Meteorologische Zeitschrift, Vol. 22, No. 3, 297–316 (October 2013)

Open Access Article

Evaluation of daily maximum and minimum 2-m temperatures as simulated with the Regional Climate Model COSMO-CLM over Africa

Stefan Krähenmann[1,*], Steffen Kothe[1], Hans-Jürgen Panitz[2] and Bodo Ahrens[1]

[1]Institute for Atmospheric and Environmental Sciences, Goethe University of Frankfurt
[2]Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology

(Manuscript received December 24, 2012; in revised form March 17, 2013; accepted April 5, 2013)

Abstract

The representation of the diurnal 2-m temperature cycle is challenging because of the many processes involved, particularly land-atmosphere interactions. This study examines the ability of the regional climate model COSMO-CLM (version 4.8) to capture the statistics of daily maximum and minimum 2-m temperatures (Tmin/Tmax) over Africa. The simulations are carried out at two different horizontal grid-spacings (0.22° and 0.44°), and are driven by ECMWF ERA-Interim reanalyses as near-perfect lateral boundary conditions. As evaluation reference, a high-resolution gridded dataset of daily maximum and minimum temperatures (Tmin/Tmax) for Africa (covering the period 2008–2010) is created using the regression-kriging-regression-kriging (RKRK) algorithm. RKRK applies, among other predictors, the remotely sensed predictors land surface temperature and cloud cover to compensate for the missing information about the temperature pattern due to the low station density over Africa. This dataset allows the evaluation of temperature characteristics like the frequencies of Tmin/Tmax, the diurnal temperature range, and the 90th percentile of Tmax.

Although the large-scale patterns of temperature are reproduced well, COSMO-CLM shows significant under- and overestimation of temperature at regional scales. The hemispheric summers are generally too warm and the day-to-day temperature variability is overestimated over northern and southern extra-tropical Africa. The average diurnal temperature range is underestimated by about 2°C across arid areas, yet overestimated by around 2°C over the African tropics. An evaluation based on frequency distributions shows good model performance for simulated Tmin (the simulated frequency distributions capture more than 80% of the observed ones), but less well performance for Tmax (capture below 70%). Further, over wide parts of Africa a too large fraction of daily Tmax values exceeds the observed 90th percentile of Tmax, particularly across the African tropics. Thus, the representation of processes controlling Tmax including cloud-solar interaction, radiation processes, and ground heat fluxes should be improved by further model developments. The higher-resolution simulation (0.22°) is on average about 0.5°C warmer with a more pronounced overestimation of the higher percentiles of Tmax, and yields no clear benefit over the lower-resolution simulation.

Keywords: Maximum temperature, minimum temperature, daily variability, Regional Climate Model, African climate, frequency disribution.

1 Introduction

There are numerous RCM evaluation studies that have focused on mid-latitudes, but only a few studies have investigated the entire African continent (e.g. Anyah and Semazzi, 2007; Druyan et al., 2008; Sylla et al., 2010; Nikulin et al., 2012). Nikulin et al. (2012) evaluated ten European Centre for Medium-Range Weather Forecasts (ECMWF) (ERA-Interim; Simmons et al., 2006; Uppala et al., 2008; Berrisford et al., 2009) driven RCMs and their ensemble average for simulating precipitation over Africa at a spatial resolution of 0.44°. They concluded that the multi-model average outperformed individual models, and should therefore be preferred. Several studies have investigated African climate on a regional basis, focusing on the western African monsoon flow (e.g. Abiodun et al., 2008; Afiesimama et al., 2006; Kothe et al., 2013; Steiner et al., 2009), on the variability of eastern African rainfall (e.g. Anyah and Semazzi, 2007; Davis et al., 2009; Segele et al., 2009), and on the climate variability over South Africa (e.g. Kgatuke et al., 2008). Sylla et al. (2010) showed that RegCM3 is able to realistically simulate the seasonal mean of zonal wind profile, temperature, precipitation, and associated low-level circulations. However, they found an underestimation of temperature over complex terrain of Guinea and an overestimation over the African tropics, which was responsible for a too strong West African monsoon flow and an increased intensity of the African Jet core

*Corresponding author: Stefan Krähenmann, Goethe Universität Frankfurt, Campus Riedberg, Altenhöferallee 1, 60438 Frankfurt,
e-mail: kraehenmann@ iau.uni-frankfurt.de

DOI 10.1127/0941-2948/2013/0468
0941-2948/2013/0468 $ 9.00

intensity. Nevertheless, they concluded that the model performance over the entire continent was of sufficient quality to be applied for climate change studies.

KOTHE et al. (2013) suggested that the RCM COSMO-CLM (CCLM; http://www.clm-community.eu, e.g. ROCKEL et al., 2008) struggles to correctly simulate the location and the inter-annual variability of the West African monsoon (WAM). They found a strong warm bias over the whole Sahara region of up to 6°C in comparison with Climate Research Unit (CRU) data (MITCHELL and JONES, 2005). They conclude that the shift of the WAM is mainly driven by the warm bias in the Sahara, which triggers an intensification of the Sahara heat low and an increased temperature gradient between the ocean and land surface. The absence of realistically simulated variability of rainfall and the strong warm bias over the Sahel introduce doubt on the reliability of climate predictions for that region.

In this study we analysed simulations by the newer CCLM version 4.8 for the entire African continent, carried out as part of the Coordinated Regional climate Downscaling Experiment-CORDEX (GIORGI et al., 2009; JONES et al., 2011). The simulations were performed at two different horizontal resolutions (0.22°: CLM-0.22 and 0.44°: CLM-0.44), and cover the time period 1989–2010. The simulations were driven at the lateral boundaries and at sea-surface with ERA-Interim reanalyses data following the so-called "perfect-boundary" approach. This approach allows isolating deficiencies in the RCM without the complexity of biases inherited from a forcing global climate model.

Primarily, the ability of the model to reproduce annual means and annual cycles of Tmin and Tmax as well as the diurnal temperature range were evaluated to detect major model biases. The diurnal temperature range is of particular interest since it is a useful measure of the counteracting effects of longwave and shortwave forcings in a climate model (MAKOWSKI et al., 2008) that are mainly controlled by water vapour, soil moisture and cloudiness (STECHIKOV and ROBOCK, 1995; DAI et al., 1999; STONE and WEAVER 2003). However, since averaging can hide model deficiencies or systematic errors inherent in daily data (e.g. KHARIN and ZWIERS, 2000), the mean diurnal temperature range (as provided in the CRU dataset) alone is insufficient for an evaluation of the entire data distribution. Even if a climate model reproduces well the observed mean it may not capture well other attributes (e.g. SCHAEFFER et al., 2005). Yet changes in the tails of the distributions are more likely to affect humans and agriculture than changes in the mean. There are indications that extremes became more frequent and intensified their amplitudes (e.g. increasing number of days exceeding the 90th percentile threshold) in recent years (BÖHM et al., 2001; KUNKEL et al., 2008). That is why extremes are more important in climate impact research. There are critical temperature values above or below which human life and agriculture may suffer damages. Very low temperatures affect several aspects of agriculture, such as survival, photosynthesis, growth and yield of crops. Very high temperatures may severely hinder plant growth through dehydration, disruption of photosynthesis and respiration, and also may kill seedlings. Therefore, the question arises: what is CCLM's performance simulating the distributions of daily Tmin and Tmax.

High-quality observational data at a suitable spatio-temporal resolution are required for the model evaluations. To the authors knowledge no such datasets are available for the entire African continent. Therefore, an observational reference (hereafter referred to as RKRK dataset) was generated using the regression-kriging-regression-kriging (RKRK) gridding algorithm, described in KRAEHENMANN and AHRENS (2013). The method applies the remotely sensed predictors land surface temperature and cloud cover and has been shown to perform strongly in data-sparse regions. The gridding algorithm involves a two-step interpolation process involving both regression and kriging. Due to limitations in satellite data availability the evaluations were restricted to the time period of 2008–2010.

A description of the CCLM model is given in section 2. The study area, the observational reference datasets, the ERA-Interim reanalysis and the statistical analysis methods used for the evaluation process are described in section 3. Section 4 presents the evaluation of the RKRK dataset and the analysis of the model performance for Africa, section 5 provides a discussion of the results and section 6 presents the main conclusions of this study.

2 Model and observational data

2.1 COSMO-CLM model description and setup

The three-dimensional non-hydrostatic regional climate model COSMO-CLM (CCLM) is the climate version of the operational weather forecast model COSMO (Consortium for Small-scale Modelling; http://cosmo-model.org; STEPPELER et al., 2003; BALDAUF et al., 2011) of the German and other Meteorological Services. It prognostically solves the compressible equations for wind, temperature, pressure, specific humidity, cloud water, and cloud ice content as well as rain, snow, and optionally graupel. The CCLM was first described by BÖHM et al. (2006). A variety of CCLM applications are gathered in a special issue of Meteorologische Zeitschrift (ROCKEL et al., 2008). Earlier versions of CCLM have been successfully applied to Africa (e.g. KOTHE and AHRENS, 2010; PAETH et al., 2011, KOTHE et al. (2013).

The equations are solved on an Arakawa-C grid (ARAKAWA and LAMB, 1977) defined on a rotated geographical coordinate system. In the vertical a hybrid coordinate is used. Close to the surface the numerical layers follow the terrain, and near the top of the model domain they are flat. The model physics contains a cloud scheme

(DOMS et al., 2011) with prognostic precipitation and four hydrometeor species (cloud droplets, raindrops, cloud ice, and snowflakes), a mass flux scheme for convection (TIEDTKE, 1989), and a delta-two-stream radiation scheme as described by RITTER and GELEYN (1992) that is invoked every hour. In the standard CCLM version, land surface processes are parameterized through the soil-surface module TERRA_ML (GRASSELT et al., 2008), constituting the lower boundary conditions to the atmospheric model by the provision of energy and water fluxes at the surface. Detailed descriptions of the dynamics, numerical methods and physical parameterizations in the model can be found in the documentations of the model (e.g. DOMS, 2011; DOMS et al, 2011 http://www.clm-community.eu).

The here evaluated Coordinated Regional climate Downscaling Experiment (CORDEX; GIORGI et al., 2009) Africa simulations were performed using CCLM version 4.8, which was recently evaluated for Europe (KEULER et al., 2012). The model configuration has been chosen based on a series of sensitivity studies, described briefly, along with the physical packages used, in PANITZ et al. (2012, in preparation). The model domain is identical to that demanded by CORDEX with an additional sponge zone of 10 grid points at each side, in which boundary data are impressed on the model. For the first phase of CORDEX a horizontal grid resolution of 0.44° is demanded. The horizontal model domain has a size of 214 grid-points from West to East and 221 grid-points from South to North, including the sponge zone. Additional simulations done with higher grid resolution of 0.22° use the same domain, so the number of grid points increases to 427 x 441. The vertical coordinate with 35 levels is identical for both cases, with the upper most level at 30 km above Sea Level. The settings of all model parameters are identical for both simulations, with the exception of the number of grid points and the numerical time step. A Runge-Kutta integration scheme has been used with time steps of 240 s for the low-resolution simulation and 120 s for the high-resolution simulation. Both simulations were driven by the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (SIMMONS et al., 2006; UPPALA et al., 2008; DEE et al., 2011). Transient simulations were carried out for the period 1989–2010.

2.2 ERA-Interim

ERA-Interim is the latest European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis product. It improves previous reanalysis data such as ERA-40 (UPPALA et al., 2005) in horizontal resolution and in using an updated atmospheric model and updated assimilation system. The reanalysis is carried out using a 4D-variational analysis on a spectral grid with triangular truncation of 22 waves (T255, approximately 80 km), a hybrid vertical coordinate system consisting of 60 vertical layers, and a 3 hourly temporal resolution. The assimilated observational data originate from diverse sources, including weather stations, radio soundings, ship measurements, and satellite observations. ERA-Interim spans the time period from January 1979 onwards, and continues to be extended forward in near-real time.

Reanalysis products such as ERA-Interim are known to exhibit biases over the African continent (TRENBERTH and CARON, 2001; DIONGUE et al., 2002; TADROSS et al., 2006), which are transmitted to the RCMs (WANG et al., 2004; SYLLA et al., 2010). Therefore, in this study the ERA-Interim reanalysis (UPPALA et al., 2008; SIMMONS et al., 2006; BERRISFORD et al., 2009; DEE et al., 2011) was not only used as the lateral boundary driving fields of the CCLM simulation, but was also included in the evaluation. For evaluation purposes the ERA-Interim reanalysis was interpolated to the grids of CLM-0.22 and CLM-0.44 using inverse distance weighting and a height correction with a constant lapse rate of 0.65°C/100m.

2.3 RKRK dataset

2.3.1 Daily gridded Tmin/Tmax

The only available set of global gridded daily temperature observations is the Hadley Global Historical Climate Network Dataset (HadGHCND; available via http://www.metoffice.gov.uk/hadobs), which spans from 1950 to the present, but has very poor spatial coverage (only partial coverage of Africa) and poor spatial resolution (2.5° latitude x 3.75° longitude). To provide a comprehensive observational reference of daily Tmin and Tmax for the evaluation of the CCLM simulated temperature fields, we created a gridded observational dataset (hereafter referred to as RKRK dataset) covering the period 2008–2010 for two spatial resolutions to match the grids used in CLM-0.22 and CLM-0.44. We therefore adopted the RKRK gridding method described by KRAEHENMANN and AHRENS (2013). They compared a number of gridding methods for daily Tmin and Tmax, and concluded that RKRK, which makes use of satellite-derived predictors, performed best in data sparse areas (here, RKRK was applied to 338 synoptic weather stations across the African continent and surrounding areas). We only provide a short summary here including a discussion of the study area, and the station and predictors data used for the gridding.

2.3.2 Study area and station data

Africa is unique in the sense that it is the only continent that almost symmetrically spans both sides of the equator, and therefore exhibits a highly variable climate with both northern and southern hemispheric climatic influences (ODADA and OLAGO, 2005). Africa stretches about 8000 km in length from Tunisia in the north to the southern tip of South Africa, and encompasses more than 20%

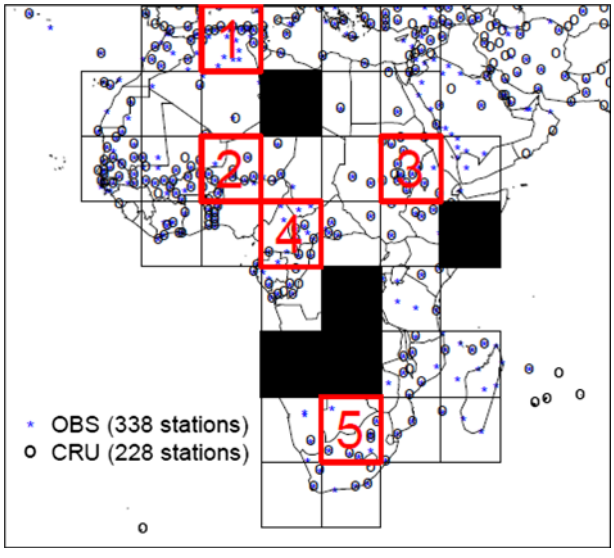

Figure 1: Locations of observing stations as used in the gridded observations (blue asterisks) and in CRU (black circles). The average number of African stations available per day in January 2009 is given in brackets. The rectangles represent the 10° regions over Africa. Regions containing fewer than two observing stations are blacked. Squares in red show those regions selected for the analysis of frequency distributions.

of the global land area. The topography is characterized by complex terrains in the southern (e.g. Table mountains, Lesotho) and eastern regions (e.g. Tanzania), with mountains higher than 1500 meters, and more localized highlands in the African tropics (e.g. Cameroon Mountains, Jos Plateau in Nigeria, and the Guinea Highlands). Owing to its location, vast size, and variable terrain features, Africa comprises eight climate zones according to the Köppen climate classification (SANDERSON, 1999). Northern Africa is strongly influenced by atmospheric subsidence due to the subtropical high-pressure system and therefore is primarily desert or arid, while the central and southern areas mainly consist of savannah and rainforests.

For the gridding we used synoptic weather stations (SYNOP) across the African continent (about 230 stations) and surrounding areas (about 100 stations). Fig. 1 illustrates the poor station coverage across Africa, with the location of these stations being unevenly distributed. The station coverage is most dense along the coastlines of the Mediterranean Sea, Western Africa, and South Africa and lowest in tropical areas. SYNOP data are available in near real time, but do not undergo a high level quality control. To remove erroneous values we excluded daily anomalies of Tmin and Tmax that deviated by more than five standard deviations from the spatially averaged daily anomalies (for each climate region separately).

Evaluations of the CCLM simulations were carried out with respect to 10° x 10° regions (see Fig. 1) and to African climate regions (shown in Fig. 2f). Because the RKRK dataset is of low quality in regions with very low station coverage, the CCLM simulations were only evaluated in 10° regions containing at least two observing stations (e.g. blacked regions in Fig. 1 not evaluated). The distance to the next observing station is therefore smaller than about 1000 km, which is approximately the variogram range (or de-correlation length) used for the gridding process. For the analysis of simulated frequency distributions of daily Tmin/Tmax we selected five 10° regions in different climate regions where the station number is higher than 10 (squares in red, see Fig. 1). Note that only land points were used for the evaluations.

2.3.3 Predictors data

The gridding process includes two linear regression steps in which the target variables (firstly the monthly averaged daily Tmin and Tmax, secondly the daily anomalies of Tmin/Tmax) were fitted against multiple predictors. We used elevation, continentality index (GORCZYNSKI, 1920), climatological zonal mean temperature, and land surface temperature for the regression of monthly averaged daily Tmin and Tmax values, and cloud cover for the regression of daily Tmin/Tmax anomalies (see Fig. 2).

Mean monthly CRU TS3.1 data (CRU: Climatic Research Unit; http://www.cru.uea.ac.uk; MITCHELL and JONES, 2005) for the period 1961–1990 have been used to derive the continentality index and the climatological zonal mean temperature (see also KRAEHENMANN et al., 2011). Land surface temperature and cloud cover were retrieved using measurements of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiometer mounted on the Meteosat Second Generation (MSG) platform. MSG is a geostationary satellite that allows high-time-resolution retrievals of land surface temperature (every 15 minutes) and cloud cover (every hour). The spatial resolution is 3 km at nadir (e.g. the point of the earth directly below the satellite) for both datasets. For cloud cover (obtained from Analysis Satellite Applications Facility Nowcasting and Very Short Range Forecasting (SAF NWC); available via http://www.nwcsaf.org/HD/MainNS.jsp) we determined averages for two time intervals, one in the morning (0:00 – 6:00 UTC), and one in the afternoon (10:00 – 16:00 UTC). Land surface temperature was obtained from Land Surface Analysis (LSA SAF; http://landsaf.meteo.pt/). Daily maximum and minimum land surface temperature was derived using the same time intervals as for cloud cover. For the gridding process we used monthly averaged daily maximum and minimum land surface temperature because daily values contain data voids where clouds prevent the retrieval of land surface temperature. For consistency we adopted the digital elevation model (DEM) data available from the LSA SAF homepage (to which land surface temperature data were geo-referenced). The original data provider of the DEM was the United States Geological Survey (USGS;

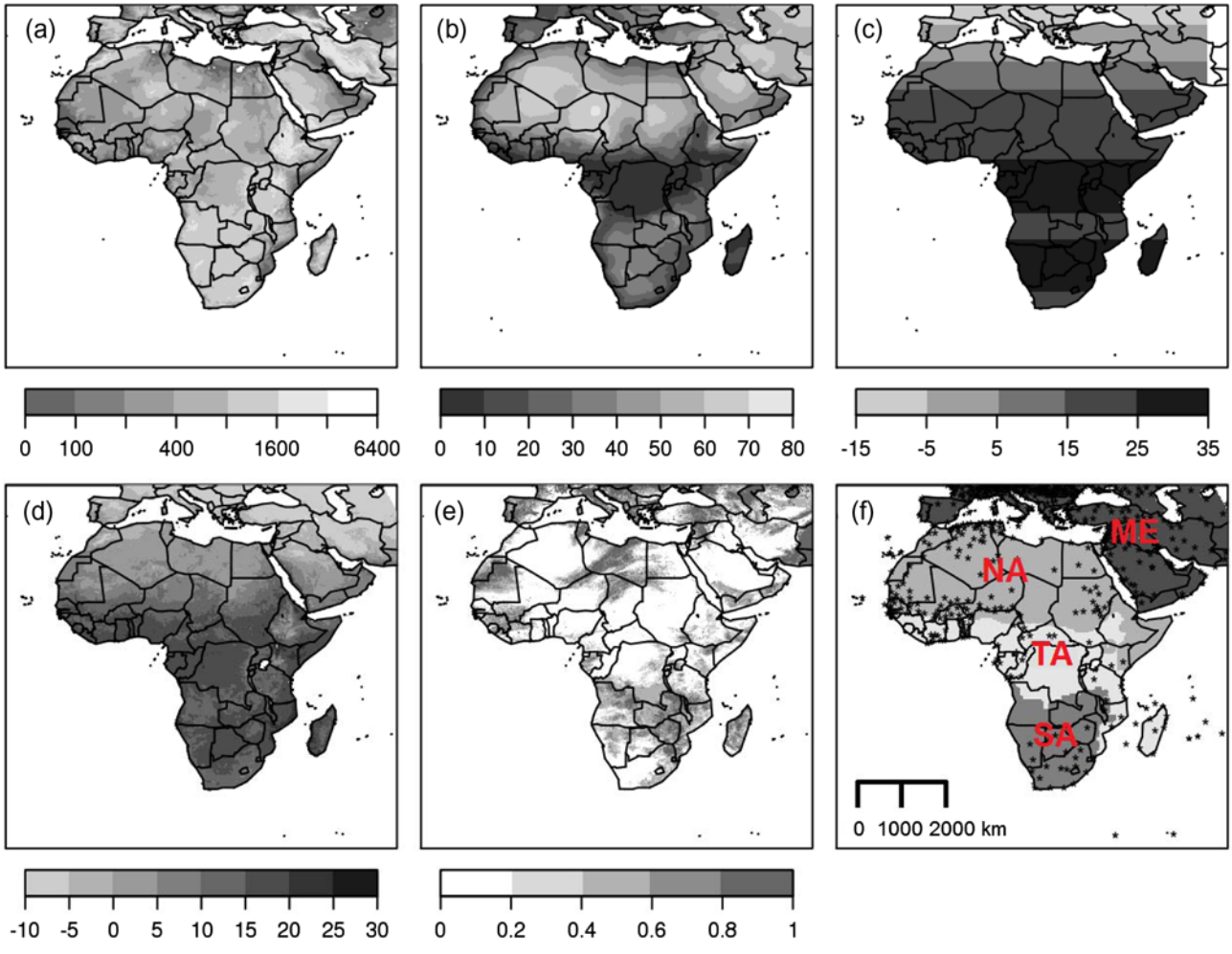

Figure 2: Values of the predictors (a) elevation [m], (b) continentality index [1], (c) climatological zonal monthly mean temperature (here for January), [°C], (d) land surface temperature, (here January 2009) [°C], and (e) cloud cover [1] for 1 January 2009. Locations of observing stations for 1 January 2009 and the climate regions ME, NA, TA and SA are shown in (f).

data can be obtained via http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30_info).

2.3.4 Gridding method

The gridding of daily 2-m maximum and minimum temperatures (Tmax/Tmin) was performed using the RKRK gridding algorithm, which was described in KRAEHENMANN and AHRENS (2013). While KRAEHENMANN and AHRENS (2013) used point kriging, the present study applied the spatially more representative block kriging (e.g. DEUTSCH and JOURNEL, 1998). Block kriging allows for estimation of grid values that are representative for areal values and not for point values (as is the case for standard kriging methods), which is preferable if the dataset is to be used for climate model evaluation.

Here the RKRK method is used for the interpolation of daily anomalies: first, block RK is applied to monthly averaged daily Tmin and Tmax values to determine the underlying spatial structure of the data; second, block RK is applied to daily anomalies with regard to the monthly mean. Multiple linear regression is used to estimate the spatial variation explained by the predictors, and kriging to account for the non-explainable variation. To ensure the data were normally distributed, a normal score transformation (DEUTSCH and JOURNEL, 1998) was applied to the regression residuals (of the monthly and daily data, respectively) prior to each kriging step.

Both regression and kriging involve solving a set of linear equations to minimize the mean squared error of the residuals from the interpolating surface. This least squares problem assumes that the data being interpolated are homogeneous in space. However, this is not the case for the stations across such a large and complex region as Africa with its many climate zones. We addressed this problem by conducting the gridding steps in four overlapping subregions (Fig. 2f), which were determined by merging Köppen-Geiger climate zones (SANDERSON, 1999) so that the number of available stations per region is large for robust regression. The subregions are: Mediterranean (ME), Northern Africa (NA), Tropical Africa (TA), and Southern Africa (SA). Finally, the subregional maps were merged by linearly weighting across an

overlapping area of 1000 km, to yield the daily prediction maps for Tmin and Tmax.

2.4 CRU dataset

Since the satellite data based RKRK dataset spans only a short time period (2008–2010), the Climate Research Unit (CRU, University of East Anglia, UK) monthly mean global gridded dataset version TS 3.1 was used as a further reference for the evaluation of monthly mean Tmin/Tmax for the time period 1990–2009. A reference for CRU TS 3.1 is in preparation (data available via http://www.cru.uea.ac.uk). However, the production of the dataset is very similar to that of TS 2.1 (MITCHELL and JONES, 2005). The dataset consists of nine climate variables, including monthly averages of daily maximum and minimum temperatures at a 0.5° x 0.5° grid resolution, covers the global land surface (excluding Antarctica) and spans the period 1901–2009.

The gridding of this dataset was based upon more than 4000 weather stations around the world (about 230 stations over the African continent, see Fig. 1). Monthly anomalies relative to the 1961–1990 normal were used for interpolation to make it less vulnerable to fluctuations in station coverage (MITCHELL and JONES, 2005). The normals were interpolated using three-dimensional thin-plate splines depending on latitude, longitude and elevation (NEW et al., 1999). For the interpolation of the anomalies angular distance weighting was applied (NEW et al., 2000), using the eight nearest stations. The station weights were determined using the correlation decay distance (JONES et al., 1997), and the angular separation of the data points from the grid point. Angular distance weighting does not account for elevation, however, NEW et al. (1999) argued that the spatial variation of monthly temperature anomalies (to the climate normals) is mainly determined by large-scale circulation patterns and relatively independent of elevation. The monthly climate data were finally obtained by applying the gridded anomalies to the mean monthly climatology of 1961–1990. For the present analysis the CRU dataset was interpolated onto the native grid of CLM-0.22 and CLM-0.44 applying inverse distance weighting and a height correction with a constant lapse rate of 0.65°C/100m.

3 Statistical analysis methods

3.1 Statistics used to quantify the interpolation performances of the gridding algorithms

The evaluation statistics for the gridding algorithms were determined using cross validation (WACKERNAGEL, 2003). As the goal was to re-estimate point observations in the evaluation of the gridding method these results were based on point interpolation, and not block interpolation. Although for block kriging the semivariogram calculates the covariance between blocks instead of points, a particular interpolation algorithm will be superior regardless whether block or point kriging is applied.

The root mean square error (RMSE) measures the deviation from the observed value, and is based on squared errors. A perfect score is indicated by a value of 0. To illustrate the performance of the interpolation method in an area (e.g. the climate regions) we used the $RMSE_A$. The $RMSE_A$ was calculated for each day from the daily prediction errors in the area and then averaged over the considered time period (e.g. 2008–2010). It is defined by:

$$\mathrm{RMSE_A} = \frac{1}{n_d}\sum_{d=1}^{n_d}\left(\sqrt{\frac{1}{n_s}\sum\nolimits_{s=1}^{n_s}\left(y_{s,d}-o_{s,d}\right)^2}\right),$$
$$s = 1\cdots n_s, d = 1\cdots n_d \qquad (3.1)$$

where $y_{s,d}$ is the predicted value and $o_{s,d}$ is the observed value at station s on day d, n_d is the number of days of the considered time period, and n_s is the number of the observing stations in the target area.

The $VARI_A$ is defined as the ratio of the variance of predictions $\sigma_d^2(y)$ to the observed variance $\sigma_d^2(o)$ of station locations in the area for each day d, averaged over the complete evaluation period:

$$\mathrm{VARI_A} = \frac{1}{n_d}\sum_{d=1}^{n_d}\left(\frac{\sigma_d^2(y)}{\sigma_d^2(o)}\right), d = 1\cdots n_d \qquad (3.2)$$

The $VARI_A$ was used to evaluate the degree of smoothing by the tested gridding algorithms over the target area. A score value close to 1 indicates that a high percentage of observed variability is retained.

3.2 Statistics used to evaluate the CCLM simulations

The $BIAS_{i,j}$ is the difference between two datasets averaged over n_d at location s:

$$\mathrm{BIAS_S} = \frac{1}{n_d}\sum_{d=1}^{n_d}\left(y_{s,d}-o_{s,d}\right), d = 1\cdots n_d \qquad (3.3)$$

The temporal standard deviation at location s is denoted σ_s. Prior to calculating σ_s the datasets were de-trended subtracting a mean annual cycle (derived from the period 2008–2010 and smoothed using an 11-day running mean). An 11-day running mean was selected after comparing preliminary results using 5-day to 31-day means.

$VARI_d$ is the ratio of simulated $\sigma_s(y)$ and observed $\sigma_s(o)$ temporal standard deviations of daily Tmin/Tmax at locations s:

$$\mathrm{VARI_d} = \frac{\sigma_s(y)}{\sigma_s(o)} \qquad (3.4)$$

A perfect agreement between observed and modelled σ_s yields a score value of 1. Score values smaller than 1 indicate an underestimation of σ_s in the model, the reverse applies for score values larger than 1.

The Taylor diagram (TAYLOR, 2001) was used to analyse the statistical relationship between the simulated fields and the reference field. The statistics were derived for daily anomalies from the mean annual cycles of the fields (derived from the period 2008–2010 and smoothed by an 11-day running mean). Note that time series of region-wise (e.g. NA, TA and SA) averages of daily Tmin/Tmax were analysed (e.g. y_A and o_A are the spatial averages of the predictions and the observations, respectively).

The Taylor diagram displays the standard deviations of the evaluated ($\sigma_A(y)$) and the reference ($\sigma_A(o)$) time series, the correlation coefficient R_A and the centred RMSE ($CRMSE_A$). The $CRMSE_A$ is equivalent to the $RMSE_A$, apart from subtracting the mean value from the predicted and observed values prior to its calculation:

$$\mathrm{CRMSE_A} = \sqrt{\frac{1}{n_d}\sum_{d=1}^{n}\left(\left(y_{A,d}-\bar{y}_A\right)-\left(o_{A,d}-\bar{o}_A\right)\right)^2}, \quad d = 1\cdots n_d \tag{3.5}$$

where the overall mean of the time series is indicated by an overbar. $CRMSE_A$ and $\sigma_A(y)$ were normalized dividing by $\sigma_A(o)$ (i.e. $CRMSE_A^{no} = CRMSE_A/\sigma_A(o)$, $\sigma_A(y)^{no} = \sigma_A(y)/\sigma_A(o)$, and $\sigma_A(o)^{no} = 1$, where $\sigma_A(y)^{no}$ is the normalized standard deviation of the evaluated time series, $CRMSE_A^{no}$ is the normalized $CRMSE_A$, and $\sigma_A(o)^{no}$ is the normalized reference standard deviation, which is by definition 1). This allows to display statistics of different time series on the same plot.

To measure the similarity between modelled and observed frequency distributions of daily Tmin/Tmax we adopted the skill score S_{FD} proposed by PERKINS et al. (2007). S_{FD} measures the mutual area between two frequency distributions by summing the minimum value of the two frequency distributions of each binned value, i.e.

$$\mathrm{S_{FD}} = \frac{1}{n}\sum_{i=1}^{n}\min\left(Z_m, Z_o\right) \tag{3.6}$$

where n is the number of bins used in the frequency distribution, and Z_m and Z_o are the respective frequencies of a value in a given bin for the model and the observational reference. A perfect agreement between observed and modelled frequency distributions yields a skill score of 1 and a poor agreement with negligible overlap of the frequency distributions yields a score close to zero.

For both the observations and the simulations we define a percentile function q_p of the frequency distribution *FD(x)* of daily Tmin/Tmax (over the period 2008–2010), which returns a Tmin or Tmax value at the boundary of a given percentile level p, by

$$\mathrm{q_P} = \min\left\{x \in R : FD(x) \geq p\right\}, p \in (0,1) \tag{3.7}$$

where *FD(x)* is a right continuous and increasing function of x that strictly increases from 0 to 1. The analysis is performed in five 10° regions in different climate regions with at least ten observing stations (see Fig. 1).

The average diurnal temperature range (ADTR) was calculated from daily Tmin/Tmax, averaged over n_d (e.g. the number of days over the period 2008–2010) at location s, i.e.:

$$\mathrm{ADTR_S} = \frac{1}{n_d}\sum_{d=1}^{n_d}\left(Tmax_{s,d} - Tmin_{s,d}\right), d = 1\cdots n_d \tag{3.8}$$

4 Results

This section starts with an evaluation of the RKRK gridding algorithm used to generate the RKRK dataset. Next, we present the analysis of the ERA-Interim driven CCLM simulations for Africa (CLM-0.22 and CLM-0.44), first in terms of mean bias of Tmin/Tmax and the diurnal temperature range, and then in terms of the model's ability to simulate frequency distributions and the temporal variability of daily Tmin and Tmax values, and the 90th percentile of Tmax. Since the accuracy of RCM simulations depends to some extent on the quality of the driving data, and in particular on how well the driving data represent observed circulation patterns for the region of interest (JACOB et al., 2007), we also evaluated the ERA-Interim reanalysis.

4.1 Evaluation of the RKRK algorithm

4.1.1 Regression analysis

We evaluated the performance of the selected predictor set, (see section 3.1.2) using multiple linear regression of monthly (daily) temperature by means of the $RMSE_A$. The $RMSE_A$ was calculated from residuals of the regression between the selected predictors and the synoptic weather stations inside 10° regions and averaged for the period 2008–2010 (the corresponding Fig. S1 is provided as supplementary online material[1]). For monthly regression (using the predictors zonal mean temperature, continentality index and land surface temperature) the $RMSE_A$ values were smallest in western and southern Africa (1 to 1.5°C). A marked annual $RMSE_A$ cycle was found over northern Africa (amplitude of ~1.5°C), with the highest values (~3°C) in summer. We attribute this to the lower reliability of land surface temperature observations for very hot surfaces - the saturation temperature (around 330 K) of the channels used for the

[1]Available on the journals website at www.schweizerbart.de/journals/metz after material.

Table 1: Average daily $RMSE_A$ for four gridding algorithms in summer (JJA) and winter (DJF) for the period 2008–2010 for Tmin and Tmax. The $RMSE_A$ is calculated region-wise (climate regions NA, TA, SA) and then averaged over all subregions.

	DJF				JJA				Average	
	Tmin		Tmax		Tmin		Tmax			
	$RMSE_A$	$VARI_A$	$RMSE_A$	$VARI_A$	$RMSE_A$	$VARI_A$	$RMSE_A$	$VARI_A$	$RMSE_A$	$VARI_A$
RKRK	2.4	0.8	2.2	0.8	2.2	0.8	2.4	0.8	2.3	0.8
RKK	2.4	0.7	2.3	0.7	2.3	0.8	2.5	0.7	2.4	0.7
OK	3.3	0.5	3.2	0.5	3.4	0.4	3.5	0.5	3.3	0.5
IDW	3.5	0.5	3.4	0.5	3.6	0.4	3.8	0.4	3.6	0.4

retrieval is frequently exceeded by Tmax in arid areas in summer (PINHEIRO et al., 2004). $RMSE_A$ values were in most cases similar for Tmin and Tmax, except in summer over northern Africa ($RMSE_A$ values were ~1°C lower for Tmin) for the abovementioned reasons.

For daily regression (cloud cover was used as predictor for the regression of daily anomalies) both Tmin and Tmax showed almost identical $RMSE_A$ values all over Africa. The lowest $RMSE_A$ values (~ 2°C) occurred over western Africa, where the station network is relatively dense. Surprisingly, we found considerably higher $RMSE_A$ values of about 3°C over northern Africa, despite a relatively dense station network. This may be related to terrain complexity and the proximity to the Mediterranean Sea. Over most of Africa the $RMSE_A$ was relatively constant throughout the year (amplitude < 0.6°C). However, in the Northeast we found considerably higher $RMSE_A$ values during winter (rising from 2°C to 3°C).

4.1.2 Cross validation of RKRK and comparison with simpler algorithms

The performance of the RKRK algorithm was evaluated using cross-validation (e.g. re-estimation of point-wise Tmin/Tmax observations leaving out each synoptic weather station once in turn) and compared to regression kriging kriging (RKK, see also KRAEHENMANN et al., 2011; KRAEHENMANN and AHRENS, 2013) and two simpler interpolation methods, namely inverse distance weighting (IDW) and ordinary kriging (OK) of daily observations. The RKK algorithm involves two major steps: regression kriging (RK) of monthly Tmin and Tmax values (using elevation, continentality index, climatological zonal mean temperature, and land surface temperature as predictors for regression of monthly mean data), followed by simple kriging of the daily anomaly.

Table 1 displays the results, where $RMSE_A$ and $VARI_A$ are given as averages of daily scores. The scores were determined separately for the climate regions NA, TA and SA and then averaged. The hybrid gridding algorithms RKK and RKRK clearly outperformed the simpler algorithms in terms of both $RMSE_A$ and $VARI_A$. The values of both score values were fairly constant over the year and differed negligibly between Tmin and Tmax (not shown). These findings are consistent with the results of the regression analysis, which indicated no annual cycle in the relationship between the predictors and target variables on the continental scale. The statistical response of using cloud cover as predictor for daily anomalies was small. However, RKRK yielded a slightly lower prediction error (smaller $RMSE_A$) and preserved a higher proportion of observed variance (higher $VARI_A$).

Despite being more complex, RKRK is more appropriate given the size and climatological diversity of the region, as well as the generally low station density. The use of cloud cover accounts for the daily cycle of differential heating and cooling processes due to clouds. In addition, the evaluations showed for RKRK a slightly better performance in preserving the observed spatial variability.

We also evaluated the predictive skills of the proposed algorithms for the 10° regions (the corresponding Fig. S2 is provided as supplementary online material[1]). The results showed very similar $RMSE_A$ patterns for all tested algorithms, with $RMSE_A$ being high in eastern Africa and the tropics and low over western Africa, yet the hybrid algorithms (RKK and RKRK) yielded considerably lower $RMSE_A$ values than both the IDW and OK (on average 2.3°C for RKRK and 3.3°C for OK). An evaluation of monthly averaged $RMSE_A$ showed negligible seasonal variations over most of Africa for the hybrid algorithms for both Tmin and Tmax. However, in western, eastern and tropical Africa the $RMSE_A$ was 0.5–1.5°C higher in June-July-August (JJA) than in December-January-February (DJF) (this was also found in the regression analysis). Meanwhile, OK and IDW yielded over eastern Africa the largest $RMSE_A$ in JJA (about 6°C). The large discrepancy in $RMSE_A$ between the algorithms during months other than JJA, with the $RMSE_A$ being around 2°C lower for the hybrid algorithms, demonstrates the strong impact of the regression on the gridding performance over eastern Africa.

4.2 Mean monthly CCLM temperature statistics

4.2.1 Mean bias in Tmin/Tmax of the CCLM simulations

Mean Tmin/Tmax fields of both CCLM simulations (CLM-0.22 and CLM-0.44) were compared to CRU for

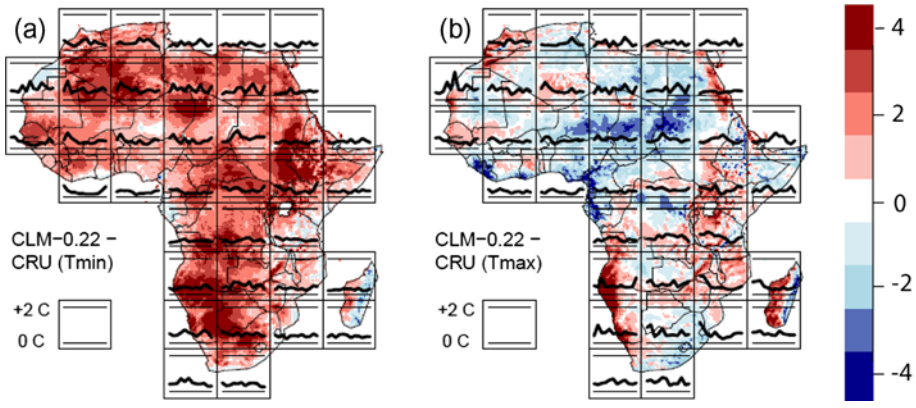

Figure 3: Mean biases for the period 1990–2009 of CLM-0.22 compared to CRU data [°C]: (a) Tmin and (b) Tmax biases. Thick lines show the annual cycles of the mean monthly bias standard deviations (the lower thin line indicates 0°C and the upper 2°C values).

the period 1990–2009 and for 17 three year long sub-periods to evaluate the information value of the sub-periods. Spatially averaged over the African continent both the complete time period and the sub-periods showed fairly similar annual bias cycles. The standard deviation of the model biases of the sub-periods was for all months smaller than 0.4°C (on average 0.2°C) and it was highest during the northern winter (DJF). The annual bias cycle had two maxima (in May and October) and two minima (in January and July), which are not easily interpreted but may be related to an erroneous simulation of the displacement of the ITCZ. CLM-0.22 was approximately unbiased relative to CRU for Tmax (annual mean difference 0°C) and positively biased for Tmin (annual mean difference around +2°C). On average over the period 1990–2009, both simulations showed consistent patterns for both Tmin and Tmax, but CLM-0.22 was around 0.5°C warmer than CLM-0.44. As a consequence the overall bias (considering both Tmin and Tmax) of CLM-0.44 was smaller than that of CLM-0.22.

Fig. 3 shows the results of evaluations of CCLM for the 10° regions. In comparison to CRU, CLM-0.22 clearly overestimated Tmin over most of Africa (*panel a*). For Tmax CLM-0.22 produced in some areas (inner Sahara, and Ivory Coast) lower and in several other areas (Madagascar, Angola, and Namibia) higher values than CRU (*panel b*). In general, CLM-0.22 and CLM-0.44 performed similarly, except that the Tmax bias in CLM 0.44 was more negative over the Sahara and the African tropics (not shown). The standard deviation of the mean monthly CCLM biases (relative to CRU) was generally between 0.5 and 1°C (extreme months with standard deviations larger than 1°C, but in respective months the bias was also large (> 5°C)). The relatively small variability in the biases of the three year long CCLM simulation sequences indicates that the evaluation shown in the next section done for daily data in the period 2008–2010 is climatologically representative.

Next, mean Tmin/Tmax fields of the simulations were evaluated against the specially generated RKRK dataset for the period 2008–2010. Inter-comparison of CCLM biases relative to both the RKRK dataset and to CRU allows to encompass uncertainties in the reference data.

Fig. 4 displays the average Tmin for the period 2008–2010, of the RKRK dataset (*panel a*) and the associated biases of CLM-0.22, CLM-0.44 and ERA-Interim (*panels b-d*). The annual cycles of the monthly mean Tmin and the biases (plotted inside the boxes) were calculated by averaging the (land only) grid square values inside the 10° regions. CLM-0.22 overestimated Tmin in most regions by 1 to 4°C (*panel b*). Only along the coasts of South Africa and Mozambique, in some parts of inner Africa (Sudan) and in Madagascar the simulation was unbiased or slightly negatively biased. CLM-0.22 was overall around 0.5°C warmer and therefore slightly more biased than CLM-0.44, yet yielded similar $BIAS_s$ patterns. ERA-Interim was on average about 1°C cooler than the RKRK dataset (*panel d*), with the strongest cold bias over the Tibesti Mountains in Chad (up to 3°C). Adversely, over northern Africa ERA-Interim was slightly warmer than the RKRK dataset.

For Tmax CLM-0.22 showed an even stronger warm $BIAS_s$ over tropical regions (Fig. 5b) than for Tmin. However, we found considerable cold $BIAS_s$ over the eastern Sahara (where there was a warm-bias for Tmin). Tmax was generally warmer by around 0.5°C in CLM-0.22 than in CLM-0.44, which was also the case for Tmin (also see section 4.2.1). The higher resolution simulation CLM-0.22 had smaller cold $BIAS_s$ than CLM-0.44 over the Sahel region, yet intensified the warm $BIAS_s$ over the tropics. The Tmax pattern in the ERA-Interim reanalysis data was very similar to that of the RKRK dataset, with temperature differences ($BIAS_s$) mostly below 2°C (*panel d*).

Although CCLM biases relative to both CRU (Fig. 3) and the RKRK dataset (Figs. 4 and 5) were comparable over northern Africa and Madagascar, they clearly differed in the African tropics and over large parts of the southern extra-tropics. These discrepancies refer to uncertainty in the reference data. Mean Tmin/Tmax fields of CRU and the RKRK dataset (period 2008–2009) differed by up to 6°C in tropical areas (not shown), due to differences in

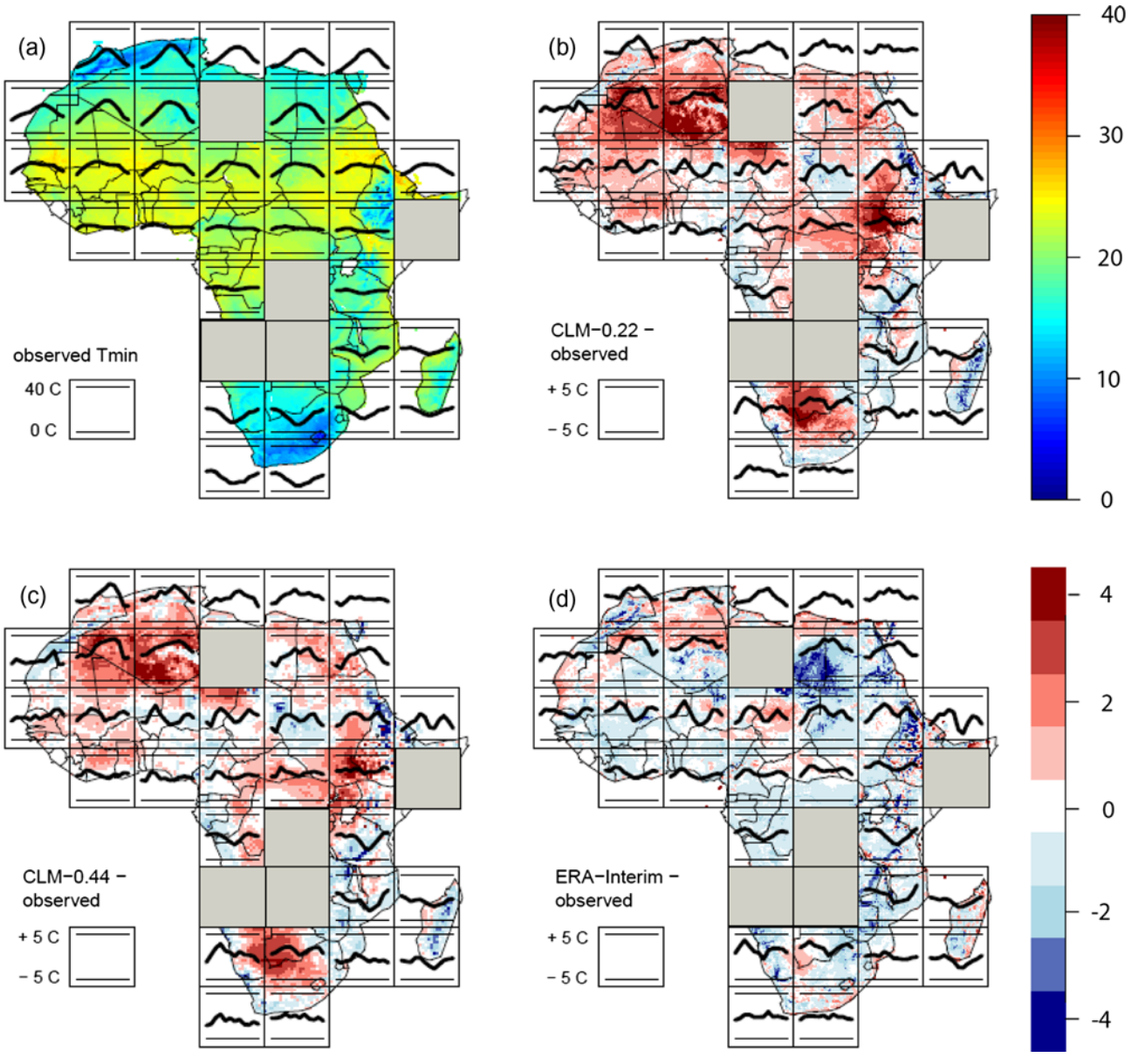

Figure 4: (a) Average daily Tmin [°C] of the RKRK dataset (colour bar top right), and (b-d) $BIAS_s$ for CLM-0.22, CLM-0.44 and ERA-Interim for the period 2008–2010 (color bar bottom right). Thick lines show the annual cycles of mean monthly Tmin (panel a) with thin lines indicating 0 and 40 °C and of the mean monthly biases (panels b-d) with lines indicating +/− 5 °C). 10° regions containing fewer than two stations were not evaluated.

station coverage (Fig. 1) and applied mapping procedures. However, the RKRK dataset bases on a larger number of observing stations than CRU (338 versus 228), applies satellite-derived predictors to compensate for the missing temperature information in data sparse regions, and has daily data coverage. Hence, the RKRK dataset provides a more useful reference than CRU.

Moreover, on average over the period 2008–2010, the evaluations revealed smaller differences between ERA-Interim and the RKRK dataset than between CCLM and the RKRK dataset, which also provides confidence into the RKRK dataset. Furthermore, this means that the stated biases of CCLM are not inherited from the global forcing but develop within the regional climate model grid.

4.2.2 Average diurnal temperature range (ADTR)

Fig. 6 displays the RKRK dataset's $ADTR_s$ (*panel a*) and the bias in $ADTR_s$ for CLM-0.22, CLM-0.44 and ERA-Interim (*panels b-d*) relative to the RKRK dataset. The thick lines inside the 10° regions indicate the annual $ADTR_s$ cycles of the RKRK dataset (*panel a*) and related biases of the validated datasets (*panels b-d*) relative to the RKRK dataset. *Panels e-f* show the bias in $ADTR_s$ for CLM-0.22 relative to ERA-Interim for the period 2008–2010 and to CRU for the period 1990–2009, respectively. *Panels e-f* were added to display biases in CLM-0.22 in the 10° regions where a lack of observing stations did not allow for an evaluation against the RKRK dataset.

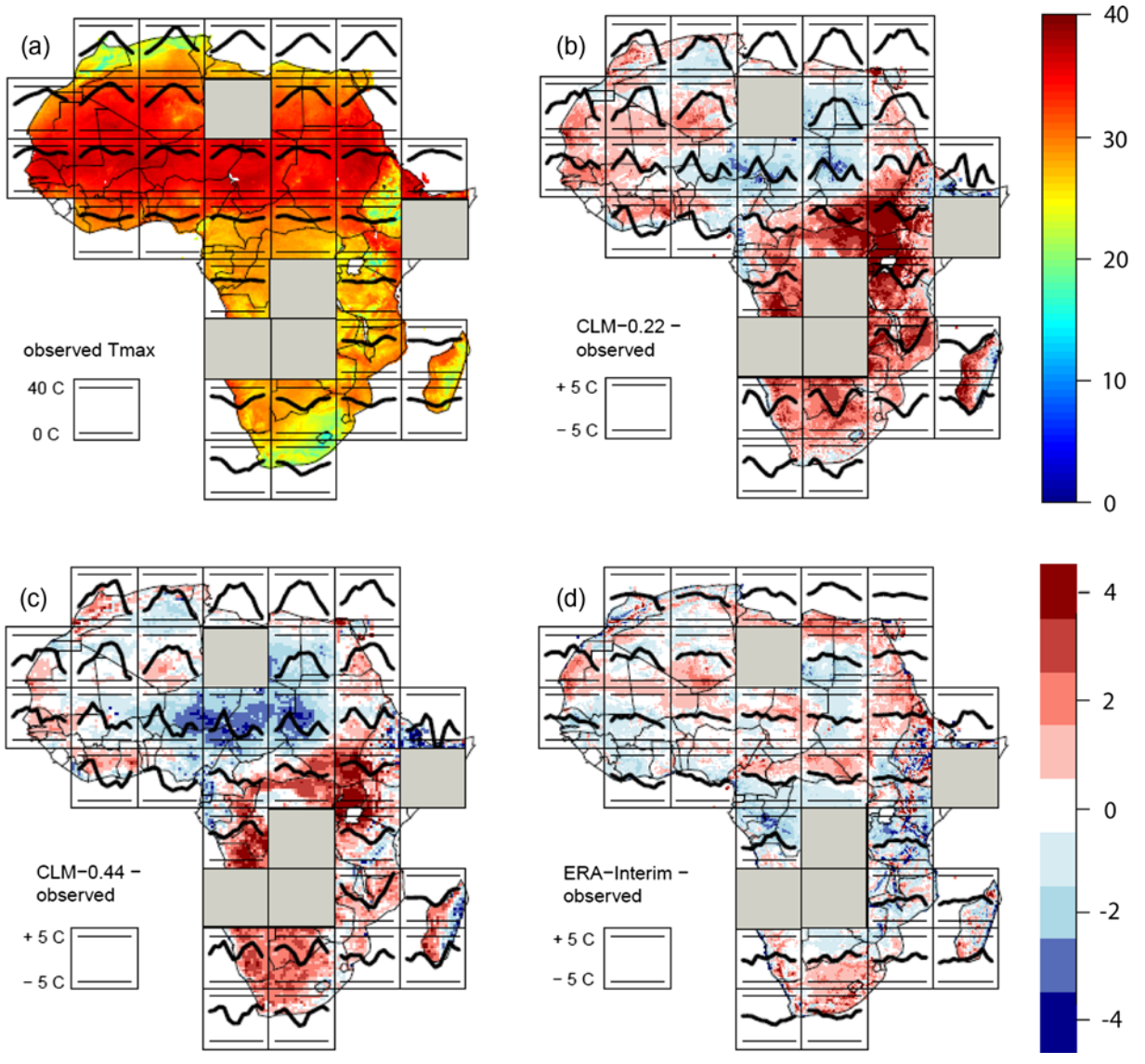

Figure 5: Same as in Figure 4, but for Tmax.

The RKRK dataset (*panel a*) indicated low $ADTR_s$ in tropical areas (around 5°C), where the humidity is high and cloudiness is predominant, and along coastlines (around 10°C), where the ocean serves as a temperature regulator. The highest $ADTR_s$ values were found in semi-arid and arid regions such as the Sahara desert and the western Kalahari desert (15-20°C). CLM-0.22 (*panel b*) considerably underestimated $ADTR_s$ over the Sahara, but overestimated $ADTR_s$ across the African tropics, particularly in Tanzania and along the African southwest coast. The $ADTR_s$ bias pattern in CLM-0.44 (*panel c*) was very similar to that in CLM-0.22, yet CLM-0.44 overestimated less strongly $ADTR_s$ over the tropics, the Ethiopian Highlands, southeast Africa, and Madagascar. Both the RKRK dataset and ERA-Interim showed small differences in $ADTR_s$ (*panel d*) in regions just north of the equator, but ERA-Interim produced particularly high $ADTR_s$ values in arid regions. CLM-0.22 showed an overall similar bias pattern in comparison to both the RKRK dataset and the ERA-Interim reanalysis. Yet, in comparison to ERA-Interim, CLM-0.22 showed a stronger positive bias over deserted areas (*panel e*). CCLM's $ADTR_s$ biases against CRU were somewhat different to those obtained relative to the RKRK dataset, with a considerable underestimation in $ADTR_s$ over most of Africa (*panel f*). This is in consistence with mean biases in Tmin and Tmax fields (see Fig. 3, *panels a* and *b*).

The RKRK dataset showed a strong annual $ADTR_s$ cycle over the transition zones (e.g. northern and southern Savannah regions), with the highest $ADTR_s$ during the dry season (Fig. 6a). Over the subsidence regions the annual $ADTR_s$ cycle was weaker and peaked during the hemispheric summers. Both CCLM and ERA-Interim considerably overestimated the annual $ADTR_s$ cycle over the African tropics. For CCLM this was related to an

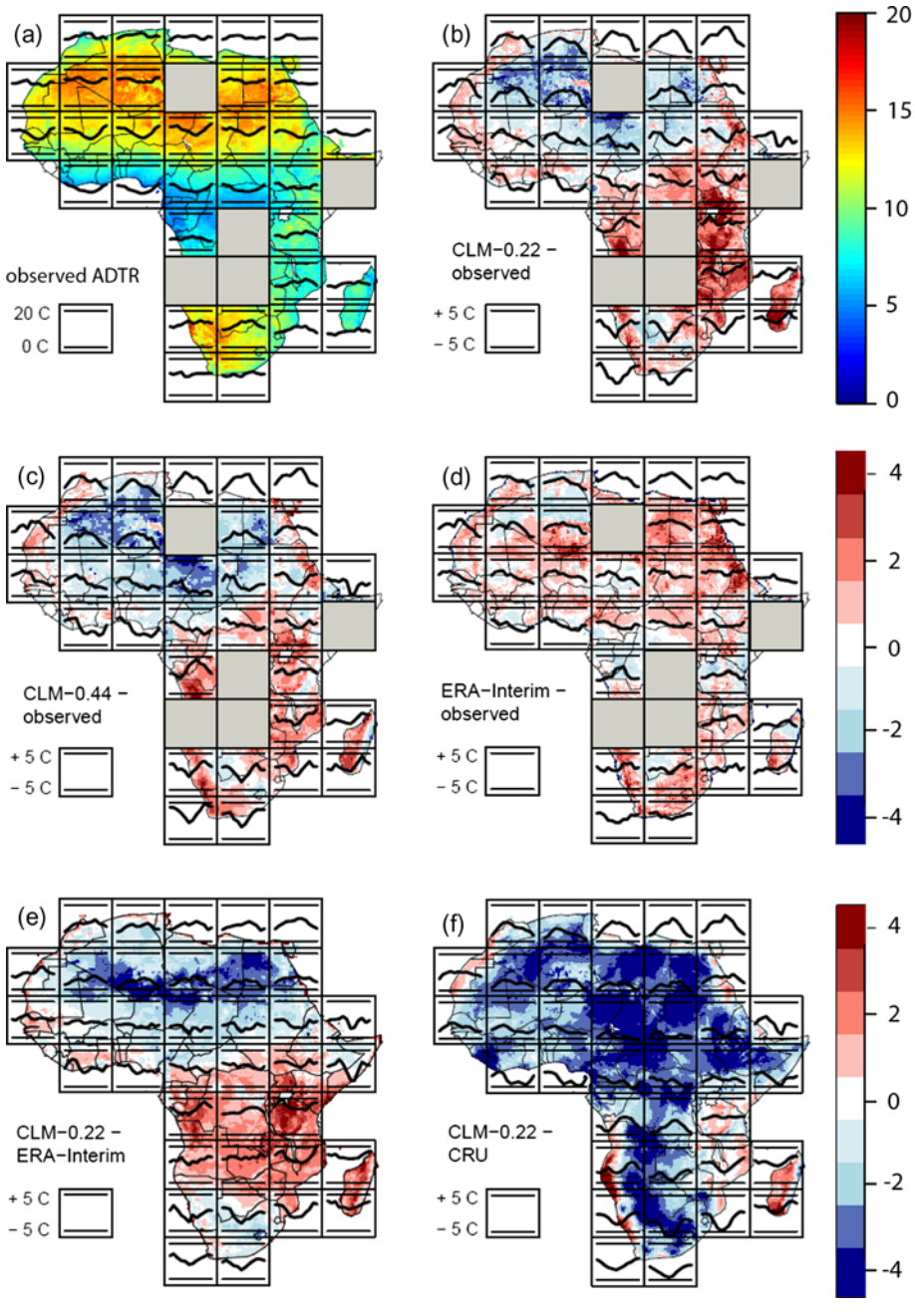

Figure 6: (a) The RKRK dataset's $ADTR_s$ [°C] (colour bar top right), (b-d) bias in $ADTR_s$ [°C] of CLM-0.22, CLM-0.44, ERA-Interim compared to the RKRK dataset, and (e-f) bias in $ADTR_s$ [°C] of CLM-0.22 compared to ERA-Interim for the period 2008–2010, and to CRU for the period 1990–2009 (color bar bottom right). Thick lines show the annual $ADTR_s$ cycles of the RKRK dataset (panel a) with thin lines indicating 0 and 20 °C and of the bias in $ADTR_s$ (panels b-f) with thin lines indicating +/− 5 °C. 10° regions containing fewer than two stations were not evaluated (panels a-d).

underestimation in Tmin (Fig. 4b) and a concurrent overestimation in Tmax (Fig. 5b) in JJA. In ERA-Interim Tmax was overestimated in JJA and underestimated in DJF (Fig. 5d), resulting in a too strong annual $ADTR_s$ cycle over the tropics (Fig. 6d).

4.3 Daily CCLM temperature statistics

In the preceding section we evaluated the ability of the model to simulate the monthly and longer term Tmin/Tmax. However, long-term averages can hide systematic errors at the daily scale. Therefore, this section focuses on the model's ability to reproduce the observed frequency distribution and the temporal variability of daily Tmin and Tmax.

4.3.1 Frequency distributions

The simulated frequency distributions of Tmin/Tmax allow assessing the model's capacity to simulate the

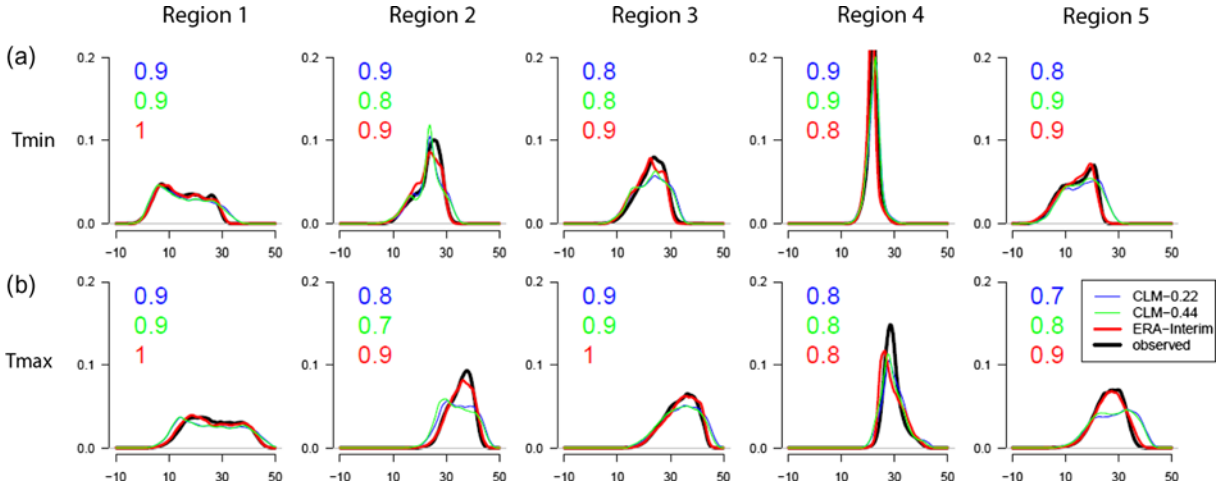

Figure 7: Frequency distributions for regions 1–5 (see Fig. 1) for: (a) daily Tmin and (b) daily Tmax [°C]. Frequency distribution-based skill scores (S_{FD}) are given in numbers (colours correspond to the respective simulation datasets).

complete range of observed values on daily time scales. For the analysis the ERA-Interim reanalysis and the CLM-0.44 simulation were interpolated to the CLM-0.22 grid using inverse distance weighting and a height correction with a constant lapse rate of 0.65°C/100m. For the RKRK dataset, the ERA-Interim reanalysis, and both model simulations, all daily Tmin/Tmax land only data within a given 10° region were used to derive the frequency distributions. The data were binned between −10°C and 50°C using a bin size of 0.5°C. For the analysis we selected five 10° regions in different climate regions where the station number is higher than 10 (see Fig. 1).

Fig. 7 shows the simulated and observed frequency distributions for five selected regions (see Fig. 1) for daily Tmin (*panel a*) and Tmax (*panel b*). Also given in Fig. 7 are the frequency distribution-based skill scores S_{FD} that measure the degree of overlap between observed and modelled frequency distributions. Both the very tightly shaped frequency distributions in region 4 and the broad frequency distribution in region 1 were well reproduced by the simulations. In general, the simulated frequency distributions were very close to those of the RKRK dataset. However, the model tended to overestimate the warmer temperatures in regions 3 and 5, while it represented the probability of lower temperatures quite well. In addition, the model considerably underrepresented the observed peak in the frequency distributions for Tmax in regions 2 and 5 (Fig. 7b). On the other hand, the frequency distributions of ERA-Interim were very similar to the observed ones in region 5, and thus provide confidence into the RKRK dataset. ERA-Interim compared worst with the RKRK dataset in region 4, underrepresenting the observed peak and producing too many low values in the frequency distribution for Tmax.

The overall skill scores for the simulations exceeded 0.8 in all regions for Tmin (Fig. 7a) and were in most cases lower for Tmax (S_{FD} > 0.6 in all regions). The worst performing areas for Tmax were regions 2 and 5

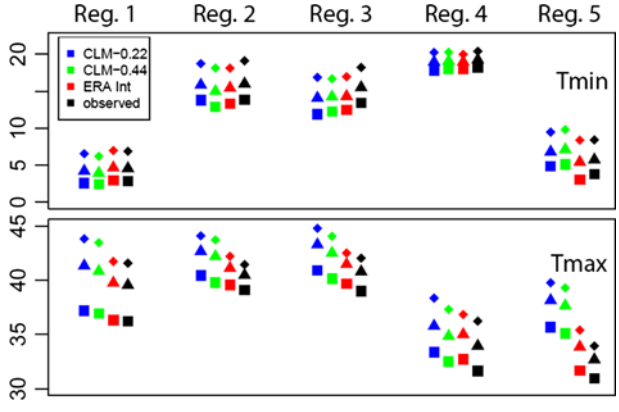

Figure 8: Percentile values for regions 1-5 (see Fig. 1) for: (a) Tmin and (b) Tmax [°C]. Squares represent the 5th (80th) percentile, triangles the 10th (90th) and diamonds the 20th (95th) for Tmin (Tmax).

with skill scores between 0.7 and 0.8. The skill scores of both model resolutions were very similar. ERA-Interim performed overall best (S_{FD} > 0.8), and was only outperformed by the CCLM simulations in region 4.

Fig. 8 allows a closer inspection of extreme percentiles (the 5th, 10th and 20th percentiles for Tmin and the 80th, 90th and 95th percentiles for Tmax). In four of the five regions (e.g. regions 1 to 4) both models were within ± 1°C of the observed for Tmin percentiles (*panel a*). There was a general trend in both models to overestimate the higher percentiles for Tmax by 1–4°C (*panel b*). The overall largest discrepancies were found in region 5 for Tmax and also to a minor extend for Tmin as it can also be seen in Fig. 7. The evaluations further revealed relatively small differences between the two CCLM simulations. The model performed equally well for the lower percentiles of Tmin at both resolutions, while CLM-0.22 was about 1°C warmer than CLM-0.44 for the higher percentiles of Tmax. As a consequence the overall bias of CLM-0.22 was slightly larger than that of CLM-0.44.

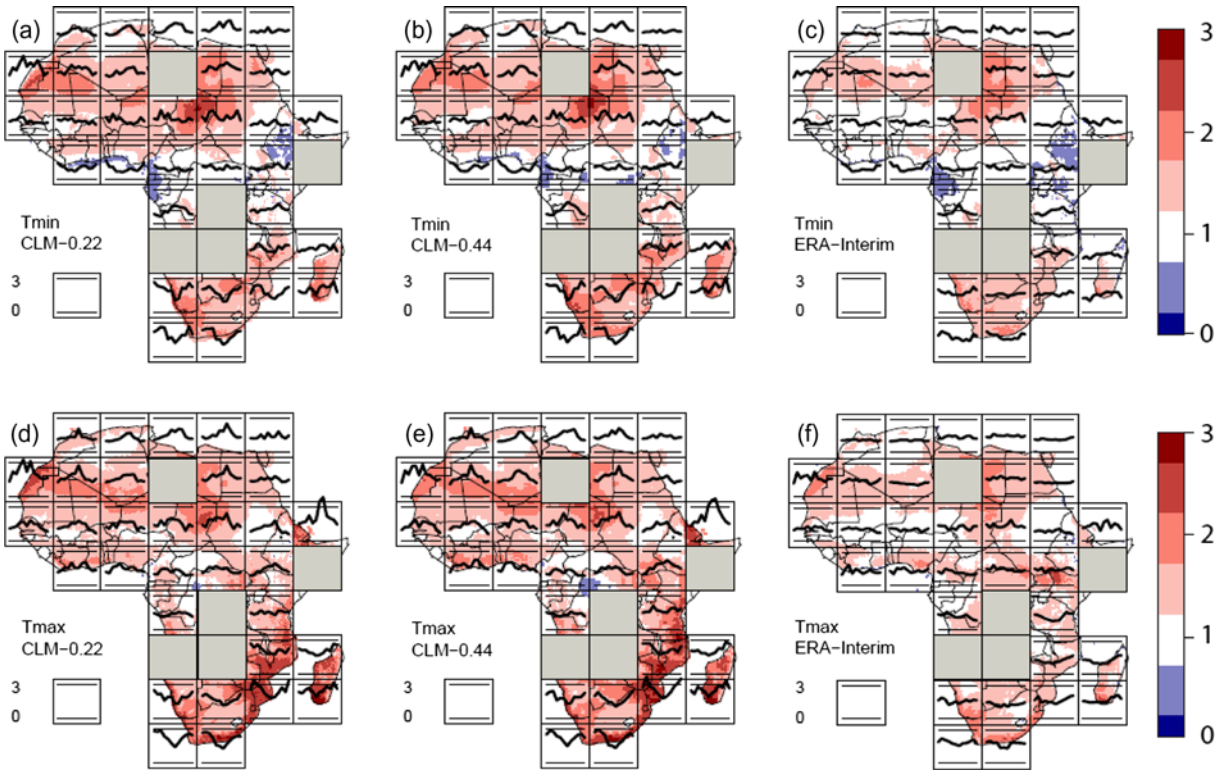

Figure 9: $VARI_d$ [-] for (a and d) CLM-0.22, (b and e) CLM-0.44 and (c and f) ERA-Interim reanalysis. Panels a-c display $VARI_d$ for daily anomalies in Tmin and d-f in Tmax. The annual cycle of $VARI_d$ was calculated from land points only inside the 10° regions and is given by thick lines inside the boxes, lower thin lines indicate 0 times and upper thin lines 3 times the observed temporal standard deviation. 10° regions containing fewer than two stations were not evaluated.

4.3.2 Temporal variability of Tmin/Tmax (σ_s)

Fig. 9 displays the ratio ($VARI_d$) of simulated and observed temporal standard deviation (σ_s) for CLM-0.22 (*panels a and d*), CLM-0.44 (*panels b and e*), and ERA-Interim (*panels c and f*) for daily anomalies in Tmin (*panels a-c*) and in Tmax (*panels d-f*). Thick lines inside the 10° regions indicate the annual cycle of $VARI_d$ (e.g. ratio for σ_s of monthly time series).

The RKRK dataset showed the largest temporal standard deviation (2-3°C) in areas of continental climate (high continentality index) and particularly small (<1°C) over the African tropics and in coastal regions where the seasonal temperature variability is reduced by the maritime effect (not shown). In tropical areas $VARI_d$ was about 1 in both CCLM simulations. Otherwise, the simulations considerably overestimated the temporal standard deviation for both Tmin and Tmax ($VARI_d$~2-3), particularly over desert areas (Sahara and Kalahari desert). Also ERA-Interim showed $VARI_d$ values greater than 1 in arid areas, yet the temporal standard deviation was less strongly overestimated than in CCLM.

The RKRK dataset indicated very similar annual cycles of the temporal standard deviation for Tmin and Tmax and the amplitude of the cycle increased towards higher latitudes (not shown). Over both the northern and southern extra-tropical regions in Africa the temporal standard deviation was about 2-3°C during the hemispheric winters and about 1°C during the hemispheric summers. Over the northern and southern Savannas off the coasts, the temporal standard deviation was largest (2°C) during the transitional seasons (spring and autumn). The modelled annual cycles of temporal standard deviations of Tmin/Tmax were generally not too different to those of the RKRK dataset. However, over the Sahel region and the southern tropics CCLM clearly overestimated the temporal standard deviation for both Tmin and Tmax, particularly in the hemispheric summers ($VARI_d$ up to 3). ERA-Interim captured the RKRK dataset's annual cycle of the temporal standard deviation generally well, yet it was slightly overestimated during the hemispheric summer months in arid areas ($VARI_d$~1.5, see Fig. 9f).

The Taylor diagram (TAYLOR, 2001) in Fig. 10 summarizes the statistics of the anomaly time series (daily Tmin/Tmax relative to smoothed mean annual cycles) for the period 2008–2010 in three climate regions for the RKRK dataset, CLM-0.22, CLM-0.44, and ERA-Interim.

Both CCLM simulations agreed well with the standard deviation $\sigma_A(o)^{no}$ of Tmax in the climate region TA depicted by the RKRK dataset, but underestimated $\sigma_A(o)^{no}$ of Tmin in the region TA and overestimated

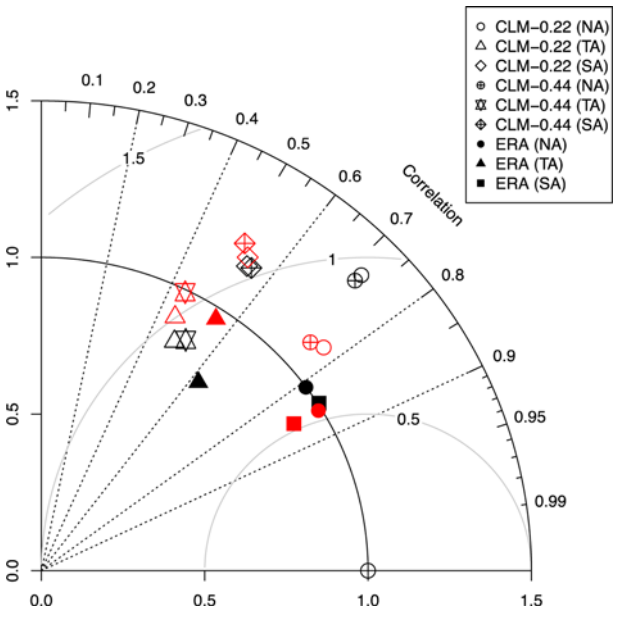

Figure 10: Taylor diagram showing the temporal patterns of daily Tmin/Tmax anomalies in three climate regions for the period 2008-2010 for the RKRK dataset, CLM-0.22, CLM-0.44, and ERA-Interim. Results for Tmin are given in black and for Tmax in red.

$\sigma_A(o)^{no}$ of Tmin and Tmax in the regions NA and SA ($\sigma_A(y)^{no}$ 1.2-1.4). This is in consistence with results presented in Fig. 9 that depict an overestimation in σ_s over the African extra-tropics for CCLM ($VARI_d$~2). The simulations captured the temporal evolution of daily Tmin/Tmax anomalies well in the region NA (R_A~0.75), but suboptimal in the regions TA and SA (R_A~0.5). $CRMSE_A^{no}$ was smallest for Tmax in the region NA (~0.8), otherwise $CRMSE_A^{no}$ was about 1. The two CCLM simulations performed similarly.

The driving dataset ERA-Interim performed consistently better than the CCLM simulations. However, in terms of R_A and $CRMSE_A^{no}$ there were large differences among the climate regions. R_A ranged from 0.5 in the region TA to 0.85 in the regions NA and SA, and $CRMSE_A^{no}$ ranged from 0.9 in TA to 0.55 in NA and SA. Although $\sigma_A(y)^{no}$ was in most cases close to 1, ERA-Interim underestimated $\sigma_A(o)^{no}$ for Tmin in the region TA ($\sigma_A(y)^{no}$~0.7). The better agreement of ERA-Interim with the RKRK dataset supports our findings for mean Tmin/Tmax fields (Figs. 4 and 5 and frequency distributions of daily Tmin/Tmax (see Fig. 7).

4.3.3 90th percentiles of Tmax

Fig. 11 displays the 90th percentile (q_{90}) of daily Tmax for the period 2008–2010, of the RKRK dataset (*panel a*) and the percentage of daily Tmax values greater or equal to the observed 90th percentile (should be 10%) for CLM-0.22, CLM-0.44 and ERA-Interim (*panels b-d*). The RKRK dataset (*panel a*) showed the highest q_{90} values over the northern subtropics (about 45°C) and the lowest q_{90} values over complex terrain, the tropics and over southern Africa (about 30°C). CLM-0.22 clearly overestimated q_{90} over wide parts of Africa (*panel b*), particularly over the Mporkoso Mountains between Congo and Zambia and the mountainous regions of Tanzania where more than 50% of daily Tmax values exceeded the observed q_{90}. Only along the coastlines of Cameroon and Guinea q_{90} was underestimated. CLM-0.44 performed similarly to CLM-0.22, yet less strongly overestimated q_{90} as the lower resolution CLM-0.44 simulation was on average about 0.5°C cooler than CLM-0.22 (see also Fig. 5). The ERA-Interim reanalysis compared overall better with the RKRK dataset than CCLM (*panel d*), particularly over the African tropics. ERA-Interim, however, deviated strongly from the observed q_{90} over the Ethiopian Highlands and over the mountainous region to the east of the Gulf of Guinea, which may be related to high uncertainties in the reference datasets over these areas due to very low station numbers (see Fig. 1).

5 Discussion

Our results showed significant differences between both the high- and low-resolution COSMO-CLM (CCLM) simulations and the RKRK dataset. In the annual mean both simulations overestimated ADTR over tropical areas, particularly over Tanzania and along the African southwest coast, related to a moderate warm bias in Tmin (~1°C) and a strong warm bias in Tmax (4°C). In contrast, ADTR was underestimated over most of the Sahara, which is likely because of uncertainties in cloud cover and aerosol parameterizations (KOTHE and AHRENS, 2010) and errors in soil thermal conductivity (J.-P. Schulz, DWD, pers. com., 2012). In CCLM soil thermal conductivity is prescribed assuming semi-wet soil conditions throughout the African continent, and is thus strongly overestimated in arid zones causing an overestimation of the ground heat flux and consequently a too small ADTR in respective regions (MURRAY, and VERHOEF, 2007; J.-P. Schulz, pers. com., 2012). The same argument applies in tropical regions but reversed. Interestingly, CCLM concurrently overestimated Tmin and Tmax during the summer months in arid regions, most likely caused by a wrong representation of the cloud diurnal cycle in the model. Indeed, a total cloud cover validation study by PFEIFROTH et al. (2012) indicated a not correctly modelled cloud diurnal cycle during summer in deserted areas, with an afternoon maximum and a morning minimum. The driving ERA-Interim reanalysis agreed overall better with the RKRK dataset. Tmin was on average 1°C too cold with a strong cold anomaly over Chad (~3°C) and Tmax biases were generally smaller than 1°C, thereby ADTR biases were smaller than ±1°C.

The evaluations revealed for CCLM, and to a lesser extend for ERA-Interim, a strong overestimation of

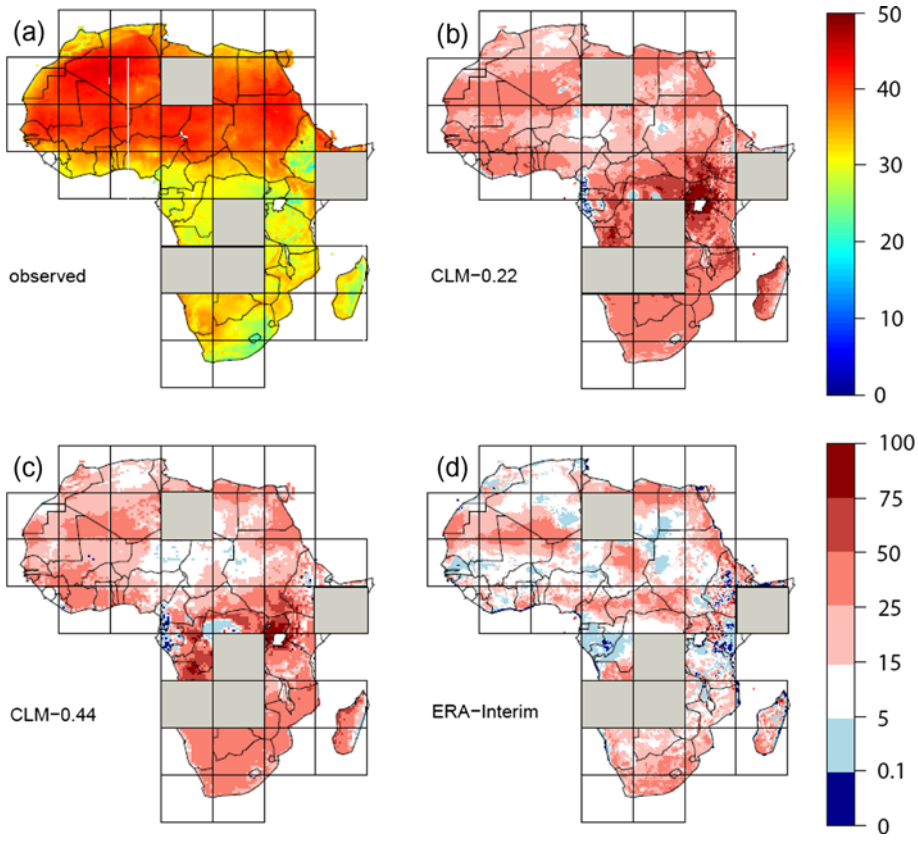

Figure 11: 90th percentile of daily Tmax [°C] of the RKRK dataset (colour bar top right), and (b-d) percentage of daily Tmax greater than or equal to the 90th percentile of the observed daily Tmax for CLM-0.22, CLM-0.44 and ERA-Interim for the period 2008–2010 (color bar bottom right). 10° regions containing fewer than two stations were not evaluated.

temporal Tmin/Tmax variability during the summer months over both the northern and southern extra-tropics ($VARI_d$ ~2–3). This can be linked with the summer drying problem (e.g. HAGEMANN et al., 2004), a well-known problem in current RCMs (BERGANT et al., 2007; CHRISTENSEN et al., 2007). Investigations of several ERA40 driven RCMs (e.g. MACHENHAUER et al., 1998; HAGEMANN et al., 2004) showed for different regions in Europe an overestimation in precipitation throughout the year except during summer. The summer drying has been associated with deficient features in the general circulation of the models where insufficient moisture is advected into the region. This leads to a too low cloud cover and thus to an overestimation in solar radiation, which eventually results in an excessive drying out of the soil due to enhanced evaporation (SCHRODIN and HEISE, 2002; SENEVIRATNE et al., 2006). However, the drier the soil becomes, the more the evaporative fluxes (that act to limit temperature variability) will decrease and the more the temperature variability will increase.

Both CCLM simulations reproduced the RKRK dataset's frequency distributions for Tmin and Tmax in five selected 10° regions reasonably well (see Fig. 7). In terms of Tmin, S_{FD} was in all regions greater than 0.8, and generally lower for Tmax (S_{FD} 0.7-0.8 in two out of five regions). An evaluation of the percentiles derived from the frequency distributions indicated high model skill in the lower percentiles for Tmin (within ± 1°C of the RKRK dataset), yet an overestimation of higher percentiles (80th to 95th) for Tmax by 2–4°C. ERA-Interim generally agreed well with the frequency distributions and the derived percentiles showed by the RKRK dataset for both Tmin and Tmax (S_{FD} > 0.8). However, over region 4 lower Tmax values were overrepresented by ERA-Interim, indicating a stronger uncertainty of the reference data in this region. The high model skill in reproducing the frequency distributions for Tmin indicates that the models are capturing the radiative cooling processes (associated with cloud cover and water vapour) generally well. The lower model skill for Tmax can be explained by the larger number of associated processes that have to be captured by the model (PERKINS et al., 2007). Besides clouds interacting with incoming solar radiation and atmospheric moisture modifying net infrared radiation, aerosols and surface albedo affect Tmax through absorption of solar radiation (KOTHE and AHRENS, 2010). Moreover, in regions 2 and 5, both located in transitional zones between non-limiting and limiting soil moisture (KOSTER et al., 2004), an erroneous modelling of soil moisture may produce flawed Tmax values through a feedback cycle with evaporation processes (COLLATZ et al., 2000).

CCLM's low skill in simulating Tmax was confirmed by an analysis of the 90th percentile of Tmax. Across wide parts of Africa q_{90} was overestimated, particularly over the African tropics where the observed q_{90} was exceeded in more than 50% of the days. However, q_{90} was underestimated in the Gulf of Guinea region, which resulted from an underestimation of Tmax in summer (Fig. 5), due to an overestimation of convective activity in CCLM (KOTHE et al., 2013). Also ERA-Interim underestimated q_{90} in the Golf of Guinea region, which complies with the overrepresentation of low Tmax values in region 4 (see Fig. 7b). However, unlike CCLM ERA-Interim underestimated q_{90} over the east African tropics, probably due to an erroneous representation of the annual precipitation pattern in ERA-Interim over the African tropics that is better represented in CCLM (see also NIKULIN et al., 2012).

Although both model simulations showed overall consistency, CLM-0.22 was on average about 0.5°C warmer in the annual means of both Tmin and Tmax, more strongly overestimated the higher percentiles for Tmax. Also the spatial variability in Tmin/Tmax was higher in CLM-0.22, which is due to increased fine-scale structures associated with better-resolved landscape and topography compared to CLM-0.44. This is consistent with a validation study by JAEGER et al. (2008) in which a comparison of two ERA40-driven CCLM simulations of differing horizontal resolutions (0.22° and 0.44°) for Europe showed similar performance by both model resolutions in terms of 2-m temperature and precipitation.

6 Summary

We present an evaluation of 2-m temperature over the African continent done with the RCM COSMO-CLM (CCLM) (version 4.8) applying two different grid-spacings (0.22° and 0.44°) and driven by the ERA-Interim reanalysis for the period 2008–2010. Performance of the model was evaluated in detail for different aspects of African 2-m temperature at daily time scales, namely: diurnal temperature range, day-to-day variability and frequency of Tmin/Tmax, and the 90th percentile of Tmax. Besides the specially generated RKRK dataset our evaluations included also the driving model, to see if CCLM biases are affected by the quality of ERA-Interim. A preliminary inter-comparison at seasonal mean time scales for the period 2008–2009 revealed that CRU and the RKRK dataset differed by as much as 6°C due to very few or no reporting stations over large parts of Africa. Yet, the RKRK dataset based on a larger number of observing stations than CRU, included satellite-derived observations to compensate for missing temperature information and has daily data coverage, and thus was considered a more useful reference than CRU.

Evaluations of the CCLM simulations detected a moderate warm bias in Tmin over the African tropics (~1°C), which coincided with a strong warm bias in Tmax (4°C), resulting in a considerable overestimation of the diurnal temperature range. In contrast, the diurnal temperature range was mainly underestimated over the Sahara, due to uncertainty in the cloud cover parameterization (KOTHE and AHRENS, 2010; PFEIFROTH et al., 2012) and soil thermal conductivity (J.-P. Schulz, pers. com., 2012). The day-to-day variability of Tmin and Tmax was overestimated, particularly during the summer over desert areas (e.g. Sahara, Kalahari). A possible explanation is the so-called summer drying problem of RCMs (HAGEMANN et al., 2004), which leads back to a wrong surface energy budget probably due to erroneous parameterizations of convection and a displacement of the general circulation. CCLM reproduced the RKRK dataset's frequency distributions for daily Tmin and Tmax with considerable skill ($S_{FD} > 0.7$). The model was able to capture the changes in the shape of the frequency distributions in different climates over Africa. Yet it overestimated the higher percentiles (80th to 95th) for Tmax by 2-4°C, while the lower percentiles (5th to 20th) of Tmin were fairly accurate (within ±1°C). The observed 90th percentile of Tmax was exceeded in more than 50% of the days in the African tropics, and underestimated over the Gulf of Guinea region. The generally lower accuracy of the simulations for Tmax can be explained by the large number of processes that have to be captured by the model. Besides the clouds and the atmospheric moisture, which are interacting with the solar radiation, also aerosols and soil albedo significantly modify Tmax through absorption of solar radiation. Although both CCLM simulations provided overall consistent results, distinct differences were found. Mean Tmin/Tmax was ~0.5°C higher in CLM-0.22, and the higher percentiles for Tmax were more strongly overestimated.

Differences between ERA-Interim and the RKRK dataset were, on average, smaller than differences between CCLM and ERA-Interim, indicating that the model biases originate to a large extent from within the model domain (e.g. JACOB et al., 2007). CCLMs biases need to be further investigated and reduced. In particular, the parameterization of aerosol and cloud radiative effects (KOTHE and AHRENS, 2010) and the soil thermal conductivity that affect the ground heat fluxes (J.-P. Schulz, pers. com., 2012) require a detailed analysis, and in general the CCLM model needs further tuning and optimization for applications on non-European domains.

Acknowledgments

Daily observations were obtained from the Deutscher Wetterdienst (DWD; German Meteorological Service, Offenbach) by Jana SCHROEDER (Goethe University Frankfurt/Main). The authors acknowledge funding from the Hessian initiative for the development of scientific and economic excellence (LOEWE) at the Biodiversity and Climate Research Centre (BiK-F), Frankfurt/Main.

This work was supported by the German Federal Ministry of Education and Research (BMBF) under grant MiKliP: DEPARTURE/01LP1129D and by DKRZ, Hamburg.

References

ABIODUN, B. J., J. M. PRUSA, W. J. GUTOWSKI, 2008: Implementation of non-hydrostatic, adaptive-grid dynamics core in CAM3. Part I: comparison of dynamics cores in aqua-planet simulations. – Clim. Dynam. **31**, 795–810.

AFIESIMAMA, E. A., J. S. PAL, B. J. ABIODUN, W. J. GUTOWSKI, A. ADEDOYIN, 2006: Simulating the West African Monsoon using RegCM3: Part 1 – Model Evaluation and Interannual variability. – Theor. Appl. Climatol. **86**, 23–37. doi: 10.1007/s00704-005-0202-8.

ANYAH, R. O., F. H. M. SEMAZZI, 2007: Variability of East African rainfall based on multiyear RegCM3 simulations. – Int. J. Climatol. **27**, 357–371.

ARAKAWA, V., V. LAMB, 1977: Computational design of the basic dynamical processes in the UCLA general circulation model. – In: Methods in Computational Physics **17**, Academic Press, 174–264.

BALDAUF, M., A. SEIFERT, J. FÖRSTNER, D. MAJEWSKI, M. RASCHENDORFER, T. REINHARDT, 2011: Operational Convective-Scale Numerical Weather Prediction with the COSMO Model: Description and Sensitivities. – Mon. Wea. Rev. **139**, 3887–3905. doi: 10.1175/MWR-D-10-05013.1.

BERGANT, K., M. BELDA, T. HALENKA, 2007: Systematic errors in the simulation of European climate (1991-1990) with RegCM3 driven by NCEP/NCAR reanalysis. – Int. J. Climatol. **27**, 455–472.

BERRISFORD, P., D. P. DEE, K. FIELDING, M. FUENTES, P. KALLBERG, S. KOBAYASHI, S. M. UPPALA, 2009: "The ERA-Interim Archive". – ERA Report Series No.1. ECMWF: Reading, UK.

BÖHM, R., I. AUER, M. BRUNETTI, M. MAUGERI, T. NANNI, W. SCHÖNER, 2001: Regional temperature variability in the European Alps 1760–1998 from homogenized instrumental time series. – Int. J. Climatol. **21**, 1779–1801.

BÖHM, U., M. KÜCKEN, W. AHRENS, A. BLOCK, D. HAUFFE, K. KEULER, B. ROCKEL, A. WILL, 2006: CLM – The climate version of LM: Brief description and long-term applications. – COSMO Newsletter **6**, 225–235, www.cosmo-model.org.

CHRISTENSEN, J. H., B. HEWITSON, A. BUSUIOC, A. CHEN, X. GAO, I. HELD, R. JONES, R. K. KOLLI, W. T. KWON, R. LAPRISE, V. MAGANA RUEDA, L. MEARNS, C. G. MENENDEZ, J. RÄISÄNEN, A. RINKE, A. SARR, P. WHETTON, 2007: Regional Climate Projections. – In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. SOLOMON, S., D. QIN, M. MANNING, Z. CHEN, M. MARQUIS, K. B. AVERYT, M. TIGNOR, H. L. MILLER (Eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

COLLATZ, G. J., L. BOUNOUA, S. O. LOS, D. A. RANDALL, I. Y. FUNG, P. J. SELLERS, 2000: A mechanism for the influence of vegetation on the response of the diurnal temperature range to a changing climate. – Geophys. Res. Lett. **27**, 3381–3384.

DAI, A., K. E. TRENBERTH, T. R. KARL, 1999: Effects of clouds, soil moisture, precipitation and water vapor on diurnal temperature range. – J. Climate **12**, 2451–2473.

DAVIS, N., J. BOWDEN, F. SEMAZZI, A. J. XIE, 2009: Customization of RegCM3 regional climate model for eastern Africa and a tropical Indian ocean domain. – J. Climate **22**, 3595–3616, doi:10.1175/2009JCI2388.1.

DEE, D. P., S. M. UPPALA, A. J. SIMMONS, P. BERRISFORD, P. POLI, S. KOBAYASHI, U. ANDRAE, M. A. BALMASEDA, G. BALSAMO, P. BAUER, P. BECHTOLD, A. C. M. BELJAARS, L. van de BERG, J. BIDLOT, N. BORMANN, C. DELSOL, R. DRAGANI, M. FUENTES, A. J. GEER, L. HAIMBERGER, S. B. HEALY, H. HERSBACH, E. V. HOLM, L. ISAKSEN, P. KAALLBERG, M. KOHLER, M. MATRICARDI, A. P. MCNALLY, B. M. MONGE-SANZ, J.-J. MORCRETTE, B.-K. PARK, C. PEUBEY, P. de ROSNAY, C. TAVOLATO, J.-N. THEPAUT, F. VITART, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. – Quart. J. Roy. Meteor. Soc. **137**, 553–597. doi: 10.1002/qj.828.

DEUTSCH, C. V., A. G. JOURNEL, 1998: GSLIB: geostatistical software library and users's guide, 2nd ed. Oxford University Press, New York.

DIONGUE, A., J. P. LAFORE, J. L. REDELSPERGER, R. ROCA, 2002: Numerical study of a Sahelian synoptic weather system: Initiation and mature stages of convection and its interactions with the large-scale dynamics. – Quart. J. Roy. Meteor. Soc. **128**, 1899–1927.

DOMS, G., 2011: A description of the nonhydrostatic regional COSMO model part 1: Dynamics and numeric. – DWD, Offenbach, Germany, 153 pp., http://www.cosmo-model.org

DOMS, G., J. FÖRSTER, E. HEISE, H.-J. HERZOG, D. MIRONOV, M. RASCHENDORFER, T. REINHARDT, B. RITTER, R. SCHRODIN, J.-P. SCHULZ, G. VOGEL, 2011: A description of the nonhydrostatic regional COSMO model part 2: Physical Parameterization. – DWD, Offenbach, Germany, 161 pp., http://www.cosmo-model.org.

DRUYAN, L. M., M. FULAKEZA, P. LONERGAN, 2008: The impact of vertical resolution on regional simulation of the west African summer monsoon. – Int. J. Climatol. **28**, 1293–1314.

GIORGI, F., C. JONES, G. R. ASRAR, 2009: Addressing climate information needs at the regional level: the CORDEX framework. – WMO Bulletin **58**, 175–183.

GORCZYNSKI, W., 1920: Sur le calcul du degré de continentalisme et son application dans la climatologie. – Geogr. Annaler. **2**, 324–331.

GRASSELT, R., D. SCHUETTEMEYER, K. WARRACH-SAGI, F. AMENT, C. SIMMER, 2008: Validation of TERRA-ML with discharge measurements. – Meteorol. Z. **17**, 763–773. doi:10.1127/0941-2948/2008/0334.

HAGEMANN, S. B., B. MACHENHAUER, O. B. CHRISTENSEN, M. DÉQUÉ, D. JACOB, R. G. JONES, P. L. VIDALE, 2004:

Evaluation of water and energy budgets in regional climate models applied over Europe. – Clim. Dynam. **23**, 547–567.

JACOB, D., L. BARRING, O. B. CHRISTENSEN, J. H. CHRISTENSEN, M. De CASTRO, M. DEQUE, F. GIORGI, S. HAGEMANN, G. LENDERINK, B. ROCKEL, E. SANCHEZ, C. SCHÄR, S. I. SENEVIRATNE, S. SOMOT, A. van ULDEN, B. van den HURK, 2007: An inter-comparison of regional climate models for Europe: model performance in present-day climate. – Clim. Change **81**, 31–52.

JAEGER, E. B., I. ANDERS, D. LÜTHI, B. ROCKEL, C. SCHÄR, S. I. SENEVIRATNE, 2008: Analysis of ERA-40-driven CLM simulations for Europe. – Meterol. Z. **17**, 349–367.

JONES, P. D., T. J. OSBORNE, K. R. BRIFFA, 1997: Estimating sampling errors in large-scale temperature averages. – J. Climate **10**, 2548–2568.

JONES, C.G., F. GIORGI, G. ASRAR, 2011: The Coordinated Regional Downscaling Experiment: CORDEX; An international downscaling link to CMIP5. – CLIVAR Exchanges No. 56, International CLIVAR Project Office, Southampton, United Kingdom, 34–40.

KHARIN, V., F. ZWIERS, 2000: Changes in the extremes in a ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. – J. Climate **13**, 3760–3788.

KEULER, K., K. RADTKE, G. GEORGIEVSKI, 2012: Summary of evaluation results for COSMO-CLM Version 4.8-clm13 (clm17): Comparison of three different configurations over Europe driven by ECMWF reanalysis data ERA40 for the period 1979–2000. – Brandenburg University of Technology (BTU), Cottbus, 45 pp., http://www.clm-community.eu.

KGATUKE, M. M., W. A. LANDMAN, A. BERAKI, M. P. MBEDZI, 2008: The internal variability of the RegCM3 over South Africa. – Int. J. Climatol. **28**, 505–520.

KOSTER, R. D., P. A. DIRMEYER, Z. GUO, G. BONAN, E. CHAN, P. COX, C. T. GORDON, S. KANAE, E. KOWALCZYK, D. LAWRENCE, P. LIU, C.-H. LU, S. MALYSHEV, B. MCAVANEY, K. MITCHELL, D. MOCKO, T. OKI, K. OLESON, A. PITMAN, Y. C. SUD, C. M. TAYLOR, D. VERSEGHY, R. VASIC, Y. XUE, T. YAMADA, 2004: Regions of coupling between soil moisture and precipitation. – Science **305**, 1138–1140.

KOTHE, S., B. AHRENS, 2010: On the radiation budget in regional climate simulations for West Africa. – J. Geophys. Res. **115**, D32120, doi:10.1029/2010JD014331.

KOTHE, S., D. LÜTHI, B. AHRENS, 2013: Analysis of the West African Monsoon system in the regional climate model COSMO-CLM. – Int. J. Climatol, DOI: 10.1002/joc.3702.

KRAEHENMANN, S., B. AHRENS, 2013: Spatial Gridding of Daily Maximum and Minimum 2-m Temperatures Supported by Satellite Observations. – Meteor. Atmos. Phys., **120**, 87–105, http://dx.doi.org/10.1007/s00703-013-0237-9.

KRAEHENMANN, S., P. BISSOLLI, J. RAPP, B. AHRENS, 2011: Spatial gridding of daily maximum and minimum temperatures. – Meteor. Atmos. Phys. **114**, 151–161.

KUNKEL, K., P.D. BROMIRSKI, H.E. BROOKS, T. CAVAZOS, A.V. DOUGLAS, D.R. EASTERLING, K. EMANUEL, P.Y. GROISMAN, G.J. HOLLAND, T.R. KNUTSON et al., 2008: Observed Changes in Weather and Climate Extremes. – In: KARL, T.R., G.A. MEEHL, C.D. MILLER, S.J. HASSOL, A.M. WAPLE, W.L. MURRAY (Eds.): Weather and Climate Extremes in a Changing Climate: Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands. – CCSP Synthesis and Assessment Product 3.3. Washington, D.C: U.S. Climate Change Science Program, 35–80.

MACHENHAUER, B., M. WINDELBAND, M. BOTZET, J.H. CHRISTENSEN, M. DTQUT, R.G. JONES, P.M. RUTI, G. VISCONTI, 1998: Validation and analysis of regional present-day climate and climate change simulations over Europe. – Max Planck-Institute for Meteorology, Report 275, Hamburg, Germany.

MAKOWSKI, K., M. WILD, A. OHMURA, 2008: Diurnal temperature range over Europe between 1950 and 2005. – Atmos. Chem. Phys. **8**, 6483–6498.

MITCHELL, T. D., P. D. JONES, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. – Int. J. Climatol. **25**, 693–712. doi:10.1002/joc.1181.

MURRAY, T., A. VERHOEF, 2007: Moving towards a more mechanistic approach in the determination of soil heat flux from remote measurements - II. Diurnal shape of soil heat flux. – Agr. Forest Meteor. **147**, 88–97. doi:10.1016/j.agrformet.2007.06.009.

NEW, M. G., M. HULME, P. D. JONES, 1999: Representing twentieths-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology. – J. Climate **12**, 829–856.

NEW, M. G., M. HULME, P. D. JONES, 2000: Representing twentieths centrury space-time climate variability. Part II: development of 1901–96 monthly grids of terrestrial surface climate. – J. Climate **13**, 2217–2238.

NIKULIN, G., C. JONES, F. GIORGI, G. ASRAR, M. BÜCHNER, R. CEREZO-MOTA, O. CHRISTENSEN, M. DÉQUÉ, J. FERNANDEZ, A. HAENSELER, E. van MEIJGAARD, P. SAMUELSSON, M. SYLLA, L. SUSHAMA, 2012: Precipitation Climatology in An Ensemble of CORDEX-Africa Regional Climate Simulations. – J. Climate **25**, 6057–6078. doi:10.1175/JCLI-D-11-00375.1.

ODADA, E., D. OLANGO, 2005: Holocene climatic, hydrological and environmental oscillations in the tropics with special reference to Africa. Climate Change and Africa, P.S. LOW (Ed). – Cambridge University Press, 3–23.

PAETH, H., N. M. J. HALL, M. A. GAERTNER, M. D. ALONSO, S. MOUMOUNI, J. POLCHER, P. M. RUTI, A. H. FINK, M. GOSSET, T. LEBEL, A. T. GAYE, D. P. ROWELL, W. MOUFOUMA-OKIA, D. JACOB, B. ROCKEL, F. GIORGI, M. RUMMUKAINEN, 2011: Progress in regional downscaling of west African precipitation. – Atmos. Sci. Lett. **12**, 75–82.

PANITZ, H.-J., P. BERG, G. SCHÄDLER, G. FOSSER, 2012: Modelling Regional Climate Change for Germany and Africa. – In: High Performance Computing in Science and Engineering '11. Transactions of the High Performance Computing Center, Stuttgart (HLRS), W.E. NAGEL, D.B. KRÖNER, M.M. RESCH (Eds.), 503–512, doi:10.1007/978-3-642-23869-7_36.

PANITZ, H.-J., A. DOSIO, M. BÜCHNER, K. KEULER, D. LÜTHI, in preparation: COSMO-CLM (C-CLM) climate simulations over CORDEX Africa Domain: Analysis of the ERA-Interim driven simulations at 0.44° and 0.22° resolution.

PERKINS, S. E., A. J. PITMAN, N. J. HOLBROOK, J. MCANENEY, 2007: Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. – J. Climate **20**, 4356–4376.

PFEIFROTH, U., R. HOLLMANN, B. AHRENS, 2012: Cloud Diurnal Cycles in Satellite Data and Regional Climate Model Simulations. – Meteorol. Z. **21**, 551–560, doi: http://dx.doi.org/10.1127/0941-2948/2012/0423.

PINHEIRO, A. C. T., J. L. PRIVETTE, R. MAHONEY, C. J. TUCKER, 2004: Directional effects in a daily AVHRR land surface temperature dataset over Africa. – IEEE Transactions on Geosciences Remote Sensing **42**, 1941–1954.

RITTER, B., J.-F. GELEYN, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. – Mon. Wea. Rev. **120**, 303–325.

ROCKEL, B., A. WILL, A. HENSE, 2008: The Regional Climate Model COSMO-CLM (C-CLM). – Meteorol. Z. **17**, 347–348. doi:10.1127/0941-2948/2008/0309.

SANDERSON, M., 1999: The classification of climates from Pythagoras to Köppen. – Bull. Amer. Meteor. Soc. **80**, 669–673.

SCHAEFFER, M., F. M. SELTEN, J. D. OPSTEEGH, 2005: Shifts in the means are not a proxy for changes in extreme winter temperatures in climate projections. – Clim. Dynam. **25**, 51–63.

SCHRODIN, E., E. HEISE, 2002: A New Multi-Layer Soil Model. – Tech. Rep. **2**, COSMO Newsletter.

SEGELE, Z. T., P. J. LAMB, L. M. LESLIE, 2009: Seasonal-to-interannual variability of Ethiopia/Horn of Africa Monsoon. Part I: associations of wavelet-filtered large-scale atmospheric circulation and global sea surface temperature. – J. Climate **22**, 3396–3421.

SENEVIRATNE, S. I., D. LÜTHI, M. LITSCHI, C. SCHÄR, 2006: Land-atmosphere coupling and climate change in Europe. – Nature **443**, 205–209.

SIMMONS, A., S. UPPALA, D. DEE, S. KOBAYASHI, 2006: ERA-interim: new ECMWF reanalysis products from 1989 onwards. – ECMWF Newsletter No. **110**, 25–35.

STECHIKOV, G. L., A. ROBOCK, 1995: Diurnal asymmetry of climatic response to increased CO_2 and aerosol: forcings and feedbacks. – J. Geophys. Res. **100**, 26211–26227.

STEINER, A. L., J. S. PAL, S. A. RAUSCHER, J. L. BELL, N. S. DIFFENBAUGH, A. BOONE, L. C. SLOAN, F. GIORGI, 2009: Land surface coupling in regional climate simulations of the West African monsoon. – Clim. Dynam. **6**, 869–892. doi:10.1007/s00382-009-0543-6.

STEPPELER, J., G. DOMS, U. SCHÄTTLER, H. BITZER, A. GASSMANN, U. DAMRATH, G. GREGORIC, 2003: Meso-gamma scale forecasts using the nonhydrostatic model LM. – Meteor. Atmos. Phys. **82**, 75–96.

STONE, D. A., A. J. WEAVER, 2003: Factors contributing to diurnal temperature range trends in twentieth and twenty-first century simulations of the CCCma coupled model. – Clim. Dynam. **20**, 435–445.

SYLLA, B. M., E. COPPOLA, L. MARIOTTI, F. GIORGI, P. M. RUTI, A. DELL'AQUILA, X. BI, 2010: Multiyear simulation of the African climate using a regional climate model (RegCM3) with the high resolution ERA-interim reanalysis. – Clim. Dynam. **35**, 231–247.

TADROSS, M. A., W. J. GUTOWSKI, B. C. HEWITSON, C. JACK, M. NEW, 2006: MM5 simulations of interannual change and the diurnal cycle of southern African regional climate. – Theor. Appl. Climatol. **86**, 63–80.

TAYLOR, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. – J. Geophys. Res. **106**, 7183–7192.

TIEDTKE, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. – Mon. Wea. Rev. **117**, 1779–1800.

TRENBERTH, K. E., J. M. CARON, 2001: Estimates of meridional atmosphere and ocean heat transports. – J. Climate **14**, 3433–3443.

UPPALA et al., 2005: The ERA-40 archive, ECMWF 2005. – Technical Report Series 17, available at www.ecmwf.int/publications/library/do/references/list/192.

UPPALA, S. M., D. P. DEE, S. KOBAYASHI, P. BERRISFORD, A. J. SIMMONS, 2008: Towards a climate adapt assimilation system: – Status update of ERA-Interim ECMWF Newsletter **115**, 12–18.

Van ULDEN, A., G. LENDERINK, B. van den HURK, E. van MEIJGAARD, 2007: Circulation statistics and climate change in Central Europe: PRUDENCE simulations and observations. – Clim. Change **81**, 179–192. doi:10.1007/s10584-006-9212-5.

WACKERNAGEL, H., 2003: Multivariate geostatistics, 3rd edn. – Springer, Berlin, 387 pp.

WANG, Y., L. R. LEUNG, J. L. MCGREGOR, D. K. LEE, W. C. WANG, Y. DING, F. KIMURA, 2004: Regional climate modeling: Progress, Challenges, and Prospects. – J. Meteor. Soc. Japan. **82**, 1599–1628.