A multi-satellite climatology of clouds, radiation and precipitation in southern West Africa and comparison to climate models

Peter G. Hill[1], Richard P. Allan[1], J. Christine Chiu[1], Thorwald H. M. Stein[1]

[1]Department of Meteorology, University of Reading, Reading, UK

Corresponding author: Peter Hill (p.g.hill@reading.ac.uk)

Key Points:

- Satellite estimates of cloud cover at all heights show large differences

- Inter-annual variability of surface precipitation and dry static energy divergence are strongly linked

- AMIP5 models are unable to reproduce the observed mean cloud, radiation and precipitation in Southern West Africa

Abstract

Southern West Africa (SWA) has a large population that relies on highly variable monsoon rainfall, yet climate models show little consensus over projected precipitation in this region. Understanding of the current and future climate of SWA is further complicated by rapidly increasing anthropogenic emissions and a lack of surface observations. Using multiple satellite observations, the ERA-Interim reanalysis, and four climate models, we document the climatology of cloud, precipitation and radiation over SWA in June-July, highlight discrepancies among satellite products, and identify shortcomings in climate models and ERA-Interim. Large differences exist between monthly mean cloud cover estimates from satellites, which range from 68 to 94 %. In contrast, differences among satellite observations in top of atmosphere outgoing radiation and surface precipitation are smaller, with monthly means of about 230 W m^{-2} of longwave radiation, 145 W m^{-2} of shortwave radiation and 5.87 mm day^{-1} of precipitation. Both ERA-Interim and the climate models show less total cloud cover than observations, mainly due to underestimating low cloud cover. Errors in cloud cover, along with uncertainty in surface albedo, lead to a large spread of outgoing shortwave radiation. Both ERA-Interim and the climate models also show signs of convection developing too early in the diurnal cycle, with associated errors in the diurnal cycles of precipitation and outgoing longwave radiation. Clouds, radiation and precipitation are linked in an analysis of the regional energy budget, which shows that inter-annual variability of precipitation and dry static energy divergence are strongly linked.

Introduction

West Africa is currently experiencing a rapid growth in population and associated air pollution; it is also subject to the substantial variability of the West African Monsoon (WAM) rains, making this region particularly vulnerable to climatic impacts relating to health and food security [Knippertz et al., 2015b]. Monsoon precipitation and related large-scale processes show variability at both intra-seasonal [Sultan et al., 2003; Gu and Adler, 2004] and inter-annual [Sylla et al., 2009; Jianping and Qingcun, 2003] scales. This rainfall variability affects regional crop yields [Sultan et al., 2005a] and has been connected to meningitis outbreaks [Sultan et al., 2005b]. The WAM is intimately linked with the global atmospheric circulation and consequently its impacts spread beyond Africa. For example, modelling studies have shown that a strong monsoon contributes to the positive phase of the North Atlantic Oscillation [Gaetani et al., 2011].

The most populous part of West Africa is predominantly urban and concentrated along the southern coastal regions [Nelson, 2004] yet there is a general lack of surface observations in southern West Africa (SWA) and the region remains understudied [Knippertz et al., 2015a]. Consequently this highly populated region remains relatively poorly understood and the complex interplay between clouds, precipitation, radiation, aerosol pollution and the WAM further motivate detailed investigations into the climate and the physical drivers of variability across this region.

The West African monsoon season runs between March and October. During this period the southern coastal regions experience two rainy seasons separated by a short dry season [Le Barbé et al., 2002]. The first wet season runs from March to June, with maximum precipitation along the coast in May. The intertropical convergence zone (ITCZ) and consequently much of the precipitation shifts further north in June/July, which leads to increased cloudiness and is the focus period for this study. With the ITCZ to the north, the coastal regions experience a short dry season during July and August, which is linked to stronger subsidence associated with outflows from deep convection to the north [Omotosho, 1988]. As the ITCZ retreats back towards the coast, bringing a second wet season from September to October.

During the monsoon season, SWA experiences a complex pattern of cloud types with a pronounced diurnal cycle. Stein et al. [2011] identified numerous distinct cloud types in this region, from CloudSat and CALIPSO, with notable differences in cloud type, frequency of occurrence and amount when present between the day and night overpasses. Moreover, low-level stratus has been found to occur frequently, with maximum extent between 0900 and 1000 UTC and minimum around 1800 UTC [van der Linden et al., 2015; Schrage et al., 2006].

General circulation models struggle to correctly simulate the complex observed cloud and precipitation climatology in SWA. Hourdin et al. [2010] showed large accumulated precipitation differences between the models that participated in the African Monsoon Multi-disciplinary Analysis Model Intercomparison Project throughout West Africa. Knippertz et al. [2011] found that Coupled Model Intercomparison Project phase 3 (CMIP3) models show little consensus on the mean July-September cloud cover in this region and on average allow too much solar radiation to reach the surface. Cloud and

precipitation errors persist in the CMIP5 models, which show large zonal mean precipitation and cloud radiative effect differences throughout West Africa and have problems with the timing and magnitude of the diurnal cycle of precipitation [Roehrig et al., 2013]. Even numerical weather prediction models at convection-permitting resolution can have large cloud errors in SWA; Stein et al. [2015] showed that the Met Office Unified Model (MetUM) overestimates low cloud cover and underestimates anvil cloud cover at all resolutions from 40 − 1.5 km. Cloud errors in SWA can have a large impact on radiation [Knippertz et al., 2011], which in turn may affect the broader circulation [Li et al., 2015; Marsham et al., 2013]. In light of these problems simulating present day climate, it is perhaps unsurprising that projections of future precipitation in West Africa show large diversity, for both regional [Paeth et al., 2011] and global coupled [Cook and Vizy, 2006] models.

The uncertainty arising from existing model deficiencies is compounded by rapid increases in anthropogenic emissions in SWA; aerosol concentrations in SWA are poorly understood, as are their interactions with clouds [Knippertz et al., 2015b]. The Dynamics–Aerosol–Chemistry–Cloud Interactions in West Africa (DACCIWA) project [Knippertz et al., 2015a] aims to improve our understanding of the physical and chemical processes in this region, through an extensive field campaign in SWA, including a multi-aircraft campaign in June-July 2016, and a broad range of modelling studies.

The present study aims to document and understand the climatology of clouds, radiation, precipitation and their interconnections by combining observations and simulations in June-July over the region of West Africa centered upon 5–10°N and 8°W–8°E, which we shall hereafter refer to as the DACCIWA region. The primary motivation is to improve understanding of the physical processes determining the climate of this region, to highlight deficiencies in the reanalysis and climate model simulations and to provide context for and interpretation of upcoming planned observational campaigns such as the DACCIWA project. While precipitation is of great importance for health and agriculture, links between the water and energy cycles and the complex interaction between cloud, aerosol and radiation require a holistic approach in evaluating the climate of this region. For this purpose we rely upon satellite measurements which are able to capture the large-scale structure of clouds and radiation and greatly add to the limited surface observations available. Advancing on previous studies [Knippertz et al., 2011; Stein et al., 2011;

Roehrig et al., 2013; van der Linden et al., 2015] we make detailed assessment of multiple satellite products and consider state of the art climate model simulations including their representation of the diurnal cycle within a focused region. Driving questions are: what are the primary characteristics of the climatology of this region, which satellite observations are most suitable for analysis, what are the primary model biases and what aspects of the climate system are poorly understood for this region? This analysis therefore also seeks to inform and prioritize future and ongoing observations and research undertaken as part of the DACCIWA project. The following sections outline the observations used and detail the climatology of clouds, radiation and precipitation and their diurnal cycle in observations, reanalyses and climate model simulations.

1. Datasets

This study considers the wider Southern West Africa region shown in Figure 1, but focuses in particular on the area $8°W - 8°E$ and $5 - 10°N$, as highlighted by the dotted boxes, which we refer to as the DACCIWA region. These boundaries were chosen to avoid the Guinea highlands to the West, the Jos plateau to the East and the Gulf of Guinea to the South. This results in a domain that is large enough to reduce statistical sampling errors, but sufficiently homogeneous for domain mean values to be meaningful and matches that used for averaging in the modelling component of the DACCIWA project. The DACCIWA region consists of forests and cropland near the coast, with wooded savannas further inland and savannas to the north of the domain. We focus on the months of June and July coinciding with the progression of the summer monsoon and for consistency with the planned DACCIWA aircraft campaign [Knippertz et al., 2015a].

As noted in the introduction, June-July is a transition period for the DACCIWA region. However, June-July differences (see supplementary material) are generally quite small. Cloud cover differences between June and July are much smaller than the differences between different satellite cloud cover products discussed in the following section. Moreover, observed differences in clouds, precipitation and top of atmosphere radiation between June and July are smaller than the errors in the climate models and ERA-Interim discussed in section 4.

We use seven observational datasets including retrievals from both active and passive sensors, and from both sun synchronous and geostationary orbiting satellites. Expanded acronyms and key properties for these datasets are summarized in Table 1. As few of the datasets exist prior to 2000, we focus on 2000-present. For ease of comparison, the observational datasets, ERA-Interim and the climate models are all re-gridded to a resolution of 1°.

Longwave cloud radiative effects are strongly dependent on the temperature and consequently height of a cloud. To facilitate our understanding of cloud-radiation interactions, we consider cloud cover at different heights, with separate values of cloud cover for high clouds (tops above 440 hPa), mid-level clouds (tops between 440 and 800 hPa) and low clouds (tops below 800 hPa). These thresholds were chosen as they are already in use in many of the datasets considered and consequently can be applied consistently across climate models, ERA-Interim and observations.

1.1. Satellite products

The DARDAR [Delanoe and Hogan, 2010] product is based on combined CloudSat radar [Stephens et al., 2002] and CALIPSO lidar [Winker et al., 2009] observations at 1:30 and 13:30 local time, providing vertically-resolved cloud properties. While the DARDAR product is based on active measurements that can detect multiple cloud layers, the other satellite cloud products use observations from passive instruments that are only able to detect the highest cloud layer and cannot detect lower clouds beneath higher clouds. Consequently only the high clouds from the passive satellite datasets can be usefully compared to DARDAR. To ensure a meaningful comparison between DARDAR and the passive observations, we remove potentially obscured clouds from DARDAR. This only affects the mid-level and low cloud cover from DARDAR. We shall refer to this as DARDAR-passive-like.

The MODIS dataset is based on radiometer observations from two polar orbiting satellites; Terra was launched in 2000 and crosses the equator at approximately 10:30 and 22:30 local time, while Aqua was launched in 2002 and crosses the equator at approximately 1:30 and 13:30 local time. We use the monthly global product, collection 6 [Platnick, 2015], which includes monthly statistics at a resolution of 1° x 1°. For total cloud cover, we use the mean cloud fraction, which is derived from the instantaneous cloud mask [Frey et al., 2008] and

includes both day and night data. For our analysis of cloud cover at different heights we calculate high, mid-level and low cloud cover from the cloud fraction-pressure histogram, which is based on the instantaneous cloud top pressure [Platnick et al., 2003] . The cloud top pressure is determined by a CO_2 slicing technique using multiple spectral bands between 13.3 and 15 μm. For low clouds, if the CO_2 slicing technique fails to give a valid solution, then the infrared window approach is used, where the brightness temperature at 11 μm is used to determine a cloud top temperature, and temperature profiles from the National Centers for Environmental Prediction Global Data Assimilation System are used to convert this to a cloud top pressure. This ensures that the cloud top pressure is determined for all cloudy points.

The CMSAF-CLAAS (Climate Monitoring Satellite Applications Facility - CLoud property dAtAset using SEVIRI (Spinning Enhanced Visible and InfraRed Imager)) product is based on instruments on the geostationary Meteosat Second Generation satellites [Stengel et al., 2014]. The geostationary orbit provides excellent diurnal sampling; the CMSAF dataset has excellent potential for understanding cloud diurnal variability. However, as a relatively new dataset, there is likely to be potential for further refinements to the algorithm. We use the instantaneous cloud mask and cloud top pressure to calculate monthly mean total, high, mid-level and low cloud cover for West Africa at a resolution of 1° x 1°. The cloud detection algorithm [Derrien and Le Gleau, 2005] is based on visible and infrared threshold tests in multiple bands combined with spatial variability tests. The threshold tests use eight bands, compared to the 22 used in the MODIS cloud mask algorithm. Like MODIS, cloud top pressure is determined using a combination of the infrared window approach and CO_2 slicing. However SEVIRI lacks the multiple CO_2 channels used in the MODIS algorithm and CO_2 slicing is used much less frequently than in MODIS. The infrared window approach is used for low and mid-level thick clouds. For semi-transparent high clouds, the infrared window intercept method is applied. This is similar to CO_2 slicing, but uses infrared and water vapor channels. CO_2 slicing is only applied for optically thick high clouds or when the infrared window intercept method fails.

Note that the cloud top pressure in the CMSAF-CLAAS product is not estimated for all cloudy pixels. If a pixel contains fractional clouds (which are mainly low-level water clouds according to Stengel et al. [2014], then no cloud top pressure is determined. Consequently the mid-level, low and high cloud cover, which are estimated using the cloud top pressure do not sum to the total cloud cover, which is based on the cloud mask.

We compare three precipitation datasets. The first dataset, the TRMM (Tropical Rainfall Measuring Mission) Multi-satellite Precipitation Analysis (TMPA) [Huffman et al., 2007] has been identified as one of the best satellite rainfall datasets over West Africa [Roca et al., 2010; Pfeifroth et al., 2016]. We use the 3B42 data that provides precipitation estimates every three hours. TMPA combines microwave precipitation estimates from a variety of satellites (which are intercalibrated to precipitation estimates directly from the TRMM satellite) with geostationary infrared precipitation estimates created using the calibrated microwave precipitation. To ensure good performance at the monthly mean scale, the precipitation estimates are subsequently rescaled to match monthly mean gauge measurements.

The second precipitation dataset considered is the Global Precipitation Measurement (GPM) mission [Hou et al., 2014], which is the follow up to TRMM. The Integrated Multi-satellitE Retrievals for GPM (IMERG) dataset [Huffman et al., 2013] used in this study uses a similar algorithm to the TMPA product. GPM was launched in February 2014 and data is available from March 2014. Consequently our comparison with TRMM is limited to June-July 2014. Nevertheless, we think this provides a useful initial examination of the diurnally-resolved differences between the datasets in this region. Initial global comparisons on a monthly scale show that TMPA-IMERG differences are small over land [Liu, 2016]. Quantifying uncertainty in merged precipitation products such as TMPA and GPM IMERG is difficult and remains an area of ongoing research [Sorooshian et al., 2011].

The third precipitation dataset considered is the Global Precipitation Climatology Centre (GPCC) precipitation reanalysis product [Schneider et al., 2014]. This is based on quality-controlled gauge records obtained from 75,000 surface stations between 1900 and 2013. The GPCC reanalysis product is used to calibrate the TRMM product over land at the monthly mean scale up to 2010. For later years, both TRMM and GPM use the GPCC monitoring product, which is based on fewer gauges. Although the number of gauges used in the GPCC analysis product in West Africa is relatively small, Nicholson et al. [2003] showed that it

compares well with a denser rain gauge network. However, the GPCC product only provides estimates at the monthly mean scale, with no diurnal cycle. Consequently we prefer the TRMM and GPM datasets, which by design match the monthly mean GPCC values, but provide sub-daily data for analyzing the diurnal cycle of precipitation.

Note that satellite precipitation products generally capture convective precipitation, but struggle to detect warm rain (precipitation produced by liquid clouds) events [Ebert et al., 2007]. The use of gauges in the TRMM product may improve total biases at the expense of increasing the precipitation amount when present to compensate for any undetected precipitation events [Tian et al., 2009]. The sparsity of surface observations may also affect the performance of both TMPA and GPM, which depend on local gauge calibration [Worqlul et al., 2014].

For top of atmosphere (TOA) radiation, we use CERES [Wielicki et al., 1996] and GERB [Harries et al., 2005] datasets. The CERES instrument measures radiances in three channels: total (0.3-200 μm), SW (0.3-5 μm) and LW window (8-12 μm). These radiances are converted to instantaneous TOA flux estimates using a scene-dependent angular distribution model (ADM) as described by Loeb et al. [2003]. CERES instruments are limited to the same two polar orbiting satellites as MODIS, and consequently only sample the diurnal cycle at four times. We use the synoptic radiative fluxes and clouds (CERES-SYN) product [Doelling et al., 2013], which capitalizes on geostationary imagers to estimate the diurnal variability between CERES observations, using coincident geostationary and MODIS/CERES observations to ensure CERES calibration is maintained. The dominant source of uncertainty in CERES TOA radiation measurements relates to absolute calibration [Loeb et al., 2009], which is estimated to be 2% for the SW channel and 1% for the total channel [Priestley et al., 2002]. Clear-sky fluxes are estimated as the mean for those CERES footprints for which the cloud fraction is less than 1%, where the cloud mask is estimated using MODIS [Minnis et al., 2011].

The GERB dataset [Dewitte et al., 2008] is based on a series of GERB instruments on the geostationary Meteosat Second Generation Satellites. GERB2 was launched in August 2002 and began operations in February 2004 while GERB1 was launched in December 2005 and began operations in May 2007. The HR (high resolution) product used in this study has a temporal resolution of 15 minutes. The GERB instruments measure radiances in two

channels: total (0.3 – 100 μm) and SW (0.3 – 4 μm), and LW radiance is calculated by subtraction SW from the total. The theoretical uncertainty for instantaneous unfiltered radiances is about 2% in the SW and 1% in the LW [Clerbaux et al., 2009].

To convert radiance to flux, the CERES ADMs in the SW and theoretical ADMs [Clerbaux, 2003] in the LW were used, along with the scene identified from SEVIRI. Comparison of instantaneous fluxes for earlier editions of the GERB and CERES products [Clerbaux et al., 2009] found reasonable agreement between GERB2 and CERES in the LW, but larger differences in the SW, which were attributed to differences in the absolute calibration accuracy; following the GERB2 data quality summary recommendations we apply an additional calibration factor which reduces this difference.

Unlike CERES products, the GERB product does not report SW fluxes for solar zenith angles larger than 80°. To complete the diurnal cycle, we fill this missing data using mean twilight values from CERES [Kato, 2003] for zenith angles between 86.5° and 104.5°. Kato [2003] does not report fluxes at zenith angles less than 86.5°, so we use linear interpolation in time to fill zenith angles between 80° and 86.5°.

Estimates of the radiative heating of the atmosphere are required to understand the atmospheric energy budget and links between radiative and latent heating. The atmospheric radiative heating is calculated by subtracting net downwards surface radiative fluxes from the net downwards top of atmosphere radiative fluxes. Surface fluxes are not measured by satellites, but are calculated using radiative transfer models in some CERES products. For consistency, atmospheric radiative heating is calculated using surface and top of atmosphere fluxes from the same product. In particular, we use the CERES-EBAF [Loeb et al., 2009] product and the CERES-COMPUTED [Rutan et al., 2015]product. We do not use the CERES-SYN product employed in the analysis of top of atmosphere fluxes as it does not have consistent surface flux estimates.

1.2. ERA-Interim

Reanalyses provide spatially and temporally complete estimates of the state of the atmosphere, and estimates of atmospheric motion that are difficult to observe. For regions of sparse observations such as southern West Africa, reanalyses are particularly useful, although caution needs to be exercised since they are likely to be less reliable due to the lack of conventional observations to assimilate.

The European Centre for Medium-range Weather Forecasts (ECMWF) ERA-Interim [Dee et al., 2011] re-analysis runs from 1979 and is continuously updated. If, as Knippertz et al. [2011] suggest, ERA-Interim performs reasonably in the DACCIWA region, then it potentially provides a useful tool for understanding errors in other models. The analysis is computed using a 2006 release of ECMWF's Integrated Forecasting System (IFS), which has a horizontal resolution of approximately 80 km and 60 vertical levels. For comparison to the satellite clouds products, we calculated high, mid-level and low cloud cover (as seen from space) from the cloud cover on model-levels, assuming maximum-random overlap [Geleyn and Hollingsworth, 1979]. The contribution of horizontal energy transport to the regional energy budget, as discussed in section 3.2.2, is approximated by the column-integrated divergence of dry static energy flux and calculated from ERA-Interim with a mass correction following Liu et al. [2015].

Note that ERA-Interim analysis fields are available at most every 6 hours and only every 12 hours for some fields. To fill in the missing points in the diurnal cycle, we use data from the ERA-Interim forecast, which is available every three hours. We also use ERA-Interim forecast data for variables that are output as accumulated fluxes and therefore not available at the analysis time (e.g. radiative fluxes and surface precipitation). Differences between ERA-Interim forecast and analysis fields as noted by Knippertz et al. [2011] are beyond the scope of this study.

1.3. Climate models

To investigate errors in state of the art climate models, we use simulations from the coupled model intercomparison project (CMIP) phase 5 experiment [Taylor et al., 2012]. CMIP5 provided the climate model simulations that form the basis of the Intergovernmental Panel on Climate Change Fifth Assessment report [Pachauri et al., 2014]. As this study focuses exclusively on atmospheric processes, we use atmospheric model intercomparison project (AMIP) simulations that employ predetermined realistic sea surface temperatures and sea ice. As SWA exhibits notable diurnal cycles of cloud, precipitation and radiation [Stein et al., 2011; van der Linden et al., 2015; Pfeifroth et al., 2016], which will affect their interactions, we focus on those AMIP5 models for which three hourly output is available: CanAM4, CNRM-CM5, HadGEM2-A and MRI-CGCM3. Further details of these models are given in Table 2.

2. Satellite observations of cloud, precipitation and radiation
2.1. Spatial variability

To understand the overall relation between cloud, precipitation and radiation, we first examine their mean regional distribution using MODIS, TRMM and CERES products, as shown in Fig. 1. Regions with extensive cloud cover and heavy precipitation are associated with high outgoing shortwave radiation (OSR) due to substantial reflectance of sunlight by clouds, and low outgoing LW radiation (OLR) due to cold cloud top temperature. Largest cloud cover and precipitation occur to the immediate East and West of the DACCIWA region and are linked to topography, in particular the Guinea highlands to the West and the Jos plateau to the East. Increased cloud cover in these regions is linked to increased OSR and decreased OLR. Between these cloud/precipitation maxima lies the DACCIWA region, highlighted by the black box. This excludes by design significant topography, but may be affected by convective systems initiated over the Jos plateau that travel westward into the region of interest [Fink et al., 2006]. While the DACCIWA region has less extreme values than observed over the topography to the East and West it shows larger precipitation, cloud cover and OSR, and smaller OLR than the Sahel to the immediate North and the Gulf of Guinea to the South.

The DACCIWA region is relatively homogeneous compared to the strong zonal variability to the immediate North and topographically induced variability to the East and West. Mean cloud, radiation and precipitation fields over the DACCIWA region and uncertainty across observational dataset is quantified in the following section in the context of inter-annual variability.

2.2. Inter-annual variability

The inter-annual variability can provide insight into the physical processes determining the climate of the region during June-July as well as contributing context and additional quantification of uncertainty in the planning for the forthcoming DACCIWA observational field campaign. Multi-annual mean and inter-annual variability are quantified in Table 3 for each variable and dataset.

2.2.1. Cloud, precipitation and radiation

Overall, Figure 2 highlights the large differences between the three satellite cloud products. There is little coherent variability or trends in cloud fraction across datasets, although the DARDAR-passive-like dataset displays an increase in high cloud cover from 71% to 83% over the period 2006–2010. Among the three cloud datasets used in the study, DARDAR-passive-like produces the largest total cloud cover, primarily due to much larger high cloud cover (Fig. 2a); this is not surprising because the lidar observations used in DARDAR can detect optically thin clouds at high altitudes that are often missed by MODIS [Sun et al., 2011; Lee et al., 2009]. Although these clouds are optically thin, their radiative effect, particularly in the LW is significant [Haladay and Stephens, 2009; Lee et al., 2009; Sun et al., 2011].

In contrast, low and mid-level cloud covers from DARDAR-passive-like are less than those from MODIS. This may be due to cases such as thin cirrus pixels over lower thicker clouds, which are classified as high cloud by DARDAR-passive-like, but as low clouds by MODIS. To account for differences in mid and low-level cloud due to obscuration by higher clouds, we calculate new weighted mid- and low-level cloud cover (Figs. 2e and 2f). To calculate the weighted mid-level cloud cover, we divide the mid-level cloud cover by the maximum possible mid-level cloud cover, which is less than one due to the presence of high cloud (i.e. we divide the mid-level cloud cover by one minus the high cloud cover). Similarly, the

weighted low-level cloud cover is calculated by dividing the low-level cloud cover by the maximum possible low-level cloud cover, which is less than one due to the presence of high and mid-level cloud. As detailed in the Appendix, if clouds at different levels are randomly overlapped, then these weighted values represent the true mid- and low-level cloud covers. Thus if clouds at different levels are randomly overlapped then the weighted cloud cover for passive satellite observations can be compared directly to the original DARDAR dataset (where we have not removed mid and low-level cloud beneath higher clouds as in 2c and 2d).

To validate the weighted cloud cover, we focus initially on the difference between the weighted DARDAR-passive-like cloud cover and the original DARDAR mean cloud cover. Agreement between the two indicates that the weighting can be applied sensibly to the passive cloud cover datasets (i.e. MODIS and CMSAF) for comparison to the original DARDAR observations. For mid-level cloud the original DARDAR mean cloud cover and weighted DARDAR-passive-like mean cloud cover are very similar (Figure 2e). However, for low-level clouds, the weighted DARDAR-passive-like cloud cover is larger than the original cloud cover (Figure 2f). This indicates that low clouds and mid-level or high clouds are not randomly overlapped in the DARDAR dataset. This may be because they are not randomly overlapped in reality, but could also be due to missing low-level cloud beneath higher cloud in the DARDAR dataset. Indeed, Schrage and Fink [2012] highlighted problems detecting low level stratus beneath higher cloud in SWA, due to attenuation of CALIPSO lidar signal by the higher cloud and contamination in CloudSat radar reflectivity by ground clutter.

Calculating weighted cloud cover for MODIS gives lower mid-level cloud cover than the DARDAR original product (Figure 2e); this confirms that the MODIS mid-level cloud cover is only larger than the DARDAR-passive-like mid-level cloud cover due to larger high cloud cover that obscures mid-level cloud in the DARDAR-passive-like dataset. The difference between DARDAR original and weighted MODIS mid-level cloud cover shows that MODIS is missing some optically thin mid-level clouds. For low clouds, the weighted MODIS cloud cover is larger than the original DARDAR cloud cover. Interpretation of this result is complicated by the lack of agreement between the DARDAR original and weighted DARDAR-passive-like cloud cover and requires a more detailed comparison of DARDAR and MODIS in this region.

Compared to DARDAR-passive-like and MODIS, CMSAF tends to detect less cloud cover at all heights, consistent with results over tropical lands reported in Reuter et al. [2009]. It matches DARDAR-passive-like mid and low level cloud covers better, but comparison of the CMSAF weighted cloud cover with the DARDAR original cloud cover shows this good match is due to increased high cloud in DARDAR-passive-like obscuring more mid-level and low cloud. The MODIS-CMSAF difference is largest for low cloud. This difference is primarily due to a lack of nocturnal low cloud in CMSAF, as discussed in Sect. 3.3.

Figure 3 (a) shows surface precipitation from TRMM, GPM, GPCC and ERA-Interim. For 2014, mean precipitation from GPM and TRMM show agreement to within 0.5 mm day^{-1} due to both being constrained at the monthly mean scale by the same surface rain gauge analysis. Given the lack of extended GPM data at this point in time, TRMM is clearly the better dataset for evaluating the climate models and ERA-Interim. TRMM mean precipitation varies substantially from year to year, ranging from ~4.5 mm day^{-1} in 2013 up to ~7.5 mm day^{-1} in 2002. The longer-term GPCC gauge-analysis product shows that this mean and inter-annual variability is consistent with the record going back to 1900 (not shown). However, figure 3(b) shows that the last few years have been consistently drier than both the TRMM mean (5.8 mm day^{-1}) and the longer-term GPCC mean (6.1 mm day^{-1}). Indeed, averaging over five year periods, the five most recent years of TRMM observations (i.e. 2011-2015) are drier than any other five-year period in either the TRMM or GPCC record, with the caveat that the GPCC full reanalysis currently extends to 2013, so more recent TRMM observations are calibrated using the GPCC monitoring product.

Figure 4 shows inter-annual variability of clear-sky and all-sky TOA outgoing fluxes from CERES-SYN and GERB. For the OSR, there is a notable contrast between the GERB-CERES differences for the GERB2 instrument (2004–2006) and the GERB1 instrument (2007–2011). GERB2-CERES differences are small, while CERES predicts larger monthly mean values than GERB1, the magnitude increasing with time up to about 5 W m^{-2} in 2011, but remains within the sum of the respective observational uncertainties (~6 W m^{-2}). Consequently the mean value for the GERB OSR (Table 3) is slightly smaller than the CERES value. The difference between GERB1 and GERB2 is also evident in the OLR; again GERB2-CERES differences are smaller, while GERB1 appears to overestimate OLR slightly by around 2 W m^{-2}, which is within the observational uncertainty for both CERES and GERB.

It is worth noting that the limited sampling of the diurnal cycle in CERES-SYN has a non-negligible effect on the mean SW fluxes. Sampling the GERB fluxes at similar points in time leads to an increase in the GERB mean outgoing SW flux of approximately 6 W m^{-2} and gives GERB2 values that are larger than CERES and GERB1 values that are a closer match to CERES. Limited sampling of the diurnal cycle can cause similar problems in climate models [Zhou et al., 2015] and will be discussed further in section 3.3.

Generally there are little obvious inter-annual changes in the cloud or radiation observations. Exceptions include the GERB SW flux, where the change from GERB2 to GERB1 leads to differences of approximately 3.5 W m^{-2}. The GERB team has more confidence in the calibration of GERB1 (Clerbaux, pers comm.) so we focus on CERES and GERB1 when considering radiation observations in the rest of this paper. Another notable change is the fall of approximately 10 W m^{-2} in OLR between 2007 and 2008, which happens to be the year used for the evaluation of the diurnal cycles in the climate models and ERA-Interim. This decrease in OLR coincides with a precipitation maximum and high cloud maxima in MODIS and CMSAF. For precipitation, the natural inter-annual variability in the DACCIWA region appears to be large and we must bear in mind that individual years (as used for example in section 4) may not be representative of the longer term.

2.2.2. Budget analysis

To compare inter-annual variability of radiation and precipitation, understand how they are linked, and identify what the sources of variability are, we consider the atmospheric energy and moisture budgets for June-July. Under the hydrostatic approximation and assuming that kinetic energy transport and changes in energy and water storage are negligible as in Muller and O'Gorman [2011], the time-mean regional atmospheric energy budget may be expressed as

$$L_c P = Q + H_d$$

where L_c is the latent heat of condensation of water vapor, P is the surface precipitation rate, H_d is the atmospheric horizontal divergence of dry static energy and Q is the atmospheric net heating from radiation and sensible heat. Similarly, assuming changes in water storage are negligible, the time-mean moisture budget may be expressed as

$$P - E = M_h$$

where E is the surface evaporation rate and M_h is the horizontal divergence of moisture.

Satellite datasets provide estimates of precipitation rates, SW column radiative heating and LW column radiative cooling, while ERA-Interim can provide sensible heating, surface evaporation, divergence of moisture and divergence of dry static energy. Figure 5 illustrates the key terms in the atmospheric energy budget of the DACCIWA region and includes June-July multi-annual mean estimates for each term. The DACCIWA region is fueled by ~110 W m^{-2} of SW radiative heating, ~170 W m^{-2} of latent energy release through precipitation, and ~25 W m^{-2} of sensible heating from the surface. This is balanced by LW net radiative cooling between –188 and –211 W m^{-2} (depending on dataset) and horizontal divergence of dry static energy of about –105 W m^{-2}. While SW and LW radiation are amongst the largest terms in the energy budget, they and the sensible heat flux show little inter-annual variability (values in parenthesis in Figure 5), and are relatively homogeneous throughout SWA (not shown). Latent heating and H_d are much more variable, both temporally and spatially. Year to year changes in latent heating and H_d are well correlated (Figure 6a), with a correlation coefficient of 0.96 when both are estimated from ERA-Interim, and 0.41 when latent heating is based on TRMM precipitation. For both latent heating and H_d these year-to-year changes are generally larger in the eastern part of the DACCIWA region and appear to be linked to changes in the magnitude and location of the precipitation maximum near the coast of Nigeria/Cameroon (see supplementary material).

Within ERA-Interim, the energy budget balances reasonably well, with cooling exceeding heating by 5.7 W m^{-2}. When we combine satellite estimates of latent heating and radiation with ERA-Interim estimates of sensible heating and divergence of dry static energy, which are not available from satellites, heating exceeds cooling by 4.3 W m^{-2} when we use CERES-EBAF for radiation and 13.9 W m^{-2} for CERES-COMPUTED radiation. These differences are primarily due to lower estimates of LW cooling from the satellites. ERA-Interim appears to overestimate LW cooling in the DACCIWA region, which is consistent with an overestimate of OLR and an underestimate of mid-level cloud cover (Table 3). Assuming that the imbalance in the combined satellite and ERA-Interim energy budget is due to errors in ERA-Interim, closing the budget requires either an increase in the divergence of dry static energy or a decrease in sensible heating in ERA-Interim.

The ERA interim moisture budget also balances reasonably well; loss of water through precipitation (5.8 mm day^{-1}) is balanced by evaporation (3.4 mm day^{-1}) and horizontal moisture convergence (2.4 mm day^{-1}). The mean difference between ERA-Interim and TRMM precipitation is small (Figure 3a), so the balance is reasonably well maintained when TRMM precipitation (5.9 mm day^{-1}) is used. Similarly to the energy budget, inter-annual variability is dominated by precipitation and horizontal divergence (Figure 6b). Increased precipitation is linked to increased horizontal convergence of moisture (i.e. negative divergence), with a Pearson correlation coefficient of –0.98 when both are estimated from ERA-Interim and –0.33 for TRMM precipitation.

2.3. Diurnal cycle

Previous studies have indicated notable diurnal cycles of cloud [van der Linden et al., 2015; Stein et al., 2011] and precipitation [Pfeifroth et al., 2016] in SWA. Interactions between clouds, precipitation and radiation are affected by their respective diurnal cycles. For instance, due to the large amplitude of the diurnal cycle of shortwave radiation, cloud-radiation interactions are strongly dependent on when in the diurnal cycle cloud occurs; shortwave cloud heating of the atmosphere tends to be larger than longwave cooling during the day, while at night when there is no shortwave radiation, longwave radiation leads to a cooling of the atmosphere. This section investigates links between the diurnal cycles of clouds, precipitation and radiation.

Figure 7 shows a number of key features of the diurnal cycle of cloud cover, precipitation and outgoing radiation at TOA. To ensure a fair comparison we focus on those years that are included in all datasets, 2007–2010. Note that DARDAR is the only dataset that uses the same method to retrieve daytime and night time cloud cover; therefore, any notable diurnal signature in cloud cover from DARDAR should not be caused by the retrieval method itself. In contrast, MODIS and CMSAF both have separate methods for daytime and night time retrievals.

Figure 7a shows that DARDAR and MODIS do not display a noticeable diurnal cycle in total cloud cover, but there is a clear diurnal difference in high cloud cover (Fig. 7b). High clouds are more prevalent at night (roughly at 0–5 and 20–24 UTC), consistent with reduced OLR (Fig. 7h) and following the diurnal maximum in precipitation (Fig. 7f), suggesting that the diurnal cycle of high cloud cover strongly links to convective systems that develop during the afternoon and evening and lead to increased anvil and cirrus cloud at night (e.g. Stein et al., 2011). In contrast, low clouds are more prevalent during the day; this cycle does not match surface synoptic observations presented by van der Linden et al. [2015], showing that low cloud cover increases overnight to a maximum at 6 UTC and decreases through the day to a minimum at 18 UTC. This discrepancy may be related to the dependence of low cloud cover from passive satellites on high cloud cover, as discussed in section 3.2 and the supplementary material; diurnal changes in MODIS and DARDAR-passive-like high cloud cover result in diurnal changes to the amount of low cloud that is obscured by high cloud, which may lead to unrealistic increases in low cloud cover when high cloud cover decreases.

The diurnal cycle of cloud cover from CMSAF behaves differently to the other datasets. For high- and mid-level clouds, the diurnal cycle from CMSAF is generally consistent with that found in DARDAR and MODIS; however, for the total cloud cover, CMSAF shows a larger cloud cover at daytime than night (Fig. 7a), which is primarily due to low-level cloud cover (as seen in Fig. 7d). This raises a question why CMSAF has similar low-level cloud cover to DARDAR and MODIS during the day, but at night has virtually no low cloud cover, which is much lower than the other two products and warrants more detailed discussions.

Focusing on night time land retrieval algorithm, MODIS relies on the $11 - 3.9$ μm brightness temperature difference (BTD) to detect low clouds, with empirically derived thresholds that depend on total precipitable water predicted by the Global Data Analysis System. Similarly, CMSAF uses the $10.8 - 3.9$ BTD, but with thresholds determined from a look-up table that depend on satellite zenith angle, surface type and NWP forecast integrated water vapor content [Derrien and Le Gleau, 2005]. This suggests that the lack of nocturnal low cloud in the CMSAF product is due to problems with the threshold applied to the $10.8 - 3.9$ BTD. Further evidence for this is provided by Van der Linden et al. [2015] who obtained better low-cloud cover over SWA from SEVIRI by simply fixing the 10.8–3.9 BTD threshold at 2K. Since a fixed threshold is unlikely to be applicable to other scenes and cloud types, this suggests further areas for CMSAF developers for improvement.

Each of the cloud datasets has strengths and weaknesses. The CMSAF product has excellent diurnal sampling, but has large cloud cover biases, particularly for low cloud cover, due to issues detecting nocturnal low cloud. On the other hand, DARDAR has the greatest sensitivity and detects the most cloud, but has fewest samples of the diurnal cycle. In terms of both cloud cover biases and diurnal sampling, MODIS lies between DARDAR and CMSAF. As a compromise between accuracy and diurnal sampling, we use the MODIS dataset for comparison to climate models and ERA-Interim in this study. Satellite simulators that emulate remotely resolved observations [Bodas-Salcedo et al., 2011] are also a useful tool for evaluating clouds in climate models, but the number of instruments and temporal resolution available from AMIP5 models is limited.

As expected, agreement between the diurnal cycle of precipitation in TRMM and GPM is good (Fig 7f, for 2014 only), with a root mean square difference of 0.03 mm hr^{-1}; both show peak precipitation in the early evening and a minimum in the morning, which is characteristic of the diurnal cycle of convective precipitation over land [Nesbitt and Zipser, 2003]. Although precipitation shows large inter-annual variability (Fig. 3), there are no obvious differences between the mean diurnal cycles for 2014 (Fig. 7f) and 2007–2010 (Fig.7e). This suggests that the large inter-annual variability is due to variability in the number and strength of convective events rather than significant variability in the rainfall type. However, this lack of variability in the diurnal cycle over a broad region does not necessarily apply at local scales; Pfeifroth et al. [2016] showed that the diurnal cycle of precipitation at Oueme in Benin showed inter-annual variability that they linked to changes in the proportion of precipitation due to local convection and westward propagating convective systems. Similar to previous studies of the diurnal cycle of convection [Tian, 2004; Chung et al., 2007; Soden, 2000], the diurnal precipitation maximum coincides with an increase in high cloud and precedes the diurnal maximum of high cloud. This may be a consequence of the use of infrared observations in TRMM, which use cloud top temperature as an indicator of surface precipitation. Pfeifroth et al. [2016] found that the TRMM product lagged the surface synoptic diurnal maximum by 1-2 hours for the Ouémé meso-site (centered on 9.5°N, 2°E) and attributed this to the use of infrared observations in TRMM.

Figure 7g shows the diurnal cycle of SW radiation from CERES-SYN and GERB1. For the CERES sample time, differences with GERB values at adjacent times are small. In fact, the mean flux for the GERB1 points closest in time to CERES is 147.1 W m^{-2}, which is slightly larger than the GERB1 diurnal mean (143.3 W m^{-2}) and closer to the 2007–2011 CERES mean (147.1 W m^{-2}). A similar result is obtained when we sample GERB2 at similar times to the CERES-SYN sampling. Differences between CERES-SYN and GERB in the DACCIWA region at the CERES-SYN sample time are consistent with a previous comparison which showed that GERB2 fluxes were generally larger than CERES [Clerbaux et al., 2009]. Diurnal mean differences between the datasets are larger due to different temporal sampling in GERB and CERES-SYN.

The amplitude of the diurnal cycle of OLR (Fig. 7h) is much smaller than the OSR. Consequently, the reduced temporal sampling in CERES has less impact. GERB gives larger values between 20 and 13.5 UTC, so the GERB diurnal cycle has smaller amplitude, but the phases show good agreement. The clear-sky OLR has virtually no diurnal cycle indicating that the all-sky OLR diurnal cycle is due to cloud. The OLR minimum at night (around 1.5 UTC for CERES) and maximum around 13.5 UTC are consistent with the high cloud diurnal cycle shown in Fig. 7b.

Since the flux differences between CERES and GERB are small, in principle, either is suitable for further analysis. GERB has advantage of better temporal resolution, but it lacks clear-sky flux estimates that are necessary to extract the radiative effects of clouds. Therefore, CERES data will be used for the following evaluation of climate models and ERA-Interim. However, bear in mind that the CERES has a potential diurnal average difference of ~5 W m^{-2} due to limitations in temporal sampling.

3. Evaluation of climate model simulations and ERA-Interim

We now undertake a preliminary assessment of CMIP5 simulations and ERA-Interim for the SWA region, focusing on atmosphere-only model AMIP simulations that provide the set of diagnostics necessary to investigate climatological and diurnal variability in radiation, precipitation and cloud cover. The motivation for this, rather than providing a comprehensive evaluation of the set of CMIP5 models, is to update previous analyses of older generation models (e.g. Knippertz et al., 2011) and to highlight major biases that warrant further

investigation. Based upon the analysis in Section 3, the most suitable observational datasets are now exploited for evaluating ERA-Interim and four climate model simulations and identifying deficiencies in their representation of cloud, radiation and precipitation.

With respect to the observed fields shown in Figure 1, Figure 8 shows maps of model mean errors for total cloud cover, precipitation, OSR and OLR respectively. Generally, the climate model errors are fairly homogeneous across the DACCIWA region and consistent with the error over SWA. Most of the spatial variability in errors in the DACCIWA region occurs because the climate models are too homogeneous and fail to capture the limited spatial variability that does exist in the observations; for example, many of the models fail to capture the observed decrease in OLR from West to East, which results in errors that increase from West to East. This is not surprising given the coarse horizontal resolution of the models (Table 2). Of note is the gross underestimation of cloud and precipitation and overestimation of OLR by the ERA Interim reanalysis and the HadGEM2-A climate model simulation at around 10–15°N, immediately to the north of the DACCIWA region. It is somewhat surprising that the ERA Interim has the underestimation of cloud by up to ~50% in coverage and of OLR of up to ~ 50 W m^{-2} given the assimilation of observations, although this bias may be partly a consequence of the paucity of observations in this region. It is also worth noting that cloud is not directly assimilated in ERA-Interim and errors in the model microphysics scheme can cause problems in cloud altitude and radiative properties [Chevallier et al., 2001]. Although generally smaller in magnitude, these cloud and radiation biases also broadly apply across all climate models considered indicating an endemic deficiency in the representation of the WAM. Zonal (10°W – 10°E) mean values (supplementary material) show that in the case of ERA-Interim (and to a lesser extent HadGEM2-A and MRI-CGCM3) these errors are linked to a failure to capture the mean northwards progression of the intertropical convergence zone from June to July and are consistent with the annual mean ERA-Interim precipitation bias identified by Paeth et al. [2011] and the ERA-Interim rainfall onset timing bias identified by Dunning et al. (submitted). This failure to capture the northward shift of the WAM precipitation may be a common feature of reanalyses as it occurs in both ERA-Interim and the MERRA reanalyses [Thorncroft et al., 2011].

Model and ERA-Interim inter-annual variability is generally of a similar magnitude to the observed inter-annual variability (Table 3 and supplementary material), the exception being for clear-sky outgoing SW flux, where the inter-annual variability in both ERA-Interim and the climate models is smaller than CERES-SYN. We suspect this is because the variability is overestimated in the CERES-SYN observations due to sampling errors in this cloudy region.

Figure 9 shows diurnal cycles for ERA-Interim and climate model total cloud cover, precipitation, and top of atmosphere radiative fluxes. For most climate models the diurnal cycle is only available for 2008, so we limit the comparison to this year. Differences in the phase of the diurnal cycles of radiative flux cannot be disentangled from the temporal sampling that the climate models use to reduce the cost of the computationally expensive radiative transfer schemes. The different climate models and ERA-Interim use different methods to account for this reduced temporal sampling, and as a result fluxes output at the same point in time may represent quite different time periods. For example, HadGEM2-A calculates radiative fluxes every three hours using atmospheric properties from the start of the radiation time step and a solar zenith angle that is the mean over the radiation time-step [Manners et al., 2009], outputting fluxes at the start of the radiative time-step. CanAM4 on the other hand applies a 1-hr radiation time-step, uses atmospheric properties and solar zenith angle from the start of the time-step and output fluxes at the end of the time-step. Such differences make comparison of the phase of the diurnal cycles of radiation problematic, so we shall restrict our discussion to the amplitude of the diurnal cycles.

There is a large spread of clear-sky OSR values amongst the models and ERA-Interim (Fig. 9e and Table 3). This is primarily due to surface albedo differences (supplementary material); MRI-CGCM3 has a particularly large albedo and is a clear outlier, but the range of albedos in the other models and ERA-Interim is greater than 0.02, which is large enough to affect regional climate [Taylor et al., 2002]. The AMIP5 models all overestimate clear-sky OSR compared to CERES-SYN and ERA-Interim, which generally offsets the underestimated SW cloud radiative effect (Table 3) caused by underestimating cloud cover. However, the CERES-SYN clear-sky fluxes may be subject to a sampling bias in this very cloudy region. The spread in surface albedo values is likely to have an impact on the simulated cloud and precipitation; many previous studies have shown the sensitivity of West African climate to small changes in albedo [Charney et al., 1977; Taylor et al., 2002].

Only CanAM4 and ERA-Interim simulate levels of low cloud cover within 0.05 of the MODIS values (Table 3). CanAM4 actually has a larger value of low cloud cover than MODIS, but this may simply be a consequence of less higher cloud in CanAM4 to obscure low cloud. The lack of low cloud cover to insulate the surface at night in ERA-Interim and the climate models leads to a larger surface temperature diurnal cycle (not shown). This ultimately leads to an overestimate of the amplitude of the diurnal cycle of clear-sky OLR (Fig. 9b), but this signal is dwarfed by the diurnal cycle of cloud in the all-sky OLR. Note that the amplitude of the diurnal cycle of clear-sky OLR in CanAM4 is a consequence of the unique treatment of solar radiation in the infrared range; solar radiation between 4 and 1000 μm (approximately 12 W m^{-2}) is treated as LW radiation [Li et al., 2010]. This diurnally varying radiation is included in the upwelling LW radiation to ensure that the net LW radiation at the TOA is unaffected.

In the DACCIWA region, the climate models generally underestimate high cloud cover (Table 3). In most cases this does not result in significant OLR biases (Table 3) due to compensating biases in mid-level cloud. However HadGEM2-A overestimates OLR due to a lack of both high and mid-level cloud. The amplitude of the OLR diurnal cycle in HadGEM2-A is much larger than both the observed OLR diurnal cycle and the HadGEM2-A clear-sky OLR diurnal cycle, which suggests that cold cloud cover is overestimated in the evening and underestimated during the rest of the day, particularly between 7.5 and 15.5 UTC.

While the ERA-Interim mean high cloud cover compares well with MODIS, ERA-Interim overestimates OLR due to a lack of mid-level cloud. The previous version of the ECMWF reanalysis overestimated OLR over tropical land due to cloud errors [Allan et al., 2004] and it seems likely that this remains true of ERA-Interim. OLR diurnal cycle errors (Fig. 9d) suggest that ERA-Interim underestimates mid-level cloud throughout the day and particularly at night.

Precipitation diurnal cycle errors in ERA-Interim and the models (Fig 9b) are typical of general circulation models (GCMs) in convective regions [Dai, 2006]; the amplitude of the diurnal cycle is too large and peak precipitation occurs too early. However, at least some of the climate models do a better job of capturing the amplitude of the diurnal cycle of OLR (Fig. 9d), which is also linked to convective cloud [Hartmann and Recker, 1986]. While CanAM4 and HadGEM2-A overestimate the amplitude, CNRM-CM5 and MRI-CGCM3

perform better in reproducing the observed amplitude of the diurnal cycle of OLR of 18.8 W m^{-2}, with amplitudes of 25.7 W m^{-2} and 16.3 W m^{-2} respectively. ERA-Interim also underestimates the amplitude of the diurnal cycle of OLR (relating to the overall underestimation in precipitation shown in Fig. 8).

Each of the AMIP models has large biases between 8°W − 8°E and 5 − 10°N in some (if not all) of the cloud cover, precipitation and radiation fields considered here. ERA-Interim generally does a better job of reproducing the mean observed values in the DACCIWA region, but Figure 8 shows this is at least partly due to a fortuitous cancellation of smaller scale errors. Moreover OLR and diurnal cycles related to convection in ERA-Interim are problematic and over the wider West African region, ERA-Interim biases are as large as the AMIP models, with particularly notable OLR and cloud cover biases over the Sahel.

4. Conclusions

We have constructed a multi-satellite climatology of clouds, top of atmosphere (TOA) radiation, and precipitation for southern West Africa during June-July to support ongoing investigations of the complex interactions within the West Africa Monsoon (WAM) and the evaluation of ERA-Interim and climate model simulations. The primary characteristics of this climatology are as follows. Mean cloud cover in the DACCIWA region ranges from 68 to 94 %, larger than the global average, which is around 68 % [Stubenrauch et al., 2013]. Consequently, the regional mean OSR of ~145 W m^{-2} is also larger than the global average of 100 W m^{-2}, while the regional mean OLR of ~230 W m^{-2} is smaller than the tropical average of, which is estimated to be ~250 W m^{-2} [L'Ecuyer and Stephens, 2003]. Additionally, surface precipitation in the DACCIWA region is ~5.87 mm day^{-1} which is around double the global mean. Diurnal cycles of high cloud cover, precipitation and OLR are typical of convective regions; high cloud cover and precipitation peak in the evening and coincide with the minimum in the OLR diurnal cycle.

Total cloud cover and TOA shortwave and longwave fluxes show little inter-annual variability; for all three the standard deviation is less than 2 % of the mean. However, the inter-annual standard deviation of the regional mean surface precipitation is much larger, around 15 % of the mean, with mean values ranging from ~4.5 mm day^{-1} in 2013 to ~7.5 mm day^{-1} in 2002. This inter-annual variability in precipitation in the DACCIWA region is linked to the regional inter-annual variability of dry static energy. From an atmospheric energy budget perspective, inter-annual changes in latent heat release from precipitation are balanced by dry static energy changes of a similar magnitude.

Comparisons between different satellite products have been used to identify inconsistencies and determine which products are most suitable for further analysis. Differences between the TRMM and GPM precipitation datasets are small, which is unsurprising given the similarity of their algorithms and their use of the same gauge data. The two radiation datasets, GERB and CERES-SYN, also show generally excellent agreement. However, CERES-SYN overestimates the diurnal mean outgoing SW flux due to insufficient temporal sampling. Differences between the satellite cloud observations are much larger. MODIS detects less high (45 % compared to 74 %) and total (89 % compared to 94 %) cloud cover than DARDAR, but more cloud at all heights than CMSAF. All three satellite cloud products have faults. Temporal sampling in MODIS is poor and even more so in DARDAR, while CMSAF and MODIS underestimate cloud cover. We use MODIS for comparison to ERA-Interim and the climate models as a compromise between temporal sampling and accuracy. Clearly a deeper understanding of the interaction between clouds and radiation in this region requires more frequent diurnal sampling than provided by MODIS yet improvements in the geostationary satellite-based CMSAF cloud products are necessary.

Based on our analysis of satellite products, we chose the products that are best suited for comparison to climate models and used these to identify a number of significant biases in the climate models. With one exception (CanAM4), the AMIP5 models considered in this study significantly underestimate the MODIS observed low cloud cover (18 %), with a multi-model mean value of 10 %. This is consistent with large low cloud errors in the previous generation of climate models over southern West Africa identified by Knippertz, et al. [2011] and indeed with a more general climate model bias, where low cloud occurrence in the tropics is systematically underestimated [Nam et al., 2012]. This is likely to have significant impacts on the surface radiation budget [Knippertz et al., 2011] and consequently the broader circulation

[Li et al., 2015]; it is planned to examine this in detail using radiation calculations in a future study. High and mid-level cloud cover errors are smaller, but there are deficiencies in the representation of their diurnal cycle and also that of precipitation, which are linked to convection parametrization in the models [Pearson et al., 2013]. The models have a surprisingly broad range of surface albedos, which leads to a substantial range of clear-sky outgoing SW flux values (62.7 W m^{-2} – 84.4 W m^{-2}) and can mask the effect of cloud errors on the all-sky outgoing SW flux.

Despite assimilating observations, ERA-Interim shares many biases with the free-running climate models. This is thought to be due to a combination of a lack of non-satellite observations to constrain ERA-Interim and the fact that cloud cover, surface precipitation and top of atmosphere radiative fluxes, which are analyzed in this article are not directly assimilated. ERA-Interim generally has smaller domain-mean biases in the DACCIWA region than the AMIP models. However, this is at least partly due to a cancellation of smaller-scale errors (Fig. 8). Moreover, ERA-Interim precipitation diurnal cycle errors are as large as any of the AMIP models, as is the overestimate of OLR, both of which are likely to be due to deficiencies in the convection parametrization used. Over the broader SWA region, ERA-Interim performs no better than the AMIP models, with particularly notable cloud and OLR errors over the Sahel, linked to a failure to capture the northwards progression of the zonal precipitation maximum from June to July.

Current satellite observations provide a wealth of useful data for understanding the climate of southern West Africa. However, many aspects of the climate system in the DACCIWA region remain poorly understood and there are a number of obvious areas that require either improved satellite products or observations from other sources. In particular, there are large differences between cloud cover datasets and those that are thought to perform better lack adequate temporal sampling in a region with marked diurnal variability. It is hoped that future releases of the CMSAF dataset will address some of the errors identified by this study as it provides the highest temporal resolution. Surface radiative flux estimates are required to understand the local atmospheric radiation budget and, consequently, the energy budget. There is a lack of surface-based radiation measurements in the DACCIWA region, but satellite-based surface radiative flux products such as the CERES-EBAF surface fluxes [Kato et al., 2013]and CMSAF-CLARA [Karlsson et al., 2013] provide an opportunity to fill this data gap. As these products rely on radiative transfer calculations and ancillary data with

associated uncertainties, further validation in the DACCIWA region is required. Similarly, satellites struggle to provide robust statistics for aerosol properties in the DACCIWA region due to the rarity of clear-sky. The forthcoming DACCIWA field campaign [Knippertz et al., 2015a], will contribute additional observations that will be used directly for improving understanding of the regional climate system and to further evaluate and suggest improvements to satellite products in this region.

Future work will focus on quantifying cloud radiative effects in SWA for individual cloud types, assessing how different cloud types contribute to the regional energy budget and consequently how they affect the WAM circulation.

Appendix A

One limitation of passive satellite observations is their inability to detect multiple cloud layers, missing low clouds that may exist beneath higher clouds. In contrast, active radar/lidar observations are less prone to this problem. Consequently, it is difficult to compare estimates of low cloud cover from different satellite observations. To minimize the impact of these obscured clouds on our intercomparison, we define a weighted cloud cover that is normalized by non-obscured areas only. For mid-level clouds, the weighted cloud cover, $C_{mid,weighted}$ is calculated by:

$$C_{mid,weighted} = \frac{C_{mid,passive}}{1-C_{high}}, \qquad\qquad (A1)$$

where $C_{mid,passive}$ is mid-level cloud fraction from passive observations, and C_{high} is high-cloud cover. As the nearest clouds to the satellite, high-cloud cover is not affected by clouds in other layers. Similarly, the weighted low cloud cover, $C_{low,weighted}$, is calculated as

$$C_{low,weighted} = \frac{C_{low,passive}}{1-(C_{high}+C_{mid,passive})}, \qquad\qquad (A2)$$

where $C_{low,passive}$ is low-level cloud fraction from passive observations. A schematic illustration of equation (A1) and (A2) can be found in Figure A1.

The difference between the true mid-level cloud cover, $C_{mid,true}$, and the passive mid-level cloud cover is the fraction of mid-level cloud that is obscured, i.e., the amount of overlap between mid- and high-level clouds. Since the amount of overlap would be the product of mid- and high-level cloud covers if they were randomly overlapped, we then obtain:

$$C_{mid,true} - C_{mid,passive} = C_{high}C_{mid,true}. \qquad (A3)$$

By rearranging equation (A3), the true mid-level cloud cover is given by

$$C_{mid,true} = \frac{C_{mid,passive}}{1-C_{high}}, \qquad (A4)$$

which is the same as $C_{mid,weighted}$ as shown in equation (A1), demonstrating that the weighted mid-level cloud cover represents the true mid-cloud cover. Similarly, if low clouds are randomly overlapped with high and mid-level clouds, then the weighted low cloud cover is equal to the true low cloud cover.

	High-level	Mid-level	Low-level
Truth	1/4	3/4	4/4
Passive-like	1/4	2/4	1/4
Weighted from passive observations	1/4	2/(4–1) = 2/3, same as 0.5/(1–0.25) shown in (A1)	1/(4–3) = 1/1, same as 0.25/(1–0.25–0.5) shown in (A2)
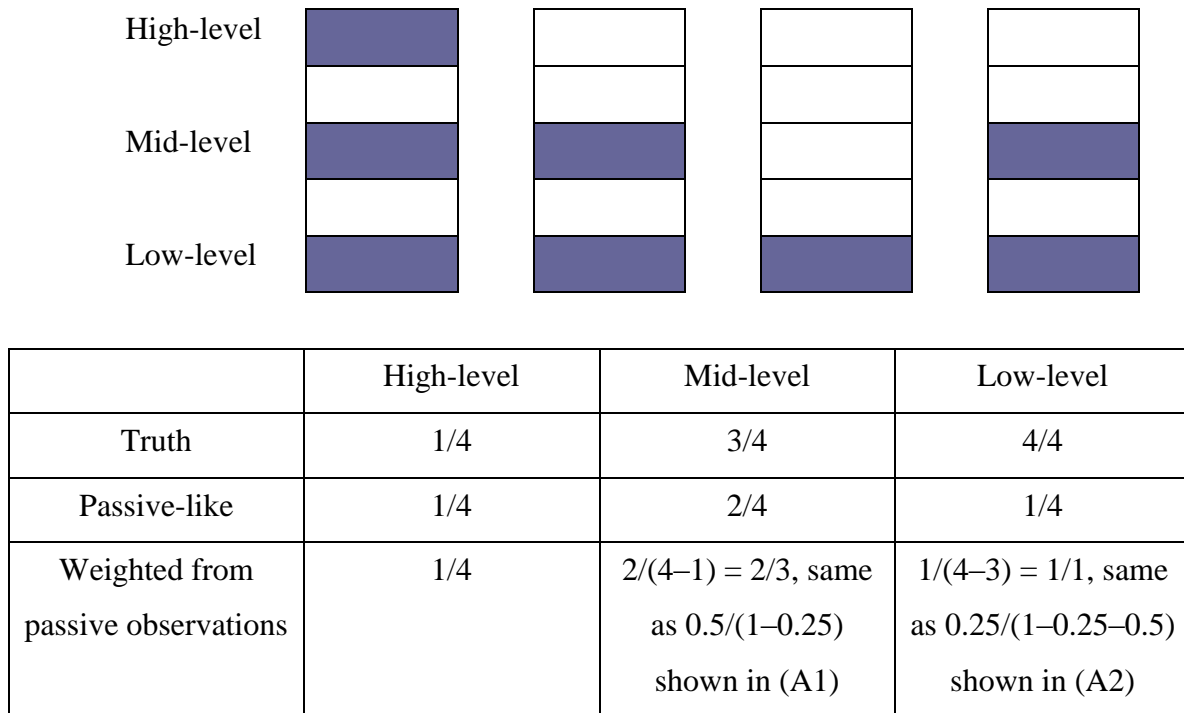
Fig. 1A. A schematic illustration explaining the definition and difference between the truth, passive-like and weighted cloud covers.

Acknowledgments and Data

References

Allan, R. P., Ringer, M. A., Pamment, J. A. & Slingo, A., 2004. Simulation of the Earth's radiation budget by the European Centre for Medium-Range Weather Forecasts 40-year reanalysis (ERA40). *J. Geophys. Res.,* 109[D18107].

Bodas-Salcedo, A. et al., 2011. COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.,* 92[8], pp. 1023-1043.

Charney, J., Quirk, W. J., Chow, S. & Kornfield, J., 1977. A Comparative Study of the Effects of Albedo Change on Drought in Semi-Arid Regions. *J. Atmos. Sci.,* 34[9], pp. 1366-1385.

Chevallier, F. et al., 2001. Model Clouds over Oceans as Seen from Space: Comparison with HIRS/2 and MSU Radiances. *Journal of Climate,* 14[21], pp. 4216-4229.

Chung, E. S., Sohn, B. J., Schmetz, J. & Koenig, M., 2007. Diurnal variation of upper tropospheric humidity and its relations to convective activities over tropical Africa. *Atmospheric Chemistry and Physics,* 7[10], pp. 2489-2502.

Clerbaux, N., 2003. Outgoing longwave flux estimation: improvement of angular modelling using spectral information. *Remote Sensing of Environment,* 85[3], pp. 389-395.

Clerbaux, N. et al., 2009. Comparison of GERB instantaneous radiance and flux products with CERES Edition-2 data. *Remote Sensing of Environment,* 113[1], pp. 102-114.

Cook, K. H. & Vizy, E. K., 2006. Coupled Model Simulations of the West African Monsoon System: Twentieth- and Twenty-First-Century Simulations. *J. Climate,* 19[15], pp. 3681-3703.

Dai, A., 2006. Precipitation Characteristics in Eighteen Coupled Climate Models. *Journal of Climate,* 19[18], pp. 4605-4630.

Dee, D. P. et al., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. Roy. Meteorol. Soc.,* 137[656], pp. 553-597.

Delanoe, J. & Hogan, R. J., 2010. Combined CloudSat-CALIPSO-MODIS retrievals. *J. Geophys. Res.,* 115[D00H29].

Derrien, M. & Le Gleau, H., 2005. MSG/SEVIRI cloud mask and type from SAFNWC. *International Journal of Remote Sensing,* 26[21], pp. 4707-4732.

Dewitte, S. et al., 2008. The Geostationary Earth Radiation Budget Edition 1 data processing algorithms. *Advances in Space Research,* 41[11], pp. 1906-1913.

Doelling, D. R. et al., 2013. Geostationary Enhanced Temporal Interpolation for CERES Flux Products. *J. Atmos. Oceanic Technol.,* 30[6], pp. 1072-1090.

Ebert, E. E., Janowiak, J. E. & Kidd, C., 2007. Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *Bull. Amer. Meteor. Soc.,* 88[1], pp. 47-64.

Fink, A. H., Vincent, D. G. & Ermert, V., 2006. Rainfall Types in the West African Sudanian Zone during the Summer Monsoon 2002. *Mon. Wea. Rev.,* 134[8], pp. 2143-2164.

Frey, R. A. et al., 2008. Cloud Detection with MODIS. Part I: Improvements in the MODIS Cloud Mask for Collection 5. *J. Atmos. Oceanic Technol.,* 25[7], pp. 1057-1072.

Gaetani, M., Pohl, B., Douville, H. & Fontaine, B., 2011. West African Monsoon influence on the summer Euro-Atlantic circulation. *Geophysical Research Letters,* 38[9].

Geleyn, J. F. & Hollingsworth, A., 1979. An econmical analytical method for the computation of the interaction between scattering and line absorption of radiation. *Contrib. Atmos Physics,* Volume 52, pp. 1-16.

Gu, G. & Adler, R. F., 2004. Seasonal Evolution and Variability Associated with the West African Monsoon System. *Journal of Climate,* 17[17], pp. 3364-3377.

Haladay, T. & Stephens, G., 2009. Characteristics of tropical thin cirrus clouds deduced from joint CloudSat and CALIPSO observations. *J. Geophys. Res.,* Volume 114.

Harries, J. E. et al., 2005. The Geostationary Earth Radiation Budget Project. *Bull. Amer. Meteor. Soc.,* 86[7], pp. 945-960.

Hartmann, D. L. & Recker, E. E., 1986. Diurnal Variation of Outgoing Longwave Radiation in the Tropics. *J. Climate Appl. Meteor.,* 25[6], pp. 800-812.

Hou, A. Y. et al., 2014. The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.,* 95[5], pp. 701-722.

Hourdin, F. et al., 2010. AMMA-Model Intercomparison Project. *Bulletin of the American Meteorological Society,* 91[1], pp. 95-104.

Huffman, G. J. et al., 2013. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). In: *Algorithm theoretical basis document, version 4.1.* s.l.:NASA.

Huffman, G. J. et al., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeor,* 8[1], pp. 38-55.

Jianping, L. & Qingcun, Z., 2003. A new monsoon index and the geographical distribution of the global monsoons. *Adv. Atmos. Sci.,* 20[2], pp. 299-302.

Karlsson, K.-G.et al., 2013. CLARA-A1: a cloud, albedo, and radiation dataset from 28 yr of global AVHRR data. *Atmospheric Chemistry and Physics,* 13[10], pp. 5351-5367.

Kato, S., 2003. Computation of domain-averaged shortwave irradiance by a one-dimensional algorithm incorporating correlations between optical thickness and direct incident radiation. *J. Atm. Sci.,* 60[1], pp. 182-193.

Kato, S. et al., 2013. Surface Irradiances Consistent with CERES-Derived Top-of-Atmosphere Shortwave and Longwave Irradiances. *J. Climate,* 26[8], pp. 2719-2749.

Knippertz, P. et al., 2015a. The DACCIWA project: Dynamics-aerosol-chemistry-cloud interactions in West Africa. *Bull. Amer. Meteor. Soc..*

Knippertz, P. et al., 2015b. The possible role of local air pollution in climate change in West Africa. *Nature Climate Change,* 5[9], pp. 815-822.

Knippertz, P. et al., 2011. Ultra-low clouds over the southern West African monsoon region. *Geophys. Res. Lett.,* 38[21].

Le Barbé, L., Lebel, T. & Tapsoba, D., 2002. Rainfall Variability in West Africa during the Years 1950-90. *Journal of Climate,* 15[2], pp. 187-202.

L'Ecuyer, T. S. & Stephens, G. L., 2003. The Tropical Oceanic Energy Budget from the {TRMM} Perspective. Part I: Algorithm and Uncertainties. *Journal of Climate,* 16[12], pp. 1967-1985.

Lee, J. et al., 2009. Distribution and Radiative Forcing of Tropical Thin Cirrus Clouds. *J. Atmos. Sci.,* 66[12], pp. 3721-3731.

Li, J., Curry, C. L., Sun, Z. & Zhang, F., 2010. Overlap of Solar and Infrared Spectra and the Shortwave Radiative Effect of Methane. *J. Atmos. Sci.,* 67[7], pp. 2372-2389.

Li, R., Jin, J., Wang, S.-Y. & Gillies, R. R., 2015. Significant impacts of radiation physics in the Weather Research and Forecasting model on the precipitation and dynamics of the West African Monsoon. *Clim Dyn,* 44[5-6], pp. 1583-1594.

Liu, C. et al., 2015. Combining satellite observations and reanalysis energy transports to estimate global net surface energy fluxes 1985-2012. *Journal of Geophysical Research: Atmospheres,* 120[18], pp. 9374-9389.

Liu, Z., 2016. Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results. *J. Hydrometeor,* 17[3], pp. 777-790.

Loeb, N. G. et al., 2003. Angular Distribution Models for Top-of-Atmosphere Radiative Flux Estimation from the Clouds and the Earth's Radiant Energy System Instrument on the Tropical Rainfall Measuring Mission Satellite. Part I: Methodology. *Journal of Applied Meteorology,* 42[2], pp. 240-265.

Loeb, N. G. et al., 2009. Toward Optimal Closure of the Earth's Top-of-Atmosphere Radiation Budget. *J. Climate,* 22[3], pp. 748-766.

Manners, J. et al., 2009. Two fast radiative transfer methods to improve the temporal sampling of clouds in numerical weather prediction and climate models. *Q. J. Roy. Meteorol. Soc.,* Volume 135, pp. 457-468.

Marsham, J. H. et al., 2013. The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations. *Geophys. Res. Lett.,* 40[9], pp. 1843-1849.

Minnis, P. et al., 2011. CERES Edition-2 Cloud Property Retrievals Using TRMM VIRS and Terra and Aqua MODIS Data; Part I: Algorithms. *IEEE Trans. Geosci. Remote Sensing,* 49[11], pp. 4374-4400.

Muller, C. J. & O'Gorman, P. A., 2011. An energetic perspective on the regional response of precipitation to climate change. *Nature Climate Change,* 1[5], pp. 266-271.

Mulmenstadt, J., Sourdeval, O., Delanoe, J. & Quaas, J., 2015. Frequency of occurrence of rain from liquid-, mixed-, and ice-phase clouds derived from A-Train satellite retrievals. *Geophys. Res. Lett.,* 42[15], pp. 6502-6509.

Nam, C., Bony, S., Dufresne, J.-L. & Chepfer, H., 2012. The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophys. Res. Lett.,* 39[21].

Nelson, A., 2004. *Population Density for Africa in 2000.* fourth ed. s.l.:UNEP/GRID Sioux Falls.

Nesbitt, S. W. & Zipser, E. J., 2003. The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements. *Journal of Climate,* 16[10], pp. 1456-1475.

Nicholson, S. E. et al., 2003. Validation of TRMM and Other Rainfall Estimates with a High-Density Gauge Dataset for West Africa. Part I: Validation of GPCC Rainfall Product and Pre-TRMM Satellite and Blended Products. *Journal of Applied Meteorology,* 42[10], pp. 1337-1354.

Omotosho, J., 1988. Spatial variation of rainfall in Nigeria during the 'little dry season'. *Atmospheric Research,* 22[2], pp. 137-147.

Pachauri, R. K. et al., 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Paeth, H. et al., 2011. Progress in regional downscaling of west African precipitation. *Atmosph. Sci. Lett.,* 12[1], pp. 75-82.

Pearson, K. J. et al., 2013. Modelling the diurnal cycle of tropical convection across the 'grey zone'. *Quarterly Journal of the Royal Meteorological Society,* 140[679], pp. 491-499.

Pfeifroth, U., Trentmann, J., Fink, A. H. & Ahrens, B., 2016. Evaluating satellite-based diurnal cycles of precipitation in the African tropics. *J. Appl. Meteor. Climatol.,* 55[1], pp. 23-39.

Platnick, S. e. a., 2015. MODIS Atmosphere L3 Daily Product.. *NASA MODIS Adaptive Processing System, Goddard Space Flight Center, USA.*

Platnick, S. et al., 2003. The MODIS cloud products: algorithms and examples from terra. *IEEE Trans. Geosci. Remote Sensing,* 41[2], pp. 459-473.

Priestley, K. et al., 2002. Early radiometric validation results of the CERES Flight Model 1 and 2 instruments onboard NASA's Terra Spacecraft. *Advances in Space Rearch,* 30[11], pp. 2371-2376.

Reuter, M. et al., 2009. The CM-SAF and FUB Cloud Detection Schemes for SEVIRI: Validation with Synoptic Data and Initial Comparison with MODIS and CALIPSO. *Journal of Applied Meteorology and Climatology,* 48[2], pp. 301-316.

Roca, R. et al., 2010. Comparing Satellite and Surface Rainfall Products over West Africa at Meteorologically Relevant Scales during the AMMA Campaign Using Error Estimates. *J. Appl. Meteor. Climatol.,* 49[4], pp. 715-731.

Roehrig, R. et al., 2013. The Present and Future of the West African Monsoon: A Process-Oriented Assessment of CMIP5 Simulations along the AMMA Transect. *J. Climate,* 26[17], pp. 6471-6505.

Rutan, D. A. et al., 2015. CERES Synoptic Product: Methodology and Validation of Surface Radiant Flux. *J. Atmos. Oceanic Technol.,* 32[6], pp. 1121-1143.

Schneider, U. et al., 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology,* 115[1-2], pp. 15-40.

Schrage, J. M., Augustyn, S. & Fink, A. H., 2006. Nocturnal stratiform cloudiness during the West African monsoon. *Meteorol. Atmos. Phys.,* 95[1-2], pp. 73-86.

Schrage, J. M. & Fink, A. H., 2012. Nocturnal Continental Low-Level Stratus over Tropical West Africa: Observations and Possible Mechanisms Controlling Its Onset. *Mon. Wea. Rev.,* 140[6], pp. 1794-1809.

Soden, B. J., 2000. The diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere. *Geophys. Res. Lett.,* 27[15], pp. 2173-2176.

Sorooshian, S. et al., 2011. Advanced Concepts on Remote Sensing of Precipitation at Multiple Scales. *Bull. Amer. Meteor. Soc.,* 92[10], pp. 1353-1357.

Stein, T. H. M. et al., 2011. The vertical cloud structure of the West African monsoon: A 4 year climatology using CloudSat and CALIPSO. *J. Geophys. Res.,* 116[D22].

Stein, T. H. M. et al., 2015. The representation of the West African monsoon vertical cloud structure in the Met Office Unified Model: an evaluation with CloudSat. *Q.J.R. Meteorol. Soc..*

Stengel, M. S. et al., 2014. CLAAS: the CM SAF cloud property data set using SEVIRI. *Atmos. Chem. Phys.,* 14[8], pp. 4297-4311.

Stephens, G. L. et al., 2002. The CloudSat mission and the A-Train. *Bull. Amer. Meteor. Soc.,* 83[12], pp. 1771-1790.

Stubenrauch, C. J. et al., 2013. Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the {GEWEX} Radiation Panel. *Bull. Amer. Meteor. Soc.,* 94[7], pp. 1031-1049.

Sultan, B. et al., 2005a. Agricultural impacts of large-scale variability of the West African monsoon. *Agricultural and Forest Meteorology,* 128[1-2], pp. 93-110.

Sultan, B., Janicot, S. & Diedhiou, A., 2003. The West African Monsoon Dynamics. Part I: Documentation of Intraseasonal Variability. *Journal of Climate,* 16[21], pp. 3389-3406.

Sultan, B., Labadi, K., Guegan, J.-F. & Janicot, S., 2005b. Climate Drives the Meningitis Epidemics Onset in West Africa. *PLoS Medicine,* 2[1].

Sun, W. et al., 2011. A study of subvisual clouds and their radiation effect with a synergy of CERES, MODIS, CALIPSO, and AIRS data. *Journal of Geophysical Research: Atmospheres,* 116[D22].

Sylla, M. B., Dell'Aquila, A., Ruti, P. M. & Giorgi, F., 2009. Simulation of the intraseasonal and the interannual variability of rainfall over West Africa with {RegCM}3 during the monsoon period. *International Journal of Climatology,* pp. n/a-n/a.

Taylor, C. M. et al., 2002. The Influence of Land Use Change on Climate in the Sahel. *Journal of Climate,* 15[24], pp. 3615-3629.

Taylor, K. E., Stouffer, R. J. & Meehl, G. A., 2012. An Overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.,* 93[4], pp. 485-498.

Thorncroft, C. D., Nguyen, H., Zhang, C. & Peyrille, P., 2011. Annual cycle of the West African monsoon: regional circulations and associated water vapour transport. *Quarterly Journal of the Royal Meteorological Society,* 137[654], pp. 129-147.

Tian, B., 2004. Diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere: Satellites versus a general circulation model. *J. Geophys. Res.,* 109[D10].

Tian, Y. et al., 2009. Component analysis of errors in satellite-based precipitation estimates. *J. Geophys. Res.,* 114[D24].

van der Linden, R., Fink, A. H. & Redl, R., 2015. Satellite-based climatology of low-level continental clouds in southern West Africa during the summer monsoon season. *Journal of Geophysical Research: Atmospheres,* 120[3], pp. 1186-1201.

Wielicki, B. A. et al., 1996. Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment. *Bull. Amer. Meteor. Soc.,* 77[5], pp. 853-868.

Winker, D. M. et al., 2009. Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *J. Atmos. Oceanic Technol.,* 26[11], pp. 2310-2323.

Worqlul, A. W. et al., 2014. Comparison of rainfall estimations by TRMM 3B42, MPEG and CFSR with ground-observed data for the Lake Tana basin in Ethiopia. *Hydrology and Earth System Sciences,* 18[12], pp. 4871-4881.

Zhou, L., Zhang, M., Bao, Q. & Liu, Y., 2015. On the incident solar radiation in CMIP5 models. *Geophysical Research Letters,* 42[6], pp. 1930-1935.

Table 1: Key features of observational datasets used in this study.

Source	Version	Period	Orbit	Resolution	References
MODIS MODerate resolution Imaging Spectroradiometer	*Atmosphere L3 Monthly Global Product; collection 6*	Dec 2001-present (Terra) June 2002 – present (Aqua)	Polar x2	Gridded to 1.0°	Platnick [2015]
DARDAR raDAR-liDAR	DARDAR mask	2006 – 2010	Polar	1.1 km horizontal 60m vertical	Delanoe & Hogan [2010]
CM-SAF The Satellite Application Facility on Climate Monitoring	CLoud property dAtAset using SEVIRI (CLAAS) version 1	Jan 2004 – Dec 2011	Geostationary	Approx. 4 km	Stengel et al. [2014] http://dx.doi.org/10.5676/EUM_SAF_CM/CLAAS/V001
GPCC	Full Data Reanalysis Version 7.0 at 1.0°	Jan 1901 – Dec 2013	N/A (gauge based)	Gridded to 1.0°	Schneider et al. [2014] http://dx.doi.org/10.5676/DWD_GPCC/FD_M_V7_100
TRMM Tropical Rainfall Measuring Mission	3B42 V7	Jan 1998 - present	N/A (merged product)	Gridded to 0.25°	Huffman et al. [2007]
GPM Global Precipitation Measurement	IMERG 3B-HHR	March 2014 – present	N/A (merged product)	0.1°	Hou et al. [2014] Huffman et al. [2013]
CERES Clouds and the Earth's Radiant Energy System	SYN1deg-3Hour Ed3A	Dec 2001-present (Terra) June 2002 – present (Aqua)	Polar x2	1.0°	Wielicki et al. [1996] Doelling et al. [2013]
GERB Geostationary Earth Radiation Budget	High Resolution Image Version 6	April 2004 - present	Geostationary	9x9 km at nadir	Harries et al. [2005] Dewitte et al. [2008]

Table 2: Details of climate models used in this study.

MODEL NAME	COUNTRY	HORIZONTAL RESOLUTION	VERTICAL RESOLUTION
CANAM4	Canada	$2.8^o \times 2.8^o$	L35
CNRM-CM5	France	$1.4^o \times 1.4^o$	L31
HadGEM2-A	UK	$1.25^o \times 1.875^o$	L38
MRI-CGCM3	Japan	$1.125^o \times 1.125^o$	L48

Table 3: DACCIWA region (8W-8E, 5-10N) mean and inter-annual variability (standard deviation of annual means, given in parenthesis) for June-July. For each dataset we use all available years between 2000 and 2015 inclusive.

| Variable	Observations			Models				Reanalysis
Cloud	*DARDAR*	*MODIS*	*CMSAF*	*CanAM4*	*CNRM-CM5*	*HadGEM2-A*	*MRI-CGCM3*	*ERA-Interim*
Total cover	0.94 (0.02)	0.89 (0.01)	0.68 (0.02)	0.82 (0.03)	0.71 (0.04)	0.54 (0.04)	0.72 (0.02)	0.74 (0.03)
High cover	0.74 (0.05)	0.45 (0.04)	0.36 (0.03)	0.31 (0.06)	0.44 (0.03)	0.29 (0.05)	0.36 (0.04)	0.45 (0.04)
Mid-level cover	0.14 (0.03)	0.21 (0.03)	0.13 (0.01)	0.28 (0.02)	0.18 (0.03)	0.16 (0.02)	0.28 (0.03)	0.10 (0.01)
Low cover	0.06 (0.01)	0.18 (0.02)	0.05 (0.01)	0.21 (0.02)	0.04 (0.01)	0.09 (0.02)	0.07 (0.01)	0.16 (0.03)
Precip. (mm day^{-1})	*TRMM*	*GPM*		*CanAM4*	*CNRM-CM5*	*HadGEM2-A*	*MRI-CGCM3*	*ERA-Interim*
	5.87 (0.89)	5.00		6.45 (0.97)	5.72 (0.74)	4.65 (0.69)	5.16 (0.59)	5.80 (0.89)
Radiation (W m^{-2})	*CERES*	*GERB*		*CanAM4*	*CNRM-CM5*	*HadGEM2-A*	*MRI-CGCM3*	*ERA-Interim*
Outgoing SW	147.6 (3.2)	144.7 (4.0)		154.6 (4.4)	137.4 (7.0)	127.7 (6.4)	163.8 (4.5)	149.4 (5.3)
Clear-sky outgoing SW	58.47 (1.2)	-		66.7 (0.5)	69.2 (0.1)	70.5 (0.6)	84.4 (0.1)	62.7 (0.1)
TOA SW CRE	88.6	-		87.9	68.2	57.1	79.4	86.7
OLR	229.8 (3.5)	230.4 (2.8)		231.1 (6.5)	234.1 (3.2)	246.6 (5.2)	234.7 (3.4)	244.4 (3.0)
Clear-sky OLR	273.0 (0.8)	-		272.4 (1.0)	272.6 (0.9)	276.6 (2.1)	272.9 (0.7)	271.4 (1.0)
TOA LW CRE	-43.2	-		-41.2	-38.5	-30.0	-38.2	-27.0
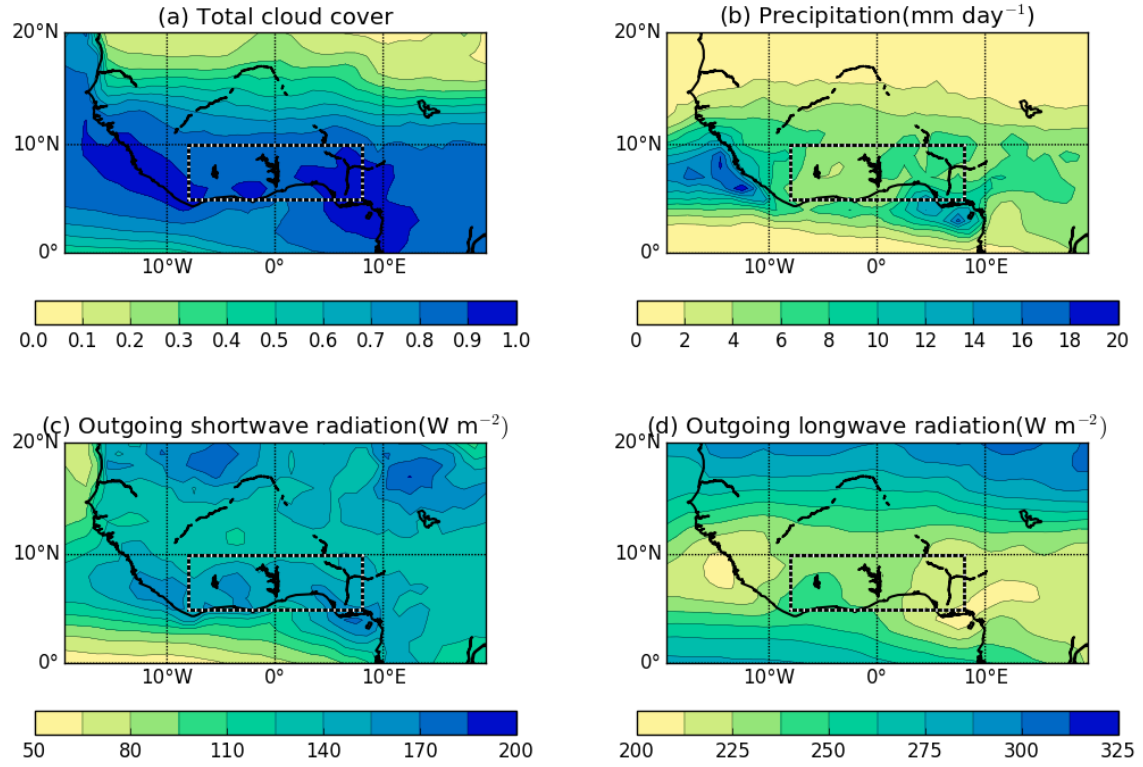
Figure 1. June-July mean properties of cloud, precipitation and radiation over year 2002–2008 in southern West Africa, including (a) total cloud cover from MODIS, (b) precipitation from TRMM, (c) outgoing shortwave radiation from CERES-SYN, and (d) longwave radiation from CERES-SYN. The dotted indicates the DACCIWA region that is used in Fig 2 onwards.
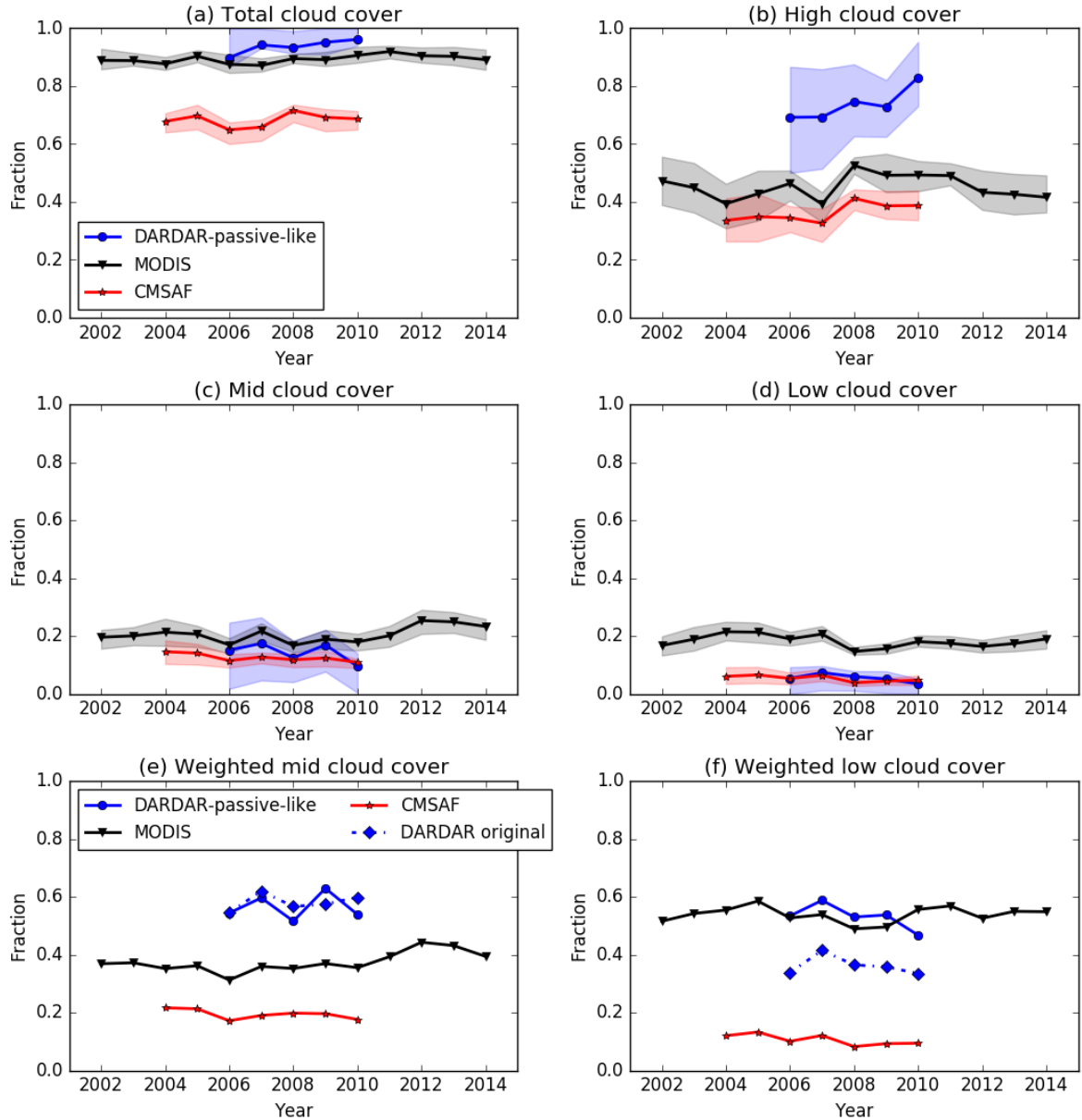
Figure 2. June-July mean cloud cover from various satellite products. (a)–(d) shows total, high-, mid-, and low-level cloud cover. Shaded regions represent spatial variability, indicated by the interquartile range of cloud cover over the DACCIWA region. Weighted cloud covers in (e) and (f) are calculated by dividing by the fraction that is not already obscured by higher cloud. The DARDAR original cloud cover is calculated directly from the DARDAR cloud mask and includes mid- and low-level cloud that is detected by DARDAR beneath higher cloud.
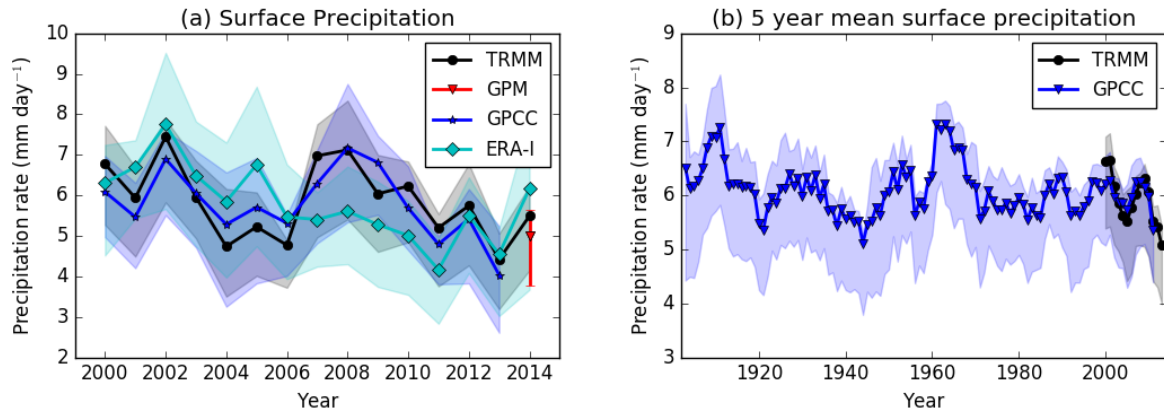
Figure 3. June-July mean precipitation (mm day^{-1}). (a) shows June-July means for individual years as observed by TRMM, GPM and GPCC and modelled in ERA-Interim. (b) shows five year running means for June-July for TRMM and GPCC. The shaded regions (error bar for GPM) shows spatial variability, indicated by the interquartile range over the DACCIWA region, based on 1° resolution.
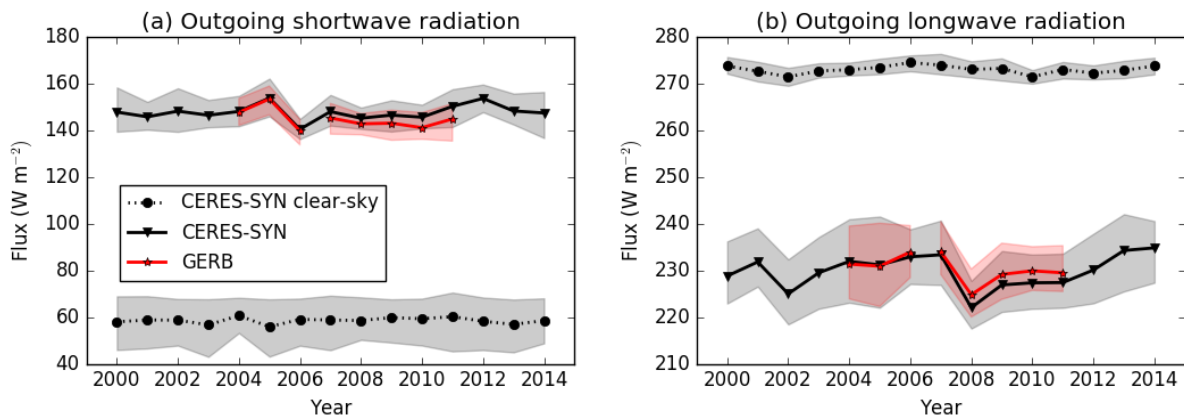

Figure 4. Mean top of atmosphere outgoing shortwave (a) and longwave (b) radiative fluxes for June-July from CERES-SYN and GERB. Shaded region shows spatial variability, in indicated by the interquartile range, over the DACCIWA region based on 1° resolution. The break in the GERB lines shows where GERB1 replaces GERB2 in 2007.
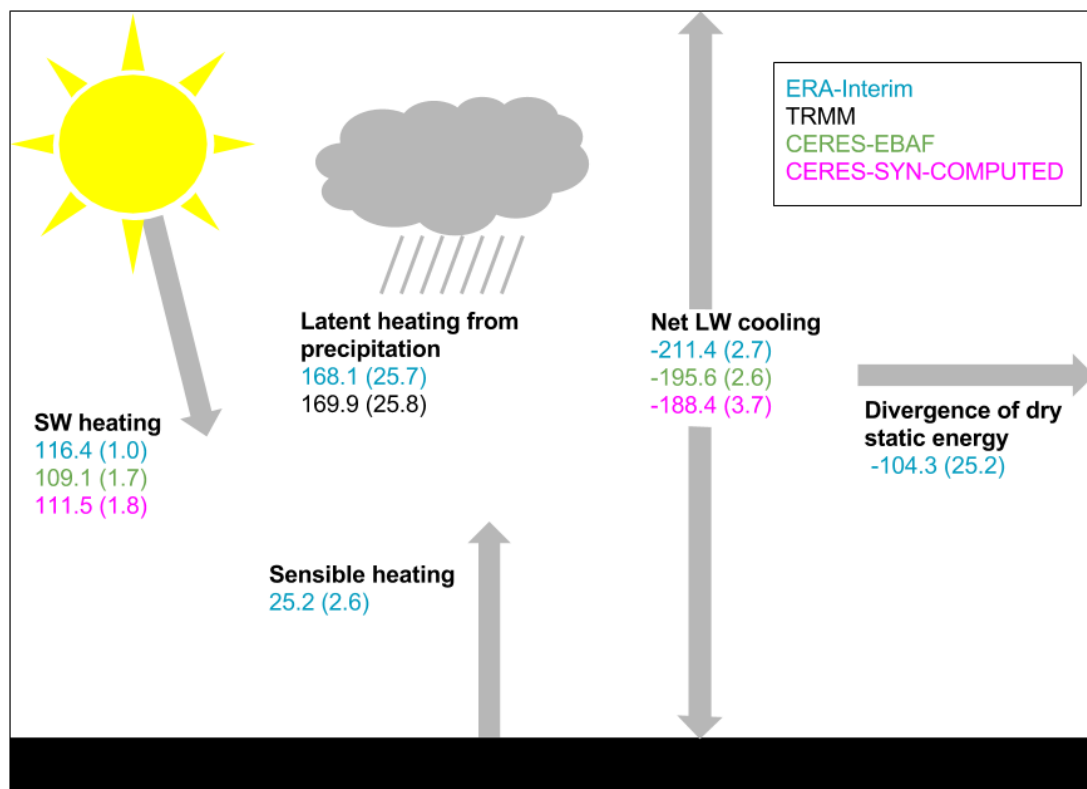
Figure 5: Key terms in the atmospheric energy budget for the DACCIWA region. Numbers shown are June-July means (standard deviations of inter-annual means in parenthesis) for 2000-2015 from the data source indicated by that colour.
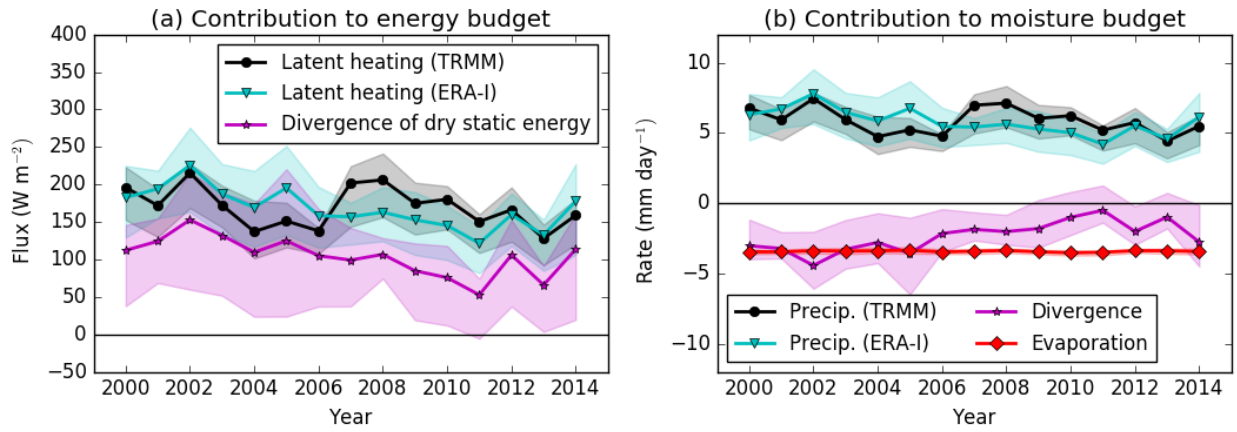
Figure 6: Mean contributions to the energy budget (a) and moisture budget (b) for the DACCIWA region for June-July. Shaded regions show spatial variability, indicated by the interquartile range, based on 1° resolution. To improve readability of the figure (a) only includes the terms that show largest inter-annual variability, namely latent heating and divergence of dry static energy.
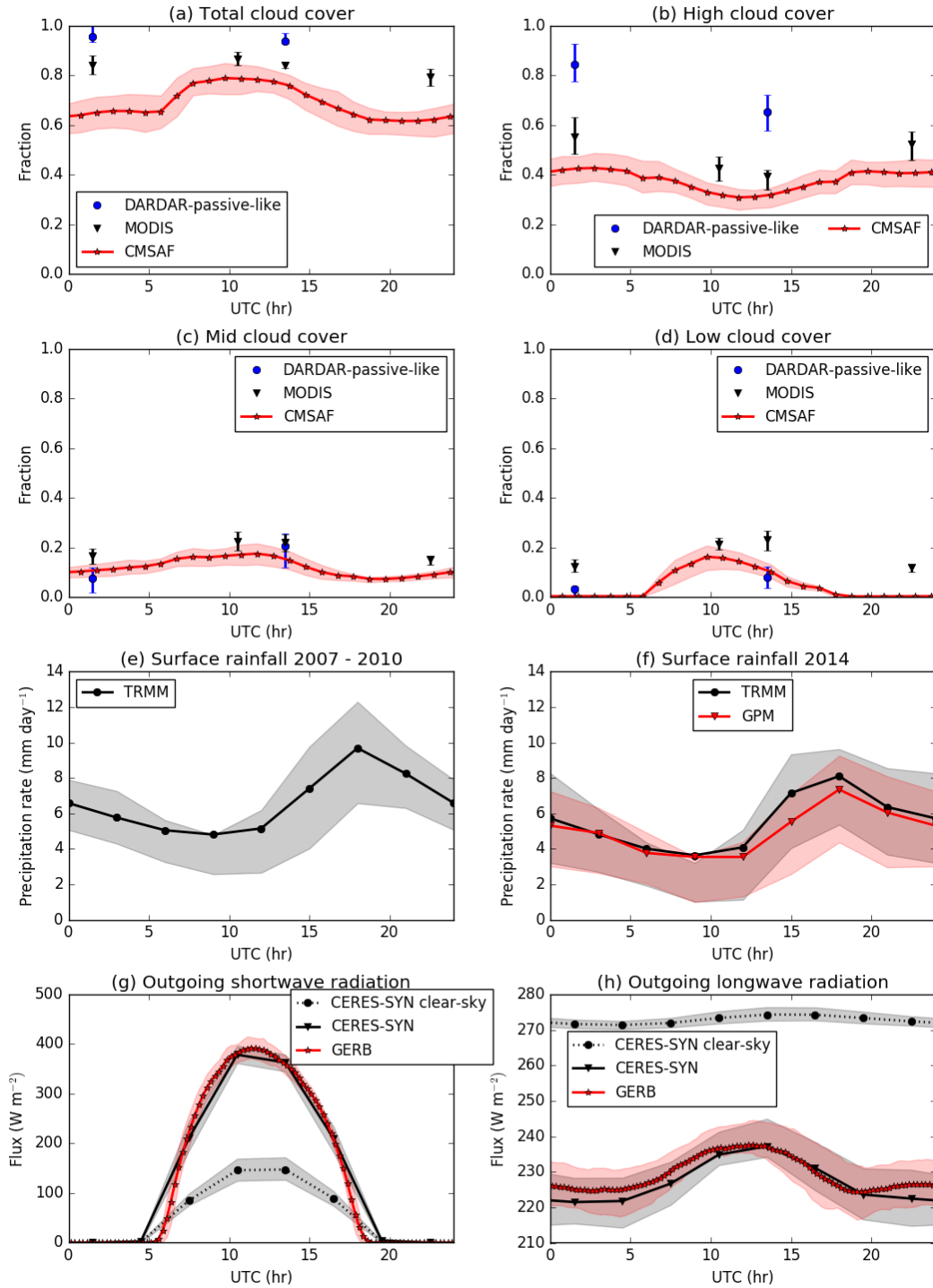
Figure 7. Diurnal cycle of cloud cover, precipitation and radiative fluxes for June-July. All panels show means for 2007-2010, except for (f) which shows 2014. Shaded regions and error bars show spatial variability, indicated by the interquartile range, based on 1° resolution.
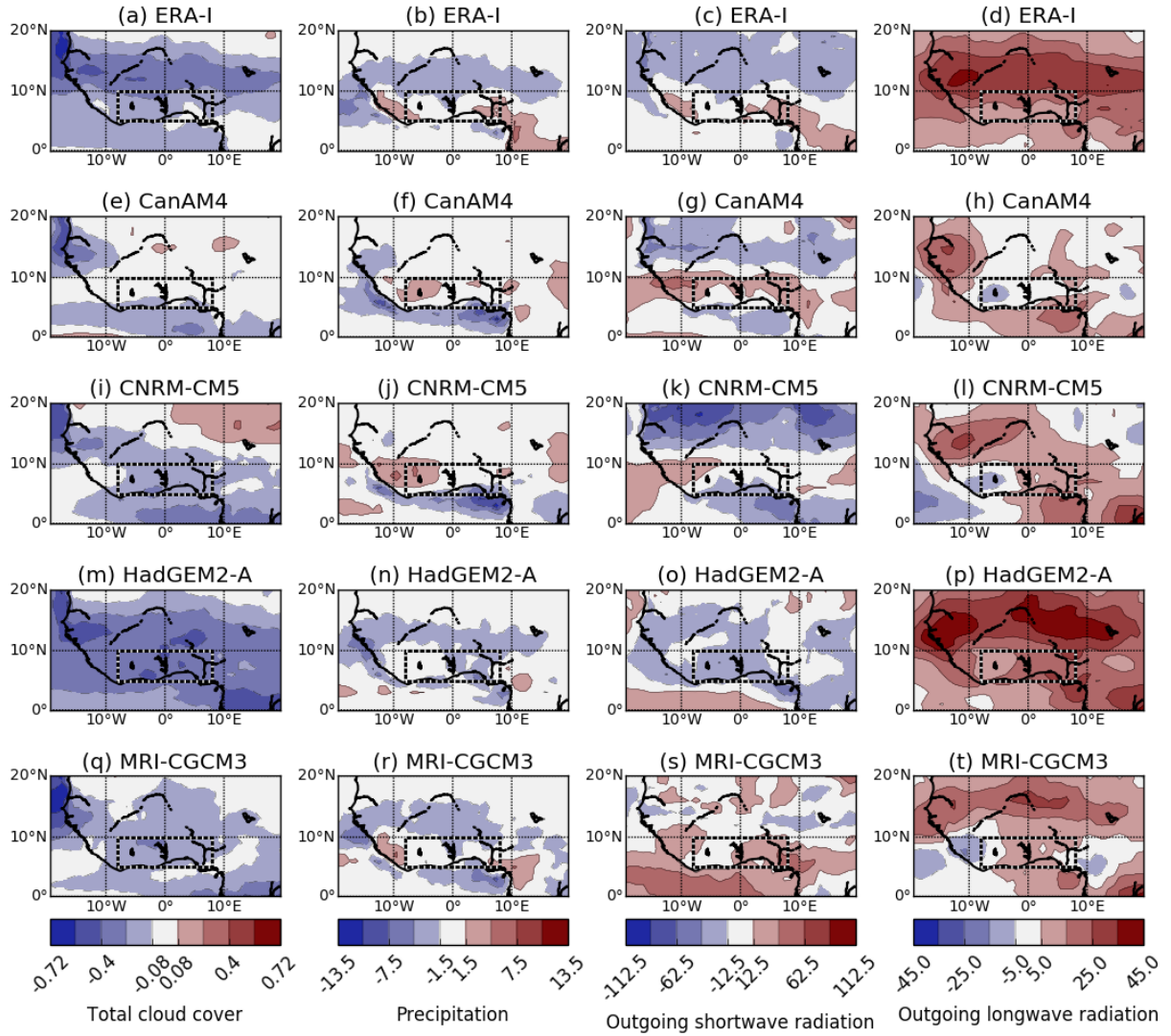
Figure 8. Maps of ERA-Interim and climate model errors for June-July 2002–2009. Panels from left to right show errors in cloud cover, precipitation and outgoing shortwave and longwave radiation, calculated with respect to MODIS, TRMM and CERES-SYN respectively. Information on various climate models can be found in Table 2.
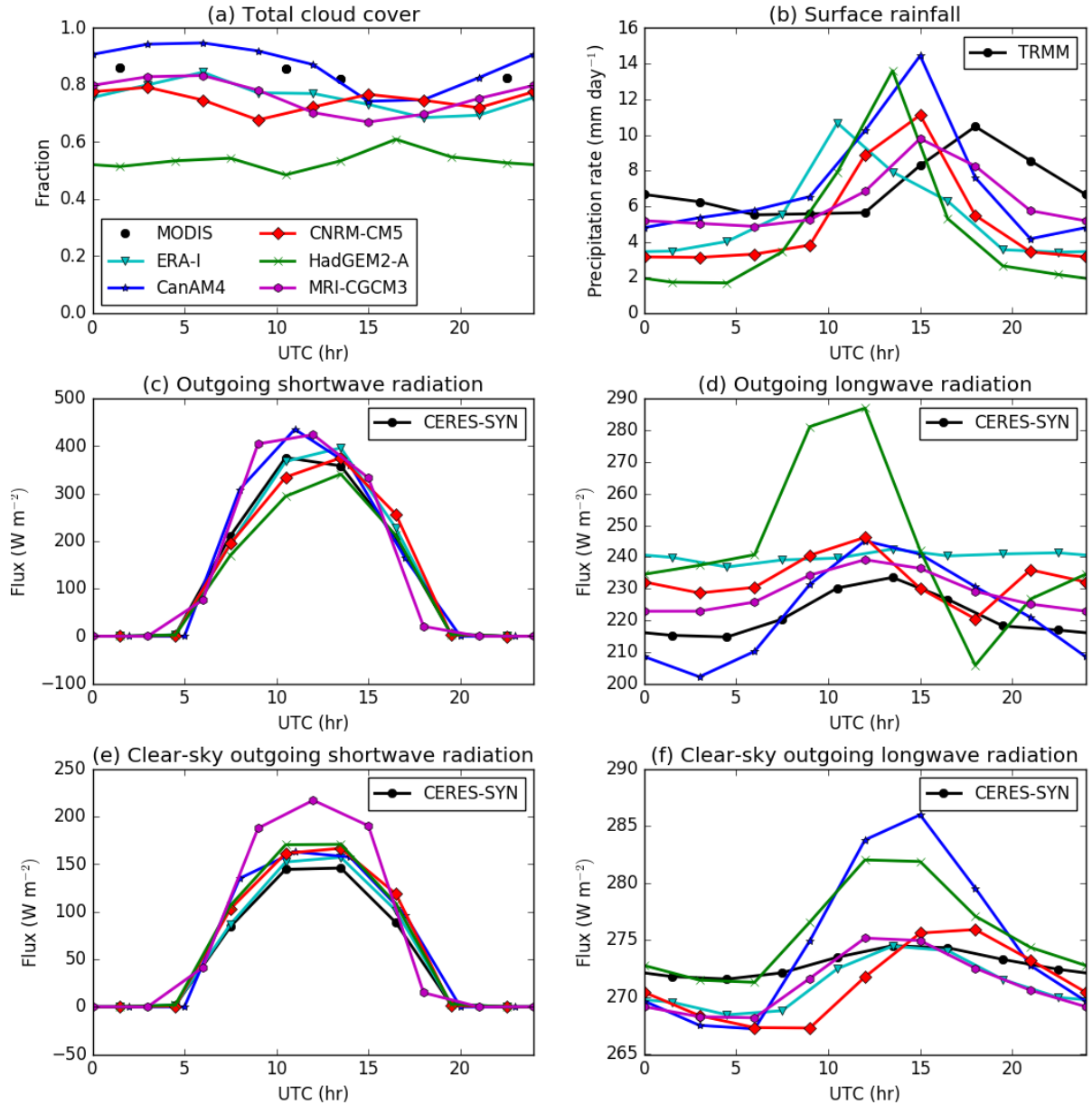
Figure 9: Comparison of diurnal cycles for June-July 2008. In each case, the black line shows the observed diurnal cycle (from the dataset indicated in the legend) and the coloured lined show the four AMIP models and ERA-Interim. For shortwave radiation (c and e), the AMIP models are shown at a time that is most consistent with the solar zenith angle used, rather than the output time