Implications of Improved Representation of Convection for the East Africa Water Budget Using a Convection-Permitting Model

DECLAN L. FINNEY

School of Earth and Environment, University of Leeds, Leeds, United Kingdom

JOHN H. MARSHAM

School of Earth and Environment, University of Leeds, and National Centre for Atmospheric Science, Leeds, United Kingdom

LAWRENCE S. JACKSON

School of Earth and Environment, University of Leeds, Leeds, United Kingdom

ELIZABETH J. KENDON, DAVID P. ROWELL, PENELOPE M. BOORMAN, RICHARD J. KEANE, RACHEL A. STRATTON, AND CATHERINE A. SENIOR

Met Office, Exeter, United Kingdom

(Manuscript received 22 June 2018, in final form 30 November 2018)

ABSTRACT

The precipitation and diabatic heating resulting from moist convection make it a key component of the atmospheric water budget in the tropics. With convective parameterization being a known source of uncertainty in global models, convection-permitting (CP) models are increasingly being used to improve understanding of regional climate. Here, a new 10-yr CP simulation is used to study the characteristics of rainfall and atmospheric water budget for East Africa and the Lake Victoria basin. The explicit representation of convection leads to a widespread improvement in the intensities and diurnal cycle of rainfall when compared with a parameterized simulation. Differences in large-scale moisture fluxes lead to a shift in the mean rainfall pattern from the Congo to Lake Victoria basin in the CP simulation—highlighting the important connection between local changes in the representation of convection and larger-scale dynamics and rainfall. Stronger lake–land contrasts in buoyancy in the CP model lead to a stronger nocturnal land breeze over Lake Victoria, increasing evaporation and moisture flux convergence (MFC), and likely unrealistically high rainfall. However, for the mountains east of the lake, the CP model produces a diurnal rainfall cycle much more similar to satellite estimates, which is related to differences in the timing of MFC. Results here demonstrate that, while care is needed regarding lake forcings, a CP approach offers a more realistic representation of several rainfall characteristics through a more physically based realization of the atmospheric dynamics around the complex topography of East Africa.

1. Introduction

East Africa is relatively arid considering its tropical latitude (Yang et al. 2015). The livelihoods of populations there are directly or indirectly tied to rainfall and the resultant water availability, which have significant interannual variations and decadal trends, and which may change with climate change (Cook and Vizy 2013; Collins et al. 2013; Thiery et al. 2016; Nicholson 2017; Vanderkelen et al. 2018). A recent publication has suggested that mean rainfall over Lake Victoria may decrease under climate change, while extreme precipitation could increase (Thiery et al. 2016).

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

Corresponding author: D. L. Finney, d.l.finney@leeds.ac.uk

Meanwhile, the large seasonal and interannual variability has spawned research into the driving factors (Rowell et al. 2015; Nicholson 2016b), but some of this research has highlighted how many models fail to capture key aspects of the climate, such as the recent trend in the long rains of March–May (Yang et al. 2014; Rowell et al. 2015). This encourages a closer look at the atmospheric water budget of East Africa to improve our process understanding. Additional motivation for understanding the water budget is the dependence of many people on inland water bodies, which are, especially for Lake Victoria, intimately coupled with the atmospheric hydrological cycle, since the lake is largely fed by on-lake rain and loses water through evaporation (Yin and Nicholson 1998).

It is known that the parameterization of moist convection is a key source of uncertainty in models (Stephens et al. 2010; Dirmeyer et al. 2012; Marsham et al. 2013; Birch et al. 2014a; Stratton et al. 2018). Models that use parameterized convection, as is the case with almost all climate models, are known to incorrectly represent moist convective processes, leading to unrealistic accumulations and timing of rainfall (Stephens et al. 2010; Dirmeyer et al. 2012; Marsham et al. 2013). Such biases affect the heating of the atmosphere and the surface, and therefore can affect the continental-scale water budget (Birch et al. 2014b). Very high-resolution models (order kilometer grid spacing) are able to represent convection explicitly without the need for a parameterization scheme. Such models are termed *convection permitting* (CP) because storms and mesoscale convective organization are permitted, even if smaller storms are not resolved (Prein et al. 2015). Convection-permitting models are able to realistically capture rainfall characteristics at hourly time scales, unlike coarser-resolution models (Stratton et al. 2018). It has also been shown that CP simulations can substantially alter the continental-scale water budget (e.g., Marsham et al. 2013; Birch et al. 2014a), providing insights into how convection couples with large-scale flows in the real atmosphere. Our work here examines how explicit convection affects the water budget of the East African region, and in particular the Lake Victoria basin, which is an area of high population density, with some of the fastest urban expansion rates in the world (Seto et al. 2012).

East Africa has a varied climate tied to its orography, which extends from the coastline up to almost 6000 m on Mount Kilimanjaro. Among the mountains are large inland water bodies, including the largest tropical lake, Lake Victoria (Fig. 1). Flanking the east and west of the lake are mountains making up the East African rift. Mesoscale flows form a coupling between the lake and surrounding mountains leading to storm formation on the mountains during the afternoon, and over the lake in the evening, as the wind direction shifts between the lake (i.e., from the lake) and land (i.e., from the land) breezes (Anyah et al. 2006). To the east of the mountains lies the Greater Horn of Africa with its coastline to the Indian Ocean. The horn and the coastline have a drier environment than the Lake Victoria basin, but nevertheless, the region experiences flooding events such as that around Mombasa in May 2017 (Standard Media 2017). The coasts, mountains, and lakes give strong mesoscale forcings, which are expected to be better represented at high resolution. These, along with the explicit representation of convection, are expected to significantly improve the triggering and evolution of convection (Morcrette et al. 2007; Birch et al. 2014a), which may then affect the continental-scale water budget, as seen in West Africa by Birch et al. (2014b).

Sources in the atmospheric water budget are evaporation and horizontal convergence of advected moisture. Over sufficiently long time scales, on the order of months, these are balanced by precipitation which is a sink of water from the atmosphere. On shorter time scales, the water held in the atmosphere in the form of vapor and clouds can vary substantially. Over East Africa two studies have gained insight by considering the variability and effect of moisture fluxes and the water budget (Anyah et al. 2006; Yang et al. 2015). Anyah et al. (2006) showed that rainfall over Lake Victoria in a coupled atmosphere–lake model reduced by 63% when moisture at the eastern boundary over the Indian Ocean was halved. This shows the dependence of the lake on moisture flux into the region. Yang et al. (2015) found that the annual cycle of moisture flux from the Indian Ocean peaked in the rainy seasons, and that the variation in this moisture flux was primarily a result of the annual cycle in dynamics, with the humidity variation a secondary factor. Using two models that differ in their representation of moist convection, we will analyze the complete atmospheric water budget of East Africa to gain an understanding of how it is affected by either an explicit or parameterized representation. Understanding the processes affecting the atmospheric water budget can highlight potential sources of error in the simulation of precipitation, identify aspects of the parameterized convection simulation which are and are not robust with a convection-permitting simulation, and can provide confidence and understanding of the models' reliability regarding their application to future climate simulations.

A first convection-permitting simulation has been performed over Africa for 10 years using a regional climate version of the Met Office Unified Model (Stratton et al. 2018). Complementary simulations of
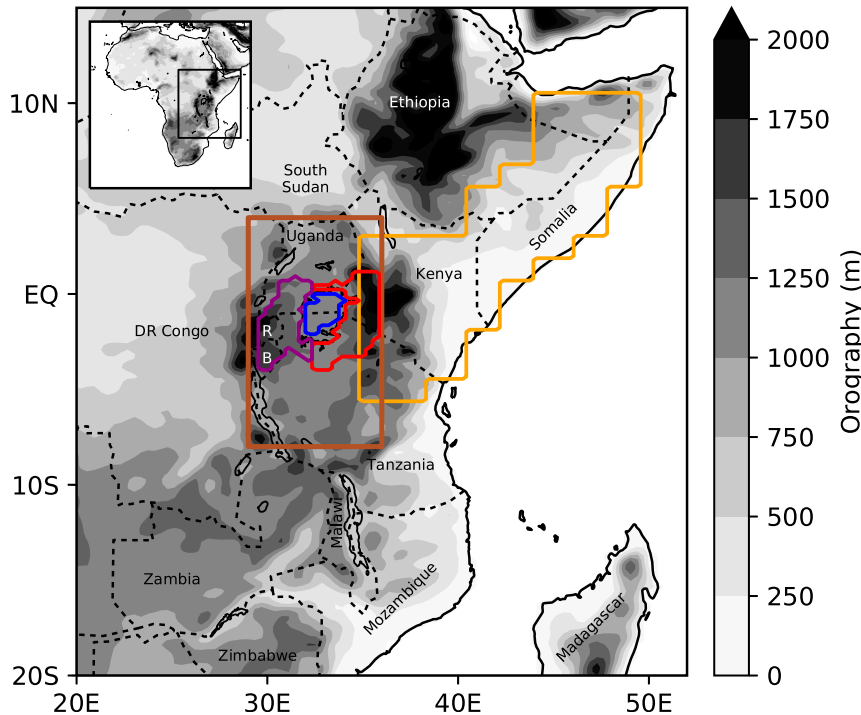
FIG. 1. East Africa region including orography, lakes, and country borders and labels. The full regional model domain used for both CP4 and P25 is shown in the inset, with the East African region box overlaid. Countries labeled R and B are Rwanda and Burundi, respectively. The brown box shows the East African Rift region used in Fig. 6. The orange, red, blue, and purple polygons show the Greater Horn of Africa, eastern Lake Victoria basin, Lake Victoria itself, and western Lake Victoria basin, respectively. These regions are used in the moisture budget analysis.

future climate (~2100; RCP8.5) have recently been completed, meaning that the results of this study will be able to inform analysis of the simulated climate change in future work. This new present-day simulation, run under the Future Climate For Africa (FCFA) Improving Model Processes for African Climate (IMPALA) project, allows us to study how an explicit representation of moist convection alters the East Africa atmospheric water budget (section 3). We then take a more detailed look at the atmospheric water budget of the Lake Victoria basin (section 4). Lake Victoria is a focus for the Integrating Hydro-Climate Science into Policy Decisions for Climate-Resilient Infrastructure and Livelihoods in East Africa (HyCRISTAL) project (Semazzi et al. 2012) due to its importance for human society, and as an interesting hydrometeorological system.

2. Data and methods

a. Regional climate simulations

Two regional climate simulations have been independently performed using the Met Office Unified

Model over an African domain (Stratton et al. 2018). Both simulations are one-way nested within an unnudged N512L85 global atmospheric model using the Global Atmosphere/Land 7.0 (GA7/GL7) configuration (Walters et al. 2017b)—the latest science configuration of the Unified Model. The global model, which includes a convection parameterization, provides lateral boundary conditions to the regional models, and is itself driven by sea surface temperature (SST) analyses (Reynolds et al. 2007). The years 1997–2007 are used from the full simulation with the global model which spans 1988–2010. The regional models use this same SST dataset at their ocean boundary for the years 1997–2007. The SST dataset has a regular spatial grid of 0.25°, and a daily temporal frequency. The Unified Model is a nonhydrostatic model based on a semi-implicit, semi-Lagrangian dynamical core. The global and regional configurations involved in this study use Even Newer Dynamics for General Atmospheric Modelling of the Environment (ENDGame) (Wood et al. 2014). The land surface is simulated using the Joint U.K. Land Environment Simulator (JULES) at the same resolution as the respective regional models, and coupled to the

atmospheric model. The simulations have been run for 10 years and 2 months (January 1997–February 2007). Unless otherwise stated all results use the 10 years of simulation following the first 2 months of spinup. Analysis here uses data archived from the simulations at hourly frequency for surface diagnostics and 3-hourly frequency for pressure-level diagnostics.

Several lakes within the model domain are represented as inland sea grid cells, many of which are located in East Africa (Stratton et al. 2018). The largest of these, Lake Victoria, is made up of 3502 grid cells in the CP simulation, and 71 grid cells in the coarser-parameterized convection simulation. Climate simulations of the lakes are rarely performed at such high resolution, so this provides a unique opportunity to study the mesoscale dynamics of their environment. Lakes included in the ARC-Lake, version 3, dataset (MacCallum and Merchant 2012; Hook et al. 2012, http://www.laketemp.net/home/) including Lakes Victoria, Tangayika, and Malawi, use a surface temperature of monthly nighttime estimates from a climatology over years 1995–2012; otherwise, the value from the nearest sea grid cell is assumed for smaller lakes (typically for lakes less than $50\,km^2$). Unlike sea surface temperatures, the lake surface temperatures do not include interannual variability. Since we limit our study to the climatological mean rainfall and dynamics, this should not greatly affect our results.

Finally, a uniform sandy soil is applied across the domain in both regional simulations. This choice is made in order to avoid introduction of rainfall biases related to unrealistic small-scale variability in soil properties (De Kauwe et al. 2013), particularly in the convection-permitting model. A detailed justification is given for the soil properties chosen in Stratton et al. (2018). For our analysis here the simulations provide a clean comparison of the effect of the representation of convection, but a caveat for our evaluation against observations is that accurate soil properties have not been used. A version of P25 with variable soil properties did not greatly affect our conclusions and we assume that both models will exhibit similar biases with regard to the use of sandy soil. Therefore, the differences between the model biases can be attributed to differences in the simulation setup.

Further details of the simulations are provided by Stratton et al. (2018). In particular, differences between the simulations are summarized in Stratton et al. 2018; see Tables 1 and 2 therein). Below we describe the relevant differences between the models for this study, including changes such as those to the boundary layer turbulence parameterization which become necessary at higher resolutions.

1) CONVECTION-PERMITTING SIMULATION OVER AFRICA AT 4.5 KM (CP4)

This regional climate simulation has no convection parameterization; instead, convection occurs explicitly based on the atmospheric equations of motion. Convection is possible without parameterization due to the high resolution used (4.5-km grid spacing at the equator). A moisture conservation scheme is applied to this simulation (Aranami et al. 2015). The large-scale cloud scheme is described by Smith (1990), as used in other convection-permitting versions of the Unified Model. The cloud scheme diagnoses liquid cloud fraction and condensed water when the gridbox mean relative humidity exceeds a critical value. Ice water content is determined by the microphysics scheme, with cloud fractions then diagnosed as in Abel et al. (2017). To explicitly simulate convection, a higher vertical resolution is used than in the parameterized convection models. The vertical grid consists of 80 levels up to 38.5 km, with higher resolution in the boundary layer and troposphere. To improve the triggering of resolved convection, the simulation applies stochastic perturbations in the subcloud layer of cumulus-capped boundary layers (following Lock et al. 2000). To account for the fact that at higher resolution some boundary layer turbulence will begin to be resolved, a scheme is used to blend between subgrid and resolved scales in order to avoid aliasing turbulent motions smaller than the grid scale onto the grid scale (Stratton et al. 2018).

For all analysis in this study, the CP4 data have been regridded using area weighting to the N512 horizontal grid. The N512 grid is more than 5 times the grid spacing of the CP4 raw grid. Therefore, this regridding also results in analyzing the CP4 data on its resolved scale. Because the convective parameterization assumes many clouds per grid cell, and produces smoother rainfall fields, it is not essential to coarsen that model data. The regridding of CP4 provides a fair comparison between the two models used here.

2) PARAMETERIZED CONVECTION SIMULATION OVER AFRICA AT 25 KM (P25)

A regional simulation with parameterized convection is used to compare against the convection-permitting simulation; this parameterized convection approach is representative of the method used in current large-scale climate models such as those used in the Intergovernmental Panel on Climate Change reports, albeit applied at a relatively high resolution. This simulation has the same grid spacing as the global driving model (approximately 26 km in latitude $\times$ 39 km in longitude at the equator). The vertical grid consists of 63 levels up to

41 km. The cloud scheme used by the model is the prognostic cloud scheme (PC2; Wilson et al. 2008) as used in the global driving model. The P25 model is similar to the global driving model but provides a better comparison to the CP4 simulation as it uses the same domain, land surface, and aerosol climatologies. This simulation allows for the isolation of the impact of convection parameterization and resolution on African climate.

The convective parameterization is that implemented in GA7 (Walters et al. 2017b). The scheme is based upon the mass flux scheme of Gregory and Rowntree (1990), with several extensions, including those described in Walters et al. (2017a). These include the addition of downdraft and convective momentum transport and the use of a convective available potential energy (CAPE) closure based on the mass-weighted, large-scale velocity in the convective column.

b. Satellite-based rainfall products

Two satellite-based rainfall products are used to evaluate the model simulations. These are the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis 3B42, version 7, and the Climate Prediction Center morphing technique (CMORPH) product (Xie et al. 2017). As with the CP4 data, all satellite data have been regridded using area weighting to the N512 resolution, in order to compare to the parameterized convection model.

The TRMM product combines passive microwave–based estimates of rainfall from multiple polar-orbiting instruments with infrared-based rainfall estimates from geostationary satellites (Huffman et al. 2007). Bias correction is then applied to the estimates based on monthly rain gauge data. Data are provided on a 0.25° horizontal grid and at 3-hourly intervals from 1998. The observation years used to compare to model simulations are noted in the relevant figure captions.

The CMORPH product uses precipitation estimates from the same set of passive microwave instruments as the TRMM dataset. However, it is different from the TRMM product in that it propagates these precipitation estimates using motion vectors, which are obtained from geostationary satellite infrared data. Finally, time-weighted linear interpolation is applied between forward- and backward-propagated precipitation estimates. Here we used the *bias-corrected* CMORPH product, which has calibrated the satellite-based precipitation estimates to gauge station data (Xie et al. 2017). Data are provided on an 8-km horizontal grid and at 30-min intervals from 1998. The observation years used to compare to model simulations are noted in the relevant figure captions.

Several previous studies have evaluated satellite-based precipitation estimates over parts of East Africa (Dinku et al. 2007; Diem et al. 2014; Cattani et al. 2016; Kimani et al. 2017). All of these studies compared the satellite rainfall products to either rain gauge data, or Global Precipitation Climatology Centre (GPCC) analysis of gauge data. The TRMM 3B42 rainfall estimates were identified in all of the studies as one of the best products at capturing the observed rainfall on seasonal and annual time scales. The CMORPH rainfall estimates were also found to perform well in the studies by Dinku et al. (2007) and by Kimani et al. (2017). Around complex orography of Ethiopia and Zimbabwe both TRMM and CMORPH were found to have a good combination of low bias (within 16%) and relatively strong correlation coefficient (at least 0.32 for Ethiopia and at least 0.47 for Zimbabwe) (Dinku et al. 2007). However, these correlation coefficients are still low, highlighting that there remain biases in rainfall estimates, especially around locations with steep orography such as Ethiopia (Dinku et al. 2007; Diem et al. 2014).

c. Subregions used in analysis

In cases where regional statistics are calculated, a selection of subregions is used (Fig. 1). Shown on this figure is the P25 orography, which is used throughout in other figures. These regions are used to broadly capture regions of East Africa with a bimodal annual cycle (Dunning et al. 2016) or those that are part of the key mesoscale system of the Lake Victoria basin.

The large region over the Greater Horn of Africa (orange) is based on the region defined by Rowell (2013), within which interannual anomalies are spatially coherent. The purple, blue, and red regions demarcate parts of the Lake Victoria basin not including grid cells with fractional land cover (i.e., $0 <$ land fraction < 1). Together they illustrate the basin as a whole, for which grid cells with fractional land are also included in the calculation. The brown box is used to provide a closer look at the Lake Victoria region, including much of the East African Rift Valley, for moisture flux analysis and extreme rainfall statistics. The full region used for the CP4 and P25 simulations encompasses all of the African continent and is shown in the Fig. 1 inset.

3. East African rainfall

The focus of this paper is to understand how the fundamentally different representation of the moist convection in the CP4 simulation affects the larger-scale dynamics and the water budget. First however, we compare the simulations to satellite-based rainfall

estimates to illustrate the model biases relative to these and the differences in rainfall between the models.

a. Evaluation against satellite-based rainfall products

Figure 2a shows the timing of peak rainfall in the TRMM rainfall product, with orange colors indicating an afternoon/evening peak, and purple indicating a nighttime/morning peak. Rainfall over the majority of East Africa typically peaks in the midafternoon (1500–1800 LT), with solar heating of the surface leading to a deepening boundary layer and reducing convective inhibition during the day. However, rainfall in several locations peaks during the night, for example, parts of Ethiopia, South Sudan, inland Somalia, and the African Great Lakes. These nighttime maxima are associated with both propagation of mesoscale convective systems and diurnal reversal of mesoscale circulations (Camberlin et al. 2018), such as that over Lake Victoria (approximately 1°S, 32°E). In addition, Lake Victoria exhibits a northeast–southwest gradient in timing, with peak rainfall to the northeast occurring around midnight and to the southwest peaking in the morning (0600–0900 LT). Figure 2e shows the difference in timing of peak rainfall between TRMM and CMORPH estimates, and indicates that over East Africa there is general agreement between the two rainfall products. In addition, Camberlin et al. (2018) have compared these features of diurnal cycle in the TRMM estimates to rain gauge measurements in East Africa, where available. They find that TRMM captures the important differences in nighttime and daytime peaks for the majority of gauge stations.

The biases in the timing of peak rainfall between each of the models and TRMM are shown in Figs. 2f and 2g. In general, biases are reduced with the CP4 simulation, particularly with regard to the broad early bias in afternoon rainfall peak across the region, evident in the P25 simulation. However, the nighttime peaks over northern Somalia and South Sudan are not captured by either model. Lake Victoria is the only location with an obvious increased bias of timing in the CP4 simulation compared to the P25 simulation. In this case, the CP4 rainfall peaks later than the TRMM rainfall. Finally, the CP4 simulation seems to reduce the P25 late bias over parts of the Kenyan and Ethiopian highlands. This may be a result of both increased resolution and explicit convection in the CP4 model.

Another focus of past study using convection-permitting simulations has been extreme rainfall rates (Kendon et al. 2014). Figure 3 shows the fractional contribution of each 3-hourly rainfall rate to total rainfall over the East Africa Rift Valley region (brown box on Fig. 1) in TRMM, CMORPH, and the two models. There is clearly some uncertainty in the satellite-based estimates of rainfall with regard to the contribution of

the different rain rates. For instance, CMORPH has a higher contribution than TRMM from rain rates up to $1 \, \text{mm} \, \text{h}^{-1}$. However, the two satellites are closer than the two models and therefore provide a useful range of the fractional contribution to evaluate the two models against.

It is clear that the contribution from extreme rain rates is much greater in the CP4 simulation than the P25 simulation, and that this is broadly consistent with the TRMM and CMORPH data. A similar picture emerges across different seasons, and over the smaller domain of Lake Victoria basin (not shown). Therefore, the results here support the conclusions of previous literature (e.g., Ban et al. 2014) regarding the improved representation of the intensity distribution of rainfall when using explicit convection simulations.

As well as a better representation of the fractional rainfall contribution than the P25 simulation, the CP4 simulation also better simulates dry 3-h periods $[<0.125 \, \text{mm} \, (3\text{h})^{-1}]$ when compared to the satellite estimates. The fraction of dry 3-h periods in the satellite estimates are 0.89 and 0.82 for TRMM and CMORPH, respectively. The fraction for the CP4 simulation is 0.85, and for the P25 simulation is 0.71. This highlights the tendency for too much light rain in the parameterized convection model, whereas CP simulations tend to have more dry periods accompanied by more intense rainfall when it does rain.

Broadly speaking, rainfall in much of East Africa is concentrated into two rainy seasons. Different names and definitions exist for these seasons including their comprising months. In a recent review, Nicholson (2017) provides discussion on this topic. Due to the variability in intraseasonal dynamical processes, the majority of our study considers individual months or annual averages. An exception is the following analysis of seasonal rainfall in which we employ a widely used definition of the rainy seasons: the "long rains" of March–May and the "short rains" of October–December.

Figure 4 shows the mean seasonal rainfall distribution in the TRMM rainfall estimates along with differences with the CMORPH rainfall product, and biases in the models. Peaks in the north [Ethiopia and South Sudan; June–September (JJAS)] and south [Tanzania and Mozambique; January–February (JF)] of the domain mark the extent of the tropical rainband. For this study we primarily consider the region in between where rainfall peaks biannually, in approximately March–May and October–December.

Potential uncertainty in the TRMM rainfall retrievals is considered by comparing to an alternative satellite-based rainfall product, CMORPH, which uses an alternative methodology for estimating rainfall. In the austral summer (JF), there seems to be uncertainty over
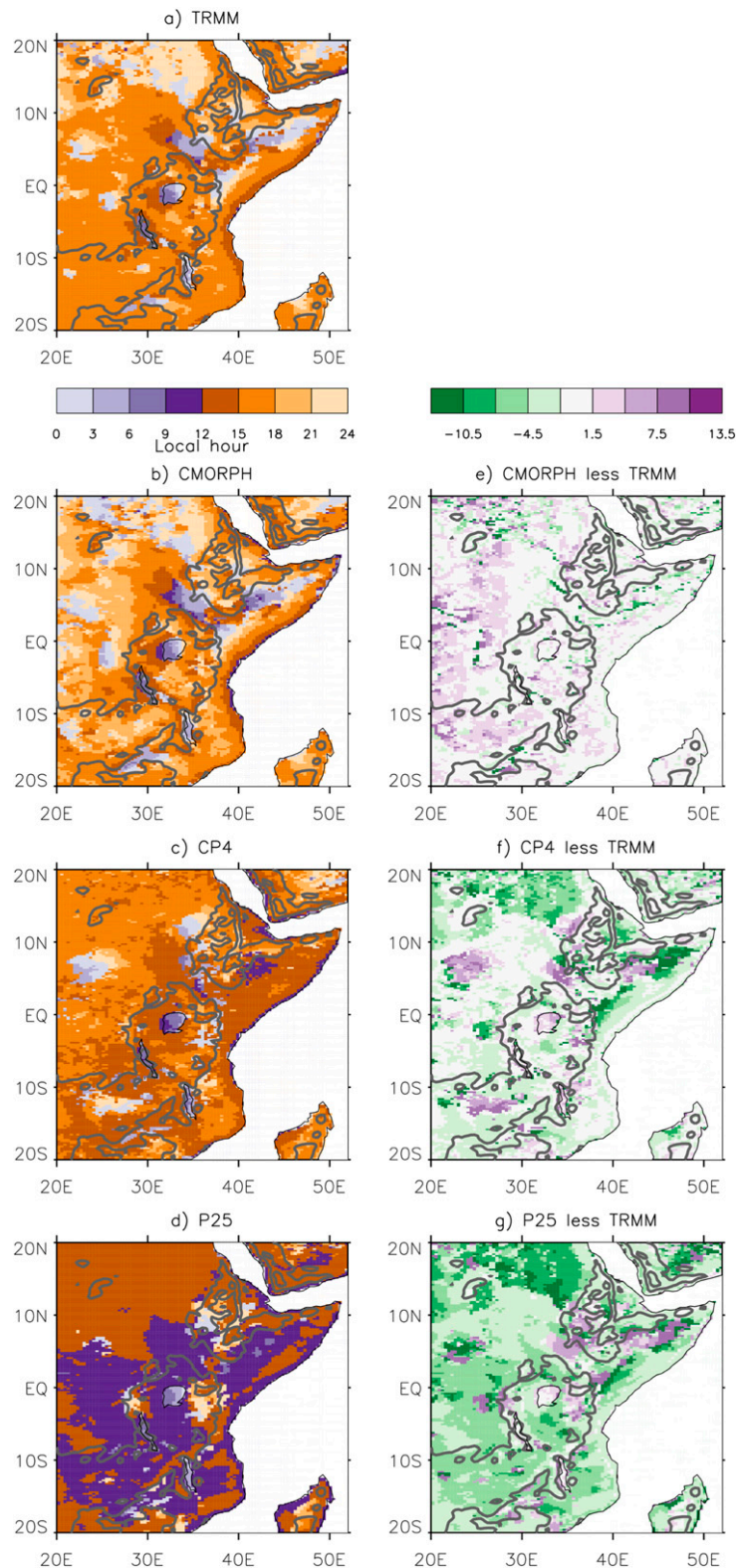
FIG. 2. Peak local time of rainfall in the mean diurnal cycle of (a) TRMM, (b) CMORPH, (c) CP4, and (d) P25 rainfall and differences in local peak time between TRMM and each of (e) CMORPH, (f) CP4, and (g) P25. Local time is calculated on a longitude basis. Orography lines are shown in gray for 1000 and 2000 m. All panels use data for January 1998–December 2006.
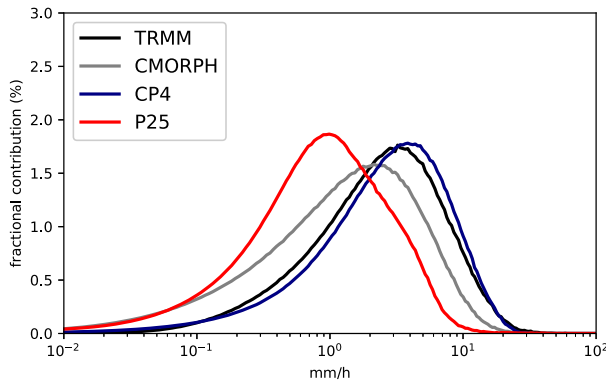
FIG. 3. Fractional contribution of each 3-hourly rainfall rate to total rainfall for all times over the East Africa brown box shown in Fig. 1. The figure uses data for March 1997–February 2007 for the models and March 1998–February 2008 for the satellite products.

the south of the region, where CMORPH suggests there could be around 1.2–6.0 mm day^{-1} less rainfall than TRMM, which shows absolute values of approximately 4.8–12.0 mm day^{-1}. On the other hand, the CMORPH data show higher rainfall, in all seasons, over the eastern Congo along the mountains of the western rift. In addition to these uncertainties between the satellite products, Dinku et al. (2007) has highlighted the challenge for satellite products to accurately determine precipitation over complex orography. The difficulty likely arises due to the occurrence of warm orographic rain. Both infrared and passive microwave detectors face challenges in providing retrievals that can accurately estimate such rainfall. Thus the above caveats need to be considered when we evaluate the seasonal rainfall in each model against TRMM.

In general, the CP4 simulation has higher rainfall than the satellite products and the P25 simulation. For instance it increases rainfall biases around Madagascar that are already apparent in the P25 simulation. On the other hand, over the western parts of Lake Victoria, the two simulations show opposite biases compared to the TRMM product, with the CP4 (P25) simulation overestimating (underestimating) rainfall. Finally, both simulations exhibit biases, compared to TRMM, over the Ethiopian highlands in JJAS, but due to the difficulty of satellite measurements in this region, we do not give much weight to this comparison. The difference between models in these areas is likely due to a combination of improved representation of convection and higher resolution of the CP4 model. Overall, there is moderate correlation of biases between the two models, with a mean Pearson spatial correlation coefficient across the seasons of 0.60. The weakest correlations are in the long rains and boreal summer (MAM: 0.53; JJAS: 0.51) and the strongest correlation is in austral summer (JF: 0.74).

While it is clear that neither simulation completely reproduces the satellite-based estimates of seasonal rainfall, they do capture the broad features across East Africa and are useful tools for understanding the role of convection. For instance, both have higher rainfall within the East African Rift Valley than over the Greater Horn of Africa, and both capture the biannual progression of the tropical rainband. The opposing biases in precipitation over Lake Victoria and the surrounding basin provide a useful modeling case study, especially due to the importance to the fast-growing population in the region.

b. Water budget and large-scale dynamics

As is apparent from Fig. 4, the local changes in representation of convection have led to large-scale changes in rainfall distribution. To understand these large-scale rainfall differences in more detail, we examine the differences in monthly mean moisture budget terms over East Africa (Fig. 5). The combination of evaporation and precipitation ($E - P$) provides a useful measure for whether a region is a source or sink of moisture, while the moisture flux illustrates the role of atmospheric dynamics in transporting moisture from source to sink. Note though that this is not equivalent to a full trajectory analysis to relate sources and sinks. The moisture flux is the product of the specific humidity and the zonal or meridional wind component, calculated at each time step and averaged each hour.

The four months used in Fig. 5 show the biannual progression of the tropical rainband where net precipitation (blue) occurs, while the majority of net evaporation (red) occurs over the Indian Ocean. In addition, the monsoonal winds of the Somali low-level jet are evident in the moisture flux, as a strong northeasterly flow in January, transitioning through April to a strong southeasterly flow in July, then completing the cycle through the November transition. There is also a strong southeasterly moisture flux, consistent throughout the year, between the East African highlands and Ethiopian highlands. This is collocated with the Turkana jet (Kinuthia and Asnani 1982; Nicholson 2016a), and highlights the clear role of orography on atmospheric dynamics over East Africa, something that has been investigated in previous studies (Slingo et al. 2005; Ogwang et al. 2014).

There are substantial differences between the two simulations. The convection-permitting scheme generates less rainfall over the Indian Ocean along the Kenyan coastline, as can be seen in April (Fig. 4). In the moisture budget this is manifested as increased net evaporation (positive $E - P$) seen in Fig. 5, which can also be seen in January, and to some extent November. Moisture transport from these locations to East Africa does not markedly increase, and in fact the difference in
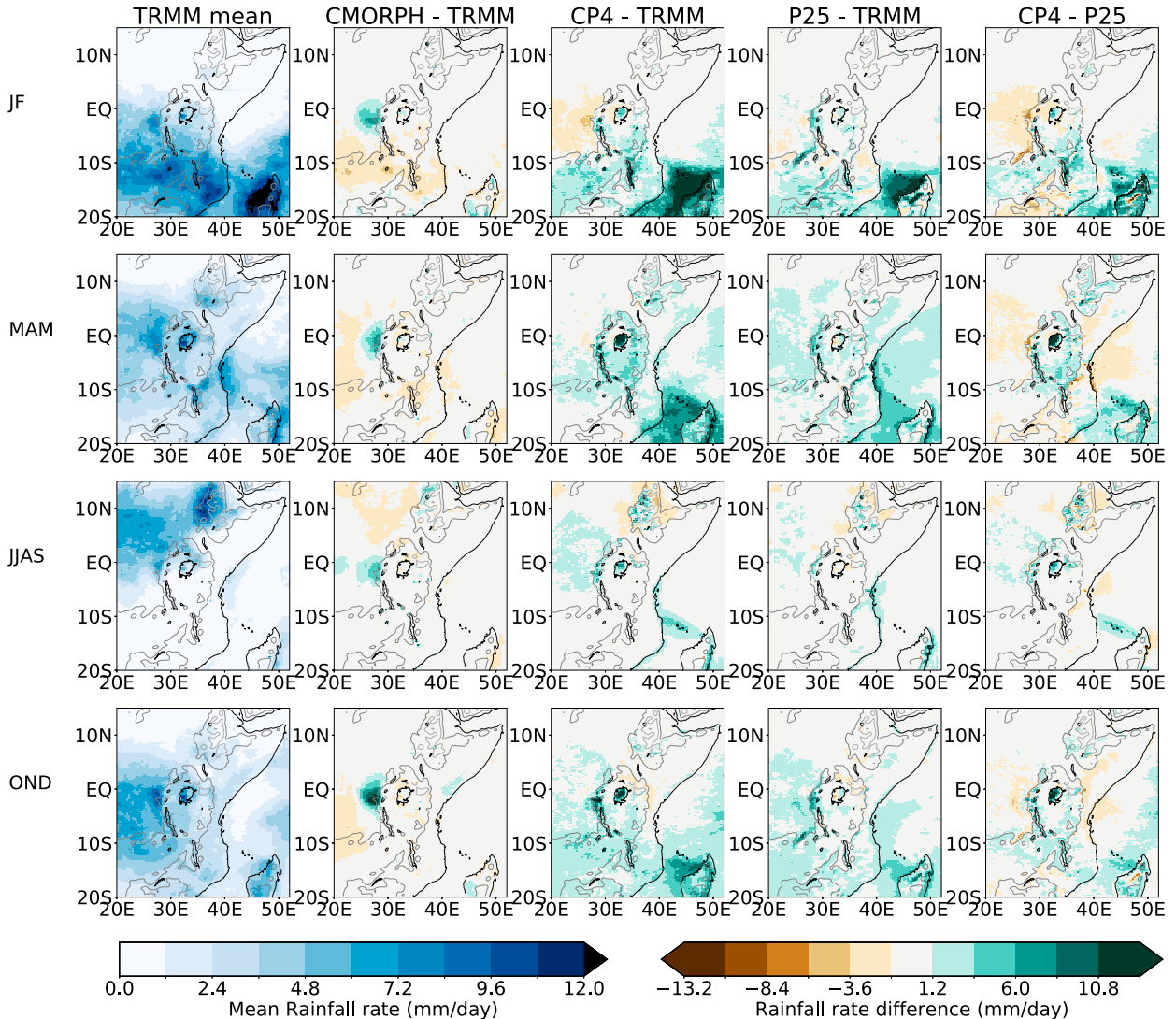
FIG. 4. Seasonal mean rainfall in TRMM and differences with CMORPH, CP4, and P25. Orography lines are shown in gray for 1000 and 2000 m. All panels use data for January 1998–December 2006.

moisture flux is offshore. In January in particular, it would seem that much of the additional net evaporation along the coastline is flowing toward Madagascar, where there is much larger net precipitation (negative $E - P$) in the CP4 simulation, possibly due to enhanced ocean rainfall biases in the CP model. Despite this, there is an increase in net precipitation over much of East Africa, specifically in the East African Rift Valley and over the East African Great Lakes. The larger net precipitation in the Rift Valley in the CP4 simulation is accompanied by a decrease in net precipitation over the Congo basin in January and April (also in absolute precipitation difference in Fig. 4).

Anyah et al. (2006) found that the rainfall over Lake Victoria was highly dependent on the easterly moisture

flux from the Indian Ocean. In Fig. 6 we show the moisture fluxes through each side of a rectangle containing the East African Rift Valley (Figs. 1 and 5, brown box). This shows that through the first half of the year the CP4 simulation has much less of an easterly moisture flux entering the region compared to the P25 simulation, as well as a much lower westerly moisture leaving the region. In April especially, which is the month showing some of the greatest differences in moisture budget terms (Fig. 5), there is a higher southerly moisture flux into the basin, which can be seen in the moisture flux arrows in Fig. 5. There is also not the same level of moisture leaving the region to the north. Differences in the easterly and westerly moisture fluxes largely cancel out. Overall, these features lead to the higher net moisture flux into the basin
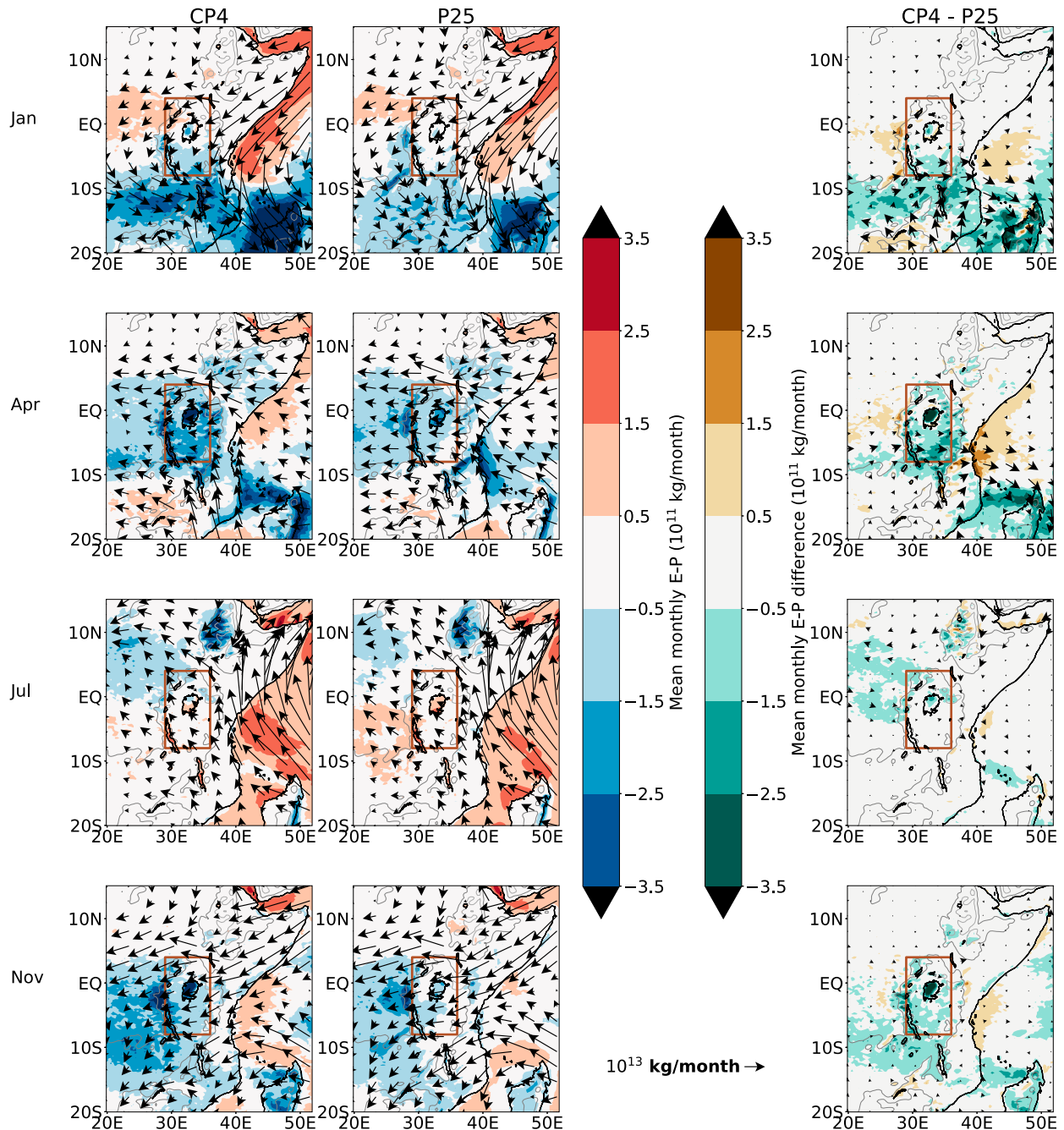
FIG. 5. (left),(center) Monthly mean $E - P$ and vertically integrated moisture flux (arrows). Orography lines are shown in gray for 1000 and 2000 m. (right) Differences between the two simulations. The brown box used is repeated from Fig. 1 and is added to assist with relating this figure with Fig. 6.

during April (black) and therefore the larger net precipitation in the CP4 simulation compared to the P25 simulation (Fig. 5). A bootstrap resampling of the 10 yearly differences in northerly, southerly, and net fluxes in April shows the mean differences of these terms to be significant at the 5% level. In the latter half of the year, the mean fluxes (Fig. 6) from the two simulations are much

more similar, with the exception of the timing of peak net moisture flux during the short rains, which occurs as a distinct November peak in the CP4 simulation, but as similar October and November fluxes in the P25 simulation. This coincides with a reduced outflow through the western boundary in the CP4 simulation. A bootstrap resampling of the 10 yearly differences in westerly and net
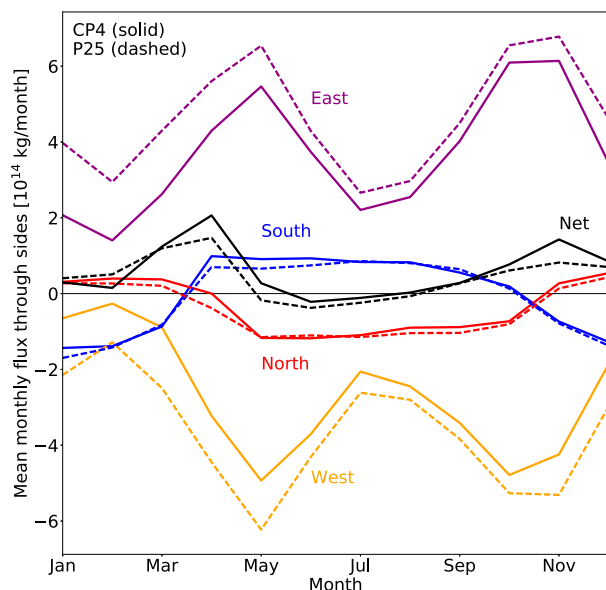
FIG. 6. Climatological monthly mean moisture flux from the directions north (red), south (blue), east (purple), and west (orange) into the East Africa Rift Valley region shown as a brown box in Figs. 1 and 5. The black line is the total of the four sides, where a positive value represents moisture convergence in the region. The CP4 simulation is given as solid lines and the P25 simulation as dashed lines.

flux in November shows the mean differences to be significant at the 5% level.

Annual total values of the budget terms for the two simulations over two East African regions are shown in Table 1. The two regions considered are 1) a region over the Greater Horn of Africa developed by Rowell (2013) as an area within which interannual variability of rainfall is highly correlated between different locations (orange region in Fig. 1), and that is also similar to the region of bimodal annual rainfall (Nicholson 2014; Dunning et al. 2016), and 2) the Lake Victoria basin (the combination of red, purple, and blue regions in Fig. 1). Values are independent of the size of the region as they have been divided by the region area. As such, we note the much greater rainfall per meter squared within the Lake Victoria basin compared to the Horn of Africa. Evaporation and moisture flux convergence (MFC) terms are also shown, as well as the imbalance of the budget. The

MFC term is calculated as the negative of the divergence of the moisture flux vectors (e.g., vectors in Fig. 5). The imbalance term demonstrates the modeling uncertainty in the three terms that should balance over the course of the year, but that do not because the semi-Lagrangian advection used in the Unified Model does not enforce moisture conservation. While a moisture conservation adjustment is made to the CP4 domain as a whole, as highlighted in section 2, it does not mean moisture will be conserved locally. The budget imbalance is not negligible. However, it is sufficiently small that broad conclusions can be drawn from the differences in the budget terms between the models. Care should be taken with absolute values of each term given the uncertainty. By relating these differences in water budgets to atmospheric processes, we can gain a greater insight into the role of moist convection in East Africa.

On the whole, the two models have similar water budgets for the Greater Horn of Africa (Table 1). This follows from the small differences seen over the Greater Horn region regarding seasonal mean precipitation (Fig. 4) and moisture fluxes (Fig. 5). The largest difference in budget terms occurs in the Lake Victoria basin, where precipitation is approximately a third higher in CP4 than P25. Differences in moisture flux convergence contribute to this difference in precipitation, though this difference is of the same order as the imbalance listed. The terms illustrate the greater role of moisture flux convergence within the CP4 simulation over the Lake Victoria basin. The difference in rainfall caused through use of explicit convection, as well as the importance of rainfall to the lake's own water budget and the societal importance, motivates the more detailed study of atmospheric processes in the basin in the following section.

4. Lake Victoria basin

The Lake Victoria basin represents a prime example of combining atmospheric drivers of convection. Namely, the diurnal lake breeze reinforces the upslope, anabatic flow generated by the surrounding mountains leading to convergence over the mountains in the east and west. There is then a tendency for storms to propagate with an easterly component driven by the African easterly jet at around

TABLE 1. Table of annual total water budget terms over two regions of East Africa for each simulation ($\times 10^9 \, \mathrm{kg \, m^{-2} \, yr^{-1}}$). Here, P is precipitation, E is evaporation, and MFC is moisture flux convergence. The regions used are shown in Fig. 1.

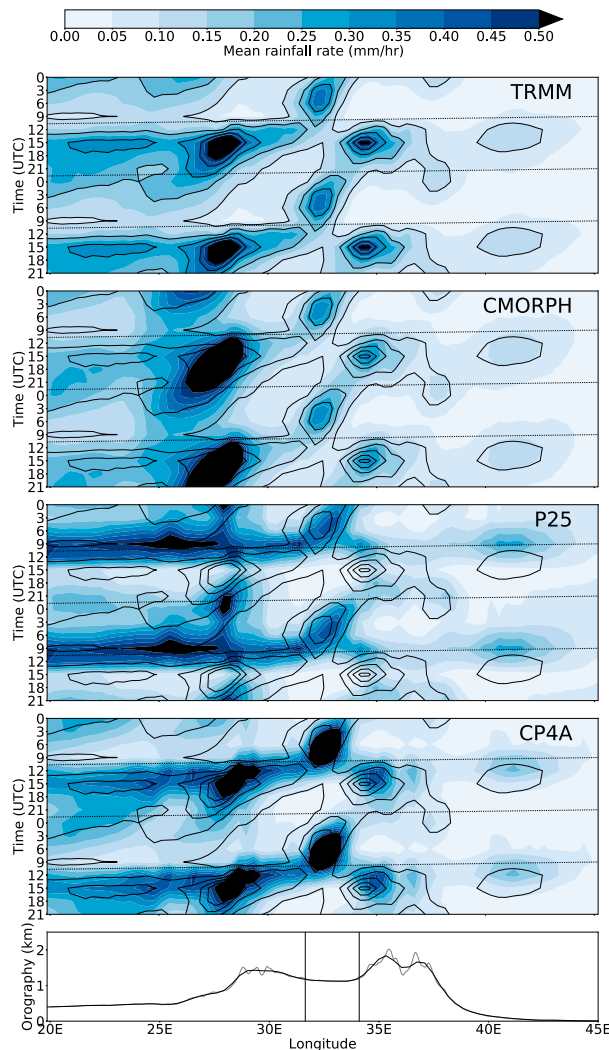| Region	Simulation	P	E	MFC	Imbalance (E + MFC − P)
East Africa–Horn of Africa (as in Rowell 2013)	CP4	0.6	0.5	0.3	0.2
	P25	0.6	0.6	0.1	0.1
East Africa–Lake Victoria basin	CP4	1.9	1.2	0.4	−0.3
	P25	1.4	1.2	0.2	0.0

FIG. 7. Mean 3-hourly rainfall rate from 3°S to 2°N. Mean orography on the N512 grid (black) and on the 4.5-km grid (gray) are shown in the bottom panel. Vertical lines in the orography panel show the approximate western and eastern longitudes of Lake Victoria. Black contour lines overlay the TRMM results in the panels for the other datasets. Dotted lines show local midday and midnight. All panels use data for January 1998–December 2006. The diurnal cycle is repeated once.

600–700 hPa. Overnight, the reverse nocturnal land breeze is reinforced by the downslope, katabatic winds. Mountains to the east and west of the lake produce katabatic winds that combine with the lake–land breeze at night, generating new storms, and influencing propagation of daytime storms from the land (Anyah et al. 2006; Williams et al. 2015). Figure 7 evaluates this formation and propagation of rainfall features within the basin with Hovmöller plots of diurnal mean rainfall.

The TRMM satellite product exhibits the east and west generation of rainfall over orography in the

afternoon with easterly propagation into the night, and peaking rainfall over the lake in the early morning. A peak in rainfall between 40° and 45° longitude can also be seen, which shows coastal rainfall associated with the sea breeze. The CMORPH rainfall product, which uses an analysis method using propagation vectors from geostationary satellite infrared measurements, generally corroborates the TRMM propagation features.

With the P25 simulation, the unrealistic midday convection and rainfall is very clear. A second peak in rainfall occurs over the eastern mountains during the evening, but this occurs later than the observed rainfall peak. The parameterized convection achieves a morning peak over the lake though this is offset to the east of the lake, as was clear from Fig. 4.

Employing the CP4 model, with explicit convection and higher resolution, allows improved timing of rainfall over the mountains as well as propagation of rainfall features. This is consistent with improved storm propagation in storm-tracking analysis of the CP4 model compared to the P25 model over West Africa (Crook et al. 2018, manuscript submitted to *Earth Space Sci.*). The CP4 model better captures the peak in rainfall over the eastern side of the basin, from where some eastward propagation of rainfall is also apparent. Over the lake, the high rainfall and slight late bias is evident in the CP4 data compared to TRMM, as identified in Figs. 2 and 4. Finally, the afternoon peak in the CP4 simulation at ~37° longitude is not as apparent in any of the other panels of Fig. 7. The gray and black orography lines show this is an area where the model resolution is significantly impacting the model orography. Meanwhile, it is possible that biases in the satellite products, or their coarser resolution than CP4, means that they also do not clearly identify the localized rainfall peak. For instance, Guilloteau et al. (2017) have shown that the effective resolution of the TRMM microwave sensor, which is used within the TRMM and CMORPH rainfall products, is 40 km or coarser around the highlands in the Lake Victoria basin. Since this is roughly the longitudinal resolution of Fig. 7, it may be pushing the ability of the satellite to resolve such small-scale features. As such, we suggest that the higher resolution of the CP4 model could be providing new information on rainfall in this mountainous region, beyond the capability of satellite rainfall products. Results here are largely independent of month, for which the main difference is in amplitude of rainfall not the direction or speed of propagation of rainfall features.

a. Annual water budget and cycle

In consideration of the complementary, but distinct, dynamical behavior in the basin, we split the basin into

TABLE 2. As in Table 1, but for annual total water budget terms over three parts of the Lake Victoria basin ($\times 10^9$ kg m^{-2} yr^{-1}). The regions used are shown in Fig. 1.

Region	Simulation	P	E	MFC	Imbalance (E + MFC − P)
Lake Victoria	CP4	4.5	1.8	2.2	−0.5
	P25	1.9	1.4	0.3	−0.2
Eastern basin land	CP4	1.5	1.0	0.2	−0.3
	P25	1.3	1.1	0.3	0.1
Western basin land	CP4	1.4	0.9	0.1	−0.4
	P25	1.3	1.1	0.2	0.0

three parts: the lake, eastern land, and western land (as shown in Fig. 1). Table 2 gives the annual water budget terms for these three regions.

Over the lake, the CP4 simulation generates more than twice as much rainfall as is evaporated from the lake. The moisture for this additional rainfall is provided through moisture convergence. In contrast, in the P25 simulation there is close to no moisture convergence over the lake, and since evaporation is similar to that in the CP4 simulation, the P25 generates much lower on-lake rainfall. Over the land, moisture convergence and evaporation within the basin are similar in the two simulations.

In Fig. 8, water budget terms for the three basin parts are broken down into monthly mean values. The dashed black line shows the sum of the budget terms in each month, this should equal the change in total column water (TCW) (black), so where this deviates indicates modeling uncertainty in the budget terms. Also, as a consequence, the dashed black line will not be visible where this matches the change in TCW.

Other than differences in magnitude already discussed, the annual cycle in rainfall is reasonably similar between models and the satellite datasets. The main timing difference between the models and satellite is the timing of peak rainfall in the long rains for the west side of the basin. Here, the models both simulate a clear April peak in rainfall, while TRMM and CMORPH show similar rainfall for March and April. This appears to be a systematic bias in the models. The main timing difference in rainfall between the two models is in the peak rainfall month during the short rains over the western basin, where P25 simulates similar October and November rainfall, while CP4, TRMM, and CMORPH all show a clear peak in November. The timing of moisture flux convergence is associated with this difference in peak month between the models, and was discussed for a larger region in Fig. 6 as the result of a shift to decreased westerly outflow in November in the CP4 simulation. This difference demonstrates how explicit convection can impact large-scale dynamics and in certain cases improve intraseasonal timing of rainfall in the model.

The cycle in evaporation is fairly similar between simulations. The P25 simulation does have a greater amplitude in the annual cycle of evaporation over land. This may be due to a greater drying out of the soils through less rainfall in the dry season and higher interception of rainfall by vegetation as a consequence of lighter rainfall intensities resulting in more rapid evaporation (Willetts et al. 2017).

In both simulations, the variation in monthly mean moisture convergence is much larger than the variation in evaporation, and similar in magnitude to the variation in monthly mean precipitation. Over the lake, the mean annual moisture convergence of the CP4 simulation spends fewer months with negative convergence than the P25. Furthermore, the moisture convergence dominates the evaporation term during the rainy seasons in the CP4 simulation but not in the simulation with parameterized convection. In the dry season (JJAS) the CP4 simulation generally generates more rainfall over western basin land areas than the P25 simulation despite a similar divergence of moisture. This higher rainfall in cases of large-scale moisture flux divergence is consistent with the ability of explicit convection to couple with mesoscale flows to overcome inhibition (Birch et al. 2014b,a).

b. Diurnal cycle in water budget

Figure 9 shows the mean diurnal cycle in budget terms over the three parts of the Lake Victoria basin. The main differences in the diurnal cycle of evaporation are in magnitude as already identified in Table 2, since the timing is largely constrained by the solar cycle. However, the substantial differences in the diurnal cycle of precipitation and moisture convergence between the simulations highlights the significant effect of explicit convection on mesoscale flows which parameterized convection struggles to capture (as seen in Fig. 7). In addition, on the diurnal time scale, changes in total column water become significant, and this can be seen in Fig. 9.

With parameterized convection (P25), the peak rainfall on the east side of the basin is mainly in time with peak evaporation. This is then followed by a secondary
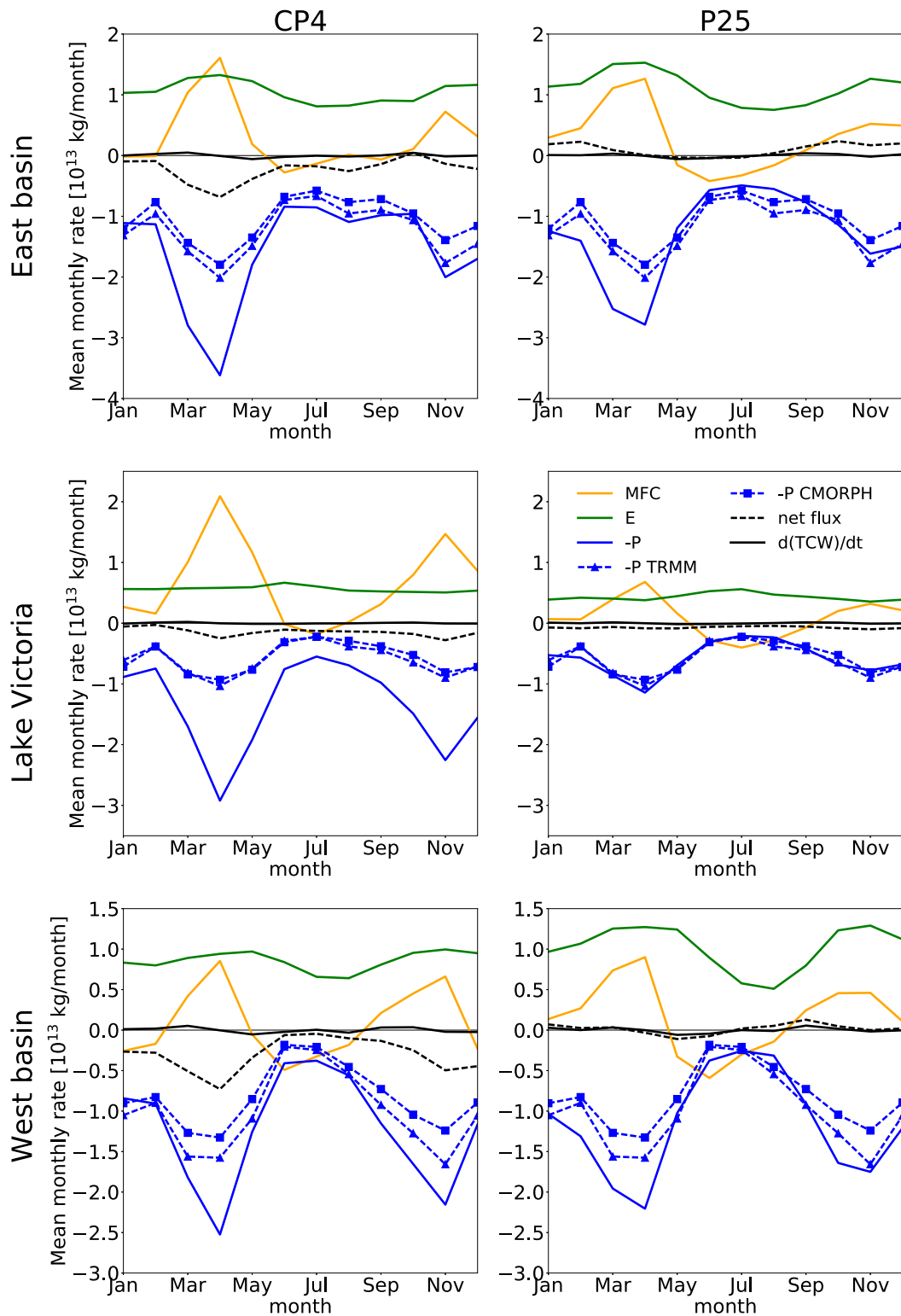
FIG. 8. Annual cycle of moisture budget terms, across the Lake Victoria basin. The budget terms are moisture flux convergence (MFC), evaporation E, and negative precipitation $-P$. Regions used are shown in Fig. 1. TRMM and CMORPH results use data for January 1998–December 2006.
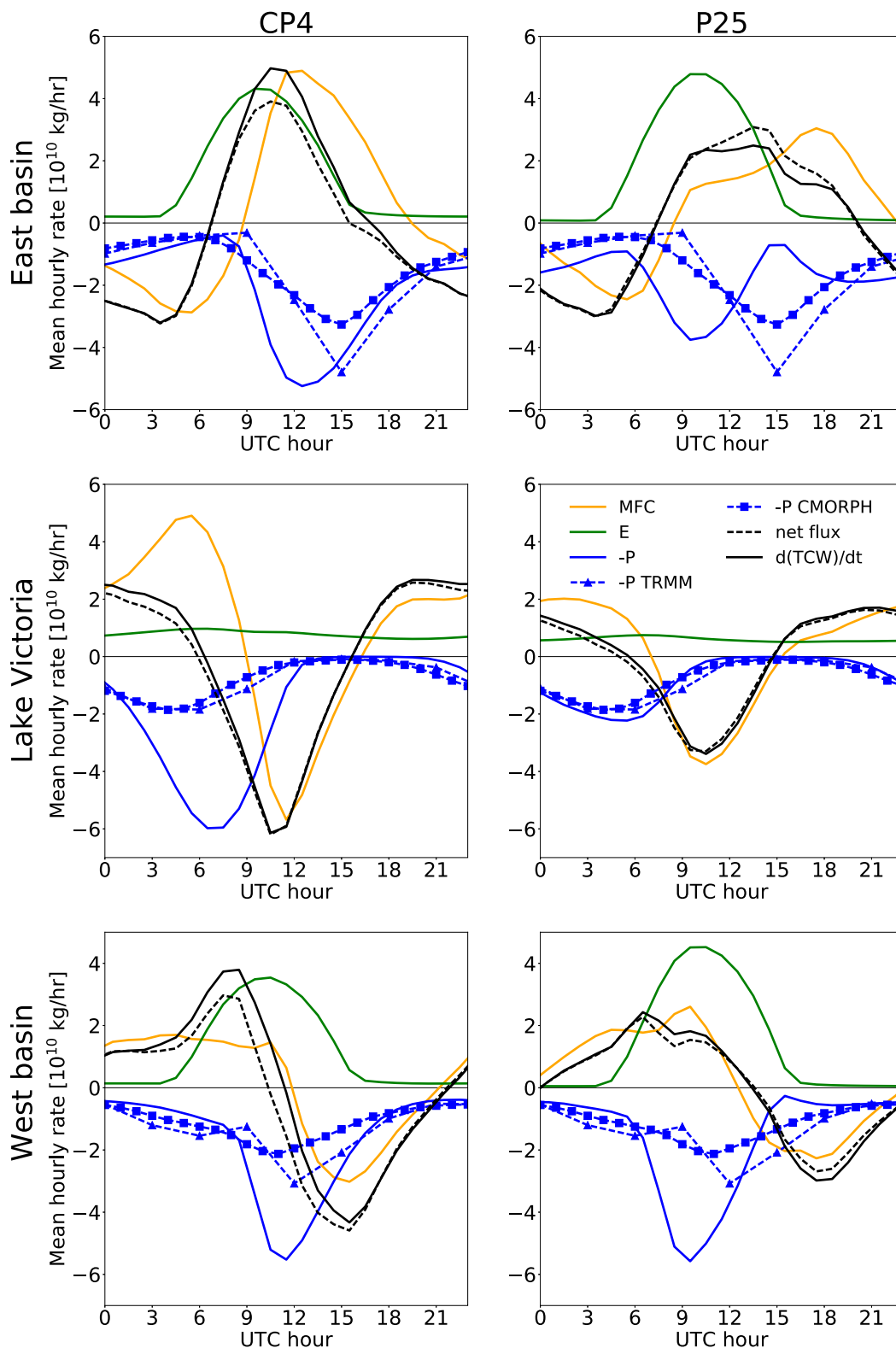
FIG. 9. Diurnal cycle of moisture budget terms, across the Lake Victoria basin region shown in Fig. 1. The budget terms are moisture flux convergence (MFC), evaporation E, and negative precipitation $-P$. TRMM and CMORPH results use data for January 1998–December 2006.

peak later in the evening, in time with peak MFC. Such biases in rainfall timing were also found in a recent study with a different model which also used a convection parameterization along with a much higher resolution of ~7-km grid spacing (Thiery et al. 2016, their supplementary Fig. 6). The simulated cycle of rainfall with the convection parameterization model is out of phase with the single afternoon rainfall peak seen in the satellite products. This suggests that the underlying timing of evaporation and MFC, or the convective triggers, are not adequate for simulating subdaily rainfall around this mountainous region to the east of Lake Victoria. Over the lake itself the rainfall diurnal cycle is much closer to the satellite estimates and can be seen to follow an increase of moisture convergence and total column water into the night (1500–0000 UTC). Over the western basin, increasing convergence in the morning in the P25 simulation does not appear to affect the timing in rainfall by as much as might be expected given the size of the increase. This possibly indicates the atmosphere is too stable to generate convection at this time of day. Instead the total column moisture increases over many hours, and rainfall occurs in time with peak evaporation, as it does over the eastern basin. Compared to satellite estimates, the rainfall peaks too high and early, followed by too little rainfall in the afternoon.

The CP4 simulation leads to two key improvements over the eastern side of the basin when comparing to the satellite estimates of rainfall: 1) a later peak in rainfall and 2) a single diurnal peak in rainfall, which is followed by a gradual reduction into the night. The increasing total column water in advance of precipitation in the CP4 simulation, as well as an earlier peak in MFC than in the P25 simulation, are prominent differences and provide sufficient moisture to maintain afternoon storms into the evening. The remaining early occurrence of the peak rainfall in CP4 compared to satellite estimates may be a consequence of early convective triggering, potentially affected by the use of stochastic perturbations of the boundary layer (section 1). Stratton et al. (2018) found in sensitivity studies without the stochastic perturbation that the diurnal precipitation cycle shifted later over land regions.

Over the lake itself, the CP4 simulation generates much higher precipitation than the satellite observations and the P25 simulation. The two simulations have similar values of MFC over the lake around midnight, but differ into the morning, with the CP4 simulation showing a large increase in MFC over the lake. This increase in convergence corresponds to increasing divergence over the eastern basin, though that is the case in both simulations. Over the western basin, the CP4 simulation does not show the increase in MFC which

occurs in the P25 simulation in the morning, suggesting that this may explain at least some of the difference in precipitation between the models over the lake. Finally, for CP4 over the western basin, evaporation seems to generate the increase in total column water over the western basin which precedes the peak rainfall. The peak rainfall within the CP4 simulation is too high, as with the P25 simulation, but occurs close to the time of peak rainfall in the satellite estimates, and does continue into the evening during which time the rainfall reduces too quickly in the P25 simulation.

As already described, the factors driving mesoscale flows in the basin are the anabatic/katabatic winds and lake/land breezes. We now consider these two drivers through analyzing the diurnal cycle in mean low-level wind convergence and virtual potential temperature θ_v (which determines the buoyancy of unsaturated air) in different parts of the basin (Fig. 10). The parts of the basin used are the lake itself, the land below 1250-m altitude (lowland), the land above 1500 m (highland), and air at a pressure level above the lake with a similar pressure (~800 hPa) to that of the high-altitude land surface (Fig. 10a). Note that the lake is already at approximately 1000 m above mean sea level. By calculating contrasts in buoyancy, we can compare the strength of the forcing for the land–lake breeze and anabatic/katabatic flows, since these mesoscale flows occur due to such contrasts. The low-level wind convergence should occur in different parts of the basin at different times of the day, depending on the timing of these mesoscale flows.

Both simulations exhibit the low-level convergence expected given the known mesoscale flows in the basin (Fig. 10b). During the nighttime, both have convergence over the lake and divergence over the mountains, then during the day convergence over the mountains and divergence over the lake. There is relatively little diurnal cycle in convergence over the land around the lake below 1500 m.

The difference in 1.5-m θ_v shows that there is relatively more buoyant air over the lake at night compared to the surrounding land below 1250 m, and vice versa during the daytime (Fig. 10c). This is a result of the diurnal cycle in cooling and heating of the land surface, while the diurnal cycle in air temperature above the lake is moderated by the higher heat capacity of the lake water compared to that of the land.

To consider the anabatic/katabatic effect we consider the 1.5-m θ_v of land above 1500 m, compared to θ_v of the air at 800 hPa above the lake and above land below 1250 m. The 800-hPa level is the closest pressure level available to the mean surface pressure of land above 1500 m (~820 hPa). We look at this quasi-horizontal
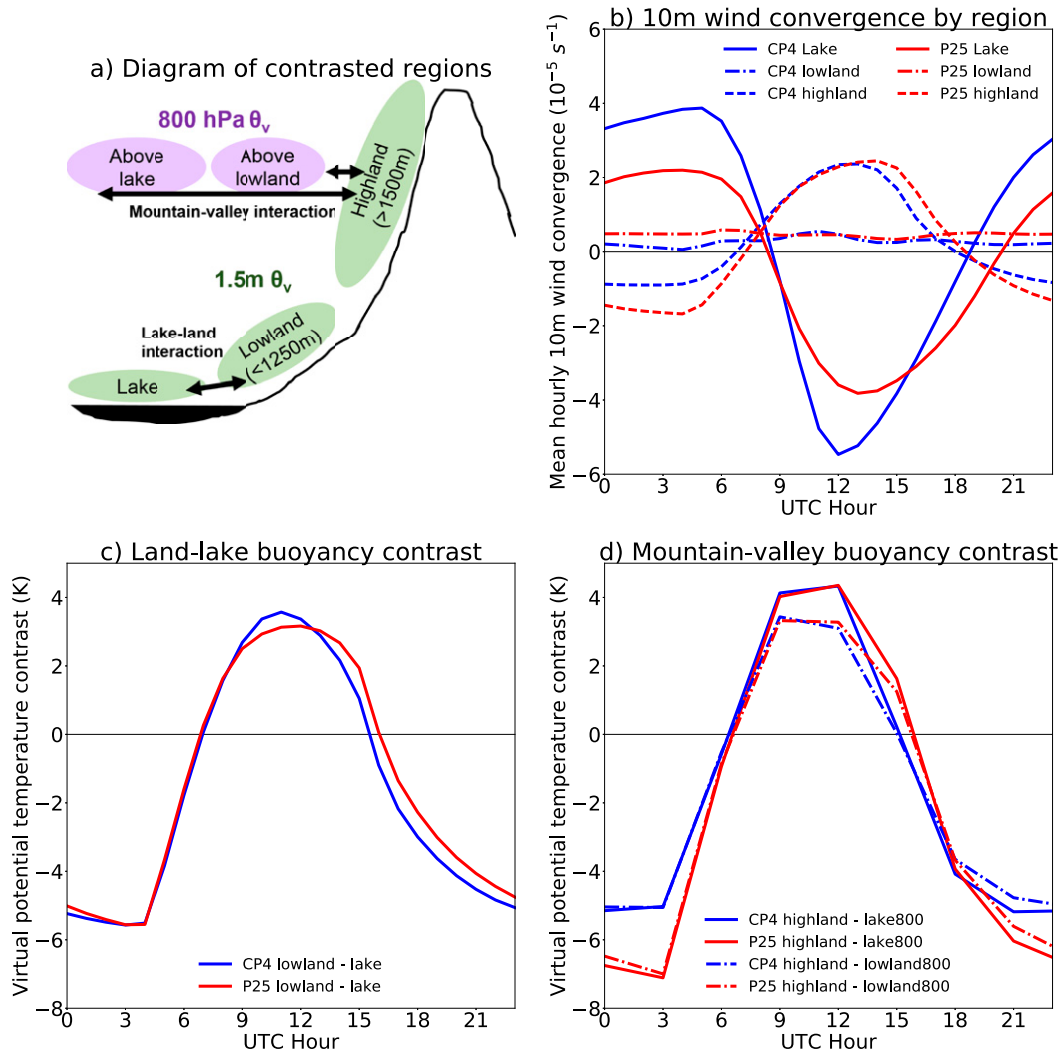
FIG. 10. (a) Diagram illustrating the relevant regions for land-lake and mountain-valley breezes, which are used in calculations for (b)–(d). (b) Diurnal cycle of 10-m wind divergence over Lake Victoria and the lowland and highland of the Lake Victoria basin and (c),(d) the contrasts in virtual potential temperature between these regions. Highland is defined as land over 1500 m above mean sea level, and lowland is below 1250 m above mean sea level. Where the virtual potential temperature has been taken at the 800-hPa level above an area, this is denoted by "800" in the legends. All panels use data for January 1998–December 2006. The temporal frequency is 3 hourly in (d), while (b) and (c) have hourly frequency.

contrast because it is the horizontal contrast in buoyancy that generates the up- and downslope winds. The contrast of the mountain 1.5-m θ_v with the 800-hPa air above the lowland, and above the lake, both show relatively more buoyant air over the mountains during the day and a reversal at night (Fig. 10d). This will drive horizontal flows into the mountain and upslope during the day, and vice versa at night. In both simulations there is also less of a buoyancy contrast during the day between the mountain and the air above the lowland, than with the air above the lake. This is because heating at the surface above all land will generate buoyant air to rise.

Over the lake the diurnal heating has less of an effect. As a result we see that daytime low-level divergence is much greater over the lake than the surrounding lowland (Fig. 10b).

While both models appear to simulate the necessary contrasts in buoyancy to generate mountain–valley and lake–land breezes, there are some clear differences. Nighttime low-level convergence over the lake is much greater in the model with explicit convection (Fig. 10b), and as a consequence generates the storms and higher precipitation seen over the lake in CP4 (Fig. 9). This additional convergence is generated by the high land–lake

contrast in θ_v (nighttime negative values on Fig. 10). The high nighttime land–lake θ_v contrast in CP4 results from generally cooler absolute temperatures over land (not shown) in the CP4 simulation, compared to the P25 simulation, which result from less incoming shortwave radiation due to reflection by clouds. Since the lake absolute temperatures are constant boundary conditions, the air above the lake is less affected by the reduced shortwave radiation. Therefore, at nighttime a greater land–lake absolute temperature contrast arises in CP4. On the other hand, the contrast in θ_v between the mountains and air above the lake and lowland, which generates the katabatic flow, is actually weaker in the CP4 simulation, so this is not contributing the additional convergence we see over the lake. The earlier drop-off in afternoon low-level convergence over the mountains in the CP4 simulation, compared to the P25 simulation, is apparent in the θ_v contrast between lake and land (Fig. 10c), and between mountain and valley (Fig. 10d). Therefore, both mechanisms are relevant for differences in daytime moisture convergence and rainfall.

Overall, the morning lake rainfall bias of the CP4 simulation compared to the TRMM satellite product (Fig. 9) implies that the convergence of the CP4 simulation at this time of day may be too strong (Fig. 10b). On the other hand, since the lake is ungauged, the satellite rainfall estimate does not have gauge-based measurements over the lake to reliably evaluate its own estimates. Subdaily observations of rainfall over the lake, and of the lake and mountain mesoscale flows are greatly needed for future work to evaluate this key aspect of regional models of East Africa. During the daytime, the CP4 simulation provides a much more realistic representation than the P25 simulation of the diurnal cycle of rainfall over the land in the Lake Victoria basin. As such, this implies that the convergence and buoyancy contrasts over land in Fig. 10 are more reliable, and this is also likely due to the improved representation of convection and the underlying increased resolution of the CP4 model.

5. Conclusions

We present the first analysis of water budgets over East Africa for an invaluable new dataset utilizing convection-permitting modeling. The data provide 10 years of high-resolution, convection-permitting simulations over the whole of Africa (Stratton et al. 2018). The analysis here evaluates the effect of representing convection explicitly and increased resolution upon rainfall and the water budget in East Africa. In addition, the effect of explicit convection on key continental and mesoscale dynamics is explored, with

particular focus on the lake–mountain flows in the Lake Victoria basin.

It is shown that, in general, findings in the literature regarding improved representation of rainfall intensities and the diurnal cycle of rainfall, when using explicit convection, hold true over much of East Africa. That includes the accurate representation of the proportion of dry 3-h periods with improved convection. In addition, reduced bias in the representation of the diurnal cycle is found in parts of the Kenyan and Ethiopian highlands suggesting that increased resolution is of benefit in these locations. However, the work highlights that nighttime rainfall peaks, where the propagation of storms is involved, are harder to capture. Future work over East Africa should look to understand the controls on these nighttime peaks in rainfall which have been highlighted here. Furthermore, it would be useful for future convection-permitting studies to distinguish between the effects on daytime and nighttime peaks in other locations too.

Despite the improvements to the features discussed above, the explicit model does not substantially reduce overall biases in the seasonal mean rainfall over East Africa. Moderate correlation between the biases of the explicit and parameterized convection model demonstrate that there are other aspects of the model which need to be addressed to remove such biases. However, some level of bias exists in the satellite observations themselves, since there is found to be some disagreement between satellite products, and there are known biases around mountainous regions. From a model perspective, these may be also related to other parameterizations that are still present in both models, such as parameterizations of the boundary layer or cloud microphysics. However, some improvements in the timing of monthly rainfall within the Lake Victoria basin have been identified. This highlights that the model is a tool, not a perfect representation of reality, which provides a new perspective when studying subdaily and storm-scale precipitation, as well as the effects of convection on the larger-scale dynamics.

The continental-scale impacts of explicitly representing convection have been explored. It is found that the explicit convection greatly modifies the sources, sinks, and transport of moisture over East Africa. For instance, a connection has been identified during the long rains (April) where the explicit convection leads to a more southerly flow into the Lake Victoria basin, associated with increased moisture convergence and rainfall. Rainfall and moisture convergence are intimately related, so disentangling cause and effect requires further investigation but our work provides an example of how local-scale changes in convection can impact the larger-scale climate. The higher April rainfall over the Lake Victoria basin reduces the moisture flux

westward to the Congo that, due to decreased rainfall, becomes a smaller sink of moisture to the atmosphere in the explicit convection model than in the parameterized. In this case, the Congo rainfall bias present with the parameterized convection simulation is not present in the explicit convection simulation. This example highlights the continental-scale effect that simulating convection explicitly can have on the water cycle and the opportunity it provides in understanding sources of model bias.

As well as continental-scale effects on the water budget, there are several key differences within the Lake Victoria basin. To comprehensively evaluate the mesoscale system in the Lake Victoria basin, much more detailed observations are needed than currently available. Since the region is a key focus for the HyCRISTAL project, surface flux towers in Kenya and boat observations on Lake Victoria are being implemented, and will be a valuable resource for future studies. However, in this study we have focused on describing and contrasting the mesoscale flows in convection-permitting and parameterized convection models, and relating this to differences in rainfall.

We show that the daily mesoscale processes, such as the land–lake breeze, that can trigger storms are greatly affected by the explicit representation of convection. The explicit convection shifts the rainfall diurnal cycle to unimodal from bimodal over the eastern Lake Victoria basin, which is much closer to the diurnal cycle in the satellite rainfall products than is the rainfall simulated with parameterized convection. On the other hand, over the lake itself, the explicit convection generates very strong convergence and likely unrealistically high rainfall in the morning. However, since there are no regular gauge measurements of rainfall over the lake, we cannot assume that satellite estimates are completely reliable in this location. In addition, there may be biases associated with the prescribed lake surface temperatures. The CP4 climate simulation already has a much improved resolution and representation of lake surface temperatures than almost all climate simulations of the region. However, the work here has highlighted that further detailed study of these aspects is warranted given the strong control that Lake Victoria has over the rainfall in the region. Other studies have also investigated the diurnal cycle in mesoscale flows within the basin (Anyah et al. 2006; Williams et al. 2015), and find that they are a key component of rainfall generation in the region. As such, careful consideration should be given to the impact of large lakes within convection-permitting simulations when modeling the Lake Victoria region, and other regions in the world in which lakes have a substantial role in mesoscale atmospheric processes. Despite the challenges around Lake Victoria, this study has shown

that for the Lake Victoria basin and wider East Africa, the convection-permitting model offers a useful alternative realization of the climate through its fundamentally more physical atmospheric dynamics, compared to a parameterized convection model. It shows very different behavior in subdaily moisture budgets and rainfall, as well as large-scale changes in moisture fluxes and rainfall in the Lake Victoria and Congo basins. As such it offers a useful tool in understanding the underlying processes of East African climate and, with the use of future climate (~2100; RCP8.5) simulations recently completed, will greatly enhance our knowledge of future climate change over East Africa.

REFERENCES

Abel, S. J., and Coauthors, 2017: The role of precipitation in controlling the transition from stratocumulus to cumulus clouds in a Northern Hemisphere cold-air outbreak. *J. Atmos. Sci.*, **74**, 2293–2314, https://doi.org/10.1175/JAS-D-16-0362.1.

Anyah, R. O., F. H. M. 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.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the new convective-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res. Atmos.*, **119**, 7889–7907, https://doi.org/10.1002/2014JD021478.

Birch, C. E., 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.

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.

Cattani, E., A. Merino, and V. Levizzani, 2016: Evaluation of monthly satellite-derived precipitation products over East Africa. *J. Hydrometeor.*, **17**, 2555–2573, https://doi.org/10.1175/JHM-D-15-0042.1.

Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis*, Cambridge University Press, 1029–1136, https://doi.org/10.1017/CBO9781107415324.024.

Cook, K. H., and E. K. Vizy, 2013: Projected changes in East African rainy seasons. *J. Climate*, **26**, 5931–5948, https://doi.org/10.1175/JCLI-D-12-00455.1.

De Kauwe, M. G., C. M. Taylor, P. P. Harris, G. P. Weedon, and R. J. Ellis, 2013: Quantifying land surface temperature variability for two Sahelian mesoscale regions during the wet season. *J. Hydrometeor.*, **14**, 1605–1619, https://doi.org/10.1175/JHM-D-12-0141.1.

Diem, J. E., J. Hartter, S. J. Ryan, and M. W. Palace, 2014: Validation of satellite rainfall products for western Uganda. *J. Hydrometeor.*, **15**, 2030–2038, https://doi.org/10.1175/JHM-D-13-0193.1.

Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor, and C. F. Ropelewski, 2007: Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.*, **28**, 1503–1526, https://doi.org/10.1080/01431160600954688.

Dirmeyer, P. A., and Coauthors, 2012: Simulating the diurnal cycle of rainfall in global climate models: Resolution versus parameterization. *Climate Dyn.*, **39**, 399–418, https://doi.org/10.1007/s00382-011-1127-9.

Dunning, C. M., E. C. L. Black, and R. P. Allan, 2016: The onset and cessation of seasonal rainfall over Africa. *J. Geophys. Res. Atmos.*, **121**, 11 405–11 424, https://doi.org/10.1002/2016JD025428.

Gregory, D., and P. R. 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.

Guilloteau, C., E. Foufoula-Georgiou, and C. D. Kummerow, 2017: Global multiscale evaluation of satellite passive microwave retrieval of precipitation during the TRMM and GPM eras: Effective resolution and regional diagnostics for future algorithm development. *J. Hydrometeor.*, **18**, 3051–3070, https://doi.org/10.1175/JHM-D-17-0087.1.

Hook, S., R. C. Wilson, S. MacCallum, and C. J. Merchant, 2012: Lake surface temperature [in "State of the Climate in 2011"]. *Bull. Amer. Meteor. Soc.*, **93** (7), S18–S19.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

Kendon, E., N. Roberts, and H. Fowler, 2014: Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nat. Climate Change*, **4**, 570–576, https://doi.org/10.1038/nclimate2258.

Kimani, M., J. Hoedjes, and Z. Su, 2017: An assessment of satellite-derived rainfall products relative to ground observations over East Africa. *Remote Sens.*, **9**, 430, https://doi.org/10.3390/rs9050430.

Kinuthia, J. H., and G. C. Asnani, 1982: A newly found jet in north Kenya (Turkana Channel). *Mon. Wea. Rev.*, **110**, 1722–1728, https://doi.org/10.1175/1520-0493(1982)110<1722:ANFJIN>2.0.CO;2.

Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Wea. Rev.*, **128**, 3187–3199, https://doi.org/10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2.

MacCallum, S. N., and C. J. Merchant, 2012: Surface water temperature observations of large lakes by optimal estima-tion. *Can. J. Remote Sens.*, **38**, 25–45, https://doi.org/10.5589/m12-010.

Marsham, J. H., N. S. Dixon, L. Garcia-Carreras, G. M. S. 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.

Morcrette, C., H. Lean, K. Browning, J. Nicol, N. Roberts, P. Clark, A. Russell, and A. Blyth, 2007: Combination of mesoscale and synoptic mechanisms for triggering an isolated thunderstorm: Observational case study of CSIP IOP 1. *Mon. Wea. Rev.*, **135**, 3728–3749, https://doi.org/10.1175/2007MWR2067.1.

Nicholson, S. E., 2014: A detailed look at the recent drought situation in the Greater Horn of Africa. *J. Arid Environ.*, **103**, 71–79, https://doi.org/10.1016/j.jaridenv.2013.12.003.

——, 2016a: The Turkana low-level jet: Mean climatology and association with regional aridity. *Int. J. Climatol.*, **36**, 2598–2614, https://doi.org/10.1002/joc.4515.

——, 2016b: An analysis of recent rainfall conditions in eastern Africa. *Int. J. Climatol.*, **36**, 526–532, https://doi.org/10.1002/joc.4358.

——, 2017: Climate and climatic variability of rainfall over eastern Africa. *Rev. Geophys.*, **55**, 590–635, https://doi.org/10.1002/2016RG000544.

Ogwang, B. A., H. Chen, X. Li, and C. Gao, 2014: The influence of topography on East African October to December climate: Sensitivity experiments with RegCM4. *Adv. Meteor.*, **2014**, 143917, https://doi.org/10.1155/2014/143917.

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.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1.

Rowell, D. P., 2013: Simulating SST teleconnections to Africa: What is the state of the art? *J. Climate*, **26**, 5397–5418, https://doi.org/10.1175/JCLI-D-12-00761.1.

——, B. B. B. Booth, S. E. Nicholson, and P. Good, 2015: Reconciling past and future rainfall trends over East Africa. *J. Climate*, **28**, 9768–9788, https://doi.org/10.1175/JCLI-D-15-0140.1.

Semazzi, F., S. Benedict, and P. J. van Oevelen, 2012: A Regional Hydroclimate Project for Lake Victoria Basin (HYVIC). *GEWEX News*, Vol. 22, No. 2, International GEWEX Project Office, Silver Spring, MD, 9–12, https://www.gewex.org/gewex-content/files_mf/1432211007May2012.pdf.

Seto, K. C., B. Guneralp, and L. R. Hutyra, 2012: Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Natl. Acad. Sci. USA*, **109**, 16 083–16 088, https://doi.org/10.1073/pnas.1211658109.

Slingo, J., H. Spencer, B. Hoskins, P. Berrisford, and E. Black, 2005: The meteorology of the western Indian Ocean, and the influence of the East African highlands. *Philos. Trans. Roy. Soc.*, **363A**, 25–42, https://doi.org/10.1098/rsta.2004.1473.

Smith, R. N., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Quart. J. Roy. Meteor. Soc.*, **116**, 435–460, https://doi.org/10.1002/qj.49711649210.

Standard Media, 2017: Death toll rises in Mombasa with more heavy rains predicted. Standard Media, accessed on 4 November 2017, https://www.standardmedia.co.ke/article/2001239159/death-toll-rises-in-mombasa-with-more-heavy-rains-predicted.

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-Africa 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.

Thiery, W., E. L. Davin, S. I. Seneviratne, K. Bedka, S. Lhermitte, and N. P. M. van Lipzig, 2016: Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.*, **7**, 12786, https://doi.org/10.1038/ncomms12786.

Vanderkelen, I., N. P. M. van Lipzig, and W. Thiery, 2018: Modelling the water balance of Lake Victoria (East Africa)—Part 2: Future projections. *Hydrol. Earth Syst. Sci.*, **22**, 5527–5549, https://doi.org/10.5194/hess-22-5527-2018.

Walters, D., and Coauthors, 2017a: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geosci. Model Dev.*, **10**, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017.

——, and Coauthors, 2017b: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. *Geosci. Model Dev. Discuss.*, https://doi.org/10.5194/gmd-2017-291.

Willetts, P. D., J. H. Marsham, C. E. Birch, D. J. Parker, S. Webster, and J. Petch, 2017: Moist convection and its upscale effects in simulations of the Indian monsoon with explicit and parametrized convection. *Quart. J. Roy. Meteor. Soc.*, **143**, 1073–1085, https://doi.org/10.1002/qj.2991.

Williams, K., J. Chamberlain, C. Buontempo, and C. Bain, 2015: Regional climate model performance in the Lake Victoria basin. *Climate Dyn.*, **44**, 1699–1713, https://doi.org/10.1007/s00382-014-2201-x.

Wilson, D. R., A. C. Bushell, A. M. Kerr-Munslow, J. D. Price, and C. J. Morcrette, 2008: PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description. *Quart. J. Roy. Meteor. Soc.*, **134**, 2093–2107, https://doi.org/10.1002/qj.333.

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.

Xie, P., R. Joyce, S. Wu, S.-H. Yoo, Y. Yarosh, F. Sun, and R. Lin, 2017: Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. *J. Hydrometeor.*, **18**, 1617–1641, https://doi.org/10.1175/JHM-D-16-0168.1.

Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2014: The East African long rains in observations and models. *J. Climate*, **27**, 7185–7202, https://doi.org/10.1175/JCLI-D-13-00447.1.

——, ——, ——, and ——, 2015: The annual cycle of East African precipitation. *J. Climate*, **28**, 2385–2404, https://doi.org/10.1175/JCLI-D-14-00484.1.

Yin, X., and S. E. Nicholson, 1998: The water balance of Lake Victoria. *Hydrol. Sci. J.*, **43**, 789–811, https://doi.org/10.1080/02626669809492173.