Evaluation for Moroccan dynamically downscaled precipitation from GCM CHAM5 and its regional hydrologic response

Tsou Jaw a,*, Jialun Li a, Kuo-lin Hsu a, Soroosh Sorooshian a, Fatima Driouech b

a Center for Hydrometeorology and Remote-Sensing (CHRS), University of California, Irvine, Irvine, CA 92697, USA
b Climate Studies Service, Direction de la Météorologie Nationale, Casablanca, Morocco

A R T I C L E   I N F O

Article history:
Received 14 January 2014
Received in revised form 26 December 2014
Accepted 6 February 2015
Available online 21 March 2015

Keywords:
Dynamical downscaling
Moroccan precipitation
Regional hydrology

A B S T R A C T

Study region: Morocco (excluding Western Sahara).
Study focus: This study evaluated Moroccan precipitation, dynamically downscaled (0.18-degree) from three runs of the studied GCM ECHAM5/MPI-OM, under the present-day (1971–2000/20C3M) and future (2036–2065/A1B) climate scenarios. The spatial and quantitative properties of the downscaled precipitation were evaluated by a verified, fine-resolution reference. The effectiveness of the hydrologic responses, driven by the downscaled precipitation, was further evaluated for the study region over the upstream watershed of Oum er Rbia River located in Central Morocco.

New hydrological insights for the region: The raw downscaled runs reasonably featured the spatial properties but quantitatively misrepresented the mean and extreme intensities of present-day precipitation. Two proposed bias correction approaches, namely stationary Quantile-Mapping (QM) and non-stationary Equidistant CDF Matching model (EDCDFm), successfully reduced the system biases existing in the raw downscaled runs. However, both raw and corrected runs projected great diversity in terms of the quantity of future precipitation. Hydrologic simulations performed by a well-calibrated Variable Infiltration Capacity model successfully reproduced the present-day streamflow. The driven flows were identified highly correlated with the effectiveness of the downscaled precipitation. The future flows were projected to be markedly diverse, mainly due to the varied precipitation projections. Two of the three flow simulation runs projected slight to severe drying scenarios, while another projected an opposite trend for the evaluated future period.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

Twentieth-century climate change, induced by anthropogenic forcings, has been recognized as one of the most significant factors influencing the development of human beings and their activities. Moreover, the impacts of climate change on future water resources, particularly over semi-arid regions of which the socioeconomic developments are facing difficulties from severe water scarcities, draw significant scientific attention (Abdulla et al., 2009; Arnell, 2004; Fang et al., 2007; Gao and Giorgi, 2008; Sanchez and Subiela, 2007; Vicente-Serrano, 2007). Among the semi-arid regions, the northwest African country of Morocco is especially notable for its vulnerability to climate change, while Moroccan agriculture and pasturing have been suffering from severe water resources deficits during long-term drought periods over the past few decades (Born et al., 2008; Chbouki et al., 1995; Sowers et al., 2011; Swearingen, 1992). To analyze the evolution of climate change and assess its impacts, model-derived climate data from General Circulation Models (GCMs) or Regional Climate Models (RCMs), as well as currently accessible limited observations, are utilized extensively. For example, Knippertz et al. (2003) linked the Moroccan long-term precipitation variability to large-scale circulation patterns using monthly precipitation from the Global Historical Climatology Network (GHCN) observations and the European Centre/Hamburg Model (ECHAM). Born et al. (2008) investigated the present-day and future precipitation variability by comparing ground observations and RCM outputs and concluded that northwestern Africa will continue to face drying and warming trends in the future. Paeth et al. (2009) analyzed the potential effects of land cover and land-use changes (LCLUs) on the regional climate over western Africa using downscaled data at 50-km spatial resolution under IPCC AR4 scenarios A1B and B1. They concluded that feedback from LCLUs is perhaps the second most important factor (i.e., minor to the changes in the tropical oceans) of regional climate change in the study area. Driouech et al. (2009) also evaluated the rainy-season precipitation variability from 1971 to 2000 over Morocco using ARPRGE-CLIMATE GCM. They found that the model is successful in reproducing precipitation frequency and interannual variability of the climate regimes, but seemed to misrepresent the precipitation quantity along the Atlantic coast and the longest dry spells over southern Morocco. The significant findings in the above studies have proven the importance and success of the numerical model applications in climate research. However, one significant question remains: How accurate are the numerical model results over the studied regions? Despite some recent efforts made to improve climate-data resolution at 50 km (Driouech et al., 2009; Paeth et al., 2009), most climate studies over Morocco and its vicinity interpreted by coarse GCM output (150–200 km) are still less capable of adequately presenting variable properties of local climate, in particular, precipitation spatial and temporal distributions over this region at fine resolutions. From a hydrological application viewpoint, fine-resolution precipitation is required to assess hydrological responses to climate changes over the regions with complex terrain and land-cover features (such as Moroccan high–mountain areas). Because current observations are insufficient (Boudhar et al., 2010), fine-resolution data to be generated from RCMs with coarse-resolution GCM forcings are considered as a good alternative. For the past several decades, downscaling techniques have been rapidly developed and made good progress toward benefiting climate studies for regions worldwide, including Morocco. For example, Vizy and Cook (2001) investigated summer rainfall variability and its connection with Sea Surface Temperatures (SST) through downscaling ECMWF data to 120 km using the fifth-generation Penn State/NCAR mesoscale model (MM5). A dynamical-statistical approach was adopted by studies (Huebener and Kerschgens, 2007a,b) to downscale future scenarios at 3-km resolution for a 15,000 km² region over the southern part of the Atlas Mountains. Summarizing from the above downscaling studies, the following conclusions can be addressed: (1) the accuracy of precipitation in space and time can be essentially improved by applying downscaling techniques; however, the improvement degrees varied case by case; (2) downscaling results might perform very diversely in accordance with the adopted schemes (e.g., downscaling approaches, configurations, models) and could also be very inconsistent, depending on the scheme selections; and (3) the resolutions of existing downscaled climate data sets over Morocco and its vicinity are still insufficient and inapplicable for effective basin-scale hydrologic assessment nationwide. As a result of the advancement of climate models with regard to physical processes, numerical algorithms, and computing power, the accuracy and efficiency of modeling results have been continuously improved. However, general performances of dynamical downscaling models need further investigation in terms
of precipitation reproduction and projection. The uncertainties inherent in GCM-forcing data and downsampling parameterization selections could result in considerable biases, even if a perfect climate model is used (Ines and Hansen, 2006). Directly applying downscaled outputs to the hydrologic applications can result in biases propagating through the processes of hydrologic simulations (Kotlarski et al., 2005). In order for the purposes to strengthen the robustness of downscaled outputs, post-correcting procedures in support of diminishing systematic biases might be helpful (Kotlarski et al., 2005; Teutschbein and Seibert, 2010). Although some statistical downsampling techniques (Schmidli et al., 2007; Themessl et al., 2011) are also capable of adjusting the raw variable (however, at the point scale where observations are available), most dynamical downsampling studies over Morocco did not take the correction steps into consideration.

Precipitation is known as one of the key elements in the hydrologic cycle, as well as a prominent indicator of climate change. Even though climate models have become indispensable tools in retrieving fine-resolution precipitation, climate-change studies still put more effort into investigating their large-scale precipitation patterns over semi-arid regions. Motivated by the intention of effectiveness verification (e.g., quantity and spatial distribution of historical reproductions) and potential change forecast of various downscaled precipitation data sets under semi-arid climates, this study conducted experiments, including precipitation downscaling and bias-corrections, for a large part of Morocco. In addition, a case study to test the effectiveness of hydrologic modeling using downscaled driving forces was performed for a regional-scale basin in central Morocco. Suggested by the above-mentioned climate downscaling study over Morocco (Driouech et al., 2010), insufficient observations for model calibration and high variability under future climates make statistical downscaling with inherent stationary hypotheses likely to derive less reliability for projecting Moroccan precipitation. For this reason, dynamical downscaling, dominated by more meaningful physics, is adopted to process future climate projections with high variability in this study. Coarse GCM data collected from three runs (here named Er1, Er2, and Er3) from ECHAM5/MPI-OM (referred to as ECHAM5; Roeckner et al., 1996) are collected and pre-processed to provide boundary and initial conditions to drive the chosen dynamical downscaling model, MM5. Downscaled precipitation data at 18-km resolution over a present-day period (1971–2000) and a future period (2036–2065) are post-processed with two bias corrections. Afterwards, using a semi-distributed hydrologic model, namely, the Variation Infiltration Capacity (VIC, Liang et al., 1994, 1996), the hydrologic responses to be forced by MM5 outputs with/without bias-correction are examined over a regional-scale basin located at the upper Oum er Rbia River in central Morocco under present-day and near-future climates.

2. Data sets

GCM climate data sets at coarse resolutions and precipitation reference data sets at fine resolutions were collected and pre-processed to be used in the precipitation downscaling, bias correction, and performance evaluation.

2.1. GCM climate data sets

The IPCC-adopted GCM, ECHAM5/MPI-OM (hereafter ECHAM5), is chosen to be the climate data source of the precipitation downscaling. Developed by the Max Planck Institute for Meteorology, ECHAM5 consists of an atmospheric component with a horizontal resolution of T63 (horizontal resolution of ~200 km) and 31 vertical layers, with the uppermost level located at 10 hPa (Roeckner et al., 1996). ECHAM5 owns three modeling runs (namely Er1, Er2, and Er3) to account for the uncertainties of model initial conditions of the year of 1860. This GCM was previously adopted by Huebener and Kerschgens (2007a, b) for a relevant study of precipitation patterns over the southern Morocco region.

2.2. Precipitation reference

Reliable reference is necessary and critically important in supporting downscaling evaluations and bias correction in this study. Currently available ground observations within Morocco, however, are relatively limited and not sufficient or fine enough to evaluate downscaled precipitation over the
extensive study domain during the designated time period. Three selected reference candidates are preliminarily evaluated. The one best agreeing with limited ground observations is determined to serve as the reference. The first data set, produced by the Climate Research Unit of the University of East Anglia TS3.0 (referred to as CRU), provides a monthly-based data set at a resolution of 0.5° interpolated by scattering ground observations (Mitchell and Jones, 2005). The second data set is a daily-based data set at a resolution of 0.5°, and is derived from observation-driven model simulations from NOAA/CPC (referred to as CPC). The last one is a 0.18-degree daily-based data set, dynamically downscaled by the fifth-generation Penn State/NCAR mesoscale model from NCEP/NCAR reanalysis 1 data (2.5°), involving climate information from historical observations (referred to as NCEP-d).

Due to observation access limits, this study evaluates reference candidates by comparing the annual magnitude and distribution patterns of all candidates with a gauge-based precipitation distribution map (Fig. 1) produced by the Moroccan Meteorological Service (Direction de la Météorologie Nationale, hereafter referred to as DMN), over the period of 1971–2000. Relatively, CRU and CPC over the entire Moroccan northern territory tend to overlook many details of the spatial distribution patterns, due mainly to their coarser 0.5° resolutions. The two data sets moreover underestimate precipitation over
mountainous watersheds (e.g., the Rif and Atlas mountains) where streamflow and snowmelt account for Morocco’s major water resources. In contrast, the downscaled NCEP reanalysis (NCEP-d) at the finest resolution (0.18°) seems to be more plausible with regard to either annual quantity or spatial distribution. The NCEP-d therefore is considered the best reference candidate and is adopted for the grid-to-grid downscaling evaluation and bias-correction implementations.

3. Downscaling configuration

The fifth-generation Penn State/NCAR mesoscale model (MM5) was selected to perform the dynamic downscaling for Moroccan precipitation in this study. By conducting initial sensitivity test runs for three years from 1981 to 1983, the MM5 physical schemes, including simple-ice explicit microphysics (Dudhia, 1989), Eta PBL (Janjić, 1994), RRTM radiation (Mlawer et al., 1997), the new Kain–Fritsch convection scheme (Kain and Fritsch, 1990) and the Noah LSM (Chen and Dudhia, 2001), were finally determined. Three nested domains (D01, D02, and D03, shown in Fig. 2), centered at 31° N and 5° W, were constructed under the consideration of the designed grid sizes and available computational resources. The largest nested domain (D01) covers most of northwestern Africa and southern Europe, with a total of 55 × 70 grid cells at the resolution of 162 km. The second domain (D02), with less coverage of the northern and the southern territories of Morocco, owns a total of 43 × 43 grid cells at the resolution of 54 km. The innermost domain (D03) only covers the northern Moroccan territory with 67 × 70 grid cells at the resolution of 18 km. In particular, D03 encompasses the entire Atlas Mountain region, which captures the dominant topographic features as well as the snowpack-covered areas of many headwater basins.

As shown in Table 1, downscaled climate data sets in the present-day period of 1971–2000 under 20C3M scenario and in the future period of 2036–2065 under SRES A1B scenario were collected for bias correction and downscaling evaluation uses. The present-day period was further divided into two sub-periods for downscaling skill verification. The earlier sub-period from 1971 to 1985 was used as the control period, while another sub-period from 1986 to 2000 was used as the validation period. Both the raw and corrected downscaling runs were then evaluated by the reference in the same period.
Table 1

Time period segmentations assigned for the downscaling evaluation. The present-day period is divided into the control (1971–1985) and validation (1986–2000) sub-periods to facilitate the evaluation and bias correction application based on the historic reference. The projection period (2036–2065) is corrected by the proposed correction approach using the assigned control period (1971–2000).

<table>
<thead>
<tr>
<th>Present-day (20C3M scenario)</th>
<th>Future (A1B scenario)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Downscaling</td>
<td>Control</td>
</tr>
</tbody>
</table>

4. Precipitation bias correction

Downscaled climate could be misrepresented due to systematical biases inherent in adopted GCMs and procedures of downscaling. Effective post-processing corrections are commonly used to reduce potential misrepresentations. Over the past few decades, various types of bias-correction approaches, which were different from their levels of complexity, have been developed and applied by many climate and hydrologic studies. A study that comprehensively compared many commonly-adopted bias corrections was conducted by Teutschbein and Seibert (2010).

Statistical stationarity is a widely-adopted assumption in either simple or advanced approaches. The statistical inference and parameters derived by reference data in a control period (e.g., historically observed events) are assumed to remain unchanged during the correction procedures done for raw data throughout another target period. Among various approaches holding this assumption, the Quantile-Mapping (referred to as QM) probably is most representative due to its simplicity and applicability. According to the main concept of QM, the distribution parameters of raw data sets (e.g., mean, variance, and other high-order moments) are modified to match those parameters derived from the selected “reference” (e.g., observations). Besides, this approach does not change the rank correlation of the original data. A general expression of this approach can be written as:

$$x_{\text{cor}} = \text{CDF}^{-1}_{\text{obs,hist}}(\text{CDF}_{\text{mod}}(x_{\text{mod}}))$$  \hspace{1cm} (1)

In Eq. (1), $x_{\text{mod}}$ and $x_{\text{cor}}$ represent the modeled and corrected data values of an interested variable $X$. The signs, $\text{CDF}_{\text{obs,hist}}$ and $\text{CDF}_{\text{mod}}$ denote the cumulative distribution functions comprised of historical observation in the “control” period and model output to be corrected in the target period, respectively. The notation $\text{CDF}^{-1}(\cdot)$ represents the inverse mode of CDF. The model output however, can be derived from either historical runs or future model runs (e.g. future projections), if the distribution parameters of observed data from the control period are still considered representative of the reality in the modeled time period. In other words, the modeled data sets, even from different modeled periods, are assumed to be generated from the same distribution functions. This assumption implies that the distribution properties of the interested variable are unchanged, even if the modeled data set is not in the adopted control period.

This simple approach has been extensively and successfully utilized for precipitation bias correction (Déqué, 2007; Ines and Hansen, 2006; Piani et al., 2010), climate downscaling (Quintana-Segui et al., 2011; Segui et al., 2010; Themessl et al., 2011) and hydrologic (Déqué, 2007; Hashino et al., 2007; Iizumi et al., 2011; Li et al., 2010; Quintana-Segui et al., 2011) studies and is employed to correct the raw downscaling runs.

Developed on the basis of QM approach (Li et al., 2010), the equidistant CDF matching model (referred to as EDCDFM) is another bias approach employed in this study. Featured with the non-stationary assumption, this approach is particularly used to adjust model output in a projection period, while another set of model output in a historical period is already available. The mathematic expression of this approach is written as:

$$x_{\text{cor}} = \text{CDF}^{-1}_{\text{obs,hist}}(\text{CDF}_{\text{mod,proj}}(x_{\text{mod}})) + x_{\text{mod}} - \text{CDF}^{-1}_{\text{obs,hist}}(\text{CDF}_{\text{mod,proj}}(x_{\text{mod}}))$$  \hspace{1cm} (2)

The subscript proj in the above equation denotes the projection period. The signs, $\text{CDF}_{\text{obs,hist}}$ and $\text{CDF}_{\text{obs,proj}}$ represent the cumulative distribution functions comprised of modeled data in the historical
period and projection period, respectively. In order to stand for the non-stationary features, this approach applies the differences of the modeled values at the same data percentile between the projection period and historical period to the original QM approach.

More detailed explanation and applications of the EDCDFm approach are provided by Li et al. (2010). Here, a schematic illustration shown in Fig. 3 conceptually addresses the differences of two bias correction approaches. Fig. 3a shows that QM simply shifts the modeled data $X_{\text{mod}}$ (e.g., daily precipitation) horizontally from $CDF_{\text{mod}}$ to $CDF_{\text{obs,hist}}$. Note that the raw modeled and corrected values are at the same percentile before and after applying the correction. EDCDFm (Fig. 3b), on the other hand, adopts a similar idea, but further applies a non-stationary correction term to Eq. (1). Shown in Eq. (2), this non-stationary correction is expressed as $[(X_{\text{mod}} - CDF^{-1}_{\text{mod},\text{hist}}(CDF_{\text{mod},\text{proj}}(X_{\text{mod}})))$ and can be schematically viewed as the horizontal differences between projection and historical model CDFs (hollow circle and hollow triangle in Fig. 3b).

To implement the bias corrections, the empirical cumulative distribution functions for each month are subjectively created by all wet-day precipitation data within the desired periods. Accordingly, a threshold of 0.5 mm/d for the observation data suggested by Driouech et al. (2009) is applied to differentiate dry/wet days. The threshold for the modeled precipitation of each grid cell on the other hand is determined individually. The threshold of a particular model cell is determined to be the value on the model CDF curve accounting for a specific percentile that 0.5 mm/d represents on the observation CDF curve. The corrections are performed individually for each month, which means that all daily data from each month are corrected by a same unique correction scheme.

5. Downscaled precipitation evaluation

5.1. Evaluation index

The precipitation of wet seasons over the high Atlas Mountains contributes the primary Moroccan water resources for the spring irrigation (Boudhar et al., 2010; Parish and Funnell, 1999) and is
Fig. 4. Moroccan topography map of at 25-km resolution. (a) Topographic distribution. (b) Two regions of low-elevation (LO, grids in blue) and high-elevation (HI, grids in red) distinguished by the grid-mean elevation of 1500 m. (For interpretation of references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 5. North Africa precipitation (mm/d) overlie topography in various spatial models. The central region (HI, hereafter) is distinguished into two groups by the grid-mean elevation of 1500 m. As shown in Fig. 4b, the high-elevation region (HI, hereafter) contains 93 grid cells, and the low-elevation region (LO, hereafter) contains the remaining 534 grid cells over the northern Moroccan territory.

Three proposed indices are used to evaluate the skills of downscaled precipitation reproduction over the HI and LO regions, separately. The wet-season intensity ($I_{\text{wet}}$) is defined as the mean intensity (mm/d) of all wet days in a wet season from October to March suggested by Driouech et al. (2010). The extreme intensity ($PQ_{95}$) is defined as the daily intensity (mm/d) of the wet-day precipitation at the 95th percentile over the evaluation period. The annual wet days ($WD$), on the other hand, represent the annual wet days over the evaluation period. To quantify this index, a threshold $0.5 \text{ mm/d}$ is set to distinguish the wet/dry days. The evaluation indices are computed for each individual grid cell over the innermost downscaling domain. Two statistics, bias error ($\%\text{BIAS}$) and determination of correlation ($R^2$) of the indices of all grid cells covering the LO and HI regions, are also computed to numerically conclude the skills of downscaled precipitation in terms of the model performance in quantity and spatial distribution, respectively.

5.2. Evaluation for the uncorrected present-day downscaled precipitation

Comparing with the reference NCEP-d, the evaluation using proposed indices for the downscaled precipitation in the control and validation periods is performed and shown in Table 2. Under the present-day climate, all raw downsampling precipitation data sets (Er1-d, Er2-d, and Er3-d hereafter) seem to reasonably capture the spatial features of Moroccan precipitation shown by the reference. Nevertheless, the raw data sets display varying degrees of disagreement regarding the index quantities. As for the index $I_{\text{wet}}$, three raw data sets display a common pattern of overestimation, especially in the validation period, and Er1-d demonstrates more serious overestimations (> +20%) over HI regions in both control and validation periods. The evaluation for the index $PQ_{95}$ indicates a similar pattern. All runs demonstrate positive biases, but the high-elevation region (HI) tends to be more seriously overestimated in both periods. Despite those significant biases in indices $I_{\text{wet}}$ and $PQ_{95}$, all runs considerably preserve their spatial patterns interpreted by the statistics $R^2$ (shown in Table 2), except for Er3-d, which shows more mismatches over the southeastern desert region (see Fig. 5). The evaluation for the annual wet days ($WD$), on the other hand, illustrates a common pattern of underestimation.
Fig. 5. Precipitation downscaling evaluation using indices $I_{wet}$, $PQ_{95}$ and WD in (a) the control period 1971–1985 and (b) the validation period 1986–2000. NCEP-d data serves as the reference of three raw (uncorrected) downscaling runs (Er1-d, Er2-d and Er3-d).
Table 2
Evaluation of raw downscaling runs in the control and validation periods. The more skilled indicators (defined by $R^2 > 0.9$ and $|\%BIAS| < 10\%$) are highlighted. The less skilled indicators (defined by $R^2 < 0.6$ and $|\%BIAS| > 20\%$) are bolded.

<table>
<thead>
<tr>
<th>Downscaling run</th>
<th>Er1-d</th>
<th>Er2-d</th>
<th>Er3-d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluation index</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period</td>
<td>Region</td>
<td>Statistics</td>
<td>$I_{wet}$</td>
</tr>
<tr>
<td>Control</td>
<td>LO</td>
<td>$R^2$</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$R^2$</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>LO</td>
<td>$%BIAS$</td>
<td>3.2%</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$%BIAS$</td>
<td>$20.5%$</td>
</tr>
<tr>
<td>Validation</td>
<td>LO</td>
<td>$R^2$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$R^2$</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>LO</td>
<td>$%BIAS$</td>
<td>7.3%</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$%BIAS$</td>
<td>38.0%</td>
</tr>
</tbody>
</table>

Table 3
Evaluation of corrected downscaling runs in the validation period. The more skilled indicators (defined by $R^2 > 0.9$ and $|\%BIAS| < 10\%$) are highlighted. The less skilled indicators (defined by $R^2 < 0.6$ and $|\%BIAS| > 20\%$) are bolded.

<table>
<thead>
<tr>
<th>Downscaling run</th>
<th>Er1-d</th>
<th>Er2-d</th>
<th>Er3-d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Evaluation index</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correction</td>
<td>Region</td>
<td>Statistics</td>
<td>$I_{wet}$</td>
</tr>
<tr>
<td>QM (Validation)</td>
<td>LO</td>
<td>$R^2$</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$%BIAS$</td>
<td>$-4.2%$</td>
</tr>
<tr>
<td>EDCDFm (Validation)</td>
<td>LO</td>
<td>$R^2$</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$%BIAS$</td>
<td>3.2%</td>
</tr>
<tr>
<td></td>
<td>LO</td>
<td>$%BIAS$</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$%BIAS$</td>
<td>16.7%</td>
</tr>
<tr>
<td></td>
<td>LO</td>
<td>$R^2$</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>$%BIAS$</td>
<td>$29.6%$</td>
</tr>
</tbody>
</table>

over both regions in both periods. However, $WD$ is least misestimated by Er1-d, and Er3-d, giving greatest underestimation (over ~20% over both regions in both periods; see Table 2), performs the least favorably.

An evaluation rule regarding modeling skills is used to comprehensively evaluate the spatial and quantitative performances among the downscaling runs. The values of determination of correlation greater than 0.9 ($R^2 > 0.9$) and absolute bias error less than 10% ($|\%BIAS| < 10\%$) are to be considered as “more-skilled” indicators (as highlighted in Table 2). In contrast, the values of determination of correlation less than 0.6 ($R^2 < 0.6$) and absolute bias error greater than 20% ($|\%BIAS| > 20\%$) are to be considered as “less-skilled” indicators (as bolded in Table 2).

Under this evaluation rule, it is observed that each downscaling run has certain skilled and less-skilled indictors, meaning no downscaling run can be concluded to be overwhelmingly superior to the others. Despite the result, uncorrected Er1-d seems to be more skilled in term of the numbers of more skilled indicators, particularly in the index $WD$. However, considerable biases in the indices $I_{wet}$ and $PQ_{55}$ still exist and suggest post-processing correction is necessary to derive realistic downscaling output.

5.3. Evaluation for the corrected present-day downscaling results

The evaluation results for the corrected downscaling runs in the validation period are shown in Table 2. Similarly, more-skilled (highlighted) and less-skilled (bold) indicators in Table 3 are used to
quantify the effectiveness of corrected precipitation. As for the indices \( l_{\text{wet}} \) and \( PO_{95} \), the common over-estimations, particularly over the HI region, are effectively mitigated by both correction approaches (Table 3). Relatively speaking, the corrections for the annual wet days (WD) are less effective. This is mainly because the rain/no-rain physics is controlled primarily by the downscaling model. In addition, the bias-correction approaches that simply modify the raw dry/wet frequencies to match the reference frequencies based on the defined dry/wet threshold in the control period are less influential regarding this index. Altogether, the more-skilled indicators (%\text{BIAS} and \( R^2 \)) suggest that the corrected downscaling runs are effectively modified in terms of not only the quantitative biases but also the misrepresented spatial features of the indices.

Overall, the proposed corrections improve the raw downscaling runs, but the improvement degrees regarding any specific index highly correlate with the selection of correction approaches and raw downscaling runs. By inter-comparing the skilled and unskilled statistics of each index among the correction combinations, corrected Er1-d and Er3-d runs are identified as the most- and least-skilled runs. However, the performance differences between the two correction approaches are highly related to the raw downscaling runs. QM-corrected Er1-d and Er3-d are generally more skilled, but two approaches seem to have similar effects on Er2-d.

### 5.4. Future (2035–2065) assessment under the SRES A1B scenario

Taking the present-day (1971–2000) NCEP-d data as the reference, the changes of future (2035–2065) precipitation under the SRES A1B scenario are assessed by the built evaluation approach and indices adopted for the present-day evaluation presented in Section 5.3. The distribution maps (Fig. 6) and statistics (Table 4) of index–percent changes between the two periods conceptually illustrate the trends of future precipitation properties. To make those significant changes stand out, notable increasing and decreasing signals are highlighted in Table 4 by bold (i.e., \( +>20\% \)) and underlined (i.e., \( <-20\% \)) numbers. The findings regarding those changes are described as follows:

Shown in the top panel in Fig. 6, greater diversities of wet-day mean precipitation \( (l_{\text{wet}}) \) are identified in both raw and corrected runs. Raw Er1-d projects the strongest decreasing trend over the entire Moroccan territory \( (-15.23\% \text{ over LO and } -7.86\% \text{ over HI}) \); raw Er3-d, on the contrary, projects increasing signals \( (10.41\% \text{ over LO and } 29.3\% \text{ over HI}) \) over the high-central Atlas Mountains and the southeastern territory. Raw Er2-d, placed in between raw Er1-d and Er3-d, shows mild reductions of \(-8.44\% \text{ over LO and } 8.89\% \text{ over HI} \). QM-corrected Er1-d and Er2-d runs project relatively weak signals of increasing \( l_{\text{wet}} \) for both HI and LO regions, in comparison with the EDCDFm-corrected runs, which show strong \( (\text{Er1-d}) \) to mild \( (\text{Er2-d}) \) decreases mainly due to the effect of non-stationary data properties. For Er3-d, both corrected runs consistently project the same signal of slightly decreasing \( l_{\text{wet}} \).
Fig. 6. Percent change maps of raw and corrected evaluation indices (wet-season intensity, extreme intensity and annual wet days from top to bottom panels) between future (A1B, 2036–2065) and present-day (reference) climate.
Shown in the middle panel in Fig. 6, the changing trends of extreme precipitation intensity, \( (PQ_{95}) \) are inconsistently projected among the raw runs in terms of their spatial distributions. Raw Er3-d projects the most notable increasing trends over the entire study domain, except for the southern coastal region. Raw Er2-d projects milder increases over the HI region (+11.65%) and insignificant reduction over the LO region (−5.22%). In contrast, the trends projected by raw Er1-d seem to be less relevant to the orography (−3.24% over LO and +1.90% over HI). QM-corrected runs tend to project much milder changes in extreme precipitation in comparison to EDCDFm-corrected runs. The decreasing trend projected by raw Er1-d is magnified by EDCDFm correction, while the other two EDCDFm-corrected runs tend to weaken the trends projected by the raw runs.

Three raw runs project a common reduction trend of annual wet days (WD). In particular, Er1-d shows the most serious reduction trend (see the bottom panels in Fig. 6). The differences between each downscaling run corrected by two corrections are insignificant, but the projections are extremely diverse among the three runs. Seriously decreasing WD projected by raw Er1-d run is only slightly modified by the bias corrections. Instead, both corrected Er3-d runs turn the raw projection into an opposite trend of increasing WD. Both corrected Er2-d runs commonly project very minor decreasing trends of the index WD between the present-day and future periods (i.e., less than 5% for both LO and HI regions).

6. Hydrologic response of downscaled precipitation: a case study for the upstream watershed of Oum er Rbia River

The evaluations for Moroccan precipitation by directly comparing downscaled precipitation with the adopted reference was elaborated in the previous sections. This section further evaluates the downscaled precipitation by assessing its hydrologic response for a regional-scale watershed located in central Morocco. Among the major river systems in Morocco, the longest is the (550 km) Oum er Rbia River (OER) originating in the high Atlas Mountains and drains nearly 50,000 km\(^2\) basin area. A series of dams were built on the river from upstream to downstream to facilitate water supplies on irrigation, hydropower generation, and domestic uses for Moroccan major cities and agricultural zones. In view of the importance of OER to Moroccan water resources, the upstream watershed of OER above the flow gauging station at Oulad Sidi Driss (shown in Fig. 7) was chosen for this evaluation.

Long-term streamflow as the indicator of hydrologic responses is simulated by the chosen Variable Infiltration Capacity model (VIC, Liang et al., 1994, 1996), which has been extensively utilized in many other large- or regional-scale hydrologic assessments (Bowling and Lettenmaier, 2010; Guo et al., 2009; Wood et al., 2002; Zhou et al., 2004). The Variable Infiltration Capacity model (VIC) is essentially a macro-scale, semi-distributed, and grid-based hydrologic model. Comprised of multi-layer infiltration, surface and subsurface runoff, and routing schemes, this model can be coupled with other climate models. The fundamental meteorological forcings, including gridded precipitation, maximum air temperature, minimum air temperature, and wind speed (from the raw or corrected downscaling runs), as well as necessary ground information, such as the parameters of soil and vegetation layers (from GLDAS) along with DEM (from the U. S. Geological Survey) are prepared.

Effective hydrologic simulations rely on appropriate model calibration. In this study, the model calibration was conducted using NCEP-d reference as the “truth forcings” in the present-day period. Although most VIC parameters are measurable or calculable by collected soil and vegetation information for the study watershed, appropriate calibration is still highly recommended for those indirectly measurable and highly sensitive parameters in order to model the target realistically.

Seven parameters, including the infiltration shape parameter \( (B_{inf}) \), maximum velocity of baseflow \( (D_{max}) \), fraction of \( D_{max} \) where non-linear baseflow begins \( (D_e) \), fraction of maximum soil moisture where non-linear baseflow occurs \( (W_s) \), and the thicknesses of three soil layers, are the commonly-suggested parameters to be calibrated (Liang et al., 1996). Streamflow observation at two gauging stations, Dechra el Oued (DEC) and Oulad Sidi Driss (OUL), within the study area were compared to the calibration model run. Thus, the calibration takes two sequential steps to account for the properties of flow contributing areas. The first step of calibration was to calibrate grid cells in the area above the upstream DEC as the first group. The remaining area above the downstream station OUL as the second group was subsequently calibrated. A popular optimization-searching method, namely, SCE-UA (Duan et al.,...
1992), was adopted to screen out the most plausible parameter set. In addition, Nash Sutcliffe Efficiency (NSE) of monthly streamflow was adopted as the indicator of goodness of fit for the simulation results.

To minimize the storage and operation effects caused by the constructed dams, the flow observation of the earliest 12 water years (1971–1982) in the entire available data period is presumed to be mostly naturalized and served as the calibration reference. As shown in Fig. 8, the evaluation statistics (i.e., root mean squared error (RMSE), Nash Sutcliffe efficiency (NSE), and percent bias (BIAS) shown above the hydrograph lines) suggest the calibrated monthly flow simulations at stations DEC and OUL (dashed lines) fairly match the long-term observation (solid lines). Though some underestimated flow peaks are identified, from long-term and regional-scale points of view, the calibrated flow satisfactorily reproduces essential hydrologic cycle properties (i.e., highest/lowest timing and quantity and rising/recession patterns).

The hydrologic simulations, using the calibrated parameters under the present-day and future climate scenarios, are subsequently performed and evaluated. To evaluate simulation skills throughout the entire present-day period (1971–2000), the flow reference at two stations driven by the most realistic meteorological forcing (NCEP-d as data) are utilized. Shown in Fig. 9, streamflow driven by three downscaled forcings under three correction conditions (uncorrected, QM and EDCDFm) at two gauging stations are compared with the flow reference. The variability of flow simulation runs basically is consistent with the results shown among the downscaled precipitation runs. QM-corrected Er1-d, which is the most realistic precipitation data set in the present-day also is the most effective flow-driving forcing. The decreasing trend of precipitation represented by Er1-d runs between present-day and future periods induces a severe flow reduction. Furthermore, great quantitative diversity is identified among streamflow simulation runs and is highly correlated to the precipitation forcing diversity. As previously discussed, the forcing diversity is greater among the selected runs, in comparison with the diversity among the correction options. Er3-d run driven streamflow in contrast projects a wetter future, particularly during the winter seasons. This conclusion is consistent with the results implied in the downscaling precipitation evaluation.
Fig. 8. Hydrologic model calibration results of two gauging stations Dechra el Oued (top) and Oulad Sidi Driss (bottom). The inner figures show observed and simulated monthly cycles.

A comprehensive evaluation using indices RMSE (shown in the left panels) and BIAS (shown in the middle panels) for present-day simulations is demonstrated in Fig. 10. It summarizes the performance of downscaled precipitation in terms of the effectiveness of driven streamflow. The evaluations for the raw flows at both gauging stations identify similar statistical characteristics among the forcing runs. Quantitatively speaking, uncorrected Er3-d featured with less-biased evaluation indices is considered the most realistic forcing among three uncorrected runs. The other two raw forcings tend to mislead flows, resulting in more mismatches presented by higher RMSE and BIAS values. In particular, raw Er1-d produces the most unrealistic flows, highly overestimating annul flows for both stations.

The corrected forcings demonstrate differential effectiveness in simulated flows. QM-corrected forcings are relatively effective in terms of flow simulation skills. In particular, the biased flows driven by raw Er1-d and E2-d are substantially reduced. The corrections therefore do not further improve Er3-d driven flow, which already outperforms other uncorrected forcings. Besides, the performance of present-day streamflow simulations implies EDCDFm-corrected forcings are generally less effective than QM-corrected forcings. As mentioned previously in the precipitation evaluation, EDCDFm-corrected precipitation is overestimated over high-elevation regions in the wet seasons (shown in Table 3). This result agrees with the conclusion made in the corrected precipitation evaluation.

Shown in the right panels of Fig. 10, isotropic trends, including and excluding the corrections, are identified in the future flow projections. Using the present-day (1971–2000) flow reference as the baseline, the trend projections suggest significant contradiction among the flow simulations driven by different forcings, no matter whether they are uncorrected or corrected. The long-term (30-year) averaged monthly flow (shown in Fig. 9) driven by projected Er1-d exhibits the driest condition in the future (2036–2065) under the A1B scenario. In this case, the annual flow production of the study area will decrease over 75% (shown in the right panel of Fig. 10), even when projected by the most
Fig. 9. Simulated flow comparison at the stations of Dechra el Oued (top) and Oulad Sidi Driss (bottom). The solid lines represent simulated flow driven by the reference (Ref), uncorrected (Uncorr), QM-corrected (QM) and EDCDFm-corrected (EDCDFm) forcings under the present-day climate (1971–2000) respectively. The dashed lines represent the results under the future climate (2036–2065, A1B).

Fig. 10. Quantitative evaluations for downscaled forcings driven flow simulations at stations DEC (top panels) and OUL (bottom panels). Left and middle panels show percent bias (%BIAS) and root-mean-squared error (RMSE) of the present-day simulations. Right panel shows the annual flow changes (%) between the future (A1B, 2036–2065) and present-day (20C3M, 1971–2000) period.
optimistic simulation driven by the uncorrected Er1-d. A common milder magnitude of the decreasing trends (−11%, −14% and −20%) projected by flow simulations driven from uncorrected, QM-corrected and EDCDFm-corrected forcings, respectively) is recognized in the analysis of flows driven by Er2-d. In contrast, flow simulations driven by Er3-d project a most optimistic water resource scenario in the future of the study area. Average streamflow during the future period (2036–2065) is projected to be increased by 30% (uncorrected) to 73% (QM-corrected). The changing precipitation patterns, particularly the annual wet days, are greatly responsible for the inconsistent projected trends of future streamflow. This discrepancy might likely result from original distinct in the GCM projections.

7. Discussion

The success of the proposed corrections in reducing the biases in downscaled precipitation have been identified in this study. Particularly, the notably misrepresented Er1-d improved by the QM approach is considered the most successful case. However, the statistical evaluation for present-day Moroccan precipitation implies the EDCDFm approach seems to be less effective for all raw downscaled runs. Real stationary/non-stationary properties cannot be effectively interpreted by the downscaled data between the control and model periods. The length of 15 years (1971–1985), adopted for the control period used to incorporate the effects of changing trends, might be too short. This result leads to two constructive points: (1) downscaled precipitation data in the present-day period tends to demonstrate stronger stationarity in our study, and the quantitative changes between the distributions of 15-year control and target periods seem less significant; and (2) more complicated correction approaches are not necessarily more effective, particularly when real stationary/non-stationary features, inherent in the changes between data distribution of the control and target periods, are not genuinely interpreted.

The precipitation projection (2036–2065) displays diverse results among uncorrected model runs. Overall, the uncorrected Er1-d projects the most serious drying trend indicated by the predicted reduction of the daily mean intensity and annual wet days. Though varying in terms of the change degrees in mean and extreme daily intensities, the projections given by the corrected Er1-d runs predict similar reduction trends of annual wet days, which agree with the conclusion suggested by other relevant studies in this region. In comparison, the projection variability caused by the correction selection is less significant in Er2-d and Er3-d runs. This implies that the projection diversities among three downsampling runs are greater comparing to the variability caused by the applications of bias corrections.

The effectiveness of present-day flow simulation is highly associated with the performance of downscaled forcings. The flow driven by the forcing of QM-corrected Er1-d run shows the most plausible simulation in spite of that the uncorrected Er1-d is featured with significant quantitative biases. Therefore, the downsampling runs featured by better mean and extreme precipitation in general seem to perform more effectively in driving short-term flow simulations. The downsampling runs simulating better wet-day reproduction tend to produce less biased annual streamflow. As discussed, higher diversities among different, downscaled forcing runs would lead to more variable flow projections. In contrast, the projection variability caused by the correction applications (i.e., no correction, QM, and EDCDFm) is less notable in relation. Greatly diverse future wet-day projections among the three corrected runs are responsible for the streamflow projections. Taking the present-day streamflow as the reference, the projection for the study watershed indicates the most significant streamflow reduction is driven by the corrected Er1-d runs highlighted by seriously decreasing wet days.

Though this study focuses on the precipitation downsampling evaluation, the effect of temperature change is another important factor when studying the hydrologic impacts on regional scales. Temperature change featured as a form of energy exchange is highly linked to atmospheric circulation and precipitation patterns. Changes in precipitation patterns further influence quantitative and temporal properties of surface runoff. By interchanging the states of water, temperature can also influence the hydrologic cycle directly. While the data was prepared for the hydrologic modeling in this study, the surface maximum and minimum temperatures were also downscaled and corrected. Shown in Fig. 11, the change trends of areal mean wet-season (October to March) temperature and precipitation of the study watershed between the present-day and future periods indicate their interconnections with
regional flow projections. The Er1-d runs (including corrected or uncorrected) overall project higher temperature increases in both daily minimum (2.0–3.1 °C) and maximum (5.2–6.5 °C). Accordingly, the areal precipitation seems to result in the driest condition (reduction of 66.8–81.7%). The total effects of increasing temperature and decreasing precipitation result in a total flow reductions as shown in Fig. 10. On the other hand, the daily minimum temperature changes (+0.8 to 1.4 °C) and daily maximum temperature changes (−0.6 to −0.2 °C) display the least significant temperature changes projected by Er3-d runs. While the precipitation is projected to increase by 29.1–47.8%, the least warm and wetter scenario responds to increasing flow projections.

8. Summary and conclusion

This study evaluated high-resolution (0.18-degree) Moroccan precipitation, dynamically down-scaled from three runs of the selected GCM ECHAM5, under the present-day (1971–2000 20C3M) and the future (2036–2065 A1B) climate scenarios by two evaluation approaches. One directly evaluated the spatial and quantitative features of downscaled precipitation at the fine resolution of 0.18-degree using NCEP reanalysis as the reference. Another indirectly evaluated the downscaled precipitation by assessing the effectiveness of hydrologic response for the study watershed. Both evaluation approaches concluded the raw present-day downscaling runs perform diversely and misrepresent quantitatively the precipitation intensities and wet-days, and the future precipitation was projected to be even more diverse.

By using proposed stationary (QM) and non-stationary (EDCDFm) correction approaches, the present-day precipitation is effectively improved. However our study illustrated that improvement skills for the future precipitation can be varying among different raw runs and extremely dependent on the data accuracy of the adopted control run and the model predictability for the changing trends. Despite the fact that accurately quantifying the magnitudes of stationary/non-stationary properties to date is still limited by modeling skills and resolutions, it is highly recommended to apply both types of correction approaches to prevent potential systematic biases inherent in the raw, downscaled data. Except for the variability that resulted from the adoption of the corrections, greater variability existing among the raw precipitation runs and driven flow simulations was identified in this study. Significant uncertainty among multi-model predictions is still a challenge toward assessing climatic and hydrologic impacts. While this study has identified that correction effectiveness is greatly related to the quality of the raw downscaling runs, the selection of adopted GCM ensembles on which the dynamic downscaling technique is applied is one of the most important factors to derive promising and reliable climate and hydrologic impact assessment.

Acknowledgements

This study is financially supported by the World Bank project “Impacts of Global Climate Change (GCC) on the Water Resources of Morocco, Phase I: Dynamical Downscaling GCM Outputs in Morocco".
Special thanks are also given to Direction de la Météorologie Nationale of Morocco for providing gauge precipitation data.

References


