

An evaluation of clouds and precipitation in convection-permitting forecasts for South Africa

Article

Published Version

Open Access

Stein, T. H. M., Keat, W., Maidment, R. I., Landman, S., Becker, E., Boyd, D. F. A., Bodas-Salcedo, A., Pankiewicz, G. and Webster, S. (2019) An evaluation of clouds and precipitation in convection-permitting forecasts for South Africa. *Weather and Forecasting*. pp. 233-254. ISSN 0882-8156 doi: <https://doi.org/10.1175/WAF-D-18-0080.1> Available at <http://centaur.reading.ac.uk/81689/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1175/WAF-D-18-0080.1>

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Ⓞ An Evaluation of Clouds and Precipitation in Convection-Permitting Forecasts for South Africa

T. H. M. STEIN, W. KEAT, AND R. I. MAIDMENT

Department of Meteorology, University of Reading, Reading, United Kingdom

S. LANDMAN AND E. BECKER^a

South African Weather Service, Pretoria, South Africa

D. F. A. BOYD, A. BODAS-SALCEDO, G. PANKIEWICZ, AND S. WEBSTER

Met Office, Exeter, United Kingdom

(Manuscript received 4 May 2018, in final form 13 January 2019)

ABSTRACT

Since 2016, the South African Weather Service (SAWS) has been running convective-scale simulations to assist with forecast operations across southern Africa. These simulations are run with a tropical configuration of the Met Office Unified Model (UM), nested in the Met Office global model, but without data assimilation. For November 2016, convection-permitting simulations at 4.4- and 1.5-km grid lengths are compared against a simulation at 10-km grid length with convection parameterization (the current UM global atmosphere configuration) to identify the benefits of increasing model resolution for forecasting convection across southern Africa. The simulations are evaluated against satellite rainfall estimates, *CloudSat* vertical cloud profiles, and SAWS radar data. In line with previous studies using the UM, on a monthly time scale, the diurnal cycle of convection and the distribution of rainfall rates compare better against observations when convection-permitting model configurations are used. The SAWS radar network provides a three-dimensional composite of radar reflectivity for northeast South Africa at 6-min intervals, allowing the evaluation of the vertical development of precipitating clouds and of the timing of the onset of deep convection. Analysis of four case study days indicates that the 4.4-km simulations have a later onset of convection than the 1.5-km simulations, but there is no consistent bias of the simulations against the radar observations across the case studies.

1. Introduction

The adoption of kilometer-scale models for numerical weather prediction (NWP) has improved the location and timing of forecast convective weather events (e.g., Clark et al. 2016). At horizontal grid lengths smaller than 5 km, NWP models can resolve the largest scales of individual convective clouds; these models are thus referred to as convection-permitting models (CPMs). Several operational forecasting centers currently run

CPM simulations multiple times per day over limited areas, typically nested inside a global climate model of coarser resolution. In the United States, these models are also referred to as convection-allowing models and have been used for several years now, for instance to assist in the forecasting of severe convective weather (Kain et al. 2006, 2008). In particular, the Warn-on-Forecast research program is exploring the use of an ensemble of CPMs to provide reliable severe weather warnings (Stensrud et al. 2009).

CPM simulations run on a global scale are within reach technologically and envisaged for operational weather prediction (Bauer et al. 2015). For research purposes, global CPMs are already used to study, for instance, the interactions between convection and tropical waves (Satoh et al. 2008). The challenge in evaluating global CPMs will be to develop diagnostic tools and verification metrics that are appropriate and reliable globally—considering the

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

^a Current affiliation: Centre for Climate Research, Singapore.

Corresponding author: T. H. M. Stein, t.h.m.stein@reading.ac.uk

discrepancies between different national observational networks—to allow consistent model verification and evaluation. As a proxy for a global CPM, a single CPM run over a limited area in several regions across the globe will meet a similar verification and evaluation challenge. The Met Office's Unified Model (UM) is used by several operational forecasting centers globally through the UM Partnership for numerical weather and climate prediction, including at convection-permitting resolutions. The UM Partnership consists of several organizations who use the UM in global and limited area configurations. Access to the latest UM science configurations and to a nesting suite enables partners to contribute to the Regional Model Evaluation and Development (RMED) process, through which the same science configuration is tested and evaluated by UM Partnership organizations in various convective weather situations. For individual partners, improvements in the forecasting of severe convective events can be showcased through the RMED process, while for the UM Partnership the development and use of diagnostic tools and verification metrics across the organizations enables succeeding scientific configurations to improve forecast skill and the fidelity of physical processes. Although the UM is not run operationally as a single global CPM, the RMED process meets similar challenges in verification as global CPMs, namely that verification tools can be applied consistently across the UM Partnership.

The South African Weather Service (SAWS) was the first partner on the African continent and has been using the UM for operational activities since 2007 and at convection-permitting scales since 2016. SAWS routinely runs limited area simulations with the UM to assist with forecast operations for South Africa and across southern Africa through the Severe Weather Forecasting Demonstration Project (SWFDP). Convection is a major cause of severe weather events in southern Africa. Laing and Fritsch (1993) identified on average 11 mesoscale convective complexes in MeteoSat imagery per summer season in southern Africa. In particular, the northeast of South Africa experiences some of the deepest thunderstorms on Earth (Zipser et al. 2006). This region is a local “hotspot” of lightning activity, producing more than 30 flashes $\text{km}^{-2}\text{yr}^{-1}$ (Albrecht et al. 2016). Weather hazards associated with these convective storms include large hail, tornadoes, and flash floods, with several such events occurring in November 2016 alone (Simpson and Dyson 2018).

The availability of ground-based radar data in South Africa (Terblanche et al. 2001) makes the SAWS forecasts a unique dataset to test the applicability and limitations of a variety of observational data to evaluate CPM simulations in challenging convective weather

situations. As part of the RMED process, this study therefore aims to test the applicability of different observational datasets to evaluate CPMs. Multiple satellite-based rainfall products are available for Africa [for an overview, see, e.g., Maidment et al. (2014)] but these typically do not resolve the subdaily scale and are therefore better suited for evaluation on longer time scales. Rainfall estimates from the Global Precipitation Measurement (GPM) mission are available every 30 min but are not yet thoroughly evaluated across Africa. Finally, *CloudSat* observations are ideally suited to study the vertical structure of clouds and precipitating systems (Bodas-Salcedo et al. 2008; Stein et al. 2011) but suffer from poor horizontal and temporal sampling. Nevertheless, the combined evaluation against these different observational data should lead to consistent conclusions regarding a model's performance in terms of convective clouds and precipitation.

This work therefore has three objectives, namely 1) to investigate whether model biases in the location and timing of clouds and precipitation are consistent across the different observational data and associated metrics, 2) to consider the limitations of the different observational data in evaluating clouds and precipitation in CPMs, and 3) to assess the performance of the model in terms of representing clouds and precipitation for southern Africa. While the latter is expected to largely confirm known differences between CPM simulations and simulations with convection parameterization in other regions of the world (e.g., Pearson et al. 2010), the first two objectives may increase our confidence in the use of individual observational datasets for evaluating convection in models when other data are not available, for instance across southern Africa more generally. The inclusion of GPM and *CloudSat* is pertinent as these—and their potential successors, including the Earth Cloud, Aerosol, and Radiation Explorer (EarthCARE)—are fundamental to the evaluation of global kilometer-scale simulations due to their near-global availability. The outcomes will therefore present recommendations for the future use of satellite and ground-based observations to evaluate convection in NWP models globally, as well as recommendations for the simulation of convection over southern Africa specifically.

The paper is organized as follows. Section 2 describes the different observational data that are used for model verification, namely the satellite rainfall estimates and the SAWS radar data. Section 3 describes the setup of the SAWS operational forecasting system and the different model configurations that are evaluated in this paper. The evaluation of monthly rainfall and cloud distribution for November 2016 as well as daily rainfall

amounts and the diurnal cycle are shown in [section 4](#). Evaluation against the SAWS radar data is presented in [section 5](#) for four case study days with heavy rainfall events and with good availability of radar data. [Section 6](#) provides a discussion and our conclusions, referring to our aforementioned objectives.

2. Observational data sources

a. Satellite rainfall estimates

In the absence of a routinely available and dense network of ground-based radar observations, satellite rainfall estimates can provide reliable observations for model evaluation. Spaceborne active remote sensing instruments provide the most detailed measurements of precipitating clouds, being able to resolve the vertical profile of hydrometeor distributions, but these instruments have narrow swaths and are as yet bound to polar-orbiting satellites, leading to poor spatiotemporal cover. We therefore consider several satellite rainfall products that use a combination of infrared and microwave measurements and that are therefore available on daily or shorter time scales. By adopting multiple satellite rainfall estimates, we can test whether our models forecast rainfall within the observed range ([Birch et al. 2014](#)). This analysis also allows us to consider the strengths and weaknesses of each product ([Jobard et al. 2011](#)), particularly for the purpose of model evaluation on different time scales.

1) TAMSAT

The Tropical Applications of Meteorology using Satellite and Ground-Based Observations (TAMSAT) Group, based at the University of Reading, Reading, United Kingdom, has provided satellite-derived rainfall estimates for Africa since the early 1980s. Estimates are produced in near-real time at 0.0375° horizontal grid spacing (approximately 4 km over South Africa) at daily, pentadal, decadal, monthly, and seasonal time steps and are available from 1983 onward. The TAMSAT method is designed to provide a temporally consistent rainfall record. It is based on cold cloud duration (CCD) maps derived from Meteosat thermal infrared (TIR) imagery and calibrated using historical rain gauge measurements ([Maidment et al. 2014, 2017](#); [Tarnavsky et al. 2014](#)). The resulting climatological calibration parameters, which vary both spatially and seasonally, are applied to the CCD in near-real time, overcoming the need for “live” rain gauge records that are typically unavailable across Africa. Recently, the TAMSAT estimation algorithm was revised to provide more accurate rainfall estimates ([Maidment et al. 2017](#); [Dinku et al. 2018](#)). It is these rainfall estimates (version 3.0) that are used here.

Because CCD itself is only loosely related to rainfall intensity (CCD is based only on the occurrence of cold clouds exceeding a prescribed temperature threshold), TAMSAT generally underestimates rainfall accumulations for extreme events over short time periods ([Maidment et al. 2017](#)). It may also miss rainfall events when these are caused by warm clouds that do not contribute to CCD. Despite these shortcomings, TAMSAT is widely used across Africa for rainfall monitoring and has informed previous NWP evaluation studies ([Maidment et al. 2013](#)).

2) IMERG

The Global Precipitation Measurement (GPM) mission is an international collection of weather satellites providing the most advanced observations of global precipitation to date. Building on from the Tropical Rainfall Measuring Mission (TRMM) era of rainfall observations (1997–2014), the Integrated Multisatellite Retrievals for GPM (IMERG) rainfall product provides high spatial (0.1° , approximately $11 \text{ km} \times 9 \text{ km}$ over South Africa) and temporal (30 min) resolution covering 60°S – 60°N ([Huffman et al. 2018](#)). At the heart of IMERG are data from the GPM *Core Observatory*, which carries the first spaceborne dual-frequency precipitation radar (DPR), as well as the multichannel GPM Microwave Imager (GMI). Rainfall estimates derived from the DPR and GMI are used to calibrate estimates from passive microwave sensors on board other low-orbit satellites. The resulting estimates are then merged with TIR-based rainfall estimates derived from geostationary TIR imagery to produce spatially complete rainfall fields. The IMERG process is run in two near-real-time modes (4 and 12 h after the observation time) and in one late mode 3 months after the observations. In the latter, the satellite-derived estimates are adjusted using monthly rain gauge analyses to produce an optimal rainfall estimate. It is this version (4.0A) that is used in this study. Despite the short heritage, validation studies of IMERG indicate that it performs better than its predecessor [TRMM Multisatellite Precipitation Analysis (TMPA)], especially over short time scales (e.g., [Dezfuli et al. 2017](#); [Manz et al. 2017](#); [Khodadoust Siuki et al. 2017](#); [Xu et al. 2017](#)).

3) CMORPH

The Climate Prediction Center morphing method (CMORPH) produces global rainfall estimates every 30 min with 8-km horizontal grid spacing and is available from December 2002 ([Joyce et al. 2004](#)). The estimates are based on both low-orbit passive microwave and geostationary TIR sensors. First, rainfall estimates are created based solely on the microwave retrievals. Next,

“cloud system advection vectors” are calculated using half-hourly TIR imagery; these vectors are used to propagate the precipitation fields forward and back in time where no direct passive microwave data are available. A time-weighted interpolation is applied to the available estimates to provide an estimate of both the rainfall distribution and intensity for the intervening missing 30-min periods; this process is referred to as “morphing.” In this study, the reprocessed bias-adjusted version 1.0 estimates were used (Xie et al. 2017). Wang et al. (2015) suggest that the bias-adjusted version of CMORPH performs well and better than the satellite-only version of the product.

4) CHIRPS

The Climate Hazards Group Infrared Precipitation with Station dataset (CHIRPS) provides quasi-global (50°S–50°N and 180°–180°) rainfall estimates at daily, pentadal, and monthly time steps with a 0.05° horizontal grid spacing and is available from 1981 (Funk et al. 2015). CHIRPS uses TIR imagery to produce maps of pentadal CCD using a constant rain/no-rain threshold of 235 K, calibrated using TMPA-3B42 rainfall estimates (Huffman et al. (2007), available between 2000 and 2013. As is done in TAMSAT, these climatological calibration parameters are then applied to the complete CCD record to produce a time series of rainfall estimates. These pentadal rainfall estimates are standardized (using their own climatology) and multiplied by a high-resolution and accurate rainfall climatology known as the Climate Hazards Group Precipitation climatology (CHPclim) to produce what is known as CHIRP. Finally, station rain gauge records are merged with CHIRP to create the CHIRPS product. CHIRPS version 2.0 was used in this study. Because a key part of the CHIRPS estimation process is the careful adjustment to the rainfall climatology, it typically has very low or negligible bias, and often has comparable or better skill than other products, particularly over longer time steps (e.g., Maidment et al. 2017; Dinku et al. 2018).

b. *CloudSat*

In this study, we use *CloudSat* observations to evaluate the spatial distribution of different cloud types and the vertical distribution of hydrometeors (Bodas-Salcedo et al. 2008; Stein et al. 2015). Errors in the vertical distribution may be used to direct model improvements in terms of microphysics parameterization (Satoh et al. 2010), while errors in the spatial distribution of different cloud types could suggest underlying issues with other parameterizations, such as the treatment of mixing in the boundary layer (Nguyen et al. 2017). Both sets of errors may have a two-way interaction with the circulation at the scale of the model domain, which Birch et al. (2014) have

shown may lead to biases in moisture availability within the first 24 h (i.e., on time scales relevant for operational forecasts of convection).

CloudSat was launched and joined the A-Train constellation of satellites in June 2006. The satellite carries a millimeter-wavelength radar, the *CloudSat* Profiling Radar (CPR), which has a sensitivity of -30 dBZ and a horizontal footprint at the surface of 1.5 km (Marchand et al. 2008). Its polar orbit leads to a 16-day revisit time, with equatorial crossings at approximately 0130 and 1330 local solar time. *CloudSat* came back online in November 2011 after a battery anomaly in April 2011; since then, it has only collected daytime observations. *CloudSat* observations are unique in probing the vertical structure of clouds and precipitation and have proven invaluable in the evaluation of weather and climate models (Bodas-Salcedo et al. 2008; Marchand et al. 2009; Delanoë et al. 2011; Kay et al. 2012; Stein et al. 2015).

The *CloudSat* data are obtained from the Radar–Lidar (DARDAR) product (Delanoë et al. 2011), in which the original *CloudSat* geometric profile product (2B-GEOPROF) mask and radar reflectivity have been interpolated onto a regular 1.5-km horizontal and 60-m vertical grid to match *CALIPSO* lidar observations. For model evaluation, it is convenient to set a radar reflectivity threshold to identify “clouds” and study their areal or volumetric fraction in the domain. For every 500 m in height, we will calculate the fraction of observations with a radar return considering thresholds of -30 and 10 dBZ; these fractions will be referred to as the “cloud fraction.” Since the *CloudSat* radar does not have Doppler capability, we make no distinction between cloud and precipitation, but we note that the -30 -dBZ threshold allows detection of nonprecipitating clouds, particularly ice clouds. The 10-dBZ threshold allows for a comparison of this analysis against the evaluation with the SAWS radars. However, it should be noted that the *CloudSat* measurements will be affected by non-Rayleigh scattering when hydrometeor sizes approach that of the transmitted wavelength (3.2 mm), which tends to happen at these high reflectivity values.

In addition to the vertical profile of cloud fraction, the occurrence of different cloud types is considered following the categorization from Stein et al. (2011). Due to the omission of a 3D pressure field as output from the model simulations, clouds will be distinguished by cloud-top height above mean sea level instead of cloud-top pressure, as used by Stein et al. (2011):

- 1) Shallow clouds have cloud tops below 3150 m (originally the 700-hPa level).
- 2) Midlevel clouds have cloud tops between 3150 and 8500 m and bases at least 3150 m above the

surface (8500 m corresponds approximately to the 350-hPa level).

- 3) Congestus clouds have cloud tops between 3150 and 8500 m and bases within 3150 m of the surface.
- 4) Deep clouds have cloud tops above 8500 m and bases within 3150 m of the surface.
- 5) Anvil clouds have cloud tops above 8500 m and bases at least 3150 m above the surface.

No distinct category is considered for cirrus clouds [identified as situated entirely above the 200-hPa level by Stein et al. (2011)], as these are rarely observed with *CloudSat* (Mace et al. 2009).

Due to the narrow footprint and long return time, *CloudSat* sampling is unsuitable for evaluating individual forecasts. For instance, for November 2016, only six daytime orbits crossed within 200 km of Johannesburg, South Africa. Therefore, in this paper, we will consider a climatology of cloud occurrences and vertical profiles of clouds and precipitation for November, based on observations from 2006 to 2016. During this period, 542 daytime and 279 nighttime orbits sampled the SAWS 4.4-km model domain and 69 daytime and 34 nighttime orbits sampled the smaller Highveld domain (for domains, see Fig. 1). The lower sample of nighttime orbits is due to the lack of nighttime observations since April 2011. To quantify the uncertainty due to the sampling limitations of *CloudSat*, we will apply a bootstrapping of the *CloudSat* orbits following Liu et al. (2010). The population of *CloudSat* orbits will be resampled with replacement 1000 times, calculating for each resampled population the mean cloud fraction for each reflectivity threshold; the profiles shown will then be the median of these 1000 mean cloud fractions. The 90% confidence interval will also be derived from these 1000 estimates of the mean. For individual cloud types, cloud occurrences are also obtained by calculating the mean for each of the 1000 resampled populations, then taking the median of these means, similar to the vertical profile of the cloud fraction.

c. SAWS radar network

In our study, we will use ground-based radar observations for model evaluation in two ways. First, a rainfall estimate from the radars will provide a fully independent observed estimate alongside the satellite rainfall estimates. Second, although weather radar networks are normally designed to detect and estimate rainfall, the radars scan at multiple elevations and can therefore detect part of the vertical extent of the precipitating clouds. We will use the radar-observed cloud fraction at various heights to evaluate the diurnal evolution of convective cloud in the convection-permitting models.

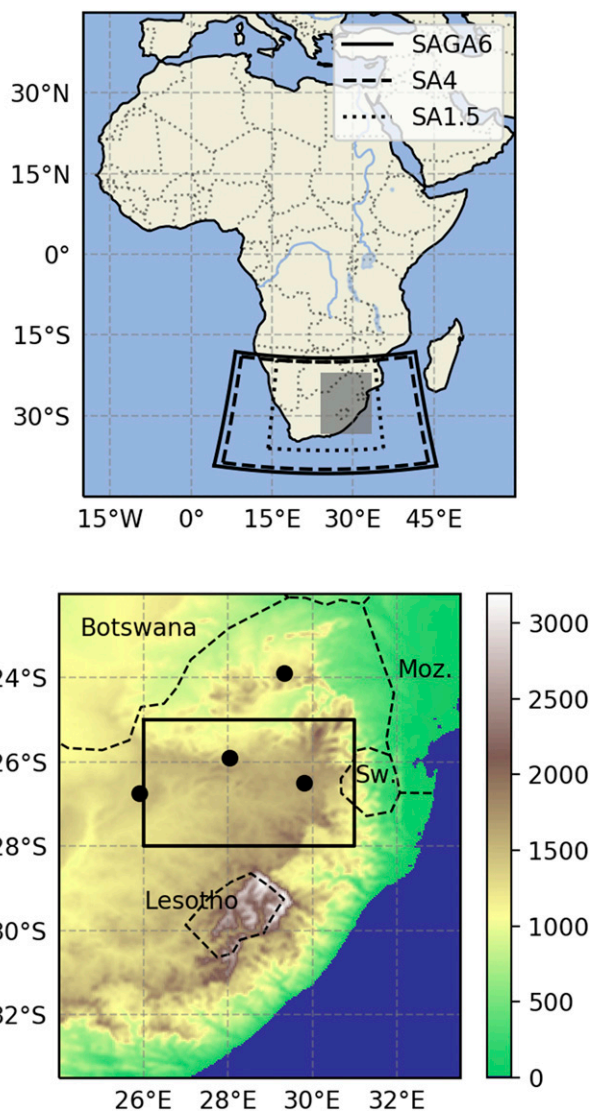


FIG. 1. (top) Map of Africa showing the domains of the nested reruns for SA1.5, SA4, and SAGA6 with country borders shown as dotted lines. The gray-shaded area shows the region of interest for the rainfall climatology in Fig. 2 and is shown in detail in the bottom panel. (bottom) Orography (m) for the region of interest, with locations of the SAWS radars used in this paper indicated by black dots. The rectangular box indicates the small Highveld domain used to analyze vertical cloud profiles and the diurnal cycle of precipitation. Neighboring countries are labeled: Sw. = Swaziland and Moz. = Mozambique.

Until 2009, the SAWS radar network consisted of 11 C-band radars (Terblanche et al. 2001). In 2010, the network underwent a significant improvement program in which 10 METEOR 600 S-band radars were purchased; 8 of these replaced the aging C-band radars and the remaining 2 were used to expand the radar network. All S-band radars considered have a maximum unambiguous range of 200 km. The radars are calibrated

once or twice each year, but power failures may affect calibration between these maintenance visits. Calibration is therefore monitored operationally using solar interference signals. Following Holleman et al. (2010), the solar interference for individual radars is compared against the S-band signal monitored at Dominion Radio Astrophysical Observatory in Kaleden, British Columbia, Canada. By removing the mean bias calculated over November 2016, we obtain a standard deviation of less than ± 1 dB for each of the SAWS radars during this period. In this paper, we will only consider the four radars in the northeast (Fig. 1), roughly covering the South African Highveld region, which experiences the heaviest annual rainfall totals and lightning activity in South Africa (Gijben 2012). Unfortunately, these radars are regularly prone to power outages (usually weather related) and therefore are not suitable for routine model evaluation and will only be considered for specific days when there is good reliability.

The SAWS radar data are available every 6 min and are provided in polar coordinates. Each radar has a 1° beamwidth, 500-m range gates, and performs scans at 12 elevations between 0.5° and 30° . The radar data are interpolated onto a regular 3D Cartesian grid using the University Corporation for Atmospheric Research (UCAR) Thunderstorm Identification Tracking Analysis and Nowcasting (TITAN) software (Dixon and Wiener 1993). For each radar in the network, TITAN was configured to perform an eight-point bilinear interpolation of the radar fields onto constant-altitude plan position indicators (CAPPIs) at 1-km horizontal and 500-m vertical grid spacing, referenced to height above mean sea level (MSL). The CAPPIs were then merged to create a single 3D Cartesian field for the entire network. For grid boxes covered by multiple CAPPIs, the merging algorithm selects the maximum radar reflectivity. For every 6-min data file, an auxiliary mask was also generated by TITAN to indicate which radars were available at that time.

The removal of ground clutter is achieved using built-in TITAN functions. For each radar, a clutter map is generated on the CAPPI grid, based on the mean radar reflectivity Z over the course of a completely dry day. If the mean Z at a point exceeds 10 dBZ, that point is considered clutter. For those clutter points, if the observed Z is within 5 dBZ of the coinciding clutter Z , it is considered to be clear of hydrometeors in our analysis.

Since we wish to evaluate the simulations against the radar data over a domain as large as possible, a threshold of 10 dBZ is selected. This is comfortably above the sensitivity for all radars of approximately 8 dBZ out to 200 km, or -38 dBZ at 1 km. However, the 10-dBZ reflectivity threshold will primarily detect precipitating

clouds only and the fractions reported should not be considered equivalent to the true cloud fraction. This fraction will be calculated at every 500 m in height as the fraction of pixels observed that has a radar reflectivity above the threshold considered.

For a qualitative comparison against the model simulations and satellite rainfall estimates, daily rainfall accumulation is derived from the radar data using the CAPPI at 3.5-km height above mean sea level. For the Highveld region, this height corresponds to approximately 2 km above ground, and a radar beam at 0.5° elevation would reach a 2-km height at a range of 125 km. Thus, a lower CAPPI would severely limit the available range for rainfall retrievals and reduce the overlap between radars. Rainfall rate is estimated from radar reflectivity using the Z - R relationship originally designed for the NEXRAD system (Fulton et al. 1998) and previously tested using the Irene radar in South Africa (Becker 2014):

$$Z = 300R^{1.4}, \quad (1)$$

where Z is in millimeters to the sixth power divided by meters cubed ($\text{mm}^6 \text{m}^{-3}$) and rainfall rate R is in millimeters per hour. Radar rainfall estimates are calculated at 6-min intervals and are capped at 103.9 mm h^{-1} , similar to the default value reported for the NEXRAD product by Fulton et al. (1998). Only daily accumulations of the radar-derived rainfall estimates are considered in this paper and only to provide a visual comparison of the spatial pattern against the satellite rainfall estimates. A quantitative precipitation estimate was not available for this study.

3. Model simulations

For the purpose of this study, SAWS provided UM simulations at 1.5-km grid length for the month of November 2016. November falls within the summer season for South Africa and typically experiences 2.6 days of significant rainfall (10 mm over 24 h) averaged over the Gauteng Province, which includes Johannesburg (Dyson 2009). For November 2016, Simpson and Dyson (2018) identified 17 severe weather events across the Highveld region.

a. SAWS operational forecasts

Through the UM Partnership, SAWS has been using regional configurations of the UM for over 10 years, starting at 12-km resolution in 2007. Since 2016, SAWS has run limited area CPMs, which include 4.4-km (SA4) and 1.5-km (SA1.5) configurations, both driven by lateral boundaries and initial conditions from the Met

Office global model run. The initial and lateral boundary conditions are received operationally four times daily from the Met Office, based on simulations using Global Atmosphere (GA) version 6.1 (GA6.1) science (Walters et al. 2017) at 0000, 0600, 1200, and 1800 UTC. The SA4 forecasts cover a 14-nation domain over southern Africa as part of SAWS's obligations as a Regional Specialized Meteorological Centre (<http://www.wmo.int/pages/prog/www/swfdp/SWFDP-SA.html>). Because of the computational expense, the SA1.5 is run over a smaller domain that covers South Africa.

During November 2016, the science configuration of the SA4 and SA1.5 was similar to the Euro4 used in the European regional model run by the Met Office (Clark et al. 2016), but both forecasts were run without additional data assimilation. The convection parameterization scheme is switched off. Subgrid mixing is parameterized using a 2D-Smagorinsky mixing scheme in the horizontal and the Lock et al. (2000) mixing in the vertical. The SA4 is initialized four times each day from the 0000, 0600, 1200, and 1800 UTC GA analyses and run out to 72, 48, 72, and 60 h, respectively, for these initialization times. SA1.5 forecasts are initialized from the same GA6.1 runs as the SA4, but always out to 36 h. Both the SA1.5 and SA4 have 70 vertical levels and a model top at 40 km and are run without data assimilation.

For the purpose of this paper, that is, to evaluate the SAWS models against ground-based and satellite-based radar observations, forward-simulated radar reflectivity diagnostics are required that are not available in the operational forecasts. Therefore, reruns were performed with both the SA1.5 and the SA4. It should be noted that the scientific configurations of these reruns differ from the SAWS operational forecast runs. However, the scientific configurations described below are those recommended for use of the UM over tropical domains, and SAWS expect to adopt these for their operational forecasts in the future.

b. SA1.5 reruns

The SA1.5 reruns were performed by SAWS on the same domain as the operational runs but with the updated tropical science. This configuration uses 80 vertical levels with a model top of 38.5 km and is used in and described by Stratton et al. (2018). The purpose of these reruns was to provide forward-simulated radar (Rayleigh) reflectivity. This diagnostic field was provided on a subset of the SA1.5 domain, namely where radar observations are available as shown in Fig. 1.

The radar reflectivity is calculated online from the model hydrometeor mixing ratios and parameters [see appendix A in Stein et al. (2014)]. The UM has a single-moment microphysics scheme (Wilson and Ballard 1999)

that treats mixing ratios of cloud ice and liquid, rain, and graupel as prognostic variables. Although there is no precipitating ice category (apart from graupel), the scheme includes a diagnostic split between ice crystals and aggregates, based on cloud-top temperature. The ice-particle number concentrations follow from Cox (1988), with the mass-diameter relationship for aggregates based on Locatelli and Hobbs (1974) and for crystals on Mitchell (1996). The rain and graupel particle size distributions are based on Abel and Boutle (2012) and Ferrier (1994), respectively.

All the SA1.5 rerun simulations were initialized with 1800 UTC initial and lateral boundary conditions from the GA6.1, identical to the operational runs. The operational forecasts and the reruns are dynamical downscalers and are consequently affected by spinup. Therefore, the first 6 h of the simulation should be ignored, and a complete 24-h period starting at 0000 UTC can be employed with greater confidence using a run initialized at 1800 UTC. Due to latency in the availability of the GA analyses at SAWS and the time required to run the operational forecasts, the 1800 UTC run also coincides with the forecast cycle that is most beneficial to the forecasters at SAWS.

c. SA4 and SAGA6 reruns

The SA4 reruns were configured for a domain slightly expanded east-, west-, and southward from the SA1.5 reruns (see Fig. 1). These were run as 36-h forecasts with the nesting suite that is available for the UM (suite u-aa753) and nested in the UM global model (GA6.1) at N320 resolution, which was initialized daily from Met Office 1800 UTC global analyses. The SA4 reruns used the same science configuration as the SA1.5 reruns, but with 70 vertical levels instead of 80. The forward-simulated radar (Rayleigh) reflectivity diagnostic was output on every hour.

The SA4 reruns were performed on the ARCHER high-performance computing facility rather than by SAWS to enable the output of diagnostics using the CFMIP Observation Simulator Package version 4.1 (COSP; Bodas-Salcedo et al. 2011). COSP is available in the UM to provide forward-simulated diagnostics for comparison against satellite measurements (Bodas-Salcedo et al. 2008). The *CloudSat* simulator in COSP requires information of hydrometeor concentrations on the subgrid scale and is therefore computationally expensive to run. In the nesting suite at convection-permitting resolutions, the UM implementation of COSP uses the Subgrid Cloud Overlap Profile Sampler to produce the subgrid distribution of clouds (Webb et al. 2001). Forward-simulated *CloudSat* reflectivities are then obtained from scattering calculations including attenuation, assuming a downward-pointing

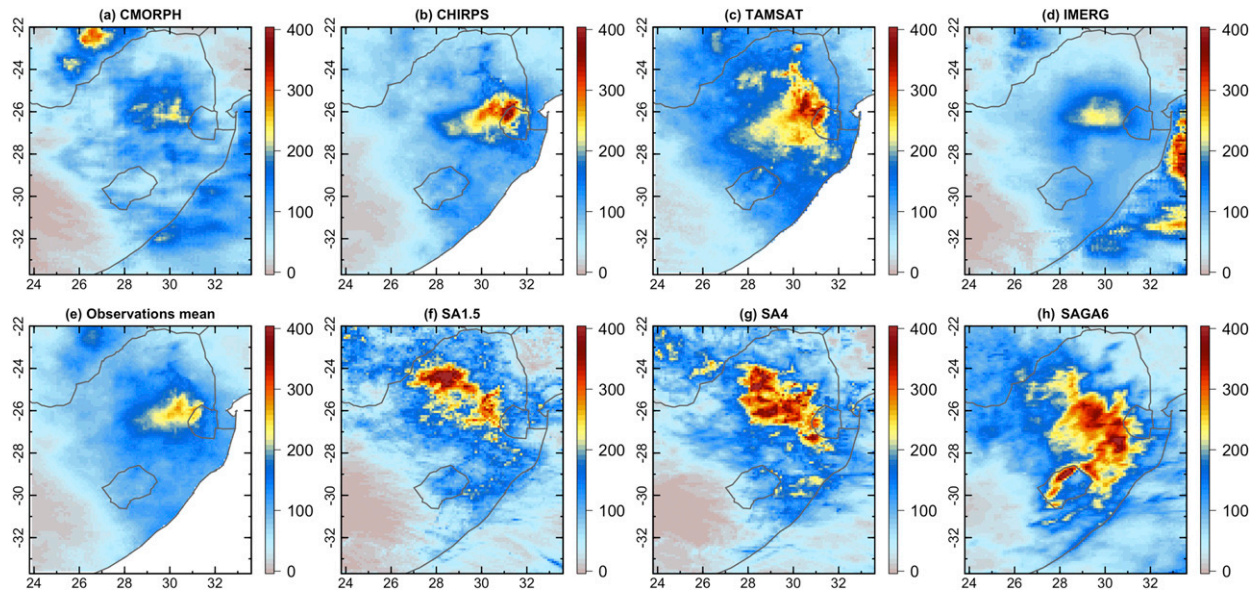


FIG. 2. Rainfall accumulation for November 2016 (mm) from the following satellite rainfall estimates and model simulations: (a) CMORPH, (b) CHIRPS, (c) TAMSAT, (d) IMERG, (e) mean of the satellite products, (f) SA1.5, (g) SA4, and (h) SAGA6. All products are interpolated onto a regular 0.1° grid. Rainfall is accumulated from 0600 UTC 1 Nov 2016 until 0600 UTC 1 Dec 2016.

radar beam. COSP output was only produced at 1200 and 0000 UTC (18 and 30 h into the simulation), to compare against the *CloudSat* observations at 1330 and 0130 local time (approximately 1130 and 2330 UTC). Cloud fraction is calculated from the forward-simulated *CloudSat* reflectivities using the same thresholds as in the observations. Rather than using the entire 3D output at each 1200 and 0000 UTC time, the model is subsampled to ensure that individual samples are not correlated. For instance, if every longitude were counted, cloud systems that span over multiple longitudes would contribute to the cloud fraction multiple times, which would not occur in the observed sample. The subsampling is performed 1000 times, each time selecting 60 longitudes and dates at random, then calculating the average cloud fraction and cloud occurrences from these 60 samples; the profiles and cloud occurrences shown will be the median of these 1000 averages.

We emphasize that the COSP simulator was only included in the SA4 reruns as this implementation required updates to the UM code, which were not available at the time of the other reruns performed for this study. However, [Stein et al. \(2015\)](#) found that UM biases in the horizontal and vertical distribution of clouds during the West African monsoon are broadly comparable across different grid lengths, especially when comparing 4- and 1.5-km simulations.

A final set of reruns was performed using a configuration with convection parameterization to understand the benefits of CPMs over the global forecasts. These

reruns will be referred to as “SAGA6” and use the Met Office GA6.1 configuration with a 10-km horizontal grid length and are nested inside the same N320 GA6.1 global run as the SA4. Although the SA4 and SAGA6 reruns use the same driving model, the SAGA6 had a slightly larger domain than the SA4 (see [Fig. 1](#)). Only total precipitation (large scale and convective combined) is evaluated from the SAGA6 reruns, since this is the output that is used by forecasters at SAWS.

4. Monthly climatology

We first consider the models’ performance for the entirety of November 2016. Monthly statistics allow us to infer model performance over a range of cases and study the consistency of forecast skill. Subtle biases, such as enhanced rainfall over orography or widespread light rainfall, may also be revealed when accumulated over longer time periods.

a. Rainfall accumulations

In [Fig. 2](#), maps of monthly rainfall accumulation show that the four satellite estimates generally agree in terms of the spatial rainfall pattern, but disagree on high rainfall totals; the latter is illustrated by the cumulative distributions shown in [Fig. 3](#). CMORPH, CHIRPS, and IMERG compare well against one another in terms of cumulative distribution up to the 95th percentile, which is lower in CMORPH. TAMSAT has a median rainfall accumulation approximately 50–60 mm higher than the

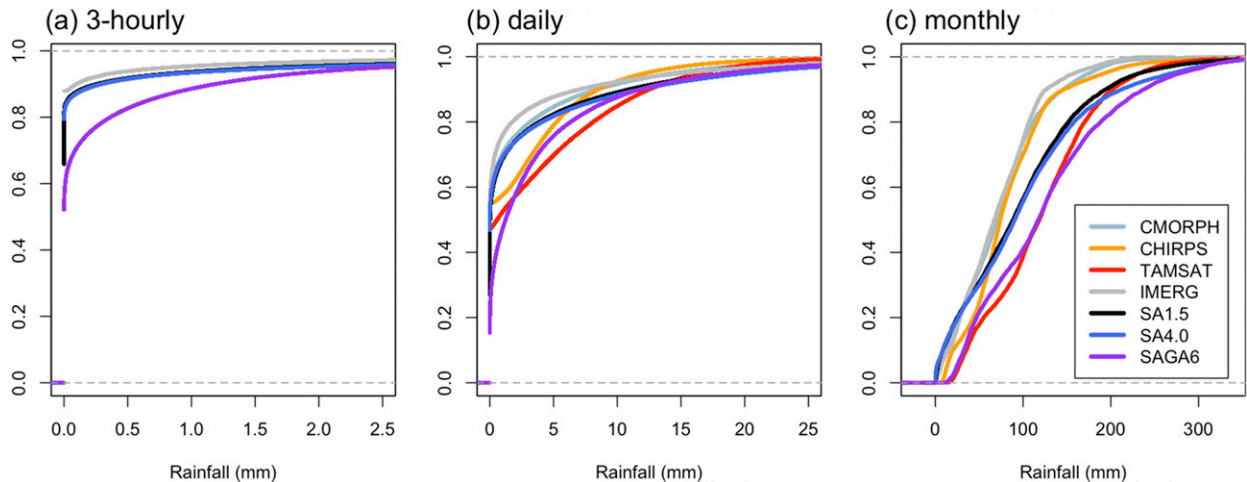


FIG. 3. Cumulative density curves for the satellite rainfall estimates and model simulations shown in Fig. 2 (excluding the observational mean) considering (a) 3-hourly accumulations, (b) daily rainfall accumulations, and (c) monthly rainfall accumulations. Only estimates over the Highveld domain in Fig. 2 are considered.

other satellite estimates, which is also evident in the rainfall map, especially in areas of low rainfall accumulations. Notably, TAMSAT is biased 10–20 mm high compared to the other estimates for its lowest decile, which is likely due to it estimating rainfall as soon as CCD is nonzero. In particular, Maidment et al. (2014) attribute this high bias for low rainfall amounts to the calibration, which for version 3.0 is done against the mean rather than the median.

All satellite products show a region of high rainfall rates west of Swaziland, around 25°–27°S and 28°–31°E (see also the map in Fig. 1). Local maxima can be identified in Botswana around 22°–23°S and 26°–27°E and (in TAMSAT and CHIRPS only) just southeast of that location at 24°–25°S and 28°E.

All three model simulations reproduce the area of high rainfall accumulation west of Swaziland in Fig. 2. The SA4.0 and SA1.5 also generate high rainfall accumulations around 24°–25°S and 28°–29°E, slightly displaced from the local maxima observed by CHIRPS and TAMSAT. The SAGA6 does not have such high rainfall accumulations in that region, but has a band of high accumulations surrounding Lesotho at 29°S and 28°E, coinciding with the Maloti mountain range. The three models have a coherent region of rainfall accumulations above 100 mm stretching from 22°S and 24°E to Swaziland. Such a band of rainfall is not obvious from the satellite products. However, this difference between models and observations may be related to the former missing the localized maximum in Botswana.

In terms of the spatial pattern and cumulative distribution of monthly rainfall accumulations (Figs. 2 and 3), the SA4 and SA1.5 simulations compare very well against each other. Both sets of simulations fall within the range

of estimates provided by CMORPH, CHIRPS, and IMERG in terms of the cumulative distribution, comparing well against these up to the 60th percentile. The SA1.5 compares well against TAMSAT above the 95th percentile, while the SA4 has slightly higher extreme rainfall accumulations. The SAGA6 simulations compare better against TAMSAT, following its cumulative distribution for the lowest two deciles, while it has the highest rainfall accumulations above the 90th percentile.

In Fig. 4, the domain-averaged daily rainfall amounts are shown, comparing the model simulations against the individual satellite products. The SA4 and SA1.5 simulations generally underestimate daily rainfall accumulation compared to TAMSAT, but compare well against CHIRPS, CMORPH, and IMERG, as noted in the monthly statistics. Conversely, the SAGA6 compares well against TAMSAT but overestimates daily rainfall compared to the other three rainfall estimates. Despite these biases, the models have a high correlation with all four satellite products at the daily time scale and at the spatial scale considered (averaged across all land points between 22.0°–33.6°S and 23.9°–33.5°E).

b. Rainfall diurnal cycle

On the subdaily scale, the models can be evaluated against CMORPH and IMERG, which provide subdaily rainfall estimates at 30-min time steps. The mean diurnal cycle for November 2016 is shown in Fig. 5, averaged for the smaller Highveld domain (Fig. 1). As on the monthly and daily time scales, the SA4 and SA1.5 results compare well against each other, but rainfall is overestimated compared to CMORPH and IMERG at the time of peak rainfall rate. The SA4 produces slightly higher rainfall rates compared to the SA1.5; this can also be discerned

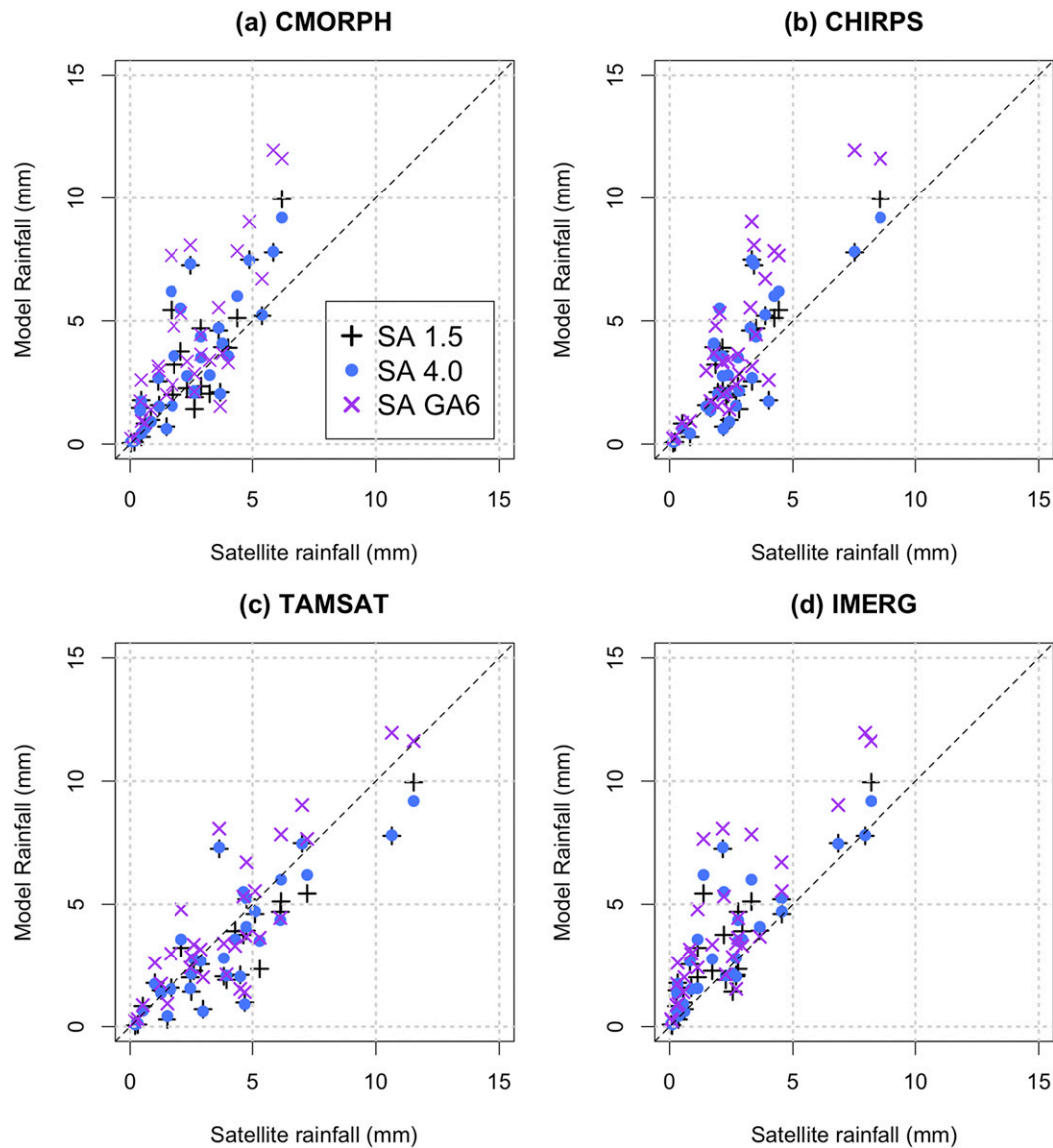


FIG. 4. Domain-averaged daily rainfall rate (mm) for the individual satellite rainfall estimates (panels and horizontal axes) vs the three model simulations (scatter points and vertical axes). All estimates over land in Fig. 2 are considered.

from the deviation between the two models at the highest percentiles of the cumulative distribution in Fig. 3. Both models have a peak around 1600–1700 UTC, which compares well against IMERG and CMORPH. A sharp peak in the IMERG diurnal cycle can be attributed to extreme rainfall rates estimated on one particular day (11 November 2016).

The SAGA6 reruns have a monthly averaged diurnal cycle that peaks 1–2 h ahead of the SA4 and SA1.5 simulations, and also ahead of the observed peak. The magnitude of the peak matches that of the observations. Comparing the SAGA6 diurnal cycle in Fig. 5 to

the accumulation in Fig. 2, the overestimated rainfall by SAGA6 appears to be due to higher rainfall rates overnight, which are not observed in the satellite retrievals and not simulated by SA4 or SA1.5.

c. Skill scores

Forecast verification of rainfall accumulation considering individual grid points allows further distinction between the performance of the three models (Figs. 6 and 7). Using 0 mm as a “rain/no rain” delineation, the SAGA6 substantially overpredicts rainfall, noted by the high bias, which is consistent with the previous findings.

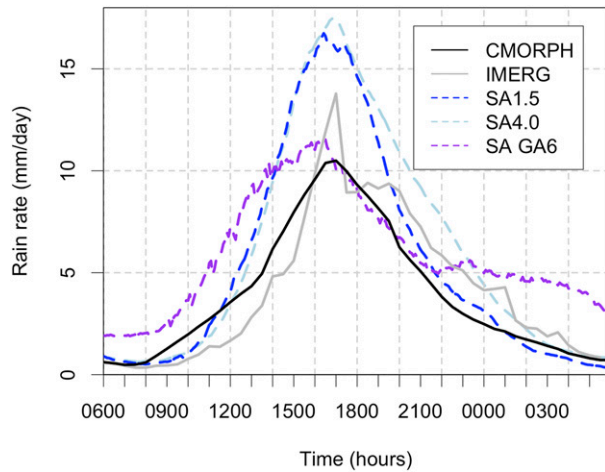


FIG. 5. Monthly averaged diurnal cycle of domain-averaged rainfall rate (mm day^{-1}) for the three model simulations and IMERG and CMORPH. The average is calculated over the Highveld domain.

The SA1.5 also overpredicts rainfall on most days, whereas the SA4 has a considerably better bias with medians against the four satellite products between 1 and 1.5. The false alarm ratio (FAR) between the three models is comparable. The good SA4 performance in bias is reflected in a much lower probability of false detection (POFD) and also in a lower probability of detection (POD). We use the IMERG 95th percentile of daily rainfall accumulations as a high rainfall threshold that should be of greater use for forecasting severe rainfall (Fig. 7). The three models slightly overpredict this rainfall occurrence, although the median against IMERG is between 1 and 1.5. The overprediction can be recognized from the CDF, where the IMERG 95th percentile value is a slightly lower percentile for the three models. The three models are comparable for all metrics considered, although we note that these standard skill scores are dominated by correct negatives for rare events.

The poorer performance of the SA4 in terms of POD suggests that while the correct cover of rainfall is predicted, it may be predicted in the wrong location. We briefly consider the fractions skill score (Roberts and Lean 2008) as a spatial verification metric to address this issue. Due to the resolution of our observational data, we only consider the “asymptotic” fractions skill score (AFSS), which is defined as

$$\text{AFSS} = 1 - \frac{(p_o - p_f)^2}{p_o^2 + p_f^2}, \quad (2)$$

where p_o and p_f are the fractions of the domain covered with rainfall above a given threshold. We consider both

fixed thresholds for all products and product-relative (percentile) values. Following Roberts and Lean (2008), we consider the mean AFSS over multiple days. However, if p_o and p_f are both 0 for one case day, the AFSS becomes undefined. Thus, we only consider those days where $p_o > 0$ for the 99th percentile of IMERG (35.6 mm), leading to 13 days in November 2016. For 3-hourly rainfall, we use the same restriction, namely the 99th percentile of IMERG (8.2 mm), leading to 138 slots of 3 h in November 2016.

From Fig. 8 we can tell that on the daily scale, the three models are nearly indistinguishable. The SAGA6 performs slightly better for the highest rainfall rate thresholds, but this may be due to the CPMs having a longer tail for the rainfall distribution compared to IMERG. When product-related percentiles are used, the SAGA6 only performs better for the 99th percentile. On the 3-hourly scale, the SA4 and SA1.5 clearly perform better than the SAGA6. This may be interpreted in relation to the diurnal cycles from Fig. 5, in which the SAGA6 produces more rainfall in the nighttime hours, when it is not observed.

d. Clouds

In Fig. 9, vertical profiles are shown of the bootstrapped median cloud fraction for November 2006–16 (*CloudSat*) and November 2016 (SA4) for the two reflectivity thresholds. The SA4 generally produces more cloud with $Z \geq -30$ dBZ at all heights in November 2016 than is found with the *CloudSat* climatology. Above 10 km during the day and above 6 km at night, the SA4 cloud fraction is within the 90% confidence interval of the *CloudSat* climatology, whereas particularly around 5 km, the SA4 generates a maximum of cloud fraction that is not observed. The simulated maximum around 5 km may be compared to the preferred detrainment of midlevel clouds just above the freezing level that was also found with the UM for West Africa by Stein et al. (2015). Notably, the *CloudSat* observations indicate a second maximum during the nighttime around 7 km, which may indicate midlevel detrainment but at a higher level than in the simulations.

For the 10-dBZ threshold, the SA4 compares very well against the *CloudSat* climatology during the nighttime (0130 local solar time), as its median profile falls within the 90% confidence interval at all heights. The nighttime profile has a peak around 6 km for both thresholds, with a second peak around 4 km for the 10-dBZ threshold, possibly due to stratiform precipitation. The daytime (1330 local solar time) 10-dBZ profile in the SA4 has a different structure than the *CloudSat* profile, with a clear maximum around 8 km,

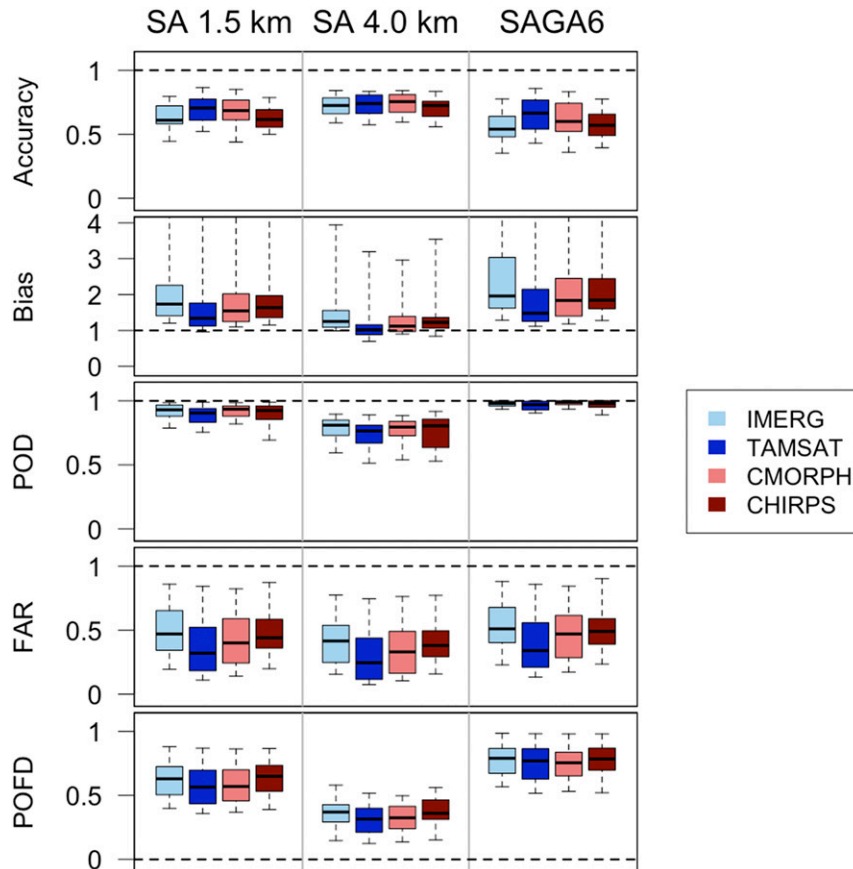


FIG. 6. Forecast verification of daily rainfall accumulation for 0.1° grid boxes in the Highveld domain using a 0-mm threshold. The (left) SA1.5, (center) SA4, and (right) SAGA6 results are verified each day in November 2016 against the rainfall estimates from IMERG, TAMSAT, CMORPH, and CHIRPS. The box plots show the median and interquartile range of the metrics over the 30 days, with whiskers indicating the 5th and 95th percentiles. Accuracy is also known as the proportion correct; POD = probability of detection (or hit rate), POFD = probability of false detection, and FAR = false alarm ratio.

several kilometers higher than the nighttime maximum. This peak of 10-dBZ occurrence at upper levels could be a sign of more active convection producing anvil clouds compared to the *CloudSat* climatology. Cloud occurrence as measured by a 10-dBZ threshold will be considered further using the SAWS radar data in section 5.

In Fig. 10, we compare the bootstrapped median cloud occurrence separated by five different types for *CloudSat* observations and the SA4 simulation. The SA4 simulates a day–night difference in the pattern and amount of shallow cloud cover that compares well against the *CloudSat* observations, although the amounts are generally too high (first row). The SA4 underestimates congestus but overestimates midlevel cloud occurrence over land (second and third rows). During the night, an area of maximum midlevel cloud cover is oriented from 23°S , 18°E to 30°S , 30°E . A similar area during the day has a slightly more

southward orientation, which can be discerned in both the SA4 and the *CloudSat* climatology. The location of the maximum midlevel cloud cover during the day in the SA4 coincides with the Highveld domain considered for the vertical profiles in Fig. 9 and could therefore explain the peak in cloud fraction at 5 km for $Z \geq -30$ dBZ. The SA4 generally lacks congestus and midlevel clouds over the ocean, which prevail in the observations both during the day and at night. Compared to the occurrence of deep clouds (fourth row), the SA4 generally produces midlevel clouds in areas with convection. As the SA4 does not produce much deep cloud or congestus over the ocean, this may be the reason it lacks midlevel clouds there. However, in the observations, the midlevel cloud pattern is not so clearly associated with the presence of deep clouds. These oceanic midlevel clouds may thus be related to synoptic patterns, and it will be informative to study these in the global UM in future work.

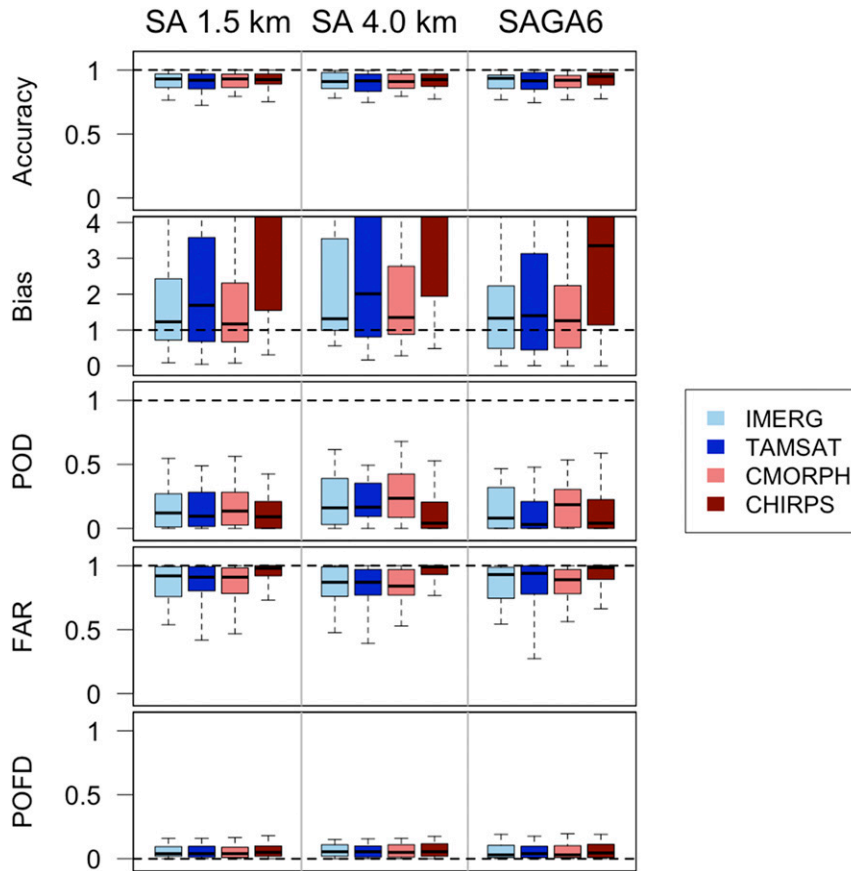


FIG. 7. As in Fig. 6, but using the 95th percentile from IMERG as the rainfall rate threshold (15.6 mm).

The deep clouds and anvil clouds (fourth and fifth rows in Fig. 10) have a similar orientation of the area of maximum occurrence during the night in the SA4 as the midlevel clouds. In the SA4, the anvil clouds occur farther east and the deep clouds occur in between the midlevel clouds and the anvil clouds. In the observations, this geographic displacement of the three cloud types is not as obvious, although the peak anvil occurrence does occur in the east. The daytime simulated pattern also shows deep clouds to the east of the midlevel clouds, without a clear maximum in anvil cover. The SA4 appears to underestimate cover from deep clouds both during the day and at night compared to *CloudSat* observations. The orientation and location of the band of anvil, altocumulus, and cumulonimbus simulated by the SA4 matches that of the enhanced rainfall in Fig. 2.

The distribution of deep convective clouds observed by *CloudSat* does not match the observed rainfall pattern very well, which is understandable given the different observation period (all Novembers 2006–16 compared to November 2016 only). This highlights the difficulty in

using *CloudSat* for model evaluation on relatively short time scales. However, the evaluation against *CloudSat* highlights issues, such as the biases in oceanic shallow and midlevel clouds. Model process studies could be directed to study these cloud biases further, including their potential impact on the local circulation and therefore the timing and location of precipitation.

5. Case study days

Four days in November 2016 have been selected to consider subdaily variations of rainfall and clouds. The primary criterion for case study selection was to have at least three of the four SAWS radars in the northeast operational for more than 90% of the time. Of those days that met this criterion, cases were selected subjectively based on whether convection developed within the radar domain. For three of these four days, severe weather events were identified in the Highveld region by Simpson and Dyson (2018), with heavy rainfall on 10 and 25 November, and a tornado and large hail occurrence on 12 November.

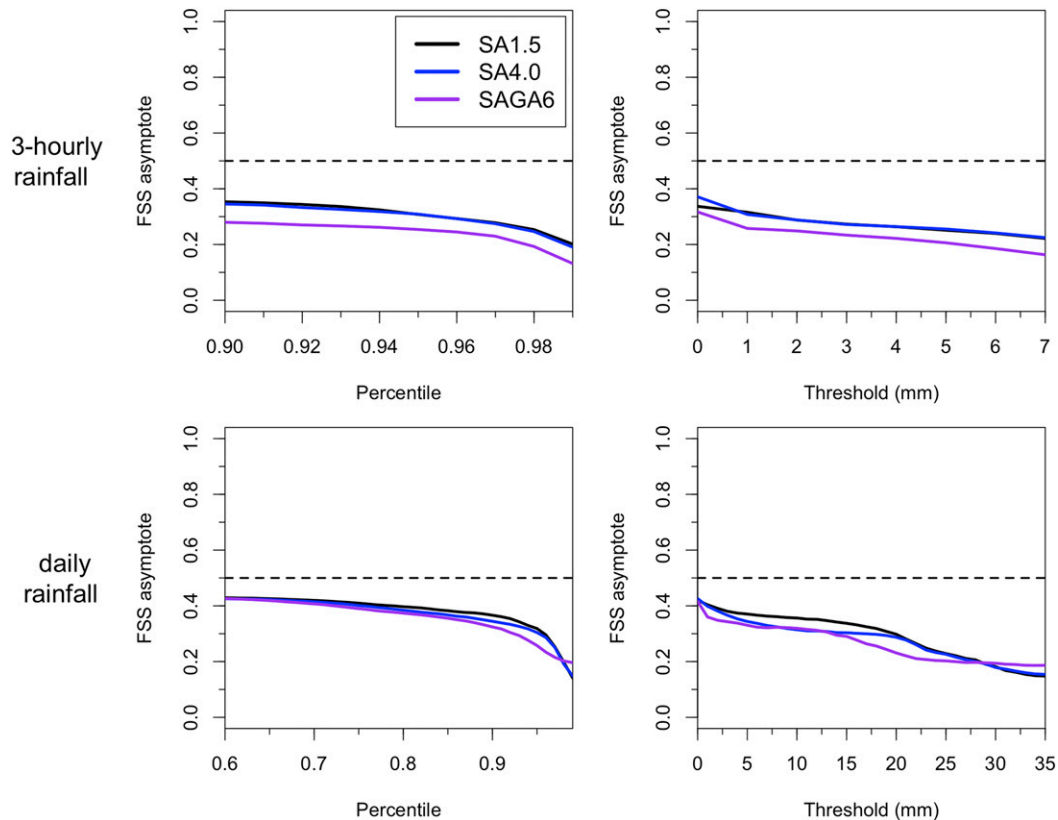


FIG. 8. Asymptote fractions skill score (FSS) for the Highveld domain, calculated using the IMERG rainfall product following Eq. (2). Percentiles are relative to each product. The dashed line at AFSS = 0.5 would be equivalent to a bias of 4 if both were calculated for a single day (Roberts and Lean 2008).

In Fig. 11, the daily rainfall accumulation is shown for the four case study days, derived from the satellite products CMORPH and IMERG, the SAWS radar observations, and the three model simulations. For these daily accumulations, IMERG compares well against the radar observations, particularly in terms of the spatial scale of rainfall features. CMORPH, despite having a higher resolution than IMERG, retrieves broader areas of rainfall than IMERG and the radar-based estimates. The SA1.5 and SA4 simulations reproduce the varying spatial scales for the different days, with small-scale storms on 8 and 25 November 2016, and larger, likely longer-lived, systems on 10 and 12 November 2016. The SAGA6 simulations fail to reproduce the isolated events of 8 and 25 November 2016, while they compare reasonably well against observations for the larger-scale event of 10 November. Interestingly, on 12 November, the SAGA6 produces bands of heavy precipitation that are oriented nearly perpendicular to the bands produced by the SA4 and SA1.5, while the latter compare well against the observations.

The diurnal cycle of precipitation is shown for the individual days in Fig. 12. The SA1.5 and SA4 simulations

behave similarly for all cases in terms of phase and amplitude, and there is no consistent difference between the two configurations. For 8 and 25 November, all three model configurations overestimate the amplitude of the diurnal cycle compared to CMORPH and IMERG, while for 12 November the models compare well against IMERG but overestimate rainfall compared to CMORPH. For 8 and 12 November, the SAGA6 produces rainfall overnight that is not observed and not simulated by the CPM configurations. In both cases, the SAGA6 rainfall maps in Fig. 11 do not match the observations or the SA1.5 and SA4 results very well. For 10 November, the SAGA6 compares very well against CMORPH and arguably performs better than the SA1.5 and SA4 since its rainfall persists overnight, as it does in the observations.

In Figs. 13 and 14, for the four case study days, the diurnal cycle of the vertical profile of the radar-observed cloud fraction is shown for the SAWS radars and for the SA1.5 and SA4 (the radar reflectivity diagnostic is not available for the SAGA6). Note that the 10-dBZ radar reflectivity threshold does not

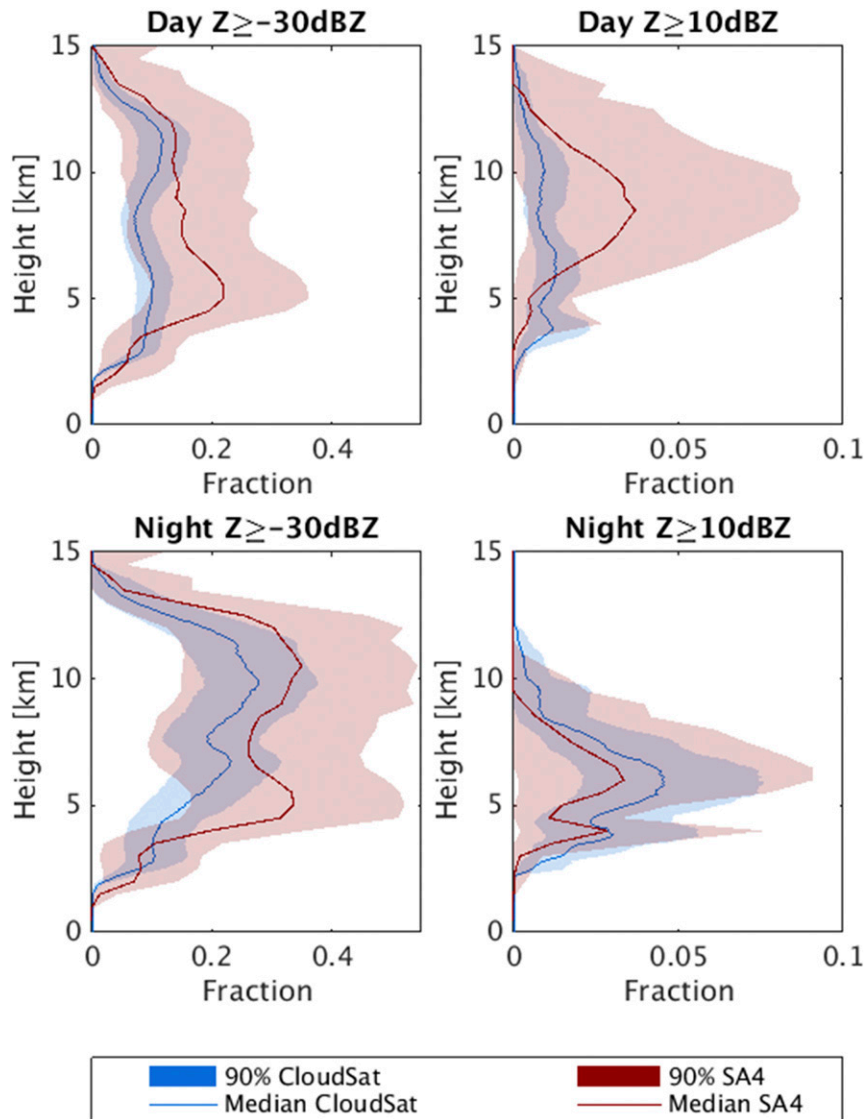


FIG. 9. Bootstrapped median and 90% confidence interval of the vertical profile of the cloud fraction derived from *CloudSat* (November 2006–16, blue) and SA4 (November 2016, red). (top) Daytime and (bottom) nighttime, for the reflectivity thresholds of (left) -30 and (right) 10 dBZ. Nighttime and daytime refer to 0130 and 1330 local solar time for *CloudSat* and to 0000 and 1200 UTC for SA4.

distinguish between clouds and precipitation. The SA4 and SA1.5 simulations behave similarly to one another in all four cases in terms of the general pattern and amounts of cloud fraction. Comparing the timing of the maximum cloud occurrence, the SA1.5 is at $-4, -3, 0,$ and -3 h compared to the observed time, and the SA4 is at $+1, -2, 0,$ and 0 h. If we only compare the timing of the lowest cloud-fraction contour at 6 km, then the SA1.5 has the timing $-1, -4, -1,$ and 0 h prior to the observed timing, whereas the SA4 has it at $+1, -2, 0,$ and $+2$ h. Thus, the SA1.5 appears to have the peak of the diurnal cycle slightly advanced compared to the observations,

whereas the SA4 is typically 2 h delayed compared to the SA1.5, which results in a better comparison against the observations.

For 8, 10, and 12 November, the domain of radar observations (see Fig. 11) can be approximated by the rectangular domain of 25° – 28° S and 26° – 31° E that has been used for calculating the diurnal cycles of rainfall in Fig. 12. For 12 November, the timing of the peak cloud fraction matches the timing of the peak rainfall rate reasonably well for both the SA1.5 and the SA4. However, in the SA1.5, for 8 and 10 November, the peak cloud fraction leads the peak rainfall rate by 2–3 h,

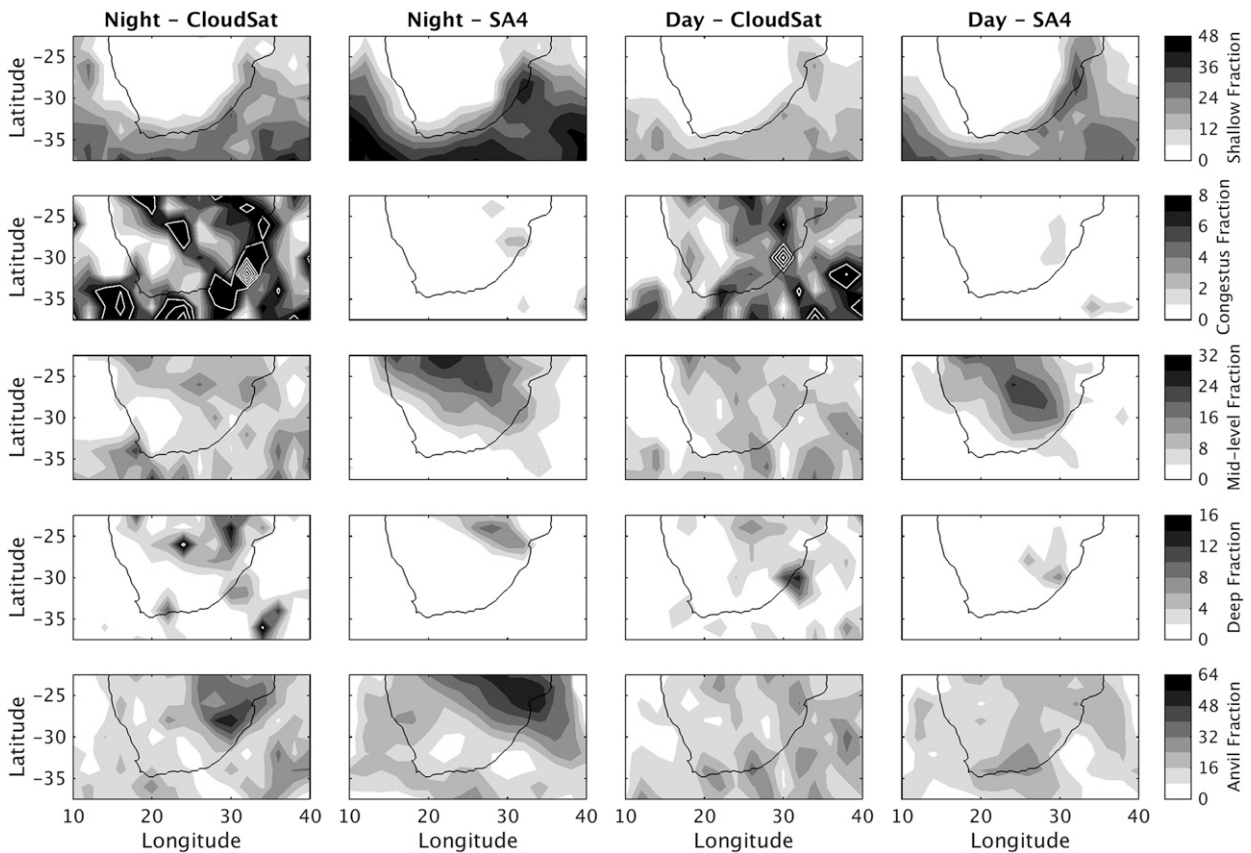


FIG. 10. Median cloud cover of different cloud types derived from *CloudSat* (November 2006–16) and SA4 (November 2016) for nighttime in the first and second columns and daytime in the third and fourth columns. (first row) Shallow clouds, (second row) congestus, (third row) midlevel clouds, (fourth row) deep clouds, and (fifth row) anvil clouds. Cloud occurrence is calculated on a $2^\circ \times 2^\circ$ latitude–longitude grid and is reported in percent. For congestus and deep clouds, white contours are spaced 4% to indicate higher occurrences in the observations.

whereas in the SA4, the two timings are still similar. In the observations, the peak cloud fraction and peak rainfall rate appear to match reasonably well in time. Since the SA1.5 and SA4 have similar scientific configurations in terms of microphysics and subgrid mixing, there is no obvious choice of parameterization that could lead the SA1.5 to display such behavior of producing a lot of (10 dBZ) cloud several hours prior to producing the highest rainfall rates. A process-oriented investigation of the physical representation of convection in the SA1.5 is planned to address this bias.

The height of individual cloud fraction contours in Figs. 13 and 14 tends to peak at the same time as the maximum cloud fraction. After this peak of convective activity, the height of the individual contours generally decreases throughout the day, but higher values of cloud fraction tend to linger beyond 0000 UTC in the observations. The SA4 and SA1.5 underestimate the height reached by the various cloud-fraction

contours. Considering the lowest cloud-fraction contour, this contour reaches a height above 14, 14, 15, and 13 km for the four different days; compared to the SA1.5 at 12, 12, 13, and 11 km; and compared to the SA4 at 13, 12, 14, and 11 km. The models possibly underestimate this height due to a general underestimation of cloud fraction, but this is not supported by the cloud fraction observed at other levels, for instance at 6 km, which compares well against the observations.

The results of Figs. 13 and 14 cannot easily be compared against the *CloudSat* results of Fig. 9. The *CloudSat* comparison is made at 1330 and 0130 local time, approximately 1200 and 0000 UTC, which misses the peak of convective activity. Also, the *CloudSat* climatology is over all November days in 2006–16 and the SA4 results in Fig. 9 are over all November days in 2016, instead of the four cases considered here. Finally, the *CloudSat*-observed and *CloudSat*-simulated reflectivities suffer from strong attenuation due to the millimeter wavelength, causing the decrease in cloud

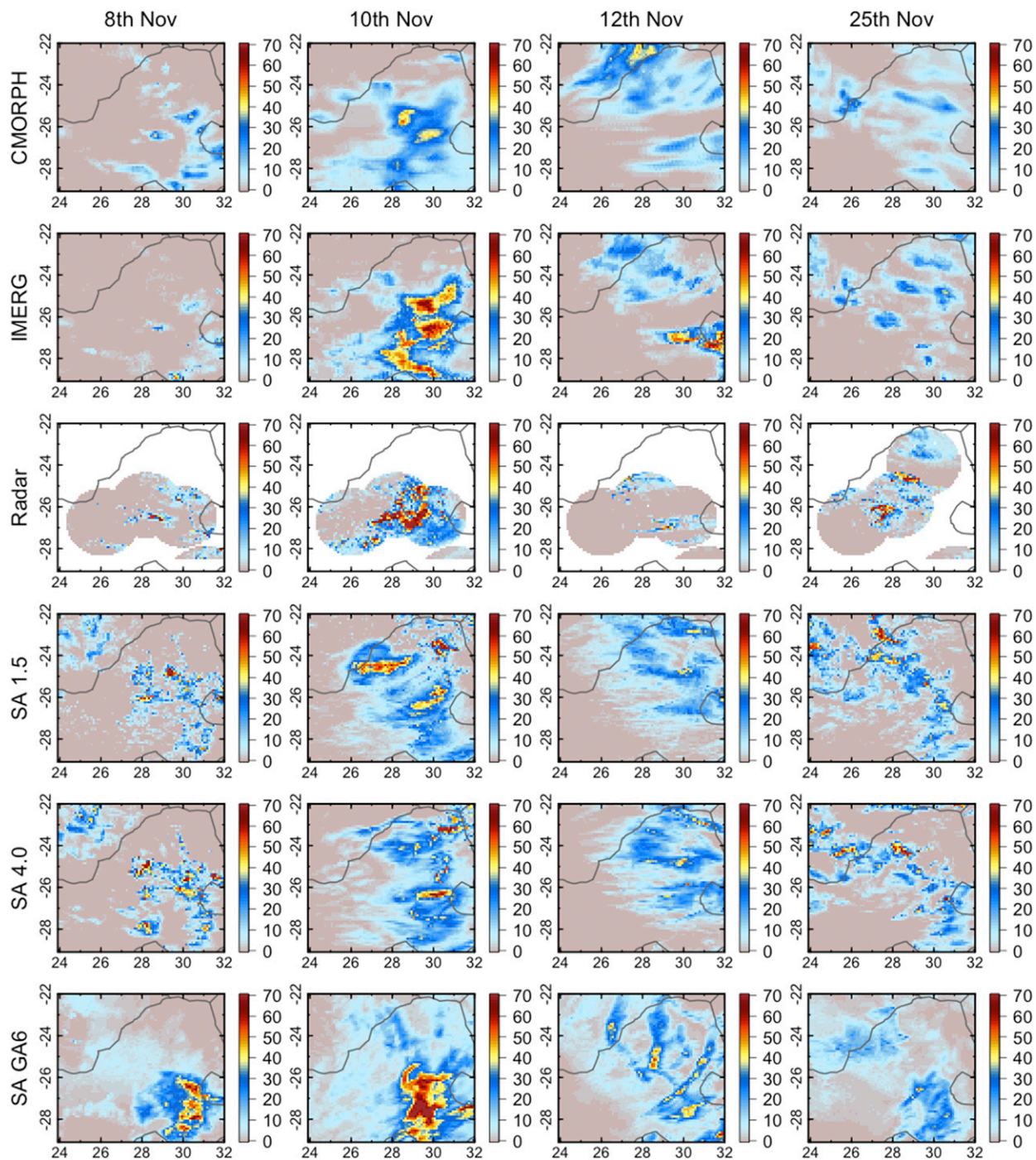


FIG. 11. Daily rainfall accumulation (mm) from the satellite products CMORPH and IMERG, from the radar observations, and from the SA1.5, SA4, and SAGA6 for the four case study days 8, 10, 12, and 25 Nov 2016. Rainfall is accumulated from 0600 UTC on the case study day until 0600 UTC on the following day.

fraction with height from about 6 to 8 km downward. Nevertheless, there is a hint in Fig. 9 that the observed daytime cloud fraction has its maximum at a higher level than the SA4, about 11 km compared to 8 km, which is consistent with Figs. 13 and 14. The

overestimate of the cloud fraction by the SA4 compared to *CloudSat* in Fig. 9 is not so evident from the four case studies, although this analysis may be sensitive to a few cases with large overestimates such as 12 November 2016.

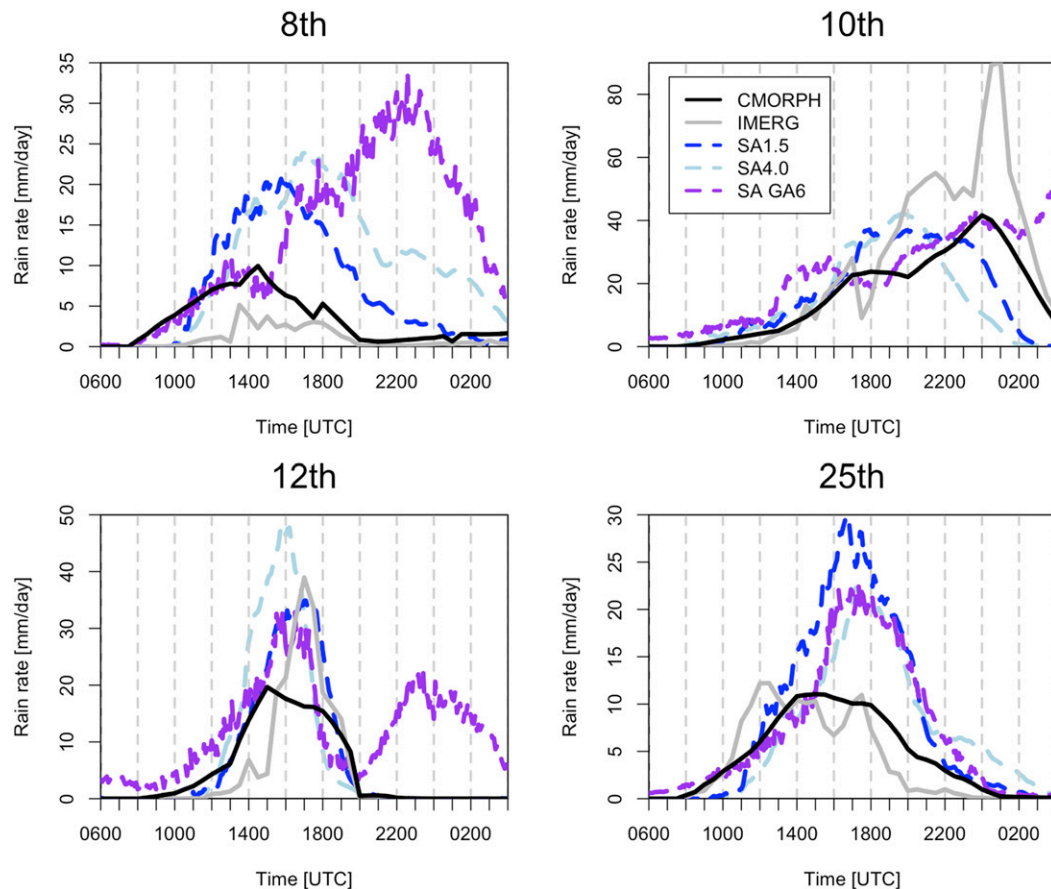


FIG. 12. Diurnal cycle of domain-averaged precipitation from the satellite products CMORPH and IMERG and from the SA1.5, SA4, and SAGA6 for the four case study days 8, 10, 12, and 25 Nov 2016. Rainfall is accumulated from 0600 UTC on the case study day until 0600 UTC on the following day. The averages are calculated for the Highveld domain.

6. Discussion and conclusions

We have shown that the Met Office Unified Model simulations run by SAWS as limited area CPMs perform well in terms of cloud and rainfall distributions across South Africa for the month of November 2016. The two CPM configurations, SA4 and SA1.5 (4.4- and 1.5-km horizontal grid lengths), show comparable behavior on the monthly and daily time scales in terms of the spatial pattern of rainfall and in terms of rainfall accumulations. At the daily scale, the SA4 shows slightly higher domain-averaged rainfall than the SA1.5, while it performs slightly better in terms of probability of detection and false alarm ratio using a range of rainfall thresholds. In terms of the diurnal cycle of clouds, the SA4 tends to be delayed by 1–2 h compared to the SA1.5, although no consistent difference is evident in the diurnal cycle of precipitation. The CPMs perform better than SAGA6, a 10-km grid-length model with convection parameterization, in terms of spatial rainfall patterns (Figs. 2, 4, and 11),

rainfall rate distribution (Fig. 3), and diurnal cycle of precipitation (Figs. 5 and 12).

We calculated the asymptote fractions skill score (Roberts and Lean 2008) to determine whether the models produced rainfall in the right place at the right time, on daily and 3-hourly scales, for the Highveld region, which is several 100 km across. The three models had comparable skill on the daily scale, which may be due to the domain-averaged rainfall being well constrained by the driving model. On the 3-hourly scale, the CPMs performed better than the SAGA6, which is expected from the better representation of the diurnal cycle of precipitation. Similar results were obtained for East Africa by Woodhams et al. (2018), who found that CPMs outperformed a global model when considering 3-hourly rainfall accumulations, but not daily accumulations. The target score for which their precipitation forecast was deemed “skillful” was not achieved until a horizontal scale of several hundreds of kilometers, similar to the scale of the Highveld domain considered in this paper.

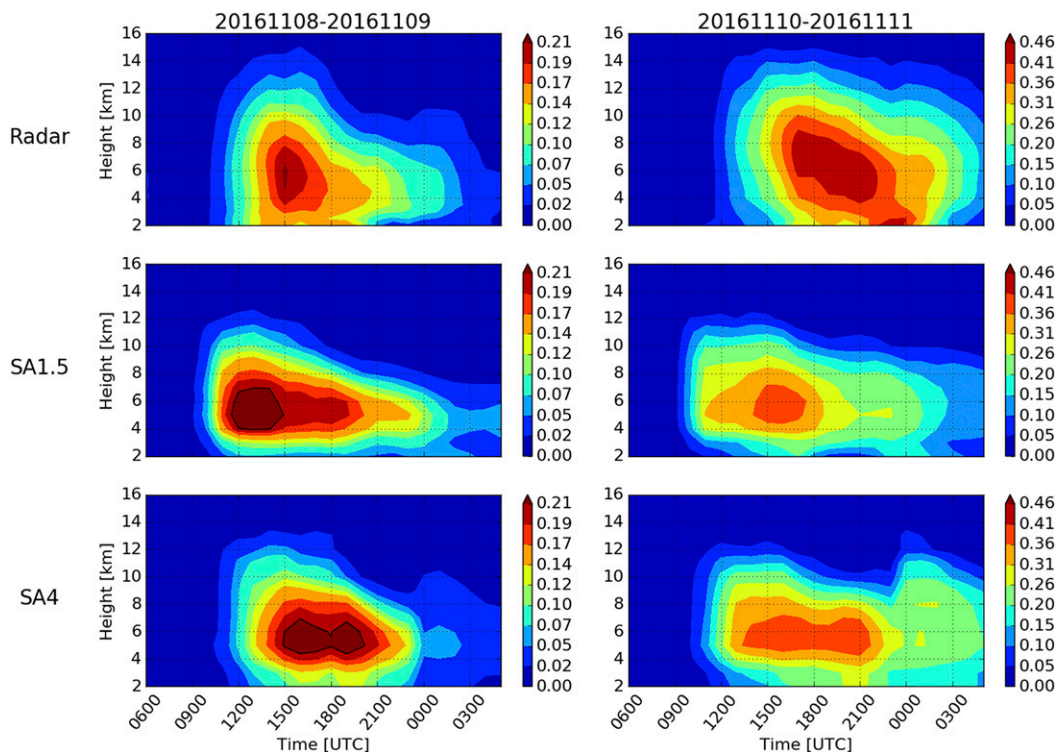


FIG. 13. Diurnal cycle of cloud and precipitation fraction for (left) 8 and (right) 10 Nov 2016 using a 10-dBZ threshold for (top) the SAWS radar observations, (middle) SA1.5, and (bottom) SA4. Radar observations are considered for each 6-min composite where all relevant radars are available for that case study day (see text); model simulations are considered hourly for that same domain. Heights are above mean sea level. The time axis starts at 0600 UTC on the first day and ends at 0600 UTC on the second day; for the model simulations, this covers hours 12–36 relative to the start of the simulation (1800 UTC on the previous day).

It is important to note that the two CPMs are largely indistinguishable only on the temporal and spatial scales considered. Our analysis does not consider individual convective clouds or rainfall extremes, for which a model at 1.5-km grid length may reasonably be expected to outperform a model at 4-km grid length and both should outperform a model with parameterized convection. To test the models on smaller temporal and spatial scales, it is necessary to obtain a verified and quality-controlled radar-based rainfall product, as used by Roberts and Lean (2008). Furthermore, Woodhams et al. (2018) suggest that running these models with data assimilation should help distinguish the performance between the CPMs.

Our conclusions are drawn from a range of satellite-based rainfall products, a multiyear *CloudSat* cloud climatology, and the SAWS ground-based radar network. This multi-instrument evaluation has allowed us to study the consistency in model biases as demonstrated by the different observational data on different time and length scales. The satellite-based rainfall products may each be used individually to evaluate the spatial pattern of rainfall on monthly time scales, and significant pattern

correlations are found even on daily time scales, but the four products considered vary significantly in terms of accumulations, with a difference in the median of approximately 100 mm on the monthly scale and 10 mm in daily accumulations. The *CloudSat* climatology could not easily be compared against the rainfall climatologies due to different sampling and different observation periods. For future model evaluation, especially on a global scale, the following should be considered:

- 1) Long-term, IR-based rainfall estimates such as CHIRPS and TAMSAT are particularly useful for assessing interannual/multidecadal variability, given their longevity and temporal stability (especially TAMSAT). Such products are also calibrated using a reliable climatology, ensuring the estimates reflect the local rainfall climate.
- 2) Rainfall estimates incorporating passive and active microwave imagery such as IMERG and CMORPH have the ability to sense the presence of raindrops and, hence, can provide a better estimate of rainfall intensity. As such, these datasets provide a reliable estimate for extreme rainfall events.

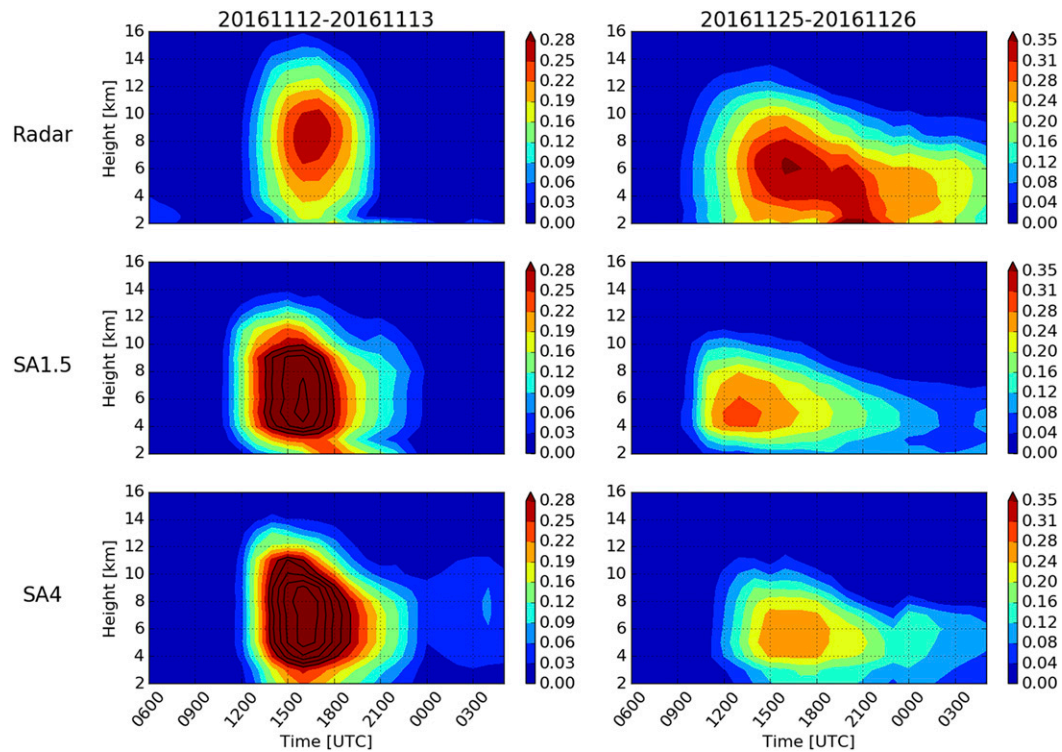


FIG. 14. As in Fig. 13, but for 12 and 25 Nov 2016.

- 3) *CloudSat* can inform on the relationship between the location of different cloud types, such as midlevel clouds upstream of cumulonimbus and anvil clouds downstream, while model biases in oceanic cloud cover suggest areas for further investigation with detailed model process studies and new and targeted observations.
- 4) A calibrated radar network provides a unique opportunity to evaluate the vertical development of precipitating clouds, to be considered alongside the diurnal cycle of precipitation.

The result that CPMs generally improve the timing of the diurnal cycle and rainfall amounts compared to a model with convection parameterization agrees with findings when running the UM over West Africa (Pearson et al. 2010; Birch et al. 2014). This is not a foregone conclusion, however, especially when considering individual forecasts. Indeed, the SAGA6 performed well and arguably better than the CPMs for one of the case studies. Our analysis is also restricted to springtime convection over a subtropical continental region, and different conclusions may be drawn when different convective regions are considered, such as tropical maritime or midlatitude convection. James et al. (2018) argue that process-oriented analysis of region-specific phenomena is essential to improving our understanding of model behavior over Africa.

The reliable physical representation of precipitating clouds and convection in the SA1.5 and SA4, as demonstrated in this paper, means that these forecasts present a valuable resource for understanding model behavior for convective-scale weather phenomena across southern Africa.

Acknowledgments. This work and its contributors (Stein, Keat, Maidment, Boyd, and Pankiewicz) were supported by the U.K.–South Africa Newton Fund through the Met Office Weather and Climate Science for Service Partnership (WCSSP) South Africa. The authors wish to acknowledge the support from the Met Office Unified Model Partnership, in particular Humphrey Lean and the Working Group on Convective-scale Modelling. This work used the ARCHER U.K. National Supercomputing Service (<http://www.archer.ac.uk>). The *CloudSat* data were analyzed using the DARDAR merged products, which were accessed through the ICARE Data and Services Center, available online (icare.univ-lille1.fr/projects/dardar).

REFERENCES

- Abel, S. J., and I. A. Boutle, 2012: An improved representation of the raindrop size distribution for single-moment microphysics schemes. *Quart. J. Roy. Meteor. Soc.*, **138**, 2151–2162, <https://doi.org/10.1002/qj.1949>.

- 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>.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Becker, E. H., 2014: Application of a quantitative precipitation estimation algorithm for the S-band radar at Irene, South Africa. M.S. thesis, Dept. of Civil Engineering, University of KwaZulu-Natal, Durban, South Africa, 125 pp., <http://hdl.handle.net/10413/12795>.
- Birch, C. E., D. Parker, J. Marsham, D. Copsey, and L. Garcia-Carreras, 2014: 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>.
- Bodas-Salcedo, A., M. J. Webb, M. E. Brooks, M. A. Ringer, K. D. Williams, S. F. Milton, and D. R. Wilson, 2008: Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities. *J. Geophys. Res.*, **113**, D00A13, <https://doi.org/10.1029/2007JD009620>.
- , and Coauthors, 2011: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.*, **92**, 1023–1043, <https://doi.org/10.1175/2011BAMS2856.1>.
- 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://dx.doi.org/10.1002/met.1538>.
- Cox, G. P., 1988: Modelling precipitation in frontal rainbands. *Quart. J. Roy. Meteor. Soc.*, **114**, 115–127, <https://doi.org/10.1002/qj.49711447906>.
- Delanoë, J., R. J. Hogan, R. M. Forbes, A. Bodas-Salcedo, and T. H. M. Stein, 2011: Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and CALIPSO data. *Quart. J. Roy. Meteor. Soc.*, **137**, 2064–2078, <https://doi.org/10.1002/qj.882>.
- Dezfuli, A. K., C. M. Ichoku, G. J. Huffman, K. I. Mohr, J. S. Selker, N. van de Giesen, R. Hochreutener, and F. O. Annor, 2017: Validation of IMERG precipitation in Africa. *J. Hydrometeorol.*, **18**, 2817–2825, <https://doi.org/10.1175/JHM-D-17-0139.1>.
- Dinku, T., C. Funk, P. Peterson, R. Maidment, T. Tadesse, H. Gadain, and P. Ceccato, 2018: Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quart. J. Roy. Meteor. Soc.*, **144** (Suppl. 1), 292–312, <https://doi.org/10.1002/qj.3244>.
- Dixon, M., and G. Wiener, 1993: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting—A radar-based methodology. *J. Atmos. Oceanic Technol.*, **10**, 785–797, [https://doi.org/10.1175/1520-0426\(1993\)010<0785:TTITAA>2.0.CO;2](https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2).
- Dyson, L. L., 2009: Heavy daily-rainfall characteristics over the Gauteng Province. *Water SA*, **35**, 627–638, <https://doi.org/10.4314/wsa.v35i5.49188>.
- Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51**, 249–280, [https://doi.org/10.1175/1520-0469\(1994\)051<0249:ADMMPF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<0249:ADMMPF>2.0.CO;2).
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting*, **13**, 377–395, [https://doi.org/10.1175/1520-0434\(1998\)013<0377:TWRA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2).
- Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, <https://doi.org/10.1038/sdata.2015.66>.
- Gijben, M., 2012: The lightning climatology of South Africa. *S. Afr. J. Sci.*, **108**, 44–53.
- Holleman, I., A. Huuskonen, M. Kurri, and H. Beekhuis, 2010: Operational monitoring of weather radar receiving chain using the sun. *J. Atmos. Oceanic Technol.*, **27**, 159–166, <https://doi.org/10.1175/2009JTECHA1213.1>.
- Huffman, G. J., 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>.
- , D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, and S.-H. Yoo, 2018: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD), version 5.2, NASA, 31 pp., https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_ATBD_V5.pdf.
- James, R., and Coauthors, 2018: Evaluating climate models with an African lens. *Bull. Amer. Meteor. Soc.*, **99**, 313–336, <https://doi.org/10.1175/BAMS-D-16-0090.1>.
- Jobard, I., F. Chopin, J. C. Berges, and R. Roca, 2011: An intercomparison of 10-day satellite precipitation products during West African monsoon. *Int. J. Remote Sens.*, **32**, 2353–2376, <https://doi.org/10.1080/01431161003698286>.
- Joyce, R. J., 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).
- Kain, J. S., S. J. Weiss, J. J. Levit, M. E. Baldwin, and D. R. Bright, 2006: Examination of convection-allowing configurations of the WRF Model for the prediction of severe convective weather: The SPC/NSSL Spring Program 2004. *Wea. Forecasting*, **21**, 167–181, <https://doi.org/10.1175/WAF906.1>.
- , and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- Kay, J., and Coauthors, 2012: Exposing global cloud biases in the Community Atmosphere Model (CAM) using satellite observations and their corresponding instrument simulators. *J. Climate*, **25**, 5190–5207, <https://doi.org/10.1175/JCLI-D-11-00469.1>.
- Khodadoust Siuki, S., B. Saghafian, and S. Moazami, 2017: Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products. *Int. J. Remote Sens.*, **38**, 558–571, <https://doi.org/10.1080/01431161.2016.1268735>.
- Laing, A. G., and J. M. Fritsch, 1993: Mesoscale convective complexes in Africa. *Mon. Wea. Rev.*, **121**, 2254–2263, [https://doi.org/10.1175/1520-0493\(1993\)121<2254:MCCIA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<2254:MCCIA>2.0.CO;2).
- Liu, Z., R. Marchand, and T. Ackerman, 2010: A comparison of observations in the tropical western Pacific from ground-based and satellite millimeter-wavelength cloud radars. *J. Geophys. Res.*, **115**, D24206, <https://doi.org/10.1029/2009JD013575>.
- Locatelli, J. D., and P. V. Hobbs, 1974: Fall speeds and masses of solid precipitation particles. *J. Geophys. Res.*, **79**, 2185–2197, <https://doi.org/10.1029/JC079i015p02185>.
- 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](https://doi.org/10.1175/1520-0493(2000)128<3187:ANBLMS>2.0.CO;2).
- Mace, G. G., Q. Zhang, M. Vaughan, R. Marchand, G. Stephens, C. Trepte, and D. Winker, 2009: A description of hydrometeor layer occurrence statistics derived from the first year of merged CloudSat and CALIPSO data. *J. Geophys. Res.*, **114**, D00A26, <https://doi.org/10.1029/2007JD009755>.
- Maidment, R. L., D. I. Grimes, R. P. Allan, H. Greatrex, O. Rojas, and O. Leo, 2013: Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda. *Meteor. Appl.*, **20**, 308–317, <https://doi.org/10.1002/met.1283>.
- , D. Grimes, R. P. Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling, and E. Black, 2014: The 30 year TAMSAT African Rainfall Climatology and Time Series (TARCAT) data set. *J. Geophys. Res. Atmos.*, **119**, 10 619–10 644, <https://doi.org/10.1002/2014JD021927>.
- , and Coauthors, 2017: A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa. *Sci. Data*, **4**, 170063, <https://doi.org/10.1038/sdata.2017.63>.
- Manz, B., S. Páez-Bimos, N. Horna, W. Buytaert, B. Ochoa-Tocachi, W. Lavado-Casimiro, and B. Willems, 2017: Comparative ground validation of IMERG and TMPA at variable spatiotemporal scales in the tropical Andes. *J. Hydrometeorol.*, **18**, 2469–2489, <https://doi.org/10.1175/JHM-D-16-0277.1>.
- Marchand, R., G. G. Mace, T. Ackerman, and G. Stephens, 2008: Hydrometeor detection using Cloudsat—An earth-orbiting 94-GHz cloud radar. *J. Atmos. Oceanic Technol.*, **25**, 519–533, <https://doi.org/10.1175/2007JTECHA1006.1>.
- , J. Haynes, G. G. Mace, T. Ackerman, and G. Stephens, 2009: A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations. *J. Geophys. Res.*, **114**, D00A20, <https://doi.org/10.1029/2008JD009790>.
- Mitchell, D. L., 1996: Use of mass- and area-dimensional power laws for determining precipitation particle terminal velocities. *J. Atmos. Sci.*, **53**, 1710–1723, [https://doi.org/10.1175/1520-0469\(1996\)053<1710:UOMAAD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<1710:UOMAAD>2.0.CO;2).
- Nguyen, H., C. Franklin, and A. Protat, 2017: Understanding the ACCESS model errors over the Maritime Continent using CloudSat and CALIPSO simulators. *Quart. J. Roy. Meteor. Soc.*, **143**, 3136–3152, <https://doi.org/10.1002/qj.3168>.
- Pearson, K. J., R. J. Hogan, R. P. Allan, G. M. S. Lister, and C. E. Holloway, 2010: Evaluation of the model representation of the evolution of convective systems using satellite observations of outgoing longwave radiation. *J. Geophys. Res.*, **115**, D20206, <https://doi.org/10.1029/2010JD014265>.
- 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**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S.-I. Iga, 2008: Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *J. Comput. Phys.*, **227**, 3486–3514, <https://doi.org/10.1016/j.jcp.2007.02.006>.
- , T. Inoue, and H. Miura, 2010: Evaluations of cloud properties of global and local cloud system resolving models using CALIPSO and CloudSat simulators. *J. Geophys. Res.*, **115**, D00H14, <https://doi.org/10.1029/2009JD012247>.
- Simpson, L.-A., and L. L. Dyson, 2018: Severe weather over the Highveld of South Africa during November 2016. *Water SA*, **44**, 75–85, <https://doi.org/10.4314/wsa.v44i1.09>.
- Stein, T. H. M., D. J. Parker, J. Delanoë, N. S. Dixon, R. J. Hogan, P. Knippertz, R. I. Maidment, and J. H. Marsham, 2011: The vertical cloud structure of the West African monsoon: A 4 year climatology using CloudSat and CALIPSO. *J. Geophys. Res.*, **116**, D22205, <https://doi.org/10.1029/2011JD016029>.
- , R. J. Hogan, K. E. Hanley, J. C. Nicol, H. W. Lean, R. S. Plant, P. A. Clark, and C. E. Halliwell, 2014: The three-dimensional morphology of simulated and observed convective storms over southern England. *Mon. Wea. Rev.*, **142**, 3264–3283, <https://doi.org/10.1175/MWR-D-13-00372.1>.
- , D. J. Parker, R. J. Hogan, C. E. Birch, C. E. Holloway, G. M. S. Lister, J. H. Marsham, and S. J. Woolnough, 2015: The representation of the West African monsoon vertical cloud structure in the Met Office Unified Model: An evaluation with CloudSat. *Quart. J. Roy. Meteor. Soc.*, **141**, 3312–3324, <https://doi.org/10.1002/qj.2614>.
- Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1500, <https://doi.org/10.1175/2009BAMS2795.1>.
- 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>.
- Tarnavsky, E., D. Grimes, R. Maidment, E. Black, R. P. Allan, M. Stringer, R. Chadwick, and F. Kayitakire, 2014: Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present. *J. Appl. Meteor. Climatol.*, **53**, 2805–2822, <https://doi.org/10.1175/JAMC-D-14-0016.1>.
- 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**, 3–25, [https://doi.org/10.1016/S0022-1694\(00\)00372-3](https://doi.org/10.1016/S0022-1694(00)00372-3).
- Walters, D., and Coauthors, 2017: 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>.
- Wang, S., and Coauthors, 2015: Evaluation of remotely sensed precipitation and its performance for streamflow simulations in basins of the southeast Tibetan Plateau. *J. Hydrometeorol.*, **16**, 2577–2594, <https://doi.org/10.1175/JHM-D-14-0166.1>.
- Webb, M., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.*, **17**, 905–922, <https://doi.org/10.1007/s003820100157>.
- Wilson, D. R., and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Quart. J. Roy. Meteor. Soc.*, **125**, 1607–1636, <https://doi.org/10.1002/qj.49712555707>.
- Woodhams, B. J., C. E. Birch, J. H. Marsham, C. L. Bain, N. M. Roberts, and D. F. A. 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>.
- 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. Hydrometeorol.*, **18**, 1617–1641, <https://doi.org/10.1175/JHM-D-16-0168.1>.
- 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>.
- Zipsper, E. J., D. J. Cecil, C. Liu, S. W. Nesbitt, and D. P. Yorty, 2006: Where are the most intense thunderstorms on Earth? *Bull. Amer. Meteor. Soc.*, **87**, 1057–1072, <https://doi.org/10.1175/BAMS-87-8-1057>.