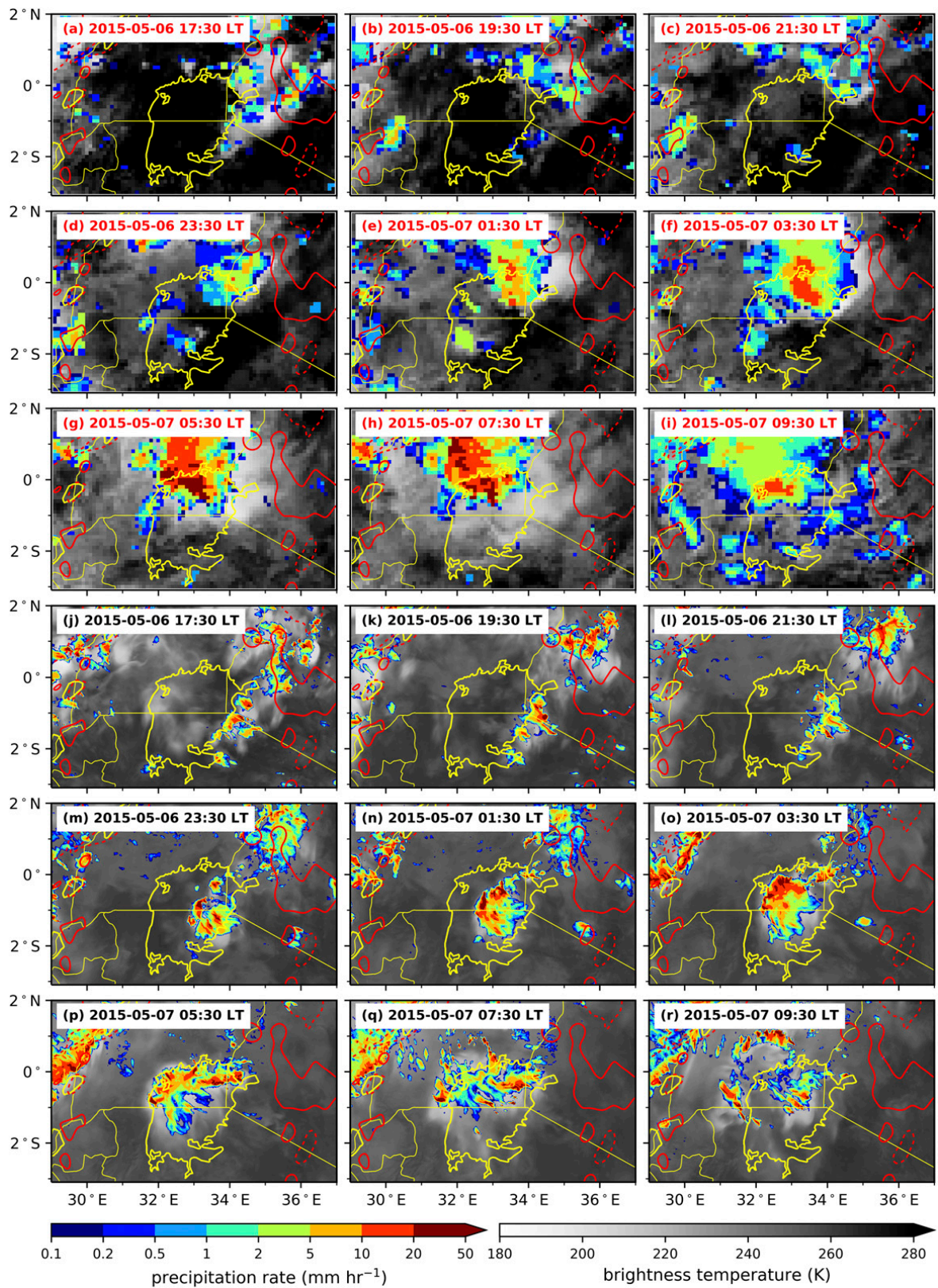


FIG. 3. As in Fig. 2, but for 6–7 May 2015 (LR storm) with (a)–(i) (red titles) showing observations and (j)–(r) (black titles) showing model data.

Although there are some large differences between the structure and location of the observed and simulated storms, the timing and propagation of the storms are similar. In particular, this simulation is appropriate to investigate the evening propagation of convection from land to lake.

### 3) DRY SEASON (DS) STORM (27–30 JULY 2016)

During this case, taken from the 2016 JJAS dry season, storms formed over the center of the lake during the early morning. Similar to the LR storm, a consistent pattern of precipitation occurred each day. The storm that initiated on the morning of 29 July is investigated in detail since it closely matched the observations (Fig. 4). Convection was initiated over the Kenyan Highlands (to the northeast of the lake) during the evening of 28 July and dissipated overnight in both the observations and simulation (Figs. 4a–d,j–o). Note that this storm persisted longer in the simulation than the observations. Convection also occurred to the northwest of the lake at this time, although over a much smaller area in the simulation compared to the observations. A significant storm also occurred in the northeast of the Democratic Republic of Congo (not shown). In both the observations and simulation, a storm initiated over the lake itself around 0630 LT (Figs. 4f,o). While the observed and simulated storms both initiated in the western sector of the lake, the simulated storm initiated approximately 100 km farther north. The storm grew and organized throughout the morning, propagating toward the southwest of the lake, with anvils sheared to the west (more so in the simulation) (Figs. 4g,h,p,q). The storm consisted of several rain cores, but these were smaller and more intense in the simulation, consistent with the bias found by Finney et al. (2019). The simulated storm had dissipated by 1530 LT, while there was still some rainfall associated with the dissipating storm in the observations at this time (Figs. 4i,r).

The good spatial and temporal agreement of the initiation and evolution of the storm over the lake between the simulation and observations make this simulation ideal for investigating an isolated storm over the lake.

While there is generally good agreement between the simulations and reality in all cases, it should be noted that excellent agreement is not essential. A greater importance is placed on sampling a good variety of conditions and types of storm, which these chosen case studies achieve.

#### *b. Case study: Dry period*

Figure 5a shows that the mean prevailing 10 m wind direction is southeasterly across the Lake Victoria basin

during the dry period case study. The near-surface air over land is generally dry, especially to the south of the lake, whereas air over the lake itself has high specific humidity.

The horizontal structure of the lake–land-breeze circulation near the surface is shown in the diurnal cycle of 10 m wind and 1.5 m specific humidity  $q$  (Fig. 6). The vertical structure of the circulation is shown in mean cross sections of the horizontal and vertical wind, virtual potential temperature  $\theta_v$  (a proxy for density), and  $q$  (Fig. 7). The transects along which the mean cross sections are computed are perpendicular to the eastern shore of Lake Victoria and shown by the lines in Fig. 1b. The lake lies between approximately 0 and  $-220$  km on the horizontal scale. In Figs. 7a–f, the along-transect wind (shading) is shown as an anomaly from the mean daily (24-h) flow at that location, to illustrate the local circulation only and make the lake–land-breeze circulation easier to identify. Arrows show the actual wind. Maps of CAPE and CIN during the morning and evening of the case study are shown in Fig. 8.

The boundary layer (BL) over the lake is characterized by high  $q$  throughout the day, due to strong evaporation from the lake surface (Figs. 6 and 7g–l). At midday, both the actual wind and the wind perturbation from the diurnal mean are southeasterly across the lake (Figs. 6a and 7a). At this time, air above the lake has lower  $\theta_v$  relative to the nearby land, despite its greater moisture content, reflecting its lower temperature (Fig. 7g).

During the afternoon, the contrast in  $\theta_v$  between the lake and land increases due to greater daytime heating of the air over land (Figs. 7h,i). Around 1500 LT, a lake breeze A forms over the lowest 1 km above the lake (Figs. 6b and 7b) and advects moisture from the lake BL onshore (Figs. 6b and 7h). The divergence over the lake results in broadscale subsidence above the lake surface (Fig. 7n). The moisture content of the air over land increases and dry convective eddies B deepen the BL to  $\sim 4$  km MSL to the east of the lake (Figs. 7h,n).

The onshore flow across the eastern shore of the lake is opposed by the prevailing southeasterly winds over the lowest  $\sim 6$  km of the atmosphere and the lake-breeze front penetrates only 20–30 km onshore at 1500 LT (Figs. 7b,h). However, the wind anomaly is upslope across the extent of the highlands, suggesting that an anabatic component of wind is working against the prevailing flow (Fig. 7b). Shallow cumulus form where the lake-breeze front penetrates almost 200 km onshore in the west (Figs. 7b,h,n). To the east of the lake, shallow cumulus form farther inland than the lake-breeze front, instead associated with the dry convective eddies (Fig. 7n).

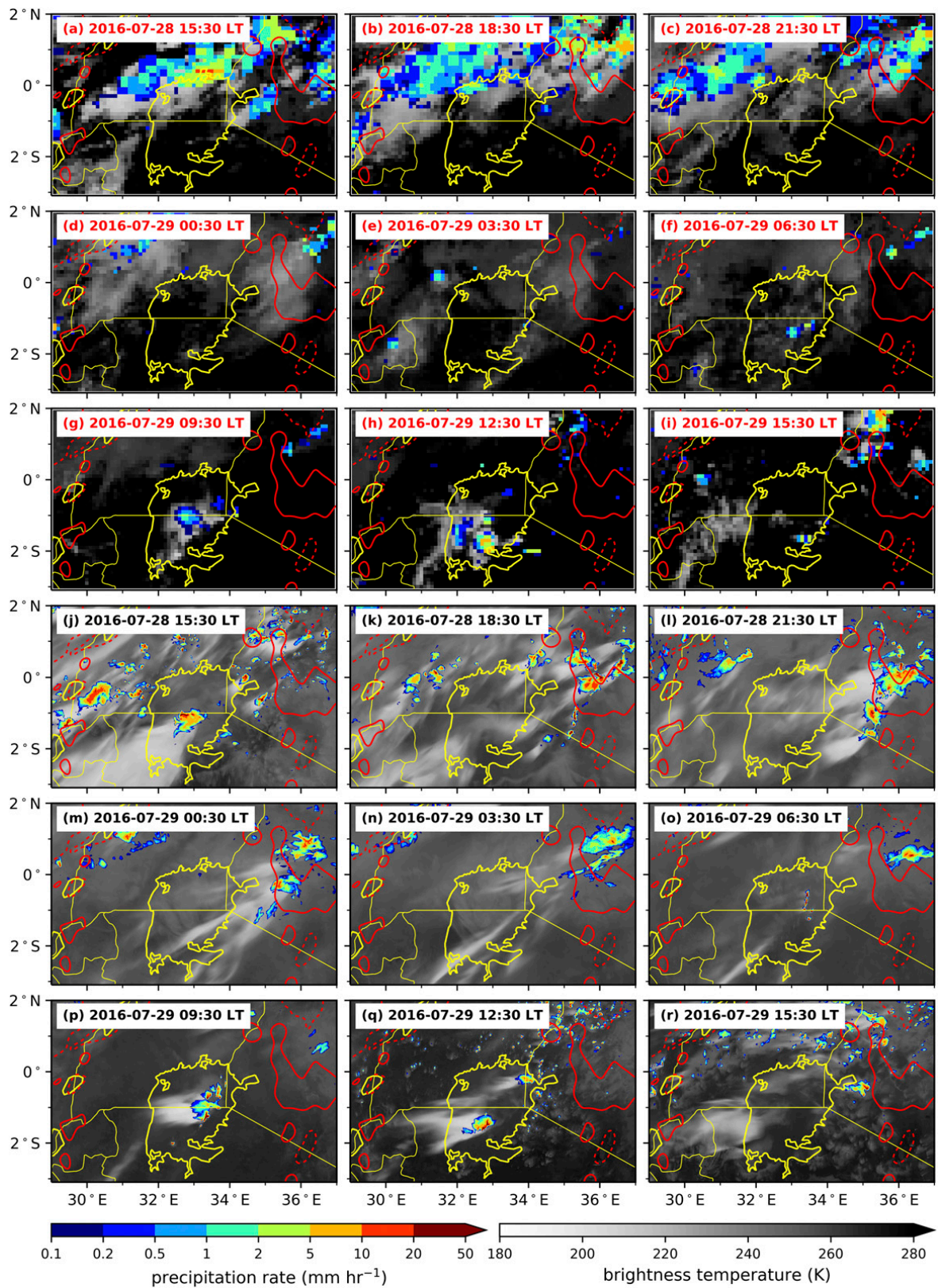


FIG. 4. As in Fig. 2, but for 28–29 Jul 2016 (DS storm) with (a)–(i) (red titles) showing observations and (j)–(r) (black titles) showing model data.

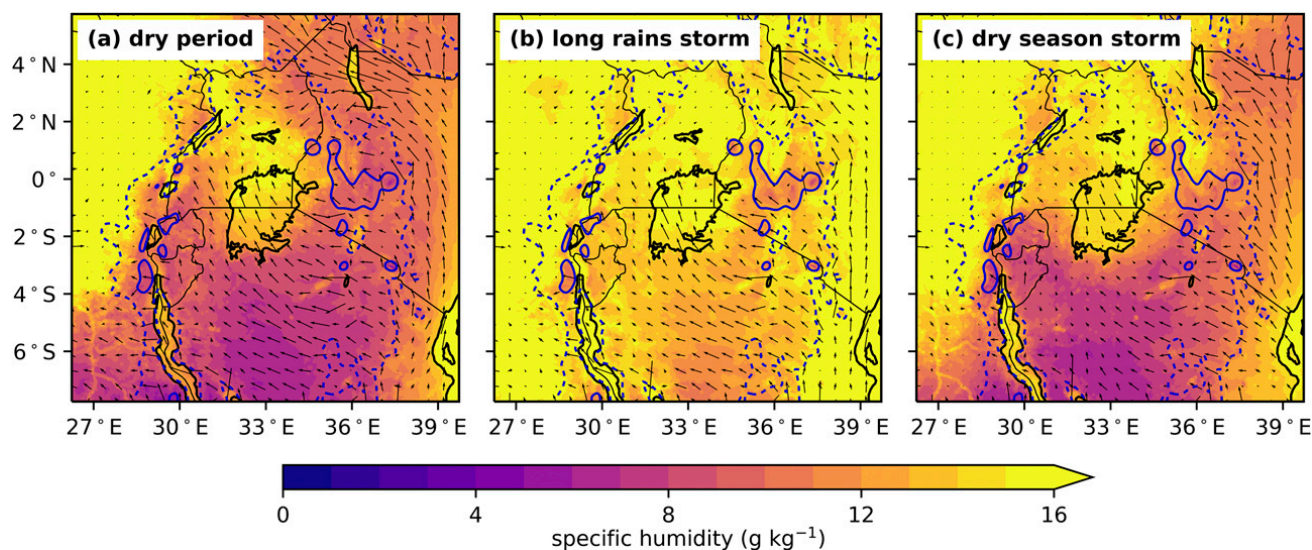


FIG. 5. Mean simulated 10 m winds (vectors) and 1.5 m specific humidity (shading) over a 24 h period from 1200 LT (a) 10 Jul 2015, (b) 6 May 2015, and (c) 28 Jul 2016 for the dry period, LR storm, and DS storm respectively. Blue contours give orography at 1 (dashed) and 2 km (solid).

Around 1800 LT,  $\theta_v$  reaches a maximum over land, and is particularly high approximately 80 km to the east of the lake (Fig. 7i). This marks the leading edge of the lake breeze (Fig. 7c), beyond which the air is well mixed up to 3.5 km MSL (Fig. 7i). As the turbulent eddies subside (Fig. 7o), the prevailing winds from the east, where air is cooler and drier, accelerate and oppose the lake breeze (Figs. 7c,i). Strong ascent occurs where the lake breeze and prevailing winds converge (Fig. 7o). While there is some CAPE to the east of the lake, CIN is also high, so deep convection is suppressed (Figs. 8a,d).

By 1800 LT, a strong return flow of the lake breezes over both shorelines is established between 3–5 km MSL; on the lake side of the eastern convergence line, a branch forms back toward the lake (C) and, on the land side, a branch flows toward the mountains (D) (Fig. 7c). The lake-side return flow advects moisture from the land BL back toward the lake (Fig. 7i).

Following sunset, the surface air cools, in particular over the land, and the gradient in  $\theta_v$  reverses between the lake and land (Fig. 7j). To the east of the lake, the southeasterly wind accelerates and low  $\theta_v$  air from east of the Kenyan Highlands is advected toward the lake over a depth of  $\sim 2$  km (Figs. 7d,j). Over the lowest few hundred meters, a land breeze E also forms at the eastern shoreline (Figs. 7e,k). By 0200 LT, the land breeze is encompassed by the prevailing southeasterly flow, which likely includes a katabatic component. The front and upper boundary of this density current are not well defined, therefore E roughly marks the land-breeze front in Figs. 7k and 7l.

Driven by the offshore flow, the convergence line to the east of the lake begins to propagate westward, crossing the lake shore at approximately 2200 LT (Figs. 6d and 7d). The return-flow pattern (C, D) between 3 and 5 km MSL is tied to the convergence and therefore also propagates toward the lake (Fig. 7d). Uplift occurs at and ahead of the convergence (Fig. 7p), where moist air over the lake is pushed up from the surface into a shallow bulge F (only a few hundred meters in depth) as the density current from the land runs into the stable layer over the lake (Fig. 7j). Although there is no actual offshore flow over the western shoreline at 2200 LT, the retreating westerly lake breeze (pushed back across the lake by the offshore flow from the east) means that the wind perturbation relative to the diurnal mean is offshore (Figs. 6d and 7d).

The convergence, and leading edge of the moisture bulge, propagate westward across the lake, reaching the lake center at  $\sim 0200$  LT (Figs. 7e,k,q). A very shallow land breeze G (limited to only the lowest  $\sim 300$  m) forms across the western shore (Figs. 6e and 7e,k). The return flows above the convergence (C, D) are now both toward land, such that there is midlevel divergence over the center of the lake (Fig. 7e). Despite vertical motion and upward transport of moisture, there is no CAPE over the lake (except for some small patches to the north) so deep convection is suppressed (Fig. 8g).

The leading upper edge of the moisture bulge F coincides with the upper boundary of the low- $\theta_v$  air advected from the east. The center of the bulge appears to propagate westward with a more well-defined density current (likely the land breeze E) at low levels, and is

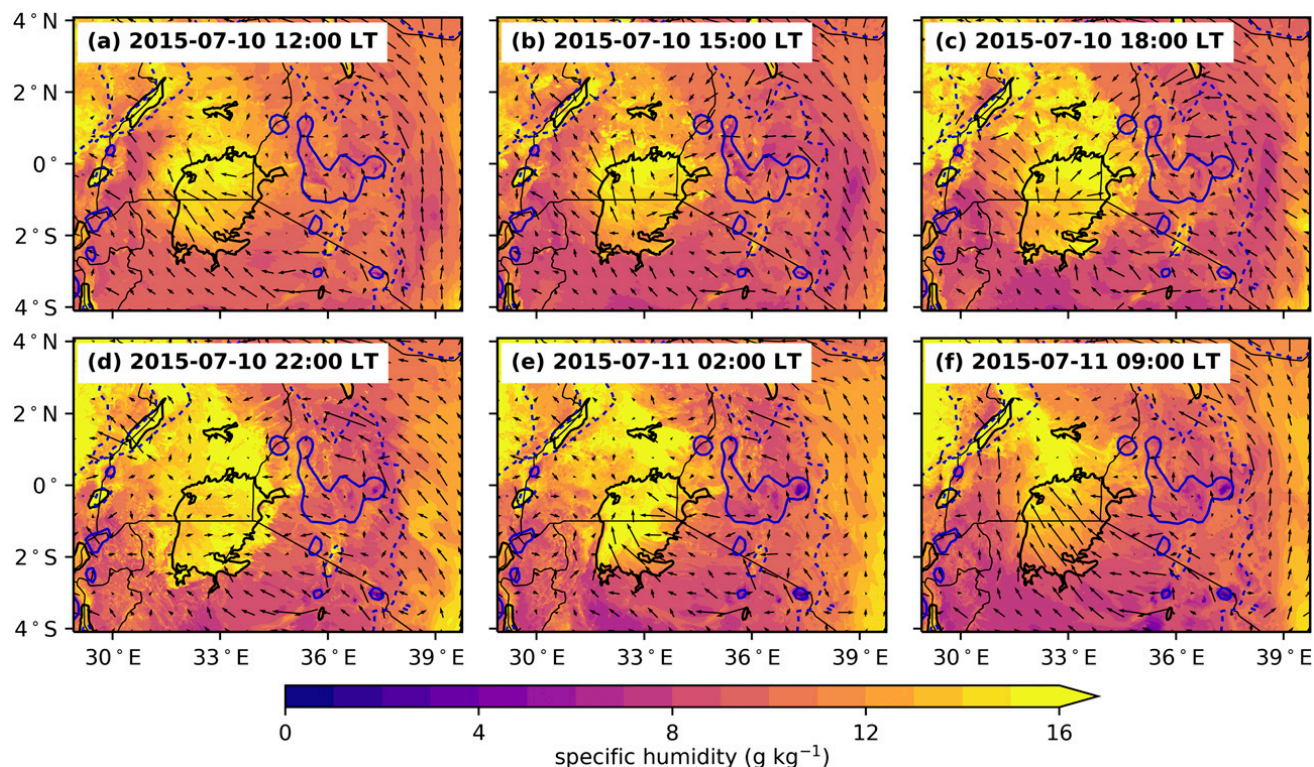


FIG. 6. Simulated 10 m winds (vectors) and 1.5 m specific humidity (shading) during the dry period case study. Blue contours give orography at 1 (dashed) and 2 km (solid).

bounded to the west by the westerly land breeze G (Fig. 7k). As such, the bulge grows horizontally. It also grows in height, in particular at the leading edge, exceeding 1 km at 0200 LT.

During the night, the specific humidity of air between ~2–5 km above sea level over the lake is also elevated relative to the surroundings (Figs. 7j–l, H). Although there is vertical motion up to this height over the lake, this increased moisture appears unrelated to the low-level moisture bulge. Rather, it exists due to westward advection of moist air from the daytime BL over land by the westward return flow of the lake breeze C on the preceding evening. Due to shear in the eastward return flow of the land breeze D (Figs. 7d–f), the top layer of moist air over the lake is advected back toward land (between 0 and +200 km, Figs. 7j–l).

Through the early morning, the easterly winds strengthen again to the west of the lake and prevail across the whole lake by 0900 LT (Figs. 6f and 7f). However, the wind perturbation across the western shore remains eastward, so there is shear across the lowest 1 km above this region and the top of the bulge F propagates ahead of the lower leading edge (Figs. 7f,l). As the bulge propagates onshore, the depth of the moist air increases, but the bulge loses its individual identity as

it merges with the daytime deepening of the land BL (Figs. 7f,l,r).

c. Case study: Long rains storm

During this case study, a storm forms inland of the eastern shore of Lake Victoria during the evening and propagates across the lake overnight. The near-surface large-scale wind field preceding the storm is very similar to that of the dry period, with prevailing southeasterlies across the lake (cf. Figs. 5a,b). However, the near-surface air has much higher specific humidity relative to the dry period case, especially over land.

The prevailing wind over the lake during the day is disrupted by a storm A over the northwest of the lake, which has persisted from the previous night. (Figs. 9a,b,m,n). However, a lake breeze B still forms over both the eastern and western shorelines around 1500 LT (Figs. 9b,h). As in the dry period case, the lake breeze across the eastern shore converges with the southeasterly prevailing winds and the lake-breeze front is prevented from penetrating far inland (Fig. 9b). Deep moist convection C initiates along the convergence line and precipitation occurs over the foothills of the Kenyan Highlands (Fig. 9n). Pockets of high CAPE (exceeding 1000 J kg<sup>-1</sup>) exist over this region, as well as areas of CIN

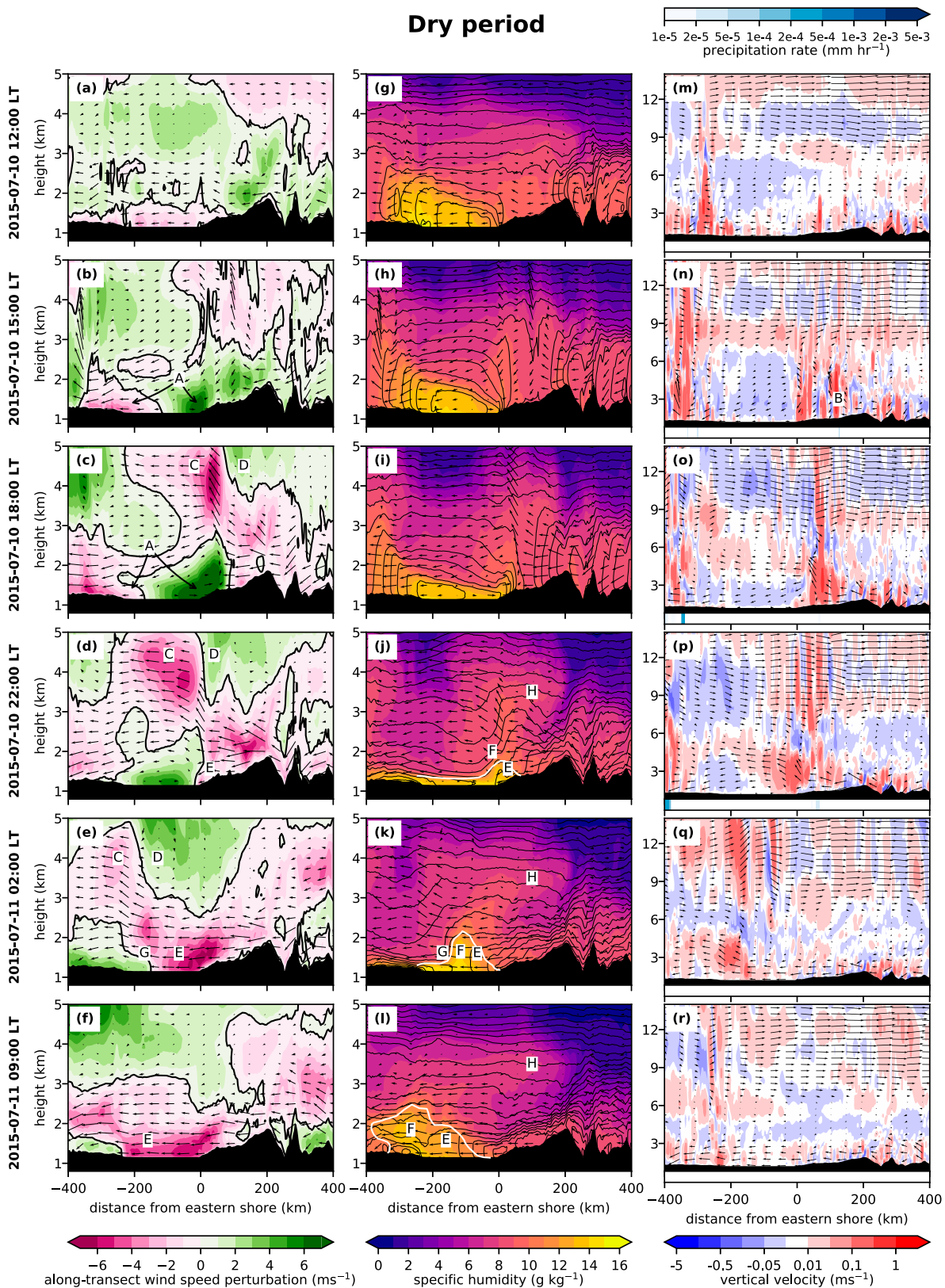


FIG. 7. Cross sections of (a)–(f) simulated along-transect wind speed anomaly, computed with respect to the 24 h mean at each grid point (shading, positive for northwesterly anomalies), (g)–(l) specific humidity (shading) and virtual potential temperature (contours), and (m)–(r) vertical wind speed (upper panel) and precipitation rate (lower panel) along a transect perpendicular to the eastern shore of Lake Victoria (averaged over the black lines in Fig. 1b) during the dry period case study. Actual wind vectors (not anomalies) in the along-transect and vertical directions are shown by the arrows. The white contour in (g)–(l) roughly marks the boundary of the moisture bulge described in section 3b. Black contours in (m)–(r) show regions where the sum of the cloud liquid water content and cloud ice content exceeds  $0.1 \text{ g kg}^{-1}$ . Labels A–H are described in the text.



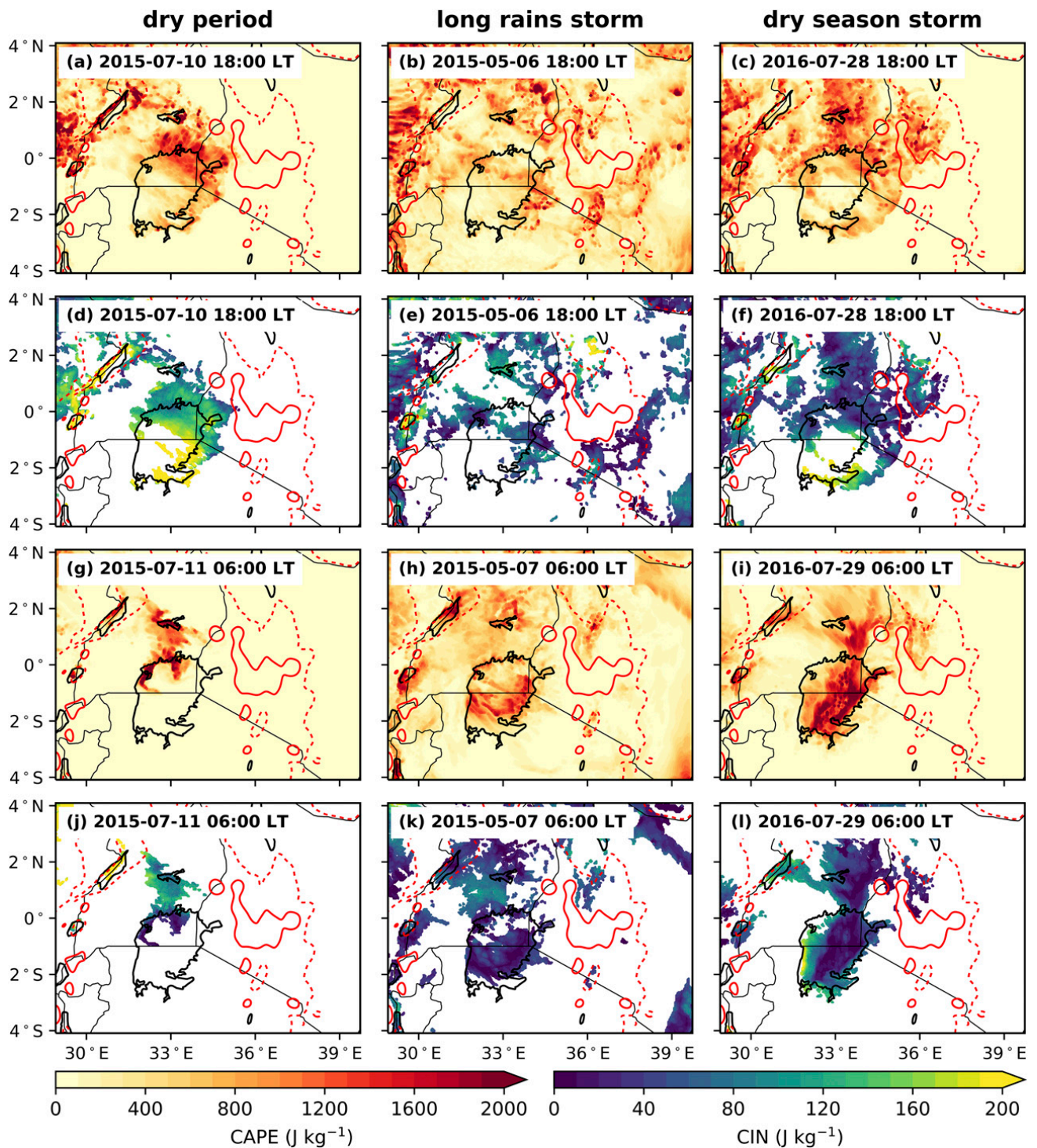


FIG. 8. Maps of (a)–(c),(g)–(i) CAPE and (d)–(f),(j)–(l) CIN (only plotted where CAPE > 400 J kg<sup>-1</sup>) for (a)–(f) 1800 LT and (g)–(l) 0600 LT for the (left) dry period, (middle) LR storm, and (right) DS storm. CAPE is computed using a parcel from 70 m above the surface. Red contours give orography at 1 (dashed) and 2 km (solid).

exceeding 100 J kg<sup>-1</sup> (Figs. 8b,e), showing that there is sufficient dynamical triggering to overcome the CIN and initiate deep convection. The convection is invigorated by the very moist air shown by the low contrast in specific humidity over the land and lake

relative to the dry case (cf. Figs. 9g–l and 7g–l). The region of convergence beneath storm C remains almost stationary into the evening, as the convection deepens, organizes and a larger, contiguous area of precipitation forms (Figs. 3l,m and 9c,d,o,p).

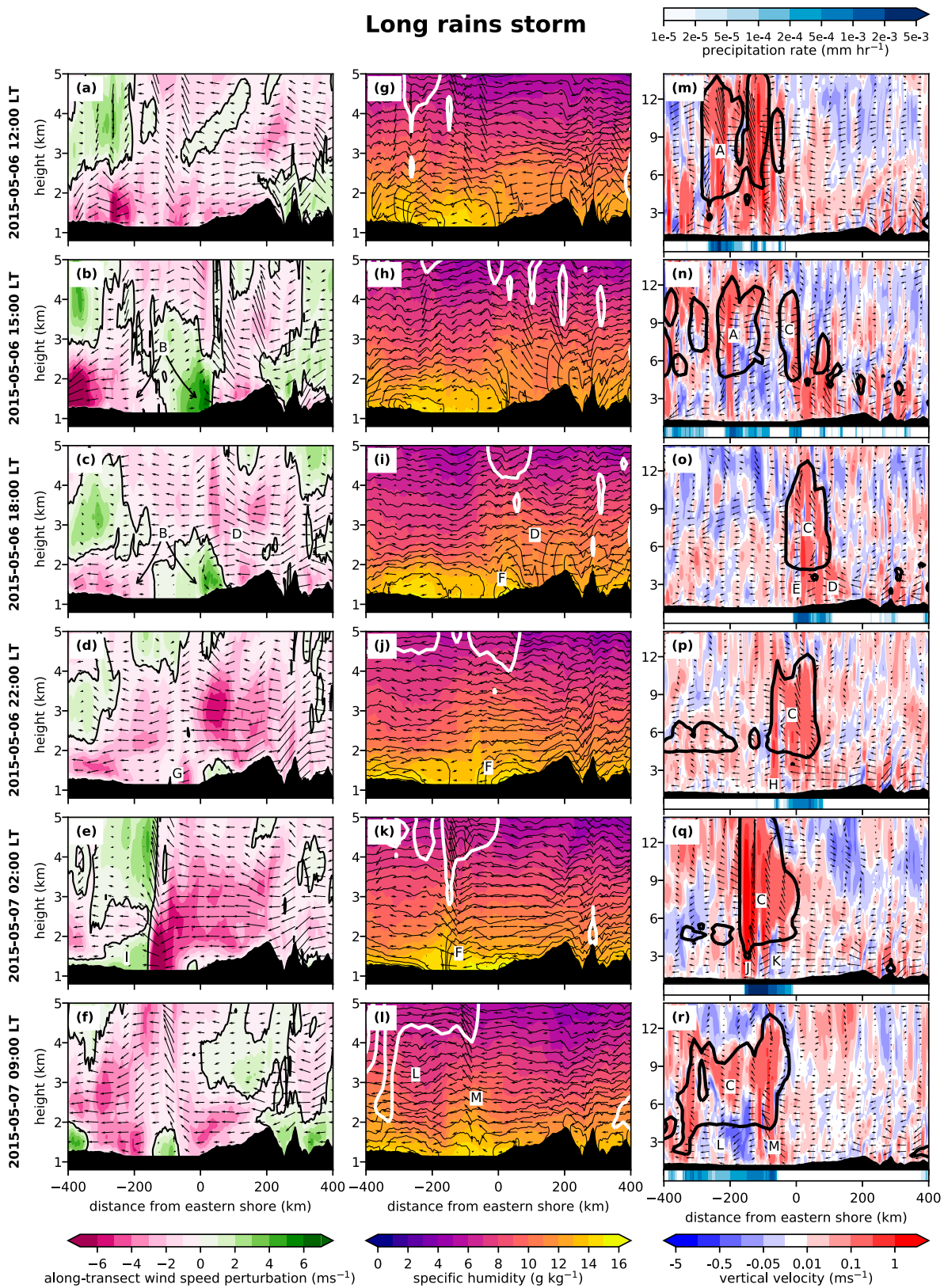


FIG. 9. As in Fig. 7, but during the LR storm case study. Note, white contours in (g)–(l) now show the same as the black contours in (m)–(r). Labels A–M are described in the text.

The deep convection disrupts the return flow of the lake breeze (between 2 and 4 km above the lake surface) such that it is less defined relative to the dry case (cf. Figs. 9c and 7c). The moist air from over land is transported vertically by the moist convection, rather than being advected back over the lake (Figs. 9h–j,n–p). During the initial stages of the storm, much of the inflow originates from the mountains to the east and flows into the rear of the storm, over a depth of approximately 4 km (Figs. 9b–d,h–j,n–p, D).

Around 1800 LT, a downdraft E begins to form beneath the storm, located just over the eastern shoreline (Fig. 9o). Beneath the downdraft, a density current F forms and propagates toward the lake (Fig. 9j). Given that  $\theta_v$  and  $q$  directly below the storm do not change, it is likely that the density current corresponds to the formation of a land breeze, rather than cold pool outflow from the storm. Similar to the dry period, this density current forms ahead of the cooler air to the east, but the two flows merge by 2200 LT, when the southeasterly winds accelerate (Figs. 9c,d,i,j).

Southeasterly winds prevail throughout the atmospheric column (i.e., there is low shear) at 2200 LT, such that the leading edge of the storm propagates along with the density current (Figs. 9d,j,p). Convergence occurs at the leading edge of the density current (G,  $\sim 70$  km west of the eastern shore) due to deceleration as the eastern land-breeze front runs into the rear end of the lake breeze across the western shore (Fig. 9d). A new updraft H forms here and lofts moist air from above the lake surface into the storm as the rear inflow (consisting of air from over the land) is reduced (Figs. 9j,p).

The storm C continues to grow in height and horizontal extent throughout the evening (Fig. 9q). At 0200 LT, a land breeze I forms across the western shore and converges with the easterlies at the leading edge of the storm (composed of the prevailing winds and eastern land breeze and strengthened by the downdraft outflow) (Figs. 9e,k,q). The convergence strengthens the updraft at the storm's leading edge, increasing the transport of moist lake BL air into the storm (Fig. 9k). High CAPE and low CIN over the lake (Figs. 8h,k) encourage storm growth, such that the top of the storm extends to a height of over 15 km MSL (Fig. 9q). Rainfall rates are greatest at the leading edge of the storm, which is situated  $\sim 150$  km from the eastern shore, and along which an arcus cloud J forms (Fig. 9q). To the rear of the storm, inflow is reduced further as a large region of downdraft K forms  $\sim 100$  km behind the leading edge (Fig. 9q).

As the storm continues to propagate westward, the land breeze from the west retreats (Fig. 9f). By 0900 LT, the storm loses organization and its cloud-top height is

greatly reduced, as are the size and strength of its updrafts (Figs. 3r and 9r). Behind the large arcus cloud centered at  $-350$  km on the horizontal scale, there is a large downdraft L extending across almost 200 km, (Fig. 9r). The downdraft forms a cold pool, shown by the low  $\theta_v$  air and reduced specific humidity below the storm (Fig. 9l). Although the storm is dissipating, a new updraft M forms at the rear, over the moist supply of air above the center of the lake (Figs. 9l,r). However, the updraft weakens during the late morning as divergence occurs over the lake due to the formation of lake breezes across both shores (not shown).

#### d. Case study: Dry season storm

During this case study, the mean 10 m wind field over the lake differs to the previous two cases, with northwesterly winds (as opposed to southeasterly) prevailing over the northwest of the basin and the lake itself (cf. Fig. 5c to Figs. 5a,b). However, to the south and east of the lake, the large-scale circulation pattern is similar to the previous cases. The near-surface moisture during the DS storm case is similar to the dry period case (although there is increased specific humidity to the north of the lake in the former case). This similarity is to be expected since both cases are taken from the same time of year.

In this case study, a storm initiates over the center of the lake during the early morning. During the preceding evening, lake breezes A occur over the northwestern and eastern shores, advecting moisture onto the land (Figs. 10a,g). Convection occurs along the lake-breeze fronts, where there is high CAPE and low CIN (Figs. 4k and 8c,f). Note that the large storm to the east (over the highlands, Fig. 4k) is not captured within the transect in Fig. 10. However, some small convective cells B do occur within the transect (Figs. 4k and 10m). The moisture for these is provided by the lake breeze, since the air to the east of the lake (at  $> +150$  km) is very dry (Fig. 10g).

During the preceding evening, a large storm also occurs in the northeast of the Democratic Republic of Congo (not captured in the transect), a region which is generally very moist as it is under the influence of the Congo air mass (not shown). The storm deposits particularly cool (and therefore low  $\theta_v$ ) air C in this region (Fig. 10h). As such, a strong land breeze D forms over a depth of  $\sim 1$  km across the western shore at 2200 LT (Figs. 10b,h). The land breeze advects the moist air from the Congo region toward the lake.

The northwesterly land breeze is opposed by weaker offshore flow across the eastern shore E, consisting of the easterly land breeze embedded within the prevailing southeasterlies (Fig. 10b). At 2200 LT, a moist bulge F begins to form at the convergence between the two land breezes, close to the eastern shore (Figs. 10b,h). As the

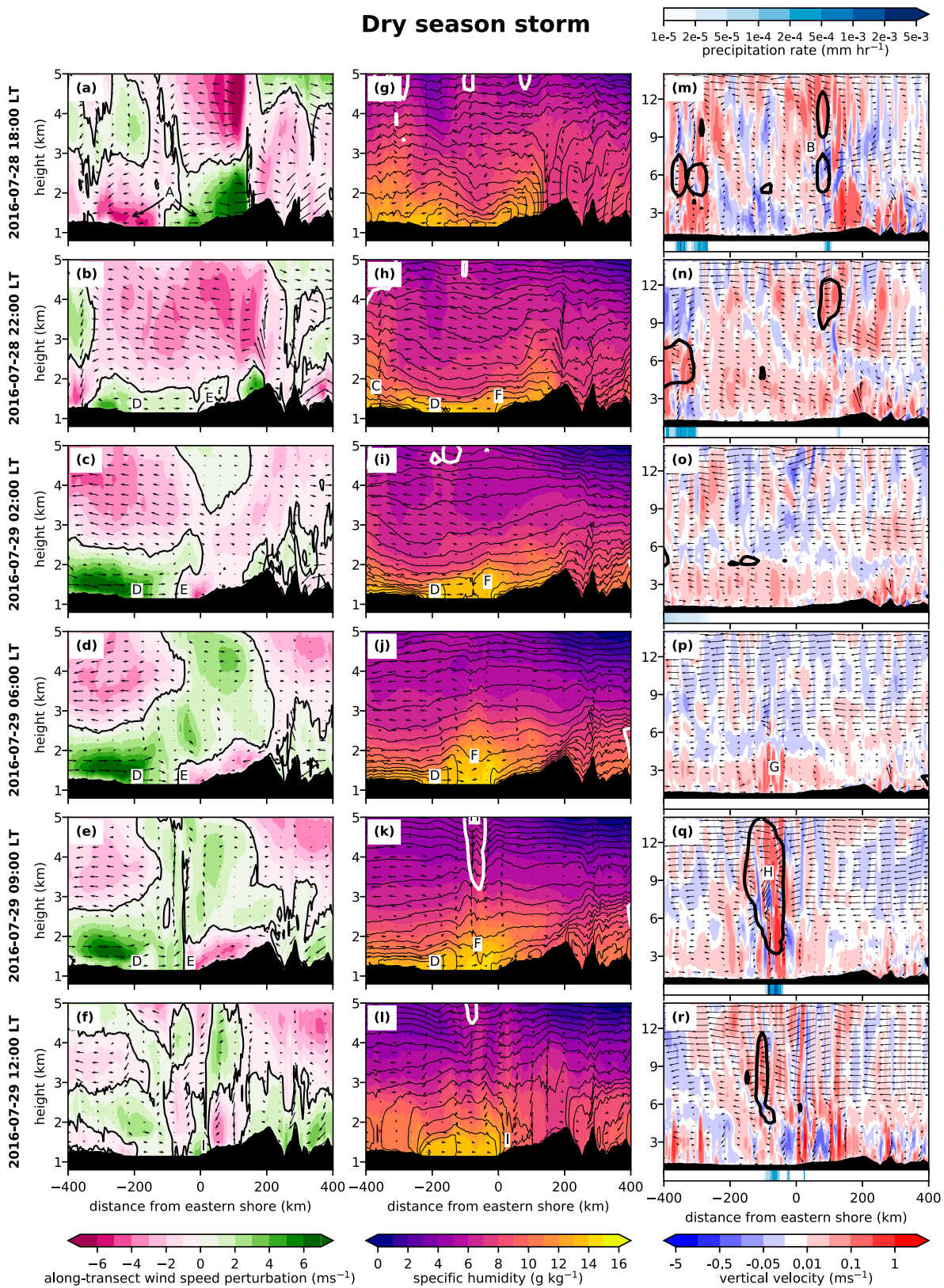


FIG. 10. As in Fig. 9, but during the DS storm case study. Labels A–I are described in the text.

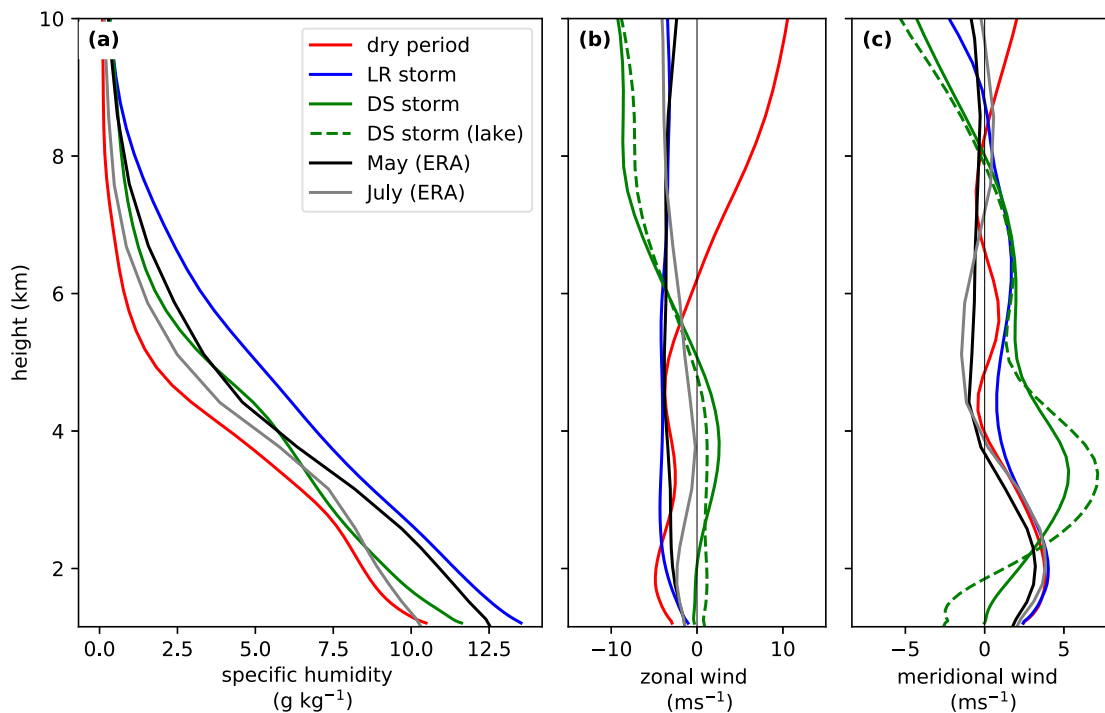


FIG. 11. (a) Mean specific humidity, (b) mean zonal wind, and (c) mean meridional wind taken over a 24 h period from 1200 LT 10 Jul 2015, 6 May 2015, and 28 Jul 2016 for the dry period, LR storm, and DS storm respectively. Climatological monthly means for May and July from ERA-Interim (1997–2016) are also shown (Dee et al. 2011). The solid lines correspond to the mean over the region within the solid box in Fig. 1a and the dashed line corresponds to the mean over Lake Victoria.

easterly flow strengthens between 2200 and 0200 LT, the convergence moves to approximately 50 km west of the eastern shore, as does the center of the moist bulge (Figs. 10c,i). This region of convergence remains stationary, and the bulge deepens, between 0200 and 0900 LT. (Figs. 10c–e,i–k). The strength of the western land breeze prevents westward propagation of the convergence, despite the southeasterly prevailing winds.

At 0600 LT, there is strong vertical motion G associated with the moist bulge F (Figs. 10j,p). High CAPE and low CIN over the lake (Figs. 8i,l) allow deep convection H to be triggered above the bulge between 0600 and 0900 LT (Figs. 10k,q).

The land breeze from the west weakens during the late morning, as the temperature gradient between the lake and the land reverses (Figs. 10f,l). The convergence line at the surface is pushed back toward the eastern shoreline by the development of a lake breeze I (Figs. 10f,l). However, the storm does not follow the low-level convergence due to the wind shear caused by the persisting southeasterlies at upper levels (Figs. 10f,l,r). As a result, the storm loses its moist air supply and the storm dissipates.

*e. Case study comparison*

Some differences between the case studies may be explained by the differing mean specific humidity

profiles of the atmosphere during each period (Fig. 11a). Climatologically,  $q$  is increased throughout the atmospheric column during May relative to July (black and gray lines respectively in Fig. 11a). The difference is greatest at the surface, where  $q$  is increased by  $>2 \text{ g kg}^{-1}$  in May. This difference is symptomatic of the former lying in the MAM long rains season and the latter during the JJAS dry season. The moisture profile for the May case study compared to the July case studies reflects these climatological differences (cf. blue line to red and green lines). However,  $q$  at the surface is  $\sim 1 \text{ g kg}^{-1}$  greater than the May mean during the LR case (cf. blue and black lines) and  $\sim 1.5 \text{ g kg}^{-1}$  greater than the July mean during the DS storm case (cf. green and gray lines), demonstrating the importance of additional moisture for storm initiation.

In general, the climatological zonal and meridional winds show little change between May and July, although the midlevel zonal component is almost zero during July (Figs. 11b,c, compare black and gray lines). Below 4 km, the wind direction is predominantly southeasterly across the region (Figs. 11b,c). The case study profiles are similar to the climatologies at mid-levels and below, apart from the DS storm. During this case, zonal winds below 5 km are predominantly from the west and, over the lake, winds in the lowest  $\sim 700 \text{ m}$

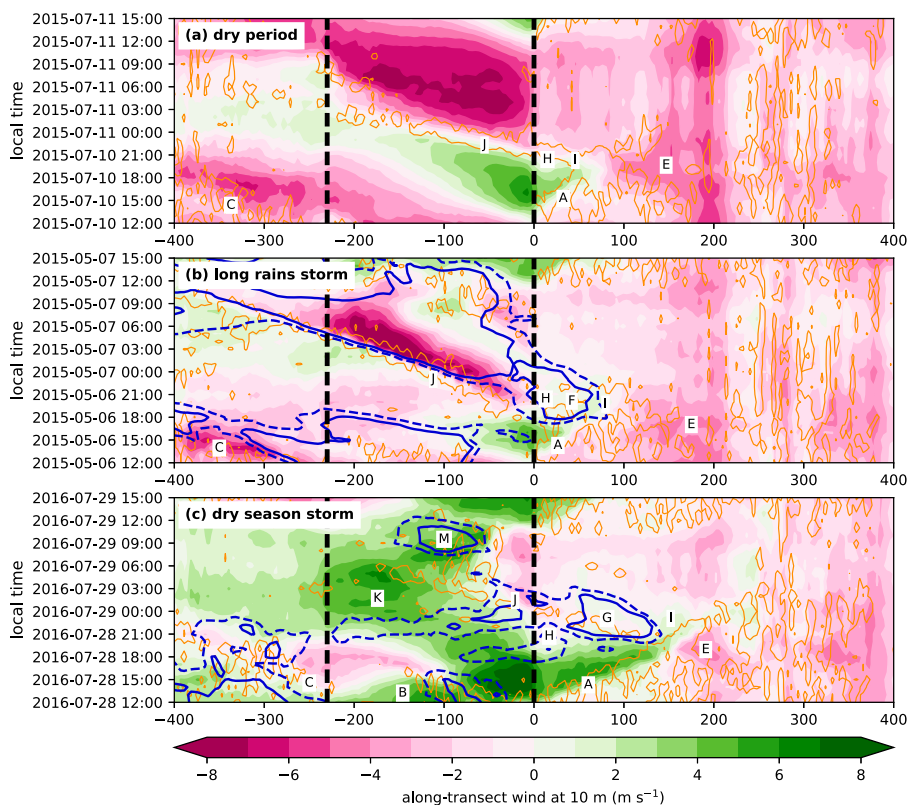


FIG. 12. Hovmöller of simulated along-transect 10 m wind speed for (a) the dry period, (b) the LR storm, and (c) the DS storm. Orange contours show regions with vertical velocity greater than  $0.05 \text{ m s}^{-1}$  at 1000 m AGL and blue contours show regions of OLR less than 200 (dashed) and 170 (solid)  $\text{W m}^{-2}$ . The black dashed lines show the boundaries of Lake Victoria. Labels A–M are described in the text.

have a component from the north. At upper levels, the wind profiles during the case studies vary significantly from one another and from the monthly climatologies, suggesting high day-to-day variability. There is almost no shear across the atmosphere during the LR storm, but high shear during the DS storm (cf. blue and green lines). The lack of shear in the former was shown to propagate the storm with the low-level convergence, increasing its longevity. In the latter case, the shear decoupled the convergence and storm, causing the storm to dissipate.

Differences between the cases are also highlighted in Hovmöller plots along the transect defined in Fig. 1b (Figs. 12 and 13). The along-transect 10 m wind shows many similarities between the dry period and LR case studies (Figs. 12a,b). Lake breezes A initiate across the eastern shore around 1200 LT and propagate  $\sim 80$  km onto the land (green shading to the right of the 0 km dashed line), advancing at  $\sim 3 \text{ m s}^{-1}$  (although the speed is more variable in time during the dry period). Preceding the DS storm, the eastern lake breeze A forms at a similar time, but propagates 160 km farther inland and at a greater, but more variable, speed

( $\sim 4.5\text{--}6.5 \text{ m s}^{-1}$ ), as it is reinforced by anomalously strong wind B across the lake from the northwest (Figs. 12c and 13c).

Lake breezes C also form across the western shore, just before 1200 LT, and propagate  $\sim 240$  km to the west in the dry period and LR storm case studies (Figs. 12a,b). The anomalously low  $\theta_v$  air to the northwest D preceding the DS storm delays, weakens, and reduces the extent of the western lake breeze C (Figs. 12c and 13c).

The lake breezes (A, C) advect moist (thick black contours), low  $\theta_v$  (green/blue shading) air away from the lake surface and onto the land (Fig. 13). In the dry period and DS storm case studies, the air over land is much drier than over the lake, creating a sharp gradient in moisture along the lake-breeze fronts (Figs. 13a,c). Air over land is already moist in the LR case study, so the gradient is weaker (Fig. 13b). There are no strong gradients in  $\theta_v$ , since the cooler lake air warms as it propagates over land, eroding the front of the density current (Fig. 13).

In all three cases, the lake breeze across the eastern shore converges with the prevailing southeasterly winds, which are particularly strong on the slopes of

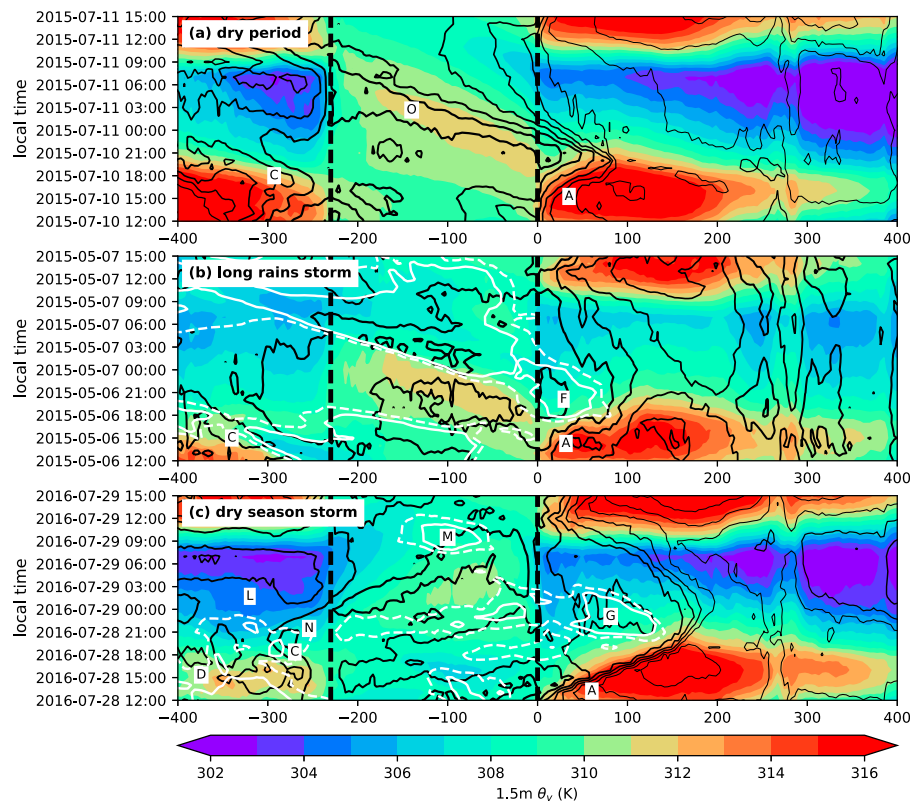


FIG. 13. Hovmöller of simulated 1.5 m virtual potential temperature  $\theta_v$ , averaged over the transects in Fig. 1b for (a) the dry period, (b) the LR storm, and (c) the DS storm. Black contours give 1.5 m specific humidity with intervals of  $1 \text{ g kg}^{-1}$  (thicker lines show higher specific humidity) and white contours show regions of OLR less than 200 (dashed) and 170 (solid)  $\text{W m}^{-2}$ . The black dashed lines show the boundaries of Lake Victoria. Labels A–N are described in the text and match those in Fig. 12.

the Kenyan Highlands E (Fig. 12). Upward motion occurs along the convergence (orange contour between 0 and  $\sim +100 \text{ km}$ , Fig. 12). In the two storm cases, the upward motion results in deep convection (low OLR, shown by blue/white contours between 0 and  $+100 \text{ km}$ , Figs. 12b,c and 13b,c). The LR storm is labeled F and the land storm during the DS storm case study is labeled G. Whereas the air is already fairly moist over land during the LR storm, the moisture source for the storm over land in the dry season is air advected from the lake by the lake breeze (Figs. 13b,c).

After sunset, air over the land begins to cool, but air over the lake continues to warm in all cases, as the lake retains its heat for longer (Fig. 13). Although the timing differs slightly between cases, a land breeze H begins to form close to the eastern shoreline around 1900–2000 LT (Fig. 12). Consequently, the leading edge I of the lake breeze detaches from the main body (Fig. 12). Flow becomes offshore at 2100 LT during the dry period and DS storm and just after 1900 LT during the LR storm. In all cases, the prevailing southeasterlies surge westward and merge with the land breeze.

Driven by the southeasterly flow, the convergence and associated upward motion propagate across the lake during the dry period and LR storm case studies (Figs. 12a,b, J). In the latter case, the leading edge of the storm propagates along with the convergence. The propagation speeds across the lake are  $\sim 10 \text{ m s}^{-1}$  in the dry period and  $\sim 6.5 \text{ m s}^{-1}$  during the LR storm (note this is also the storm propagation speed). Behind the convergence in the dry period, the wind speed rapidly increases to greater than  $8 \text{ m s}^{-1}$ , despite a lack of storm. The storm G over land during the DS storm case dissipates and does not propagate over the lake. The propagation of convergence J is blocked  $\sim 40 \text{ km}$  offshore by the strong westerlies K across the lake (Fig. 12c), which occur due to the low  $\theta_v$  air to west L (Fig. 13c). A new storm M forms along this land–breeze convergence over the lake around 0800 LT (Figs. 12c and 13c). In addition to the lake moisture source, additional moisture N is advected onto the lake from the land overnight to further fuel this storm (Fig. 13c).

During the dry period, the leading edge of the offshore flow advects the moist, high  $\theta_v$  air accumulated on

the eastern shore back across the lake. To the east of the moisture and  $\theta_v$  maximum O, which marks the center of the moist bulge described in section 3b,  $q$  and  $\theta_v$  decrease as cool, dry air from the east is advected onto the lake by the land-breeze density current. However, there is no sharp decrease in these variables because the air moistens and warms as it moves across the lake. Over time, the horizontal extent of the moist bulge increases as the moisture gradient between the land and the lake weakens further, shown by the increased spacing of the specific humidity contours with time in Fig. 13a.

The LR storm F propagates continuously onto the lake from over land, although there is a temporary reduction in intensity as it crosses the eastern shore of the lake (narrowing of OLR contours across the 0 km dashed line in Figs. 12b and 13b). Once over the lake, the rear of the storm remains almost stationary and the storm grows at its leading edge. In contrast, the storms over the land and lake (G and M respectively) in the DS case arise from separate initiations (distinct minima in OLR contours over land and lake in Figs. 12c and 13c). The storm which forms over the lake during the morning shows little propagation, as the land breeze from the east is prevented from propagating across the lake by the anomalously strong land breeze K from the west (Fig. 12c).

#### 4. Summary

Two case studies of storms over the Lake Victoria basin, and a further case of a dry period, were simulated using a convection-permitting configuration of the MetUM with 1.5 km horizontal grid spacing. This is the first time that the role—and interaction of—the large-scale conditions, lake–land-breeze circulation, and mountain flow in convergence and convection initiation/propagation over the Lake Victoria basin have been described and analyzed in such detail.

Key characteristics of each case study and their differences are summarized in Table 2 and the schematics in Fig. 14 (dry period) and Fig. 15 (LR and DS storms). In general, the large-scale circulation patterns are very similar between all three cases, as shown in Fig. 5; southeasterlies prevail across the region, except across Lake Victoria and to the northwest during the DS storm. The mean wind speed over the lowest 5 km does not vary much between cases, but is smallest during the dry period ( $5.1 \text{ m s}^{-1}$ , compared with  $5.7$  and  $5.5 \text{ m s}^{-1}$  in the LR and DS storms respectively, first column, Table 2). However, the wind speed is actually greatest at the surface during the dry period (not shown). While the large-scale circulation is very similar between cases, Fig. 5 clearly shows large-scale differences in near-surface

moisture between the cases. Mean specific humidity in the lowest 5 km over the region is  $\sim 3 \text{ g kg}^{-1}$  greater during the LR case study than the dry period case (fifth column, Table 2). Climatologically this is expected, as the mean monthly  $q$  is  $1.1 \text{ g kg}^{-1}$  higher in May compared to July (fourth column, Table 2). Specific humidity is greater during the DS storm case than the dry period case, since additional moisture is advected into the region from the Congo basin by the anomalous northwesterly wind. Table 2 (second and third columns) shows that mean potential temperature over the lowest 5 km does not change much between cases. However, the temperature in the DS storm case is lower than the dry period, likely because of the cold air associated with the storm to the northwest of the lake.

The LR storm initiates over land off the eastern shore of the lake during the afternoon, persists for  $\sim 21$  h and has a lower mean OLR, whereas the DS storm initiates in the early morning over the lake itself, persists for about half as long and has a lower amount of cold, high cloud tops (8th, 9th, 10th, and 16th columns, Table 2). Mean accumulated rainfall over the storms' lifetime is much larger in the LR storm, when accumulations are averaged over the whole lake or just over the storm area (14th and 15th columns, Table 2). This is due to a combination of a larger storm area, longer storm duration and more intense rainfall rates.

Figures 14a,d and 15a,d show the strong similarities in the eastern lake breeze (blue arrows) and its evening convergence with the easterly prevailing winds, which are strengthened as they flow down the slopes of the Kenyan Highlands (red arrows), in the dry period and LR storm case studies. The reason that a storm forms along the convergence in the latter case is the moisture-laden air and high CAPE over land to the east of Lake Victoria, which allows deep convection to be triggered. Initiation is prevented in the dry period case by high CIN over the region. In both case studies, there is also subsequent overnight propagation of the convergence line over the lake as the easterly winds (including a land breeze and katabatic component) accelerate. The LR storm propagates along with the convergence line (Figs. 14b,c,e,f and 15b,c,e,f).

In the DS storm case, the surface winds across the lake are dominated by an anomalously strong land breeze from the northwest (Figs. 15g–i, purple arrows). This breeze originates from a region of particularly cool, moist air to the northwest of the lake (purple ellipses), remnants of an intense storm on the preceding day. Consequently, the overnight temperature gradient between the lake and land is strengthened in the northwest, forming a strong land breeze from this direction. The northwesterly land breeze converges with



TABLE 2. A comparison of various dynamic and thermodynamic variables and descriptive factors for the three case studies. Wind speeds,  $\theta$ , and  $q$  are averaged over the lowest 5 km of the atmosphere above the region in Fig. 1b, and over a 24 h period from 1200 LT 10 Jul 2015, 6 May 2015, and 28 Jul 2016 for the dry period, LR storm, and DS storm, respectively. Where a monthly mean is shown, this is given by the ERA-Interim mean between 1997 and 2016. CAPE and CIN are the mean values from the region and times in Fig. 8, computed over lake/land grid points as indicated. CIN is only calculated where CAPE > 400 J kg<sup>-1</sup>. Accumulated rainfall and mean OLR for the duration of the storm are computed over a box containing Lake Victoria and mean rainfall is also computed only over the storm area (where rainfall exceeds 0.1 mm h<sup>-1</sup>). Other relevant factors are also noted.

Case	(1) Wind speed (ms <sup>-1</sup> ) + direction	(2) $\theta$ (K) (ERA-Interim monthly mean)	(3) $\theta$ (K) (model 24 h mean)	(4) $q$ (g kg <sup>-1</sup> ) (ERA-Interim monthly mean)	(5) $q$ (g kg <sup>-1</sup> ) (model 24 h mean)	(6) CAPE/CIN over land at 1800 LT (J kg <sup>-1</sup> )	(7) CAPE/CIN over lake at 0600 LT (J kg <sup>-1</sup> )	(8) Location of storm initiation	(9) Time of storm initiation (LT)	(10) Storm duration	(11) Reason for low-level convergence	(12) Moisture source	(13) Mountain flow	(14) Accumulated rainfall over lake over storm life (mm)	(15) Accumulated rainfall over storm area over storm life (mm)	(16) Mean OLR over storm life (W m <sup>-2</sup> )
Dry period	5.1 (SE)	314.5	316.0	5.9	4.9	CAPE: 140 CIN: 111	CAPE: 163 CIN: 21	Eastern shore of lake	—	—	Convergence between lake breeze and easterly prevailing winds	—	Downslope	—	—	—
LR storm	5.7 (SE)	315.6	315.8	7.0	7.8	CAPE: 318 CIN: 53	CAPE: 642 CIN: 18	Eastern shore of lake	1530	21	Convergence between lake breeze and easterly prevailing winds	Land initially, lake once storm is over lake	Downslope	14.3	95.5	217.1
DS storm	5.5 (SE, NW over lake)	314.5	314.9	5.9	6.1	CAPE: 226 CIN: 50	CAPE: 1051 CIN: 39	Center of lake	0530	10	Convergence between east and west land breezes. A particularly strong land breeze from the west is caused by cold air to the NW associated with convection from previous day	Lake and northwest	Downslope	1.17	41.3	251.1

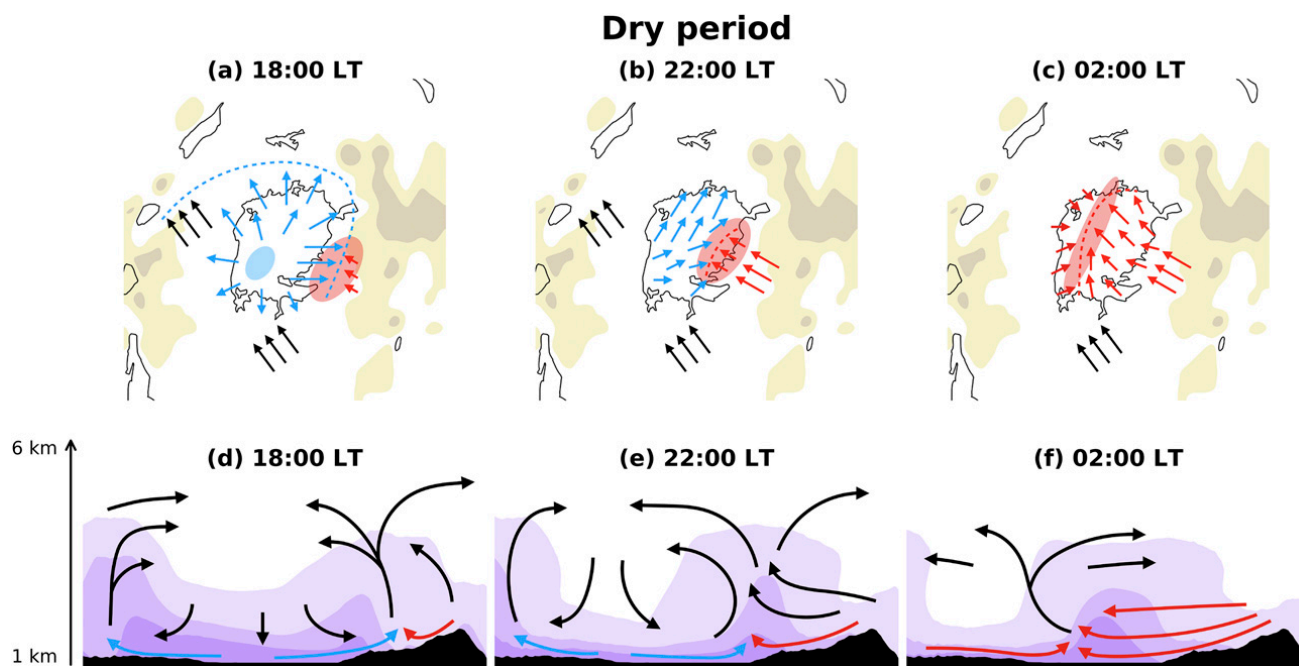


FIG. 14. Schematics showing maps and cross sections of the low-level wind, convergence, and moisture over Lake Victoria during the dry period case study. The blue dashed lines represent the lake-breeze front and the red dashed lines represent the leading edge of the offshore flow. Black arrows on the maps represent the prevailing wind, but represent the flow of air from mixed origins on the cross sections. Blue arrows represent lake breezes and red arrows represent offshore flow (encompassing land breezes, downslope flows, and prevailing winds). Blue and red ellipses represent areas of strong divergence and convergence, respectively. Purple contours on the cross sections show specific humidity.

the offshore flow from the east (red arrows) over the center of the lake at 0600 LT and causes a bulge of moist air over the center of the lake (Figs. 15h,k). High CAPE over the lake results in deep convection. Westward propagation of the convergence line (and storm) across the lake is blocked by the strong northwesterlies.

The moisture to initiate the LR storm originates mainly over land, but the moisture source changes to lake BL air once the storm propagates over the lake (black arrows, purple contours, Figs. 15d–f). In contrast, moist lake BL air is predominantly responsible for the initiation of the DS storm, with the advection of additional moisture onto the lake from the northwest (purple arrows and contours, Figs. 15k,l).

## 5. Discussion and conclusions

The lake–land–breeze circulation was shown to be the major control on the initiation location, timing and propagation of convection over the Lake Victoria basin, affirming its importance for forecasting convection in this region. Many aspects of the case studies—in particular the diurnal cycle in convergence and divergence over Lake Victoria—agree with the mean picture described by Mukabana and Piekle (1996), Ba and Nicholson (1998), Song et al. (2004), Anyah et al. (2006),

Thiery et al. (2015), and Camberlin et al. (2018). However, the differences between the cases demonstrate the importance of daily to seasonal variability in local winds and large-scale moisture availability in storm initiation and propagation, and show that individual storms cannot be understood using the mean diurnal cycle alone.

The lake–land–breeze circulations were remarkably similar during the dry period and LR case study. A moist and high-CAPE/low-CIN atmosphere was shown to be the important factor for triggering a storm in the LR case. The additional moisture was shown to be, at least in part, a seasonal effect. However, moisture content may also vary locally on subseasonal time scales.

The LR storm forms over land to the east of Lake Victoria and subsequently propagates onto the lake. Camberlin et al. (2018) used a mean diurnal cycle of satellite observations to suggest that such a propagation may occur. However, their study was not conclusive as to whether the mean rainfall over the land and lake result from the same storms, due to a discontinuity in propagation speeds over the Kenyan Highlands (slower) and lake (faster). The LR case shows that continuous propagation from land to lake may indeed occur with a discontinuity between phase speeds over the different surfaces; the evening convection to the east of the lake is

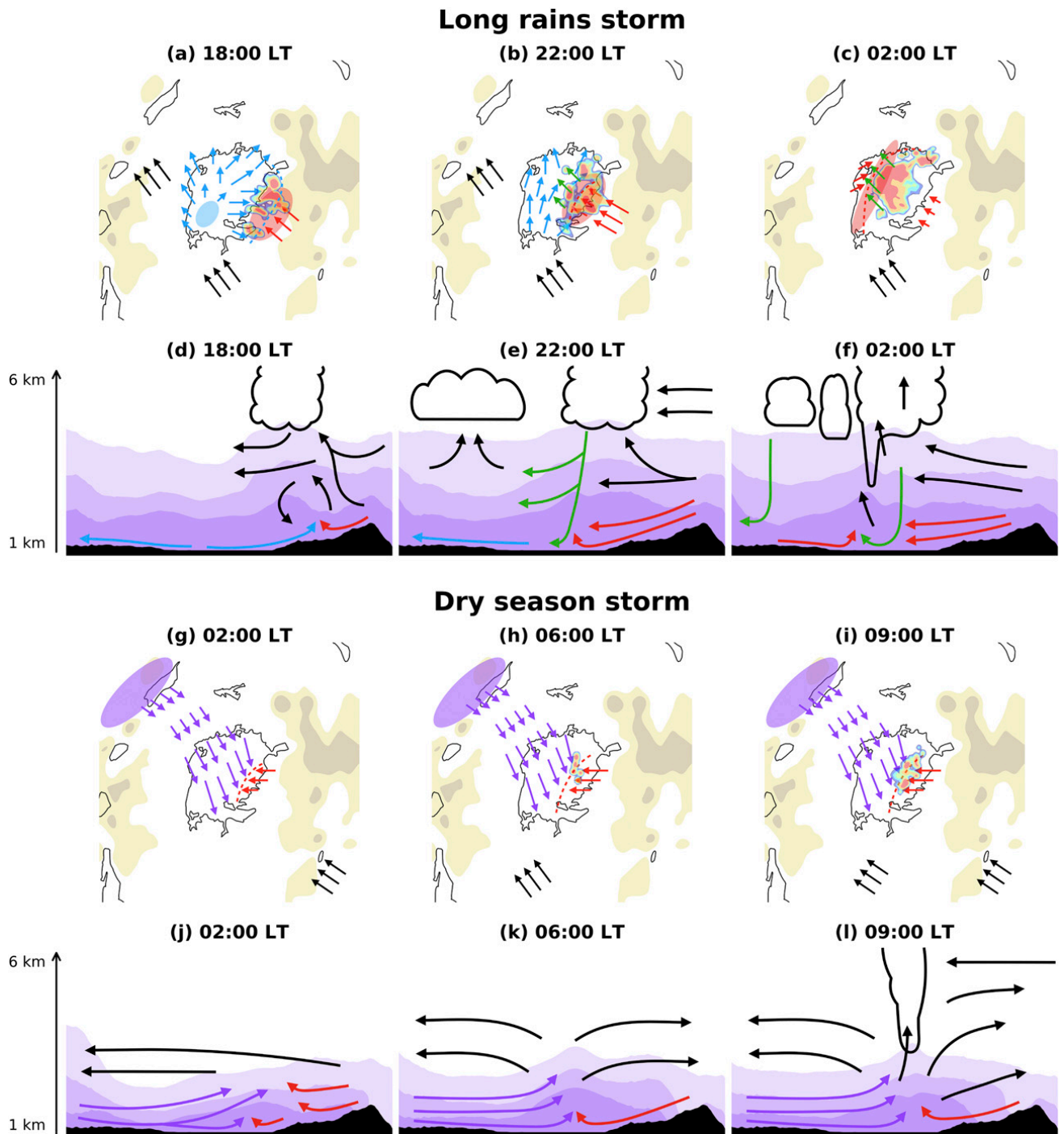


FIG. 15. Schematics showing maps and cross sections of the low-level wind, convergence, and moisture over Lake Victoria during the LR and DS storm case studies. Symbols as in Fig. 14. Green arrows represent storm outflow and purple arrows represent the anomalously strong land breeze from the northwest. Purple ellipses represent regions of cool, moist air. The smoothed precipitation field is also plotted on the maps.

effectively stalled until it moves onto the lake overnight, after which its propagation speed increases.

The DS storm demonstrates the importance of convection from the previous day in altering the local atmosphere—in this case a long-lived cold pool to the northwest of the lake strengthening the land breeze

from the west—and creating an environment conducive to storms over the lake. This result supports a study by Thiery et al. (2016), which noted a strong correlation between intense storms over land during the afternoon and intense nocturnal storms over Lake Victoria, and in part attributed this to long-lived cold

pools. Thiery et al. (2017) used this correlation in a statistical storm predictor [Lake Victoria Intense Early Warning System (VIEWS)] for short-range forecasting. For longer-range forecasts, skill could be enhanced by improving cold pool simulation in dynamical models.

Analysis of the two dry season cases has highlighted the formation of a bulge of moist air above the convergence at the surface of Lake Victoria. Such a feature has not been described in previous literature on collisions of density currents, partly because few studies focus on collisions over a water body, and partly because humidity is neglected in such studies. In the DS storm case, CAPE is high and CIN is low, so a storm is initiated above the moist bulge. There is no clear moist bulge during the LR storm, likely because there is less contrast in specific humidity between air near the surface over the lake and land and because strong circulations associated with the storms may disrupt its development.

Birch and Reeder (2013) and Birch et al. (2014) showed that the collision of offshore flow with a sea breeze can result in waves being excited and traveling upstream. Studies of density current collisions over land also show the excitation of waves (Kingsmill and Crook 2003; Goler and Reeder 2004). While no such waves were diagnosed in this study (the model data hourly time resolution was insufficient), it should not be ruled out that such waves may exist and may be responsible for upstream triggering of convection in some conditions.

During the dry period, sudden strong winds were associated with density current propagation from the east, despite the lack of storm. This suggests that the lake could pose a hazard to fishermen, even on a dry day. In general, the westward offshore propagation requires further study to elucidate the relative strengths of the land breeze, downslope flow and prevailing winds, to better understand the moisture bulge and assess potential hazards.

It is acknowledged that the conclusions drawn in this study are based solely on model simulations. This is due to the lack of high-resolution observations, in particular of low-level wind and moisture fields in this region. Moreover, the case studies presented here show just two examples of storms which occurred over the Lake Victoria basin, and their representativeness is unknown. In particular, it is acknowledged that the large-scale circulation pattern does not differ much between cases, and so its effect cannot be investigated in this study.

Given the very different characteristics of the two storms—and their correspondingly different triggers—it is very likely that other storm types and triggers also exist. It would be a natural extension of this work to classify storms over the lake—for example based on their time and location of initiation—and identify corresponding triggers in the local- and large-scale atmosphere. That

the storms analyzed in this study occurred in different seasons raises the question as to whether certain storm types and triggers are more likely in a given season.

Understanding the key processes involved in storm formation and propagation is necessary to assess their representation in forecast models, and therefore understand model behavior and make improvements where necessary. Ultimately, improved forecasts will lead to increased safety and reduced deaths on the lake.

*Acknowledgments.* This work was supported by the NERC SPHERES DTP (Grant NE/L002574/1). Woodhams was the grateful recipient of the Australian Bicentennial Award from the Menzies Centre at King's College London and the Rupert Ford Award from the Royal Meteorological Society to fund travel to the University of Melbourne. Marsham was also funded by the HyCRISTAL project (Grant NE/M02038X/1) and the NCAS ACREW project. Lane is supported by the Australian Research Council's Centres of Excellence Scheme (CE170100023). This work was also supported by U.K. Research and Innovation as part of the Global Challenges Research Fund, Grant NE/P021077/1 (GCRF African SWIFT). This project used the MetPy package developed by UCAR/Unidata (Unidata 2018). The GPM IMERG data were provided by the NASA Goddard Space Flight Center's Precipitation Measurement Missions Science Team and Precipitation Processing System, which develop and compute the GPM IMERG as a contribution to GPM, including data archiving at the NASA GES DISC. Brightness temperature data were provided by EUMETSAT and distributed by the SATMOS service (Météo-France/Center de Meteorologie Spatiale). We thank the ICARE Data and Services Center for providing access to this data. The authors thank Alexander Roberts for regridding the brightness temperature data and Martin Jucker for code to compute model level heights. The authors would also like to thank the anonymous reviewers for improving the quality and clarity of the paper.

#### REFERENCES

- Albrecht, R. I., S. J. Goodman, D. E. Buechler, R. J. Blakeslee, and H. J. Christian, 2016: Where are the lightning hotspots on Earth? *Bull. Amer. Meteor. Soc.*, **97**, 2051–2068, <https://doi.org/10.1175/BAMS-D-14-00193.1>.
- Anyah, R. O., F. H. Semazzi, and L. Xie, 2006: Simulated physical mechanisms associated with climate variability over Lake Victoria Basin in East Africa. *Mon. Wea. Rev.*, **134**, 3588–3609, <https://doi.org/10.1175/MWR3266.1>.
- Aranami, K., T. Davies, and N. Wood, 2015: A mass restoration scheme for limited-area models with semi-Lagrangian advection. *Quart. J. Roy. Meteor. Soc.*, **141**, 1795–1803, <https://doi.org/10.1002/qj.2482>.

- Ba, M. B., and S. E. Nicholson, 1998: Analysis of convective activity and its relationship to the rainfall over the Rift Valley lakes of East Africa during 1983–90 using the Meteosat infrared channel. *J. Appl. Meteor.*, **37**, 1250–1264, [https://doi.org/10.1175/1520-0450\(1998\)037<1250:AOCAAI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1998)037<1250:AOCAAI>2.0.CO;2).
- Ballentine, R. J., 1982: Numerical simulation of land-breeze-induced snowbands along the western shore of Lake Michigan. *Mon. Wea. Rev.*, **110**, 1544–1553, [https://doi.org/10.1175/1520-0493\(1982\)110<1544:NSOLBI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<1544:NSOLBI>2.0.CO;2).
- Birch, C. E., and M. J. Reeder, 2013: Wave-cloud lines over northwest Australia. *Quart. J. Roy. Meteor. Soc.*, **139**, 1311–1326, <https://doi.org/10.1002/qj.2043>.
- , —, and G. Berry, 2014: Wave-cloud lines over the Arabian Sea. *J. Geophys. Res. Atmos.*, **119**, 4447–4457, <https://doi.org/10.1002/2013JD021347>.
- , M. J. Roberts, L. Garcia-Carreras, D. Ackerley, M. J. Reeder, A. P. Lock, and R. Schiemann, 2015: Sea-breeze dynamics and convection initiation: The influence of convective parameterization in weather and climate model biases. *J. Climate*, **28**, 8093–8108, <https://doi.org/10.1175/JCLI-D-14-00850.1>.
- Camberlin, P., W. Gitau, O. Planchon, V. Dubreuil, B. M. Funatsu, and N. Philippon, 2018: Major role of water bodies on diurnal precipitation regimes in Eastern Africa. *Int. J. Climatol.*, **38**, 613–629, <https://doi.org/10.1002/joc.5197>.
- Cannon, T., F. Krüger, G. Bankoff, and L. Schipper, 2014: Putting culture at the centre of risk reduction. *World Disasters Report 2014: Focus on Culture and Risk*, T. Cannon and L. Schipper, Eds., International Federation of Red Cross and Red Crescent Societies, 185–209.
- Chamberlain, J., C. Bain, D. Boyd, K. McCourt, T. Butcher, and S. Palmer, 2014: Forecasting storms over Lake Victoria using a high resolution model. *Meteor. Appl.*, **21**, 419–430, <https://doi.org/10.1002/met.1403>.
- Datta, R., 1981: Certain aspects of monsoonal precipitation dynamics over Lake Victoria. *Monsoon Dynamics*, J. Lighthill and R. Pearce, Eds., Cambridge University Press, 333–349.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- EUMETSAT, 2012: The conversion from effective radiances to equivalent brightness temperatures. EUM/MET/TEN/11/0569. v1. Tech. Rep., EUMETSAT, 49 pp.
- Fiedler, E. K., M. J. Martin, and J. Roberts-Jones, 2014: An operational analysis of lake surface water temperature. *Tellus*, **66A**, 21247, <https://doi.org/10.3402/tellusa.v66.21247>.
- Finney, D. L., and Coauthors, 2019: Implications of improved representation of convection for the East Africa water budget using a convection-permitting model. *J. Climate*, **32**, 2109–2129, <https://doi.org/10.1175/JCLI-D-18-0387.1>.
- Flohn, H., and K. Fraedrich, 1966: Tagesperiodische zirkulation und niederschlagsverteilung am Victoria-See (Ostafrika) (The daily periodic circulation and distribution of rainfall over Lake Victoria, in German). *Meteor. Rundsch.*, **19** (6), 157–165.
- , and T. Burkhardt, 1985: Nile runoff at Aswan and Lake Victoria: A case of a discontinuous climate time series. *Z. Gletscherk. Glazialgeol.*, **21**, 125–130.
- Fraedrich, K., 1972: A simple climatological model of the dynamics and energetics of the nocturnal circulation at Lake Victoria. *Quart. J. Roy. Meteor. Soc.*, **98**, 322–335, <https://doi.org/10.1002/qj.49709841606>.
- Goler, R. A., and M. J. Reeder, 2004: The generation of the morning glory. *J. Atmos. Sci.*, **61**, 1360–1376, [https://doi.org/10.1175/1520-0469\(2004\)061<1360:TGOTMG>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1360:TGOTMG>2.0.CO;2).
- Gregory, D., and P. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev.*, **118**, 1483–1506, [https://doi.org/10.1175/1520-0493\(1990\)118<1483:AMFCSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2).
- Hastings, D. A., and P. K. Dunbar, 1999: Global Land One-kilometer Base Elevation (GLOBE) digital elevation model, documentation, volume 1.0. Key to Geophysical Records Documentation (KGRD), Doc. 34, National Oceanic and Atmospheric Administration, National Geophysical Data Center, <https://repository.library.noaa.gov/view/noaa/13424>.
- Hjelmfelt, M. R., and R. R. Braham Jr., 1983: Numerical simulation of the airflow over Lake Michigan for a major lake-effect snow event. *Mon. Wea. Rev.*, **111**, 205–219, [https://doi.org/10.1175/1520-0493\(1983\)111<0205:NSOTAO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<0205:NSOTAO>2.0.CO;2).
- Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteor. Climatol.*, **43**, 1834–1853, <https://doi.org/10.1175/JAM2173.1>.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Huffman, G. J., 2017: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree × 0.1 degree V05. Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, accessed 20 December 2017, <https://doi.org/10.5067/GPM/IMERG/3B-HH/05>.
- , and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , and Coauthors, 2018: Algorithm Theoretical Basis Document (ATBD), version 5.2. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Tech. Rep., 35 pp., [https://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V5.2.pdf](https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2.pdf).
- Joyce, R. J., and P. Xie, 2011: Kalman filter-based CMORPH. *J. Hydrometeorol.*, **12**, 1547–1563, <https://doi.org/10.1175/JHM-D-11-022.1>.
- , J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503, [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Keen, C. S., and W. A. Lyons, 1978: Lake/land breeze circulations on the western shore of Lake Michigan. *J. Appl. Meteor.*, **17**, 1843–1855, [https://doi.org/10.1175/1520-0450\(1978\)017<1843:LBCOTW>2.0.CO;2](https://doi.org/10.1175/1520-0450(1978)017<1843:LBCOTW>2.0.CO;2).
- Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts, 2012: Realism of rainfall in a very high-resolution regional climate model. *J. Climate*, **25**, 5791–5806, <https://doi.org/10.1175/JCLI-D-11-00562.1>.
- Kim, K., J. Park, J. Baik, and M. Choi, 2017: Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.*, **187**, 95–105, <https://doi.org/10.1016/j.atmosres.2016.12.007>.
- Kingsmill, D. E., and A. N. Crook, 2003: An observational study of atmospheric bore formation from colliding density currents. *Mon. Wea. Rev.*, **131**, 2985–3002, [https://doi.org/10.1175/1520-0493\(2003\)131<2985:AOSOAB>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<2985:AOSOAB>2.0.CO;2).
- Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions

- of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.*, **136**, 3408–3424, <https://doi.org/10.1175/2008MWR2332.1>.
- Lumb, F., 1970: Topographic influences on thunderstorm activity near Lake Victoria. *Weather*, **25**, 404–410, <https://doi.org/10.1002/j.1477-8696.1970.tb04129.x>.
- Lyons, W. A., 1972: The climatology and prediction of the Chicago lake breeze. *J. Appl. Meteor.*, **11**, 1259–1270, [https://doi.org/10.1175/1520-0450\(1972\)011<1259:TCAPOT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1972)011<1259:TCAPOT>2.0.CO;2).
- , and L. E. Olsson, 1972: Mesoscale air pollution transport in the Chicago lake breeze. *J. Air Pollut. Control Assoc.*, **22**, 876–881, <https://doi.org/10.1080/00022470.1972.10469725>.
- Marshall, J. H., N. S. Dixon, L. Garcia-Carreras, G. Lister, D. J. Parker, P. Knippertz, and C. E. Birch, 2013: The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophys. Res. Lett.*, **40**, 1843–1849, <https://doi.org/10.1002/grl.50347>.
- Mukabana, J. R., and R. A. Pielke, 1996: Investigating the influence of synoptic-scale monsoonal winds and mesoscale circulations on diurnal weather patterns over Kenya using a mesoscale numerical model. *Mon. Wea. Rev.*, **124**, 224–244, [https://doi.org/10.1175/1520-0493\(1996\)124<0224:ITIOSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<0224:ITIOSS>2.0.CO;2).
- Nicholson, S., 1996: A review of climate dynamics and climate variability in Eastern Africa. *The Limnology, Climatology and Paleoclimatology of the East African Lakes*, T. C. Johnson and E. O. Odada, Eds., Gordon and Breach Publishers, 25–56.
- , and X. Yin, 2002: Mesoscale patterns of rainfall, cloudiness and evaporation over the great lakes of East Africa. *The East African Great Lakes: Limnology, Palaeolimnology and Biodiversity*, E. O. Odada and D. O. Olago, Eds., Kluwer Academic Publishers, 93–120.
- Okeyo, A. E., 1986: The impact of Lake Victoria on the convective activities over the Kenya Highlands. *J. Meteor. Soc. Japan*, **64**, 689–695.
- Passarelli, R. E., Jr., and R. R. Braham Jr., 1981: The role of the winter land breeze in the formation of Great Lake snow storms. *Bull. Amer. Meteor. Soc.*, **62**, 482–491, [https://doi.org/10.1175/1520-0477\(1981\)062<0482:TROTWL>2.0.CO;2](https://doi.org/10.1175/1520-0477(1981)062<0482:TROTWL>2.0.CO;2).
- Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bull. Amer. Meteor. Soc.*, **83**, 977–992, [https://doi.org/10.1175/1520-0477\(2002\)083<0977:AITMSG>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2).
- Schneider, U., A. Becker, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2008: Global precipitation analysis products of the GPCC. Tech. Rep., Deutscher Wetterdienst, 12 pp., [ftp://ds1.iap.ac.cn/ftp/ds088\\_GPCC5-full\\_0.5\\_1month\\_ascii/GPCC\\_intro\\_products\\_2008.pdf](ftp://ds1.iap.ac.cn/ftp/ds088_GPCC5-full_0.5_1month_ascii/GPCC_intro_products_2008.pdf).
- Semazzi, F., 2011: Enhancing safety of navigation and efficient exploitation of natural resources over Lake Victoria and its basin by strengthening meteorological services on the lake. Tech. Rep., North Carolina State University Climate Modeling Laboratory, 104 pp.
- Sills, D., J. Brook, I. Levy, P. Makar, J. Zhang, and P. Taylor, 2011: Lake breezes in the southern Great Lakes region and their influence during BAQS-Met 2007. *Atmos. Chem. Phys.*, **11**, 7955–7973, <https://doi.org/10.5194/acp-11-7955-2011>.
- Song, Y., F. H. Semazzi, L. Xie, and L. J. Ogallo, 2004: A coupled regional climate model for the Lake Victoria basin of East Africa. *Int. J. Climatol.*, **24**, 57–75, <https://doi.org/10.1002/joc.983>.
- Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, [https://doi.org/10.1175/1520-0477\(2000\)081<2035:EOPSS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSS>2.3.CO;2).
- Stratton, R. A., and Coauthors, 2018: A Pan-African Convection-Permitting Regional Climate Simulation with the Met Office Unified Model: CP4-Africa. *J. Climate*, **31**, 3485–3508, <https://doi.org/10.1175/JCLI-D-17-0503.1>.
- Sungmin, O., and P.-E. Kirstetter, 2018: Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. *Quart. J. Roy. Meteor. Soc.*, **144**, 270–281, <https://doi.org/10.1002/qj.3218>.
- , U. Foelsche, G. Kirchengast, J. Fuchsberger, T. Jackson, and W. A. Petersen, 2017: Evaluation of GPM IMERG early, late, and final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrol. Earth Syst. Sci.*, **21**, 6559–6572, <https://doi.org/10.5194/hess-21-6559-2017>.
- Tang, G., Y. Ma, D. Long, L. Zhong, and Y. Hong, 2016: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *J. Hydrol.*, **533**, 152–167, <https://doi.org/10.1016/j.jhydrol.2015.12.008>.
- Tang, Y., H. W. Lean, and J. Bornemann, 2013: The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteor. Appl.*, **20**, 417–426, <https://doi.org/10.1002/met.1300>.
- Thiery, W., E. L. Davin, H.-J. Panitz, M. Demuzere, S. Lhermitte, and N. Van Lipzig, 2015: The impact of the African Great Lakes on the regional climate. *J. Climate*, **28**, 4061–4085, <https://doi.org/10.1175/JCLI-D-14-00565.1>.
- , —, S. I. Seneviratne, K. Bedka, S. Lhermitte, and N. P. van Lipzig, 2016: Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.*, **7**, 12786, <https://doi.org/10.1038/ncomms12786>.
- , L. Gudmundsson, K. Bedka, F. H. Semazzi, S. Lhermitte, P. Willems, N. P. van Lipzig, and S. I. Seneviratne, 2017: Early warnings of hazardous thunderstorms over Lake Victoria. *Environ. Res. Lett.*, **12**, 074012, <https://doi.org/10.1088/1748-9326/aa7521>.
- Unidata, 2018: MetPy: A Python Package for Meteorological Data, version 0.4.3. UCAR/Unidata, Boulder, CO, accessed 15 March 2019, <https://doi.org/10.5065/D6WW7G29>.
- Virts, K. S., J. M. Wallace, M. L. Hutchins, and R. H. Holzworth, 2013: Highlights of a new ground-based, hourly global lightning climatology. *Bull. Amer. Meteor. Soc.*, **94**, 1381–1391, <https://doi.org/10.1175/BAMS-D-12-00082.1>.
- Wang, R., J. Chen, and X. Wang, 2017: Comparison of IMERG Level-3 and TMPA 3B42V7 in estimating typhoon-related heavy rain. *Water*, **9** (4), 276, <https://doi.org/10.3390/w9040276>.
- Wood, N., and Coauthors, 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quart. J. Roy. Meteor. Soc.*, **140**, 1505–1520, <https://doi.org/10.1002/qj.2235>.
- Woodhams, B., C. E. Birch, J. Marshall, C. Bain, N. Roberts, and D. Boyd, 2018: What is the added value of a convection-permitting model for forecasting extreme rainfall over tropical East Africa? *Mon. Wea. Rev.*, **146**, 2757–2780, <https://doi.org/10.1175/MWR-D-17-0396.1>.
- Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.*, **122**, 910–924, <https://doi.org/10.1002/2016JD025418>.
- Yin, X., S. E. Nicholson, and M. B. Ba, 2000: On the diurnal cycle of cloudiness over Lake Victoria and its influence on evaporation from the lake. *Hydrol. Sci. J.*, **45**, 407–424, <https://doi.org/10.1080/02626660009492338>.

## Erratum

Page numbers refer to the page numbers of the published paper (shown on the top outer corner).

- Note that the shallow cumulus where ‘the lake breeze front penetrates almost 200 km onshore in the west’ which are referred to on p3372, col 2, are not visible as black contours (enclosing cloud liquid water and ice content exceeding  $0.1 \text{ g kg}^{-1}$ ) in Fig. 3.7n because they occur further west than shown in the figure.
- Dynamical triggering refers to processes which mechanically lift air, such as upward motion at convergence (e.g. large-scale convergence lines, sea/lake/land breezes, density currents) and orographic lift (p3377, col 1).





## Chapter 4.

# Aircraft observations of the lake–land breeze circulation over Lake Victoria from the HyVic pilot flight campaign

Prepared for submission to *Quarterly Journal of the Royal Meteorological Society*.

Contributions from: John H. Marsham, Cathryn E. Birch, Caroline L. Bain, Andrew J. Hartley, Paul Barrett, Jennifer K. Fletcher, Stuart Webster and Uganda National Meteorological Authority (UNMA)

### Abstract

The lake–land breeze circulation over Lake Victoria was observed in unprecedented detail using the FAAM BAe-146 aircraft during the HyVic pilot flight campaign. Observations were collected at various heights along a transect over the lake and dropsondes were used to profile the lower troposphere. An evening and morning flight were carried out to observe the lake and land breezes respectively. The lake breeze front was observed to have a horizontal extent of  $\sim 5$  km, suggesting that low-resolution models would struggle to represent this feature. At least one region of elevated moisture, coincident with cloud and increased turbulence, was observed over the lake surface during the early morning, but the mechanism for its formation was unclear. Small-scale variability in wind and moisture was observed across the lake surface, which may have important consequences for the initiation of deep convection. Motivation is provided for a larger field campaign in the future.

## 1. Introduction

Storms and high winds over Lake Victoria are estimated to contribute to 5,000 fatalities on the lake every year (Cannon et al. 2014). An estimated 3.5 million people rely on the lake for their livelihoods, including 200,000 fishermen (Semazzi 2011). The lake also supports transport and trade routes, as well as hydroelectric power. However, forecasting severe weather in East Africa remains a great challenge for numerical weather prediction (NWP) models and operational meteorologists alike. This is despite the relatively recent introduction of convection-permitting (CP) forecast models over East Africa by the UK Met Office (Chamberlain et al. 2014; Woodhams et al. 2018).

The lake–land breeze circulation has long been understood to play a key role in modulating and enhancing convection over the Lake Victoria basin. Flohn and Fraedrich (1966) first noted the existence of a diurnal circulation system, driven by temperature and moisture gradients between the lake and land. Their study linked the early morning maximum of rainfall over the lake to convergence produced by the nocturnal land breeze. Conversely, a divergent lake breeze forms during the day, suppressing convection over the lake, but enhancing convection along the lake breeze front over land (Datta 1981; Ba and Nicholson 1998). The lake and land breezes are also reinforced by anabatic and katabatic flows respectively, especially on the steep slopes of the Kenyan Highlands to the east (Lumb 1970; Okeyo 1986; Mukabana and Pielke 1996; Anyah et al. 2006; Thiery et al. 2015).

Previous studies of Lake Victoria have tended to focus on the mean diurnal cycle, thereby neglecting the impact of daily variability, and smoothing out small-scale details. For the first time, Woodhams et al. (2019) looked at individual case studies of lake–land breeze circulation and storm events over Lake Victoria, using a CP version of the Met Office Unified Model (MetUM) with 1.5 km horizontal grid-spacing. This study noted that afternoon convergence over land to the east of the lake (at the lake breeze front) is pushed back over the lake at night as the lake breeze weakens and the land breeze develops (reinforced by the prevailing easterlies and likely a katabatic component). In other words, the daytime land convergence and nocturnal lake convergence are not two separate features, but caused by a persisting line of convergence which propagates from land to lake. This propagation was also shown to be responsible for the lakeward propagation of a storm which formed to the east of the lake. Woodhams et al. (2019) also noted the formation of a bulge of moisture over the lake, formed along the convergence line as it propagated westward, but reinforced later as a land breeze also formed across the western shore. It was suggested that the properties of this moisture

bulge could determine whether or not a storm formed.

It should be noted that the processes and features described in Woodhams et al. (2019) were based almost entirely on model simulations, given the lack of in-situ observations in the region. Existing in-situ observations of the lake–land breeze circulation have been obtained from weather stations with fixed locations (e.g. Lumb 1970; Datta 1981), with which it is difficult to build a full picture of the circulation. Data from such stations is generally recorded with a maximum frequency of 15 minutes, too low to fully capture the passage of the lake or land breeze fronts. In addition, such stations can only sample the circulation at the surface. Upper-air observations are particularly lacking in the region, and observations over the lake itself present an exceptional challenge. As such, the altitudinal structure of the lake–land breeze and moisture bulge have not been observed. An additional challenge is that many weather stations in the region are owned by private companies or the national meteorological services, and data is not easily accessible to researchers.

For forecasting severe weather over the region, Woodhams et al. (2018) showed that a CP version of the MetUM does add value to a global (parametrised) model forecast, in particular on sub-daily timescales, although skill remains limited. A substantial amount of storms over Lake Victoria are not forecast and false alarm rates are also high. A likely reason for poor model skill in such configurations (likely over the tropics as a whole) is a lack of observations, especially upper air, both for data assimilation and model verification and development.

In January 2019, the HyVic pilot flight campaign took place, using the Facility for Airborne Atmospheric Measurements (FAAM) BAe-146 aircraft to observe the lake–land breeze circulation over Lake Victoria. The campaign consisted of an evening and morning flight, both with a duration of approximately 4 hrs, to sample the lake and land breeze components of the circulation respectively. In particular, these flights aimed to investigate some of the features simulated in Woodhams et al. (2019), such as the moisture bulge; characterise the lake and land breeze fronts; and collect observations to be used for model verification. In this paper, the two flights will be introduced and a detailed description of the observations from the lake and land breeze regimes will be presented. High-resolution CP MetUM simulations were run for the HyVic period and model output is presented alongside the observations.

This short campaign was designed as an add-on to the MOYA campaign based in Entebbe, Uganda (measuring methane over tropical Africa, Tollefson 2019) and to be a pilot for a potential extended campaign in the future. Given the high seasonal and

sub-seasonal variability in moisture availability and circulation in the region (e.g. Yang et al. 2015; Woodhams et al. 2019), it is noted that the two flights presented in this paper cannot be used to draw robust conclusions about the lake–land breeze circulation on all days. However, this novel set of observations can still provide a snapshot of the lake–land breeze circulation in unprecedented detail, and can be used to evaluate high-resolution simulations and inform future field campaigns.

## 2. Methods

### 2.a. Flights

This study was performed using the FAAM BAe-146 aircraft, operating out of Entebbe (white star, Figs. 4.1a,e). The campaign consisted of two flights: one in the evening to observe the lake breeze, and one the following morning to observe the land breeze. It was important that the morning flight directly followed the evening flight; Woodhams et al. (2019) showed that onshore convergence to the east of the lake during the evening propagates across the lake overnight, therefore the same ‘system’ could be sampled in both flights. The flights were timed to sample the mature lake and land breezes, whilst also taking into account constraints on aircraft and crew turnaround between flights, and the minimum safe altitude when flying in the dark. Details of the flights are summarised in table 4.1.

Table 4.1.: A summary of the flights performed as part of the HyVic pilot field campaign.

Flight	Date	Take Off	Landing	Entebbe Sunset/Sunrise
C130	26 Jan 2019	1234 UTC/ 1534 LT	1615 UTC/ 1915 LT	1606 UTC/ 1906 LT
C131	27 Jan 2019	0208 UTC/ 0508 LT	0637 UTC/ 0937 LT	0359 UTC/ 0659 LT

Both flights were based around flying an approximately northwest to southeast transect between Entebbe (on the northwest shore of the lake) and approximately 130 km onshore from the eastern shore in Tanzania (Fig. 4.1). This transect was flown at several altitudes in order to observe the lake and land breezes in two dimensions. During both flights, six sondes were dropped from the highest leg of the transect, to obtain full profiles throughout the lower troposphere. The transect was chosen to be similar

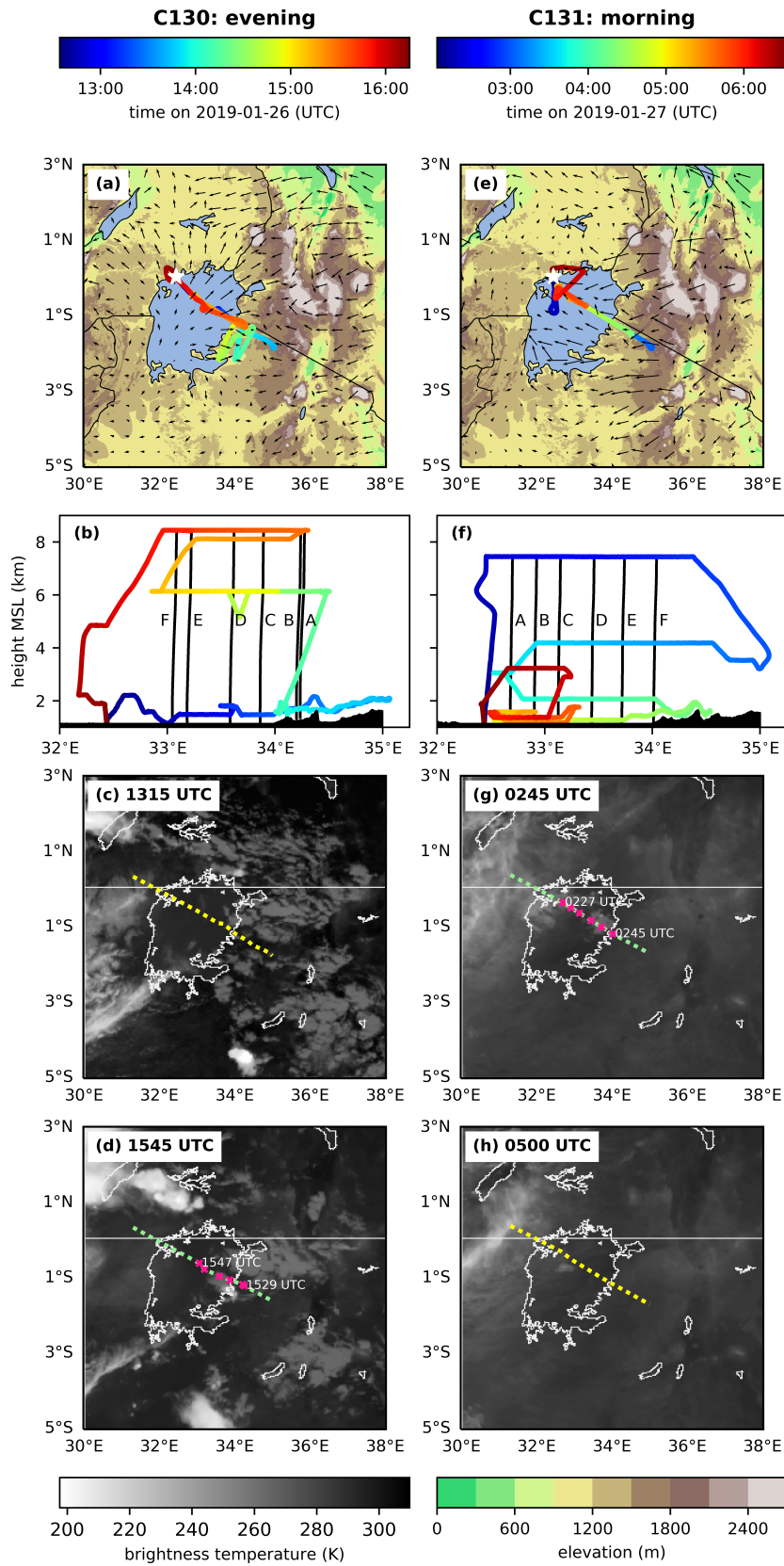


Figure 4.1.: (a, e) Map view of flight track showing terrain (shading) and mean simulated wind (vectors) during the flight time and (b) cross-section view of the flight track for flight C130 (evening) C131 (morning). Colours along the flight track show the time of day. (c, d, g, h) METEOSAT brightness temperatures closest to the times of the sonde drops/runs. Yellow and green dashed lines mark transect along which model cross-sections are computed for flight runs and sonde transects respectively. Pink crosses show where the sondes were dropped, with the time of the first and last drop labelled.

to the transect analysed in Woodhams et al. (2019), whilst also choosing a navigable path over terrain to the east of the lake. Note that the lake sits at 1,135 m above mean sea level (MSL).

### **2.a.1. C130: Evening flight**

A map view and along-transect vertical cross-section view of the evening flight are shown in Figs. 4.1a and b respectively. The flight began with a terrain-following leg at  $\sim 300$  m above ground level (AGL) ( $\sim 1450$  m MSL over the lake) which passed from the lake onto the land, to sample the lake breeze near to the surface. Note that this leg was briefly interrupted over the lake whilst awaiting air traffic control clearance. The low-level leg was followed by a return leg at  $\sim 6000$  m MSL to sample the mid-level return flow. Between these along-transect legs, two legs were flown approximately perpendicular to the transect (parallel to the lake breeze front) at the lower and upper altitudes, but these are not included in the analysis. The aircraft then ascended to  $\sim 8500$  m MSL and dropped six sondes from east to west, including two over land and four over the lake (Figs. 4.1b,d).

### **2.a.2. C131: Morning flight**

The morning flight track is shown in Figs. 4.1e,f. This flight began with the highest leg ( $\sim 7500$  m MSL), along which six sondes were dropped from west to east, the first five over the lake and the final sonde just on the shoreline (Figs. 4.1f,g). In this case, the highest leg was completed first given altitude restrictions in the dark. The aircraft then performed two further legs at  $\sim 4000$  m and  $\sim 2000$  m MSL. As the latter leg reached the shoreline, the aircraft descended to  $\sim 300$  m AGL ( $\sim 1450$  m MSL over the lake), until turning  $180^\circ$   $\sim 75$  km inland and continuing back toward the lake, following the terrain at this height. Once over the lake, the aircraft descended to 150 m AGL ( $\sim 1300$  m MSL) to complete the return leg. From the sonde drops, an approximate location of a moisture bulge was identified and this was then sampled at various heights between 30 and 500 m AGL (1165–1635 m MSL) and with two profiles. Profiles were also performed over the centre of the lake, to compare profiles inside and outside the bulge. A transect perpendicular to the main transect was also flown at  $\sim 250$  m AGL ( $\sim 1400$  m MSL), before recovery to Entebbe.

## 2.b. Aircraft data

Table 4.2.: Aircraft observations used in this paper.

Measured variable	Instrument	Notes
pressure	aircraft RVSM system	(Reduced Vertical Separation Minima)
temperature	non-deiced Rosemount loom temperature sensor	Loom instrument known for drift (personal communication, FAAM). Deiced temperature unavailable at time of submission for comparison.
humidity	Water Vapour Sensing System II (WVSS-II) (Fleming and May 2004), flush inlet	Near-IR tunable diode laser absorption spectrometer. Fast response compared to chilled mirror hygrometers e.g. Buck CR-2 (Vance et al. 2015). WVSS-II is not calibrated, but data were compared to the calibrated Buck and showed good agreement. Flush inlet used due to issues with Rosemount-fed WVSS-II. Although flush inlet samples air from within aircraft boundary layer, Vance et al. (2015) note faster response time and no susceptibility to liquid water or ice with flush inlet compared to Rosemount.
winds	GPS Inertial Navigation (GIN) unit and turbulence probe	Comparisons between wind measurements from different probes show signatures of offset issue with turbulence probe. Issue was raised with FAAM, but not resolved at time of submission. Along-heading winds show better agreement between different probes, so are used in main body of analysis. Wind data processed from the GIN and turbulence probe still use-able and very unlikely to change conclusions of study, but will be updated in future.

Pressure, humidity, temperature and wind measurements were collected on the aircraft itself. Details of the instruments used are presented in Table 4.2. Note that all measurements from the aircraft were sampled at a frequency of 1 Hz. The plane takes measurements at an indicated air speed (IAS) of 200 kts ( $\sim 100 \text{ ms}^{-1}$ ). Although the true air speed varies with altitude, the sampling frequency corresponds to a measurement approximately every 100 m. In addition, the same variables were measured by

Vaisala RD94 dropsondes with a frequency of 2 Hz and transmitted to the Airborne Vertical Atmospheric Profiling System (AVAPS) receiver on board the aircraft. The fall speed of the sonde varied from  $\sim 10\text{--}15\text{ ms}^{-1}$ , therefore measurements were taken every  $\sim 5\text{--}8\text{ m}$ .

### 2.c. MetUM simulations

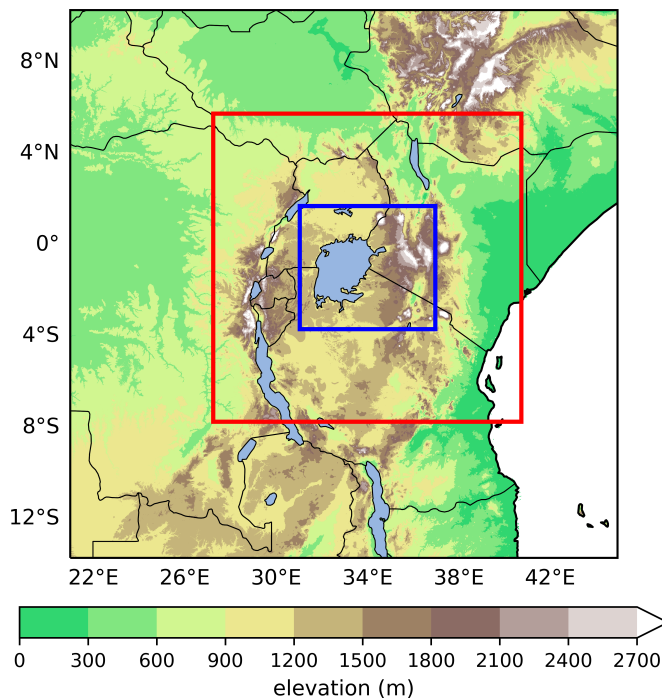


Figure 4.2.: Full domain, red line and blue line correspond to 4.4 km, 1.5 km and 300 m nests respectively for CP MetUM runs during the HyVic period. Elevation data are from the Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model (Hastings and Dunbar 1999).

CP MetUM simulations were run for the campaign period. The model setup was in general the same as that described in Woodhams et al. (2019) (chapter 3), except with the new Regional Atmosphere 1 for the Tropics (RA1T) (Bush et al. 2019) configuration. Simulations were triply one-way nested, with horizontal grid-spacings of 4.4 km, 1.5 km and 300 m (Fig. 4.2). The 4.4 km nest was driven by boundary conditions from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) model. Runs initialised at 2019/01/25 0000 UTC and 2019/01/25 1200 UTC were used to compare to aircraft data from the evening and morning flights respectively.



### 3. Results

#### 3.a. General conditions

In the weeks leading up to the campaign, the Lake Victoria basin received high rainfall totals. Note that periods of convection are not unusual in the dry season, and this enhancement of rainfall was likely due to an active MJO in phases 3 and 4 (Pohl and Camberlin 2006a,b; Hogan et al. 2015) between 14 Jan and 21 Jan, as well as a Kelvin wave passage across the region (Fig. 4.3a). By the time of the campaign, the MJO had moved into phase 6, causing suppression of rainfall activity. However, with high CAPE values over the region, isolated convective events continued to occur over the lake. In particular, significant storms occurred on the three nights preceding the first flight.

During the campaign itself (26–27 Jan), a ridge extended from the Arabian Peninsula across Ethiopia and Somalia (Fig. 4.3b), associated with anomalously strong northeasterly winds at 700 hPa (Figs. 4.3d). These 700 hPa winds were dry (Fig. 4.3f), causing a dry anomaly over much of the Lake Victoria basin (Fig. 4.3g). At the surface, winds were generally north-easterly across the basin, typical of the time of year. At 200 hPa, there was a clockwise upper-level vortex over the Somali coast (Fig. 4.3e), associated with an upper-level ridge (not shown). Over the Lake Victoria basin, 200 hPa winds were generally as expected from the climatology, but there were anomalously strong westerly jets north of 3°N and south of 20°S.

#### 3.b. Lake breeze

Figures 4.4 and 4.6 show data along the transects defined in Figs. 4.1c,d,g,h, which are approximately from the northwest to southeast. For the remainder of the paper, ‘northwesterly’ and ‘southeasterly’ will be used to describe along-transect winds, but the reader should bear in mind that these descriptions are approximate (the bearing of the transect is  $\sim 120^\circ$ ). Figures 4.4a–d show that the lake breeze front at  $\sim 300$  m AGL ( $\sim 2000$  m MSL) was located close to 45 km inland from the eastern shore at approximately 1300 UTC in both the observations and model (this is the closest model output time to the transect flight time). In the observations, the along-transect wind reverses direction across the front, changing from approximately  $+3 \text{ ms}^{-1}$  (northwesterly) to  $-5 \text{ ms}^{-1}$  (southeasterly) across  $\sim 5$  km (Fig. 4.4a). Over the same distance, specific humidity decreases by  $\sim 5 \text{ g kg}^{-1}$  (Fig. 4.4b). The specific humidity continues to decrease at a lower rate ahead of the front, in total falling  $7 \text{ g kg}^{-1}$  over  $\sim 20$  km.

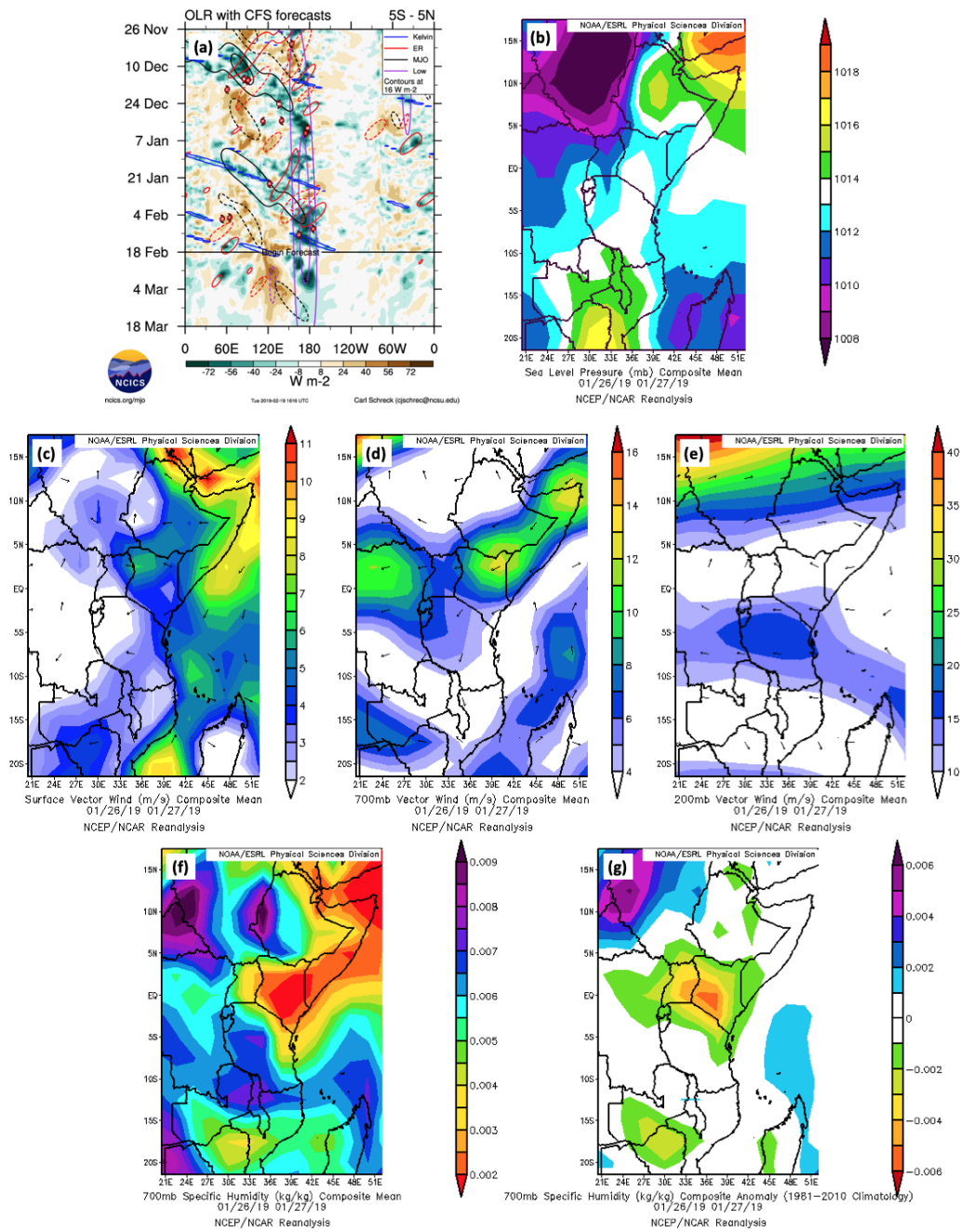


Figure 4.3.: General conditions during HyVic flights. (a) Hövmoller of OLR anomalies and tropical waves approximately two months either side of HyVic, produced by Schreck (n.d.) (<https://ncics.org/portfolio/monitor/mjo/>). Mean (b) sea level pressure, (c) surface wind, (d) 700 hPa wind, (e) 200 hPa wind, (f) 700 hPa specific humidity and (g) 700 hPa specific humidity anomaly from NCEP/NCAR reanalysis (Kalnay et al. 1996) across 26–27 Jan 2019. Images in (b-g) provided by the NOAA/ESRL Physical Sciences Division, Boulder Colorado from their website at <http://www.esrl.noaa.gov/psd/>.

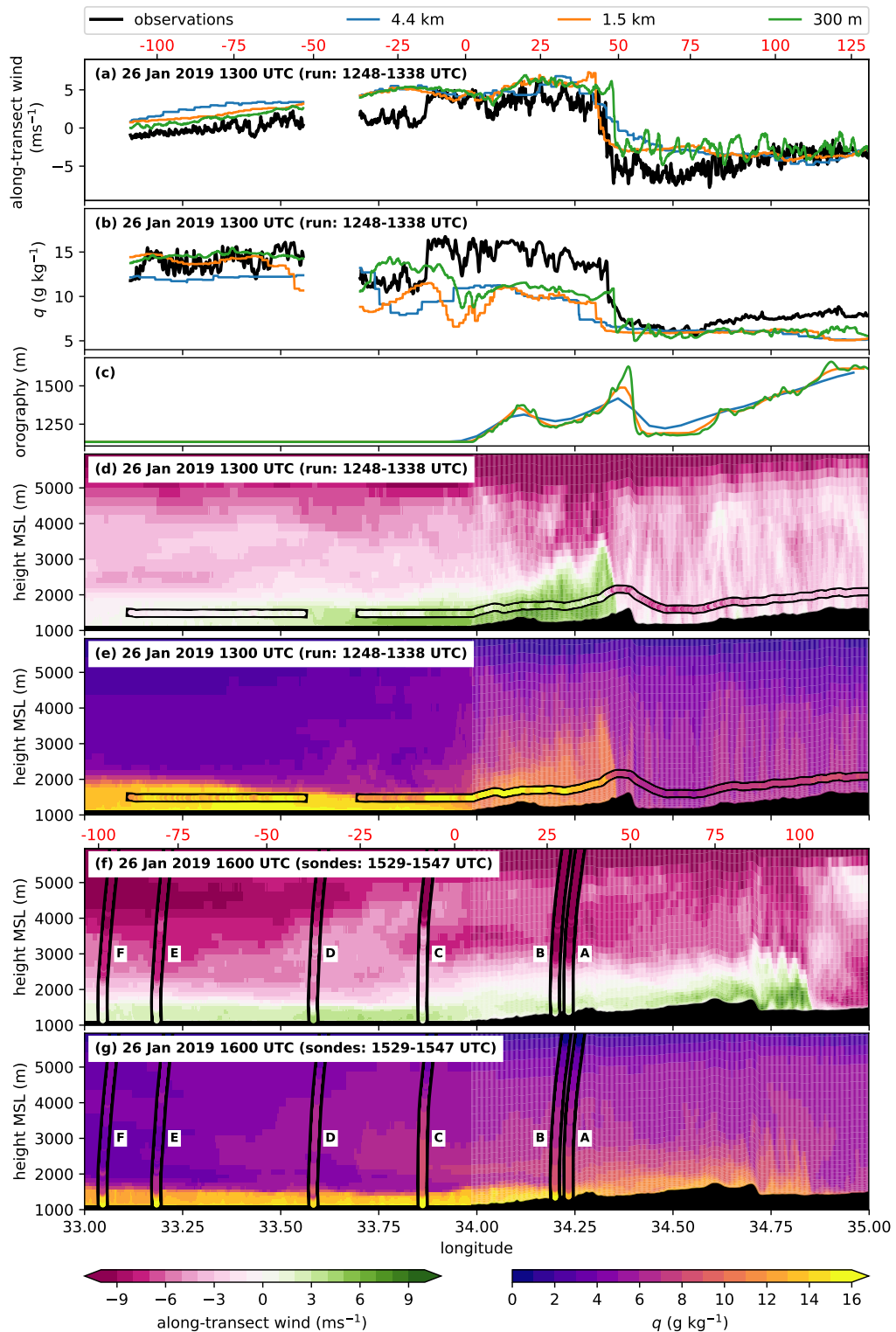


Figure 4.4.: Aircraft observations (black) of (a) along-transect wind and (b) specific humidity along a ~300 m AGL (~1400 m MSL when above the lake) run moving from the lake (northwest) to land (southeast). Simulated variables from CP MetUM with three different resolutions are also plotted (colours), obtained from a virtual fly-through of the model along the aircraft track. (c) Orography from the three model resolutions. Cross-sections of (d),(f) along-transect wind and (e),(g) specific humidity from the (d-e) aircraft and (f-g) sonde observations. The simulated variables (from the 300 m model) from the closest model output time to the observations are plotted underneath. All plots share the same x-axis (longitude) but the distance in km from the eastern shore of the lake (red numbers) differ between (a-e) and (f-g) due to slightly different flight tracks.

The location of the front in the 300 m simulation at 1300 UTC is  $\sim 3$  km further inland than the observations. Given that the aircraft actually crossed the front at 1327 UTC, this suggests that the rate of inland propagation of the front is too high in the model. The greater inland propagation may be explained by the weaker southeasterly prevailing winds ahead of the lake breeze front in the model (Fig. 4.4d), which offer less opposition to the advancement of the lake breeze (Estoque 1962; Simpson et al. 1977; Arritt 1993).

All three model resolutions capture the wind reversal and moisture decrease across the front, but the rate of change across the front is too low in the 4.4 km model, which shows a more gradual change over  $\sim 30$  km (Fig. 4.4a). Milton et al. (2017) suggest that only features with a horizontal extent greater than 7 grid-points can be fully resolved in a model. Given the observed width of the lake breeze front, it is not expected that the 4.4 km could accurately simulate this feature. The representation of orography at different model resolutions may also play a role in the gradient across the front. Figure 4.4c shows how the orographic peak approximately 50 km inland is over 200 m higher in the model with 300 m grid-spacing compared to the model with 4.4 km grid-spacing, and occurs  $\sim 5$  km further inland. The representation of the orography of the Lake Victoria basin has been shown to affect circulation and precipitation over the region (e.g. Mukabana and Pielke 1996; Song et al. 2004; Anyah et al. 2006; Finney et al. 2019), therefore it is suggested that differences in the location, gradient and height of the orography in the model may also be responsible for differences between model resolutions.

Approximately 15 km offshore, the aircraft observations show another sharp front in both wind and moisture. In the direction of the shoreline, specific humidity increases by  $\sim 5$  g kg $^{-1}$ , and the wind increases from nearly stagnant to  $\sim +4$  ms $^{-1}$  (northwesterly). The cause of this front remains unclear, but one suggestion is that the aircraft passed into a region where the lake breeze was deeper. A deeper lake breeze further inland may be the result of the transport of near-surface moisture by turbulent eddies in the convective boundary layer over land (Reible et al. 1993). Deepening may also occur if the air in the lake breeze is decelerated at lower levels due to increased surface friction (as the flow moves onshore), or if the flow is blocked by the orography, causing it to pile up. However, since the model does not capture the apparent front, it is difficult to comment on the mechanism of its formation. Instead of a jump in magnitude, the simulated along-transect winds become gradually stronger between  $\sim 100$  km offshore and the lake breeze front, such that their magnitude is similar to observed at the front (Figs. 4.4a,c). Simulated specific humidity is variable across the lake, but is not

enhanced over the leading  $\sim 60$  km of the lake breeze as observed (Figs. 4.4b,e).

Figures 4.4f,g and Fig. 4.5 (solid lines) show that the depth of the lake breeze (consisting of cooler, moister air and northwesterly winds) at 1600 UTC is  $\sim 100$ – $200$  m greater in the models compared to the profiles measured by the dropsondes over both lake and land. In addition, specific humidity within the lake breeze layer is lower in the model than the observations (Figs. 4.5b,f), despite the model being warmer (Figs. 4.5a,e). This suggests that the lake breeze is too dilute in the model as a result of its greater vertical extent. The model is also generally drier than observed immediately above the lake breeze layer ( $\sim 1600$ – $2200$  m MSL, Figs. 4.5b,f), therefore the height of the lake breeze layer in the model may be caused by excessive entrainment drying at the upper boundary, also contributing to the reduced moisture compared to observations. Given that the near-surface potential temperature is  $\sim 2$  K warmer in the model than observed over the lake (Fig. 4.5a), another contributing factor may be that surface heat fluxes are too large, causing excessive deepening through turbulent eddies. Note that the observed and modeled surface temperatures are very similar between the model and observations over land (Fig. 4.5e), suggesting inadequacies in the representation of the lake surface temperature or the interaction of the lake surface with the atmosphere in the model. It is known that stably stratified boundary layers (BLs) are difficult to simulate in numerical models, related to the parametrisation of turbulent diffusion (e.g. Sandu et al. 2013; Holtslag et al. 2013; Fiedler et al. 2013), but further analysis is required to fully address this issue.

At 1300 and 1600 UTC, the 300 m model shows southeasterlies above the lake breeze, which are particularly strong above 5000 m MSL (Fig. 4.4d,f) and corresponds to the lake breeze return flow (Thiery et al. 2015; Woodhams et al. 2019). This is also confirmed in the sonde profiles at 1600 UTC (Fig. 4.4f). At 1300 UTC, the 300 m model shows that the region of enhanced southeasterlies extends between the lake breeze front and the shoreline (Figs. 4.4d). Figure 4.4e shows enhanced specific humidity in this region, where the return flow advects moisture from the lake breeze, that has been mixed upwards over land, back over the lake. At 1600 UTC, the lake breeze front has moved further onshore and the return flow has extended further back over the lake, increasing the horizontal extent of the return flow region (Figs. 4.4f,g). The return flow is clearly visible in sondes A and B over land, where there are strong southeasterlies and high specific humidity above  $\sim 1500$  m MSL (Figs. 4.4f,g). Sondes C and D (15 and 35 km offshore respectively) also show enhanced specific humidity up to  $\sim 4500$  m MSL (compared to E and F), but the south easterlies are reduced compared to A and B, in agreement with the model.

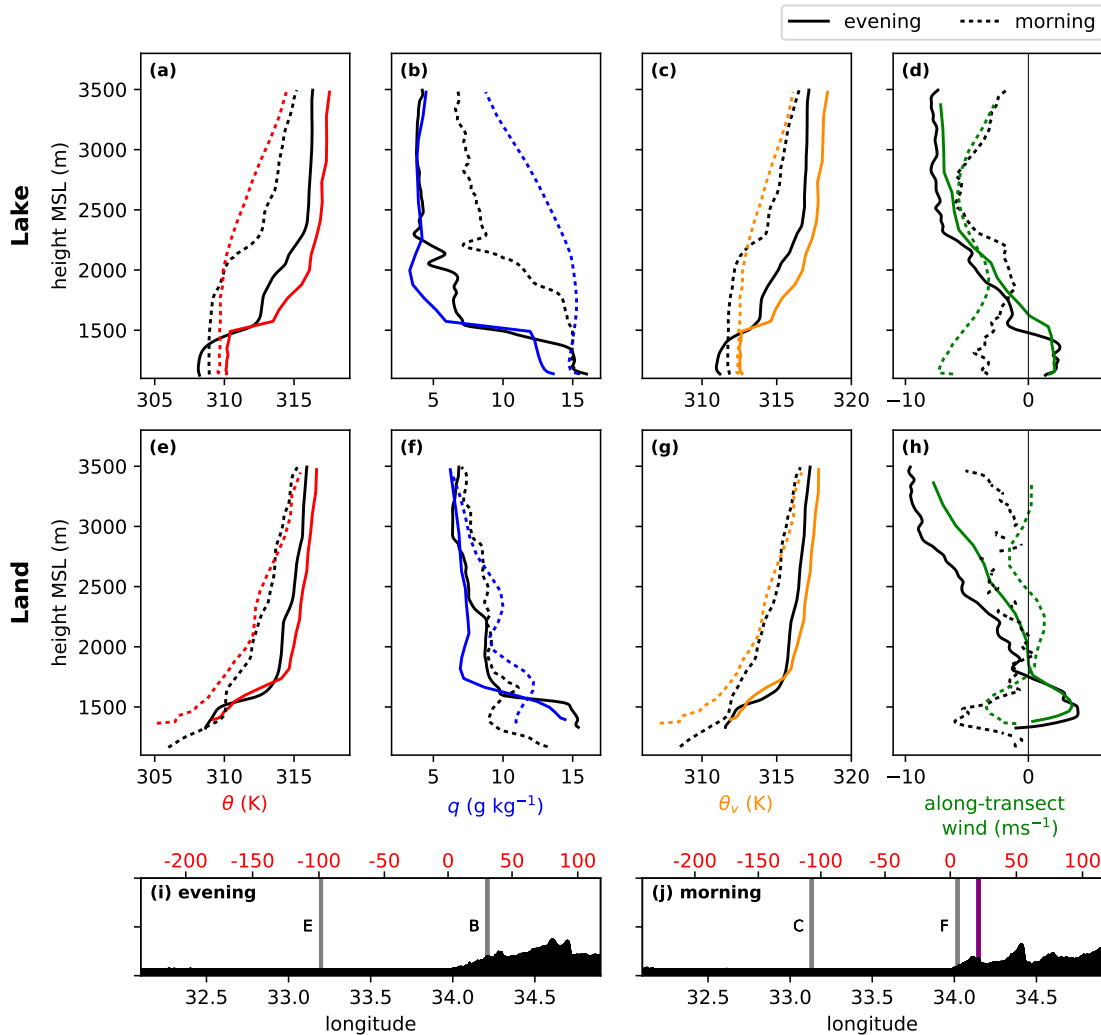


Figure 4.5.: Profiles of observed (black, from dropsondes) and simulated (colours, from the 300 m model) (a),(e) potential temperature, (b),(f) specific humidity, (c),(g) virtual potential temperature and (d),(h) along-transect wind (positive winds correspond to northwesterlies) over (a-d) Lake Victoria and (e-h) land to the east of the lake. The location of the sonde drops are shown for (i) the evening flight and (j) the morning flight by the black lines. Given that the sonde over land during the morning flight was very close to the lake shore, the simulated profile is computed further inland on the purple line in (j). In (i),(j) the distance in km from the eastern shore of the lake is given by the red numbers and the letters correspond to the sonde labels in Figs. 4.4 and 4.6.

The effect of the lake breeze return flow on the vertical profile is shown more clearly in the vertical profiles of sondes B and E in Fig. 4.5 (solid black lines). Figures 4.5d,h show southeasterlies above 1500 m MSL ( $\sim 350$  m AGL) over the lake (E) and 1750 m MSL ( $\sim 425$  m AGL) over the land (B). Over land, observations show that the effect of the return flow is strongest in a distinct layer of air between  $\sim 1500$ – $2200$  m MSL ( $\sim 350$ – $1050$  m AGL), just above the lake breeze), which is drier and warmer than the lake breeze below, but cooler and more moist than the air above (Figs. 4.5a,b). Even in sonde E (75 km offshore), there is evidence of this layer with increased moisture between 1600– $2200$  m MSL ( $\sim 350$ – $950$  m AGL) (Fig. 4.5f).

At all three model resolutions, the prevailing southeasterly along-transect winds in the 50 km ahead of the lake breeze are too weak compared to the observations (Figs. 4.4a,d). This could be related to the temperature difference between the lake and the land (Simpson 1994), or the ability of the model to reproduce mountain flows. Goler (2004) showed that sea breezes are deeper when not opposed by environmental winds, so the weaker winds ahead of the lake breeze front may explain some of the increased lake breeze depth in the models.

The 300 m model shows a pronounced ‘head’ ( $\sim 10$  km wide) at the leading edge of the lake breeze, which is approximately double the depth of the flow behind (Figs. 4.4d,e). According to density current theory, a head forms at the leading edge of a density current and is approximately  $2\times$  deeper than the following flow (Simpson 1982). Waves can also form and break at the top of the head (Simpson 1982). Figures 4.4a,b, show oscillations in observed along-transect wind and specific humidity at approximately  $-25$  and  $-35$  km, which could correspond to waves. However, it is more likely that these oscillations correspond to turbulent eddies in the convective BL over land (Reible et al. 1993), which are responsible for the deepening of the lake breeze over land relative to the lake (Figs. 4.4d,e).

Along the 300 m AGL transect, cumulus congestus were visible from the aircraft, close to where the lake breeze front was observed over land. Low clouds just onshore of the eastern shore are also visible in the satellite image at 1315 UTC (Fig. 4.1c). By 1345 UTC, the congestus in the vicinity of the lake breeze front had developed into deep convection, likely triggered by convergence at the front. This convection lasted  $\sim 1.5$  hours, remaining as a small isolated cumulonimbus, before decaying shortly before sunset. Relatively cold cloud is shown close to the flight track in the satellite image at 1545 UTC, likely remnants of the observed storm (Fig. 4.1d).

### 3.c. Land breeze

The along-transect winds, specific humidity and virtual potential temperature  $\theta_v$  measured from the sonde curtain during the morning flight (between 0227-0245 UTC) are shown in Figs. 4.6a-c. Only one sonde F was dropped over land, very close to the eastern shoreline. Compared to the sondes over the lake, the profile observed by sonde F is much drier, especially near the surface (Figs. 4.6b). In the lowest 200 m, the virtual potential temperature of the air is low, and there are strong southeasterly winds between 1300 and 1500 m MSL (150–350 m AGL), suggesting a shallow land breeze across the eastern shoreline (Figs. 4.6a,c and brown lines in Fig. 4.7). However, this land breeze does not penetrate far offshore (<30 km), shown by the high moisture and increase in  $\theta_v$  near the surface in sonde E (Figs. 4.6b,c, c.f. brown and purple lines in Fig. 4.7a). There is also no signal of the band of strong southeasterlies between 1300 and 1500 m MSL in sonde E (Fig. 4.6a, c.f. brown and purple lines in Fig. 4.7b). The offshore extent of the lake breeze across the eastern shore in the model is consistent with the observations (Figs. 4.6a,c). Sondes D and E show a band of increased southeasterlies between  $\sim$ 1600-2600 m MSL ( $\sim$ 450–1450 m AGL), whereas sonde F shows decreased southeasterlies at this level (Fig. 4.6a and c.f. purple and red lines with brown line in Fig. 4.7b), which could signal a divergent return flow above the land breeze front. Such a return flow is seen in the model, centred at  $\sim$ 30 km offshore (Fig. 4.4a).

Figure 4.1g shows that there was some relatively warm cloud ( $>260$  K) above the region where sondes D and E were dropped. At this time, solid cloud cover was also reported beneath the aircraft, estimated at  $\sim$ 6000 m MSL. A Skew-T produced by the sonde data (not shown) implies a cloud top of 5550 m MSL. According to the Global Precipitation Measurement (GPM) IMERG product (Huffman 2017; Huffman et al. 2018), there was no rainfall associated with this cloud. However, a precipitating cloud may explain the cooler, moister air near the lake surface in sonde D (Figs. 4.4b,c, red line in Fig. 4.7a).

Over the west of the lake, sonde A shows lower  $\theta_v$  air near the surface compared to the sondes over the centre of the lake (Fig. 4.6c, c.f. blue line to orange and green lines in Fig. 4.7a). Low  $\theta_v$  air could indicate a land breeze across the western shore, although this would have had to propagate over 70 km across the lake (against the prevailing wind). In addition, there are no northwesterlies near the surface in sonde A (Fig. 4.6a, blue line in Fig. 4.7b). Note that the model predicts a land breeze which extends only 10–15 km offshore from the west (Figs. 4.6a-c). There may have been a northwesterly



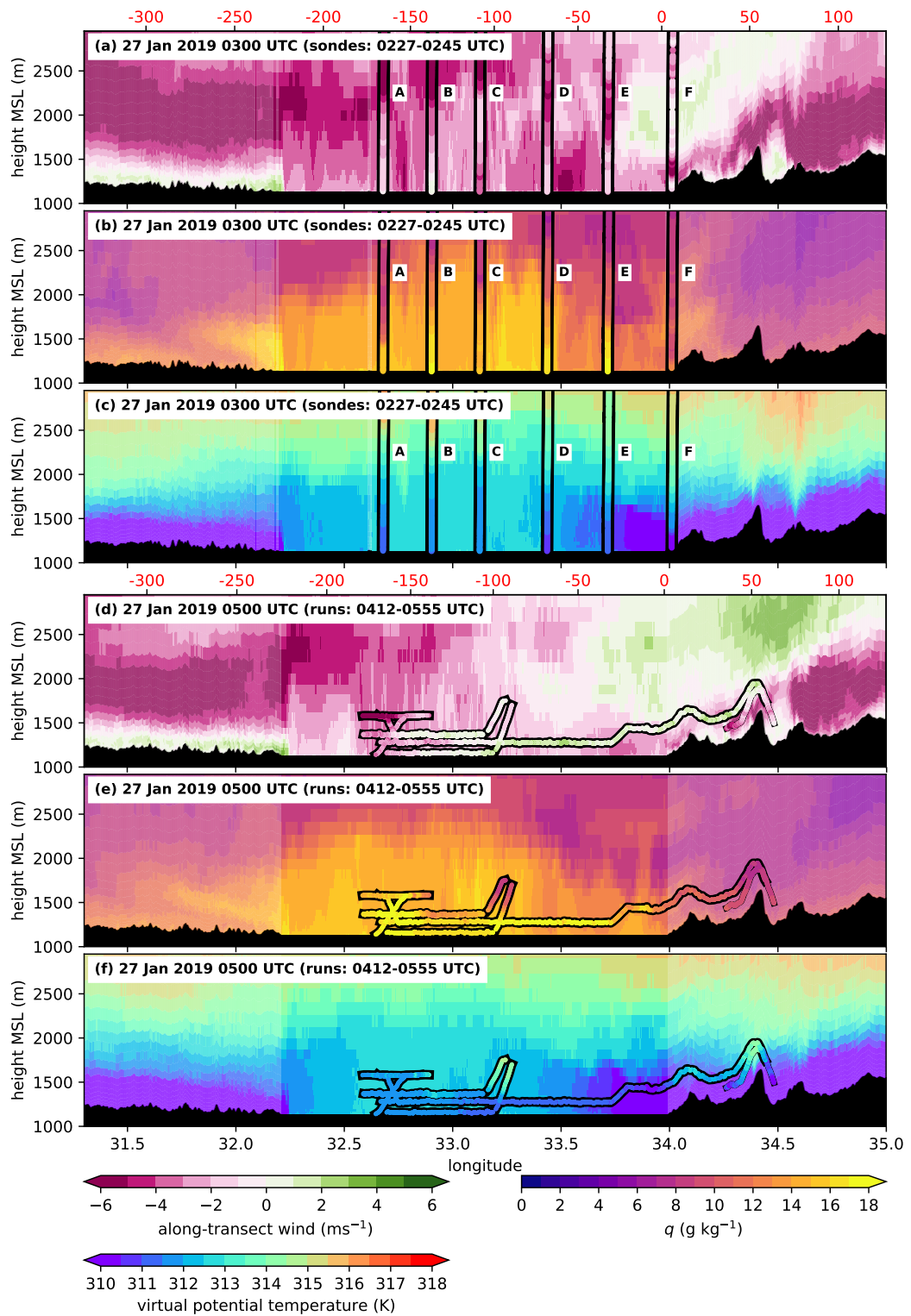


Figure 4.6.: Cross-sections of (a),(d) along-transect wind, (b),(e) specific humidity and (c),(f) virtual potential temperature from (a-c) sondes and (d-f) the aircraft. The simulated variables (from the 300 m model) from the closest model output time to the observations are plotted underneath the observations. All plots share the same x-axis (longitude) but the distance in km from the eastern shore of the lake (red numbers) differ between (a-c) and (d-f) due to slightly different flight tracks.

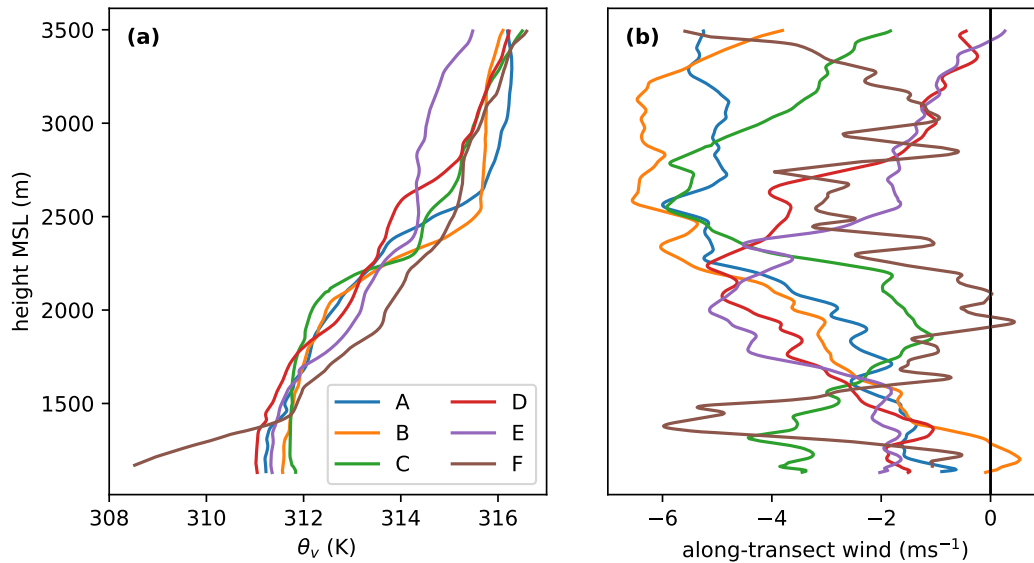


Figure 4.7.: (a) Virtual potential temperature and (b) along-transect wind profiles measured by sondes during the morning flight. Positive winds correspond to northwesterlies. The letters correspond to the labels in Fig. 4.6.

anomaly relative to the prevailing flow in the observations, but there is insufficient data coverage to support this.

Despite no northwesterlies in sonde A, weak northwesterlies were observed over the lowest  $\sim 100$  m of sonde B ( $\sim 90$  km from the western shore), the profile of which also has higher and deeper moisture than the two sondes either side (Figs. 4.6a,b, orange line in Fig. 4.7b). To the east, sonde C shows stronger southeasterlies at the surface than any other sonde over the lake (green line in Fig. 4.7b). Low-level northwesterlies in sonde B and southeasterlies in sonde C suggest low-level convergence between the two sondes. The increased and deeper moisture in sonde B supports the suggestion of convergence (Fig. 4.6b). However, it is unclear how this convergence originated given that the land breezes from the east and west coasts did not reach this part of the lake. In general, the sonde curtain suggests high variability on small spatial scales across the lake, which cannot be explained with the sparse data coverage, but which may be important for cloud formation and, in certain cases, initiation of deep convection.

Following the sonde drops, transects were also flown at  $\sim 3000$  m and  $\sim 1000$  m AGL ( $\sim 4000$  m and  $\sim 2000$  m MSL respectively). Along the 1000 m leg, clouds were observed along the track, in particular around  $32.6^\circ\text{E}$  (where the cloud base was observed approximately 50 m below, and the surface of the water beneath was ‘rippled’) and  $33.5^\circ\text{E}$  (where the aircraft entered cloud). Figure 4.1g also shows two regions of low

cloud over the lake (one to the west and one to the east) corresponding to the observed locations on the track.

Figures 4.6d-f show aircraft observations from low-level flying between approximately 0400 and 0600 UTC. Along the transect  $\sim 300$  m AGL ( $\sim 1450$  m MSL), there are two regions of increased specific humidity ('bulges'), where measurements of  $>17$  g  $\text{kg}^{-1}$  were taken. These approximately correspond to the locations where clouds were observed on the mid-level legs (note there will be some difference given the time delay between the various legs). Given time constrictions on the flight, the eastern bulge was only sampled at one height. The western bulge was chosen to be sampled in detail because, according to the operational forecasts, this was the expected location of the moisture bulge shown in Figs. 7j-l and 14d-f in Woodhams et al. (2019). This region was sampled at four heights and along two profiles, between  $\sim 30$ – $500$  m AGL, and between  $32.6$ – $32.9^\circ\text{E}$ . Over this region, specific humidity exceeding  $17$  g  $\text{kg}^{-1}$  extended to a depth of at least  $500$  m AGL ( $\sim 1650$  m MSL). At the same height at  $33.25^\circ\text{E}$  ( $\sim 60$  km east), specific humidity was only  $8$ – $10$  g  $\text{kg}^{-1}$ . The along-transect horizontal extent of the high moisture decreases with height, suggesting a dome shape. Inside the bulge, (marked by higher specific humidity, Figs. 4.8i-k) the air was generally more turbulent (marked by greater fluctuations in vertical motion, Figs. 4.8m-o), which likely explains the 'rippled' surface of the water observed from the aircraft. On the leg closest to the lake surface ( $\sim 30$  m AGL, Figs. 4.8a,e,i,m) the standard deviation of the vertical velocity was  $0.17$   $\text{ms}^{-1}$  inside the bulge and  $0.13$   $\text{ms}^{-1}$  outside the bulge. At  $\sim 240$  m AGL (Figs. 4.8c,g,k,o), the standard deviation in vertical velocity was  $0.20$  and  $0.10$   $\text{ms}^{-1}$  inside and outside the bulge respectively.

At the time of these transects, large-scale convergence was anticipated near the surface of the lake, between southeasterly flow (consisting of prevailing wind, land breeze and katabatic flow) and a land breeze from the northwest. However, northwesterlies were generally observed to the east of the western bulge, and southeasterlies were observed inside the bulge, resulting in large-scale divergence across the lake surface (Figs. 4.6d, and 4.8e-h). Although a divergent return flow would be expected above the convergent land breeze, large-scale divergence was observed at just  $30$  m AGL (Fig. 4.8e), suggesting that the aircraft was not sampling the return flow. It is unclear whether there was still a land breeze across the eastern shore at this time, since the aircraft could not get low enough to sample this region, but no land breeze was detected  $30$  km offshore at  $150$  m AGL (Figs. 4.6d,f).

Given the lack of large-scale convergence, the mechanism for bulge formation is unclear.

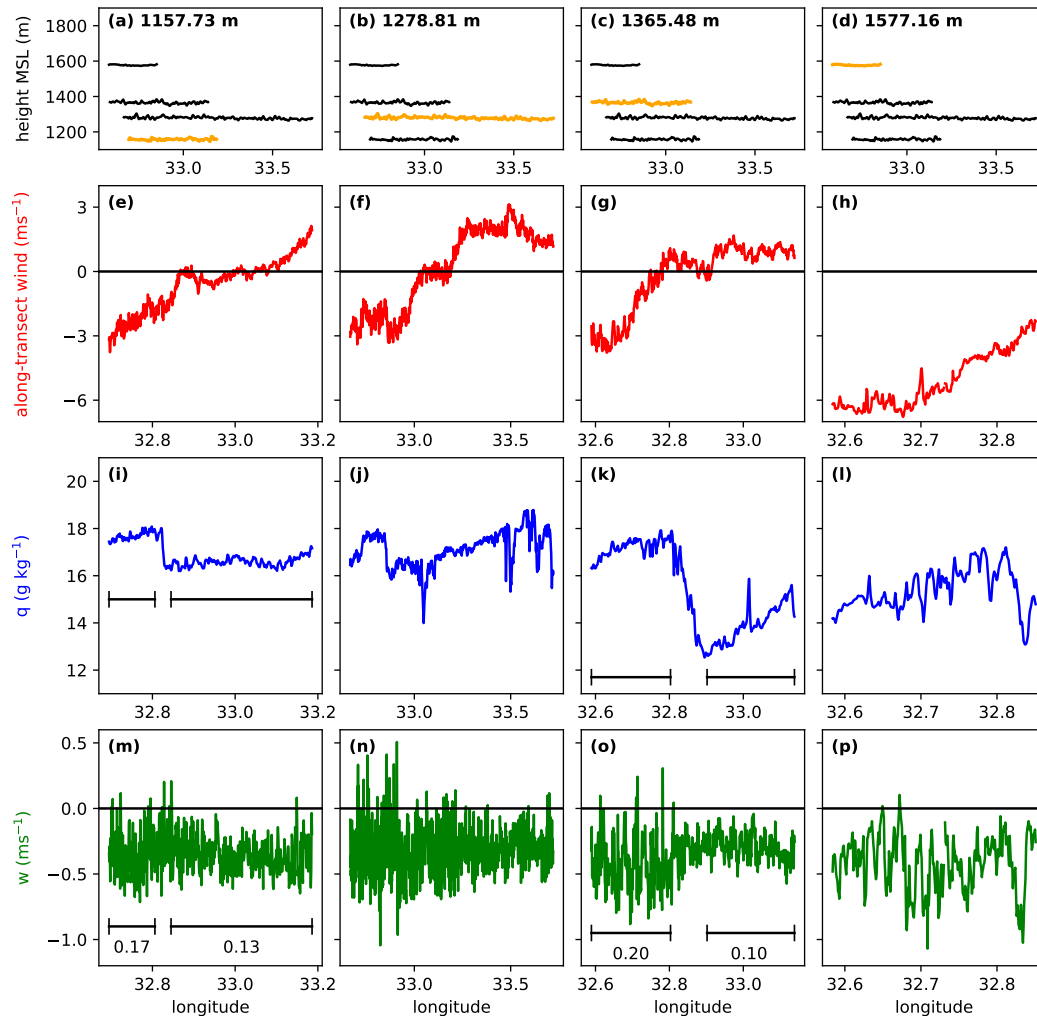


Figure 4.8.: (e-h) Along-transect wind, (i-l) specific humidity and (m-p) vertical velocity along over-lake transects plotted in orange in (a-d) during the morning flight. Regions of negative gradient of the lines in (e-h) correspond to along-transect convergence. The ‘bulge’ of moist air identified in Fig. 4.6e lies approximately to the west of 32.9°E. The numbers in (m) and (o) refer to the standard deviation of the vertical wind speed within the marked regions (inside and outside the bulge).

One possibility is that the observed divergence shows the formation of the lake breeze very early in the day, with the signal of elevated moisture—produced by overnight convergence—persisting into the morning. Note that sunrise was at 0359 UTC, so the lake breeze would have had to form almost immediately after surface warming began. In addition, Datta (1981) showed that the land breeze tends to be strongest between 0300 and 0600 UTC, however this is likely to vary based on prevailing winds and other synoptic conditions. In the model, the divergent lake breeze does not begin to form until 1000 UTC (not shown). A possible explanation for the longer-term maintenance of the bulge (without continued convergence) is that it exhibits wave-like characteristics, such that it may have propagated away from its source region. If the wave were propagating westward, then divergence would be expected in its wake.

Another suggestion is that the bulge formed from a region of very small-scale convergence. Figures 4.8e-h show along-transect wind measured along the runs over the lake. Sections of the runs with a negative gradient of along-transect wind show wind decelerating toward the northwest, therefore convergence along the transect. In several of the transects, convergence occurs over a horizontal extent of  $\sim 10$  km slightly to the east of the bulge (the eastern edge of which is located around  $32.9^\circ\text{E}$ , identified in specific humidity, Figs. 4.8i-l). Further analysis is required to determine whether this small region of convergence could be responsible for the observed bulge. The small scale of this convergent region again suggests that small-scale perturbations occur over the lake, for which the processes responsible are unknown, but which may be important for cloud and convection formation.

Note that the model also simulates two bulges of enhanced moisture (between  $32.3\text{--}32.5^\circ\text{E}$  and between  $33.0\text{--}33.1^\circ\text{E}$ ), within the lake BL. However, it is unclear whether these correspond to the observed bulges given the very different winds in the model and observations.

The moist layer above the lake in the model is approximately double the depth of the observations (differences of  $\sim 1000$  m) and has a less well-defined boundary at the top (Figs. 4.6b,e, 4.5a-c). Given the depth of the BL in the model, the air inside is more dilute, especially above 1600 m MSL ( $\sim 500$  m AGL) (Figs. 4.6b,e and 4.5b.) Fig. 4.5a shows that the potential temperature in the model is slightly warmer ( $< 1$  K) than the observations at the surface, suggesting increased surface heat fluxes may be responsible for the deeper BL, likely attributed to a warmer lake surface. However, issues with the parametrisation of sub-grid mixing in the model may also play a role.

## 4. Discussion and conclusions

A pilot field campaign using the FAAM BAe-146 aircraft to sample the lake–land breeze circulation over Lake Victoria successfully carried out two flights, observing the lake and land breezes at their respective times in the diurnal cycle. High-resolution CP MetUM simulations were performed over the HyVic period. Some preliminary model evaluation was performed and, where appropriate, the simulations were used to fill gaps in the aircraft data.

During the evening flight, the lake breeze front across the eastern shore was sampled at  $\sim 300$  m AGL, and exhibited a wind reversal in which the velocity changed by  $\sim 8$   $\text{ms}^{-1}$  and specific humidity decreased by  $5$   $\text{g kg}^{-1}$  over just  $5$  km. Model verification was performed showing that, for CP MetUM, a horizontal grid-spacing of at least  $1.5$  km is required to simulate the strong gradient in wind and moisture across the lake breeze front. At a horizontal grid-spacing of  $4.4$  km, this gradient is too weak, spread over  $\sim 30$  km instead of  $\sim 5$  km. Given that the current operational CP MetUM model over tropical Africa has a grid-spacing of  $4.4$  km, it is important to understand how this gradient may affect convective initiation. This could be achieved by performing idealised simulations with different horizontal grid-spacings, or by artificially imposing synoptic conditions or lake–land contrasts to change the strength of the lake breeze front. The position and gradient of the lake breeze front may also be sensitive to orography. An isolated cumulonimbus was observed to form around  $1400$  UTC in the region of the lake breeze front, with an estimated life cycle of  $\sim 1.5$  hours.

During the early morning flight, a land breeze was observed across the eastern shore at  $0245$  UTC. It is unclear how long this land breeze persisted. Between  $0400$  and  $0600$  UTC, large-scale divergence was observed close to the surface of the lake, when convergence was expected. However, two moist bulges—with some similarities to the bulge feature described in Woodhams et al. (2019)—were observed. Whereas the bulge in Woodhams et al. (2019) was formed and maintained by large-scale convergence over the lake, the mechanism for bulge formation in this case is unclear. It is suggested that very small-scale ( $\sim 10$  km) localised convergence near the lake surface may be responsible, or that the bulge persisted from convergence earlier in the night. Idealised modelling of the collision of land breezes over a water body is required to investigate whether the propagation of the bulge feature is controlled by density current or wave dynamics, in addition to collecting observational evidence. Turbulence was shown to increase inside the bulge and there were visible disturbances to the surface of the lake, which may pose a danger to small boats.

Sonde drops over the lake during the early morning showed high variability in wind, moisture and temperature between adjacent sondes, despite the drops being spaced only 25–50 km apart. The details and causes of this variability cannot be deduced given the limited data coverage. Fairly large differences between the model and observations during this time mean that it is unclear whether high resolution CP MetUM is able to capture the variability. A larger and denser number of observations is required to understand whether these small-scale features are common over Lake Victoria. Modelling studies are required to investigate whether the inability of even the 300 m horizontal grid-spacing model to simulate these features is a robust bias. In addition, Large Eddy Simulation (LES) models could be used to determine whether this small-scale variability is sufficient to initiate deep convection.

The observations have also shown that, at 300 m horizontal grid-spacing, the depth of the simulated lake breeze is too great and its moisture content over land is too low in CP MetUM. Similarly, the simulated depth of the BL over the lake during the morning flight is too deep and specific humidity close to the surface is too low. It is suggested that sub-grid turbulent mixing may be too strong in the models. Inaccurate lake surface temperatures may also play a role. An investigation into the effect of lake surface temperatures on the strength and timing of lake and land breezes in the models is required, which may be achieved by artificially changing the input temperatures from the OSTIA analysis used to define the water temperatures in the model (Fiedler et al. 2014). The lake surface temperatures in the model are only updated once per day, so the effect of applying a (realistic) diurnally varying temperature perturbation should be investigated. A more thorough model evaluation is planned with the observations; given the limited number of radiosonde releases across Africa, the upper-air observations provide a rare opportunity to look at how well CP MetUM simulates the atmospheric column.

Although it is difficult to draw robust conclusions from just two flights, the HyVic pilot flight campaign has provided direction and motivation for a future extended aircraft campaign over the region. Such a campaign should include increased sampling near the surface and a higher density of sonde drops to observe and understand small-scale variability over the lake. Aircraft can sample at many levels, but the atmosphere may quickly evolve between different legs, which makes it difficult to attribute differences to time or location. Accordingly, ground observations—including automatic weather stations (AWSs), wind profilers, doppler lidars and radiometers—are required in conjunction with the aircraft. An ideal tool—both for scientific study and operational forecasting—would be radar, which can complete a 3D scan within a relatively small

time window, such that multiple scans can be used to study the 3D evolution of the circulations or storm structure.

Given the devastating impacts of storm occurrence over the lake, improving understanding of the processes responsible for storm initiation is key to improving safety on the lake. HyVic has provided observations of the baseline lake–land breeze circulation in unprecedented detail, but an extended campaign is necessary to fully observe the system. In particular, it remains unclear how properties of the lake breeze front and nocturnal moisture bulge may vary on seasonal and synoptic timescales, and how such variations may lead to deep convection, therefore observations are required throughout the year.

## Acknowledgements

Airborne data was obtained using the BAe-146-301 Atmospheric Research Aircraft (ARA) flown by Airtask Ltd and managed by FAAM Airborne Laboratory, jointly operated by UKRI and the University of Leeds. Many thanks to various members of FAAM for their help and guidance during the campaign. Flight hours were provided by the National Centre for Atmospheric Science (NCAS). This research was conducted with financial support from UKAID under the WISER Programme through the HIGHWAY (HIGH impact Weather lAke sYstem) Project. The HIGHWAY project is an international collaboration led by the WMO, in cooperation with Kenya Meteorological Department, Meteo Rwanda, Uganda National Meteorological Authority, Tanzanian Meteorological Agency, UK Met Office and National Centre for Atmospheric Research, USA. This work was also supported by U.K. Research and Innovation as part of the Global Challenges Research Fund, Grant NE/P021077/1 (GCRF African SWIFT). This work was also supported by the U.K. Department for International Development (DFID)/Natural Environment Research Council (NERC) Future Climate for Africa (FCFA) HyCRISTAL project (NE/M019985/1). Woodhams was also supported by the NERC SPHERES DTP (Grant NE/L002574/1). Marsham was also supported by the National Centre for Atmospheric Science via the NERC/GCRF programme ACREW: Atmospheric hazard in developing Countries: Risk assessment and Early Warning. Many thanks to the PIs of the various projects involved for working together to deliver the campaign at such low cost. Thanks are owed to Festus Luboyera, Solomon Mangeni and others at Uganda National Meteorological Authority (UNMA) for their hard work in ensuring HyVic could take place. Thank you and credit to the NERC Methane



Observations and Yearly Assessments (MOYA) project for funding the transits and supporting HyVic as an addition to their campaign.

## References

- Anyah, R. O., F. H. Semazzi, and L. Xie, 2006: Simulated physical mechanisms associated with climate variability over Lake Victoria Basin in East Africa. *Mon. Wea. Rev.*, **134** (12), 3588–3609.
- Arritt, R. W., 1993: Effects of the large-scale flow on characteristic features of the sea breeze. *Journal of Applied Meteorology*, **32** (1), 116–125.
- Ba, M. B., and S. E. Nicholson, 1998: Analysis of convective activity and its relationship to the rainfall over the Rift Valley lakes of East Africa during 1983-90 using the Meteosat infrared channel. *J. Appl. Meteor.*, **37** (10), 1250–1264.
- Bush, M., and Coauthors, 2019: The first Met Office Unified Model/JULES Regional Atmosphere and Land configuration, RAL1. *Geosci. Model Dev. Discuss.*
- Cannon, T., F. Krüger, G. Bankoff, and L. Schipper, 2014: Putting culture at the centre of risk reduction. *World Disasters Report 2014: focus on culture and risk*, T. Cannon, and L. Schipper, Eds., International Federation of Red Cross and Red Crescent Societies, 185–209.
- Chamberlain, J., C. Bain, D. Boyd, K. McCourt, T. Butcher, and S. Palmer, 2014: Forecasting storms over Lake Victoria using a high resolution model. *Meteor. Appl.*, **21** (2), 419–430.
- Datta, R., 1981: Certain aspects of monsoonal precipitation dynamics over Lake Victoria. *Monsoon Dynamics*, J. Lighthill, and R. Pearce, Eds., Cambridge University Press, 333–349.
- Estoque, M. A., 1962: The sea breeze as a function of the prevailing synoptic situation. *Journal of the Atmospheric Sciences*, **19** (3), 244–250.
- Fiedler, E. K., M. J. Martin, and J. Roberts-Jones, 2014: An operational analysis of lake surface water temperature. *Tellus A*, **66** (1), 21–247.
- Fiedler, S., K. Schepanski, B. Heinold, P. Knippertz, and I. Tegen, 2013: Climatology of nocturnal low-level jets over North Africa and implications for modeling mineral dust emission. *J. Geophys. Res.: Atmospheres*, **118** (12), 6100–6121.
- Finney, D. L., and Coauthors, 2019: Implications of improved representation of con-

- vection for the East Africa water budget using a convection-permitting model. *J. Climate*, **(2019)**.
- Fleming, R. J., and R. D. May, 2004: The 2nd generation water vapor sensing system and benefits of its use on commercial aircraft for air carriers and society. Tech. rep., University Corporation for Atmospheric Research Document. Available at <https://www.eol.ucar.edu/system/files/spectrasensors.pdf>. Last accessed: 11 Sep 2019.
- Flohn, H., and K. Fraedrich, 1966: Tagesperiodische zirkulation und niederschlagsverteilung am Victoria-See (Ostafrika) (The daily periodic circulation and distribution of rainfall over Lake Victoria, in German). *Meteorologische Rundschau*, **19 (6)**, 157–165.
- Goler, R., 2004: Numerical model of cloud lines over Cape York Peninsula. Ph.D. thesis, Ph. D. thesis, Centre for Dynamical Meteorology and Oceanography, Monash University, 210pp.
- Hastings, D. A., and P. K. Dunbar, 1999: Global Land One-kilometer Base Elevation (GLOBE) digital elevation model, documentation, volume 1.0. Key to Geophysical Records Documentation (KGRD), Doc. 34, National Oceanic and Atmospheric Administration, National Geophysical Data Center. Available at: <https://repository.library.noaa.gov/view/noaa/13424>.
- Hogan, E., A. Shelly, and P. Xavier, 2015: The observed and modelled influence of the Madden–Julian Oscillation on East African rainfall. *Meteor. Appl.*, **22 (3)**, 459–469.
- Holtzlag, A., and Coauthors, 2013: Stable atmospheric boundary layers and diurnal cycles: challenges for weather and climate models. *Bull. Amer. Meteor. Soc.*, **94 (11)**, 1691–1706.
- Huffman, G., 2017: GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V05, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Last accessed: 20 Dec 2017, DOI: 10.5067/GPM/IMERG/3B-HH/05.
- Huffman, G. J., and Coauthors, 2018: Algorithm Theoretical Basis Document (ATBD) Version 5.2. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Tech. rep., Available at: [https://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V5.2.pdf](https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2.pdf).

- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77** (3), 437–472.
- Lumb, F., 1970: Topographic influences on thunderstorm activity near Lake Victoria. *Weather*, **25** (9), 404–410.
- Milton, S., A. Diongue-Niang, B. Lamptey, C. Bain, and C. Birch, 2017: Numerical Weather Prediction over Africa. *Meteorology of tropical West Africa: The forecasters' handbook*, D. J. Parker, and M. Diop-Kane, Eds., 1st ed., John Wiley & Sons, chap. 10, 380–422.
- Mukabana, J. R., and R. A. Pielke, 1996: Investigating the influence of synoptic-scale monsoonal winds and mesoscale circulations on diurnal weather patterns over Kenya using a mesoscale numerical model. *Mon. Wea. Rev.*, **124** (2), 224–244.
- Okeyo, A. E., 1986: The impact of Lake Victoria on the convective activities over the Kenya Highlands. *J. Meteorolog. Soc. Japan Ser. II*, **64**, 689–695.
- Pohl, B., and P. Camberlin, 2006a: Influence of the Madden-Julian Oscillation on East African rainfall. Part I: intraseasonal variability and regional dependency. *Quart. J. Roy. Meteor. Soc.*, **132** (621), 2521–2539.
- Pohl, B., and P. Camberlin, 2006b: Influence of the Madden-Julian Oscillation on East African rainfall. Part II: March-May season extremes and interannual variability. *Quart. J. Roy. Meteor. Soc.*, **132** (621B), 2541–2558.
- Reible, D., J. Simpson, and P. Linden, 1993: The sea breeze and gravity-current frontogenesis. *Quarterly Journal of the Royal Meteorological Society*, **119** (509), 1–16.
- Sandu, I., A. Beljaars, P. Bechtold, T. Mauritsen, and G. Balsamo, 2013: Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models? *J. Adv. Model. Earth Syst.*, **5** (2), 117–133.
- Schreck, C., n.d.: North Carolina Institute for Climate Studies: Tropical Monitoring. Available at: <https://ncics.org/portfolio/monitor/mjo/>. Last accessed: 19 Aug 2019.
- Semazzi, F., 2011: Enhancing safety of navigation and efficient exploitation of natural resources over Lake Victoria and its basin by strengthening meteorological services on the lake. Tech. rep., North Carolina State University Climate Modeling Laboratory.
- Simpson, J., D. Mansfield, and J. Milford, 1977: Inland penetration of sea-breeze fronts.

- Quart. J. Roy. Meteor. Soc.*, **103** (435), 47–76.
- Simpson, J. E., 1982: Gravity currents in the laboratory, atmosphere, and ocean. *Annu. Rev. Fluid Mech.*, **14** (1), 213–234.
- Simpson, J. E., 1994: *Sea breeze and local winds*. Cambridge University Press.
- Song, Y., F. H. Semazzi, L. Xie, and L. J. Ogallo, 2004: A coupled regional climate model for the Lake Victoria basin of East Africa. *Int. J. Climatol.*, **24** (1), 57–75.
- Thiery, W., E. L. Davin, H.-J. Panitz, M. Demuzere, S. Lhermitte, and N. Van Lipzig, 2015: The impact of the African Great Lakes on the regional climate. *J. Climate*, **28** (10), 4061–4085.
- Tollefson, J., 2019: Tropical Africa could be key to solving methane mystery. Available at: <https://www.nature.com/articles/d41586-019-00457-7>. Last accessed: 11 Sep 2019.
- Vance, A., S. Abel, R. Cotton, and A. Woolley, 2015: Performance of WVSS-II hygrometers on the FAAM research aircraft. *Atmos. Meas. Tech.*, **8** (3), 1617–1625.
- Woodhams, B. J., C. E. Birch, J. H. Marsham, C. L. Bain, N. M. Roberts, and D. F. Boyd, 2018: What Is the Added Value of a Convection-Permitting Model for Forecasting Extreme Rainfall over Tropical East Africa? *Mon. Wea. Rev.*, **146** (9), 2757–2780.
- Woodhams, B. J., C. E. Birch, J. H. Marsham, T. P. Lane, C. L. Bain, and S. Webster, 2019: Identifying key controls on storm formation over the Lake Victoria Basin. *Mon. Wea. Rev.*, **147** (9), 3365–3390.
- Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2015: The Annual Cycle of East African Precipitation. *J. Climate*, **28** (6), 2385–2404.



## Chapter 5.

# Conclusions

### 1. Overview

The purpose of this thesis was to improve the understanding and forecasting of severe weather over East Africa, in particular the Lake Victoria basin, through model verification and detailed study of storm processes and local circulations. This work is vital for improving safety on the lake, where thousands of fatalities are attributed to storms and their associated high winds and waves (Cannon et al. 2014). East Africa also has a high socio-economic dependence on Lake Victoria; an estimated 3.5 million inhabitants of its surrounding shores are dependent on the lake for their livelihoods (Semazzi 2011), including fishermen and farmers. In addition, the lake is an important trade and transport link between countries. Water levels in the lake are strongly tied to precipitation over the lake (e.g. Flohn and Burkhardt 1985; Piper et al. 1986; Yin and Nicholson 1998; Vanderkelen et al. 2018), of which around half may be attributed to convective rainfall (Flohn 1983).

This thesis considered current state-of-the-art operational numerical weather prediction (NWP) in East Africa by performing a detailed verification of the (then) operational convection-permitting (CP) East Africa model run by the Met Office (chapter 2). Over the last couple of decades, CP NWP has provided a step-change in the capability of meteorological services to forecast convection (Clark et al. 2016). In recent years, CP models have become operational across tropical domains, with the East Africa model one of the first of its kind. Although much work has been done to evaluate the skill of CP NWP in the mid-latitudes (e.g. Lean et al. 2008; Baldauf et al. 2011; Seity et al. 2011), verification over the tropics is lacking.

The ability of the East Africa model to forecast the most severe rainfall events was investigated and compared to global (parametrised) Met Office Unified Model (MetUM). This study was the first assessment over an extended period (two years) of CP MetUM in the tropics. Results from this work can be used to provide more guidance to local forecasters on how to use the model. Although the model had existed since 2012, personal communication with forecasters made it apparent that its benefits relative to the global model were not well understood. As a result, the model was not being used optimally.

A new variant of the Fractions Skill Score (FSS) (Roberts and Lean 2008) was introduced in chapter 2, termed the Localised Fractions Skill Score (LFSS). This metric reflects model skill at each grid-point, rather than at each timestep, as in the traditional FSS. By determining how skill varies within the model domain, knowledge of the geography of the region can provide insight into reasons behind good or bad model performance.

In general, the CP model showed greater skill for the prediction of severe rainfall compared to the global model. In particular, the model had superior performance on sub-daily timescales, especially between 1500–2100 LT. The difference in skill for 24 hour rainfall accumulations (most often used by forecasters) was minimal. The improved skill in the CP model likely reflects the improved diurnal cycle of rainfall. The CP model showed the best performance over land, especially around mountains and near the coast. Improved performance in these regions likely demonstrates the ability of the CP model to represent processes such as sea/lake breeze circulations and orographic flows, and respond to the resulting local-scale convergence (Birch et al. 2014, 2015). Over Lake Victoria, the model showed an improved hit rate for storm prediction compared to the global model, but more false alarms were recorded. Using the FSS, a ‘skilful’ scale was identified at greater than 350 km for the top 1.5% of rainfall events. Similar results were found for CP MetUM in South Africa by Stein et al. (2019). This large scale demonstrates that further improvement of the model is required, and that forecasters must be made aware of the model limitations in order to use it successfully. Model improvements should focus on reducing biases in the timing and amount of rainfall, as well as developing CP ensembles to capture forecast uncertainty. The skill of the forecast showed a greater dependence on time of day than on forecast lead time, unlike in a midlatitude study by Roberts (2008), suggesting that adjacent model initialisations could be used as a lagged ensemble.

Looking ahead to the future of CP NWP in the tropics, work is underway in the



Global Challenges Research Fund (GCRF) African Science for Weather Information and Forecasting Techniques (SWIFT) project to run and evaluate CP ensembles with 4.4 km horizontal grid-spacing over tropical Africa. A number of CP ensembles have already been run in the mid-latitudes (e.g. Gebhardt et al. 2008; Clark et al. 2011). It is anticipated that an ensemble approach could help forecasters address the large spatial uncertainty that remains in deterministic CP models. Clark et al. (2016) also suggest that CP ensembles are useful tools for investigating biases in CP models. However, biases which are passed down to CP scales from driving models persist, and require either an improvement to parametrisation in global models, or sufficient computing resources to run a global CP model, as suggested by Bauer et al. (2015). CP models are also increasingly being used for climate projections (e.g. Prein et al. 2015; Stratton et al. 2018), therefore it is very important that their limitations in predicting the current climate are well-understood.

Results from chapter 2 demonstrated that the CP model does show improved performance over the parametrised global model, but still leaves much room for improvement. Where a forecast model is unreliable in predicting rainfall, a good understanding of the local- and large-scale conditions for severe storm formation is essential. In this case, observations and other model fields (if trusted) can be used to inform a forecast. Improved physical understanding of processes that the model must accurately represent for a successful forecast, can inform a targeted approach to model development. The aim of chapters 3 and 4 was therefore to enhance fundamental understanding of controls on severe weather in the region, through high-resolution modelling studies and observations, respectively. Chapters 3 and 4 focused only on the Lake Victoria basin, given its importance in the region.

Studies of rainfall over Lake Victoria, in particular its link to the lake–land breeze circulation, date back to Flohn and Fraedrich (1966). However, many previous studies were performed with coarse climate models and analysed only the mean diurnal cycle (e.g. Song et al. 2004; Williams et al. 2015; Thiery et al. 2015). Such approaches neglect details on small spatial scales (and model errors are likely large) and obscure processes which arise from day-to-day variability. The work presented here took advantage of state-of-the-art CP MetUM, down to a horizontal grid-spacing of just 300 m in chapter 4. A horizontal grid-spacing of 1.5 km was used in chapter 3 to investigate two occurrences of storms (one in the long rains and one in the dry season), and one occurrence of the lake–land breeze circulation without a storm, to identify the important controlling factors in each case.

The importance of the land breeze in the initiation (dry season storm) and propagation and sustenance (long rains storm) of convection over the lake itself was affirmed in chapter 3. It was shown that convergence over land to the east of the lake (between the lake breeze and the prevailing easterlies) propagated back across the lake overnight as the land breeze formed and prevailing easterlies strengthened. This process was responsible for the westward propagation of the long rains storm—which initiated over land to the east of the lake—across the lake. The lake–land breeze circulation alone is not enough to trigger a storm; there must also be high environmental moisture (the likelihood of which is increased during the rainy seasons, but may vary on daily timescales) and high CAPE/low CIN. The daily variability of these variables, and the thresholds needed for a storm to occur, require further investigation. In the dry season storm case, convection to the north west of the lake on the previous day was shown to alter the environment—air in the vicinity of the convection was cooler and more moist due to cold pool formation—such that the land breeze from the north west was strengthened by an increase in temperature gradient between the lake and land. Such a process has also been noted by Thiery et al. (2016). The strength of the westerly land breeze matched the strength of the easterly land breeze (combined with prevailing easterlies), creating a stationary line of convergence over the lake, above which a storm initiated.

Chapter 3 demonstrated the importance of the lake–land breeze circulation for storm initiation using only high-resolution CP simulations. Observations of this circulation, especially in the vertical, are sorely lacking. The HyVic pilot flight campaign took place over Lake Victoria using the FAAM aircraft in January 2019. Results from chapter 3 were used to inform flight plans. Two flights took place, to observe the lake and land breezes during the evening and morning respectively. Observations from this campaign were presented in chapter 4, alongside model runs with 300 m horizontal grid-spacing. In some instances, the model was used to aid the interpretation of the observations (since the aircraft data coverage was limited) but, in other instances, the observations illustrated model limitations.

Around 1330 UTC, the lake breeze front was observed  $\sim 45$  km inland, with an approximate horizontal extent of 5 km. Across the front, the wind reversed direction, with a change of magnitude of  $\sim 8$   $\text{ms}^{-1}$ , and specific humidity reduced by  $\sim 5$   $\text{g kg}^{-1}$ . A model resolution of 4.4 km could not capture the abruptness of this change, but a sharp gradient was simulated at 1.5 km and 300 m horizontal grid-spacing. Given that models can only accurately represent features with a horizontal extent greater than  $7\times$  the horizontal grid-spacing, it is unsurprising that the 4.4 km could not simulate

this feature. However, this is a concern given that the operational Met Office Tropical Africa model has a resolution of 4.4 km. During the morning flight, small-scale variability in wind, moisture and temperature—currently unexplained due to the limited data coverage—was observed over the lake surface. How important these small-scale perturbations are for the initiation of deep convection, and how well they are represented in models, requires further investigation. Comparison of the observations to CP MetUM showed that the lake breeze over land during the evening, and the boundary layer over the lake during the morning, were too deep in the model, likely due to the issues with surface fluxes or sub-grid mixing.

The dry season simulations in chapter 3 showed that a bulge of moisture formed over the lake at the easterly land breeze front, and was reinforced when the westerly land breeze formed. During the HyVic campaign, two bulges were likely observed near the eastern and western shores, and were associated with cloud and more turbulent conditions. Chapter 3 showed that strong winds, and abrupt changes in wind, may be associated with the bulge. These winds and turbulence may pose a hazard to users of the lake, even if no storm forms. The convergence responsible for these bulges was unclear from the observations. It is anticipated that the properties and behaviour of such moisture bulges may determine whether or not a storm will initiate over the lake. Therefore, further work is required to understand the mechanisms behind their formation and propagation.

## 2. Wider impact of the work

Model verification in chapter 2 could be adapted for use in training materials for forecasters in East Africa. CP NWP as a tool for forecasting is still fairly novel in Africa. Local forecasters require training and time to familiarise themselves with strengths and weaknesses of a new model. The long period over which the model evaluation was performed (two years) should give increased confidence to forecasters that its outcomes are robust. Work from this thesis will also be incorporated into the East Africa forecaster's handbook, which is currently in development, led by the SWIFT project.

Some results from the model verification are likely applicable to CP NWP across the whole tropics. As CP NWP becomes more routine in the tropics, some of the verification techniques in chapter 2 could be replicated across different domains. The LFSS is already being used in a paper in preparation by the Met Office to evaluate the new Tropical Africa model. A standardised comparison between different domains will allow

more robust conclusions to be drawn about the effect of CP models on forecast skill. Such a comparison may also highlight where poor skill is a regional effect, due to biases in the global driving model, and therefore cannot be improved by CP modelling. In addition, it is important to monitor how model skill changes as configurations are updated.

All three results chapters can be used as evidence to push for increased and improved observations across East Africa. Chapter 2 highlights the need for more observations for data assimilation in NWP, which could likely have a large impact on model performance. Increased observations would also allow a more thorough evaluation of the forecast model. Chapter 3 relied on high-resolution modelling to study processes responsible for severe weather, but it can be argued that observations are required to substantiate the findings. As a result, chapter 4 presented novel observations of the lake–land breeze circulation from an aircraft campaign. However, there is a clear need for a more intensive observational period to capture the variability and small-scale structures responsible for severe weather.

An improved understanding of processes responsible for convection, and their representation in CP models, can also provide insight into the results of climate predictions over the region, in particular by the CP4 dataset described in Stratton et al. (2018).

### **3. Limitations of the work**

One of the major limitations on the work was the lack of observations. Generally, the available observations were limited to satellite products. Satellite products benefit from being available over large regions, and can be interpolated onto regular grids. However, products may involve large biases as a result of the measurements being indirect and remotely-sensed. In particular, issues can arise around complex orography (Tang et al. 2016; Kim et al. 2017; Xu et al. 2017; Sungmin and Kirstetter 2018). The GPM IMERG product was the primary satellite product used in this study, and its limitations are discussed in detail in chapter 2, section 2.b.

The limited availability of observations ultimately restricted the possible verification performed for the CP model in chapter 2. Only the ability of the model to predict precipitation was chosen to be verified. Fortunately, this variable is particularly important since it is widely used by forecasters. Another hazard in the region, especially over Lake Victoria, is high winds. Given the lack of observational data, especially over the centre

of the lake, the ability of the model to predict this hazard was not investigated. Although some surface satellite wind data does exist (for example, National Oceanic and Atmospheric Administration (NOAA) synthetic aperture radar (SAR)-derived high-resolution wind products, which use backscattered microwave radar returns from the surface of the water to infer wind speed and direction (Pichel et al. 2015)), the spatial and temporal coverage is often low.

A lack of observations also impeded the ability to study and understand the processes controlling storms over the Lake Victoria basin. As a result, chapter 3 relied almost exclusively on model simulations, even though CP models have known biases and are not able to represent many small-scale processes (Clark et al. 2016). Although satellite observations were used to ensure that the simulations were reproducing rainfall and outgoing longwave radiation (OLR) sufficiently similar to the observations, the results of this study should still be treated with some caution.

Across East Africa, there do exist a significant number of automatic weather stations (AWSs), either owned by the national meteorological services or by private companies. To the detriment of research, this data is not widely shared, and in many cases is not even uploaded to the Global Telecommunications System (GTS). Although one hindrance in the sharing of data may be a lack of appropriate infrastructure, an increased awareness of the value of sharing observations is required in the region.

The CP model verification conducted in chapter 2 primarily considered the most intense (top 1.5%) precipitation events. Although these are likely the events with the greatest impact, this method is unable to consider events where moderate rainfall may occur over a long period. In addition, the verification in chapter 2 is performed within an Eulerian framework, whereby only rainfall at a fixed point is considered. To fully evaluate the ability of a model to simulate realistic convection and rainfall, a Lagrangian approach should also be applied. This could be achieved using storm tracking to follow a storm throughout its lifetime, comparing its properties, such as size, structure, lifetime and track, between the model and observations. However, given that the East Africa model only ran out to T+48 h, storm-tracking statistics may be affected if storms began before initialisation, or dissipated after the end of the run.

The limited number of case studies used in chapter 3 (to identify important controls on storms) does impact on the robustness of the results. Two storms with clearly different characteristics (one large storm forming over land and propagating onto the lake, and one smaller storm initiating over the lake itself) were chosen to capture a variety of processes, but do not provide a complete picture. The choice to focus on

a few, highly-detailed, case studies was based on the lack of detailed individual case studies of convection over Lake Victoria in the literature, along with computational constraints. It is acknowledged that results from chapter 3 should be scaled up and investigated in a statistical framework.

There were also various limitations associated with the field campaign presented in chapter 4, although this was to be expected given that the campaign was designed as a pilot. As with chapter 3, representativeness is an issue; two flights are far from sufficient to sample the wide range of conditions over Lake Victoria. The great size of Lake Victoria also makes it difficult to obtain good coverage of the lake within the limits of the flight time. More than 1 hour could pass between the beginning and end of a leg of the transect, by which time the atmosphere may have evolved.

## **4. Recommendations for future work**

### **4.a. Evaluation of convection-permitting models**

Since the publication of the second chapter of this thesis, the East Africa CP model has been retired and replaced with the Tropical Africa model. The Tropical Africa model includes an updated science configuration Regional Atmosphere 1 for the Tropics (RA1T) (Bush et al. 2019) and also spans West Africa. The horizontal grid-spacing (4.4 km) remains the same as the East Africa model. It is important that there is continued monitoring of CP models in the region, to ensure that upgrades are indeed improving the quality of forecasts. Accordingly, ongoing weaknesses should be fed back to model developers. Ongoing changes to models (for example, it is likely that RA2T will become operational in the near future), and the corresponding changes in performance, must be monitored. As forecasters become more familiar with models, they learn their strengths, weaknesses and biases. If any of these change, forecasters must be kept informed, since they may need to adapt the way in which they use the model.

Chapter 2 provided a quantitative verification of various aspects of precipitation forecasts from the East Africa model, providing guidance about model performance for forecasters. Although some reasons for limitations in predictability were suggested, these were not examined in detail or at a process level. Such analysis is difficult given the lack of observations, especially upper-air observations. Recently, the new Aeolus satellite has been launched and will provide profiles of wind (European Space Agency

n.d.), which could be used to investigate whether the model is correctly simulating circulation patterns. Although radiosonde launches are very infrequent in East Africa, those that do exist could be collated and used for model evaluation, especially given the current focus of the Met Office on understanding pre-convective environments in CP models. There are also more radiosonde launches in West Africa, which could be used to evaluate the Tropical Africa model in that region.

The importance of initial conditions in the skill of a forecast is well known (e.g. Guichard et al. 2010; Vié et al. 2011; Melhauser and Zhang 2012; Birch et al. 2013; Schumacher et al. 2013; Luo and Chen 2015). Neither the East Africa nor Tropical Africa model include data assimilation (other than in the driving global model). Additional data assimilation is applied to CP models at various operational centres around the world (Gustafsson et al. 2018), including the Meteorological Service Singapore (MSS), which runs a CP version of the MetUM (Huang et al. 2019). MSS assimilate surface observations, radiosondes and data from aircraft, as well as radiances, derived upper-air and surface winds and cloud fraction information from satellites. Huang et al. (2019) showed that data assimilation does improve model performance over the Singapore domain, with the greatest improvement coming from assimilation of surface observations and satellite radiances. Barring computing resource issues, assimilation of satellite data could be implemented over Africa in a similar way. Surface observations are relatively scarce over the region, although it is known that many exist which are not reported to the GTS. Improving infrastructure to report the data from these stations should be a priority in order for them to be incorporated into data assimilation.

#### **4.b. Controls on storms over the Lake Victoria basin**

Chapter 3 identified some key factors responsible for storm initiation and propagation over Lake Victoria using three case studies. As discussed in chapter 3, the representativeness of these cases is unknown. However, the findings provide a starting point to investigate the processes responsible for storms in more detail. There is now a need to scale up and look at a larger number of cases, to investigate how important and frequent various triggering mechanisms and local- and large-scale atmospheric conditions are for storm initiation. One method could be to use storm tracking to obtain statistics on where and when storms form, including information about their lifetime and propagation. Using these dimensions, a machine learning algorithm—such as K-means clustering (Lloyd 1982)—could be used to define ‘types’ of storm. Reanalysis data could then be used to build composites of these different types. It would also be

interesting to examine how the frequency of each storm type varies by season, since seasonal shifts in circulation and moisture availability are likely responsible for much of the variability in storm formation (chapter 3). In particular, it would be interesting to look in more detail at storms which initiate over the lake itself; such storms are likely more dangerous to users of the lake than those which propagate from land onto the lake, as there may be no visible prior warning.

Improved understanding of the local- and large-scale conditions and important triggers for each storm type would be of great use to forecasters, possibly providing ‘rules of thumb’ for predicting convection in the region. An additional step would be to investigate how predictable each storm ‘type’ is in models (in particular CP models) and whether this predictability is related to the ability of the model to correctly predict, and respond to, the local- or large-scale features. Such a predictability study could be performed using convective-scale ensemble simulations.

#### **4.c. Idealised modelling of land breezes**

In the scientific literature, much focus has been given to studies of sea or lake breezes, whereas land breezes have generally been neglected. In particular, there are no idealised studies of land breezes. Simulations in chapter 3 showed the importance of the land breeze in storm formation over Lake Victoria, as well as the formation of a moisture bulge. The properties of the moisture bulge, also identified in observations in chapter 4, are anticipated to be important in determining whether a storm forms, but have not been investigated in detail. For example, it remains unclear whether this bulge exhibits any wave-like properties which may affect convection downstream.

A study by Goler and Reeder (2004) investigated the collision of two sea breezes over an idealised Cape York Peninsula. A similar study is suggested, where the collision of two land breezes over an idealised Lake Victoria is simulated. Importantly, this simulation should include moist processes, which the Goler and Reeder study did not. A Large Eddy Simulation (LES) could be used for this purpose, but would need to include orography. Such a study should produce simulations of lake breeze collisions under different environmental conditions, such as environmental humidity, prevailing wind (which affects the strength/asymmetry of land breezes) and land–lake temperature contrasts. These simulations could be used to understand the conditions in which the bulge of moisture contributes to storm formation over Lake Victoria. In addition, the winds at the leading edge of the bulge, and the potential hazard which they pose, could



be investigated using the idealised model.

#### **4.d. Lake surface temperatures**

Comparison of temperature and moisture profiles near the lake surface in the models and observations in chapter 4 suggest that there may be issues with the representation of the lake surface. To what extent lake surface temperatures (LSTs) may be responsible for poor model performance in chapter 2 remains unclear. Incorrect LSTs could affect storm formation in two ways: firstly, by changing the strength of the lake–land breeze circulation by changing the temperature gradient between the lake and land; secondly, by altering the sensible and latent heat fluxes at the surface. In Met Office models, the most recently observed LSTs are prescribed at the start of a run, are not updated, and have no diurnal variation (Fiedler et al. 2014). A first step would be to find or collect observations which show how much daily and sub-daily variation there is in lake surface temperatures. If there is significant variation, then sensitivity tests must be performed using either an idealised model or full CP MetUM, for example by imposing a diurnally varying perturbation on LSTs based on the observations.

If there is large day-to-day variability in LSTs (for example, wind and temperature changes accompanying storms could affect the lake surface), a lake model coupled to the atmospheric model may be required. Previous studies (e.g. Song et al. 2004; Anyah et al. 2006; Sun et al. 2014) have suggested that a coupled 3D lake model is required in order to correctly predict convection over Lake Victoria. Such studies were performed on seasonal timescales, and a 3D lake model is likely unnecessary for NWP timescales. However, it is suggested that some sensitivity tests could be performed with an East African CP MetUM domain coupled to a 1D lake model—such as the Freshwater lake (FLake) model—to investigate whether any additional skill is gained on NWP timescales. FLake has been used by Thiery et al. (2014) over other African Great Lakes and was found to acceptably predict LSTs. Rooney and Bornemann (2013) previously coupled FLake to the MetUM UKV configuration over the UK, and showed some modification of local weather.

#### **4.e. Extended field campaign**

Chapter 4 provided a novel dataset of observations of the lake–land breeze circulation in unprecedented detail. However, just two flights and aircraft observations alone can-

not be used to provide a robust description of the dynamics and thermodynamics of the circulation. In particular, data coverage was inadequate to explain the small-scale variability in winds and moisture near the lake surface during the early morning. Although a moisture bulge was identified, the mechanism of its formation was unclear from the observations.

An extended observations field campaign, ideally over several seasons to observe different conditions and variability, would be of great value. Such a campaign could exist of an intense period of radiosonde launches, or an extended aircraft campaign. Multiple aircraft or newer technologies, such as unmanned aerial vehicles (UAVs), could be deployed to achieve greater data coverage. On the ground, AWSs, wind profilers, doppler lidars and radiometers (for temperature and moisture profiles) would increase data coverage. A priority would also be obtaining observations over the lake itself, which could be achieved by fitting boats or buoys with meteorological equipment. Radar data would also be invaluable and provide observations over the lake. Data from this campaign could be used to build on the understanding of the lake–land breeze circulation and storm formation in chapters 3 and 4. In addition, the observations could be used to build on the model verification in chapter 2, by allowing a process-based evaluation of the dynamics and thermodynamics over the basin. Such an evaluation would help to elucidate reasons for model performance—good and bad—and provide more insight into aspects of the model which should be improved by model developers.

## 5. Summary

The work presented in this thesis contributes to the advancement of forecasting severe weather over East Africa, in particular over the Lake Victoria basin. This was achieved through verification of a current convection-permitting forecast model, and through modelling and observational studies to improve the fundamental understanding of severe weather occurrence.

Convection-permitting modelling was key to this thesis, with an assessment of its ability to predict intense precipitation over the region, and its use as a tool to study storm processes over Lake Victoria in the absence of suitable observations. The CP MetUM was shown to be generally more skilful than the global parametrised MetUM, but large uncertainties in the location of convection remain, and a better understanding of correct usage of the model is also required amongst forecasters. Results from this study are likely applicable to CP models throughout the tropics. This is particularly important

as their operational use becomes more routine, CP ensembles are developed, and CP models are used for future climate predictions.

Detailed studies of individual storm cases studies over the Lake Victoria basin, using CP modelling, revealed some key factors important for storm formation; in particular, the behaviour of the convergence associated with the lake–land breeze circulation, moisture availability on daily and seasonal timescales, and the effect of convection on the preceding day. Given the lack of observations in the region, an aircraft campaign was designed and conducted to provide a new observational dataset of the lake–land breeze circulation. A very sharp lake-breeze front was observed during the evening. In the early morning, some unexplained small-scale variability was observed above the lake surface, which may be important for understanding storm initiation. The small-scale processes and features identified in chapters 3 and 4 may provide some insight into why NWP, including the models in chapter 2, lack skill in the region. Observations in chapter 4 support some of the conclusions made in chapter 3, but there is a clear need for a higher density of observations (to build a more complete picture) and a longer observational period (to sample different conditions).

## References

- Anyah, R. O., F. H. Semazzi, and L. Xie, 2006: Simulated physical mechanisms associated with climate variability over Lake Victoria Basin in East Africa. *Mon. Wea. Rev.*, **134** (12), 3588–3609.
- Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Mon. Wea. Rev.*, **139** (12), 3887–3905.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525** (7567), 47.
- Birch, C., D. Parker, A. O’Leary, J. Marsham, C. Taylor, P. Harris, and G. Lister, 2013: Impact of soil moisture and convectively generated waves on the initiation of a West African mesoscale convective system. *Quart. J. Roy. Meteor. Soc.*, **139** (676), 1712–1730.
- Birch, C. E., J. H. Marsham, D. J. Parker, and C. M. Taylor, 2014: The scale dependence and structure of convergence fields preceding the initiation of deep convection. *Geophys. Res. Lett.*, **41** (13), 4769–4776.
- Birch, C. E., M. J. Roberts, L. Garcia-Carreras, D. Ackerley, M. J. Reeder, A. P. Lock, and R. Schiemann, 2015: Sea-Breeze Dynamics and Convection Initiation: The Influence of Convective Parameterization in Weather and Climate Model Biases. *J. Climate*, **28** (20), 8093–8108.
- Bush, M., and Coauthors, 2019: The first Met Office Unified Model/JULES Regional Atmosphere and Land configuration, RAL1. *Geosci. Model Dev. Discuss.*
- Cannon, T., F. Krüger, G. Bankoff, and L. Schipper, 2014: Putting culture at the centre of risk reduction. *World Disasters Report 2014: focus on culture and risk*, T. Cannon, and L. Schipper, Eds., International Federation of Red Cross and Red Crescent Societies, 185–209.
- Clark, A. J., and Coauthors, 2011: Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. *Mon. Wea. Rev.*, **139** (5), 1410–1418.
- Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-

- permitting models: a step-change in rainfall forecasting. *Meteor. Appl.*, **23** (2), 165–181.
- European Space Agency, n.d.: Aeolus. Available at: [https://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Aeolus](https://www.esa.int/Our_Activities/Observing_the_Earth/Aeolus). Last accessed: 27 July 2019.
- Fiedler, E. K., M. J. Martin, and J. Roberts-Jones, 2014: An operational analysis of lake surface water temperature. *Tellus A*, **66** (1), 21–247.
- Flohn, H., 1983: Das Katastrophenregen 1961/2 und die Wasserbilanz des Viktoria-See-Gebietes. (The catastrophic rainfall of 1961/1962 and the water balance of the Lake Victoria region, in German). *Ber. Meteor. Inst. U. Karlsruhe*, **4**, 17–34.
- Flohn, H., and T. Burkhardt, 1985: Nile runoff at Aswan and Lake Victoria: a case of a discontinuous climate time series. *Z. Gletscherk. Glazialgeol*, **21**, 125–130.
- Flohn, H., and K. Fraedrich, 1966: Tagesperiodische zirkulation und niederschlagsverteilung am Victoria-See (Ostafrika) (The daily periodic circulation and distribution of rainfall over Lake Victoria, in German). *Meteorologische Rundschau*, **19** (6), 157–165.
- Gebhardt, C., S. Theis, P. Krahe, and V. Renner, 2008: Experimental ensemble forecasts of precipitation based on a convection-resolving model. *Atmos. Sci. Lett.*, **9** (2), 67–72.
- Goler, R. A., and M. J. Reeder, 2004: The Generation of the Morning Glory. *J. Atmos. Sci.*, **61** (12), 1360–1376.
- Guichard, F., and Coauthors, 2010: An intercomparison of simulated rainfall and evapotranspiration associated with a mesoscale convective system over West Africa. *Wea. Forecasting*, **25** (1), 37–60.
- Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, **144** (713), 1218–1256.
- Huang, X.-Y., and Coauthors, 2019: SINGV—the Convective-Scale Numerical Weather Prediction System for Singapore. *ASEAN Journal on Science and Technology for Development*, **36** (3), 81–90.
- Kim, K., J. Park, J. Baik, and M. Choi, 2017: Evaluation of topographical and seasonal

- feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.*, **187**, 95–105.
- Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Wea. Rev.*, **136** (9), 3408–3424.
- Lloyd, S., 1982: Least squares quantization in PCM. *IEEE Trans. Inf. Theory*, **28** (2), 129–137.
- Luo, Y., and Y. Chen, 2015: Investigation of the predictability and physical mechanisms of an extreme-rainfall-producing mesoscale convective system along the Meiyu front in East China: An ensemble approach. *J. Geophys. Res.: Atmospheres*, **120** (20), 10 593–10 618.
- Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69** (11), 3350–3371.
- Pichel, W. G., F. M. Monaldo, C. R. Jackson, J. Sapper, and R. Hatteberg, 2015: NOAA high resolution sea surface winds data from Synthetic Aperture Radar (SAR) on the RADARSAT-2 satellite. NOAA National Centers for Environmental Information. Dataset. Available: <https://doi.org/10.7289/v5m906ns>. Last accessed: 12 Sep 2019.
- Piper, B., D. Plinston, and J. Sutcliffe, 1986: The water balance of Lake Victoria. *Hydrol. Sci. J.*, **31** (1), 25–37.
- Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53** (2), 323–361.
- Roberts, N., 2008: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteor. Appl.*, **15** (1), 163–169.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136** (1), 78–97.
- Rooney, G. G., and F. J. Bornemann, 2013: The performance of FLake in the Met Office Unified Model. *Tellus A*, **65** (1), 21 363.

- Schumacher, R. S., A. J. Clark, M. Xue, and F. Kong, 2013: Factors influencing the development and maintenance of nocturnal heavy-rain-producing convective systems in a storm-scale ensemble. *Mon. Wea. Rev.*, **141** (8), 2778–2801.
- Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson, 2011: The AROME-France convective-scale operational model. *Mon. Wea. Rev.*, **139** (3), 976–991.
- Semazzi, F., 2011: Enhancing safety of navigation and efficient exploitation of natural resources over Lake Victoria and its basin by strengthening meteorological services on the lake. Tech. rep., North Carolina State University Climate Modeling Laboratory.
- Song, Y., F. H. Semazzi, L. Xie, and L. J. Ogallo, 2004: A coupled regional climate model for the Lake Victoria basin of East Africa. *Int. J. Climatol.*, **24** (1), 57–75.
- Stein, T., and Coauthors, 2019: An evaluation of clouds and precipitation in convection-permitting forecasts for South Africa. *Wea. Forecasting*, **34** (1), 233–254.
- Stratton, R. A., and Coauthors, 2018: A Pan-African Convection-Permitting Regional Climate Simulation with the Met Office Unified Model: CP4-Africa. *J. Climate*, **31** (9), 3485–3508.
- Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95** (3), 409–426.
- Sungmin, O., and P.-E. Kirstetter, 2018: Evaluation of diurnal variation of GPM IMERG derived summer precipitation over the contiguous US using MRMS data. *Quart. J. Roy. Meteor. Soc.*
- Tang, G., Y. Ma, D. Long, L. Zhong, and Y. Hong, 2016: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *J. Hydrol.*, **533**, 152–167.
- Thiery, W., E. L. Davin, H.-J. Panitz, M. Demuzere, S. Lhermitte, and N. Van Lipzig, 2015: The impact of the African Great Lakes on the regional climate. *J. Climate*, **28** (10), 4061–4085.
- Thiery, W., E. L. Davin, S. I. Seneviratne, K. Bedka, S. Lhermitte, and N. P. van Lipzig, 2016: Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.*, **7**, 12786.

- Thiery, W., A. Martynov, F. Darchambeau, J.-P. Descy, P.-D. Plisnier, L. Sushama, and N. P. van Lipzig, 2014: Understanding the performance of the FLake model over two African Great Lakes. *Geosci. Model Dev.*, **7** (1), 317–337.
- Vanderkelen, I., N. P. van Lipzig, and W. Thiery, 2018: Modelling the water balance of Lake Victoria (East Africa)—part 1: observational analysis. *Hydrol. Earth Syst. Sci.*, **22** (10), 5509–5525.
- Vié, B., O. Nuissier, and V. Ducrocq, 2011: Cloud-resolving ensemble simulations of Mediterranean heavy precipitating events: uncertainty on initial conditions and lateral boundary conditions. *Mon. Wea. Rev.*, **139** (2), 403–423.
- Williams, K., J. Chamberlain, C. Buontempo, and C. Bain, 2015: Regional climate model performance in the Lake Victoria basin. *Climate Dyn.*, **44** (5-6), 1699–1713.
- Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res.: Atmospheres*, **122** (2), 910–924.
- Yin, X., and S. E. Nicholson, 1998: The water balance of Lake Victoria. *Hydrol. Sci. J.*, **43** (5), 789–811.