

Project No: 603502

DACCIWA

"Dynamics-aerosol-chemistry-cloud interactions in West Africa"

Deliverable

D7.4 Model Evaluation

Due date of deliverable: 31/08/2017

Completion date of deliverable: 01/09/2017

Start date of DACCIWA project: 1st December 2013 **Project duration:** 60 months

Version: [V1.0]

File name: [D7.4_Model_Evaluation_DACCIWA_v1.0.pdf]

Work Package Number: 7

Task Number: 4

Responsible partner for deliverable: KIT

Contributing partners: KIT, ECMWF, ETHZ

Project coordinator name: Prof. Dr. Peter Knippertz

Project coordinator organisation name: Karlsruher Institut für Technologie

Dissemination level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Nature of Deliverable		
R	Report	
P	Prototype	
D	Demonstrator	
O	Other	

Copyright

This Document has been created within the FP7 project DACCIWA. The utilization and release of this document is subject to the conditions of the contract within the 7th EU Framework Programme. Project reference is FP7-ENV-2013-603502.

DOCUMENT INFO

Authors

Author	Beneficiary Short Name	E-Mail
Anke Kniffka	KIT	anke.kniffka@kit.edu
Tanja Stanelle	ETH	tanja.stanelle@env.ethz.ch
Peter Knippertz	KIT	P.Knippertz@leeds.ac.uk
Angela Benedetti	ECMWF	Angela.Benedetti@ecmwf.int
Malcolm Brooks	UK MO	malcolm.e.brooks@metoffice.gov.uk
Peter Hill	Uni Reading	p.g.hill@reading.ac.uk

Changes with respect to the DoW

Issue	Comments
Originally envisaged analysis of output from the UK consortium Cascade is already slightly outdated by now.	We analysed YoTC data instead that are multi-model, cover longer time-periods and provide diabatic tendencies

Dissemination and uptake

Target group addressed	Project internal / external
Public	Internal and external

Document Control

Document version #	Date	Changes Made/Comments
0.1	13.07.2017	Template
0.2	14.08.2017	First version for approval by the General Assembly
1.0	01.09.2017	Final version approved by the General Assembly

Table of Contents

1	Introduction	5
2	Evaluation strategy	5
2.1	Choice of models and data sources	6
2.2	Planning of the evaluation procedure	6
2.3	Data processing script and further analysis	7
3	Description of participating models	8
3.1	ECHAM (ETHZ)	8
3.1.1	The climate model ECHAM6	8
3.1.2	The aerosol climate model ECHAM6-HAM2	8
3.1.3	Adaptation for DACCIWA	9
3.2	IFS seasonal runs (ECMWF)	9
3.2.1	The Coupled Ensemble Prediction System	9
3.3	UNIFIED MODEL (MET OFFICE)	10
3.4	CMIP5 (as analysed by KIT)	11
3.5	YoTC (as analysed by KIT)	12
4	Evaluation data sets	13
4.1	Radiosonde data	13
4.2	Ground-based observations	13
4.2.1	GPCP	13
4.3	Satellite data	14
4.3.1	GERB	14
4.3.2	SEVIRI	14
4.3.3	ATOVS	14
5	Results	15
5.1	Climatology: vertical structure of the atmosphere	15
5.2	Cloud properties and radiation	21
5.3	Average climate at DACCIWA stations	26
5.4	Moisture and precipitation	32
6	Conclusions	34
7	References	37
8	Appendix	42

1 Introduction

In this deliverable, we present results from an evaluation of the representation of the West African monsoon (WAM) system in widely used global and regional climate and seasonal model datasets following a similar strategy as in D7.3 for weather forecasting. Considered model datasets include (a) long-term climate simulations from the Coupled Model Intercomparison Project Phase 5 (CMIP5) coordinated by the Intergovernmental Panel on Climate Change (IPCC), (b) climate simulations conducted as part of the Years of Tropical Convection (YoTC; Waliser et al. 2012), which provides additional tendency terms and diurnally resolved data, (c) own climate simulations using the ECHAM6 (run by ETHZ) and Unified Model (run by MO), and (d) seasonal forecasts using the ECMWF Integrated Forecast System (IFS). As using the YoTC data was not anticipated in the proposal, we decided to drop the originally envisaged analysis of output from the UK consortium Cascade, which is already slightly outdated by now. The output from the various models was compared to satellite and ground-based observations and re-analysis data. The analysis of the ECHAM6 and Unified Model control runs will be a preparation of the scenario experiments to be conducted as part of Task 7.5. The analysis presented here complements work in WP5 concentrating specifically on the radiation budget (Hill et al., 2016). Some first results focusing on low clouds in YoTC and CMIP5 data have already been published (Hannak et al., 2017), updating earlier results based on CMIP3 by Knippertz et al. (2011).

This report is structured in the following way: Section 2 details the evaluation strategy containing information about chosen parameters, and spatial and temporal averages. Section 3 contains information about the employed models and datasets, while section 4 describes the observations for evaluation. Results of the multi-model evaluation are presented in section 5 followed by conclusions and an outlook in section 6.

2 Evaluation strategy

For the present model evaluation, we adapted in large parts the evaluation strategy from D7.3, but with an important difference. For D7.3 simulations were carried out specifically for the campaign period. In contrast to this, here we analysed climatological simulations that were already existing or were performed in the scope of DACCIWA but not necessarily specifically for this deliverable due to the immense processing times of climatological simulations. Therefore, the simulation periods are not all the same but vary about a few years around the target simulation period from 2006–2015. For each dataset the respective period is denoted in Table 1.

Table 1: Time periods that were available for this study from models and measurements. Unless otherwise indicated, the months June–September were processed.

Source (model or measurement)	Time frame
ECHAM	1996 – 2009
IFS	2003 – 2015
UM	1989 – 2008
YoTC	1991 – 2010
CMIP5	1979 – 2008
GERB	2006 – 2015
GPCP	2006 – 2015
SEVIRI	2006 – 2015
ATOVS	2006 – 2012
Radiosondes	2016 (June & July only)

2.1 Choice of models and data sources

As in D7.3, we firstly identified all models suitable for the evaluation procedure. Our selected models are either used for research within DACCIWA (e.g. ECHAM and IFS seasonal aerosol runs) or stem from coordinated climate simulation experiments (CMIP5 and YoTC). The latter has the advantage of having a large number of models running with the same settings, therefore the CMIP5 and YoTC experiments serve as benchmark simulations against which the DACCIWA models are compared. Figure 1 shows a list of all models considered at the planning stage.

Model evaluation D7.4 (global & regional models, concentrate on 5–10°N, 8°W–8°E, JJAS separately)									
Available models and time frames									
type	model	delta x	delta t output	Time span	Tendencies	Int. aerosol	SST	status	Contact
Own climate and seasonal runs									
research	ECHAM (version 6.3)	T63 (200km)	6h (monthly means)	1996–2015 (20 yr)	T,q	N	AMIP	done	T. Stanelle
research	ECHAM-HAM (6.3, HAM 2.3)	T63 (200km)	6h (monthly means)	1996–2015 (20 yr)	T,q	Y, rad & clou	AMIP	ongoing	T. Stanelle
research	MO UM			20 yr				done	M. Brooks
seasonal	ECMWF IFS	T255 (80 km)	12h (plus monthly means)	2003-2015	X	Y	Coupled ocean	done	A. Benedetti
Already existing external climate runs									
CMIP5	multi (12)	variable	daily - monthly	1979–2008 (30yr)	X	N	AMIP	available	A. Kniffka
YoTC	multi (8)	variable	6h	1991–2010 (20 yr)	T,q (u,v)	N	AMIP	available	A. Kniffka

Figure 1: List of models initially considered for this evaluation exercise.

Together with researchers from WP5 (Radiative Processes) we discussed the advantages and disadvantages of possible observational data for comparison such as radiosondes or satellite sensors in the context of suitability for southern West Africa.

Figure 2 provides a list of different data sources considered (first column) and the meteorological variables they could provide (following columns). After some testing, not all considered products were used for the evaluation. A drawback for many satellite sensors for example is the persisting and thick high cloud cover that obscures clouds close to the surface and handicap the derivation of solar surface radiation. This has to be taken into account for the discussion. We chose data sources that were (a) particularly suitable for the evaluation and (b) had a sufficiently long time-series so that climatological representative averages could be constructed with sufficient quality.

model evaluation		precip	cloud fraction	LWP/IWP	CTT	TPW	surface sol.	TOA rad	LWC/IWC
JJAS 2006–2015	TRMM 3B42	(1,3)(A/B,C/D)	x	x	x	x	x	x	x
	GPCC/P	(A,B)					x		x
	CLAAS	x	(1,3)(A/B,C/D)	(1,3)(A/B,C/D)	(1,3)(A/B,C/x	x	x	x	x
	NASA-Langley-Geo		?	?	?		?	x	x
	GERB	x	x	x	x	x	x	(1,3)(A/B,C/x	x
	MODIS (Terra/Aqua)		(1,3)(C/D)	(1,3)(C/D)		(1,3)(C/D)		x	x
	GEOPROF	x	(1,3)(C/D)	1,3(C/D)	1,3(C/D)	x	x	x	(1,2,4,5)(C/
	ATOVS	x	x	x	x	(1,3)(A/B,C/x	x	x	x
	Rad. Stations	x	x	x	x	x	(1,3)(A/B,C/x	x	x

Figure 2: Sources of possible evaluation data considered during initial discussions (first column) and parameters of interest (following columns). The numbers and letters in the individual cells stand for temporal and spatial averages as detailed in the Appendix. Not all data listed here were used in the end, sources marked with a "?" were excluded due to unavailability and some turned out not to be suitable for the evaluation.

2.2 Planning of the evaluation procedure

The approach to compare multi-scale models over a common region was developed and applied successfully in the forecast evaluation (D7.3), therefore it is used here once more but slightly adapted to fulfil the requirements of a climate model evaluation. Again, we distributed the workload amongst all project partners in order to reach the maximum number of comparable models and parameters. In general, one person was responsible for all parameters from one model and/or one observational source. This also depended on the already existing archived model or satellite data. We developed a processing guideline, which all partners used. In this guideline, common metrics, horizontal grid

resolution, vertical levels, and other parameters were prescribed as well as a naming convention. Together with the guideline, an evaluation software package was developed to ensure easy comparability.

The partners applied the software to reformat the data sets so that they all have the same structure and follow a specific naming convention, which makes automatic comparison possible. All the data sets that were prepared in this manner were then delivered to KIT for the actual evaluation. This approach was found to be suitable to compare multi-scale models over a common region.

2.3 Data processing script and further analysis

As mentioned in Section 2.2., to process data from the different sources in an efficient and consistent way, a script was developed and shared between the involved partners. This shell script uses the Climate Data Operators that can process different kinds of model data given in netcdf format. The script performs all necessary preparatory tasks such as re-gridding, geographical truncation, averaging in time and space etc. The processed datasets possess the following properties:

Area: 8°W–8°E, 5–10°N

Grid: 1.0° x 1.0°

Pressure levels: 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 925, 950, 975, 1000

Runs: Time-series had to be at least ten years long, preferably from 2006 – 2015, 6 hourly output, for the months June, July, August and September

Temporal averages: monthly mean and monthly mean diurnal cycle plus seasonal mean (June, July, August and September) and seasonal mean diurnal cycle including standard deviations

Spatial averages: DACCIWA region as box average (purple box in Fig. 1) and individual stations (Fig. 1)

Naming convention: Detailed information on the automatic filename assignment can be found in the appendix of D7.3.



Figure 3: Geographical overview of DACCIWA focus region (taken from recently published overview paper of Flamant et al., 2017). The purple box shows the area used for spatial averaging (8°E–8°W, 5–10°N). Ground-sites, aircraft base, radiosonde stations and other geographical features of interest are marked.

After the pre-processing described, all data were delivered to KIT, where they were further analysed using python scripts as described in Section 5. It needs to be mentioned that no specific runs were carried out for this evaluation. Hence, not all variables were available from each model. In case some models are missing in the figures in Section 5, this is due to the unavailability of variables from the respective models unless otherwise noted.

3 Description of participating models

3.1 ECHAM (ETHZ)

3.1.1 The climate model ECHAM6

ECHAM6 is the sixth generation of the atmospheric general circulation model ECHAM. It is described in detail in Stevens et al. (2013). ECHAM6 employs a spectral transform dynamical core and a flux-form semi-Lagrangian tracer transport algorithm from Lin and Rood (1996). Vertical mixing occurs through turbulent mixing, moist convection (including shallow, deep, and mid-level convection), and momentum transport by gravity waves arising from boundary effects or atmospheric disturbances. Sub-grid scale cloudiness (stratiform clouds) is represented using the scheme of Sundqvist (1989) which calculates diagnostically the grid cell cloud fraction as a function of the relative humidity in the given grid cell, once a threshold value is exceeded. Liquid (cloud water) and solid (ice water) condensates are treated prognostically following Lohmann and Roeckner (1996). Radiative transfer in ECHAM6 is represented using the radiation transfer broadband model (RRTMG), which considers 16 and 14 bands for the shortwave (820 to 50000 cm^{-1}) and longwave (10 to 3000 cm^{-1}) parts of the spectrum, respectively (Iacono et al., 2008).

Radiative transfer is computed based on compounds present in the atmosphere and their related optical properties. Trace gas concentrations of long-lived greenhouse gases are specified in the model (except for water vapour). Optical properties of clouds are calculated for each band of the RRTMG scheme using Mie theory and the concentration of liquid water and ice condensates as computed by the 1-moment scheme.

3.1.2 The aerosol climate model ECHAM6-HAM2

The ECHAM6-HAM2 model is a global aerosol climate model. The aerosol module HAM was first implemented in the 5th generation of ECHAM (ECHAM5; Roeckner et al., 2003) by the Max Planck Institute for Meteorology (Stier et al., 2005). Over the past years, the HAM module has been improved and completed with new processes (HAM2) as described in Zhang et al. (2012). The HAM2 module is now coupled to the ECHAM6.

Aerosol microphysics is simulated using the M7 module (Vignati et al., 2004), which accounts for sulphate, black carbon, particulate organic matter, sea salt, and dust. The atmospheric aerosol population is described as a superposition of seven lognormal distributed modes for which standard deviations are prescribed. The total number concentration and masses of the different chemical components are prognostic variables in the model. The modes are divided into soluble, internally mixed modes (containing sulphate) and insoluble, externally mixed modes, which are assigned to different size ranges. The modal diameters can vary and are calculated at each time step from the mass and number concentrations for each mode. Dust particles are considered as part of the soluble and insoluble accumulation and coarse modes. Sedimentation and dry and wet deposition are parameterized as functions of the aerosol size distribution, composition, and mixing state and depend on the ECHAM6 meteorology. The emission fluxes of dust, sea salt, and dimethyl sulfide from the oceans (DMS) are calculated online, based on the model meteorology. Anthropogenic emissions are prescribed.

The optical properties of the modelled aerosol concentrations are employed in the radiative transfer calculations. The aerosol activation and ice nucleation parameterizations of the two-moment stratiform cloud scheme of Lohmann et al. (2007) provide links between the simulated aerosol population and the number concentrations of cloud droplet and ice crystal. So the model accounts for aerosol effects on cloud microphysics (droplet number and size) and macrophysics (liquid water path).

3.1.3 Adaptation for DACCIWA

After a brief evaluation of first test simulations, we decided to use the ECHAM6.3, respectively ECHAM6.3-HAM2.3, model version in the framework of the DACCIWA project. Anthropogenic and biomass burning emissions are prescribed by (1) the ACCMIP (Atmospheric Chemistry and Climate Model Intercomparison Project) emission inventory (Lamarque et al., 2010), (2) ACCMIP + GFAS (Global Fire Assimilation System; Heil et al., 2010) for biomass burning and (3) the HTAPv2.0 emission inventory + GFAS.

All model simulations were performed with a T63L47 resolution (approx. 200 km x 200 km horizontal resolution). The simulation period is from January 1996 until December 2009. The output frequency of variables of interest was 6 hourly, presented are monthly averages for 00 UTC, 06 UTC, 12 UTC, and 18 UTC. For the purpose of the following comparison the output is interpolated to 1° x 1° horizontal resolution and to 21 vertical pressure levels. The following simulations were conducted (Table 2).

Table 2: Available model runs using the ECHAM system with and without sophisticated aerosol treatment.

Model	Anthropogenic aerosol	Biomass burning aerosol
e1: ECHAM6.3-HAM2.3	ACCMIP emission inventory	ACCMIP
e2: ECHAM6.3-HAM2.3	ACCMIP	GFAS (* 3.4)
e3: ECHAM6.3-HAM2.3	HTAP2.0	GFAS (* 3.4)
ea: ECHAM6.3		

3.2 IFS seasonal runs (ECMWF)

3.2.1 The Coupled Ensemble Prediction System

Monthly forecasts have been produced routinely at ECMWF since March 2002, and operationally since October 2004 (Vitart, 2014). In the current configuration, the monthly forecasts are generated by extending the 15-day ensemble integrations to 32 days twice a week (at 00 UTC on Mondays and Thursdays). Forecasts are based on the medium range/monthly ensemble forecast (ENS) which is part of ECMWF's Integrated Forecasting System (IFS). ENS includes 51 members run with a horizontal resolution of Tco639 (about 16 km) up to forecast day 15, and Tco319 (about 32 km) thereafter. The atmospheric model is coupled to an ocean model (NEMO) with a 1 degree horizontal resolution. Initial perturbations are generated using a combination of singular vectors and perturbations generated using the ECMWF ensemble of data assimilations, and model uncertainties are simulated using two stochastic schemes. The climatology (re-forecasts) used to calibrate the real-time forecasts is computed using the re-forecast suite that includes only 5 members of 32-day integrations with the same configuration as the real-time forecasts, starting on the same day and month as the real-time forecast over the past 20 years.

3.2.2 The IFS aerosol model

An atmospheric composition prediction system based on the IFS meteorological model has been developed by ECMWF under various EU-funded initiatives. Currently the Copernicus Atmosphere Monitoring Service (CAMS) is responsible for the maintenance and development of the atmospheric composition model. The version used in this work corresponds to cycle 41R1 of the IFS for which a detailed description can be found at <https://software.ecmwf.int/wiki/display/IFS/CY41R1+Official+IFS+Documentation>. The IFS version used with coupled chemistry is referred to as Composition IFS (C-IFS). The only difference between IFS and C-IFS is the set-up and the resolution. Generally IFS is not used with the full coupled chemistry, which is too expensive to be run at the resolution of the operational NWP model. When IFS is run with the coupled

chemistry, the resolution is lower, generally around 80km with 60 vertical level up to 0.1 hPa as opposed to the operational high resolution of 9km and 137 vertical levels up to 0.01 hPa. Aerosols are forecasted within the global system by a bulk/bin scheme (Morcrette et al., 2009, based on earlier work by Reddy et al., 2005 and Boucher et al., 2002) that includes five species: dust, sea salt, black carbon, organic carbon and sulfates.

Dust aerosols are represented by three prognostic variables that correspond to three size bins, with bin limits of 0.03, 0.55, 0.9 and 20 μm in radius. Sea-salt aerosols are also represented by three size bins with limits of at 0.03, 0.5, 5 and 20 microns in radius. For all other tropospheric aerosols, emission sources are defined according to established inventories (Lamarque et al, 2010). Fire emissions are prescribed from the GFAS (Kaiser et al, 2012). Removal processes include, sedimentation of all particles, wet and dry deposition and in-cloud and below cloud scavenging. For organic matter and black carbon, two components, hydrophobic and hydrophilic, are considered. A very simplified representation of the sulphur cycle is also included with only two variables, sulphur dioxide (SO_2) and sulphate (SO_4), this latter one in the particulate phase. Overall, a total of 12 new prognostic variables for the mass mixing ratio of the different components (bins or types) of the various aerosols are used. Several revisions of the dust emission schemes have been undertaken as well as developments to include more aerosol species, such as nitrates and secondary organic aerosols. These will be included in the next release of the C-IFS (Rémy, private communication).

In the operational version of the global CAMS system, the radiative impact of aerosols is taken into account using the aerosol monthly climatology of Tegen et al. (1997). In an experimental version of the model, the aerosol direct effect can be computed from the mass mixing ratio of the prognostic aerosols provided by the aerosol module. We make use of this capability to set-up experiments with ENS to investigate the importance of the direct radiative impact of the prognostic aerosols relative to a run that uses the Tegen climatology.

3.2.3 Set-up of the experiments

Two experiments were run to assess the aerosol impacts: one control run in which all settings are similar to the operational run including the use of the Tegen climatology for the radiative aerosol effects, but run at lower horizontal resolution (T255 corresponding to 80km), and an interactive aerosol run with the same resolution in which the prognostic aerosols are activated. Both runs were conducted with 91 vertical levels. Several technical modifications were necessary in order to include the aerosols in the ENS suite. This included several script changes to activate the aerosol settings, such as the fetching of the (prescribed) emissions over the years of interest (2003 to 2015) as well as any other changes required to run the interactive aerosols. For computational cost the size of the ensemble was limited to 10 members with start date May 1. Both runs were set-up to be 7 months long, although only statistics for the months of June and July are presented here in the context of the DACCIWA model evaluation exercise. The aerosol was initialized using the MACC reanalysis up to 2012 and the NRT MACC/CAMS experiments were used for the remaining years. The meteorological variables were initialized using ERA-Interim, whereas the initial conditions for soil variables were taken from a separate experiment.

3.3 UNIFIED MODEL (MET OFFICE)

The Unified Model (UM) addressed here is the Met Office Unified Model with its scientific configuration according the Global Atmosphere (GA) model at version 7.0. The UM is a general circulation model that covers a wide range of spatial and temporal scales and can be applied for numerical weather prediction or seasonal forecasting as well as climate predictions. It can be operated either in global or in regional mode. When used in a global mode, the scientific, as opposed

to technical, configuration is defined and developed according to the GA model. The GA7.0 configuration used in these simulations is the most recently defined GA version, and is intended for the Met Office's contribution to the CMIP6 experiments (Eyring et al., 2016).

The GA7 configuration includes further developments to the GA6.0/6.1 as described in Walters et al. (2017) which has been used as the operational global numerical weather prediction model since July 2014. Here the complete dynamical core underwent a refreshment and consists now of the ENDGame (Even Newer Dynamics for General atmospheric modelling of the environment). The dynamical core uses a semi-implicit semi-Lagrangian formulation to solve the non-hydrostatic, fully compressible deep-atmosphere equations of motion (Wood et al., 2014). The three-dimensional (3-D) wind components, virtual dry potential temperature, Exner pressure and dry density were chosen as the main prognostic variables. Mass mixing ratio of water vapour and prognostic cloud fields are advected as free tracers like other atmospheric constituents. The prognostic equations are solved horizontally on a regular longitude–latitude grid with Arakawa C-grid staggering (Arakawa and Lamb, 1977), the vertical discretization is realized with terrain-following hybrid height coordinates and Charney–Phillips staggering (Charney and Phillips, 1953).

The new dynamical core together with updates in the physics package increase mid-latitude variability and, even more important for the present model evaluation, it increases variability in the tropics. Here tropical cyclones as well as other tropical phenomena can be modelled more accurately with the new model version. Relative to GA6, as described in Walters et al (2017), the changes included in GA7 are less significant with the two main changes being a change to numerical representation of the termination level of convection, which allows convection to terminate at a higher level, and migration to a new aerosol scheme based on the Global Model of Aerosol Processes (GLOMAP, Mann et al., 2010).

The model simulations in this analysis were performed with a n216L85 resolution (approx. 90 km x 60 km horizontal resolution in the tropics). The simulation period is from January 1981 until December 2009. The output frequency of variables of interest was monthly means, calculated during the model run. For the purpose of the following comparison the output is interpolated to 1° x 1° horizontal resolution and to 21 vertical pressure levels

3.4 CMIP5 (as analysed by KIT)

In order to establish an assessment of the current state-of-the-art in climate modelling, against which our models can be compared, we analysed the outputs from 18 models of CMIP5. Within the Coupled Model Intercomparison Project Phase 5 (CMIP5) several coordinated experiments were realized with climate models from all over the world (Taylor et al., 2012) to analyse the capability of the models to represent the past climate and subsequently provide answers about the near and the far future climate. With these coordinated simulations, the scientific community seeks to isolate some of the factors responsible for differences in model projections, with a specific focus on key feedbacks involving clouds and the carbon cycle (Taylor et al., 2011). We chose the AMIP-like simulations that were conducted to simulate the recent past (1979–2008). The AMIP runs were all forced with the same observation-based sea surface temperatures. Differences in model output arise from varying realizations of the atmospheric dynamics and the representations of the earth's surface processes such as land-use or vegetation growth and soil moisture. The differences are also caused by the varying horizontal and vertical resolution of the models. Horizontal grid spacings range from 0.25–2.8125° and the number of vertical levels varies between 18 and 95. The model output is given as monthly averages. For our comparison, we re-gridded the fields to 1°x1° horizontal resolution and to 21 vertical levels ranging from 100–1000 hPa. Original resolutions and references can be found in Table 3.

Table 3: CMIP5 and YoTC models selected for the present study: native resolutions and references. YoTC models are highlighted in red, models that took part in both projects in orange and CMIP5 models in white.

Model	Institute	Lat x Lon	Lev	Reference
CAM5	U.S. National Center for Atmospheric Research (NCAR)	0.9 x 1.25	30	Neale et al. (2012)
CAM5-ZM	NCAR	0.9 x 1.25	30	Song and Zhang (2011)
CCSM4	NCAR	1.25 x 1.25	26	Gent et al. (2011)
CMCC-CM	Centro Euro-Mediterraneo per I Cambiamenti Climatici	0.75 x 0.75	31	Scocciamarro (2011)
CNRM-AM	Centre National de Recherches Météorologiques and Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	1.406 x 1.406	31	Voltaire et al. (2013)
CNRM-CM5	Centre National de Recherches Météorologiques and Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	1.406 x 1.406	31	Voltaire et al. (2013)
FGOALS-s2	The State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, The Institute of Atmospheric Physics	1.67 x 2.8125	26	Bao et al. (2013)
GEOS-AGCM	Global Modeling and Assimilation Office of the National Aeronautics and Space Administration (NASA)	0.5 x 0.625	72	Molod et al. (2012)
GISS-E2-R	NASA Goddard Institute for Space Studies	2 x 2.5	29	Schmidt et al. (2014)
INM-CM4	Russian Institute for Numerical Mathematics	1.5 x 2	21	Volodin et al. (2010)
IPSL-CM5A-MR	Institut Pierre Simon Laplace	1.27 x 2.5	39	Dufresne et al. (2012)
MIROC5	University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1.4 x 1.4	40	Watanabe et al. (2010)
MPI-ESM-MR	Max Planck Institute for Meteorology	1.875 x 1.875	95	Stevens et al. (2013)
MRI-CGCM3	Meteorological Research Institute (MRI)	1.125 x 1.125	35	Yukimoto et al. (2012)
MRI-AGCM3	Meteorological Research Institute (MRI)			Yukimoto et al. (2012)
NAVEM01	U.S. Naval Research Laboratory	T359	42	Hogan et al. (2014)

3.5 YoTC (as analysed by KIT)

The main source of comparison in this study is a selection of models that took part in the Year of Tropical Convection (YoTC) effort. The research program YoTC was implemented by the World Climate Research Program (WCRP) and The Observing System Research and Predictability Experiment (THORPEX) of the World Weather Research Program (WWRP) in order to address specific challenges related to simulating prominent phenomena of the tropical atmosphere (Waliser et al., 2012). The main YoTC focus period runs from May 2008 to April 2010. As a joint research activity between GEWEX Atmosphere System Study (GASS) and YoTC, the Vertical Structure and Physical Processes Multimodel Experiment was conducted using 24 global atmospheric models (Jiang et al., 2015; their Table 1). In Hannak et al. (2017), we analysed these models and found a selection of 8 models particularly suitable for the DACCIWA study area which are shown here as well. We use 20-yr climate simulations covering the period 1991–2010. The models were all run with equal SSTs and sea ice concentrations (NOAA Optimum Interpolation V2 product, Reynolds et al., 2002) and comparable prescribed aerosol. The output of all models was archived every 6 h on a standard horizontal (2.58°x2.58°) grid with 22 vertical pressure levels (nine below 700 hPa), regardless of the model's native resolutions (for original resolutions and references see Table 3).

4 Evaluation data sets

4.1 Radiosonde data

During the DACCIWA campaign in June–July 2016 the upper-air stations network, which is usually not very dense in this area, was extended by several stations with standard radiosonde observations plus additional high frequency boundary layer soundings at the super-sites (Flamant et al., 2017). For this report, only the standard observations were taken into account since they deliver profiles of the complete atmosphere up to 25 km (in the tropics).

Some stations were built up using existing infrastructure and personnel composed of German and local students, for example in Accra, Ghana where the Ghanaian Met Service hosted a station at their headquarters and in Lamto in Ivory Coast, where the Lamto Geophysical Observatory and the Université Félix Houphouët-Boigny run a research facility. Other soundings were conducted at the supersites in Kumasi (Ghana), Savé (Benin) and Ile-Ife (Nigeria). Several stations were built-up during the AMMA campaign (Parker et al., 2008) and were operated during DACCIWA, too. Moreover, other operational stations mainly situated in Nigeria contributed with sounding data to the DACCIWA radiosonde campaign. The management of soundings at Abidjan, Cotonou and Parakou was subcontracted to a private company with a success rate of almost 100%. In Fig. 3, the location of sounding stations is given. Soundings were performed 4 times daily during most of the campaign period, additional soundings were possible. Altogether some 772 radiosondes were launched.

The retrieved data from the soundings were sent to the GTS (Global Telecommunication System) via several pathways. Email real-time transmissions of FM 35 TEMP Mobilemessages to the GTS were made for Lamto, Kumasi, Accra and Parakou. FM 94 BUFFR TEMP mobile messages from Savé were emailed too, but in this case to Météo France, where they were included in the numerical weather prediction models. The subcontracted company and ASECNA (Agence pour la Sécurité de la Navigation Aérienne en Afrique et à Madagascar) ensured real-time transmission of Cotonou and Abidjan soundings to the GTS.

4.2 Ground-based observations

4.2.1 GPCP

The rainfall distribution in our models was compared to rainfall measurements from the Global Precipitation Climatology Project (GPCP). It was established by the World Climate Research Programme to quantify the distribution of precipitation around the globe on climatological time scales (Adler, 2003). In GPCP, various sources of measurements are combined to give a merged product on global precipitation with high accuracy. Therefore, individual strengths and weaknesses of the individual measurement types are considered.

There are mainly two types of sources: (a) ground-based rain gauge measurement from synoptical stations and (b) satellite data. The ground-based rain gauge measurements are taken from the Global Precipitation Climatology Centre (GPCC) monitoring product of the German weather service which combines all available station measurements to a ground-based monitoring product including error characteristics. The satellite data are merged from a vast range of sensors, mainly using infrared and microwave radiance derived rainfall estimates. The estimates are a combination of measurements from geostationary as well as polar orbiting satellites. Infrared measurements are taken from GOES, Meteosat and GMS (geostationary) plus NOAA (polar orbiting). Additional infrared precipitation estimates from the Geostationary Satellite Precipitation Data Centre (GPSPDC) is merged into the final product. Over ocean and land, microwave measurements from SSM/I (Special Sensor Microwave/Imager, polar orbiting) from two channels are used together with AIRS

(Atmospheric Infrared Sounder on Aqua, polar orbiting) and TOVS (TIROS Operational Vertical Sounder on the polar orbiting NOAA satellites).

4.3 Satellite data

4.3.1 GERB

Outgoing longwave and shortwave radiation from our models were validated with GERB products. GERB is the geostationary earth radiation budget instrument onboard the EUMETSAT's Meteosat Second Generation satellites (MSG) (see Harries et al., 2005). It is a broadband radiometer that is designed to measure the total emitted and solar reflected radiances over the earth with high temporal resolution (5 min) and 50 km spatial resolution. From the measurements, the top of atmosphere radiation, divided in short and longwave contributions can be estimated to give an accurate estimation of the earth's radiation budget (or rather the budget in MSG's field of view which covers almost half of the earth's surface). For the present model evaluation, we used the "Standard High-resolution Image" (DeWitte et al., 2008), in which the data are processed using cloud observations from SEVIRI to convert radiance measurements to radiative fluxes at approximately 10 km resolution, every 15 min.

4.3.2 SEVIRI

Liquid water path (LWP) and Ice water path (IWP) was evaluated with CLAAS 2 (CM SAF CLOUD property dAtAset using SEVIRI – Edition 2) data which is a climate data record from the Satellite Application Facility on Climate Monitoring (CM SAF) based on SEVIRI measurements on the geostationary MSG satellites (Stengel et al., 2014). The MSG satellites do not only carry the broadband radiometer GERB but also the spectrally resolved radiometer SEVIRI. SEVIRI is the Spinning Enhanced Visible and InfraRed Imager, it scans the earth in 12 spectral channels, ranging from the visible to the near infrared part of the electromagnetic spectrum. The covered wavelength range of SEVIRI corresponds to that of GERB, where GERB covers up for the gaps in between SEVIRI's channels but has a coarser horizontal resolution. Every 15 minutes a new image is produced by SEVIRI. The horizontal pixel size varies from 3 km at nadir to about 11 km towards the rim of SEVIRI's field of view.

LWP and IWP are two variables in a larger suite of SEVIRI-derived products that are all dedicated to clouds such as cloud fraction, cloud top height pressure and temperature, cloud type and phase, effective droplet radius and cloud optical thickness. The advantage of such a dataset is the temporal stability in terms of retrieval algorithms and input data. First, the dataset is compiled in one go, consistently using one retrieval version. Second, the input data from MSG 1, 2 and 3 were intercalibrated with MODIS Aqua to ensure a homogeneous data basis (Meirink et al., 2013). The whole dataset covers 12 years (2004–2015), we used in this validation the years 2006–2015. The products are available in various temporal formats, we used the monthly averages with a horizontal resolution of $0.05^\circ \times 0.05^\circ$ and the monthly mean diurnal cycles with $0.25^\circ \times 0.25^\circ$ horizontal grid spacing in hourly resolution.

4.3.3 ATOVS

The data that we used to compare the total precipitable water (TPW) stem from another CM SAF dataset. It is the vertical integrated water vapour derived from ATOVS measurements (Courcoux und Schröder, 2013). ATOVS is mounted on the polar orbiting NOAA satellites (15–19) and on Metop-A. It is not a single instrument but a system of three sounders, two microwaves sounders, AMSU-A (Advanced Microwave Sounding Unit-A) and AMSU-B (Advanced Microwave Sounding Unit-B), later replaced by MHS (Microwave Humidity Sounder) and an infrared sounder HIRS (High-

resolution InfraRed Sounder). AMSU-A and B are both cross track scanning instruments with a nominal field of view of 3.3° and 1.1° , respectively, with a horizontal resolution of 48 and 16 km. AMSU-A has 15 channels in the microwave spectrum located between 23.8 and 89 GHz in order to deliver mainly temperature soundings of the atmosphere and give information on water vapour as well as precipitation over ocean and sea ice coverage. AMSU-B has five channels between 89 and 183.31 ± 7 GHz where the primary goal is to measure water vapour and precipitation over land and ocean. HIRS scans the earth cross track in the infrared in 20 channels with an instantaneous field of view of 1.3° providing a nominal spatial resolution of 18.9 km, or 10 km for newer instruments. HIRS performs atmospheric temperature sounding and measures water vapour, ozone, N₂O plus cloud and surface temperatures.

The water vapour information from the CM SAF ATOVS data set we used for this study was temporally averaged to monthly means on a cylindrical equal area projection of $90\text{km} \times 90\text{km}$. The vertically integrated water vapour (or TPW) is the total atmospheric water vapour contained in a vertical column of unit cross-sectional area extending between the surface and 100 hPa.

5 Results

5.1 Climatology: vertical structure of the atmosphere

First of all, we compare the models in terms of their general climatology which are profiles of cloud cover, temperature, relative humidity and horizontal wind speed averaged over the DACCIWA region 8°W – 8°E and 5 – 10°N . The considered months are June, July, August and September which include the wet monsoon season. The onset of the monsoon is connected to the installation of the Atlantic cold tongue that develops during March to May and brings upwelling cold water to the coast of southern West Africa. Due to the enhanced insolation on the northern hemisphere, the Saharan heat low develops which together with the Atlantic cold tongue leads to a north-south pressure gradient. The onset of the monsoon in our DACCIWA box in June is indicated by the northward shift of the Intertropical Discontinuity and a change in circulation pattern (south westerly inflow of wind from the Atlantic) and consequently the northward migration of the rain band. The rain band shifts further north until the maximal displacement is reached in August/September, thereafter the driving forces of the rainy season become weaker. Consequently, the Intertropical Discontinuity shifts southward again and the north-easterly wind pattern with dry air from the Sahara returns.

In Fig. 4 the cloud cover profiles of all YoTC models averaged over the box is shown together with the profiles of the ECHAM runs for the months June, July, August and September. Altogether the models show a remarkable spread of about 10–20% for June and even 10–25% for the other months. All models have a high cloud maximum between 200 and 300 hPa indicating cirrus clouds and partly the high convective systems connected to the ITCZ. For most YoTC models there is no distinct midlevel cloud maximum visible except for MRI-AGCM in all months and CAM5ZMMicroCAPT and NCAR_CAM5 in July and August. In contrast to this, all ECHAM simulations possess a pronounced maximum in midlevel cloud cover ranging from 18 to 30% between June and September with a maximum in July. The realization ec1 has the highest midlevel cloud fraction with about 30%. Eca, which is the pure ECHAM run, has on average the highest cloud cover at low levels. As analysed already in more detail in Hannak et al. (2017), most models substantially underestimate the lowest clouds (about 200–300 m above the surface) accompanied by an overestimation of clouds at the top of the monsoon layer. This indicates systematic problems in vertical exchange processes, which are also reflected in large errors in the nocturnal low-level jet speed (see below). Consequently, many models show a too flat diurnal cycle in cloudiness.

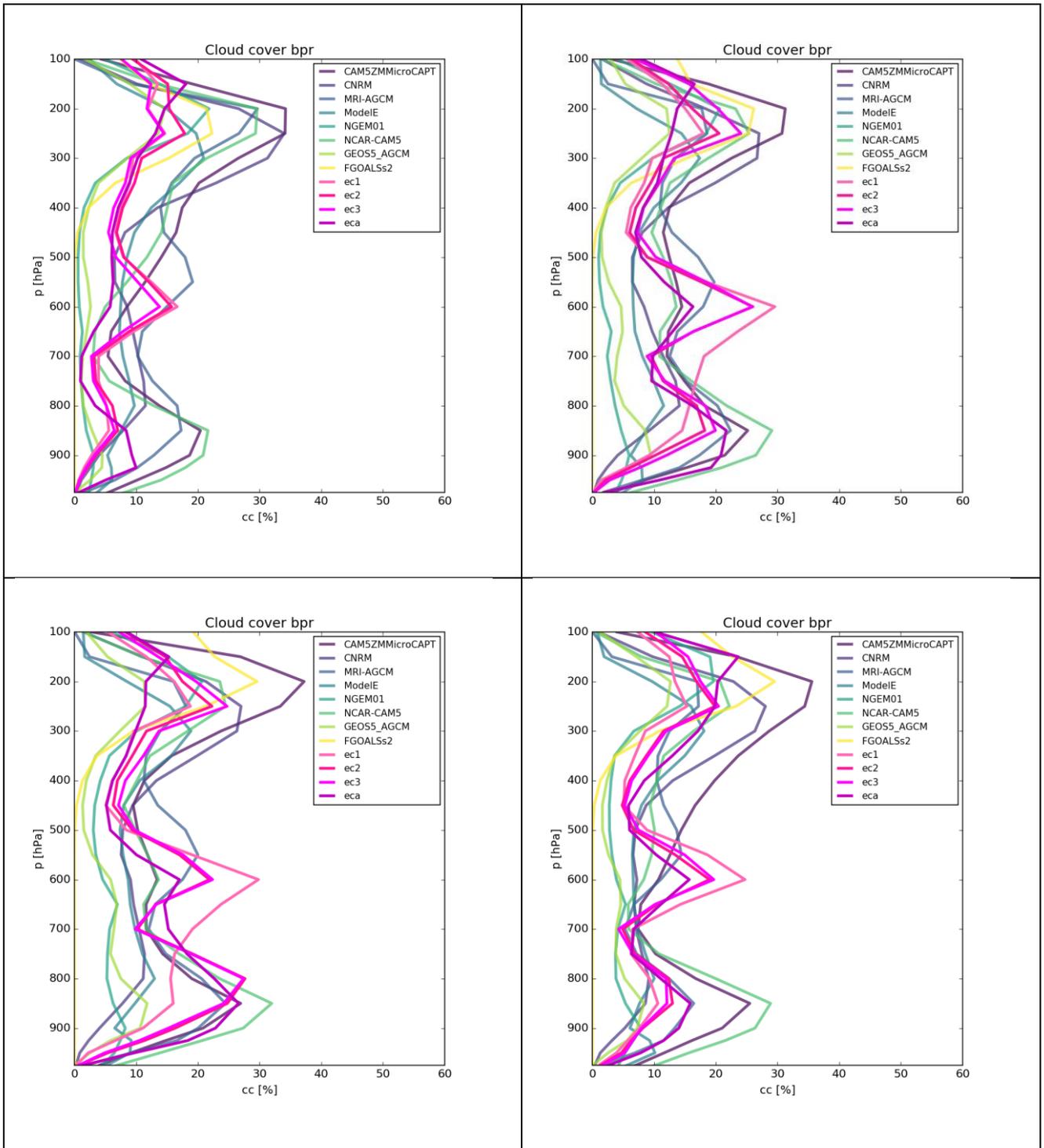


Figure 4: Box average ($8^{\circ}W-8^{\circ}E$ and $5-10^{\circ}N$) of cloud cover profiles for the YoTC models and the DACCIWA research models for the months June, July, August and September (clockwise starting from top left).

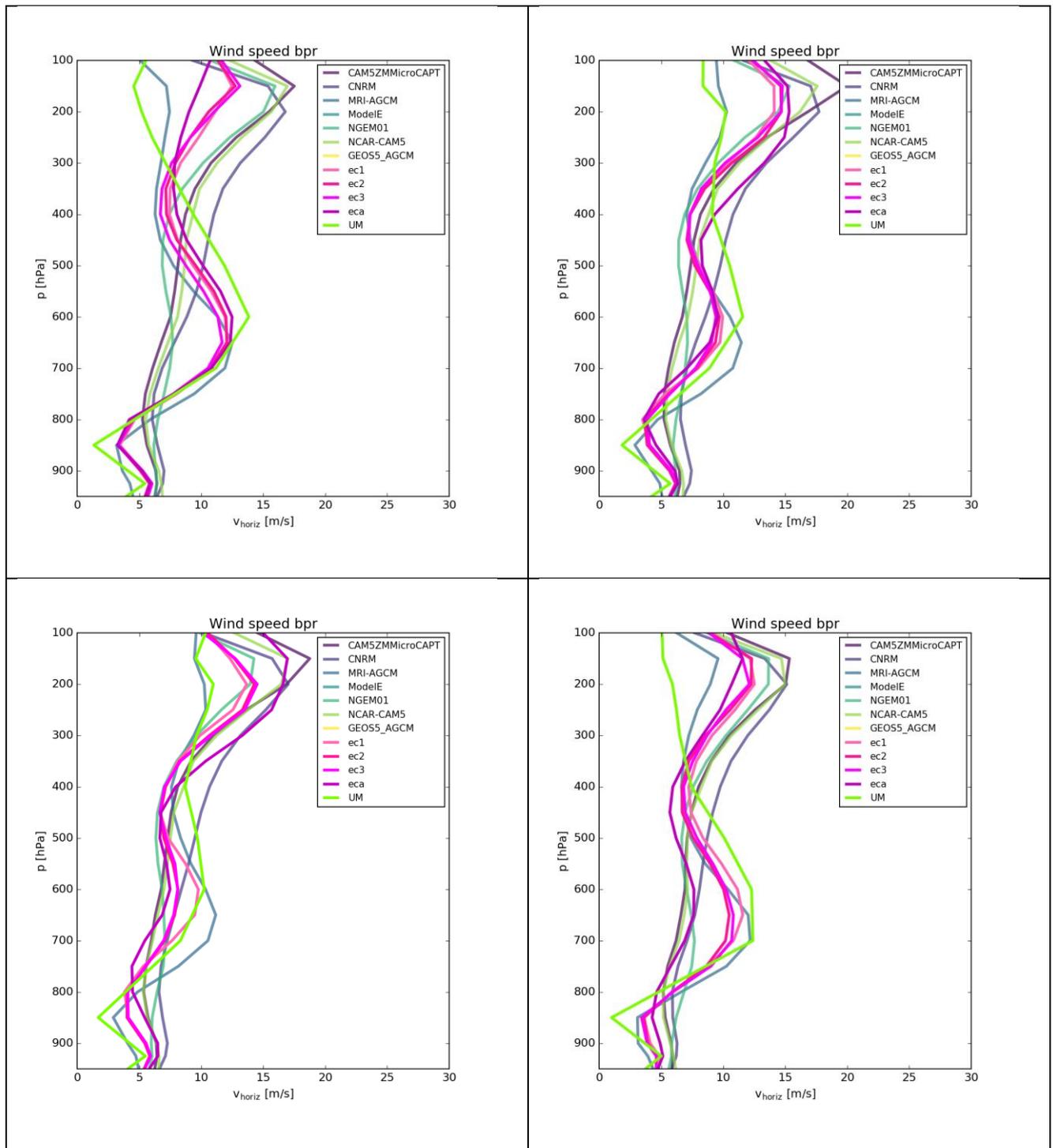


Figure 5: Box average ($8^{\circ}W-8^{\circ}E$ and $5-10^{\circ}N$) of horizontal wind speed profiles for the YoTC models and the DACCIWA research models for the months June, July, August and September (clockwise starting from top left).

Figure 5 shows too high horizontal wind speeds in the lowest levels (again see Hannak et al., 2017) for the YoTC models. However, the different versions of the ECHAM and the UM exhibit weaker winds on average. ECHAM, UM and MRI-AGCM (from YoTC) exhibit a pronounced mid-level wind speed maximum at approximately 650 hPa. This maximum weakens in August but otherwise stays equally strong about 10 m/s for the other months. At this height, the African

Horizontal wind speed seasonal mean

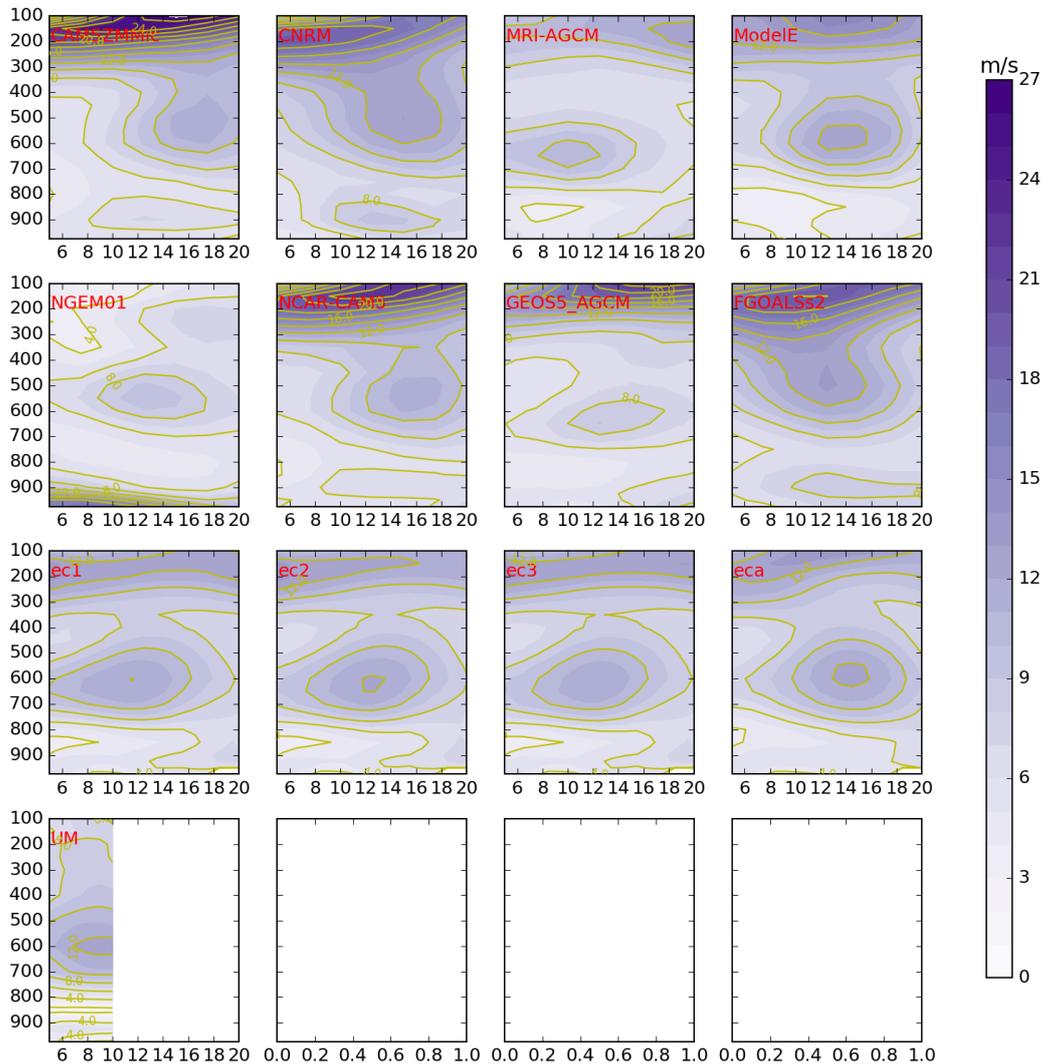


Figure 6: Zonal (8°W–8°E) average of the horizontal wind speed as seasonal mean. Latitude is expressed on the x-axis as degrees north while pressure is expressed on the y-axis in hectopascal. Please note that UM data were only available for the DACCIWA box (up to 10°N).

Easterly Jet (AEJ) is situated (Thorncroft, 2003), but its maximum typically occurs further north at about 12–15 degrees. Thus, it appears that the YoTC models place the jet at a more northward position than our research models and therefore the maximum is not reproduced in the average profiles over the DACCIWA box. The confirmation can be found in Fig. 6, where a zonal average of the horizontal wind speed is depicted. The AEJ is present in most models, but with marked longitudinal and intensity differences. For the UM, only the data from our DACCIWA box was available, which does not allow a further interpretation.

Again in Fig. 5, the other YoTC models show a more or less linear increase in wind speed up to the tropopause with no visible peak in the mid-levels. However, around 200 hPa all models show a

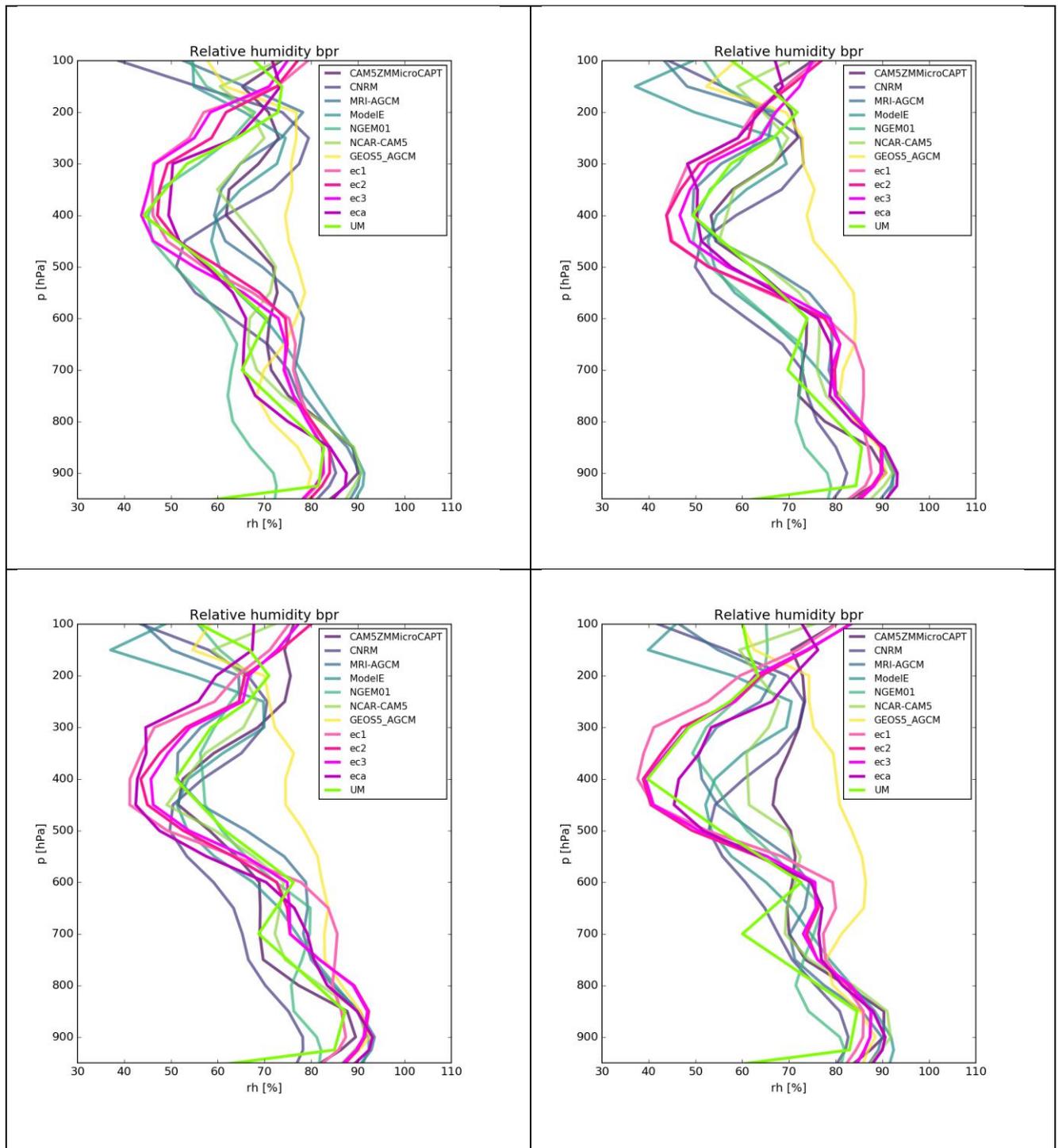


Figure 7: Box average ($8^{\circ}W-8^{\circ}E$ and $5-10^{\circ}N$) of relative humidity profiles for the YoTC models and the DACCIWA research models for the months June, July, August and September (clockwise starting from top left).

maximum corresponding to the Tropical Easterly Jet (TEJ; Nicolson and Grist, 2003) that has its maximal strength located at approximately $7-10^{\circ}N$. Noticeably, the UM has a strikingly weak TEJ, in June there is even a local minimum in wind speed. Also the ECHAM eca run does not show the jet but solely for the month of June.

The profiles of relative humidity, shown in Fig. 7, exhibit for all models, YoTC as well as the ECHAM versions and the UM, a comparable structure, with the exception of GEOS-AGCM that does not possess a pronounced vertical variation. The models show a maximum of relative humidity at about 850–900 hPa where the low-level clouds form. Around the top of the boundary layer and above, the atmosphere dries in all models. At about 600–650 hPa a small secondary

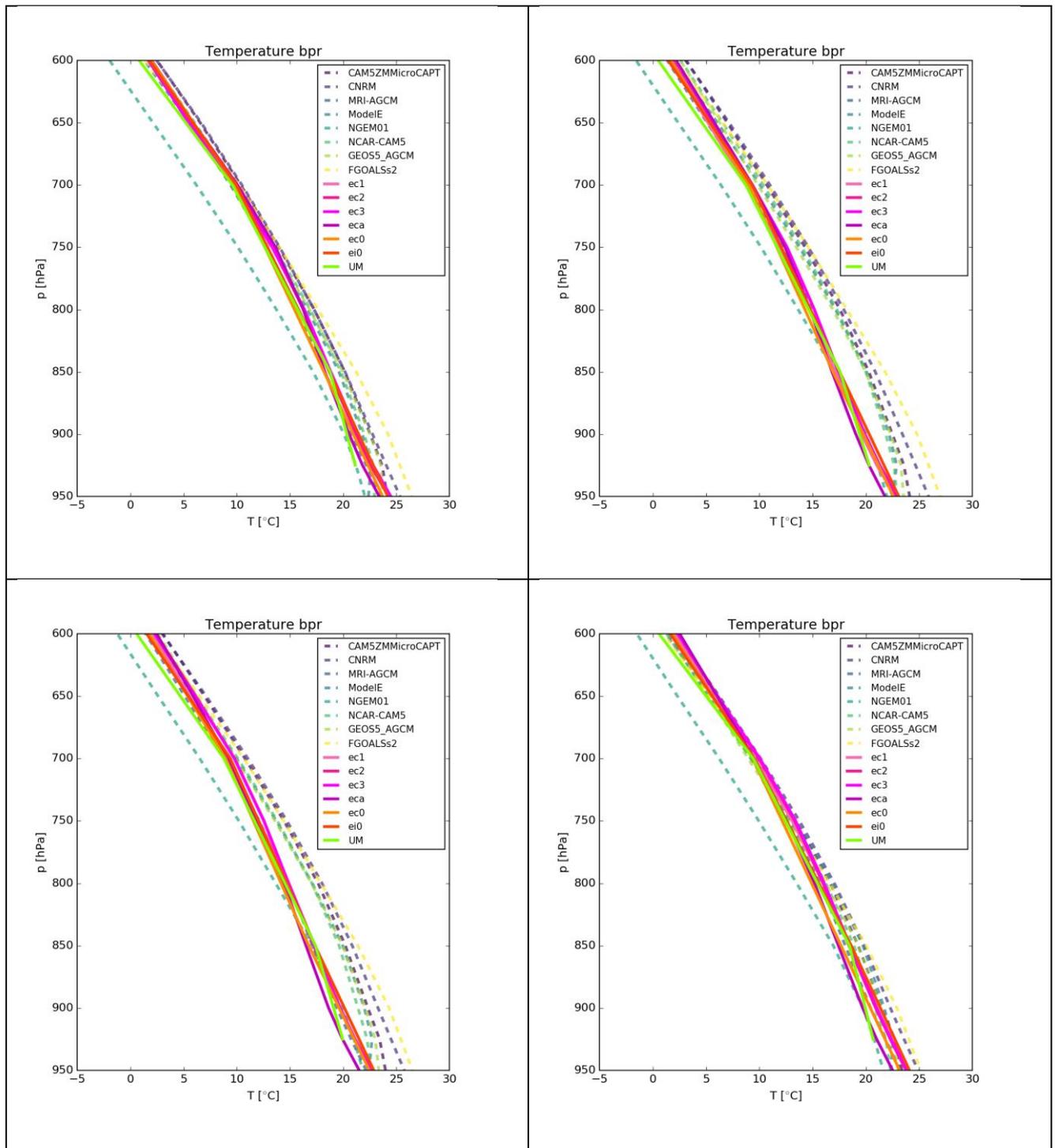


Figure 8: Box average ($8^{\circ}W-8^{\circ}E$ and $5-10^{\circ}N$) of temperature profiles for the YoTC models and the DACCIWA research models for the months June, July, August and September (clockwise starting from top left).

maximum appears, which is stronger in September than in the other months. This is not seen by all models, CNRM does not show a mid-level maximum of relative humidity and no mid-level cloud peak. Around 400 hPa all models simulate a minimum in relative humidity, which may come from horizontal advection of drier air from the north and less effective convective moistening. This is followed by an increase in relative humidity at higher levels because the fractional detrainment rate increases as the convective mass flux decreases strongly (Roms, 2014).

The temperature profiles for the YoTC models, ECHAM, UM and IFS are depicted in Fig. 8 for the lower part of the atmosphere (from 600 hPa to surface) since this resolves the most interesting features best. The models exhibit various degrees of temperature lapse rates, which allow more or

less convection and therefore upward transport of moisture to higher layers. In July and August there seem to be two regimes of stratification which divide the group of models. Our research models (plus 2 YoTC models) tend to have a comparably large temperature gradient from the surface to 850 hPa and a smaller one from 850 to 700 hPa. For the remaining YoTC models this is reversed which could explain the differences in mid-level cloud cover by changed upward transport of moisture. The ECHAM-HAM runs exhibit very similar temperatures, but the ECHAM run (denoted as eca in the figure) has lower temperatures. This explains partly the higher cloud fraction, as low temperatures at constant specific humidity lead to enhanced relative humidity, which triggers the formation of clouds in the climate models (cloud formation is parameterized with the Sundqvist scheme that depends only on relative humidity).

The final Fig. 9 in this section shows a zonal mean of the fractional cloud cover. Apart from the already mentioned vertical stratification, only the very small meridional variations in the climate models attract attention. Only in the low cloud layers some variation in the models is visible, mainly with more low levels clouds close to the coast and lesser so farther north. CAM5ZMMicroCAPT, NGEM01 and NCAR-CAM5 show an opposite gradient. Values in the lower layers range from virtually nothing (FGOALS) to 40 and more %, while observational products range around 20% (see Fig. 7 in Hill et al., 2016). However, due to obscuring by higher clouds, satellite based estimates are generally problematic.

5.2 Cloud properties and radiation

In this section, the horizontal cloud cover distribution and its influence on outgoing longwave and shortwave radiation is briefly discussed. The seasonal average of LWP and IWP is shown in Fig. 10 for some of the YoTC models plus the ECHAM and UM model runs. They are compared to LWP and IWP measurements derived from SEVIRI. Since a shortwave channel for the retrieval of LWP and IWP from SEVIRI has to be used, the averages are compiled for 12 UTC. The SEVIRI data were compiled from the dataset CLAAS 2 which includes a monthly mean diurnal cycle product. We created the mean from the 12 UTC slot by averaging over the respective time-frame. Not all YoTC models had data available for 12 UTC, therefore we could pick only a few for this figure. Also, no diurnal resolution was available for the UM; here only the seasonal mean averaged over all points in time is shown for a general comparison.

Cloud fraction seasonal mean

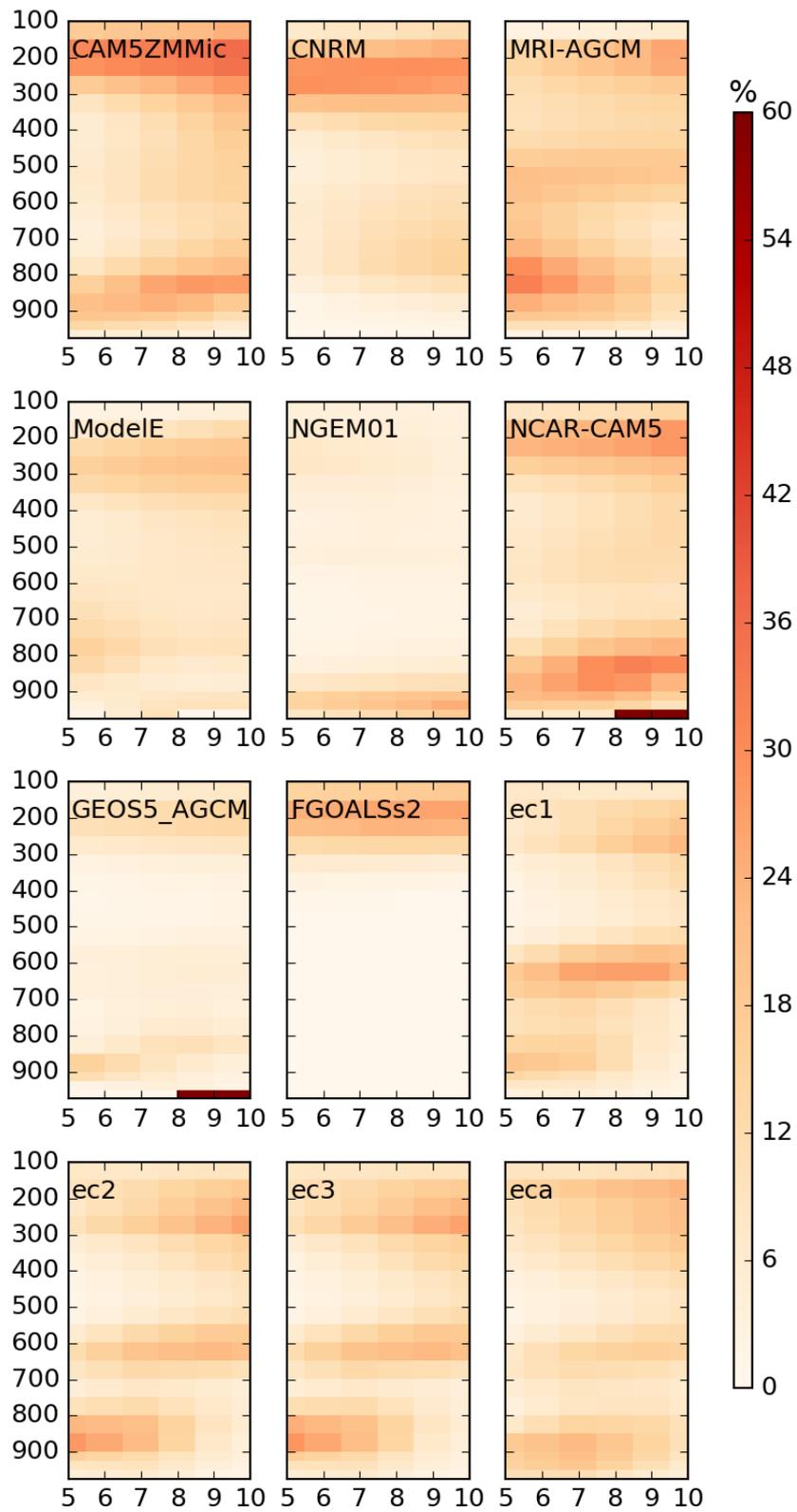


Figure 9: Zonal mean (8°W–8°E) of cloud cover for the DACCIWA box. The latitude is shown on the x-axis as degrees north and the pressure in hectopascal on the y-axis.

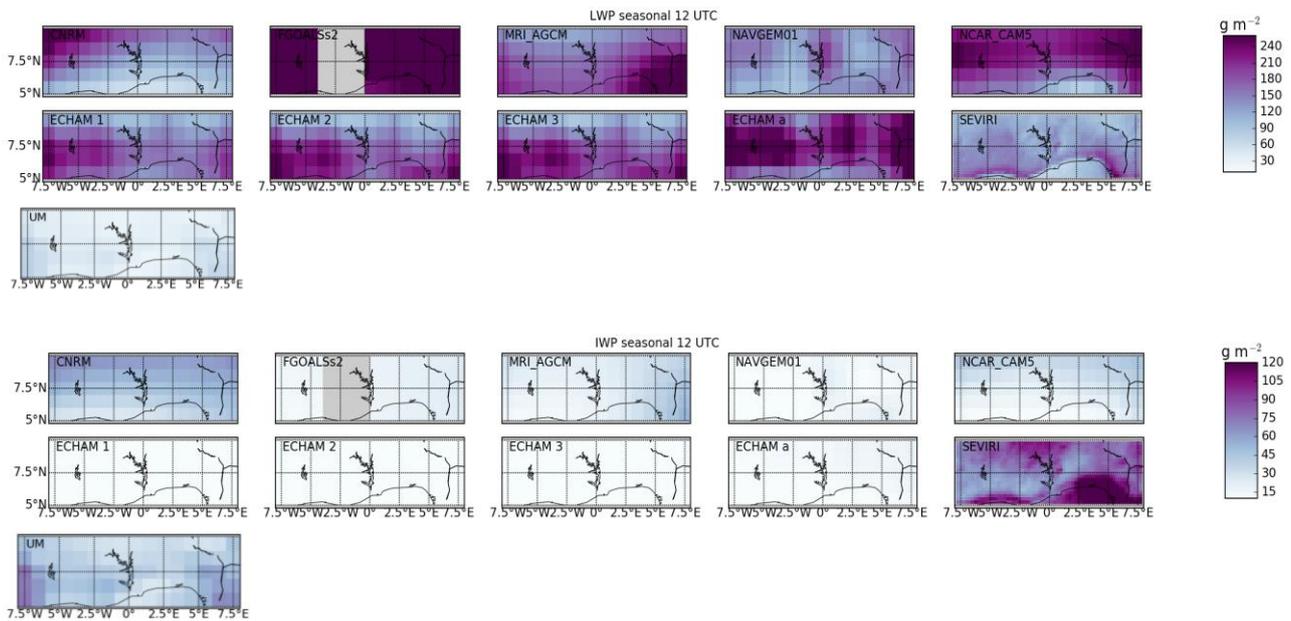


Figure 10: Seasonal mean of liquid water path (upper panel) and ice water path (lower panel) in the DACCIWA box for those YoTC models where the 12 UTC slot was available and the ECHAM runs plus UM averages. The latter were only available as all-day-average. The models are compared to SEVIRI-derived LWP and IWP 12 UTC averages from the years 2006–2015.

The models vary considerably in both, amount of liquid water and horizontal distribution of it. Even among the ECHAM runs large differences can be found. All models show more liquid water and less ice water than the SEVIRI measurements suggest. Since SEVIRI can observe only the cloud tops, it is naturally biased towards higher IWC and lesser LWC, but the results are already filtered for mixed-phase clouds, so the error caused by the measurement principle should be smaller. CNRM and NCAR CAM5 place the largest LWPs towards the northern edge of the box, while MRI AGCM and the ECHAM versions ec1, ec2 and ec3 place the thickest clouds more towards the box's southern edge. FGOALS possesses the highest LWP values even though most clouds are concentrated to the 150–250 hPa levels and there the cloud cover is not very much pronounced (compare Figs. 4 and 9).

ECHAM produces comparable results for ec1, ec2 and ec3, where ec1 exhibits the lowest LWP values but the horizontal distributions are conserved. Eca however differs in both amount and distribution. The largest values of LWP are produced in this run and the location of the maxima is shifted to the middle of the box. The ec1-ec3 simulations differences are related to the use of a 2-moment scheme and different tuning in ECHAM-HAM as well as difference in aerosol. Ec1 has the lowest aerosol particle concentration, while ec3 has the highest. Together with increasing LWP, we can state that at this point no secondary aerosol effect is visible, since this should lead to decreasing average LWP values (while increasing the droplet number concentration). Without a more detailed analysis, no decisive conclusions can be drawn.

The UM produces more ice water than liquid water which is more in accordance with SEVIRI than all other models, but the absolute values are very low. A likely explanation for this discrepancy is that the IWP and LWP diagnostics do not directly include the water in the cores of active convection, only the cloud water detrained into the large scale cloud. More cannot be said because of the lack of diurnal resolution in the UM data that were available for this study.

Overall, Fig. 10 demonstrates the substantial diversity between models in details of cloud amount and cloud internal structure.

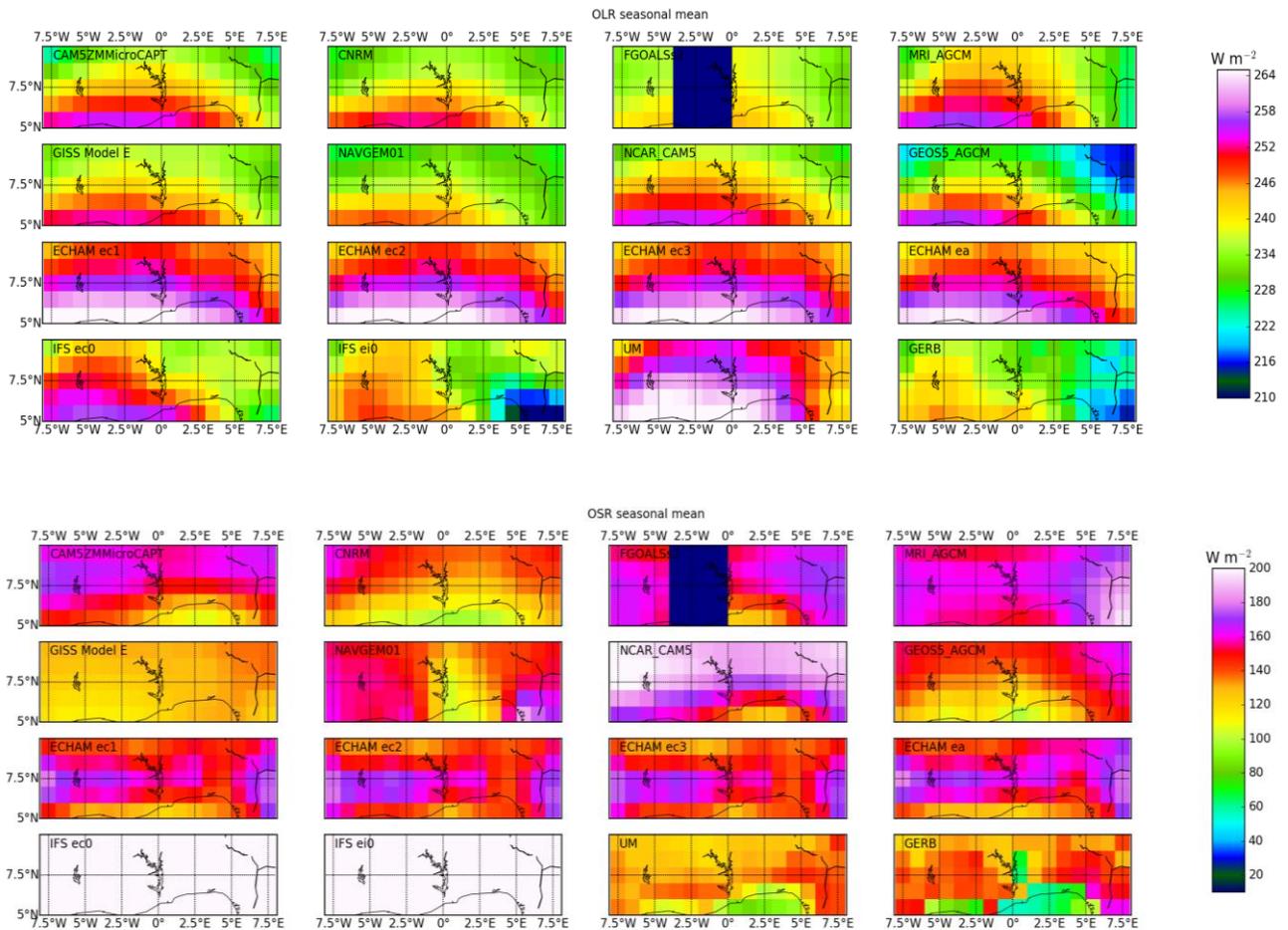


Figure 11: Outgoing longwave radiation (upper panel) and outgoing shortwave radiation (lower panel) at top of the atmosphere as seasonal mean for the YoTC models, the ECHAM and the IFS runs and the UM compared to GERB measurements from 2006–2015. OLR is influenced by the thickness of clouds (in a geometrical sense as well as optical) while OSR is more depending on the scattering properties of the cloud tops. Please note that for IFS only net radiation was available, therefore shortwave fields cannot be shown and for the longwave fields the incoming longwave radiation was assumed to be exactly zero.

The influence of clouds on outgoing shortwave and longwave radiation (see Fig. 11) also varies significantly from model to model. The UM has by far the lowest amount of liquid water and not too much ice water, and therefore it is not surprising that the seasonal mean of OLR for the UM is largest. For the other models the effect of clouds is not easy to describe without additional information, e.g. NAVGEM does not have a lot of liquid water or ice but at the same time not a very high OLR. This could either mean that the diurnal cycle is very strong and no general conclusions can be drawn from comparing a figure for 12 UTC with all-day averages, or that the microphysical properties of the clouds contribute strongly to the observed variation.

The models' OLR can be compared to the satellite measurements of GERB. The highest values of OLR from GERB are found in the southwest corner of the DACCIWA box, the smallest in the southeast corner, while a triangle-shaped region with medium values can be found in between. This structure is very well repeated by the IFS runs and most other models show at least the maximum in approximately the right position. The absolute values are too high from all models but IFS. Hence, the radiative effect of clouds is not modelled well. The GERB OSR field places the smallest values at the southern edge in the centre with larger values immediately northwards from it. All models show the same structure but the absolute values vary in a range of 190 W/m², which is much more than for OLR (55 W/m²).

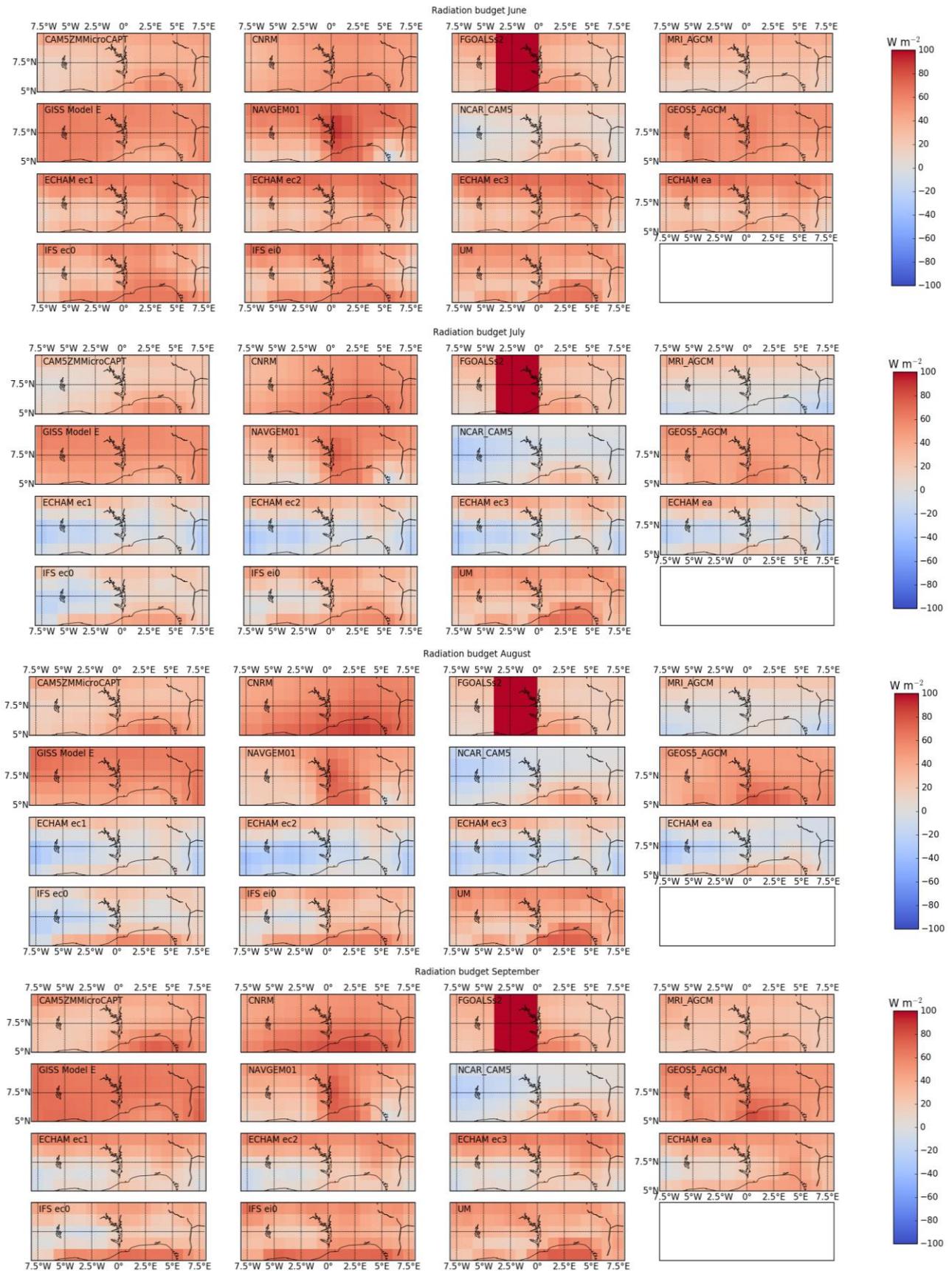


Figure 12: Radiation budget for the months June (upper panel), July (2nd panel), August (3rd panel) and September (lowest panel) for the YoTC models, the ECHAM and the IFS runs plus UM. Unfortunately, GERB data are missing since no top of atmosphere incoming solar radiation was available for this product.

In the following, the earth radiation budget is analyzed for the months June, July, August and September. To compile the averages for the YoTC datasets, we used the outgoing shortwave and longwave radiation fields and subtracted them from the incoming solar radiation from the MRI-AGCM3 model, since all models were run with the same climatic conditions. Due to lack of data, we applied the same procedure for ECHAM and UM, while the IFS runs already contained net radiative fluxes. In June, all models show a positive radiation budget which can be expected for this region (Stephens, 2015). The budget shows a north-south gradient, which is caused by the interplay of variation in incoming solar radiation and the partly zonal orientation of the cloud cover. Only NCAR-CAM5 has a neutral or slightly negative budget in some pixels. For the months July and August, the budget becomes negative for some models, but not for all. MRI-AGCM shows a north-south gradient with the negative values in the south and NCAR-CAM5 has now mainly negative values in the whole box. Our research runs exhibit negative values in all ECHAM runs in a zonal stripe from 6–8°N, The IFS runs show both negative values in the western part of the box, where the ec0 member exhibits a more negative budget than ei0. UM shows positive budget for all months with only slight variations. In general, the budget terms change to the positive for September, but not quite as much as in June for all models. The month to month changes in the budgets are caused by the variation in insolation and to a large extent by modifications in cloud cover. The liquid water path in the DACCIWA box has for all models a month to month variation, where in general an increase in LWP can be observed for the month July and August compared to June and again a decrease for September (see Appendix). This effect is roughly confirmed by SEVIRI measurements, but the average values are much lower as already discussed above.

5.3 Average climate at DACCIWA stations

Evaluation of numerical atmospheric models often lacks vertically resolved measurement data to complete the discussion. Particularly in southern West Africa the density of those measurements, let it be radiosonde, radar measurements or other, is sparse. The DACCIWA campaign therefor forms a positive exception since more than 750 radiosondes distributed over several sites were launched during June and July 2016. In the following, the climate models shall be compared to these data. Of course, it has to be taken into account that two single months of measurements cannot be representative for the long-term climate in this region, but the exceptional availability and particularly the accuracy of the soundings outperform any other kind of vertically resolved measurements in this area which makes it a valuable source of validation data. Results are shown for July 2016; the same figures for June 2016 can be found in the Appendix.

The temperature soundings depicted in Fig. 13 are arranged from east (Ivory Coast) to west (Benin). The radiosonde soundings from coastal stations Abidjan, Accra and Cotonou do not vary considerably on average from the more inland located stations Lamto, Kumasi, Savé and Parakou. The soundings compiled from the climate models, however, do show some variations where the inland stations tend to have a slightly bigger temperature gradient between 900 and 750 hPa. Due to the grid box size of the climate models during run times, coastal profiles will be “contaminated” by ocean to some extent. There is no such grouping of profiles shapes (steeper gradient in lower levels followed by a flatter one at higher levels and vice versa) as was observable for the box-averaged profiles from Fig. 8, therefore the radiosondes cannot help to decide which course of profile is more realistic.

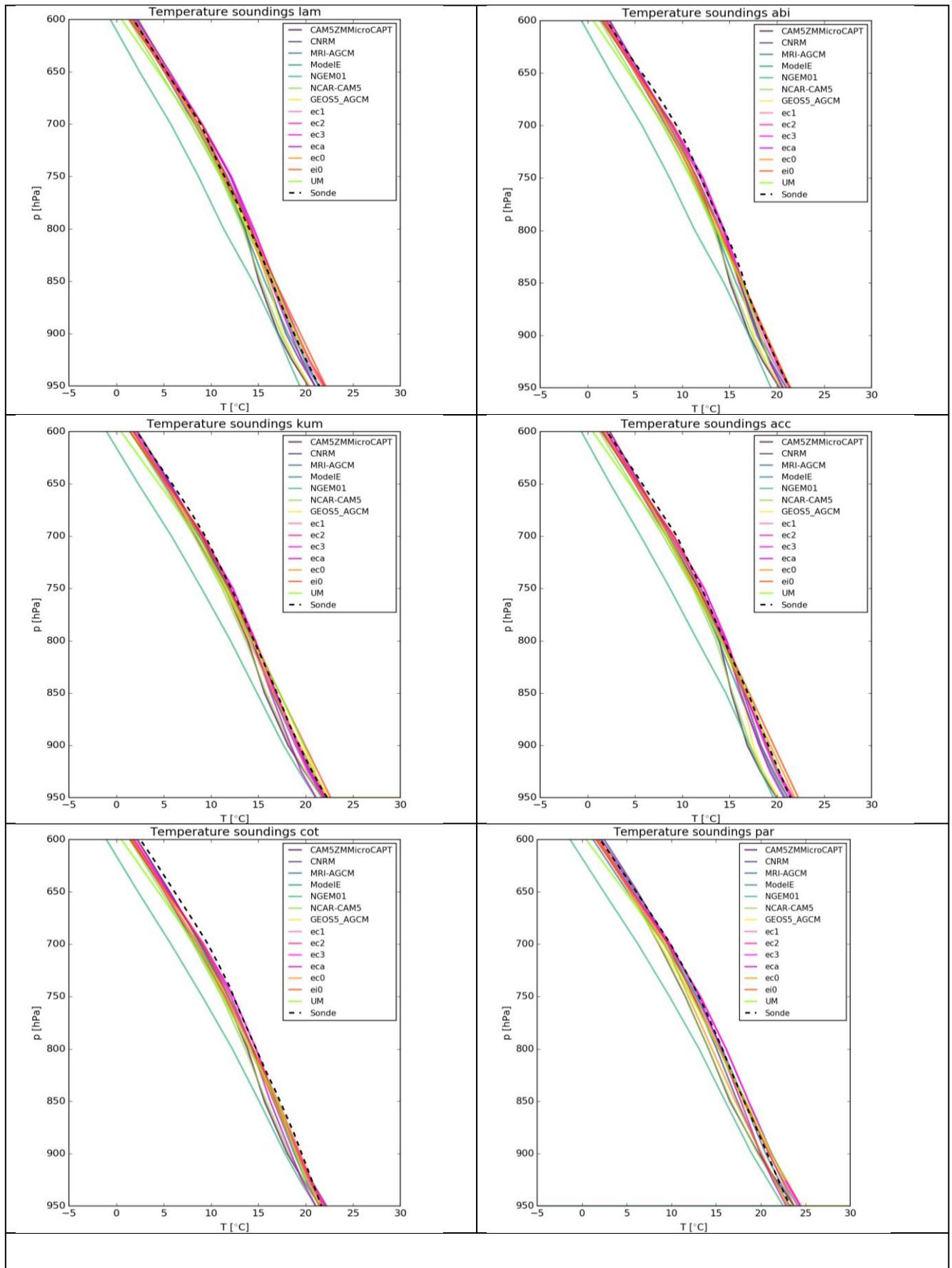
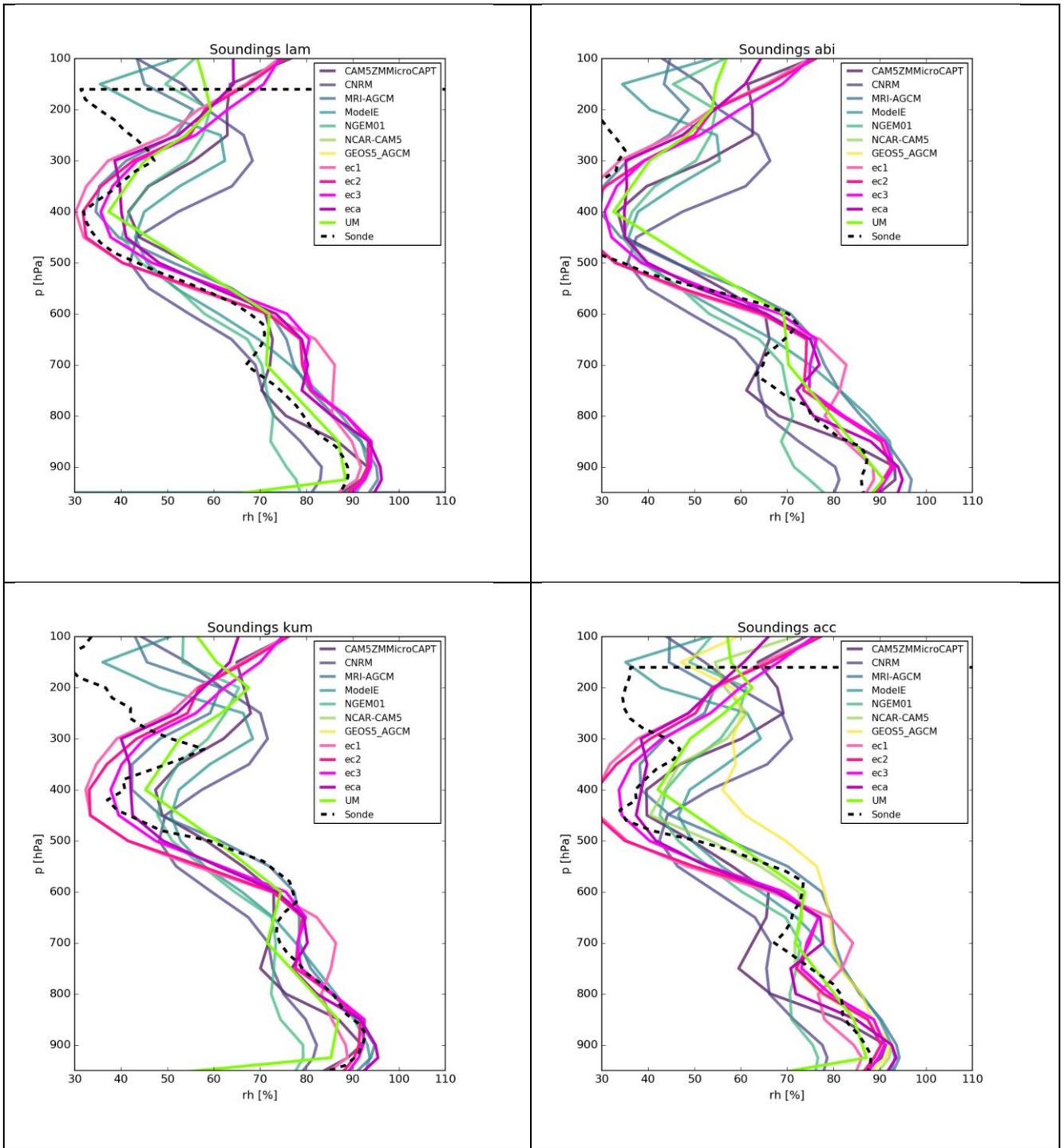


Figure 13: Temperature profiles for July averaged over all years for the YoTC models and our research models compared to the radiosonde measurements from July 2016 at the sites Abidjan, Accra, Cotonou, Lamto, Parakou Kumasi and Savé.



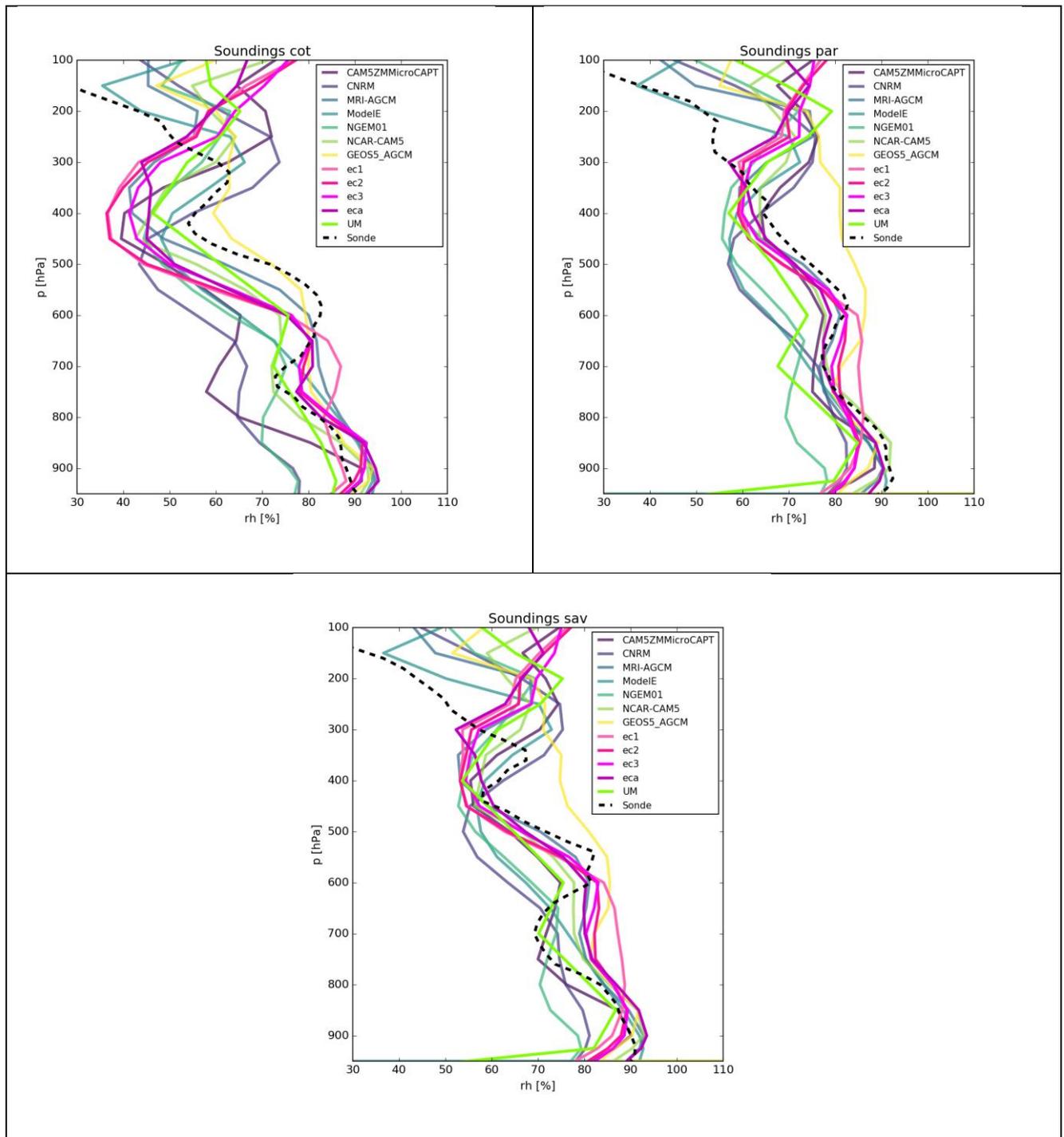
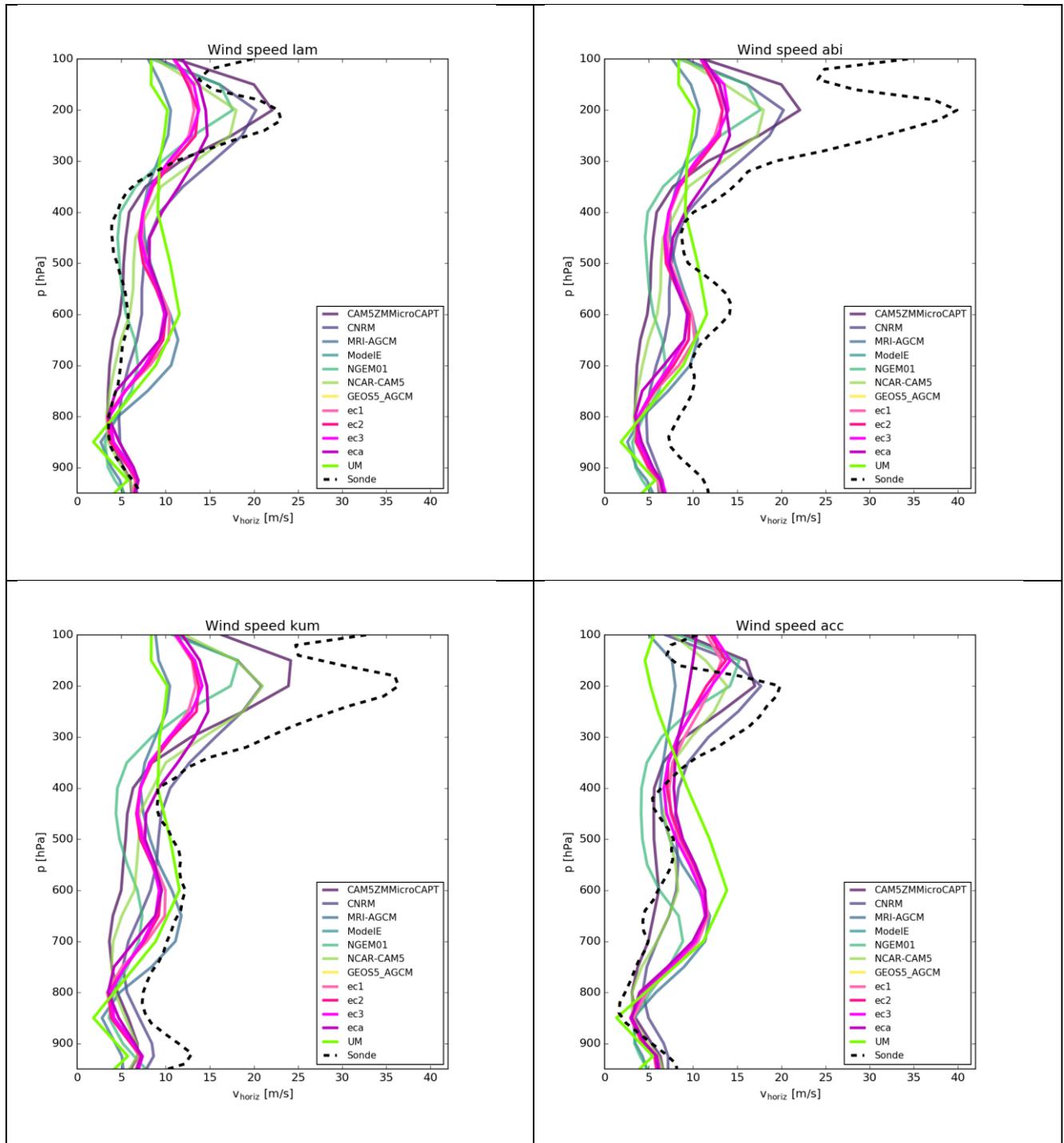


Figure 14: Relative humidity profiles for July averaged over all years for the YoTC models and our research models compared to the radiosonde measurements from July 2016 at the sites Abidjan, Accra, Cotonou, Lamto, Parakou Kumasi and Savé.

In contrast to the temperature profiles, variation is more easily observed in the profiles of relative humidity in Fig. 14. Generally, the radiosonde profiles are placed well in between the model profiles, no systematic shifts appear. Second, there seems to be an east-west gradient in relative humidity range which can be observed for both, the models and the radiosonde soundings. The vertical relative humidity gradient from the surface to about 400 hPa is much more pronounced in the western stations than in the eastern stations and thus the minimum around 400 hPa is smaller.

Still there are some major differences between models and radiosonde soundings: the soundings all show a second maximum of relative humidity at about 600 hPa, that is not produced by all models (which therefore lack mid-level cloud cover) and if it is produced, it is placed too low. The latter is

particularly true for the ECHAM simulations. The UM-profile places the second maximum in the same vertical position as the radiosondes. The last maximum, roughly about 300 hPa is, in contrast to the second one, placed too high by the models, which means that the tropopause is modelled too high compared to the measurements. However, it needs to be taken into account that radiosonde measurements become increasingly uncertain at very low humidities.



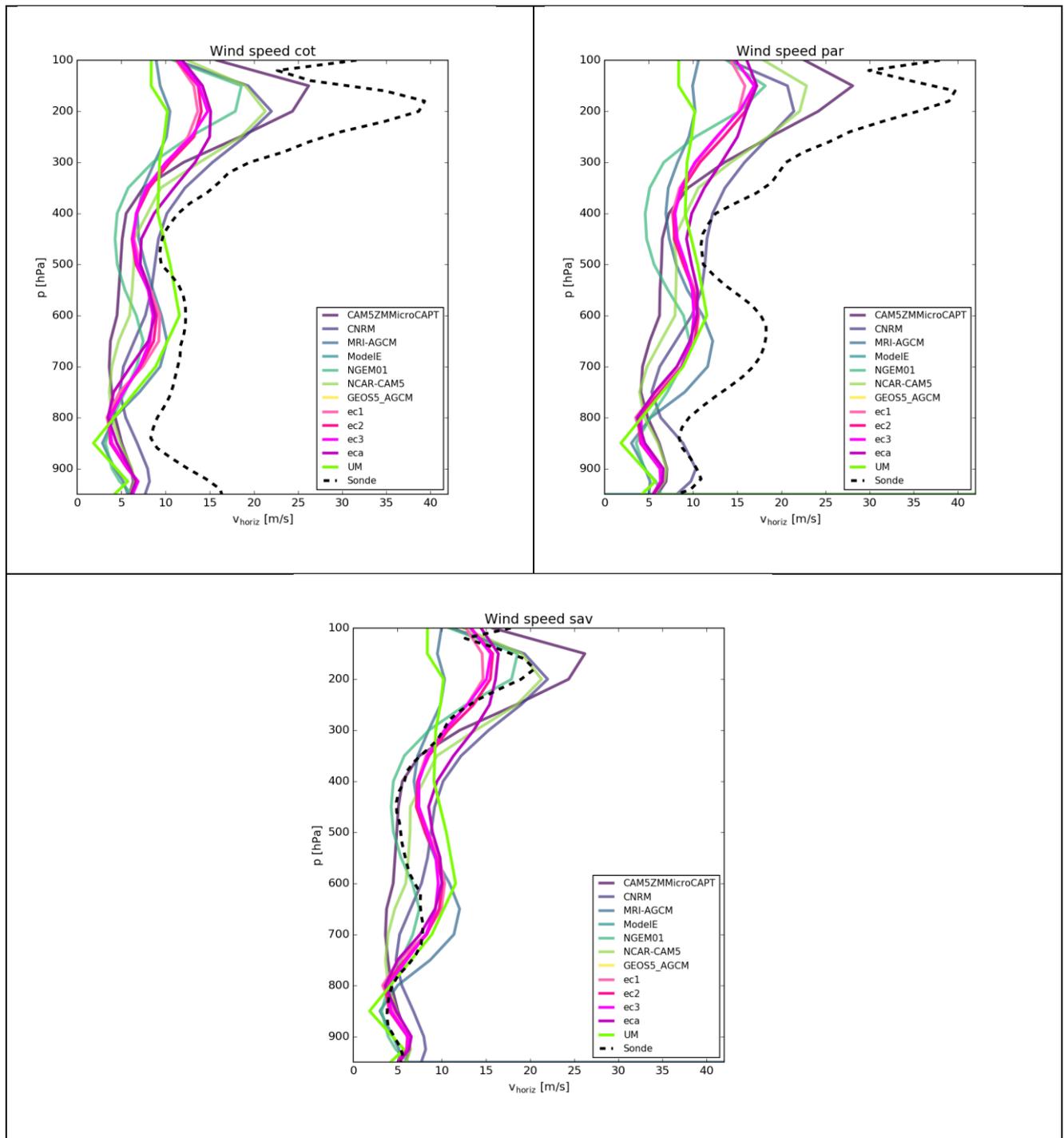


Figure 15: Profiles of horizontal wind speed for July averaged over all years for the YoTC models and our research models compared to the radiosonde measurements from July 2016 at the sites Abidjan, Accra, Cotonou, Lamto, Parakou Kumasi and Savé.

The horizontal wind-speed for July 2016 is displayed in figure 15. The sonde-derived wind speed is considerably higher for some stations than for the models, but the overall shape of the profiles is reproduced. Again, the highest peak of wind speed is placed too high in the models, if there is any peak at all. The sondes do not show a mid-level peak (or at least not a very pronounced one) in horizontal wind speed in contrast to ECHAM. This indicates, that the African Easterly Jet was not located as far to the south as ECHAM produces it (compare Fig. 6). This is corroborated by the north-south gradient in “peakiness” with the northernmost stations Savé and Parakou exhibiting this peak.

5.4 Moisture and precipitation

Tropical West Africa and therefore also our DACCIWA box is influenced by a large inflow of moisture from the Atlantic Ocean. Soil moisture is high in the region resulting in strong evapotranspiration. Dry regions like the Sahel zone north of the DACCIWA box depend strongly on the transport of moisture from the south. Additionally, the dynamics of convective systems are influenced by water vapour through the release of latent heat from condensation or evaporation. In this way, water vapour does not only transport precipitable water but also energy. Therefore, it is of interest to compare the models' ability to provide water vapour in the atmosphere with a measured climatology. The means of comparison at this point is the water vapour dataset ATOVS from CMSAF. We consider the months June, July, August and September for the years 2006–2012 (Fig. 16).

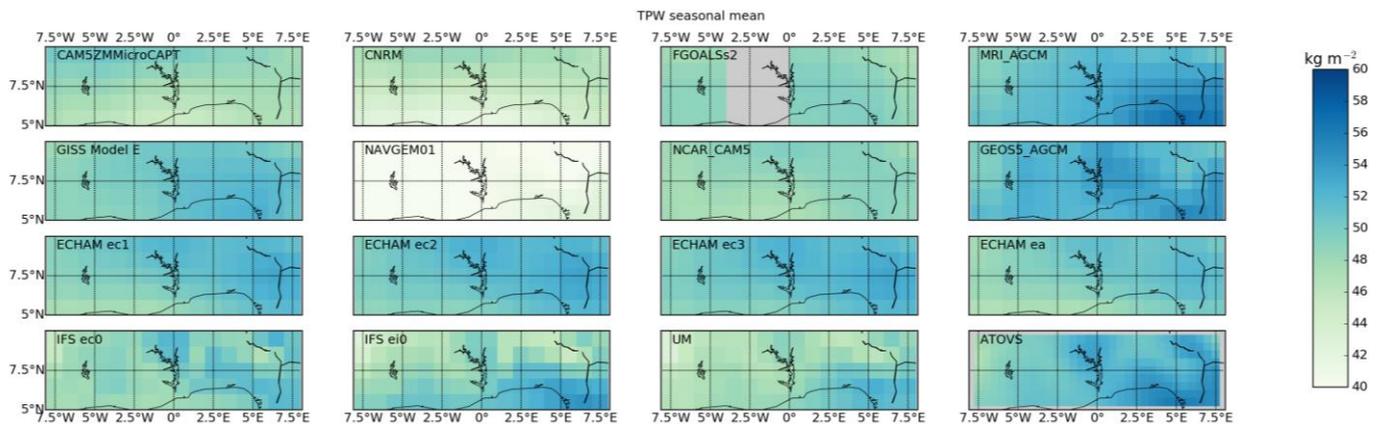


Figure 16: Seasonal mean of total precipitable water from the YoTC models, the ECHAM and IFS runs and the UM compared to ATOVS from the years 2006–2012. TPW allows a comparison of the absolute water content in the models' atmosphere which complements the considerations of the relative humidity.

All models and the ATOVS climatology show total precipitable water in the range of 46 to 60 kg/m². Only NAVGEM01 produces similar horizontal patterns, but the overall amount is much lower with 40–44 kg/m². In ATOVS, high water vapour fields are concentrated to the southeast but also north of lake Volta and around the Niger enhanced values can be found. Altogether, they form a north-west to south-east gradient that is reproduced by the models GISS ModelE and MRI-AGCM from the YoTC group and by all of our research models. In particular IFS and UM exhibit the enhanced values north of lake Volta. The ECHAM runs ec1–ec3 have on average more water vapour, than the eca run. So for ECHAM, an additional input of aerosol and different microphysics lead to changes in dynamics so that the atmosphere can store more water. ECA has lower temperatures which lead to higher cloud formation (at constant specific moisture) and precipitation. This again will lead to a removal of water vapour. After a while, the model atmosphere could adjust to a dryer regime compared to ec1–ec3. The exact mechanisms that produce this feature have to be analysed in further studies.

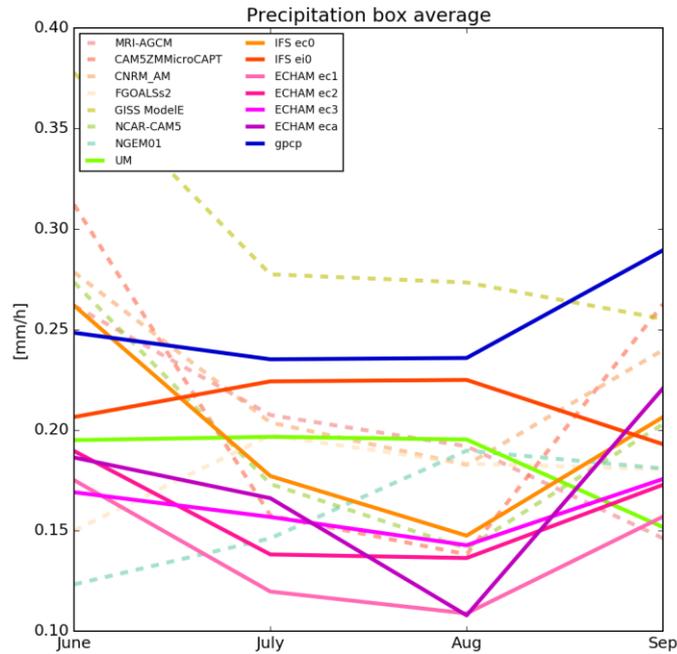


Figure 17: Box average of precipitation rate for all months, YoTC models (dashed lines), ECHAM runs (solid pinkish lines), IFS runs (solid orange shades) and UM (solid green) compared to GPCP from the years 2006–2015.

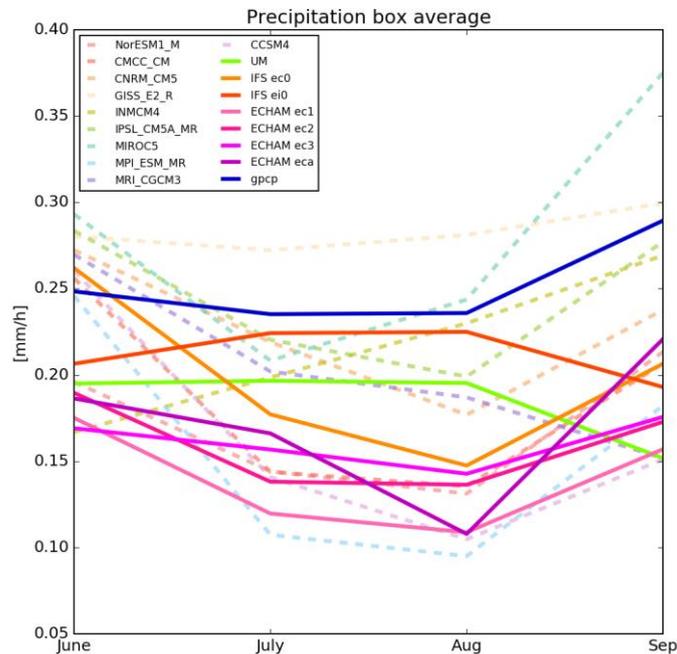


Figure 18: Box average of precipitation rate for all months, CMIP5 models (dashed lines), ECHAM runs (solid pinkish lines), IFS runs (solid orange shades) and UM (solid green) compared to GPCP from the years 2006–2015.

The seasonal cycle of precipitation averaged over the DACCIWA box, as displayed in Fig. 17 varies immensely from model to model. Our source of comparison is GPCP, which takes all available sources of data into account, either satellite-derived or ground-based measured. Its seasonal cycle starts with 0.25 mm/h precipitation in June, decreases to 0.24 in July and further to 0.235 in August to a final increase up to 0.285 in September. This behaviour reflects the northward shift of the rain band, which moves halfway out of the DACCIWA box during the monsoon season. Most YoTC

models are capable of reproducing this pattern, but the overall amounts are too high in June and too low in all other months. Only GISS ModelE has too much rain in all months. The same method of processing was applied to CMIP5 models, the intraseasonal cycle can be found in Fig. 18. The results correspond to the findings for the YoTC models. The ECHAM model versions show the lowest rain rates from our research model group, but the seasonal cycle is well reproduced. The UM features the opposite cycle, with low rains in June and September and enhanced rain rates in July and August. The most interesting feature in this figure are the IFS runs. The control run ec0 follows the GPCP diurnal cycle but the run ei0 with prognostic aerosol has a reversed course with high rain rates in July and August and lower rates in June and September. For ei0 more water vapour was available than for ec0 and the radiation budget was more positive with a less steep vertical temperature gradient. Unfortunately, not more variables were available from IFS for this study in the form of monthly means which could have helped in the discussion.

For a better comparison, the rain rate frequency distributions are shown in Fig. 19. The GPCP distribution is slightly negatively skewed with its mode in the 2.75 mm/h bin. The distribution is clipped at the sides and more narrow than other datasets that merge satellite and gauge data (Pendergrass and Deser, 2017). For CMIP5, the distributions largely resemble those of our research models. They show rain rates distributed between 0.75 mm/h and 3.75 mm/h, most of them are positively skewed and some negatively. The ECHAM versions are ordinarily distributed compared to the other models and GPCP, but the ECHAM run eca on the contrary shows a very broad distribution with many high rainfall events. IFS control and aerosol run differ in a way that the control run possess a comparably broad distribution where the aerosol run has a narrower distribution with a high frequency of rainfall events concentrated at 2 mm/h. The UM has a distribution that is comparable in width to GPCP but with a lower mode.

6 Conclusions

In this model evaluation, we compared several climate models concerning their ability to reproduce the climate of southern West Africa in the wet monsoon season June–September. We chose our research models based on the results of a testing procedure and compared them firstly to a group of climate models from the YoTC experiment plus a group of CMIP5 models and secondly to measurement data from various sources. The focus was placed on long-term availability of the sources but in case of radiosonde measurements an exception was made because of the uniqueness of the number of radiosonde launches in this area. From our research models, we chose several instances that differed regarding the representation of atmospheric aerosol.

The inter-model variation concerning relative humidity, temperature and wind speed is very large among the YoTC and CMIP 5 models and therefore the development of clouds and precipitation varies strongly, too. Not only the regional scale near-surface dynamics of the development of nocturnal low-level clouds is not necessarily realistically represented in all models as we already showed in Hannak et al. (2017) but also the representation of dynamical features at higher levels of the atmosphere is affected. The TEJ for instance is rather weak in some models and the African Easterly Jet is located too far south or is developed too strong.

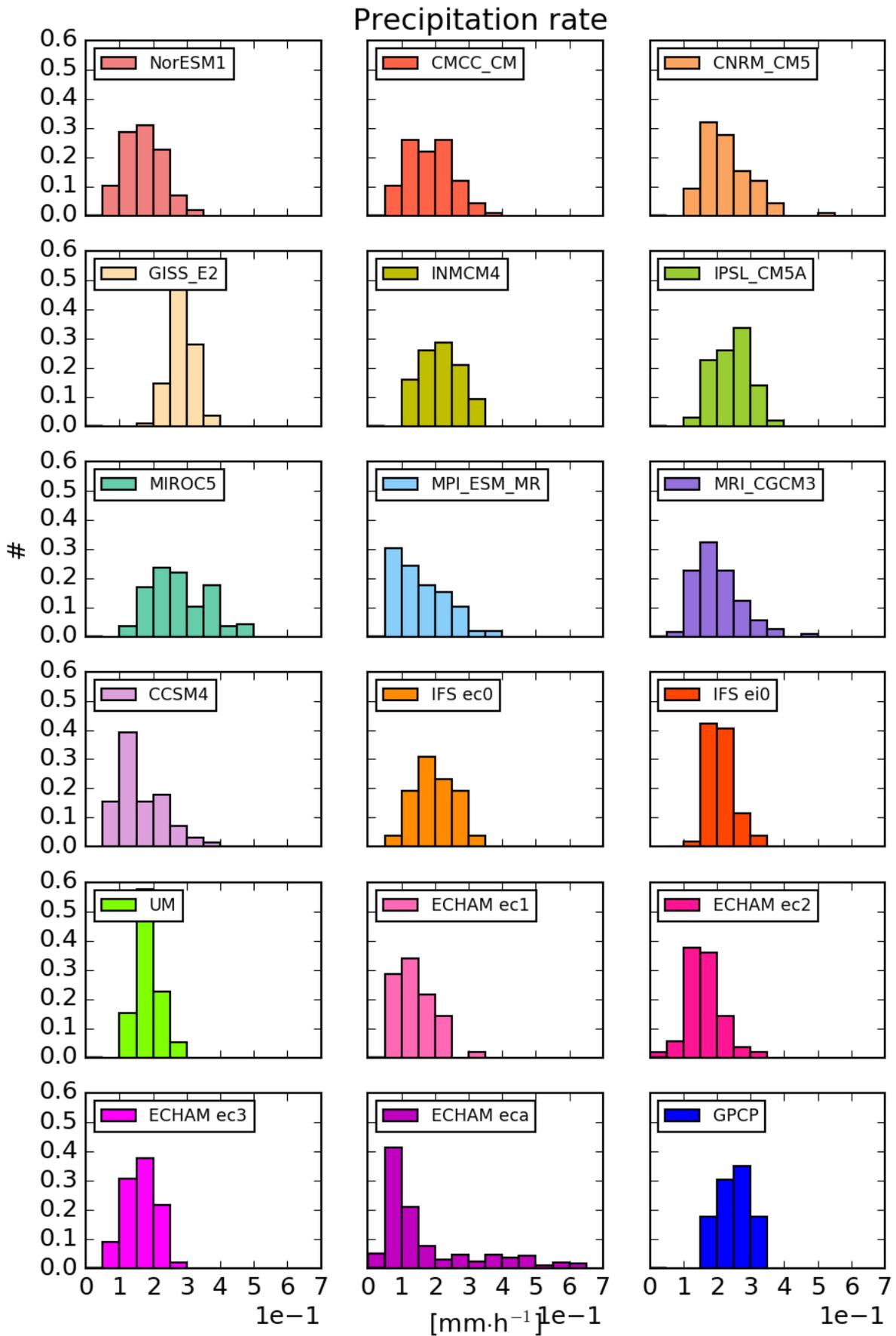


Figure19: CMIP5 models: Frequency distributions of the seasonal precipitation rates made of the box averages for individual months.

The varying aerosol concentrations and parameterization schemes in our research models produce atmospheric average states, that can differ strongly amongst each other. Enhanced aerosol concentration lead to higher temperatures, less cloud cover, reduced precipitation and enhanced water vapour amounts via a complex process chain. This chain will have to be explored in future research on the individual models.

Compared to the radiosonde measurements from the DACCIWA campaign in 2016, the relative humidity deviations at around 200 hPa show, that the tropopause height is placed too high in most models. This feature has to be analysed further, since differences may occur from the very short measurement period compared to climatic time-scales. A source of comparison could be the total precipitable water vapour product in 5 layers or the free tropospheric relative humidity product from CM SAF.

The most striking feature is the top of atmosphere radiation budget. This budget can be considered as the summation of all possible effects on atmospheric dynamics and aerosol physics as well as thermodynamic processes that influence cloud formation. The models do not only show variations amongst each other, but the radiation budgets do have different signs. In some models, the DACCIWA region has a positive radiation budget and is therefore a region with uptake of energy while in other models the region loses energy during the monsoon months. Our research models do not have comparable budgets either. The ECHAM runs have a partly negative budget but positive values towards the northern and southern domain edges. The UM always has a positive budget. The two IFS runs differ so much, that the radiation budget turns from negative to positive for September and partly August. Unfortunately, the GERB dataset used in this study did not provide incoming solar radiation measurements, therefore no conclusive comparison can be made at this point. We will repeat the study with another GERB dataset and possibly CERES measurements.

Most models underestimate precipitation compared to GPCP. Even though the total precipitable water is comparable to ATOVS and some model atmospheres even contain more water than the measurements indicate, this does not simply lead to enhanced precipitation due to the involved thermodynamic processes.

To sum up, the present study delivers a comprehensive overview of the climatic state of the monsoon season in southern West Africa as it is represented in state of the art climate models. Significant discrepancies could be found among the models for several atmospheric variables and some general explanations could be given. However, a more detailed analysis is needed in most cases which is not the scope of this report. Hopefully this overview will give an impetus for further investigations on the monsoon in southern West Africa.

7 References

- Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, 4,1147-1167.
- Arakawa, A. and Lamb, V. R., 1977: Computational design of the basic dynamic processes of the UCLA general circulation model, *Methods in Computational Physics: Advances in Research and Applications*, 17, 173–265.
- Bao, Q., and Coauthors, 2013: The Flexible Global Ocean– Atmosphere–Land System model version: FGOALS-s2, *Adv. Atmos. Sci.*, 30, 561–576, doi:10.1007/s00376-012-2113-9.
- Boucher, O., Pham, M., and Venkataraman, C., 2002: Simulation of the atmospheric sulfur cycle in the LMD GCM: Model description, model evaluation, and global and European budgets, Note 23, 26 pp., Inst. Pierre-Simon Laplace, Paris, France.
- Charney, J. G. and Phillips, N. A., 1953: Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows, *J. Meteorol.*, 10, 71–99, doi:10.1175/1520-0469(1953)010<0071:NIOTQG>2.0.CO;2.
- Courcoux, Nathalie; Schröder, Marc. (2013): Vertically integrated water vapour, humidity and temperature at pressures levels and layers from ATOVS - Daily Means / Monthly Means. Satellite Application Facility on Climate Monitoring. DOI:10.5676/EUM_SAF_CM/WVT_ATOVS/V001, http://dx.doi.org/10.5676/EUM_SAF_CM/WVT_ATOVS/V001.
- DeWitte, S., Gonzalez, L., Clerbaux, N., Ipe, A., Bertrand, C., & De Paepe, B., 2008: The geostationary earth radiation budget edition 1 data processing algorithms, *Adv. in Spac. Res.*, 41(11), 1906-1913.
- Dufresne, J.-L., and Coauthors, 2012: Climate change projections using the IPSL-CM5 Earth system model: From CMIP3 to CMIP5. *Climate Dyn.*, 40, 2123–2165, doi:10.1007/ s00382-012-1636-1.
- Eyring, V., Bony, S., Meehl, G.A., Senior C., Stevens B., Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, European Geosciences Union, 2016, 9 (5), pp.1937-1958. <https://doi.org/10.5194/gmd-9-1937-2016>.
- Flamant, C., P. Knippertz, A.H. Fink, A. Akpo, B. Brooks, C.J. Chiu, H. Coe, S. Danuor, M. Evans, O. Jegede, N. Kalthoff, A. Konaré, C. Lioussé, F. Lohou, C. Mari, H. Schlager, A. Schwarzenboeck, B. Adler, L. Amekudzi, J. Aryee, M. Ayoola, A.M. Batenburg, G. Bessardon, S. Borrmann, J. Brito, K. Bower, F. Burnet, V. Catoire, A. Colomb, C. Denjean, K. Fosu-Amankwah, P.G. Hill, J. Lee, M. Lothon, M. Maranan, J. Marsham, R. Meynadier, J. Ngamini, P. Rosenberg, D. Sauer, V. Smith, G. Stratmann, J.W. Taylor, C. Voigt, and V. Yoboué, 0: The Dynamics-Aerosol-Chemistry-Cloud Interactions in West Africa field campaign: Overview and research highlights, *Bull. Amer. Meteor. Soc.*, 0, <https://doi.org/10.1175/BAMS-D-16-0256.1>.
- Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, 24, 4973–4991, doi:10.1175/ 2011JCLI4083.1.
- Hannak, L.; Knippertz, P.; Fink, A. H.; Kniffka, A.; Pante, G., 2017: Why do global climate models struggle to represent low-level clouds in the West African summer monsoon? *J. Climate*, 30(5), 1665–1687. [doi:10.1175/JCLI-D-16-0451.1](https://doi.org/10.1175/JCLI-D-16-0451.1).
- Harries, J. E., J. E. Russell, J.A. Hanafin, H. Brindley, J. Fytan, J. Rufus, S. Kellock, G. Matthews, R. Wrigley, A. Last, J. Mueller, R. Mossavati, J. Ashmall, E. Sawyer, D. Parker, M. Caldwell, P. M. Allan, A. Smith, M. J. Bates, B. Coan, B. C. Stewart, D. R. Lepine, L. A. Cornwall, D. R. Corney, M. J.Ricketts, D. Drummond, D. Smart, R. Cutler, S. Dewitte, N. Clerbaux, L. Gonzalez, A. Ipe, C. Bertrand, A. Joukoff, D. Crommelynck, N. Nelms, D. T. Llewellyn-Johnes, G. Butcher, G. L. Smith, Z. P.Szewczyk, P. E. Mlynczak, A. Slingo, R. P. Allan and M. A. Ringer, 2005: The Geostationary Earth Radiation Budget Project, *Bulletin of the American Meteorological Society*, Volume 86, No. 7, pp 945-960.

Heil, A., Kaiser, J. W., van der Werf, G. R., Wooster, M. J., Schultz, M. G., and Dernier van der Gon, H., 2010: Assessment of the Real-Time Fire Emissions (GFASv0) by MACC, Tech. Memo, 628, ECMWF, Reading, UK.

Hill, P. G., R. P. Allan, J. C. Chiu, and T. H. M. Stein, 2016: A multisatellite climatology of clouds, radiation, and precipitation in southern West Africa and comparison to climate models, *J. Geophys. Res. Atmos.*, 121, 10,857–10,879, doi:10.1002/2016JD025246.

Hogan, T. F., and Coauthors, 2014: The Navy Global Environmental Model, *Oceanography*, 27, 116–125, doi:https://doi.org/10.5670/oceanog.2014.73.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.*, 113, D13103, doi:10.1029/2008JD0099443.

Jiang, X., and Coauthors, 2015: Vertical structure and physical processes of the Madden-Julian oscillation: Exploring key model physics in climate simulations, *J. Geophys. Res. Atmos.*, 120, 4718–4748, doi:https://doi.org/10.1002/2014JD022375.

Jones, A., Roberts, D. L., and Slingo, A., 1994: A climate model study of indirect radiative forcing by anthropogenic sulphate aerosols, *Nature*, 370, 450–453, doi:10.1038/370450a0.

Jones, A., Roberts, D. L., Woodage, M. J., and Johnson, C. E., 2001: Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle, *J. Geophys. Res.*, 106, 20293–20310, doi:10.1029/2000JD000089.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R., 2012: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, *Biogeosciences*, 9, 527–554, https://doi.org/10.5194/bg-9-527-2012.

Knippertz, P.; Fink, A. H.; Schuster, R.; Trentmann, J.; Schrage, J. M.; Yorke, C., 2011: Ultra-low clouds over the southern West African monsoon region, *Geophys. Res. Lett.*, 38, L21808, doi:10.1029/2011GL049278.

Lamarque, J.-F., T. C. Bond, V. Eyre, C. Granier, A. Heil, Z. Klimont, D. Lee, C. Liousse, A. Mieville, B. Owen, M. G. Schultz, D. Shindell, S. J. Smith, E. Stehfest, J. Van Aardenne, O. R. Cooper, M. Kainuma, N. Mahowald, J. R. McConnell, V. Naik, K. Riahi, and D. P. van Vuuren, 2010: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, *Atmos. Chem. Phys.*, 10, 7017–7039, doi:10.5194/acp-10-7017-2010.

Lin, S. J. and R. B. Rood, 1996: Multidimensional flux-form semi-Lagrangian transport schemes, *Mon. Weather Rev.*, 124, 2046–2070.

Lohmann, U. and E. Roeckner, 1996: Design and performance of a new cloud microphysics scheme developed for the ECHAM general circulation model, *Climate Dyn.*, 12, 557–572.

Lohmann, U., Stier, P., Hoose, C., Ferrachat, S., Kloster, S., Roeckner, E., and Zhang, J., 2007: Cloud microphysics and aerosol indirect effects in the global climate model ECHAM5-HAM, *Atmos. Chem. Phys.*, 7, 3425–3446, doi:10.5194/acp-7-3425-2007.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, *Geosci. Model Dev.*, 3, 519–551, doi:10.5194/gmd-3-519-2010, 2010.

Meirink, J. F., Roebeling, R. A., and Stammes, P., 2013: Inter-calibration of polar imager solar channels using SEVIRI, *Atmos. Meas. Tech.*, 6, 2495–2508, https://doi.org/10.5194/amt-6-2495-2013.

Molod, A., L. Takacs, L. M. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann, 2012: The GEOS-5 atmospheric general circulation model: Mean climate and development from MERRA to Fortuna, NASA Tech. Rep. NASA TM-2012-104606, Vol. 28, 117 pp.

Morcrette, J.-J., Boucher, O., Jones, L., Salmond, D., Bechtold, B., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A., 2009: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, *J. Geophys. Res.*, 114, D06206, doi:10.1029/2008JD011235.

Neale, R. B., and Coauthors, 2012: Description of the NCAR Community Atmosphere Model: CAM 5.0, NCAR Tech. Rep. NCAR/TN-4861STR, 274 pp.

Nicholson, S.E. and J.P. Grist, 2003: The Seasonal Evolution of the Atmospheric Circulation over West Africa and Equatorial Africa, *J. Climate*, 16, 1013–1030, [https://doi.org/10.1175/1520-0442\(2003\)016<1013:TSEOTA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<1013:TSEOTA>2.0.CO;2).

Parker, D.J., A. Fink, S. Janicot, J. Ngamini, M. Douglas, E. Afiesimama, A. Agusti-Panareda, A. Beljaars, F. Dide, A. Diedhiou, T. Lebel, J. Polcher, J. Redelsperger, C. Thorncroft, and G. Ato Wilson, 2008: The Amma Radiosonde Program and its Implications for the Future of Atmospheric Monitoring Over Africa, *Bull. Amer. Meteor. Soc.*, 89, 1015–1027, <https://doi.org/10.1175/2008BAMS2436.1>.

Pendergrass, A.G. and C. Deser, 2017: Climatological Characteristics of Typical Daily Precipitation, *J. Climate*, 30, 5985–6003, <https://doi.org/10.1175/JCLI-D-16-0684.1>.

Reddy, M. S., Boucher, O., Bellouin, N., Schulz, M., Balkanski, Y., Dufresne J.-L., and Pham, M, 2005: Estimates of global multicomponent aerosol optical depth and direct radiative perturbation in the Laboratoire de Meteorologie Dynamique general circulation model, *J. Geophys. Res.*, 110, D10S16, doi:10.1029/2004JD004757.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate, *J. Climate*, 15, 1609–1625, doi:[https://doi.org/10.1175/1520-0442\(2002\)015<1609:AISAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1609:AISAS>2.0.CO;2).

Roeckner, E., G. Baeuml, L. Bonventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, U. Rhodin, U. Schlese, U. Schulzweida, A. Tomkins, 2003: The atmospheric general circulation model ECHAM5. Part I: Model description, Report 349, Max Planck Institute for Meteorology, Hamburg, Germany.

Romps, D.M., 2014: An Analytical Model for Tropical Relative Humidity, *J. Climate*, 27, 7432–7449, <https://doi.org/10.1175/JCLI-D-14-00255.1>.

Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, *J. Adv. Model. Earth Syst.*, 6, 141–184, doi:10.1002/2013MS000265.

Scoccimarro, E., and Coauthors, 2011: Effects of tropical cyclones on ocean heat transport in a high resolution coupled general model, *J. Climate*, 24, 4368–4384, doi:10.1175/2011JCLI4104.1.

Song, X., and G. J. Zhang, 2011: Microphysics parameterization for convective clouds in a global climate model: Description and single-column model tests, *J. Geophys. Res.*, 116, D02201, doi:10.1029/2010JD014833.

Stengel, M., Kniffka, A., Meirink, J. F., Lockhoff, M., Tan, J., and Hollmann, R., 2014: CLAAS: the CM SAF cloud property data set using SEVIRI, *Atmos. Chem. Phys.*, 14, 4297–4311, <https://doi.org/10.5194/acp-14-4297-2014>.

Stevens, B., M. Giorgetta, M. Esch, T. Mauritsen, T. Crueger, S. Rast, M. Salzmann, H. Schmidt, J. Bader, K. Block, R. Brokopf, I. Fast, S. Kinne, L. Kornblueh, U. Lohmann, R. Pincus, T. Reichler, and E. Roeckner, 2013: Atmospheric component of the MPI-M Earth System Model: ECHAM6, *James*, 5, 146–172, doi:10.1002/jame.20015.

Stephens, G. L., T. L'Ecuyer, 2015: The Earth's energy balance, *Atmos. Res.*, 166, 195–203.

Stier, P., J. Feichter, S. Kinne, S. Kloster, E. Vignati, J. Wilson, L. Ganzeveld, I. Tegen, M. Werner, Y. Balkanski, M. Schulz, O. Boucher, A. Minikin, and A. Petzold, 2005, The aerosol-climate model ECHAM5-HAM, *Atmos. Chem. Phys.*, 5, 1125-1156.

Sundqvist, H., E. Berge, and J. E. Kristjansson, 1989, Condensation and Cloud Parameterization studies with a mesoscale numerical weather prediction model, *Mon. Weather Rev.*, 117, 1641-1657.

Taylor, K., R. J. Stouffer, and G. A. Meehl, 2011: A Summary of the CMIP5 Experiment Design, http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf.

Taylor, K.E., R.J. Stouffer, and G.A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.*, 93, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.

Tegen, I., Hoorig, P., Chin, M., Fung, I., Jacob, D., and Penner, J., 1997: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results, *J. Geophys. Res.*, 102, 23895–23915.

Thorncroft, C.D., D.J. Parker, R.R. Burton, M. Diop, J.H. Ayers, H. Barjat, S. Devereau, A. Diongue, R. Dumelow, D.R. Kindred, N.M. Price, M. Saloum, C.M. Taylor, and A.M. Tompkins, 2003: The JET2000 Project: Aircraft Observations of the African Easterly Jet and African Easterly Waves, *Bull. Amer. Meteor. Soc.*, 84, 337–351, <https://doi.org/10.1175/BAMS-84-3-337>.

Vignati, E., J. Wilson, and P. Stier, 2004, An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models, *J. Geophys. Res.*, 109, D22202, doi:10.1029/2003JD004485.

Vitart F. 2014: Evolution of ECMWF sub-seasonal forecast skill scores, *Quart. J. Roy. Meteor. Soc.*, 140, 1889-1899.

Voltaire, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation, *Climate Dyn.*, 40, 2091–2121, doi:10.1007/s00382-011-1259-y.

Volodin, E. M., N. A. Dianskii, and A. V. Gusev, 2010: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, *Izv. Atmos. Ocean. Phys.*, 46, 414–431, doi:10.1134/S000143381004002X.

Waliser, D. E.; Moncrieff, M.; Burridge, D.; Fink, A.; Gochis, D.; Goswami, B. N.; Guan, B.; Harr, P.; Heming, J.; Hsu, H.-H.; Jakob, C.; Janiga, M.; Johnson, R.; Jones, S.; Knippertz, P.; Marengo, J.; Nguyen, H.; Pope, M.; Serra, Y.; Thorncroft, C.; Wheeler, M.; Wood, R.; Yuter, S., 2012: The "Year" of Tropical Convection (May 2008 to April 2010): Climate variability and weather highlights, *Bull. Amer. Meteor. Soc.*, 93(8), 1189–1218. [doi:10.1175/2011BAMS3095.1](https://doi.org/10.1175/2011BAMS3095.1).

Walters, D., I. Boutle, M. Brooks, T. Melvin, R. Stratton, S. Vosper, H. Wells, K. Williams, N. Wood, T. Allen, A. Bushell, D. Copley, P. Earnshaw, J. Edwards, M. Gross, S. Hardiman, C. Harris, J. Heming, N. Klingaman, R. Levine, J. Manners, G. Martin, S. Milton, M. Mittermaier, C. Morcrette, T. Riddick, M. Roberts, C. Sanchez, P. Selwood, A. Stirling, C. Smith, D. Suri, W. Tennant, P. L. Vidale, J. Wilkinson, M. Willett, S. Woolnough, and P. Xavier, 2017: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, *Geosci. Model Dev.*, 10, 1487-1520, <https://doi.org/10.5194/gmd-10-1487-2017>.

Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity, *J. Climate*, 23, 6312–6335, doi:10.1175/2010JCLI3679.1.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M., and Thuburn, J., 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations, *Q. J. Roy. Meteorol. Soc.*, 140, 1505–1520, doi:10.1002/qj.2235.

Yukimoto, S., and Coauthors, 2012: A new global climate model of the Meteorological Research Institute: MRI-CGCM3-model description and basic performance, *J. Meteor. Soc. Japan*, 90A, 23–64, doi:10.2151/jmsj.2012-A02.

Zhang, K., D. O'Donnell, J. Kazil, P. Stier, S. Kinne, U. Lohmann, S. Ferrachat, B. Croft, J. Quaas, H. Wan, S. Rast, and J. Feichter, 2012: The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations, *Atmos. Chem. Phys.*, 12, 8911-8949, doi:10.5194/acp-12-8911-2012.

8 Appendix

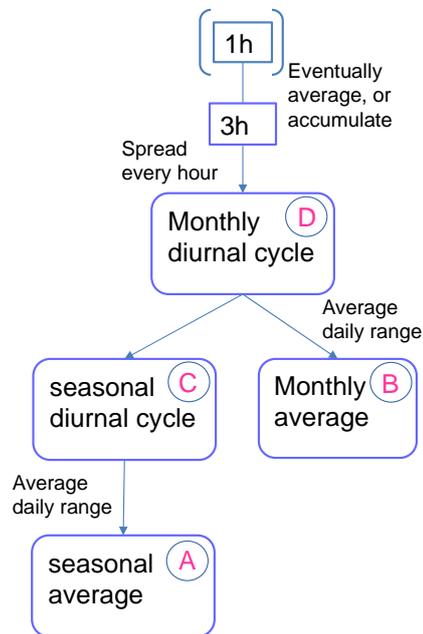
The processing flow charts, excerpt from the processing guideline:

Average modes									
Spatial					Temporal				
	Dimension	Type	approx. # of values		Type	Month	Time span	approx. # of values	
1.)	0D	Box average	2D surface fields	1	A	Seasonal	(JAS)	2006-2015	1
2.)	1D	Box vertical profiles	3D fields	nz=20	B	Monthly	J,J,A,S	2006-2015	nm=4
3.)	2D	Field horizontal	2D surface height averaged	nlon*nlat=10 0	C	Seasonal Diurnal cycle	(JAS)	2006-2015	nh=8
4.)	2D	Lat-height lon-lat profiles	3D fields	nlat*nz=100	D	Monthly diurnal cycle	J,J,A,S	2006-2015	nh*4=32
5.)	1D	stations	3D fields	5					

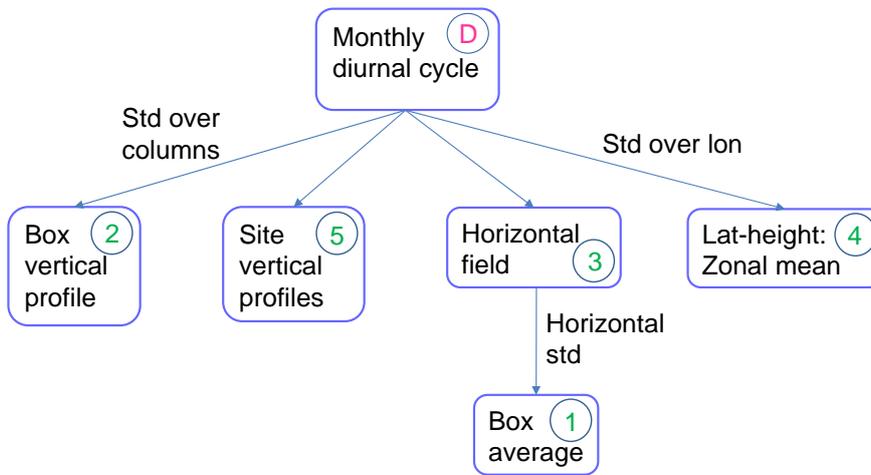
Model evaluation

Temporal

Data flow post processing



Spatial



More figures:

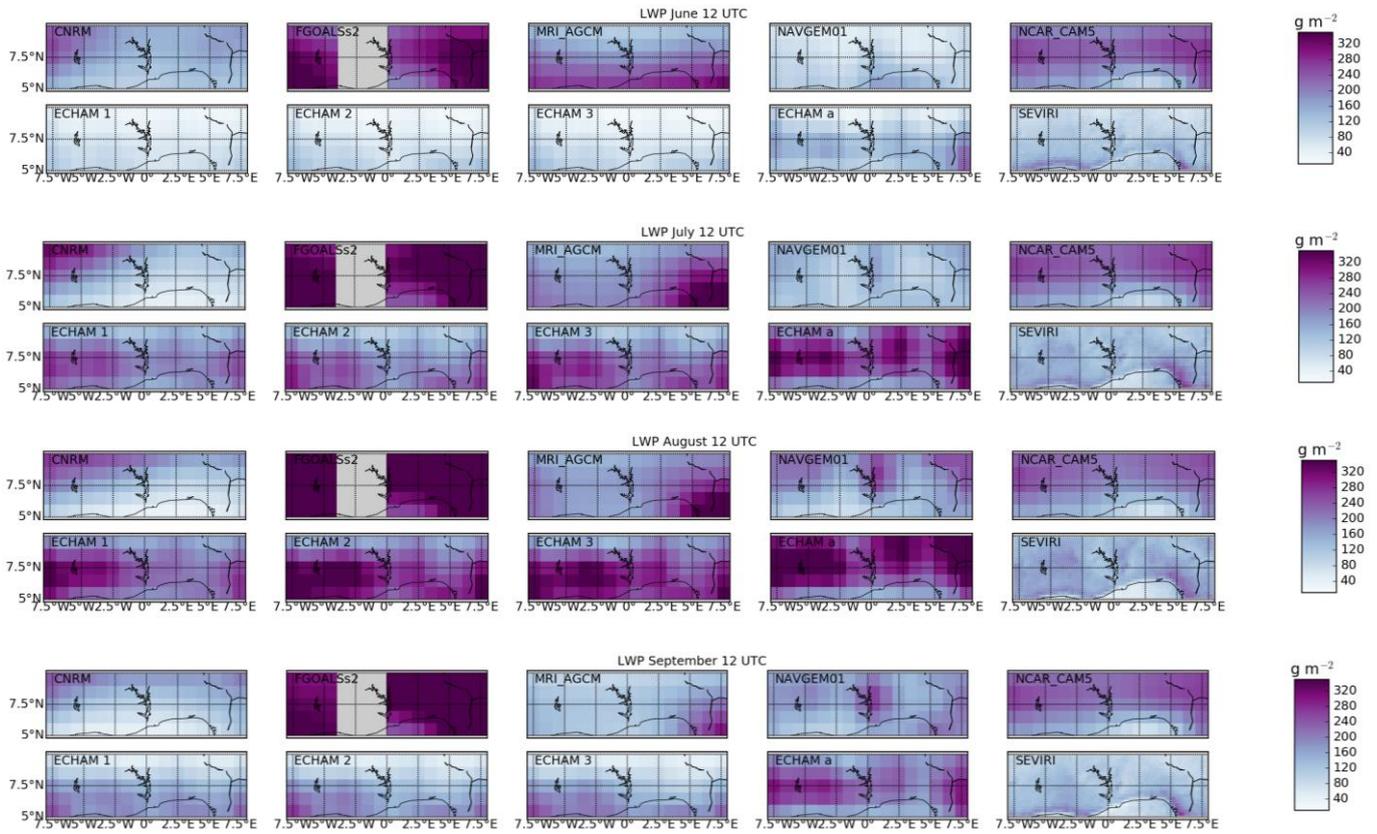
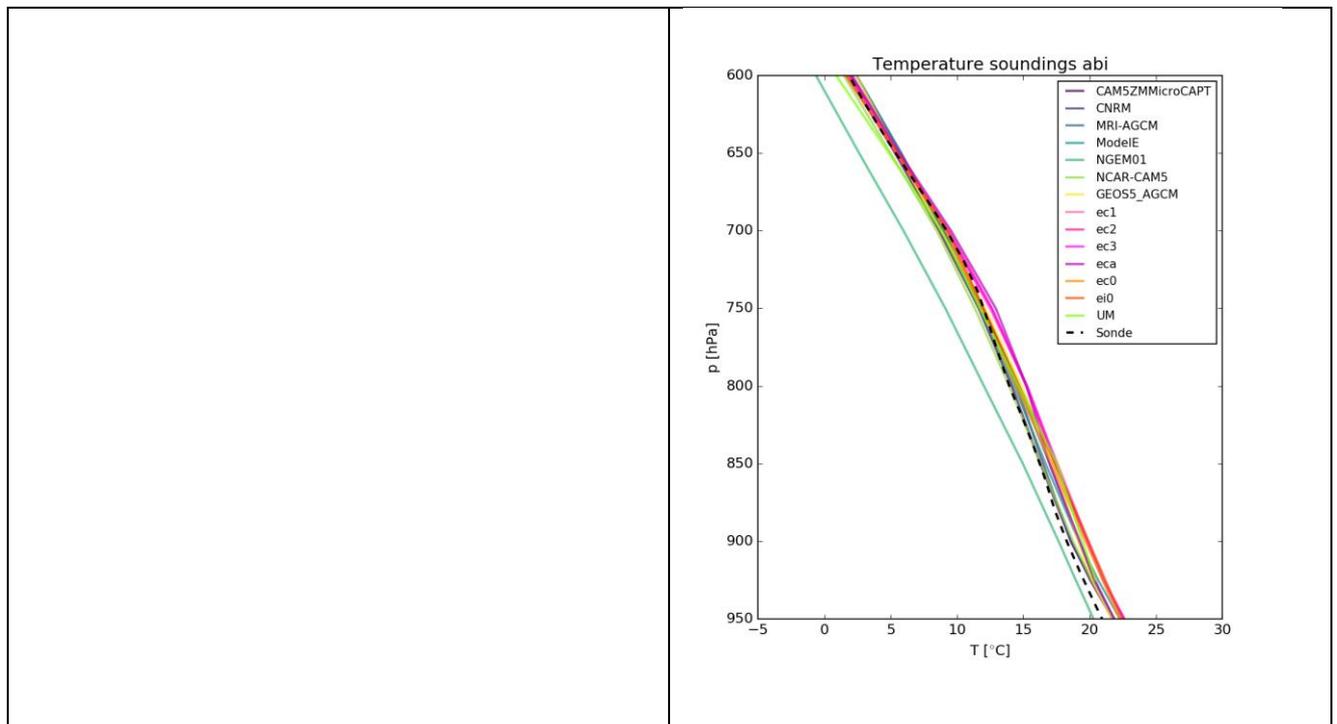


Figure 19: Liquid water path monthly means for June – September compared to SEVIRI, averages were produced from the 12 UTC fields.



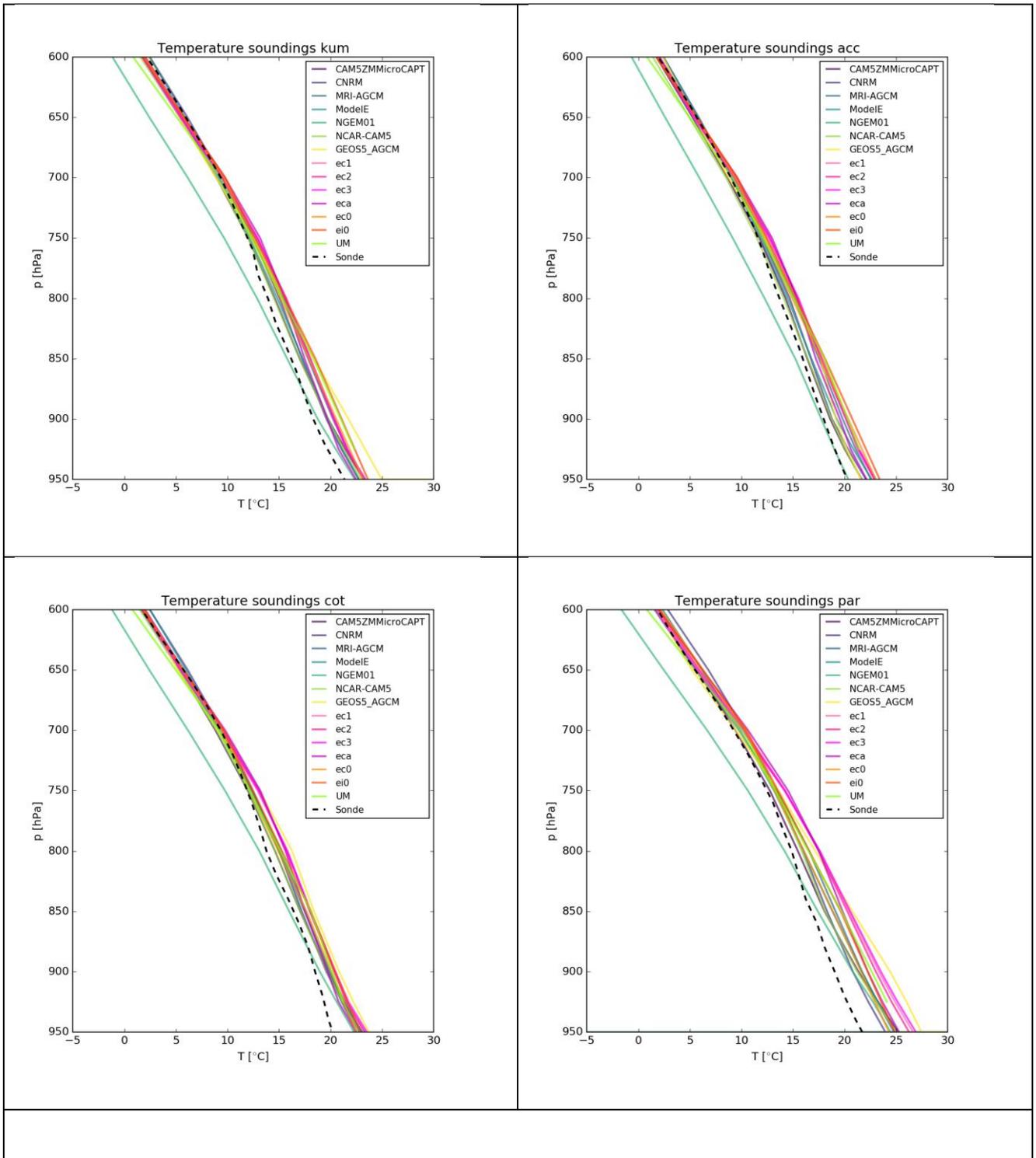
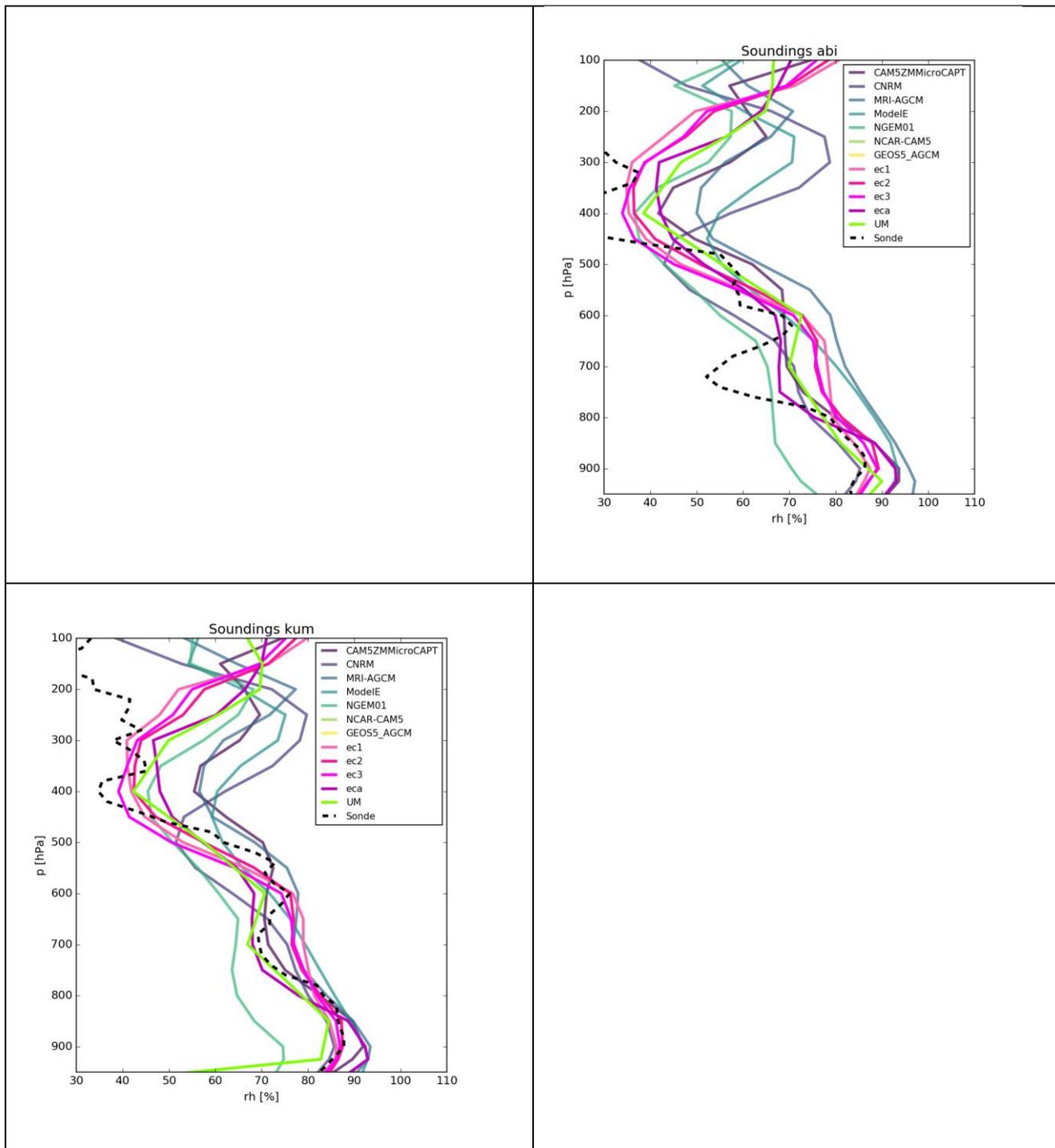


Figure 20: Profiles of temperature for June averaged over all years for the YoTC models and our research models compared to the radiosonde measurements from June 2016 at the sites Abidjan, Accra, Cotonou, Lamto, Parakou Kumasi and Savé.



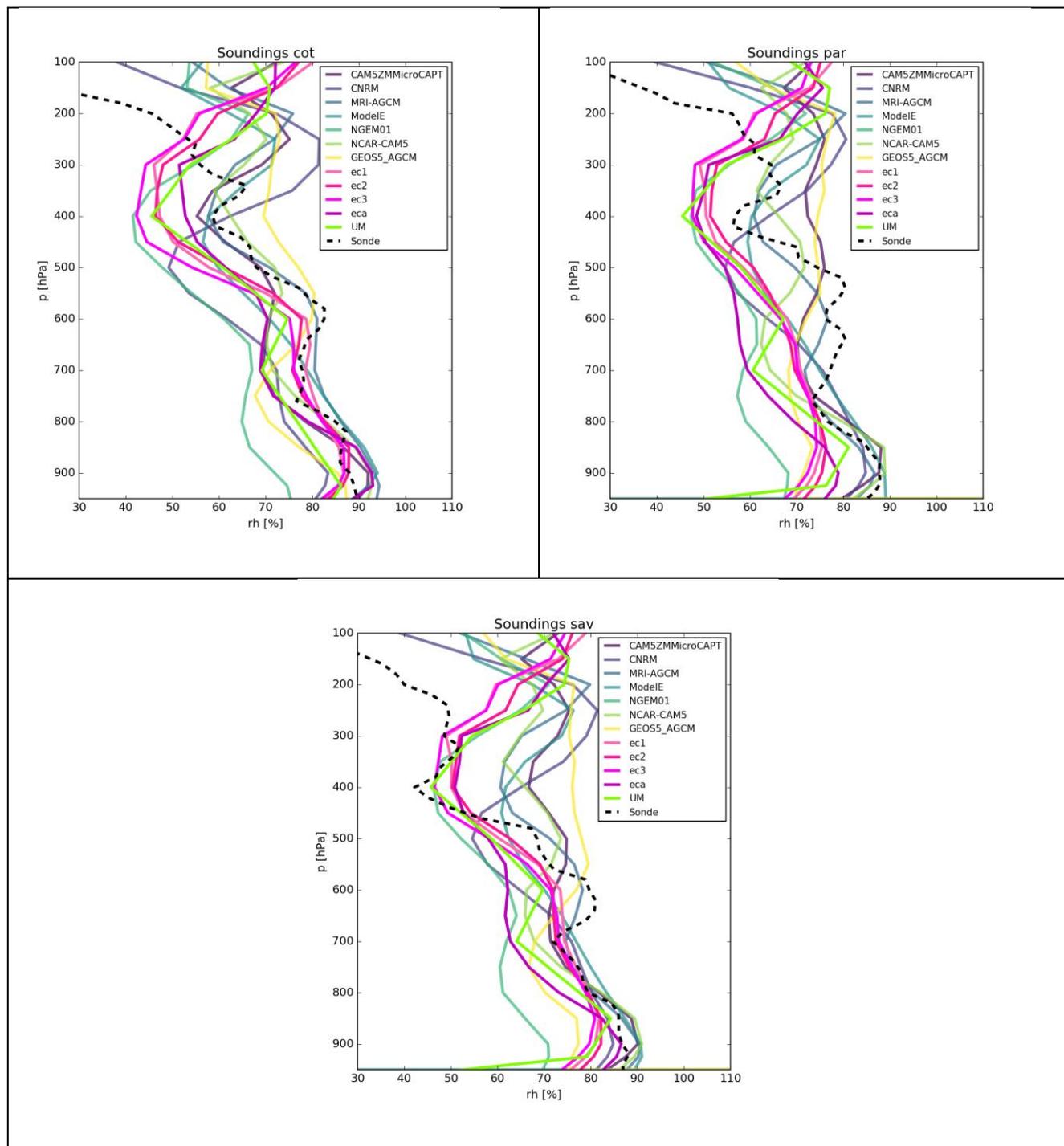
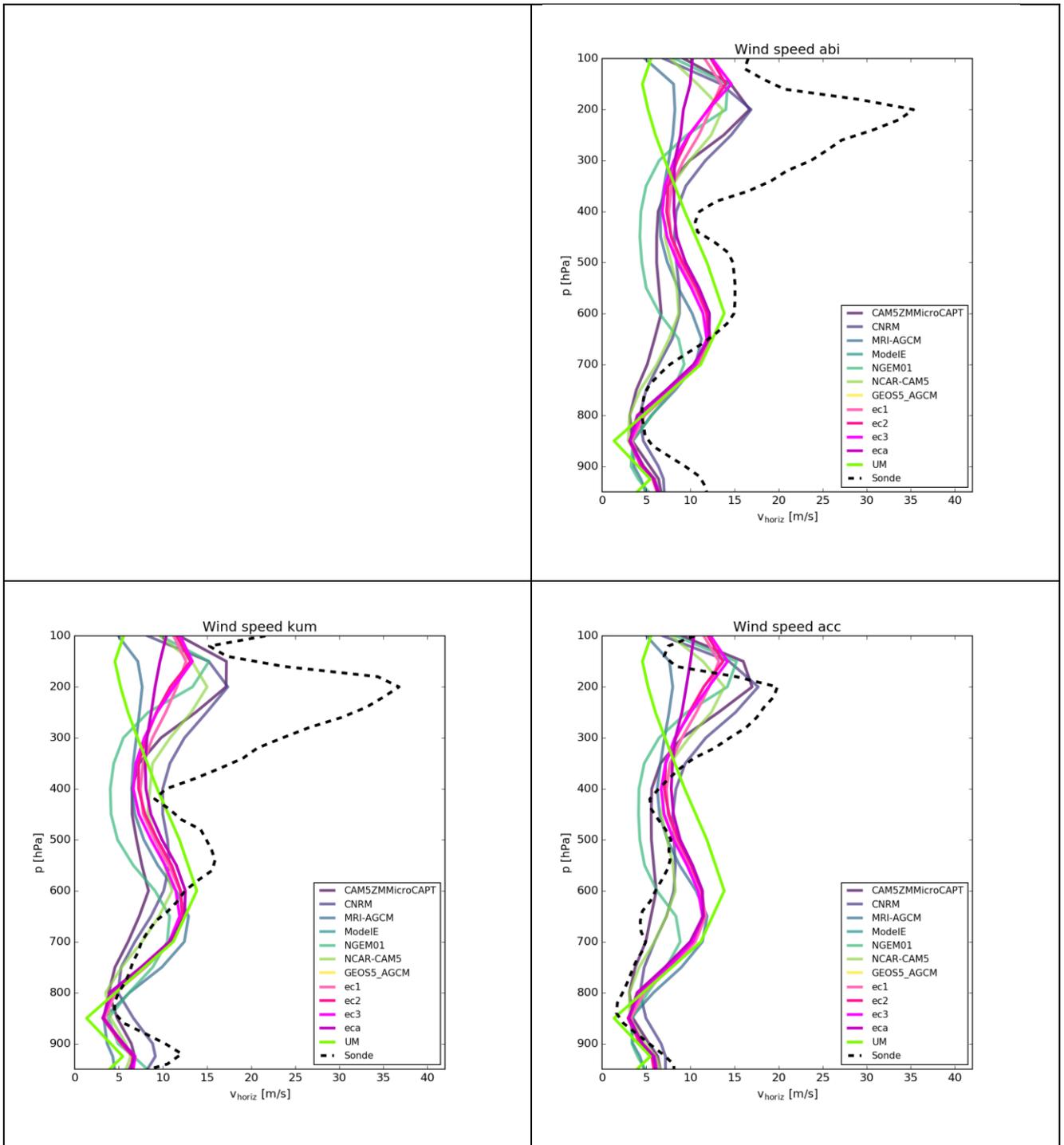


Figure 21: Profiles of relative humidity for June averaged over all years for the YoTC models and our research models compared to the radiosonde measurements from June 2016 at the sites Abidjan, Accra, Cotonou, Lamto, Parakou Kumasi and Savé.



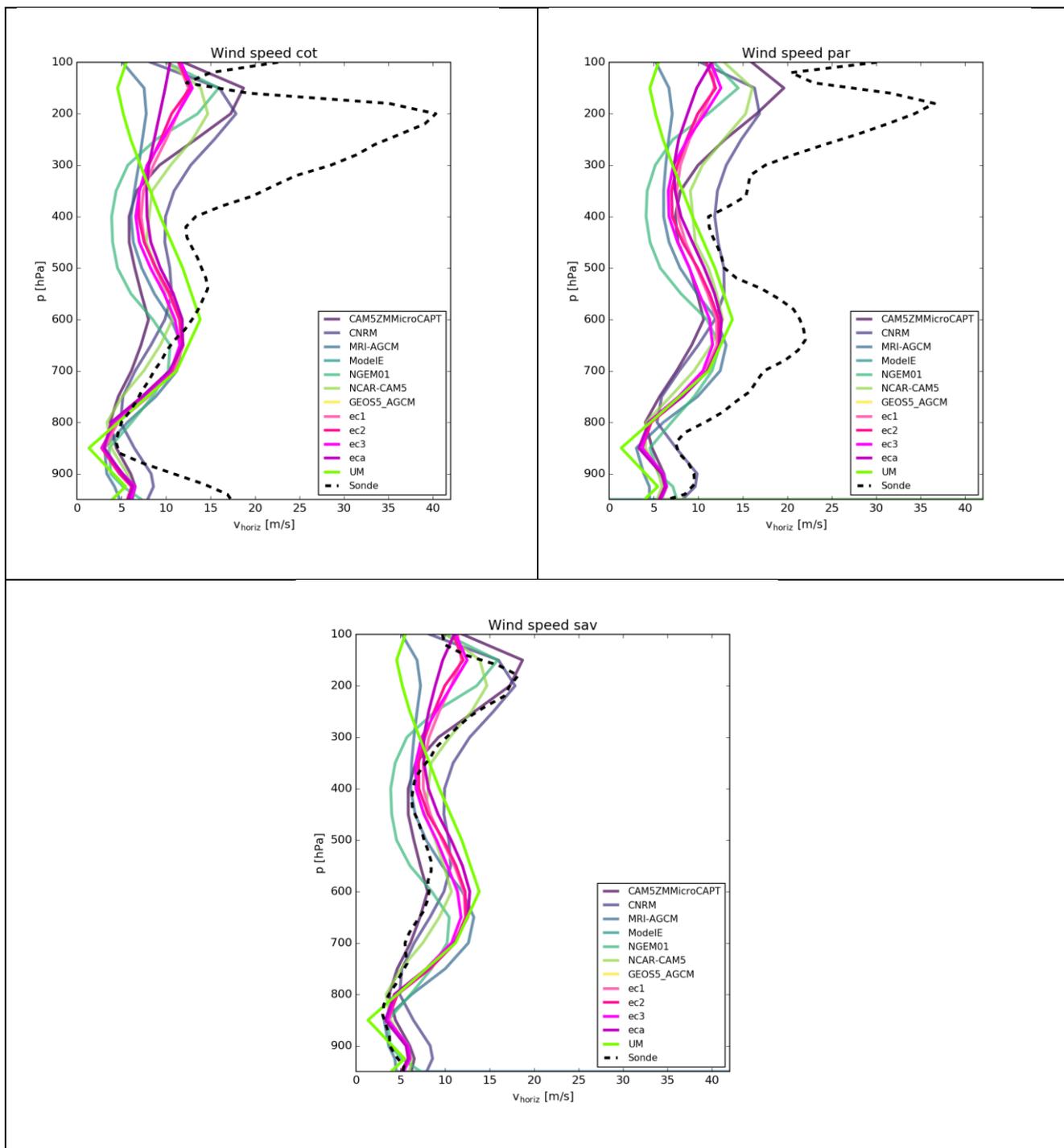


Figure 22: Profiles of horizontal wind speed for June averaged over all years for the YoTC models and our research models compared to the radiosonde measurements from June 2016 at the sites Abidjan, Accra, Cotonou, Lamto, Parakou Kumasi and Savé.