ORIGINAL PAPER

CrossMark

Customized framework of the WRF model for regional climate simulation over the Eastern NILE basin

Mohamed Abdelwares[1] · Mohammed Haggag[1] · Ahmad Wagdy[1] · Jos Lelieveld[2,3]

Abstract

Different configurations of the Weather and Research Forecasting (WRF-ARW) regional climate model, centered over the Eastern Nile Basin, have been investigated. Extensive sensitivity analyses were carried out to test the model performance in simulating precipitation and surface air temperature, focusing on the horizontal extent of the simulation domain, the mesh size and the parameterizations of the boundary layer, radiation, cloud microphysics, and convection. A simulation period of 2 years (1998–1999) was used to assess the model performance during the rainy season (June–September) and the dry season (December–March). Three sets of numerical experiments were conducted. The first tested the effects of changing the horizontal extent of the simulation domain; three domains have been examined to investigate, e.g., the effect of including a larger part of the Indian Ocean, for which no significant impact was found. The second set of experiments tested the sensitivity of WRF to the horizontal mesh size (about 16, 12, and 10 km). It was found that increased resolution results in a more accurate simulation of precipitation and surface temperature. The third set of experiments was designed to select the optimal combination of physics parameterizations. All simulations were forced by ERA-Interim reanalysis data to provide initial and boundary conditions, including sea surface temperature, and the Noah land surface model (NPAH) was used to simulate land surface processes. To rate the model performance, we used a range of statistical metrics, summarized with a scoring technique to obtain a single index that ranks different alternatives. The simulated precipitation was found to be much more sensitive to the choice of physics parameterization compared to the surface air temperature. Precipitation was most sensitive to changing the cumulus and the planetary boundary layer schemes, and least sensitive to changing the microphysics scheme. Modifying the long-wave radiation scheme led to more significant changes compared to the short-wave radiation scheme.

1 Introduction

Global climate models (GCMs) are main tools for projecting future changes in the earth's climate and provide the data products necessary for impact, adaptation, and mitigation studies in many societal sectors. A major limitation of GCMs is their horizontal mesh resolution which is quite coarse relative to the scale of processes that control, e.g., clouds and precipitation, as well as exposure metrics in most impact assessments. Most GCMs do not provide information on scales less than about hundred kilometers and commonly several hundreds of kilometers.

Thus, estimating climate changes on the local scale from the global scale modeling products requires some form of regionalization. It means downscaling from GCM output from global to local climate conditions, i.e., from a coarse to a high spatial resolution. Downscaling of climate data can be carried out by either statistical or dynamical methods. Statistical downscaling is based on the development of empirical relationships between historical large-scale atmospheric and local climate characteristics. Statistical downscaling incorporates a heterogeneous group of methods that varies in sophistication and applicability. The main statistical downscaling categories include linear methods (Hay and Clark 2003; Hay et al. 2000; Zorita and von Storch 1999), weather classifications methods (Yin 2011; and Benestad 2008), and weather generator methods (Wilby and Dawson 2013; Ahmed et al. 2013; Jones and Thornton 2013; UNFCCC 2013; Semenov 2012; Wilby et al. 2009). Dynamical downscaling refers to the use of high-resolution regional climate models (RCMs), driven by

✉ Mohamed Abdelwares
   moh_sherif20@yahoo.com

[1] Department of Irrigation and Hydraulics, Faculty of Engineering, Cairo University, Dokki, Giza 12613, Egypt

[2] Atmospheric Chemistry Department, Max Planck Institute for Chemistry, Mainz, Germany

[3] Energy, Environment and Water Research Center, Cyprus Institute, Nicosia, Cyprus

boundary conditions provided by GCMs. The RCM is characterized by a higher spatial resolution and better representation of regional information, which more realistically represent the local topography and atmospheric processes. The adequacy of the RCM simulation results depends to a large extent on the used boundary conditions from the GCM (Wilby et al. 2009). In addition, each RCM contains different dynamical schemes and physical parameterizations for the grid resolved and sub-grid scale variables. The RCMs need to be customized through adjusting the configuration and optimizing tunable parameters.

The weather research and forecasting model (WRF) is a state-of-the-art regional climate model, i.e., a mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting needs. The model serves a wide range of meteorological applications across scales from tens of meters to thousands of kilometers (Michalakes et al. 2001). Parameterization of sub-grid scale variables is one of the challenging problems in the application of WRF (García-Díez et al. 2013; Dudhia 2014). Physics parameterizations are representations of important physical processes that cannot be directly resolved by the model, based on simplified physical or statistical, i.e., empirical representations. Parameterizations are needed because the processes may occur on a sub-grid scale, are too complex, and computationally costly to be resolved explicitly, or when a process is not sufficiently well understood to be represented through mathematical equations, hence parameterizations aim at predicting the effects of sub-grid scale process using only information at the resolved grid.

WRF offers a series of physics options that can be combined into different configurations. These options range from simple and efficient to complex and computationally costly. The WRF system categorizes physics parameterization schemes into (1) shortwave radiation (SWR), (2) longwave radiation (LWR), (3) planetary boundary layer (PBL), (4) cumulus convection scheme (CUM), (5) microphysics (MIC), and (6) land surface model (LSM).

The Eastern Nile Basin (ENB), which encompasses the Blue Nile, Tekeze river basin, Atbara river basin, Sobat river basin, and Baro-Akobo river basin, is the main source of Nile inflow at Aswan in Egypt. It contributes ~85% of the total water flow at the High Aswan Dam. Egypt comprises about 4% of the total basin area, Sudan 13%, South Sudan 61%, and Ethiopia 22%. The basin has both tropical and sub-tropical climates, with seasonal precipitation, while most of the northern area of the basin is considered arid to semi-arid. Regarding the orography of the region, the elevation within the basin is found to range from zero meter above sea level (ASL) at the northern part of the basin to 4300 m ASL in the Ethiopian highlands.

Many previous studies have tested the WRF sensitivity to alternative physics parameterizations for different simulation timescales and a variety of locations worldwide (Ruiz et al. 2010; Zittis et al. 2014; Flaounas et al. 2011). Zittis et al. (2014) used the WRF model over the MENA–CORDEX domain (MENA is Middle East–North Africa; CORDEX is COordinated Regional climate Downscaling Experiment). They investigated the performance of 12 different physics combinations at 50 km grid resolution, assessing the model performance in simulating total precipitation, minimum and maximum surface air temperature, and compared with gridded observational data and station measurements. The results showed that the simulated surface air temperature is most sensitive to the choice of the microphysics parameterization selection, while the precipitation is more sensitive to the cumulus parameterization, and also that precipitation is difficult to be captured by the model, especially in areas with pronounced topography. Finally, they found that the obtained configuration should be considered only for the MENA region, while for other locations with different topography and prevailing weather patterns, the results may differ.

Argent et al. (2015) customized the WRF model for the Lake Victoria basin in order to capture precipitation patterns in the region. Comparison of 13 different physics parameterizations was made to obtain the best combination to be used, in addition to different sea surface temperature regimes. Also they compared the results for extreme years with the climatology year used for the initial analysis. Finally, their work provided a method for comprehensively customizing the model for a particular region that can be used for future work over any region. Mooney et al. (2013) evaluated the sensitivity of the WRF model to parameterization schemes over Europe. They used the WRF model to downscale the ERA-Interim reanalysis data over a domain covering Europe for 12 different physics parameterization combinations. The results showed that the model can simulate the surface air temperature adequately with high correlation and low bias. With respect to precipitation it seems that it is not well modeled by the WRF model, with a low correlation coefficient and large bias. Mean sea level pressure modeled by the WRF model showed no significant bias with a high correlation coefficient.

Regarding our study area, climate change impacts are poorly understood and rarely investigated. Studies with regional climate models over the ENB are needed to demonstrate the effects of climate change on the basin, and especially the water resources in the region. In order to study the climate change impacts, a first step is to customize the RCM over the area, which is our main objective here, to be able to use the model for climate change projections. To obtain a recommended configuration for the WRF model over the ENB, many numerical experiments were conducted. The experimental design is descried briefly in Section 2. Results and discussion are given in Section 3. The WRF recommended configuration is detailed in Section 4. The summary and main conclusions are presented in Section 5.

2 Data and methodology

2.1 Model and domain

ENB hindcast climate simulations were performed with the non-hydrostatic fully compressible WRF (WRF-ARW version 3.5; Skamarock et al. 2008). Two nested domains were used; the parent simulation domain is the MENA–CORDEX domain which covers the Middle East and North Africa, with a horizontal resolution of 0.44° (≈ 50 km) and 30 vertical levels. The extents of the MENA–CORDEX domain, in addition to a higher resolution nested domain centered over the ENB are shown in Fig. 1a.

All simulations were forced by ERA-Interim reanalysis data (Dee et al. 2011) over a period of 2 years (1998–1999), with additional 6-month spin-up period starting in July 1997. The Noah land surface model (NPAH) was used to represent the land surface processes in all simulation experiments (Chen et al. 1996). Other settings that were common in all simulations are the pre-processing and implementation of the forcing fields in the simulations such as the relaxation zone, the setting of vertical layering, land use databases, and sea surface temperatures.

Three different sets of simulations were conducted to test the sensitivity of the WRF model, in order to obtain a recommended configuration for the model to be used in climate studies in the ENB. The first set of experiments was designed to investigate the effect of different horizontal extents of the domain. The second set of experiments was designed to investigate the effect of different horizontal mesh resolutions for the model domain over the ENB. The last set of experiments tested the performance of different combinations of physics parameterization schemes to pursue the optimal combination to be used in an RCM assessment.

Table 1 shows the coordinates of three different horizontal extents used to test the sensitivity of WRF to different ENB domain definitions. These simulations were performed to test the effect of including the Indian Ocean on hindcasted precipitation and air temperature over the ENB domain. Figure 1c shows the extent and the orography of the nested domain over the ENB relative to the parent domain.

Using the outcomes of the domain horizontal extent simulations, the second set of experiments was performed to investigate the mesh resolution of the nested domain over the ENB. As the parent domain mesh resolution is 0.44°, three ratios for the mesh size of the parent domain to the mesh size of the nested domain were tested. Table 2 shows the different tested ratios: 1:3, 1:4, and 1:5 which are the common nesting ratios used for the WRF model.
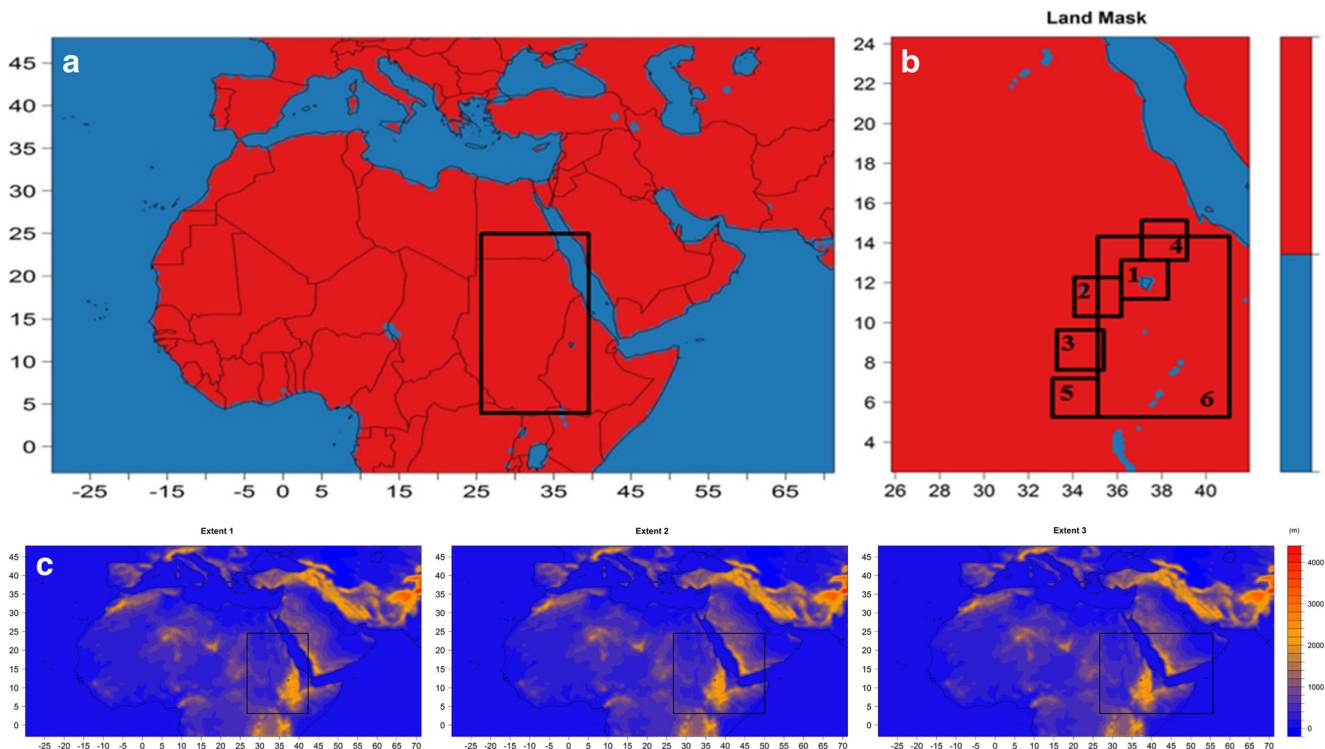
Fig. 1 a Extent and mask of the MENA–CORDEX domain in addition to the ENB domain extent, with red for land and blue for water. b The six sub-regions defined for model evaluation: 1 Lake Tana, 2 Grand Ethiopian Renaissance Dam (GERD), 3 Sobat, 4 Tekeze, 5 Akobo, and 6 Ethiopian Highlands (EH). c The extent and orography of the three different alternatives, i.e., from left to right extent 1, 2, and 3

Table 1 Longitude, latitude, and number of grids for the domain horizontal extents numerical experiments

	Extent 1	Extent 2	Extent 3
Longitude	25.5° E–42° E	25.5° E–50.0° E	25.5° E–56.0° E
Latitude	2.5° N–24.5° N	2.5° N–24.5° N	2.5° N–24.5° N
No. of grid cells	112 × 151 × 30	160 × 151 × 30	208 × 151 × 30

2.2 Physics parameterization

Subsequent to selecting the horizontal domain extents and the mesh resolution, a third set of simulations was performed to determine the most appropriate ensemble of physics parameterization schemes. Sixteen different WRF configurations of physics parameterizations were applied (Table 3).

Three cumulus convection schemes were examined, the Kain-Fritsch scheme (KF; Kain 2004), the Betts-Miller-Janjic scheme (BMJ; Janjic 1994), and the Grell 3D scheme (Grell 3D; Grell 1993; Grell and Devenyi 2002). The KF scheme is a deep and shallow sub-grid scheme using a mass flux approach with downdraft and CAPE removal time scale; it includes condensed and gaseous water detrainment. The BMJ scheme is an adjustment scheme for deep and shallow convection, and relaxing is applied towards variable temperature and humidity profiles which determined from thermodynamic consideration. Grell 3D is a multi-closure, multi-parameter, ensemble method that explicitly accounts for updrafts and downdrafts, designed for higher resolution, allowing for subsidence between neighboring columns.

Three different planetary boundary layer schemes were used, the Mellor-Yamada-Janjic scheme (MYJ; Janjic 1994), the Yonsei University scheme (YSU; Hong et al. 2006), and the symmetric convective model V2 (ACM2; Pleim 2007). The MYJ is a local scheme with total kinetic energy-based vertical mixing in boundary layer and free atmosphere. The YSU is a nonlocal mixing scheme with explicit treatment of entrainment, suitable for weather

Table 2 Attributes of the numerical experiments for testing the domain mesh size

	Experiment 1	Experiment 2	Experiment 3
Longitude	25.5° E–42.0° E	25.5°E – 42.0°E	25.5° E–42.0°E
Latitude	2.5° N–24.5° N	2.5° N–24.5° N	2.5° N–24.5° N
Parent grid ratio	1:3	1:4	1:5
Resolution	0.147° (≈ 16 km)	0.110° (≈ 12 km)	0.088° (≈ 10 km)
No. of grid cells	112 × 151 × 30	149 × 201 × 30	186 × 251 × 30

Table 3 Combination of physics parameterization for the 16 different simulations

Simulation	SWR	LWR	PBL	CUM	MIC	LSM
#1	Dudhia	RRTM	YSU	KF	WSM3	Noah
#2	CAM	CAM	YSU	KF	WSM5	Noah
#3	Dudhia	RRTM	ACM2	Grell 3D	Eta	Noah
#4	Dudhia	RRTM	YSU	BMJ	WSM5	Noah
#5	CAM	CAM	MYJ	BMJ	Lin	Noah
#6	Dudhia	RRTM	ACM2	BMJ	Eta	Noah
#7	Dudhia	CAM	ACM2	Grell 3D	Eta	Noah
#8	CAM	CAM	YSU	BMJ	WSM5	Noah
#9	CAM	CAM	ACM2	Grell 3D	Eta	Noah
#10	CAM	CAM	ACM2	BMJ	Eta	Noah
#11	CAM	CAM	ACM2	BMJ	WSM6	Noah
#12	CAM	CAM	YSU	BMJ	WSM6	Noah
#13	CAM	CAM	MYJ	BMJ	Eta	Noah
#14	Dudhia	RRTM	ACM2	KF	Eta	Noah
#15	CAM	CAM	MYJ	BMJ	WSM6	Noah
#16	CAM	CAM	MYJ	BMJ	GCE	Noah

forecasting and climate prediction models. ACM2 is a nonlocal mixing upwards from the surface layer and local mixing downwards. Six different microphysics options were used, the WRF single-moment 3, 5, and 6-class schemes (WSM3/WSM5/WSM6; Hong et al. 2004; Hong and Lim 2006), the Lin scheme (Lin et al. 1983), the Eta microphysics (Eta; NOAA 2001), and the Goddard microphysics scheme (GCE; Tao et al. 1989).

The radiation schemes evaluated were the Community Atmosphere Model (CAM; (Collins et al. 2004) for shortwave and longwave, Dudhia scheme for shortwave radiation (Dudhia 1989), and the Rapid Radiative Transfer Model scheme (RRTM;Mlawer et al. 1997) for longwave radiation. CAM is a spectral scheme with eight longwave bands, allows for interaction with aerosols, clouds, and trace gases, and the ozone profiles are a function of month and latitude. Dudhia is a simple downward integration efficiently representing cloud and clear sky absorption and scattering. The RRTM scheme is an accurate spectral scheme, accounts for multiple bands, trace gases, and microphysical properties, as it also interacts with clouds, while the ozone profile is specified. The Noah land surface scheme (NOAH LSM; Tewari et al. 2004) was kept common in all WRF configurations. The NOAH LSM represents soil temperature and soil moisture in four layers, and fractional snow cover, frozen soil physics, and vegetation effects are included, while also it provides heat and moisture fluxes to the planetary boundary layer. NOAH LSM is the recommended land surface model to be used with MODIS land category data, used in this study.

The ERA-Interim reanalysis dataset (Dee et al. 2011) was used to provide initial and boundary conditions, including sea surface temperature. The boundary conditions were updated every 6 h. ERA-Interim is a global atmospheric reanalysis data including global atmospheric and surface parameters available from 1979 to present at a spectral resolution ≈ 80 km and 60 vertical levels from the surface up to 0.1 hpa. The Noah-modified 21-category IGBP-MODIS land-use scheme with 30 arc second resolution was used to represent the land mask of the domain, different vegetation categories, topography, different soil categories, surface albedo, green fraction, etc.

2.3 Observations

To evaluate the model performance, we focused on two surface variables: precipitation (convective and non-convective precipitation) and the surface air temperature 2 m above the ground. For precipitation, the Global Precipitation Climatology Centre (GPCC) version 6 (Schneider et al. 2011) was used. The GPCC is a monthly total precipitation dataset available from 1901 to present, with a spatial resolution of $0.5° \times 0.5°$. The University of Delaware (UDEL) dataset was used in assessing the model performance in simulating the surface air temperature. The UDEL dataset is a monthly global gridded high-resolution station (land) data for air temperature and precipitation available from 1901 to 2010, with a spatial resolution of $0.5° \times 0.5°$. Precipitation and air temperature fields were re-gridded and interpolated to the grid of the nested domain mesh size to facilitate the evaluation of the simulated WRF results.

2.4 Evaluation methodology

The model performance was assessed for two seasons, the rainy season from June through September (JJAS), and the dry season from December through March (DJFM). In addition, the analysis was made using only grid points over the identified sub-regions, shown in Fig. 1b. In order to evaluate the WRF output, we used a range of statistical metrics, focusing on the two surface variables, i.e., the total precipitation and surface air temperature, and compared with the re-gridded observation fields. The statistical metrics were divided into three major categories: the standard regression statistics, error indices, and dimensionless techniques.

Standard regression statistics were used to determine the strength of the linear relationship between the simulated data and corresponding observed data, where the slope and y-intercept of the best fit regression line and Pearson's correlation coefficient were calculated. The slope indicates the relative relationship between simulated and observed values and the y-intercept indicates the presence of a lag or lead between model results and observations; also it indicates if there is an over- or underestimation in the model results compared to the observations. A slope of 1 and y-intercept of 0 indicate that the model perfectly reproduces the observed data. The Pearson's correlation coefficient is a measure of the strength of a linear association between the simulation and observations. The Pearson's correlation coefficient is given by the Eq. (1), which varies between zero and one, where the value of zero indicates that there is no correlation between the model simulations and the corresponding observations, while a value of one indicates that the model simulations and observations are perfectly correlated by a straight line:

$$COR = \frac{\sum_{i=1}^{n}\left(Obs._i - \overline{obs.}\right)\left(Sim._i - \overline{Sim.}\right)}{\sigma_{Obs.}\sigma_{Sim.}} \quad (1)$$

where:

n	the total number of the grid points,
$\sigma_{Obs.}, \sigma_{Sim.}$	the standard deviation of the observations and the model simulations, respectively, and
$\overline{Obs}, \overline{Sim}$	the mean values of the observations and the model simulations, respectively.

Error indices were used to quantify systematic deviations in the data of interest. We used the mean absolute error (MAE) given by Eq. (2) and the percent bias (Pbias) given by Eq. (3). A perfect fit is achieved with an MAE value of zero. The optimal value of Pbias is also zero, with low-magnitude values indicating accurate model simulations. Positive values indicate a model overestimation bias, while negative values indicate a model underestimation bias. The dimensionless modified index of agreement (MIA), given by Eq. (4), was used as a standardized measure of the degree of model prediction error (Willmott 1981; Legates and McCabe 1999). MIA varies between 0 and 1, with a MIA value of 1 indicating perfect agreement between the simulated and observed values, and a MIA value of zero indicates no agreement at all.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(Sim._i - Obs._i) \quad (2)$$

$$Pbias = 100 \times \left(\frac{\sum_{i=1}^{n}(Sim._i - Obs._i)}{\sum_{i=1}^{n}Obs._i}\right) \quad (3)$$

$$MIA = 1 - \frac{\sum_{i=1}^{n}(Obs._i - Sim._i)}{\sum_{i=1}^{n}\left|Sim._i - \overline{Obs.}\right| + \left|Obs._i - \overline{Obs.}\right|} \quad (4)$$

where:

n	the total number of the grid points, and
$\overline{Obs}, \overline{Sim}$	the mean value of the observation and simulation fields, respectively.

In order to facilitate the comparison process between each model configuration against the observations, a single weighted average index measure is computed from the previously described statistical measures. A total index score of 100 points is divided evenly among the five different statistical metrics such that each one can achieve 20 points out of the total 100. For each statistical measure, the allocated weight is divided between the minimum and maximum values of the measure, where the lowest range of the statistical measure takes a value of 0 and the highest range a value of 20, with linear interpolation applied in between.

For example, in the Pearson's correlation coefficient the minimum value, which is zero, means no correlation between the simulation data and observations, so we allocate a minimum score of 0, while the maximum value of the Pearson's correlation coefficient, which is one, means perfect correlation between the simulation data and observations, thus, it is allocated a maximum score of 20. For Pearson'scorrelation coefficients between 0 and 1, a linear interpolation was applied to determine the corresponding score that varies from 0 to 20. Because the y-intercept and the percent bias may be negative, we used the absolute values of both measures such that the minimum value becomes zero and the maximum is the maximum of absolute value. In this way, for a certain WRF configuration experiment, by adding the weights from the different statistical metrics, we obtain a final score out of 100 that represent the relative accuracy of each WRF configuration against the observations.

Taylor diagrams are used to provide a brief statistical summary of how well different WRF configurations and observations match in view of their correlation and standard deviation (Taylor 2001). The correlation and standard deviation are indicated with a single point on a 2-D polar coordinates plot. The standard deviation is represented by the polar distance from the origin and the azimuthal position refers to the correlation. The reference point with a correlation and standard deviation of 1 is also indicated on the Taylor plot. This diagram helps identify the model configuration with optimal performance and to distinguish between errors due to limitations in the simulation results.

3 Results

3.1 Domain extents and mesh resolution

The WRF experiments that were performed to test the horizontal extent of the ENB nested domain did not result in apparent differences in the simulated precipitation and surface temperature fields for the different three tested domains. Accordingly, it was decided to use the smallest ENB domain extent (25.5° E–42.0° E, 2.5° N–24.0° N) in all other model configurations. This approach helps minimize computational cost because of the relatively smaller size of the nested domain. The WRF experiments that were performed to select the optimal mesh size of the nested domain indicated that higher mesh resolutions result in more realistic simulations of the precipitation fields. Accordingly, all further WRF configuration are performed with the highest mesh resolution of 10 km. All 16 WRF configurations for the physics parameterizations were performed using the selected nest horizontal domain (25.5° E–42.0° E, 2.5° N–24.0° N) and the selected nest mesh size (10 km).

3.2 Precipitation

Figure 2 shows the daily mean precipitation bias (WRF–GPCC) during the wet season months (JJAS) for different WRF configurations of physics parameterizations during the period 1998–1999. All simulations overestimated the observed precipitation amounts over the pronounced topography of the Ethiopian highlands. In simulations 5, 13, 15, and 16 the overestimations are relatively smallest. In these four simulations, the common physics schemes were CAM scheme for the short and the longwave radiation, BMJ scheme for the cumulus convection, MYJ scheme for the planetary boundary layer, and Noah land surface model scheme. Other simulations showed larger overestimations of precipitation over the southwestern part of the domain, such as simulations 1, 2, 3, and 14. Figure 3 shows the daily mean precipitation bias during the dry season months (DJFM). The same pattern can be found during the dry season where precipitation was overestimated in the same simulations.

In order to investigate the effects of each physics parameterization individually and to identify how the different schemes affect precipitation, we selected simulations where only a single parameterization scheme is changed. To check the effect of changing the shortwave radiation schemes, simulations 7 and 9 were used, and for the longwave radiation schemes, simulations 3 and 7 were used. To check the impact of changing the cumulus convection schemes simulations 3, 6, and 14 were used. Simulations 5, 13, 15, and 16 showed the consequences of changing the microphysics scheme. Finally, simulations 11, 12, and 15 showed the effects of changing the planetary boundary layer schemes. By comparing the above simulations, it was found that the precipitation simulation is most sensitive to changing the cumulus parameterization scheme and the planetary boundary layer scheme, and less sensitive to changing the microphysics scheme. As for the effect of radiation the precipitation seems to be more sensitive to changing the longwave than the shortwave radiation scheme.
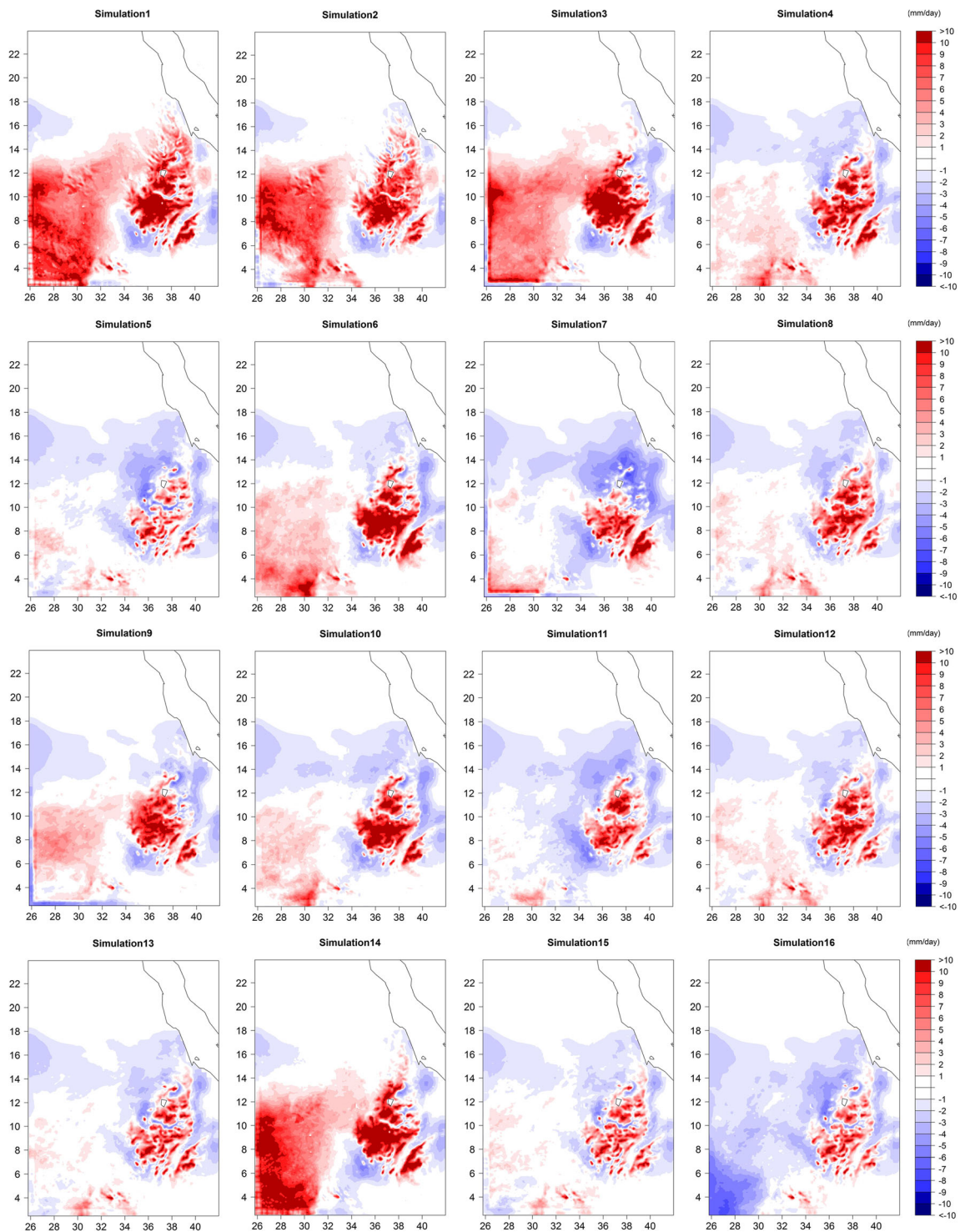
Fig. 2 Daily mean precipitation bias (WRF–GPCC) for the wet season for combinations of physics parameterizations, during the period 1998–1999

To compare the 16 WRF configurations and to evaluate the performance of the model with each physics combination, the different statistical metrics have been calculated at the 10 identified sub-regions, and the scoring technique was applied. In the following, the results at the GERD sub-region were presented as a sample of model results. Figure 4a, b presents the

Taylor diagrams for the 16 WRF configurations for both the rainy and dry seasons. Most simulations achieve a correlation coefficient varying between 0.45 and 0.65, and the centered root mean square error ranged between 2 and 3. For the wet season, it was found that some simulations have a high correlation coefficient, such as simulation 9, when compared to
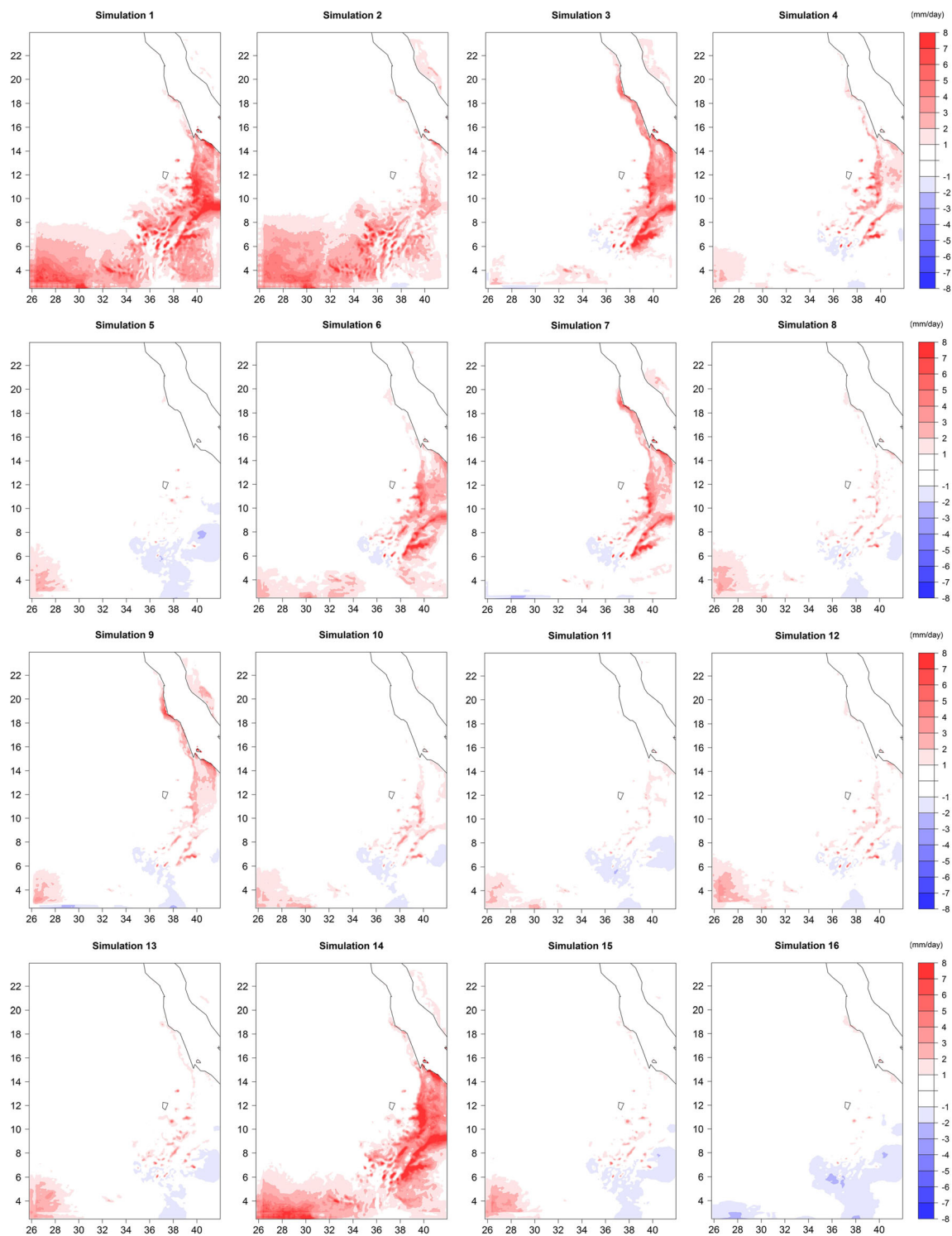
Fig. 3 Daily mean precipitation bias (WRF – GPCC) for the dry season for combinations of physics parameterizations during the period 1998–1999

other simulations such as simulation 11. Further, simulation 11 has a lower standard deviation than simulation 9. Figure 4c, d shows the probability density function for the 16 simulations and the GPCC data during the wet and dry seasons.

Focusing on the wet season, many simulations could not capture the general pattern of the GPCC, while some

simulations captured the general pattern but with slight changes in the density of some small values of precipitation. For some simulations such as 1, 2, and 3, it was noticed that the relative likelihood of large precipitation amounts was higher than in GPCC, which means overestimation of precipitation in these simulations. Figure 4e presents the daily mean

Fig. 4 **a** Taylor diagrams showing the correlation coefficients, RMSE, and standard deviations of precipitation during the wet season relative to GPCC for the 16 simulations over the GERD sub-region, **b** same for the dry season, **c** the PDF of precipitation for the 16 simulations for the GERD sub-regionin the wet season, **d** for the dry season, and **e** daily mean precipitation by month, over the GERD sub-region for the 16 simulations and GPCC during the period 1998–1999
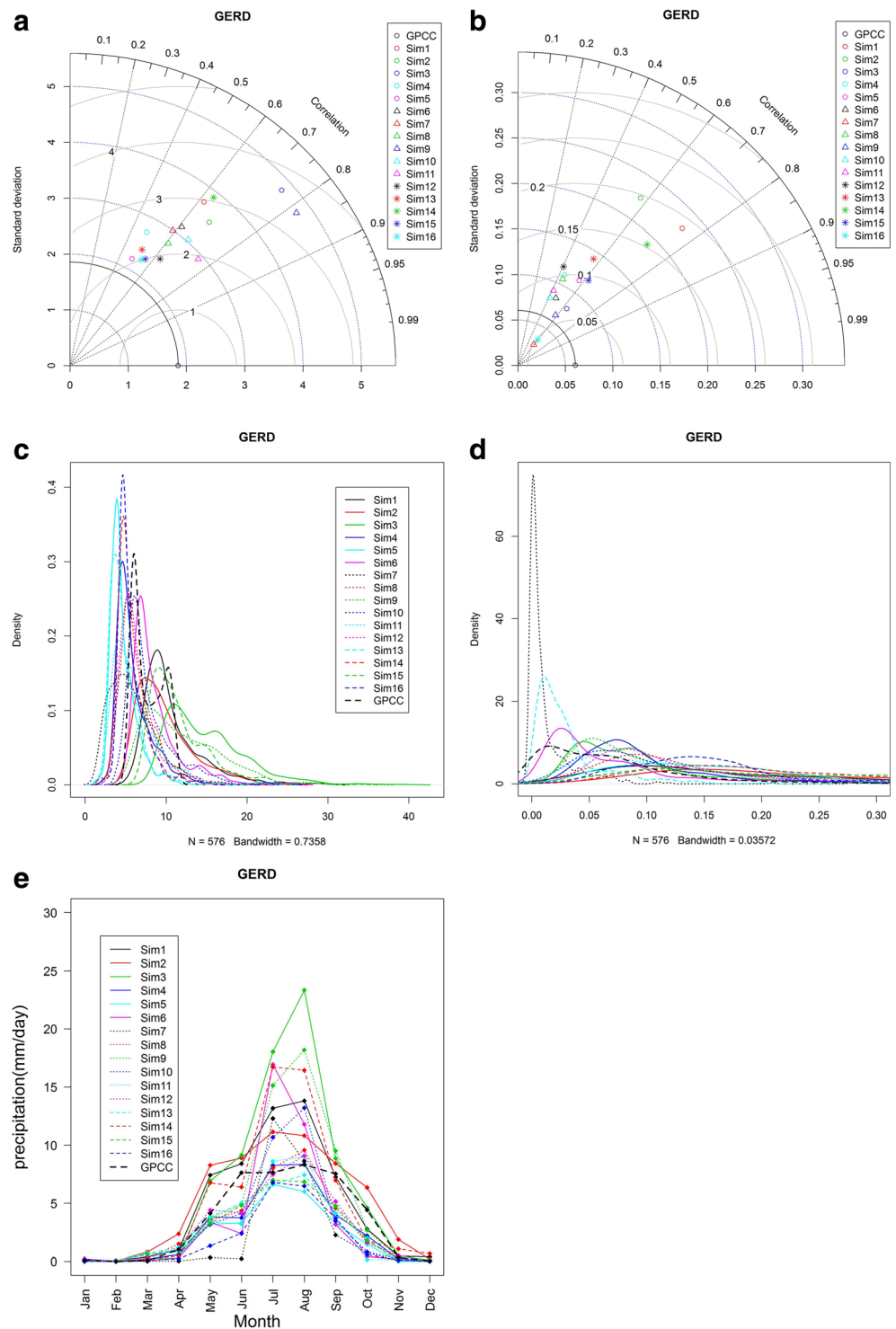
precipitation by month during the whole simulation period at the GERD sub-region. Some simulations were able to capture the annual cycle of precipitation, e.g., simulations 15, 12, and 8, with a small bias during the wet season. Other simulations yielded high biases, such as simulation 3 where the bias reached about 17 mm/day. Also simulations 1, 2, 8, and 9 gave rise to strong overestimations during the rainy months.

Figure 5 shows the scatter plots between the simulated precipitation on the *y*-axis and GPCC observations on the *x*-axis during the wet season. It clearly shows the precipitation biases, i.e., either overestimation or underestimation for the different WRF configurations. Comparing the simulated precipitation from the different WRF physics configurations against the observations underscores the
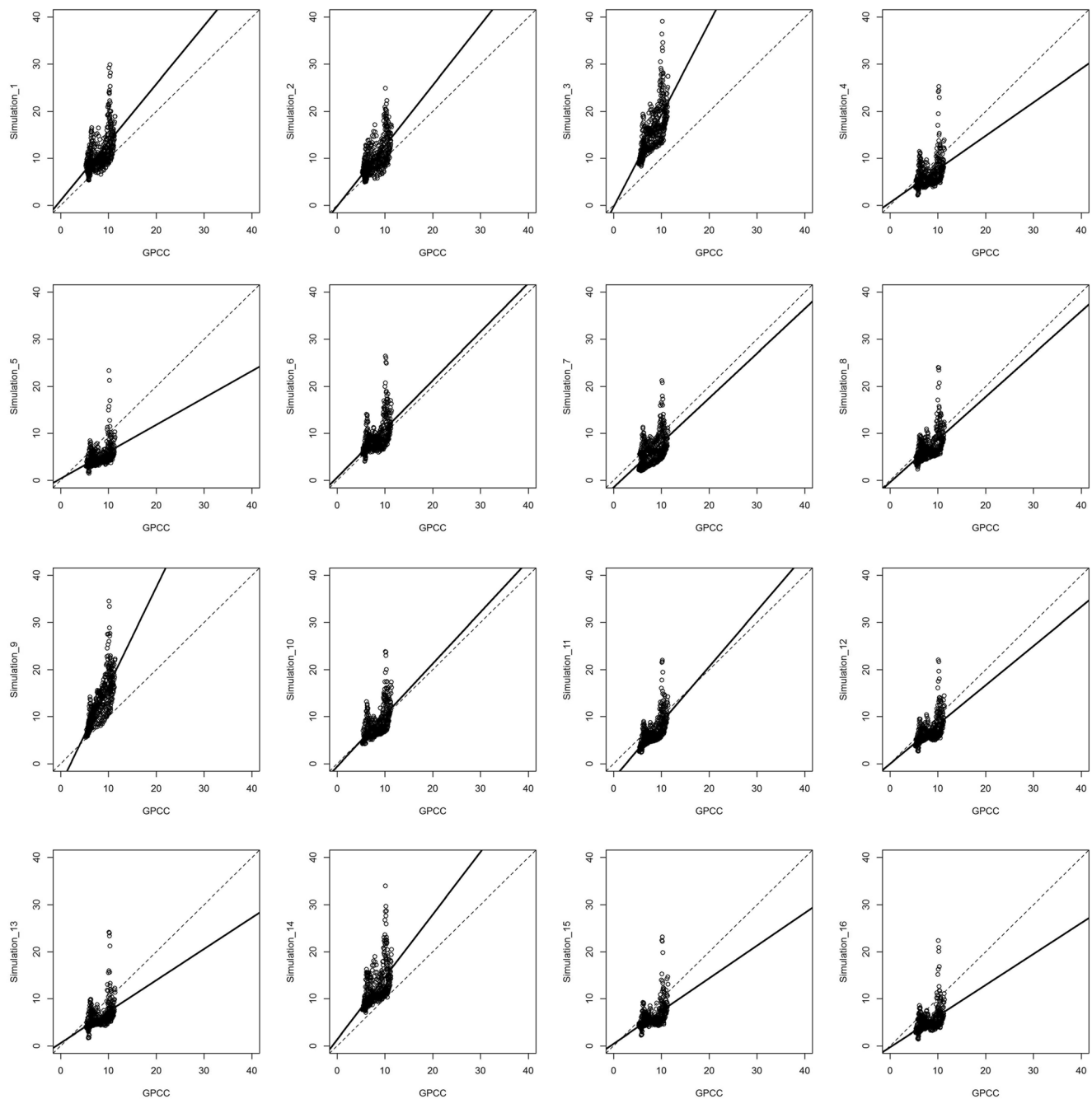
Fig. 5 Scatter plots between the simulated precipitation in the wet season on the *y*-axis and GPCC on the *x*-axis, over the GERD sub-region for the 16 simulations

strong sensitivity of precipitation results to the adopted physics parameterization schemes. A detailed selection of the recommended combination of physics scheme will be presented later in the paper.

3.3 Surface air temperature

Figure 6 shows the mean daily bias of the surface air temperature (WRF vs. UDEL) in the wet season during the simulation period 1998–1999. Unlike precipitation, there are no major

differences between the 16 different WRF configurations in the surface air temperature. Some differences were found in simulations 7 and 16, which showed an apparent underestimation compared to other simulations during the wet season.

The statistical metrics were computed at the 10 identified sub-regions with the results at the GERD presented hereafter as a sample of model results. Figure 7a shows a Taylor diagram for the 16 simulations for the surface air temperature in the wet season, where all simulations achieve similar correlation coefficients, central RMSE,
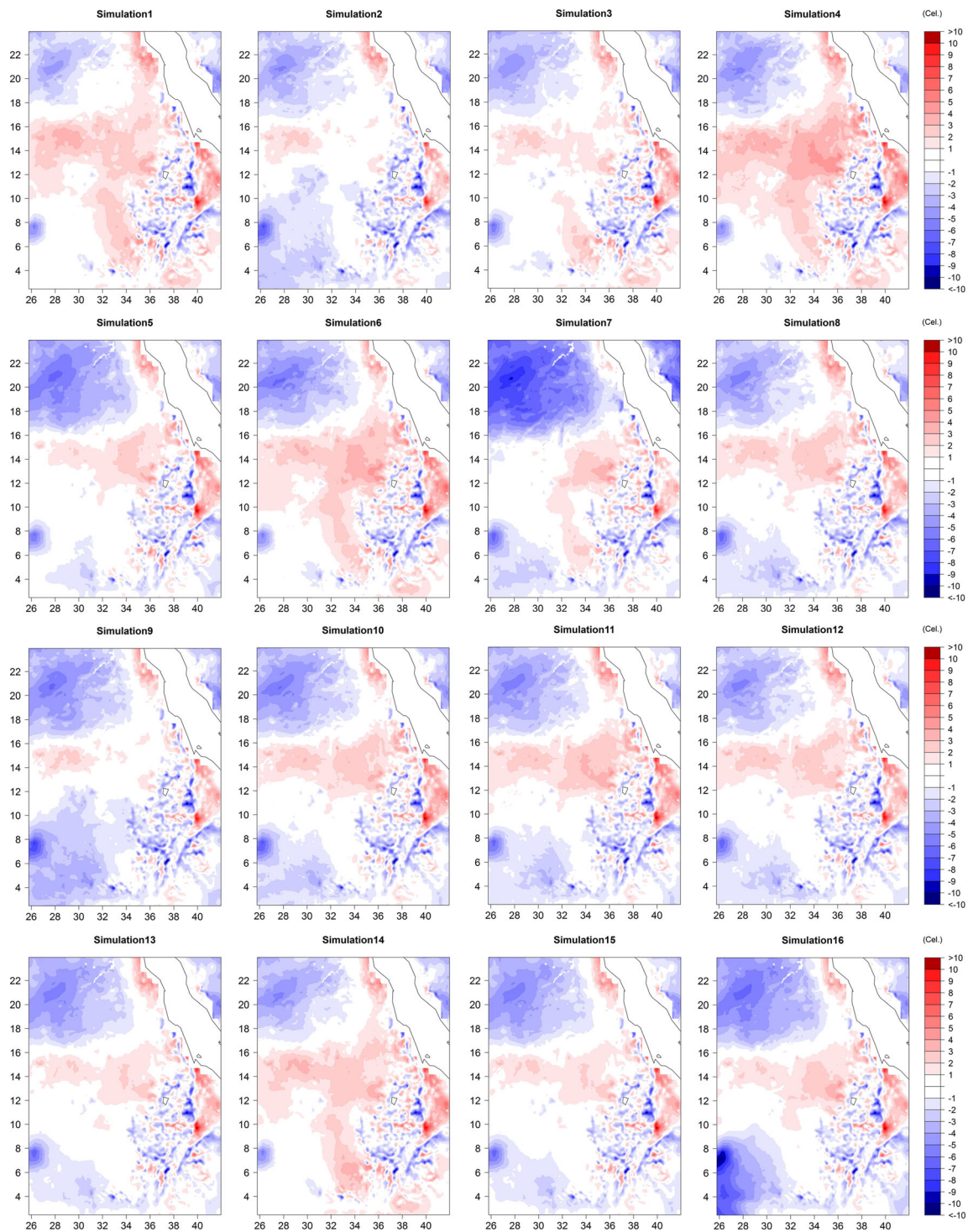
Fig. 6 Bias (WRF–UDEL) for the mean surface air temperature in the wet season for combinations of physics parameterizations during the period 1998–1999

and standard deviations. Figure 7b shows the mean surface air temperature on a monthly basis during the whole simulation period, and the bias between the simulations and the observations is typically nearly 3 °C, while the annual cycle of surface air temperature is captured well by all simulations. Figure 7c presents the probability density function between the simulated surface air temperature

and the UDL data, where all simulations approximately capture the general pattern of the observations. Figure 8 shows the scatter plots between the simulated surface air temperature during the wet season on the y-axis and UDL on the x-axis over the GERD sub-region for the 16 simulations. It indicates no major differences between the different WRF configurations.
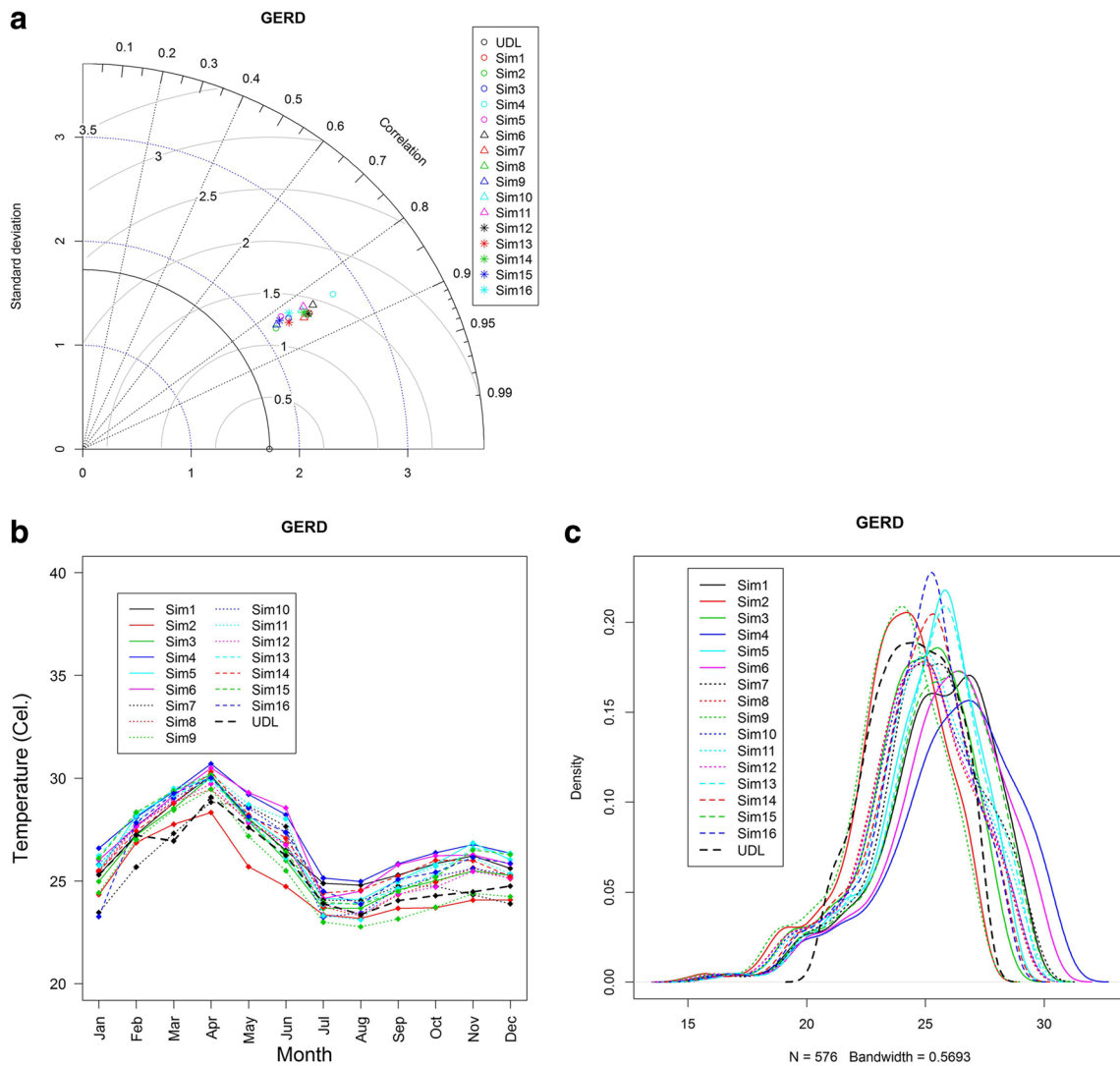
Fig. 7 **a** Taylor diagram showing correlation coefficients, RMSE, and standard deviations of surface air temperature during the wet season relative to UDL for the 16 simulations over the GERD sub-region, **b** daily mean surface air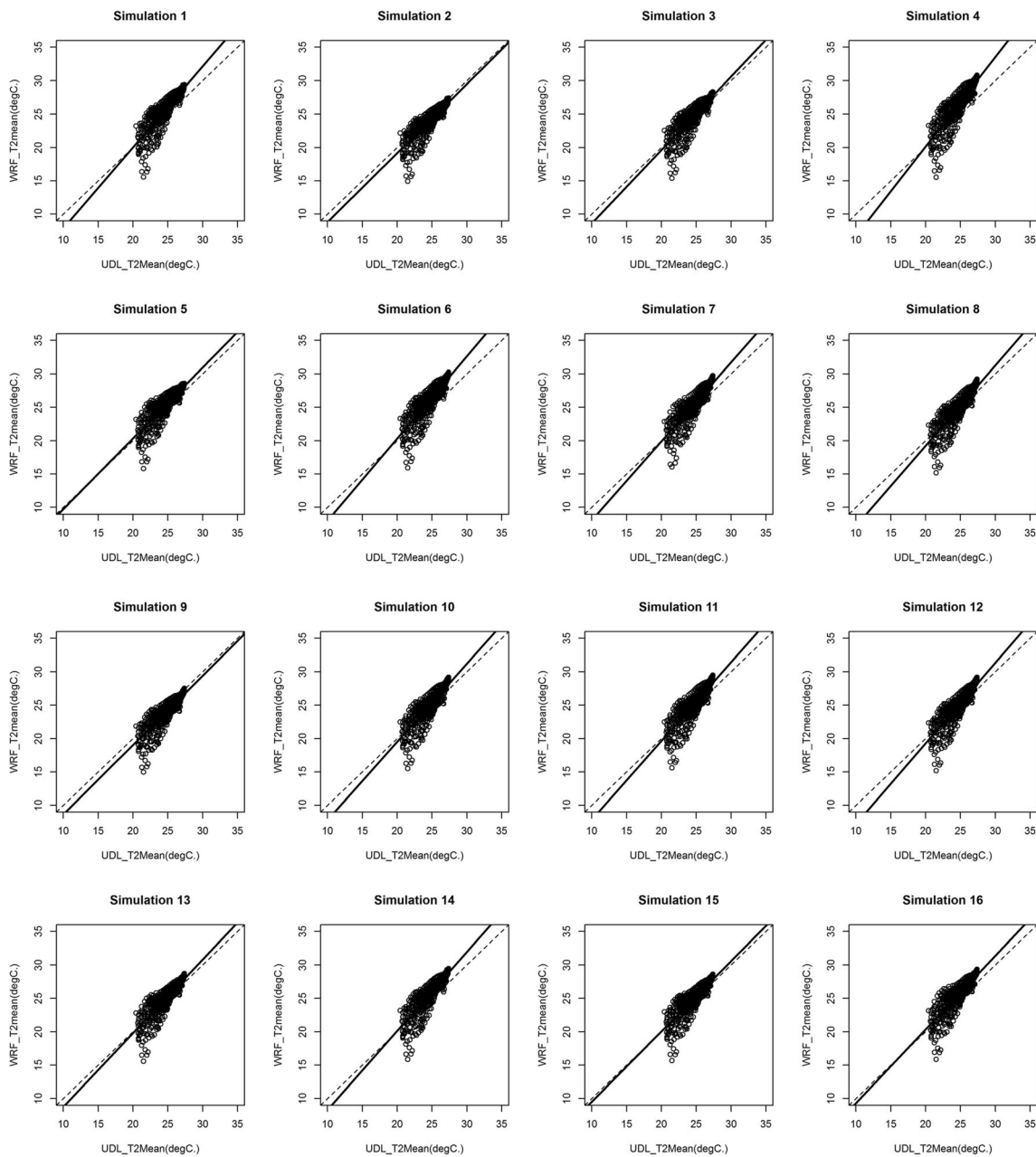 temperature by month over the GERD sub-region for the 16 simulations and UDL during the period 1998–1999, and **c** PDF for surface air temperature in the wet season for the 16 simulations in the GERD sub-region

4 Recommended WRF configuration

Previous investigations of the simulated precipitation and surface air temperature compared to observational data emphasized that changing the physics parameterization schemes can have major effects on the simulated precipitation fields, which is less the case for the surface air temperature over the ENB. To select the recommended WRF physics parameterization configuration over the ENB, the total score of the different statistical metrics at different sub-regions as well as the whole simulation domain were computed for the precipitation and surface air temperature in both the wet and dry seasons. Figure 9a shows a color matrix of the total scores (out of 100) for simulated precipitation and Fig. 9b for the simulated

surface air temperature. For precipitation, the Lake Tana sub-region showed a low score for all simulations, but by considering the whole domain the score increases. For surface air temperature the majority of the simulations achieve a high score, except at the GERD sub-region.

Table 4 presents the integral scores at all sub-regions for different WRF configurations. WRF configurations1 and 2 resulted in the lowest total score among the different configurations. Configurations 8, 12, and 15 attained the highest scores with minor differences between the three configurations. We find that any of these three configurations could be used adequately over the ENB domain. Focusing on the total score of precipitation during the wet season among these three, it was found that configuration #15 is optimal for the region.

Fig. 8 Scatter plots between the simulated surface air temperature on the *y*-axis and UDL on the *x*-axis, during the wet season over the GERD sub-region for the 16 simulations

Zittis et al. (2014) recommended a configuration for the whole MENA domain at 50 km resolution, with the YSU planetary boundary layer scheme, KF as cumulus scheme, the WSM6 microphysics scheme, CAM for long- and short-wave radiation, and the NOAH land surface model. The differences between his configuration and the recommended one here include the planetary boundary layer and the cumulus scheme, which appear to have major effects on the simulated precipitation compared to other parameterizations. This is mainly related to the nature of our study area, which is wetter than most of the MENA domain regions.

An additional experiment was performed to assess the suitability of the recommend WRF configuration in simulating the precipitation and surface air temperature during a drought period in the ENB. The simulation period of this experiment extended over 2 years (1984–1985) with additional 6 months of spin-up period, thus starting in July 1983. The overestimation in simulated precipitation (results not shown) over the Ethiopian highlands was still noticeable, though smaller compared to the results from the simulation period of the configuration experiments. This was mainly because precipitation was

a
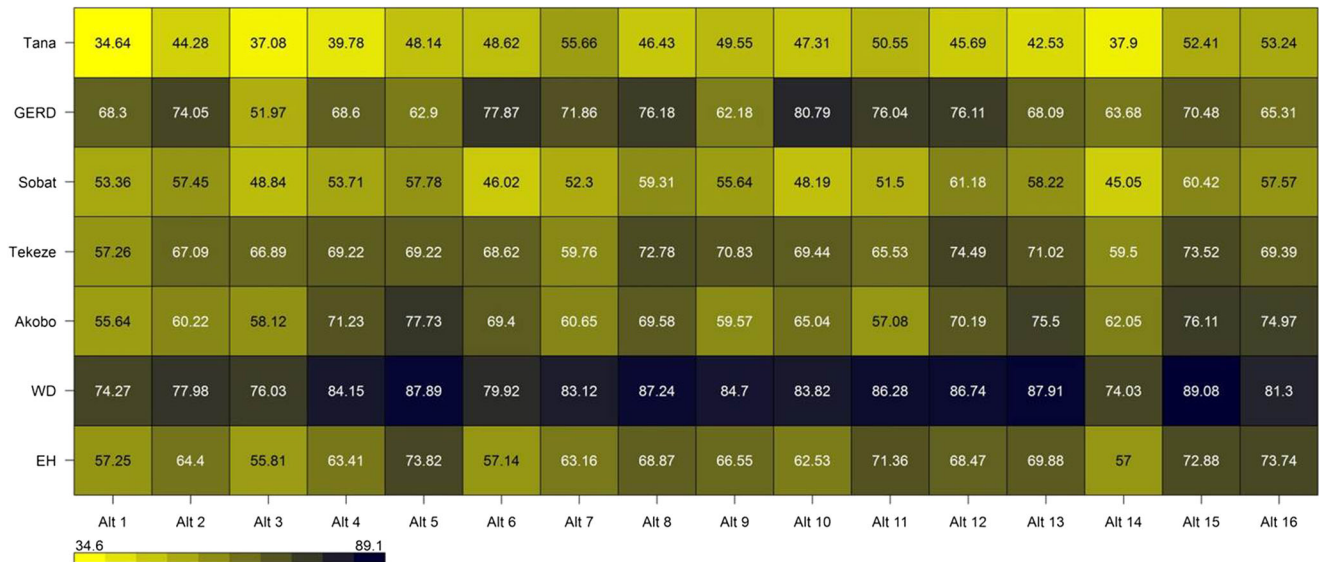
b
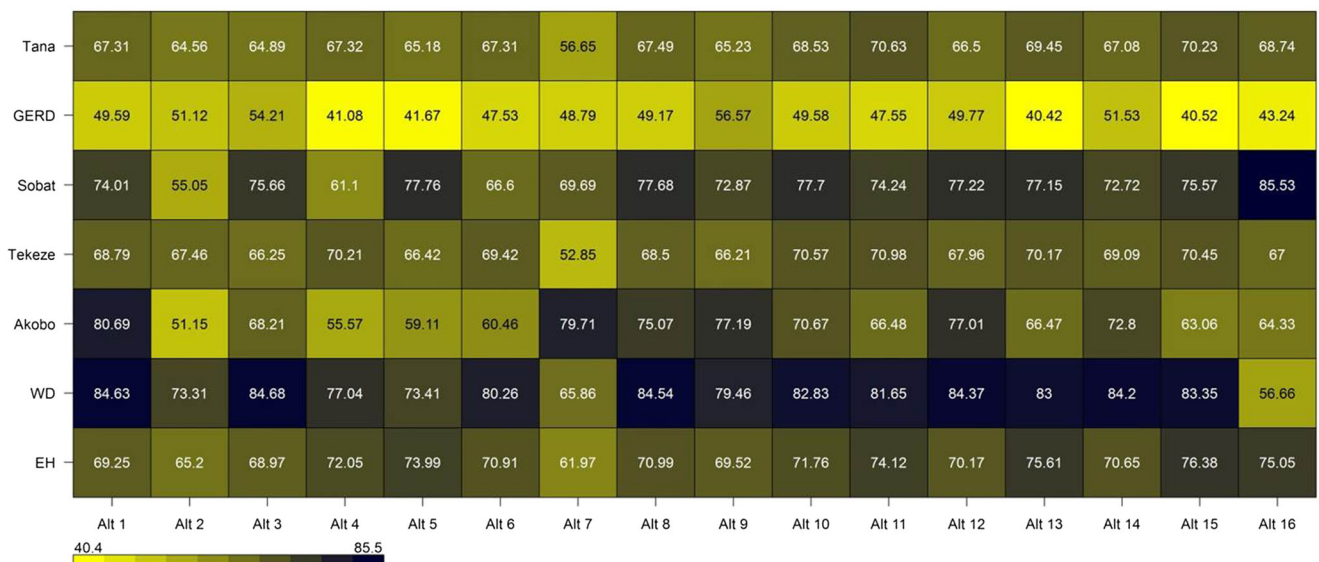
Fig. 9 **a** Color matrix showing the score (/100) for simulated precipitation during the wet season in each sub-region for the 16 simulations. **b** Score (/100) for simulated surface air temperature during the dry season in each sub-region for the 16 simulations

minimal during this period. The results for the surface air temperature appeared to be satisfactory with some overestimation during the dry season (results not shown). The results of the statistical metrics for each sub-region during the drought period indicate higher scores for surface air temperature and acceptable values for precipitation.

The relative enhancement in simulating the precipitation and surface air temperature by using the recommend WRF configuration (#15) over the default set of physics parametrizations (#1) was calculated and is presented in Fig. 10. It shows the enhancement in percent, which is the improvement in the simulated precipitation and surface air

temperature in the recommended and the default physics configuration relative to the observations. During the wet season of the simulation period, enhancements in the simulated precipitation from the recommend configuration reached up to 300% at a few grid points, while 64% of the domain grid points had an enhancement between 0 and 50%, and 16% of the domain grid points had an enhancement between 50 and 100%. Over the whole domain, the precipitation enhancement was found to be 47.5%. The enhancement was much less for the surface air temperature, with a maximum value of 12% during the rainy season at some grid points, and with an average enhancement value of 4% over the whole domain.

Table 4 Score (/700) for both seasons, for precipitation and surface air temperature for the 16 simulations

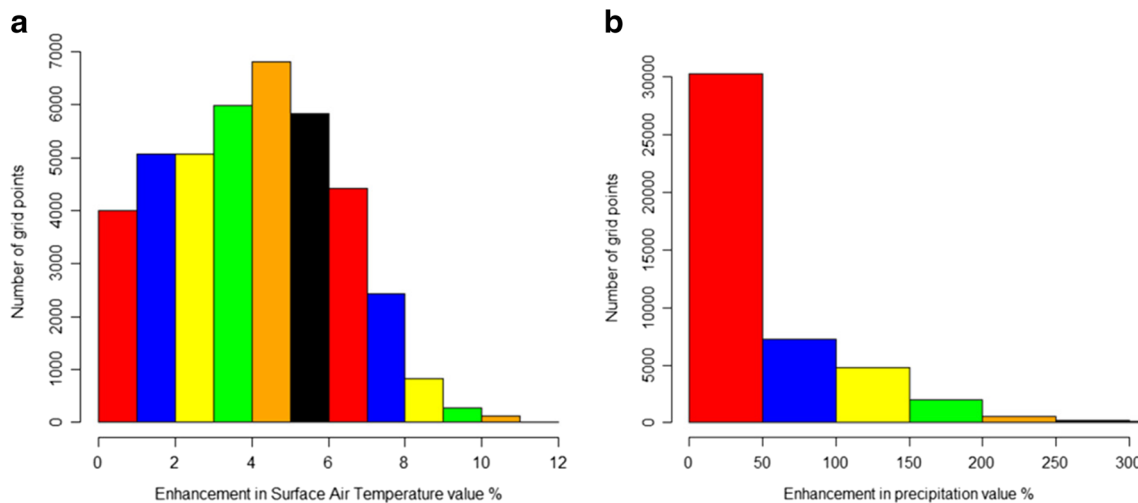|	Season	Simulation #															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Precipitation	Wet	400.7	445.5	394.7	450.1	477.5	447.6	446.5	480.4	449	457.1	458.4	482.9	473.1	399.2	494.9	475.5
	Dry	411.6	431.3	503.7	507.6	494.1	509.3	497	524.2	522.5	519.2	514.2	521.6	493.3	452.5	507.4	492.5
Surface Air Temperature	Wet	479	507.3	512.6	458.2	531.3	475.6	488.2	528.5	487.4	527.5	519.1	530.3	537.4	479.5	543.1	524.8
	Dry	494.3	427.8	482.9	444.4	457.5	462.5	434.5	493.4	487.1	491.6	485.6	493	482.3	488.1	479.6	460.6
Sum		1786	1812	1894	1860	1960	1895	1866	2027	1946	1995	1977	2028	1986	1819	2025	1953

5 Conclusions

With the objective of developing a regional climate model for the Eastern Nile Basin, our study focused on testing the performance of different parameterization configurations of the WRF-ARW model in hindcasting precipitation and surface air temperature. A simulation period over 2 years (1998–1999), with an additional 6-month spin-up period (thus starting in July 1997), was used to assess the performance of different WRF configurations in simulating precipitation and surface air temperature over the ENB during the wet season (June–September) and the dry season (December–March). Three sets of numerical experiments were conducted.

The first set tested the effects of changing the horizontal extents of a nested domain centered over the ENB within the CORDEX-MENA region as the parent domain. Three extents for the nested domain were tested to investigate the effect of including a larger area of the Indian Ocean into the simulation domain. No significant impact was found from increasing the horizontal extent of the nested domain, so it was decided to use the smallest domain extent considering the high computational cost of increasing the domain size. The second set of experiments tested the sensitivity of WRF to the horizontal mesh size; three experiments were performed for the resolutions 16, 12, and 10 km. Increasing resolution improved the simulation of precipitation and surface air temperature. Accordingly, it was decided to proceed with the highest resolution tested (10 km).

The third set of experiments was designed to select a recommended combination of physics parameterizations. A total of 16 WRF configurations with different physics parameterization combinations were tested to derive the optimal combination over the ENB. Generally, it was found that precipitation is most difficult to model realistically, as the performance of the WRF in simulating surface air temperature fields was generally more realistic than the simulated precipitation fields. The biases in the simulated surface air temperature fields were much less than the precipitation biases. Simulation of precipitation over the Ethiopian highlands, where the topography is most pronounced within the ENB, showed a significant positive bias.

Simulating precipitation was much more sensitive to the change in physics parameterization compared to the surface air temperature. Precipitation was most sensitive to changing the cumulus parameterization and the planetary boundary layer schemes, and least sensitive to changing the microphysics scheme. Modifying the longwave radiation scheme leds to more significant changes compared to the shortwave radiation scheme. The recommended WRF configuration for the ENB consists of a CORDEX-MENA parent domain with a higher resolution nested domain extending from 25.5° E to 42.0° E and 2.5° N to 24.0° N with a mesh size of 10 km. The recommended physics parameterizations include NOAH for the land

Fig. 10 a Relative enhancement in simulating the surface air temperature by using the recommend WRF configuration (#15) over the default set of physics parameterizations (#1) during the wet season. b Relative enhancement in simulating precipitation by using the recommend WRF configuration (#15) over the default set of physics parameterizations (#1) during the wet season

surface scheme, CAM for the longwave and shortwave radiation, the BMJ for the cumulus scheme, MYJ for the planetary boundary layer schemes, and WSM6 for the cloud microphysics. The recommended physics parameterization combination enhanced the simulated precipitation and surface air temperature over the ENB by an average of 47.5 and 4%, respectively, compared to the WRF default physics parameterizations.

References

Ahmed KF, Wang G, Silander J, Wilson AM, Allen JM, Horton R, Anyah R (2013) Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. northeast. Glob Planet Chang. https://doi.org/10.1016/j.gloplacha.2012.11.003

Argent R, Sun X, Semazzi F, Xie L, Liu B (2015) The Development of a Customization Framework for the WRF Model over the Lake Victoria Basin, Eastern Africa on Seasonal Timescales. Adv Meteorol. https://doi.org/10.1155/2015/653473

Benestad RE (2008) Empirical-Statistical Downscaling. World Scientific Publishing Co., Singapore. https://doi.org/10.1142/6908

Chen F, Mitchell K, Schaake J, Xue Y, Pan HL, Koren V, Duan QY, Ek M, Betts A (1996) Modeling of land surface evaporation by four schemes and comparison with FIFE observations. J Geophys Res 101(D3):7251–7268. https://doi.org/10.1029/95JD02165

Collins WD, Rasch PJ, Boville BA et al (2004) Description of the NCAR Community Atmosphere Model (CAM3.0). NCAR Technical Note TN-464 + STR, National Center of Atmosphere Research

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597. https://doi.org/10.1002/qj.828

Dudhia J (1989) Numerical study of convection observed during the winter monsoon experiment using a mesoscale two–dimensional model. J Atmos Sci 46(20):3077–3107. https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2

Dudhia J (2014) A history of mesoscale model development. Asia-Pac J Atmos Sci 50(1):121–131. https://doi.org/10.1007/s13143-014-0031-8

Flaounas E, Bastin S, Janicot S (2011) Regional climate modelling of the 2006 West African monsoon: sensitivity to convection and planetary boundary layer parameterisation using WRF. Clim Dyn 36(5–6):1083–1105

García-Díez M, Fernández J, Fita L, Yagüe C (2013) Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. Q J R Meteorol Soc 139(671):501–514. https://doi.org/10.1002/qj.1976

Grell GA (1993) Prognostic evaluation of assumptions used by cumulus parameterizations. Mon Weather Rev 121(3):764–787. https://doi.org/10.1175/1520-0493(1993)121<0764:PEOAUB>2.0.CO;2

Grell GA, Devenyi D (2002) A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. Geophys Res Lett 29(14):38-1–38-4. https://doi.org/10.1029/2002GL015311

Hay LE, Clark MP (2003) Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. J Hydrol 282(1–4):56–75

Hay LE, Wilby RL, Leavesley GH (2000) A comparison of delta change and downscaling GCM scenarios for three mountainous basins in the United States. J Am Water Resour Assoc 36(2):387–397. https://doi.org/10.1111/j.1752-1688.2000.tb04276.x

Hong SY, Lim JOJ (2006) The WRF single–moment 6–class microphysics scheme (WSM6). J Korean Meteor Soc 42(2):129–151

Hong SY, Dudhia J, Chen SH (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. Mon Weather Rev 132(1):103–120. https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2

Hong SY, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev 134(9):2318–2341. https://doi.org/10.1175/MWR3199.1

Janjic ZI (1994) The step–mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure

schemes. Mon Weather Rev 122(5):927–945. https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

Jones PG, Thornton PK (2013) Generating downscaled weather data from a suite of climate models for agricultural modelling applications. Agric Syst 114:1–5. https://doi.org/10.1016/j.agsy.2012.08.002

Kain JS (2004) The Kain–Fritsch convective parameterization: an update. J Appl Meteorol 43(1):170–181. https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Legates DR, McCabe GJ (1999) Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation. Water Resour Res 35(1):233–241. https://doi.org/10.1029/1998WR900018

Lin YL, Richard DF, Harold DO (1983) Bulk parameterization of the snow field in a cloud model. J Clim Appl Meteorol 22(6):1065–1092. https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2

Michalakes J, Chen S, Dudhia J, Hart L, Klemp J, Middlecoff J, Skamarock W (2001) Development of a next generation regional weather research and forecast model, developments in teracomputing: proceedings of the ninth ECMWF workshop on the use of high performance computing in meteorology, vol 1. World Scientific, Singapore, pp 269–276

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res: Atmos 102(D14):16,663–16,682. https://doi.org/10.1029/97JD00237

Mooney PA, Mulligan FJ, Fealy R (2013) Evaluation of the sensitivity of the weather research and forecasting model to parameterization schemes for regional climates of Europe over the period 1990–95. J Clim 26(3):1002–1017. https://doi.org/10.1175/JCLI-D-11-00676.1

NOAA (2001) National Oceanic and Atmospheric Administration changes to the NCEP Meso Eta analysis and forecast system: increase in resolution, new cloud microphysics, modified precipitation assimilation, modified 3DVAR analysis. http://www.emc.ncep.noaa.gov/mmb/mmbpll/eta12tpb/

Pleim JE (2007) A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: model description and testing. J Appl Meteorol Climatol 46(9):1383–1395. https://doi.org/10.1175/JAM2539.1

Ruiz J, Saulo JC, Nogués-Paegle J (2010) WRF model sensitivity to choice of parameterization over South America: validation against surface variables. Mon Weather Rev 138(8):3342–3355. https://doi.org/10.1175/2010MWR3358.1

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011) GPCC full data reanalysis version 6.0 at 0.5: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. doi:https://doi.org/10.5676/DWD_GPCC/FD_M_V6_050

Semenov EM (2012) LARS-WG a stochastic weather generator for use in climate impact studies. User's manual, Version 3.0. http://www.rothamsted.ac.uk/mas-models/download/LARS-WG-Manual.pdf

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Wang W, Powers JG (2008) A description of the Advanced Research WRF Version 3. NCAR Technical Note TN-464 + STR, National Center of Atmosphere Research

Tao WK, Joanne S, Michael M (1989) An Ice–Water Saturation Adjustment. Mon Weather Rev 117(1):231–235. https://doi.org/10.1175/1520-0493(1989)117<0231:AIWSA>2.0.CO;2

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res. https://doi.org/10.1029/2000JD900719

Tewari M, et al (2004) Implementation and verification of the unified NOAH land surface model in the WRF model. 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction (formerly paper number 17.5) pp. 11–15. https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm

United Nations Framework Convention on Climate Change (2013) Statistical Downscaling Model. Compendium of methods and tools to evaluate impacts of, and vulnerability and adaptation to, climate change. http://unfccc.int/adaptation/nairobi_work_programme/knowledge_resources_and_publications/items/5487.php

Wilby RL, Dawson CW (2013) The statistical downscaling model: insights from one decade of application. Int J Climatol 33(7):1707–1719. https://doi.org/10.1002/joc.3544

Wilby RL, Troni J, Biot Y, Tedd L, Hewitson BC, Smith DM, Sutton RT (2009) A review of climate risk information for adaptation and development planning. Int J Climatol 29(9):1193–1215. https://doi.org/10.1002/joc.1839

Willmott CJ (1981) On the validation of models. Phys Geogr 2(2):184–194

Yin C (2011) Applications of self-organizing maps to statistical downscaling of major regional climate variables. Dissertation, Environmental Science, University of Waikato

Zittis G, Hadjinicolaou P, Lelieveld J (2014) Comparison of WRF model physics parameterizations over the MENA-CORDEX domain. Am J Clim Chang 3(05):490–511. https://doi.org/10.4236/ajcc.2014.35042

Zorita E, Von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. J Clim 12(8):2474–2489. https://doi.org/10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2