An evaluation of CORDEX regional climate models in simulating precipitation over Southern Africa

Mxolisi E. Shongwe,1,2* Chris Lennard,3 Brant Liebmann,4 Evangelia-Anna Kalognomou,5 Lucky Ntsangwane1 and Izidine Pinto2

1 South African Weather Service, Private Bag X097, Pretoria, South Africa
2 University of Pretoria, South Africa
3 Climate Systems Analysis Group, University of Cape Town, 7945, South Africa
4 NOAA-CIRES Climate Diagnostics Center, Boulder, CO, USA
5 Laboratory of Heat Transfer and Environmental Engineering, Aristotle University, Thessaloniki, Greece

*Correspondence to: M. E. Shongwe, South African Weather Service, Private Bag X097, Pretoria, South Africa. E-mail: mxolisi.shongwe@weathersa.co.za

Abstract

This article evaluates the ability of the Coordinated Regional Downscaling Experiment (CORDEX) regional climate models (RCMs) in simulating monthly rainfall variation during the austral summer half year (October to March) over southern Africa, the timing of the rainy season and the relative frequencies of rainfall events of varying intensities. The phasing and amplitude of monthly rainfall evolution and the spatial progression of the wet season onset are well simulated by the models. Notwithstanding some systematic biases in a few models, the simulated onset and end of the rainy season and their interannual variability are highly correlated with those computed from the reference data. The strongest agreements between the reference and modelled precipitation patterns are found north of about 20°S in the vicinity of the Inter Tropical Convergence Zone. A majority of the RCMs adequately capture the reference precipitation probability density functions, with a few showing a bias towards excessive light rainfall events.

Keywords: CORDEX – Africa; regional climate model evaluation; rainfall characteristics

1. Introduction

As the Earth’s climate continues to change, the characteristics of regional precipitation and extreme events may change without necessarily being reflected in seasonal totals (Seneviratne et al., 2012; Trenberth et al., 2003). For example, changes in wet season timing may not affect calendar season totals, and changes in wet season duration may be balanced by changes in the intensity or frequency of daily precipitation during that season. Changes in the character of precipitation have important implications for a number of climate-sensitive sectors such as agriculture, forestry, water resources, ecosystem services and disaster risk management. For instance, steady, soaking, moderate rains are generally better for agriculture than the same amount of rainfall over a short period, which may result in rapid runoff and flash flooding, leaving the deeper soil layers dry. The timing and duration of the rainy season inter alia determine the planting dates and the selection of crop types. Despite their obvious importance, however, the changes in rainfall characteristics are seldom analysed in observations (e.g. Reason et al., 2005) or regional-scale climate simulations in Africa.

Changes in the scale and shape of the rainfall distribution, which may alter tail probabilities, affect the physical and natural systems more than changes in its central tendency (Easterling et al., 2000). An evaluation of the ability of climate models in simulating the entire precipitation probability distribution at regional scales (e.g. Perkins et al., 2007) while clearly warranted has not yet been done for southern Africa. The present study attempts to address these issues by assessing the ability of the Coordinated Regional Downscaling Experiment (CORDEX) regional climate models (RCMs) in capturing monthly rainfall evolution, selected rainfall characteristics and the observed rainfall probability density functions (PDFs) in their control simulations of daily rainfall events over predefined rainfall regions (Shongwe et al., 2009) in Africa south of 10°S.

Since the launch of the CORDEX Africa programme (Jones et al., 2011), a few studies have evaluated the RCM performance over parts of the continent, including southern Africa (Nikulin et al., 2012; Kalognomou et al., 2013). Monthly and seasonal data have been used to assess the RCM’s ability to simulate the main features of seasonal mean rainfall distribution and the annual rainfall cycle. Notwithstanding biases in some regions and seasons, such as a wet bias (dry bias) close to Lesotho (over northern Mozambique) during austral summer months, the models were found to adequately simulate precipitation patterns (Kalognomou et al., 2013). However, monthly and seasonal averages or totals can conceal systematic biases in the simulated climate (e.g. Tadross et al., 2005). Also, given that cumulative effects of weather events on daily time scales have a direct impact on natural systems and...
human activities, an assessment of model ability to simulate the characteristics of daily rainfall events is clearly valuable.

Over southern Africa, seasonal rainfall characteristics have become a subject of interest in recent years/decades, particularly their relevance for agriculture (Tadross et al., 2005, 2009). Tadross et al. (2005), using observation-based data up to 1997, showed that the mean onset occurs earlier (September or October) over eastern South Africa and later (November to December) over northern Mozambique and Botswana. A trend towards a later onset was found over northeastern South Africa. Using seven Coupled Model Intercomparison Project Phase 3 (CMIP 3) global climate models statistically downscaled to station data, Tadross et al. (2009) projected a reduction in mid-twenty-first century (2046–2065) austral spring precipitation, and an increase in the autumn, suggesting a shift of the rainy season to later dates over southern Africa north of about 20°S. No attempt was made to quantify the uncertainty associated with the downscaling procedure. The CORDEX Africa programme produces large volumes of data necessary to provide future projections of high-frequency precipitation statistics and a quantification of the inherent uncertainties. This paper provides the necessary foundation for such analyses by assessing the models’ ability to replicate some rainfall statistics in their control simulations.

2. Data and methodology

Monthly and daily rainfall simulated by 10 CORDEX RCMs over the common 17-year period (1991–2007), driven by the ERA-Interim reanalysis (Simmons et al., 2006), are used. The RCMs include: (1) the Université du Québec à Montréal fifth-generation Canadian Regional Climate Model (CRCM5), (2) the Universidad de Cantabria Weather Research and Forecasting Model, version 3.1.1 (WRF3.1.1), (3) the Sveriges Meteorologiska och Hydrologiska institut (SMHI) Rossby Centre Regional Atmospheric Climate Model, version 3.5 (RCA3.5), (4) the Max Planck Institute Regional Model (REMO), (5) the Consortium for Small-scale Modelling (COSMO) Climate Limited-Area Model, version 4.8 (CCLM4.8), (6) the Centre National de Recherches Météorologiques Action de Recherche Petite Echelle Grande Echelle, version 5.1 (ARPEGE5.1), (7) the Abdus Salam International Centre for Theoretical Physics RCM, version 3 (RegCM3), (8) the University of Cape Town Providing National Centres for Theoretical Physics RCM, version 3 (SMHI), (9) the Danmarks Meteorologiske Institut HIRHAM, version 5 (HIRHAM5; the HIRHAM5 has days with missing data and has been omitted in the analysis of onset and withdrawal of the rainy season) and the Koninklijk Nederlands Meteorologisch Instituut Regional Atmospheric Climate Model, version 2.2b (RACMO2.2b). The RCM setup details and relevant references are presented in Nikulin et al. (2012) and Kalognomou et al. (2013).

Owing to the dearth and/or inaccessibility of observational data over much of the study area, the ERA-Interim reanalysis is used as the reference data set to assess model performance. Throughout this paper, reference data refer to the ERA-Interim precipitation. Given the uncertainty associated with observation-based data sets (Kalognomou et al., 2013; Sylla et al., 2013), the Global Precipitation Climatology Project (GPCP; Huffman et al., 2009) data, which are available from 1997, are also used for comparison.

The homogeneous rainfall regions over southern Africa defined by Shongwe et al. (2009) are adopted in this study. CORDEX simulation of monthly spatially averaged precipitation is assessed using Taylor diagrams (Taylor, 2001), which graphically synthesize the degree of correspondence between RCMs and the reference data in terms of the phase and amplitude of their evolution, measured by Pearson correlation coefficients, the centred root-mean-square error (RMSE) and a comparison of their variances. Taylor diagrams are widely used to evaluate the multiple aspects of complex models and gauging the relative skill of many different models (e.g., Kalognomou et al., 2013).

Onset and cessation of the wet season is defined for each region in Figure 1, assuming spatial coherence of the timing across the individual regions, from anomalous precipitation accumulation in a given day [A (day)] as

\[ A(\text{day}) = \sum_{n=1}^{\text{day}} R(n) - \bar{R} \]  

where \( R(n) \) is the daily precipitation and \( \bar{R} \) is the long-term annual daily mean (Liebmann et al., 2007; Rauscher et al., 2007). The calculation started on 1 July (climatologically, the driest month; see Figure 2 of Shongwe et al., 2009). The onset of the rainy season is defined as the date on which the curve reaches a minimum, since after that date precipitation exceeds the annual daily climatology and before that date precipitation is less than the annual climatology. Prior to onset, there are often brief periods of precipitation causing the curve to move upward, but these are considered ‘false’ onsets, because the curve ultimately falls to its absolute minimum for the year. Similarly, cessation is defined as the date on which the curve reaches the maximum, since after that date precipitation is less than climatology (e.g., Figure 2 of Liebmann et al., 2007). Pearson correlation coefficients and RMSE are computed between the simulated and reference onset and cessation dates. Statistical significance of the computed correlation is assessed using a parametric Student’s \( t \)-test.

Regional PDFs are constructed for the reference data sets and CORDEX RCMs by considering all the grid points falling within a region, omitting daily precipitation values <0.1 mm day\(^{-1}\) at any given grid. The PDFs are compared visually.
and progresses northward. Earlier onset over eastern South Africa (Region II) is associated with moisture advection from the warm Agulhas system and instabilities induced by mid-latitude disturbances (Tyson and Preston-Whyte, 2000). The spatial progression of the regional rainfall onset in the RCMs is broadly consistent with the reference and the GPCP data, which often present similar estimates of the rainy season timing for the common period 1997/1998–2007/2008, a notable exception being the 2002/2003 season. In ERA-Interim, the mean (median) rainfall onset date in Region II is 22 October (18 October), with an interannual standard deviation (σ) of 15 days. In the GPCP data, the mean onset date and σ for eastern South Africa (i.e. Region II) are 23 October and 17 days. Albeit not statistically significant, the trend towards later onset of 1 day year−1 in the GPCP data over eastern South Africa is consistent with Tadross et al. (2005) and the broadly held notion that in parts of southern Africa, the start of the rainy season has shifted to later dates in recent years/decades. In the ERA-Interim, a weaker delayed onset (4 days decade−1) is only found in the late twentieth century.

Most of the CORDEX RCMs adequately capture the mean rainfall onset in Region II, falling within ±7 days of the reference data (Figure 3(b)). To the southwest (Region I), the mean start of the rainy season occurs around 8 November and 16 November in the ERA-Interim and GPCP reference data sets, respectively. A trend towards a delayed onset (3 ± 2 days year−1) is found during the last decade of the twentieth century. With reference to ERA-Interim, RegCM3 has the overall early mean onset bias of ≈14 days, while the onset in the CCLM4.8 is delayed by about 11 days. Except these two and the RCA35, the mean onset dates in the rest of the RCMs fall within a 7 day margin from the ERA-Interim estimate in Region I (Figure 3(a)). To the north, where moisture recycling ratios are relatively higher during austral summer months (Trenberth, 1999), the seasonal rainfall onset occurs slightly later, around 12 (09), 19 (14) and 11 (12) November in Regions III, IV and V in the ERA-Interim (GPCP) data, respectively, with a relatively low interannual variability (Figure 3(c)–(e)). In Region III, all the RCMs except the RegCM3 have onset dates between 07 and 19 November (Figure 3(c)). In the RegCM3, the mean rainfall onset occurs around 25 October. Over northern Zimbabwe and central Mozambique (Region IV), the onset bias is low, with the CCLM4.8 showing a slight late bias (Figure 3(d)). The onset bias is similarly quite low over the northern-most region (Region V) will almost all RCMs simulating mean onset dates within ±7 days of 11 November.

Correlations between reference and simulated onset dates are shown in Table 1. It is evident from the table that the RCMs capture the interannual variability of the rainfall onset reasonably well. In particular, the ARPEGE5.1 attains statistically significant correlations and lowest RMSE almost everywhere. There are a few cases when the RCM simulated onset dates are out of
Figure 2. Taylor diagrams for area-averaged monthly precipitation in Regions (a) I, (b) II, (c) III, (d) IV and (e) V shown in Figure 1. The reference (ERA-Interim) data are shown by the grey square along the horizontal axis. The individual RCMs are shown by the solid circles and the GPCP by the open circle. The radial coordinate shows the standard deviation. The azimuthal axis shows the correlation between the RCMs and the reference data. The centred root-mean square error is indicated by the dashed grey semi-circles about the reference point.
phase with the reference data (e.g. CRCM5 and REMO in Region II, RCA3.5 in Region III and WRF3.1.1 in Region IV). The weak correlation between the GPCP and ERA-Interim in Regions I and II demonstrates how different observational-based data sets may disagree on their representation of precipitation characteristics (e.g. Kalognomou et al., 2013).

In the reference data, rainfall withdrawal does not show any notable spatial migration. Almost everywhere, except in Region III, the mean end date occurs between 28 March and 05 April (27 March and 06 April) in ERA-Interim (GPCP). In Region III, which has long been known to have strong ENSO tele-connections (Ropelewski and Halpert, 1987; Rocha

Figure 3. ERA-Interim (grey rectangles), GPCP from 1997/1998 (orchid rectangles) and CORDEX-simulated (individual points) interannual onset (solid circles) and cessation (solid squares) dates of the wet season for Regions (a) I, (b) II, (c) III, (d) IV and (e) V. In each figure, the onset (cessation) date is plotted at the lower (upper) ordinate running from 01 October to 15 December (01 February to 01 June). For the reference plots, the rectangles extend to ±1σ of interannual variability from the estimated date, while the line segments within each rectangle extend to ±7 days.
and Simmonds, 1997), the mean cessation, which occurs around 08 and 06 March in the reference data and GPCP, is influenced by six early cessation seasons (1991/1992, 1993/1994, 1994/1995, 1997/1998, 2001/2002 and 2004/2005). In these extreme years, four (in bold font) of which coincided with a warm ENSO phase, the seasonal rainfall ends either late in January (e.g. 2004/2005) or early to mid-February. Statistically significant correlations (at the 10% level) have been found between ENSO and early rainfall cessation over a number of stations in Zimbabwe, southern Mozambique and Malawi (Tadross et al., 2009). Noteworthy, the southeastern regions (Regions II and III) have a high interannual variability ($\sigma \geq 25$ days) of the rainfall withdrawal (Figure 3(b) and (c)).

The pattern of rainfall cessation in the RCMs closely resembles that of the reference data, with some models showing an excellent agreement even during the unusual years. In Region I, all except two of the RCMs (RegCM3 and CRCM5) have end dates within $\pm 7$ days of the reference data. In RegCM3, the mean end date occurs quite early (04 March), while in CRCM5, the average cessation occurs around 13 March, 16 days earlier than in ERA-Interim. Over eastern South Africa (Region II), the largest biases are found in REMO and RACMO2.2 where the end of the rainy season is delayed by 21 and 14 days, respectively. In almost all the RCMs, the cessation dates vary greatly from year to year, in close agreement with the reference data (Figure 3(b)). In Region III, early withdrawals in the

---

**Table 1. Correlation between the RCMs and the reference ERA-Interim onset dates.**

<table>
<thead>
<tr>
<th>RCM</th>
<th>Region I</th>
<th>Region II</th>
<th>Region III</th>
<th>Region IV</th>
<th>Region V</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPCP</td>
<td>0.48 (17)</td>
<td>0.13 (18)</td>
<td>0.67** (8)</td>
<td>0.94** (4)</td>
<td>0.60** (10)</td>
</tr>
<tr>
<td>ARPEGE5.1</td>
<td>0.73** (12)</td>
<td>0.48 (13)</td>
<td>0.59** (12)</td>
<td>0.69** (10)</td>
<td>0.49** (13)</td>
</tr>
<tr>
<td>CCLM4.8</td>
<td>0.34 (16)</td>
<td>0.45 (13)</td>
<td>0.29 (21)</td>
<td>0.16 (18)</td>
<td>0.15 (15)</td>
</tr>
<tr>
<td>CRCM5</td>
<td>0.16 (18)</td>
<td>−0.08 (21)</td>
<td>0.29 (23)</td>
<td>0.39 (13)</td>
<td>0.28 (14)</td>
</tr>
<tr>
<td>PRECIS</td>
<td>0.22 (16)</td>
<td>0.40 (15)</td>
<td>0.04 (15)</td>
<td>0.46 (13)</td>
<td>0.78** (7)</td>
</tr>
<tr>
<td>RCA3.5</td>
<td>0.56** (13)</td>
<td>0.46 (17)</td>
<td>0.34 (20)</td>
<td>0.41 (13)</td>
<td>0.17 (15)</td>
</tr>
<tr>
<td>RACMO2.2b</td>
<td>0.56** (13)</td>
<td>0.32 (14)</td>
<td>0.12 (23)</td>
<td>0.26 (15)</td>
<td>0.21 (15)</td>
</tr>
<tr>
<td>REMO</td>
<td>0.38 (17)</td>
<td>−0.09 (18)</td>
<td>0.69** (10)</td>
<td>0.50** (13)</td>
<td>0.62** (9)</td>
</tr>
<tr>
<td>WRF3.1.1</td>
<td>0.67** (12)</td>
<td>0.46 (17)</td>
<td>0.53** (12)</td>
<td>−0.18 (22)</td>
<td>0.06 (18)</td>
</tr>
</tbody>
</table>

The RMSE (days) for each RCM is shown in brackets. Correlations that are significant at the 10% (5%) level are in bold (shown by two asterisks).
Table 2. Correlation between the RCMs and the reference ERA-Interim cessation dates.

<table>
<thead>
<tr>
<th>RCM</th>
<th>Region I</th>
<th>Region II</th>
<th>Region III</th>
<th>Region IV</th>
<th>Region V</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPCP</td>
<td>0.84**</td>
<td>0.88**</td>
<td>0.93**</td>
<td>0.54</td>
<td>0.75**</td>
</tr>
<tr>
<td>ARPEGE5.1</td>
<td>0.82**</td>
<td>0.79**</td>
<td>0.84**</td>
<td>0.78**</td>
<td>0.81**</td>
</tr>
<tr>
<td>CCLM4.8</td>
<td>0.46 (19)</td>
<td>0.60**</td>
<td>0.36 (28)</td>
<td>0.23 (16)</td>
<td>−0.29 (14)</td>
</tr>
<tr>
<td>CRCM5</td>
<td>−0.09 (29)</td>
<td>0.33 (31)</td>
<td>0.43 (28)</td>
<td>0.65**</td>
<td>0.55** (7)</td>
</tr>
<tr>
<td>PRECIS</td>
<td>0.56**</td>
<td>0.44 (21)</td>
<td>0.04 (7)</td>
<td>0.48 (7)</td>
<td></td>
</tr>
<tr>
<td>RACMO2.2b</td>
<td>0.12 (18)</td>
<td>0.72**</td>
<td>0.29 (30)</td>
<td>0.47 (14)</td>
<td>0.25 (10)</td>
</tr>
<tr>
<td>RCA3.5</td>
<td>0.20 (18)</td>
<td>0.46 (24)</td>
<td>0.07 (35)</td>
<td>0.17 (24)</td>
<td>0.17 (11)</td>
</tr>
<tr>
<td>RegCM3</td>
<td>0.36 (24)</td>
<td>0.50**</td>
<td>0.53**</td>
<td>0.47 (15)</td>
<td>0.55** (8)</td>
</tr>
<tr>
<td>REMO</td>
<td>0.25 (27)</td>
<td>0.61**</td>
<td>0.16 (38)</td>
<td>0.64**</td>
<td>−0.19 (15)</td>
</tr>
<tr>
<td>WRF3.1.1</td>
<td>0.47 (17)</td>
<td>0.38 (25)</td>
<td>0.82**</td>
<td>0.37 (21)</td>
<td>0.32 (9)</td>
</tr>
</tbody>
</table>

The RMSE (days) for each RCM is shown in brackets. Correlations that are significant at the 10% (5%) level are in bold (shown by two asterisks).

RegCM3 occur often, which results in an average negative bias of 18 days. Further north (Region IV), where the ability to simulate monthly rainfall is highest, negative biases in RCM end dates in excess of 18 days are found in PRECIS, CCLM4.8, RCA3.5 and RegCM3. Over the north-most Region V, disagreements between average reference and models’ end dates still persist (Figure 3(e)).

A much stronger agreement is found between the ERA-Interim and GPCP in terms of interannual variability of rainfall cessation (Table 2). The ARPEGE5.1 has the highest correlations in all the regions. With only a few exceptions (e.g. CCLM4.8 and REMO in Region V, and CRCM5 in Region I), the RCM simulated rainfall withdrawal is positively correlated with the reference dates indicating that the models are able to capture the interannual variation of the rainy season cessation.

5. Rainfall probability density functions

Daily precipitation PDFs for each region are shown in Figure 4(a)–(e). The x axes of the plots terminate at about 40 mm day$^{-1}$ because at higher values exceeding 20 mm day$^{-1}$, the individual curves become almost indistinguishable. We show that very intense events ($\leq 40$ mm day$^{-1}$) are simulated in both the reference data and in the RCMs, demonstrating that the models can provide useful information in relation to flood risks.

Although drizzle and light rainfall events $\leq 3$ mm day$^{-1}$ occur too often in a few of the RCMs (CCLM4.8, REMO and RCA3.5), in most regions, the majority have
low-tail probabilities comparable with the reference data. In most cases, the CCLM4.8, and in some cases, the REMO overestimate the probability of light rainfall by about 2–3 times. The RACMO2.2, the ARPEGE5.1 and the HIRHAM5 get very close to capturing the secondary peak probabilities located around 5–10 mm day$^{-1}$ in the ERA-Interim data in Region V.

6. Summary and conclusions

The present study has enabled us to get to grips with how well the CORDEX RCMs simulate monthly rainfall amounts and some regional rainfall characteristics over southern Africa. The models have been evaluated against the ERA-Interim reanalysis, whilst the GPCP data have been used for comparison. In simulating monthly rainfall variability, the overall performance of the RCMs is good, particularly to the north of the study domain. The stretched grid ARPEGE5.1 appears to be the best overall performer, in qualitative agreement with Kalognomou et al. (2013), where the individual RCM setups are presented. This RCM’s correlation with the ERA-Interim data is highest almost everywhere in southern Africa and the biases are least. No attempt is made in this paper to identify the model physics and dynamics responsible for the differences in RCM performance. Other aspects of the simulated precipitation that are of practical significance such as rainy season timing and the relative frequencies of a spectrum of rainfall intensities have also been analysed. The spatial migration of the seasonal rainfall onset is well captured by the RCMs. The reference and simulated average onset and withdrawal dates are within a few days of each other, and the interannual variability is captured by the RCMs, a few notable biases notwithstanding. A visual comparison of the reference and simulated PDFs shows that a majority of the models do quite well in representing the probabilities of certain rainfall events.

There are prospects for the CORDEX RCMs to provide usable information on regional monthly precipitation, rainfall characteristics and probabilities of events of differing intensities over southern Africa, and their likely future changes, for climate change impacts assessments where impacts are related to precipitation. The uncertainties associated with the modelling process from the driving climate to impacts are yet to be quantified. One of the goals for future work is to investigate and compare the physical mechanisms underlying rainfall variability and changes in rainfall character in observations and in the models.

Acknowledgements

Review comments by Dr. Rachel James and two anonymous reviewers are greatly appreciated. This paper is a contribution of the Southern African Analysis group to the CORDEX-Africa programme. The CORDEX-Africa programme was supported by the Global Change System for Analysis, Research, and Training (START) through the Climate Systems Analysis Group of the University of Cape Town. Support from the World Climate Research Program (WCRP), the Climate and Development Knowledge Network (CDKN), the International Centre for Theoretical Physics (ICTP), the Swedish Meteorological and Hydrological Institute (SMHI) and the European Union Seventh Framework Programme is gratefully acknowledged. Special thanks to the modelling groups contributing to the CORDEX-Africa program and to Grigory Nikulin from SMHI for post-processing the data to a common grid, data format and domain size to enable direct comparison in the analyses.

References


