RegCM4 in climate simulation over CORDEX-MENA/Arab domain: selection of suitable domain, convection and land-surface schemes

Mansour Almazroui

Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT: In order to find out the optimal setting for downscaling the Coupled Model Inter Comparison Project Phase 5 (CMIP5) database, this study describes the most suitable domain, convection and land-surface schemes within Regional Climate Model version 4 (RegCM4) over the Middle East and North Africa (MENA) domain. The activity is carried out in compliance with COordinated Regional climate Downscaling EXperiment (CORDEX) initiative, where domain suitability is assessed through seven simulations for the period 2001–2005. After the selection of domain, eight simulations are carried out to find better performing convection and land-surface schemes within RegCM4 for the same time period of 5 years. These experiments are conducted using ERA-Interim reanalysis datasets; however, the selection of optimal settings is also validated using other reanalysis datasets. Finally, using the same settings, data of five CMIP5 general circulation models (GCMs) are also downscaled to assess the applicability of RegCM4. Statistical measures, such as correlation coefficient, bias, root mean square difference and standard deviation, are taken into consideration to compare the simulated data against the gridded Climatic Research Unit (CRU) and Tropical Rainfall Measuring Mission (TRMM) datasets. RegCM4-simulated rainfall and temperature show better spatial distributions and magnitudes against observations, as compared to ERA-Interim data, which may be considered as an added value of the model. In analysing rainfall and temperature data for the 11 sub-domains (each an 8°×8° box), the region encompassing the area 7°S–45°N and 27°W–76°E, named CORDEX-MENA/Arab domain, is found to be the most cost effective and suitable domain. Further analysis over this domain shows Biosphere and Atmosphere Transfer Scheme (BATS) to be the better performing land-surface scheme, whilst Grell with Fritsch–Chappell closure (GFC) comes up as the most suitable option for convection scheme in the model. Of five CMIP5 datasets downscaled using RegCM4, rainfall and temperature patterns of four are similar to the reanalysis and observations. Therefore, it is recommended that RegCM4 may be employed with the aforementioned settings in downscaling the CMIP5 multi-model database for climate change impact and adaptation studies for the region in future.

KEY WORDS rainfall; temperature; RegCM4 configuration; suitable domain; CORDEX; MENA region

Received 26 August 2014; Revised 13 March 2015; Accepted 16 March 2015

1. Introduction

There is a unified consensus among the scientific community across the globe that the climate has been changing as a consequence of rapidly increasing concentrations of atmospheric greenhouse gases (GHGs). The Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) states that mean global temperatures will probably rise between 0.3 and 4.8°C by the end of this century (2081–2100) relative to 1986–2005 (IPCC, 2013). However, at the regional level, the magnitude of the climate change signal (in terms of different variables such as temperature, precipitation and wind) and associated extremes could be substantially different compared with the global average. The Arab region, also called Middle East and North Africa (MENA), is specifically vulnerable to the adverse effects of climate change. The IPCC AR5 report shows a significant and robust precipitation decrease (about 50%, RCP8.5) in the rainy season, particularly in the months of DJF over densely populated areas of the region by the end of 21st century (2021–2100) with respect to 1986–2005. Therefore for the MENA region, which has remained under pressure to adapt to water scarcity and heat throughout the course of history, the adverse impacts of climate change are likely to aggravate these problems in any future climate.

Recognizing the vulnerability of the Arab region to climate change, the Arab League and the United Nations jointly established RICCAR (Regional Initiative for the assessment of Climate Change impacts on water resources and socio-economic vulnerability in the Arab Region; http://www.escwa.un.org/RICCAR). RICCAR, along with its partner institutes including the United Nations Economic and Social Commission for Western Asia (ESCWA), the League of Arab States (LAS), the Arab Center for the Studies of Arid Zones and...
Dry lands (ACSAD), the Swedish Meteorological and Hydrological Institute (SMHI), the World Meteorological Organization (WMO) and the United Nations International Strategy for Disaster Risk Reduction (UNISDR), took a regional initiative that aims to assess the impacts of climate change on the freshwater resources of the region.

The first challenge within the RICCAR initiative was to construct information pertaining to future climate change for the region using state-of-the-art climate modelling techniques. Today, coupled atmosphere-ocean general circulation models (AOGCMs) are considered to be the most advanced numerical tools to understand the dynamics of the physical components of the climate system, and for making projections based on future GHGs and aerosol forcing (Meehl et al., 2007; IPCC, 2013). Beside their use for seasonal to decadal climate prediction applications, the high-resolution (sometimes variable-resolution) runs of AOGCMs are often used in process studies with a focus on a particular region. In general, the typical resolution of AOGCMs used in AR5 is about 100–200 km, however many regional-scale climate processes go beyond the scope of AOGCMs (IPCC, 2013). Therefore, downscaling coarse resolution AOGCM outputs for a better understanding of regional climatic phenomena through regional climate models (RCMs) is a technique that is now well accepted in the scientific community for climate change impact and adaptation studies (Giorgi and Hewitson, 2001; Jones et al., 2004; Pal et al., 2007). RCMs can provide finer patterns in terms of the climatic parameters for a region with physical consistency, an arguably more realistic exploration of the uncertainty range compared to GCMs and spatial patterns which follow the observations. It is worth mentioning that dynamical downscaling through an RCM is based on fundamental physical principles and is thus supposed to be (at least to some extent) transferable from one region to another. Importantly, dynamical downscaling is useful for incorporating features such as the orography, land use and vegetation of a region (Brown et al., 2008). Hence, under the RICCAR platform, it was decided to produce regionally downscaled future climate projections for the MENA/Arab region by using a regional climate modelling technique.

The next challenge for RICCAR was to establish an RCM domain for conducting long-term future climate simulations, which was by no means a trivial task. It is well established that large-scale circulation and rainfall distributions are more sensitive to domain size than to horizontal resolution (Shiaof and Juang, 2006). Recently, Almazroui (2012) concluded that domain size is highly influential in the reproduction of precipitation events across the Arabian Peninsula. He also demonstrated that the use of high resolution does not systematically improve the simulation of precipitation events if the signal in the driving forcings is not sufficiently clear. However, the domain within an RCM must be large enough to allow for the full development of fine-scale features over the region of interest (Jones et al., 1995). Also, to avoid an unrealistic response to internal forcings, the lateral boundaries must be placed well outside the region of interest, whilst a smaller domain confines the interior solution more towards the coarse driving fields, which may result in an unrealistic response to internal higher resolution forcings (Giorgi and Mearns, 1999). Therefore, domain size is a critically important issue in the performance of an RCM. Hence, after a few preliminary meetings, the partners in RICCAR agreed upon following two criteria for the selection of the domain:

I It should cover all the 22 Arab League countries. However, this criterion could not be met because Comoros was at too far to the South, and hence was left out.

II Those regions from which fresh water rivers flow into the MENA region should also be included within the domain. Therefore, the Nile and the Tigris-Euphrates should define the southern and northern extent of the CORDEX-MENA domain, respectively.

In the meantime, the World Climate Research Program (WCRP) initiated a coordinated effort with the international regional downscaling community to build a coordinated framework, in order to produce an improved generation of regional climate change projections over different domains across the world. This initiative was named the COordinated Regional climate Downscaling EXperiment (CORDEX) (http://www.meteo.unican.es/en/projects/CORDEX). The aim of CORDEX is to produce an ensemble of multiple dynamical and statistical downscaling models that consider multiple forcing GCMs from the Coupled Model Inter Comparison Project Phase 5 (CMIP5) archive. Although, the CORDEX domains cover most of the land areas of the world, certain regions, such as MENA/Arab (which has similar climate characteristics over a wide area), was not included in a single domain. The CORDEX-Africa domain terminates just on the eastern edge of the Arabian Peninsula (for instance, see http://www.meteo.unican.es/en/projects/CORDEX), and thus does not explore the influence of Indian Ocean large-scale activities, particularly the effects of the Indian monsoon on the summer climate of the Arabian Peninsula. On the other hand, the CORDEX-South-Asia domain terminates just on the western edge of the Arabian Peninsula, and thus does not investigate the Mediterranean high and Sudan low regions, which heavily influence the winter climate of the Arabian Peninsula. This latter domain also does not cover the MENA countries. Therefore, identification of a new model domain is required in order to cover the above mentioned issues and to downscale the climate of the MENA/Arab region from the CMIP5 multi-model data. Initially, the RICCAR’s regional climate modelling activity began independently of CORDEX; however, it was realized by the partners within RICCAR that their regional climate modelling activity could benefit a great deal from joining CORDEX. Hence, it was decided to carry out further activities as per CORDEX protocols.

Besides the selection of an appropriate domain, the selection of better performing land-surface and convection schemes is also crucial to the performance of an RCM over a particular region. Within RegCM4 [Regional Climate Model version 4.3.4 developed by the Abdus Salam
International Center for Theoretical Physics (ICTP), Trieste, Italy), there are four convection schemes available which are Modified-Kuo scheme (Anthes, 1977), the Grell scheme (Grell, 1993), the MIT-Emanuel scheme (Emanuel, 1991; Emanuel and Živković-Rothman, 1999) and a new mixed scheme combining Grell and Emanuel. Therefore, one of them has to be selected to drive RegCM4, due to the fact that magnitude of simulated climatic variables varies from convection scheme to another. Moreover, in a climate model, a Land Surface Model (LSM) is the interface between the earth’s surface and the planetary boundary layer. LSM in fact describes the role of vegetation and interactive soil moisture in modifying the surface-atmosphere exchanges momentum, energy and water vapour (Dickinson et al., 1993). Changes in land-surface types have important consequences for the climate system, as particular types play a crucial role in the local surface energy and water budget (Wei and Fu, 1998).

The inclusion of Biosphere and Atmosphere Transfer Scheme (BATS) and the Community Land Model (CLM) into the latest version of RegCM4 permits an exploratory study of the possible effects of changes in land cover and use (Giorgi et al., 2003).

As mentioned earlier, the selection of the RCM domain is one of the more important tasks in the RICCAR initiative. Although SMHI was given a leading role in the selection of an appropriate domain for CORDEX-MENA activity, it is worth mentioning here that during this process, the Center of Excellence in Climate Change Research (CECCR), King Abdulaziz University, Saudi Arabia, played a vital role by conducting various test runs using RegCM4. This study aims to identify the suitable domain, as well as better performing convection and land-surface schemes within RegCM4 for the MENA region. For this purpose, several simulations are performed using European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim (hereafter referred to ERA-Int) 6 hourly boundary forcings. Furthermore, the data of five GCMs are also downscaled and analysed in order to test the RegCM’s optimal settings for its use in long-term simulations.

2. Data and methodology

2.1. Model description

Regional Climate Model (RegCM) is the first limited-area model developed for long-term climate simulation and is applied by a large community of researchers for a wide range of regional climate studies. RegCM4 (Giorgi et al., 2012) is a hydrostatic and compressible model that uses sigma-p vertical coordinates, run on an Arakawa B-grid. RegCM4 has the dynamical core of the mesoscale model version 5 (MM5) developed by the National Center for Atmospheric Research (NCAR) and Pennsylvania State University (Grell et al., 1994). It incorporates the radiation scheme of the NCAR Community Climate System Model Version 3 (CCSM3) (Kiehl et al., 1996; Collins et al., 2006) for radiation and infrared spectra, following Solmon et al. (2008). The BATS for land-surface processes follows Dickinson et al. (1993) and Gao et al. (2006). The CLM (Oleson et al., 2008) is also available as the land-surface scheme in RegCM4. To represent urban and suburban environments, two new land-use types have been added to BATS in RegCM4. Urban development modifies the surface albedo and alters the surface energy balance; it also heavily affects runoff and evapotranspiration. In RegCM4, version CLM3.5 (Tawifiq and Steiner, 2011) uses a series of biogeophysically based parameterizations to describe the land–atmosphere exchanges of energy, momentum, water and carbon. Moreover, RegCM4 uses other physical processes, including those of Holtslag et al. (1990) for the Planetary Boundary Layer (PBL) and Zeng et al. (1998) for ocean flux parameterization approaches.

As previously mentioned that multiple cumulus convection schemes are available in RegCM4. The new scheme allows the user to select either Grell or Emanuel in the function of the ocean-land mask. The Arakawa–Schubert type closure or the Fritsch and Chappell (1980a, 1980b) type closure are available for use in the Grell convective parameterization scheme. In this study, the initial and boundary conditions to the model are provided from ERA-Interim reanalysis 1.5° × 1.5° gridded 6 hourly data (http://www.ecmwf.int/products/dataarchive). RegCM4 is forced with Optimum Interpolation Sea Surface Temperature (OISST) and Extended Reconstructed Sea Surface Temperature (ERSST) where applicable. The OISST is the data at weekly timescale, which are obtained from the National Oceanic and Atmospheric Administration (NOAA). ERSST is obtained from National Climate Data Centre (NCDC). It is worth mentioning that RegCM4 can be employed at further higher resolution; however, 50 km grid spacing is the initial preference within CORDEX which is favoured by the wider community, and therefore is used in this study.

2.2. Experiment set-up

2.2.1. Experiment set-up for domain selection

In order to select the CORDEX-MENA/Arab domain in employing RegCM4 (with 50 km horizontal resolution) in downsampling the CMIP5 data, a number of test runs with ERA-Int driving forcings with OISST are performed, using the Grell convective scheme with Fritsch–Chappell closure (GFC) for different domains, viz: (1) D1 bounds 10°S–55°N and 27°W–76°E, (2) D2 bounds 7°S–45°N and 27°W–76°E, (3) D3 bounds 10°S–55°N and 21°W–71°E, (4) D4 bounds 7°S–45°N and 21°W–71°E, (5) D5 bounds 7°S–45°N and 27°W–71°E, (6) D6 bounds 7°S–45°N and 21°W–76°E and (7) D7 bounds 7°S–45°N and 21°W–66°E (see Figure 1) for the period 2000–2005 (2000 is the spin-up year). Note that GFC is the initial choice for the selection of domain following Almazroui (2011, 2012) who employed RegCM3 for the simulation of rainfall and temperature across the Arabian Peninsula. Another run is performed for the decade 1997–2007 (1997 is the spin-up) for the selection of suitable domain. This decade is selected for the availability of Tropical Rainfall Measuring Mission (TRMM) data, which starts
from 1998. Two simulation periods (2000–2005 and 1997–2007) are used in order to understand the effect of run length in reproducing rainfall and precipitation information. In the context of the existing literature on similar subjects, the use of a 30-year base period is required; however, 10 years is also acceptable because shorter averaging periods (such as 10 years) often perform as adequately as 30-year averaging periods for many climatic parameters (WMO, 2007). In order to verify the performance of RegCM4 with another driving forcings, one run is completed with NNRP2 (NCEP/NCAR Reanalysis Project) with OISST. Moreover, one ERA-Int run with ERSST is also performed in order to understand the sensitivity of RegCM4 results to SST.

2.2.2. Experiment set-up for convection and land-surface schemes selection

For the selection of suitable convection as well as land-surface scheme, a total of eight experiments for RegCM4 at 50 km horizontal resolution are carried out for the better performing CORDEX-MENA/Arab domain (7°S–45°N and 27°W–76°E) for the period 2000–2005 (2000 is the spin-up year). In this work, the convective schemes of GFC, Grell with Arakawa–Schubert closure (GAS), Emanuel (EMAN) and Grell over land and Emanuel over ocean (GLEXO) are used. The land-surface scheme selection sensitivity experiments are the following:

I RegCM4 runs performed for GFC, GAS, EMAN and GLEXO using land-surface scheme BATS. Therefore, a total of four runs with BATS are completed for the period 2000–2005.

II Step (I) is repeated to complete four runs with land-surface scheme CLM for the period 2000–2005.

2.2.3. Experiment set-up for CMIP5 data downscale

Finally, by using the optimal settings (e.g. domain, convection and land-surface schemes identified earlier), a total of five CMIP5 experiments are downscaled with RegCM4 for the period 2000–2005 (2000 is the spin-up year). The CMIP5 downscale sensitivity experiments are the following:

I RegCM4 run performed with MPI-ESM-MR (ECHAM6) from Germany for the period 2000–2005.

II Step (I) is repeated with GFDL-ESM2M from United States.

III Step (I) is repeated with HadGEM2-ES from UK.

IV Step (I) is repeated with CNRM-CMS from France.

V Step (I) is repeated with CanESM2 from Canada.

2.3. Data analysis

The 3 hourly output data from RegCM4 simulations are transformed into daily, monthly, seasonal and annual scales. Simulated rainfall and temperature biases are calculated with respect to the observations, such as Climatic Research Unit data (CRU; New et al., 2000; Mitchell and Jones, 2005) and TRMM data (Kummerow et al., 1998). The CRU/TRMM dataset has the spatial resolution 0.5°/0.25° grid-mesh. The ERA-Int driving field Initial Condition and Boundary Conditions (ICBC) data are also used in the evaluation of RegCM4’s performance in the simulation of rainfall and temperature.

For statistical measures, 11 sub-domains (see Figure 1) across the analysis domain, each an 8°×8° box, are used. The sub-domains are arbitrarily selected across the MENA region to cover various of climatic conditions such as warm and cold places, complex topography (hilly and flat) and land-surface heterogeneity (deserts and wet areas). In other studies, such as in Endris et al. (2013), the sub-domains have different sizes and they cover areas not only over land but also over oceans. In order to avoid this inconsistency, equal-sized sub-domains placed over only the land areas are used, which gives better observed data coverage than over the oceans. The average from all 11 sub-domains is considered as the representative of the CORDEX-MENA/Arab domain. In order to select the suitable domain, at first D1 is considered for analysis, then it is reduced north and south for D2; then again D1 and D2 are further reduced east–west for other domains (Figure 1). The better performing land-surface scheme option in RegCM4 for each sub-domain as well as for the entire domain is decided with the aid of statistical measures such as mean, bias, correlation (r), root mean square difference (RMSD) and standard deviation (SD) against CRU values. The lognormal probability density function is used for the rainfall and temperature time series in order to understand the performance of RegCM4 over the long-term (1998–2007) simulation relative to the short-term (2001–2005) simulation. RegCM4 output for the GFC convective scheme is analysed in detail, and the other three convective scheme options are used for discussion purposes.

3. Results

3.1. Domain selection

In order to obtain a reasonably well performing convection and land-surface scheme options within RegCM4 for CORDEX-MENA/Arab domain, seven test simulations using ERA-Int boundary conditions are performed for the seven domains shown in Figure 1. Domain D1 (10°S–55°N and 27°W–76°E) includes the Indian Ocean, the Mediterranean Sea and the Sudan low region. Domain D2 bounds 7°S–45°N and 27°W–76°E and represents a reduction of D1 in the north and south. Domain D3 (10°S–55°N and 21°W–71°E) represents a reduction of D1 in both east and west. Domain D4 (7°S–45°N and 21°W–71°E) is the reduction of D2 in both east and west, whilst D5 (D6) is the reduction of D2 by 5° only in the east (west). Domain D7 (7°S–45°N and 21°W–66°E) is the reduction of D1 by 6° in the west and 10° in the east. Comparisons of RegCM4-simulated rainfall and temperature (areal averages) for the 11 sub-domains with the CRU values are considered in the evaluation of the most suitable domain.
The rainfall amounts for the different sub-domains obtained from the simulations using D1–D7 are summarized in Figure 2(a). The simulated annual rainfalls for the different domains vary slightly from domain to domain at each sub-domain. Overall, the simulations overestimated (underestimated) the rainfalls for the sub-domains C and E (F and I), where the rainfall amounts are large. For the entire domain, the simulated rainfall averages for all 11 sub-domains (all, Figure 2(a)) were slightly underestimated within a range 1.10–3.58% with respect to CRU data. The small overestimation is observed with respect to TRMM for all domains. Hence, simulated rainfall lies between the two estimates of the observations (CRU and TRMM). However, underestimation with respect to CRU is within a small range and below 4% for all the domains; and the lowest is for the domain D2 (0.87%). This clearly indicates that the larger domain D1 does not improve the results, whilst the small domain D2 performs even better, which makes the downscaling of long-run simulations (under various scenarios) more cost effective.

In general, temperature is one of the parameters that climate models can simulate more accurately. The simulated temperatures using domains D1, D2, … and D7 together with their comparisons with the CRU data (for all sub-domains) are shown in Figure 2(b). As stated previously, the simulated temperature reveals a cold bias for all sub-domains. The domain D2 provides similar results to D1, however, D2 shows better performance in the majority of sub-domains (A, B, D, E and G). For the average from all 11 sub-domains (all, Figure 2(b)), the simulated temperature has a cold bias of 1.04, 0.70, 0.91, 0.88, 0.84, 0.93 and 0.88 °C compared with CRU for D1, D2, D3, D4, D5, D6 and D7, respectively. Although this represents 0.34 °C less bias for the small domain D2 compared with the larger domain D1, for long-term climate integrations this difference could be noticeable and thus should be considered in climate change impact studies. This result is again in favour of using the relatively small domain D2 instead of the larger domain D1 in downscaling CMIP5 data using RegCM4.

Simulated rainfall and temperature compared with CRU data revealed that the overall performance of domain D2 is better than other domains, particularly superior to the larger domain D1. Moreover, D2 is cost effective and satisfies the coverage of large-scale circulations from the Indian Ocean, the Mediterranean Sea and the Sudan low region, thus making it favourable for downscaling GCM outputs with the aid of RegCM4. Therefore, based on the RegCM4-simulated rainfall and temperature comparison with CRU values for the 11 sub-domains, and through discussions at a series of CORDEX meetings organized by ESCWA, the domain 7°S–45°N and 27°W–76°E is decided to be the effective CORDEX-MENA/Arab domain in downscaling the CMIP5 multi-model database.

3.2. Convection scheme selection
In order to investigate the performance of RegCM4 in terms of convection schemes, the results of simulated rainfall and temperature for four convection schemes for the period 2000–2005 (2000 is spin-up year) for domain D2 are presented here.

Monthly rainfall amounts for four convective schemes with BATS and four convective schemes with CLM simulations reveal that the EMAN option totally overestimated the rainfall compared with CRU and TRMM for BATS and slightly overestimated for CLM (Figure 3(a)). It is obvious from the figure that GAS and GLEO options tend to underestimate the rainfall, respectively, for both the LSMs,
Figure 2. (a) Annual rainfall (mm) and (b) Temperature (°C) at different sub-domains and their averages from all 11 sub-domains (all) for domains D1, D2, D3, D4, D5, D6 and D7 with CRU. Rainfall and temperatures are averaged over the period 2001–2005.

whereas EMAN generally tends to overestimate the rainfall. Although the results of EMAN with CLM stay closer to the observations as compared to EMAN with BATS, the sporadic structure of the peaks in EMAN with CLM makes the results look unreasonable. GFC with CLM generally tends to underestimate the rainfall as well, whereas GFC with BATS produces the results much closer to the observations. In summarizing a few statistical measures for all four convective schemes in both the BATS and CLM land-surface schemes (Table 1), it is clear that the BATS option with the GFC convective scheme delivers the lowest rainfall bias and the strongest correlation (with a slope close to 1) with the CRU data. This option (BATS_GFC) also has the lowest RMSD and SD 19.37; the latter is close to the CRU SD of 17.28. For temperature, the convective schemes GLEO and GAS are the candidates, however GFC is in the middle of the range in statistical measures, but all of them are with BATS and not with CLM. Rainfall is one of the climatic parameters difficult to simulate (Islam, 2009); therefore, giving priority to rainfall and considering temperature from the middle of the range, the GFC is suggested as a preferred convection scheme in downscaling the CMIP5 database for the CORDEX-MENA/Arab domain.

3.3. Land-surface scheme selection

From the previous sub-section, GFC scheme with BATS comes up as a suitable option for the domain D2. However, there is still a need to look at the performance using spatial analysis as well as incorporating additional statistical measures to get an idea about its performance. Therefore, the performance of RegCM4 is first evaluated from the climatological point of view using the ERA-Int data for the CORDEX-MENA/Arab domain for both BATS and CLM land-surface schemes for the period 2001–2005 over the selected domain D2.
Space distributions of simulated annual rainfall for the GFC convective scheme option with BATS and CLM for land-surface schemes are shown in Figure 4. The simulated rainfall shows a dry belt from West Africa to the Arabian Peninsula (except in the south-western and northern areas of the Peninsula). The rainfall amounts are large in the southern and northern areas of the analysis domain. In the Arabian Peninsula, the rainfall amounts are relatively large in the northern and south-western areas. In comparison with the ERA-Int and observations (CRU and TRMM), the simulated rainfall distribution (Figure 4(a) and (b)) is relatively dry over the Arabian Peninsula. At this stage, one cannot measure the magnitude of wetness or dryness, and not much is known about which simulation option in terms of land surface is to be considered closer to the observations. At a glance, the CLM simulation is drier compared with BATS, especially towards the equatorial African region. The ERA-Int driving field rainfall
distribution (Figure 4(c)) does not show detailed patterns, compared with the observations (CRU and TRMM, Figure 4(d) and (e)). However, the distribution of rainfall downscaled by RegCM4 (Figure 4(a) and (b)) shows a finer structure compared with the driving ERA-Int rainfall (Figure 4(c)), which is somewhat similar to observations (Figure 4(d) and (e)), although variations are observed for different options (e.g. BATS and CLM) and regions.

In order to find the better performing land-surface scheme in RegCM4, the simulated rainfalls extracted from each sub-domain and averaged over all 11 sub-domains for both BATS and CLM, and are displayed in Figure 5. The observed (CRU and TRMM) and the ERA-Int driving field (ERA-Int_ICBC) rainfalls are also displayed. The time sequences of the simulated monthly rainfalls follow well the annual cycle of the observed data for each year from 2001 to 2005. RegCM4 with BATS definitely reduces the wet bias in driving ERA-Int for the summer months. In contrast, RegCM4 with CLM substantially underestimates the monthly rainfall amounts compared with the CRU and TRMM in every year. On the other hand, the BATS amounts more or less approximate the observations. The monthly average rainfall amount is 38.8 and 14.7 mm for RegCM_BATS and RegCM_CLM, respectively, whilst it is 38.2 (34.5) and 49.1 mm for CRU (TRMM) and ERA-Int_ICBC, respectively. This indicates clearly that RegCM4 (with the GFC convective scheme) simulates rainfall better with the BATS land-surface scheme than with CLM in the CORDEX-MENA/Arab domain.

Figure 4. Annual rainfall averaged over the period 2001–2005 for (a) RegCM_BATS, (b) RegCM_CLM, (c) ERA-Int driving, (d) CRU and (e) TRMM data. Simulated data are obtained with GFC scheme for the domain D2.
In obtaining the correlation between the simulated and the observed rainfall, and in determining which land-surface scheme provides better correlation with the observation data, a scatter plot of monthly rainfall comparing the simulation with the observation is displayed in Figure 6. It is evident that the regression slope with BATS (1.07) is better than the slope with CLM (0.33), and the correlation value with BATS (0.96) is higher than with CLM (0.82). The data range for BATS (CLM) is 80.71 (25.28), whilst it is 70.52 for CRU. The SD for BATS (CLM) is 19.37 (7.07), whilst it is 17.25 for CRU. The RMSD for BATS (CLM) is 4.83 (23.45) with reference to CRU. These statistical measures indicate that BATS performs better than the CLM in the simulation of rainfall for the CORDEX-MENA/Arab region.

In the analysis domains, particularly in the CORDEX-MENA/Arab domain, the simulated surface air temperature (at 2 m) is large in West and in East Africa, and in the eastern Arabian Peninsula (Figure 7). The spatial distributions of mean air temperature simulated by RegCM4 with BATS in the GFC convective scheme option (Figure 7(a)) show somewhat warmer temperatures compared with the corresponding CLM (Figure 7(b)). For example, the temperatures over the eastern Arabian Peninsula are above 30°C for the simulations with BATS but such temperatures are almost absent in the simulations with CLM (mostly below 30°C). Although some differences are visible region to region and option to option, the overall simulated temperature distribution patterns are similar to the ERA-Int driving field (Figure 7(c)) and the CRU observation (Figure 7(d)). However, the simulated temperatures show finer details than the ERA-Int driving field and are closer to CRU. This indicates that RegCM4 improves the temperature simulation over CORDEX-MENA/Arab domain. Further investigations on the rank and cause of these improvements are out of scope of this paper and are discussed in a subsequent study.

The simulated temperatures extracted from 11 sub-domains (and averaged from all of them) show that temperature follows well the annual cycle of the observed data in each year during 2001–2005 (Figure 8). It is evident that the simulation with CLM has a large cold bias in each month for all years during 2001–2005, whilst the simulation with BATS has a small cold bias, with respect to CRU. Importantly, RegCM4 has downgraded the temperature with CLM, compared with the ERA-Int driving forcings. Although the simulation with BATS shows slight enhancement, nevertheless, it is closer to the observation and performs better than CLM in the simulation of temperature for the CORDEX-MENA/Arab domain.

The scatter plot for temperature comparing the simulated with the observed data (for the entire domain) shows that temperature follows well the annual cycle of the observed data in each year during 2001–2005 (Figure 9). The regression slope with BATS (1.07) is close to unity, whilst for CLM (1.12) it is steeper than unity (1.0). The data range with BATS (CLM) is 14.46 (15.66), which is closer to (further from) CRU (13.66). SD with BATS (CLM) is 4.67 (4.91),
Figure 7. Daily temperature averaged over the period 2001–2005 for (a) RegCM_BATS, (b) RegCM_CLM, (c) ERA-Int driving and (d) CRU data. Temperatures are obtained for the domain D2.

Figure 8. Time sequences of monthly mean daily temperature averaged over 11 sub-domains obtained for observations (CRU) and simulations (BATS and CLM) with ERA-Int driving field (ERA-Int ICBC) and GFC convective scheme for the period 2001–2005. Temperatures are obtained for the domain D2.

whilst it is 4.35 for CRU. Also, the RMSD for BATS (CRU) is 0.88 (1.52) with reference to the CRU data. These statistical measures clearly indicate that the BATS land-surface scheme is better than CLM in the simulation of temperature over the CORDEX-MENA/Arab domain.

For further confirmation in determining which land-surface option (BATS or CLM) in RegCM4 is better for the CORDEX-MENA/Arab domain, a few statistical measures (such as bias, correlation between observed and simulated values, RMSD and SD) are considered for each sub-domain, as displayed in Figures 10 and 11.

As shown in Figure 10, the rainfall bias (in %) indicates that the simulations with CLM are underestimated by about 50% for most of the sub-domains, whilst the simulations with BATS are very close to the observation data for many sub-domains; however, overestimations for a few
sub-domains (e.g. C and E) and underestimation for a few sub-domains (e.g. A, H and K) are noticeable. When averaged from all sub-domains, the simulations with BATS show a dry bias of about 13%, whilst the simulations with CLM show a dry bias of about 47%, with reference to the CRU rainfall. Note that the rainfall amounts are large for southern sub-domains C, E, F and I but small for the other sub-domains (see Figure 2). The rainfall bias with BATS is low for F and I compared with CLM; indeed, the rainfall simulation generates small bias for BATS compared with CLM for the entire domain (all, Figure 10(a)). The correlation coefficients (Figure 10(b)) between the observed and the simulated rainfall show strong relationships for all sub-domains, except for insignificant ones (<0.40) for A (BATS) and G and H (CLM). RMSD is relatively low with BATS compared with CLM for all sub-domains, which is also reflected in the data for the entire domain (all, Figure 10(c)). The SDs for the rainfall time series indicate that the BATS values are closer to the CRU data when compared with CLM (Figure 10(d)).

For temperature, the simulations with BATS show a small bias (in °C) for all sub-domains, compared with CLM (Figure 11(a)). The correlation coefficients between the observed and the simulated temperature show strong relationships for both BATS and CLM for all sub-domains, except for a weak relationship for southern region F (Figure 11(b)). However, smaller RMSDs are observed for all sub-domains for the temperature simulations with BATS compared with CLM, except for large values for B, C, E and I; thus, RMSD is lower with BATS when averaged across all sub-domains (Figure 11(c)). The SDs for the temperature time series for BATS are closer to CRU compared with CLM (Figure 11(d)). In the summary of statistical measures on the rainfall and temperature time series, as presented in Table 2, the BATS land surface performs better than CLM for most of the sub-domains, except for a few sub-domains (A for rainfall, and C, E and I for temperature) wherein CLM is better than BATS. Hence, Table 2 provides a suitable land-surface scheme for each sub-domain/region. However, when averaged over all 11 sub-domains, the BATS land surface performs realistically well compared to the CLM in the simulation of rainfall and temperature.

3.4. Downscaling of CMIP5 AOGCMs

After the selection of optimal settings which include the domain D2, land-surface scheme BATS and convection scheme GFC, the data of another reanalysis (i.e. NNRP2) and five GCMs, downscaled with RegCM4, are analysed in order to further examine these settings. Moreover, in order to check the sensitivity of RegCM4 with SST, an additional simulation is carried out using ERA-Int with ERSST. The results presented in Figure 12 indicate that RegCM4 has done a very reasonable job in downscaling NNRP2 dataset with results for both rainfall and temperature staying closer to the observations. The model has also shown satisfactory results in downscaling ERA-Int with both ERSST and OISST, with the similar results to the observations. The downscaled data of four CMIP5 models of five closely follow the patterns of the observed monthly rainfall and temperature as well. One CMIP5 model (CanESM2) substantially underestimate rainfall (Figure 12(a)) and overestimated temperature (Figure 12(b)); however, the patterns of rest of the 4 CMIP5 models data are quiet promising in downscaling the long-term and projection period data using RegCM4.

Figure 13 shows the patterns of surface level pressures and wind fields of reanalysis datasets (ERA-Interim and NNRP2) and one AOGCM (ECHAM6), against the downscaled datasets using RegCM4 with optimal settings. The patterns of pressure and winds stay closer among the reanalysis and AOGCM datasets. Moreover, RegCM4 has done a very reasonable job in capturing the behaviour of pressure and wind fields, giving us further confidence in the simulation with the aforementioned settings.

3.5. Analysis of the length of simulation

In order to address the issues concerning analysis of data over a shorter period of 5 years, a longer simulation of 10 years is also conducted with the similar RegCM4 setting (D2, GFC and BATS). Therefore, the results of the rainfall and temperature simulations for the period 1997–2007 (1997 is spin-up year) are also discussed.

The statistical measures for mean, bias, correlation coefficient (r), RMSD and SD against the CRU values for the two simulation periods (2001–2005 and 1998–2007) are summarized in Table 3. In case of rainfall, the CRU mean is 38.32 (37.79) and SD is 18.87 (17.54) for the period 1998–2007 (1997 is spin-up year) are also discussed. The ERA-Int driving field values presented in Table 3, it can be concluded that almost all of the statistical measures show improved results for RegCM4.
Figure 10. Statistical measures of rainfall at 11 sub-domains (A, B ... K) and averaged over all sub-domains (all) for (a) bias, (b) correlation, (c) RMSD and (d) standard deviation. Data are obtained for the domain D2.

Figure 11. Statistical measures of temperature at 11 sub-domains (A, B ... K) and averaged over all sub-domains (all) for (a) bias, (b) correlation, (c) RMSD and (d) standard deviation. Data are obtained for the domain D2.
Table 2. Statistical measures for rainfall and temperature with the GFC option for both BATS and CLM land-surface processes, for the period 2001–2005 in the best domain (D2).

<table>
<thead>
<tr>
<th>Sub-domain</th>
<th>Rainfall</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>Correlation</td>
</tr>
<tr>
<td>A</td>
<td>CLM</td>
<td>CLM</td>
</tr>
<tr>
<td>B</td>
<td>BATS</td>
<td>SIMILAR</td>
</tr>
<tr>
<td>C</td>
<td>BATS</td>
<td>SIMILAR</td>
</tr>
<tr>
<td>D</td>
<td>BATS</td>
<td>SIMILAR</td>
</tr>
<tr>
<td>E</td>
<td>BATS</td>
<td>SIMILAR</td>
</tr>
<tr>
<td>F</td>
<td>BATS</td>
<td>BATS</td>
</tr>
<tr>
<td>G</td>
<td>BATS</td>
<td>BATS</td>
</tr>
<tr>
<td>H</td>
<td>CLM</td>
<td>BATS</td>
</tr>
<tr>
<td>I</td>
<td>BATS</td>
<td>BATS</td>
</tr>
<tr>
<td>J</td>
<td>BATS</td>
<td>SIMILAR</td>
</tr>
<tr>
<td>K</td>
<td>BATS</td>
<td>BATS</td>
</tr>
<tr>
<td>All</td>
<td>BATS</td>
<td>CLM</td>
</tr>
</tbody>
</table>

The best values corresponding to each sub-domains are marked with their land-surface schemes (BATS or CLM) e.g. highest values for correlation and lowest values bias. If both land-surface options provide similar values, then it is named as SIMILAR.

Figure 12. Time sequences of (a) monthly rainfall and (b) monthly mean temperature averaged over 11 sub-domains obtained for RegCM4 simulations with ECHAM6, GFDL, HadGEM2, CNRM, CanESM2 and observations (CRU and TRMM) with Era-Int and NNRP2 data. The simulated data are for GFC convection scheme with BATS in D2 for the period 2001–2005.

© 2015 The Authors. International Journal of Climatology published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

4. Summary and conclusions

The aim of this paper is to find suitable domain, as well as better performing convection and land-surface schemes within RegCM4 to downscale CMIP5 database for climate change impact studies in the entire Arab region. For this purpose first, seven sensitivity experiments for the period (2001–2005) are performed over seven domains all covering the entire Arab region. In this case, the initial and lateral boundary conditions to RegCM4 are provided from ERA-Interim dataset. The results indicate that among the seven domains the model simulations show small differences compared to observations in terms of rainfall and temperature simulations. However, the domain D2 (7°S–45°N and 27°W–76°E) is providing small bias for both rainfall and temperature simulation which may be attributed to its better coverage from Atlantic Ocean to Indian Ocean along the longitudes, and Mediterranean Sea to equatorial Africa along the latitudes. This domain (D2) is named CORDEX-MENA/Arab domain, which is considered to be the most cost effective and suitable domain. Further, by conducting eight experiments for suitable convection and land-surface schemes in RegCM4, statistical measures indicate that between the two land-surface schemes available in RegCM4, the simulations with BATS are found to be better for most of the 11 sub-domains.

Figure 13. Surface level pressure superimposed with wind field for driving ERA-Int, NNRP2, ECHAM6 (left panels; top, middle and bottom, respectively), and RegCM4 ERA-Int, NNRP2, ECHAM6 (right panels; top, middle and bottom, respectively). The data are averaged for the period 2001–2005.
and for the entire CORDEX-MENA/Arab domain. Moreover, among the four convective parameterization schemes tested in this study, the GFC is found to perform better over the CORDEX-MENA/Arab domain. These optimal settings are also validated by forcing the model with another reanalysis NNRP2, and by changing the Era-Interim SSTs from OISST to ERSST. These additional simulations further supported the previous findings. The issue concerning the length of the data is also addressed in the study, where the optimal settings are also validated against a relatively longer time scale of ten years, producing satisfactory results. Finally, five CMIP5 global climate simulations are downscaled by using RegCM4 with above optimal domain, convection and land-surface schemes settings. Of five downscaled simulations, four shows similar patterns compared to the observations in terms of monthly rainfall and temperature variations. However, one AOGCM (CanESM2) largely underestimated (overestimated) rainfall (temperature) for all through the period of analysis.

Therefore, the proposed CORDEX-MENA/Arab domain (7°S–45°N and 27°W–76°E) with BATS land-surface and GFC convection schemes is a recommended settings for downsampling the CMIP5 long-term database using RegCM4 for future studies. However, the detail investigation on convection scheme for wetter and dry regions as well for different seasons in downsampling CMIP5 database using RegCM4 for the CORDEX-MENA/Arab may be interesting to discuss in future studies.

Acknowledgements

The author acknowledge the grant by the NSTIP strategic technologies program in the Kingdom of Saudi Arabia – Project No. 12-ENV3197-03 to complete this work – and the Science and Technology Unit, King Abdulaziz University for technical support. RegCM4 group of ICTP, Trieste, Italy, is acknowledged for providing the model with LBCs. The League of Arab States (LAS) and the United Nations Economic and Social Commission for Western Asia (UN-ESCWA) are also acknowledged for leading and facilitating the “Regional Initiative for the Assessment of Climate Change Impacts on Water Resources and Socio-economic Vulnerability in the Arab Region (RICCAR)”. Carol Chouchani Cherfane and Tarek Sadek of UN-ESCWA are acknowledged for their valuable efforts in making RICCAR platform effective.

References


M. ALMAZROUI


© 2015 The Authors. International Journal of Climatology published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.