ORIGINAL ARTICLE

Assessment of the utility of dynamically-downscaled regional reanalysis data to predict streamflow in west central Florida using an integrated hydrologic model

Syewoon Hwang · Wendy D. Graham ·
Alison Adams · Jeffrey Geurink

Abstract The goal of this study was to evaluate the ability of dynamically downscaled reanalysis data to reproduce local-scale spatiotemporal precipitation and temperature data needed to accurately predict streamflow in the Tampa Bay region of west central Florida. In particular, the Florida State University Center for Ocean-Atmospheric Prediction Studies CLARReS10 data (NCEP DOE 2 reanalysis data (R2) downscaled to 10-km over the Southeast USA using the Regional Spectral Model (RSM) were evaluated against locally available observed precipitation and temperature data and then used to drive an integrated hydrologic model that was previously calibrated for the Tampa Bay region. Resulting streamflow simulations were evaluated against observed data and previously calibrated model results. Results showed that the raw downscaled reanalysis predictions accurately reproduced the seasonal trends of mean daily minimum temperature, maximum temperature and precipitation, but generally overestimated the monthly mean and standard deviation of daily precipitation. Biases in the temporal mean and standard deviation of daily precipitation and temperature predictions were effectively removed using a CDF-mapping approach; however, errors in monthly precipitation totals remained after bias correction. Monthly streamflow simulation error statistics indicated that the accuracy of the streamflow produced by the bias-corrected downscaled reanalysis data was satisfactory (i.e., sufficient for seasonal to decadal planning), but that the accuracy of the streamflow produced by the raw downscaled reanalysis data was unsatisfactory for water resource planning purposes. The findings of this study thus indicate that further improvement in large-scale reanalysis data and regional climate models is needed before dynamically downscaled reanalysis data can be used directly (i.e. without bias correction with local data) to drive hydrologic models. However, bias-corrected dynamically downscaled data show promise for extending local historic climate observation records for hydrologic simulations. Furthermore, results of this study indicate that similarly bias-corrected dynamically downscaled retrospective and future GCM projections should be suitable for assessing potential hydrologic impacts of future climate change in the Tampa Bay region.

Keywords Dynamically-downscaled reanalysis · Regional climate model · West central Florida · Bias correction · Integrated hydrologic model

S. Hwang (✉) · W. D. Graham
Water Institute, University of Florida, PO Box 116601,
Gainesville, FL, USA
e-mail: aceace111@ufl.edu

S. Hwang · W. D. Graham
Department of Agricultural and Biological Engineering,
University of Florida, Gainesville, FL, USA

A. Adams · J. Geurink
Tampa Bay Water, Clearwater, FL, USA

Introduction

Development of accurate hydrologic model predictions for current and potential future climatic conditions requires reliable spatially and temporally distributed climatic information (Hay and Clark 2003; Jasper et al. 2002). However, the availability and quality of climate data varies from region to region, and observations are sparse and

irregularly distributed in many areas (Kanamitsu and Kanamaru 2007; Haberlandt and Kite 1998; El-Sadek et al. 2011). To overcome these limitations, reanalysis studies, that merge climate model predictions and observation data, have been undertaken to develop long-term historical spatially and temporally distributed climatic variables such as precipitation, temperature, solar radiation, humidity and wind speed (Bromwich and Fogt 2004; Jonathan et al. 2011).

Reanalysis data provide a reconstruction of the historical global atmospheric circulation patterns through assimilation of available observation data into Global Climate Model (GCM) predictions. Thus, reanalysis data are widely considered to be a robust proxy of historic atmospheric observations that reproduce observed climate dynamics and are therefore useful for climate application studies where there are inadequate meteorological input data. Furthermore, verifying accurate prediction of historic climatic and hydrologic behavior using reanalysis data is an essential first step before using retrospective and future GCM projections to predict potential hydrologic impacts of future climate change.

Previous studies have shown that the coarse resolution of large-scale reanalysis data (commonly greater than 100 km grid size) precludes its direct application to assess hydrologic impacts at the local to regional watershed scales (Beck et al. 2004; Iizumi et al. 2011). A number of techniques have been developed to bridge the gap between the resolution of large-scale climate information and the resolution required for local applications (Fowler et al. 2007; Wilby and Wigley 1997). Dynamical downscaling uses Regional Climate Models (RCMs) driven by boundary conditions from large-scale GCMs to simulate detailed regional climate information (Giorgi et al. 2001; Mearns et al. 2003). Many studies have shown that dynamical downscaling has potential to effectively capture nonlinear mesoscale features that are absent in the global reanalysis (e.g. Lim et al. 2011; Lo et al. 2008). This approach produces physically consistent climate information at higher resolution than global reanalysis data or GCM outputs and thus has been applied to produce regional reanalyses at finer ($\sim$10–50 km) spatial resolutions.

Mesinger et al. (2006) developed the North America Regional Reanalysis (NARR), a 25-year regional reanalysis with a horizontal resolution of 32 km. They used the National Center for Environmental Prediction Department of Energy 2 (NCEP DOE 2) 200 km reanalysis data (Kanamitsu et al. 2002, hereinafter referred to as R2) as lateral boundary conditions for the NCEP Eta regional scale model (Mesinger et al. 1988) and assimilated available local-scale observations into the regional model predictions. Kanamitsu and Kanamaru (2007) pointed out that spatial resolution of NARR (i.e., 32 km) may still not be

satisfactory for application requirements and provided dynamically downscaled regional reanalysis at 10 km spatial resolution over the state of California using the Regional Spectral Model (RSM). Subsequently, Stefanova et al. (2011) conducted dynamically downscaled reanalysis over the Southeast United States at 10 km spatial resolution using both R2 and the ECMWF's reanalysis data (European Centre for Medium-range Weather Forecasts— 40 year reanalysis, hereafter referred to as ERA40; Uppala et al. 2005). These regional reanalysis data, namely CLARReS10 (the COAPS (Center for Ocean-Atmospheric Prediction Studies) Land–Atmosphere Regional Reanalysis downscaling for the Southeast United States at 10-km resolution) provides hourly data for various atmospheric variables including precipitation and temperature for use in local agricultural, hydrologic and coastal applications. Stefanova et al. (2011) examined the ability of CLARReS10 outputs to reproduce realistic temporal structures of precipitation (e.g., average annual cycle, diurnal cycle, and summertime precipitation variability) for several subregions (approximate size 300 km by 200 km each) over the Southeast USA from 1979 to 2001. They found that the dynamically downscaled reanalysis results were in good agreement with station and gridded observations in terms of both the relative seasonal distribution and the diurnal structure of precipitation, although total precipitation amounts were systematically overestimated.

In this study, the ability of the CLARReS10 R2 results to accurately predict historic streamflow in the Tampa Bay region, west central Florida was evaluated. The specific objectives of this study were to (1) evaluate the raw CLARReS10 R2 products at hydrologically significant space and time scales over the Tampa Bay region (2) investigate the utility of bias correction to improve the CLARReS10 regional reanalysis results, and (3) investigate the ability of raw and bias-corrected CLARReS10 results to reproduce historic streamflow using an integrated hydrologic model previously developed and calibrated for the Tampa Bay Region (Ross et al. 2004, 2005; Geurink and Basso 2012). The study is an essential first step toward the long-term goal of using COAPS's dynamically downscaled CMIP5 GCM model projections (called CLAREnCE10) to predict potential hydrologic impacts of future climate change in the Tampa Bay region.

Study area

Tampa Bay is the largest estuary in Florida and extends about 50 km inland from the Gulf of Mexico. In this region, a diverse water supply system that includes regional well fields, river withdrawals from the Hillsborough and Alafia Rivers and a seawater desalination plant are

operated by Tampa Bay Water. Accurate prediction of seasonal, interannual, and decadal climate variability, as well as potential long-term climate change, is important to water resources planning and management in the region. Average annual rainfall for the study period from 1989 to 2001 is approximately 1,230 mm, and evapotranspiration (ET) is estimated to be approximately 70 % of annual precipitation in the area (Ross et al. 2005). A variety of land cover types are present in the study area, including urban, grassland, forest, agricultural, mined land, water, and wetlands. Of these, open water and wetlands cover 25 % of the region.

Hydrologic model

There are strong interactions among surface, subsurface, and ET processes in the Tampa Bay area due to complex geology and relatively flat topography in the region (Geurink and Basso 2012). In order to understand and predict the dynamic surface–groundwater interactions in this complex system, two regional water management agencies (Tampa Bay Water and Southwest Florida Water Management District (SWFWMD)) jointly developed an integrated surface/subsurface hydrologic model for the area (Integrated Hydrologic Model, IHM). IHM couples the EPA Hydrologic Simulation Program-Fortran (HSPF; Bicknell et al. 2001) and the USGS MODFLOW96 (Harbaugh and McDonald 1996) for surface and groundwater modeling, respectively (Geurink et al. 2006). The model is characterized as a deterministic, semi-distributed, and semi-implicit with variable time steps and spatial discretization (Ross et al. 2004). Subsequently, the Integrated Northern Tampa Bay model application (hereafter refer to as INTB) was developed using IHM to improve hydrologic assessment capabilities of west central Florida by Tampa Bay Water. The hydrologic model domain for INTB is bordered by the Gulf of Mexico in the west (Fig. 1), the north and east boundaries of the model domain follow Floridan aquifer flow lines, and the southern boundary is placed far enough from the area of interest for this study to minimize the influence of the general head boundary (Geurink et al. 2006).

For the surface water component, the model domain is discretized into 172 basins based on surface drainage (Fig. 1). For each basin, hydrologic processes are simulated within hydrologic response units (land segments) based on five upland land-use categories (Ross et al. 2004). Basin-specific meteorological data are required as climate inputs. To construct the input data for INTB calibration, areally averaged daily precipitation was estimated from more than 300 point observations available over the study area and distributed to the 172 basin centroids using the Thiessen polygon method (Geurink and Basso 2012). Within INTB

preprocessing routines, daily precipitation values for each basin are disaggregated to 15 min precipitation totals using the observed pattern of the nearest observation station with historic 15 min observations that approximates the daily rainfall volume (Geurink and Basso 2012). For temperature, maximum and minimum temperature data at 6 stations over the domain were used to develop reference ET (Hargreaves and Samani 1985) time series, then those were spatially assigned to the nearest neighbor basins to define the potential evaporative demand for each basin. The INTB model was calibrated and verified for the Northern Tampa Bay Region using hydrologic observations from 1989 to 2006 (Geurink and Basso 2012). However for this study, we are only able to evaluate INTB simulations from 1989 to 2001 because the CLARReS10 dynamically downscaled reanalysis data are only available through 2001.

CLARReS10 products

CLARReS10 consists of downscaled data from two different reanalyses: the ERA40 and R2. The dynamical model used for the regional downscaling in CLARReS10 was the regional spectral model (RSM developed at NCEP; Juang and Kanamitsu 1994). The global reanalysis fields were supplied as lateral boundary conditions to the RSM at 6-hourly intervals. The resulting products cover the Southeast United States at 10-km resolution for the period from 1979 to 2001. More details about model setup and configuration for CLARReS10 are included in Stefanova et al. (2011).

CLARReS10 provides grid-based daily maximum and minimum temperature data and hourly climate predictions (for two-dimensional variables such as temperature and precipitation). Additionally, data for three dimensional variables (i.e., wind speed, humidity, etc.) are also available at three-hourly temporal resolution. All CLARReS10 data are archived in netCDF format and freely available at http://coaps.fsu.edu/CLARReS10/index.shtml. This study focused on evaluating the CLARReS10 R2 downscaled results. Hourly precipitation and daily maximum and minimum temperature data were retrieved, and daily precipitation data were assembled by aggregating hourly precipitation data.

Methodology

Estimation of basin-based precipitation from gridded data

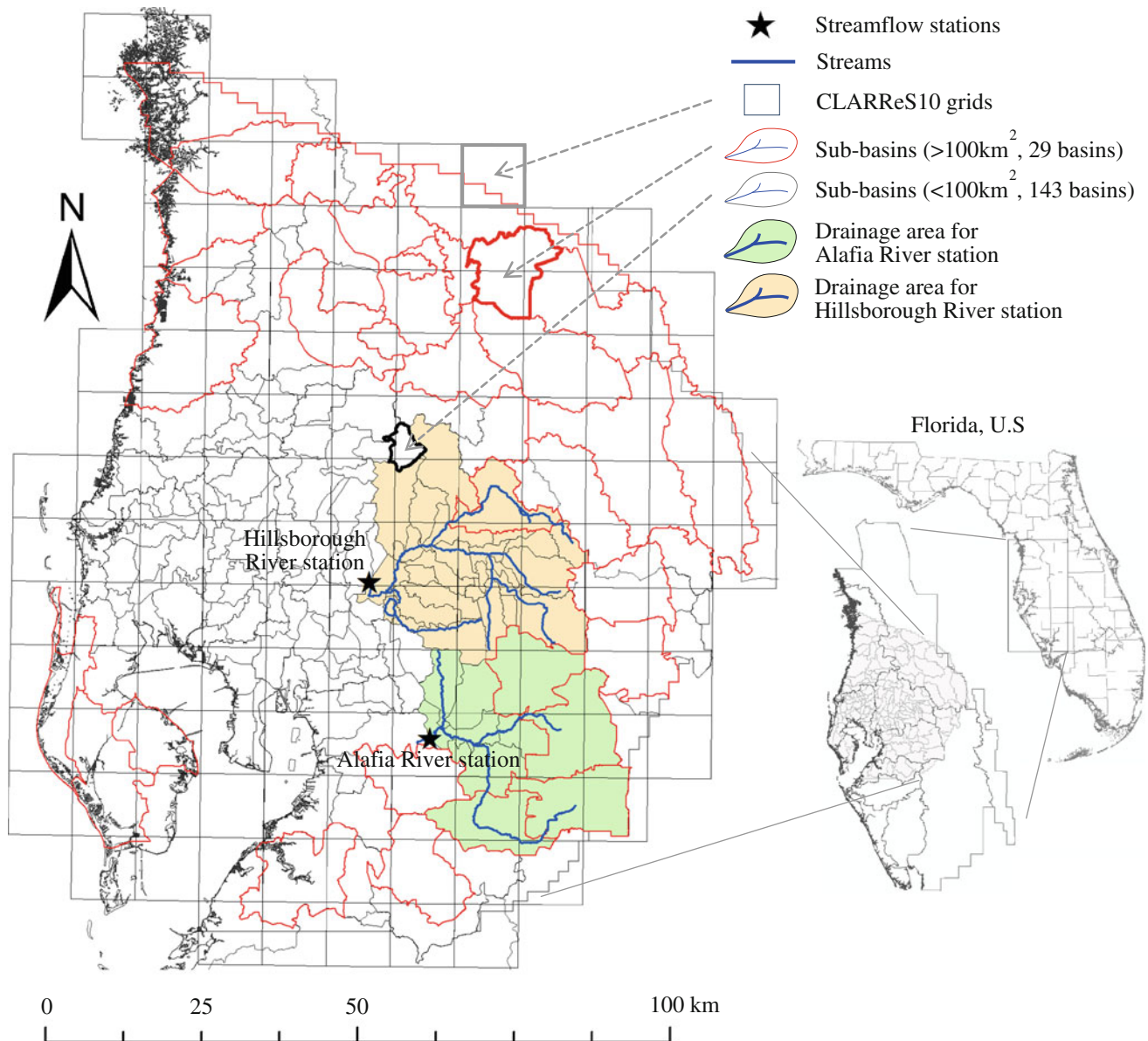CLARReS10 predicts precipitation on a 10-km rectangular grid, whereas INTB requires specification of areally

Fig. 1 Map of study domain including the extent of the 172 sub-basins, the locations of the streamflow target stations and the CLARReS10 grid for regional reanalysis. *Colored* areas indicate the contributing areas for the streamflow target stations, and *red* polygons indicate the sub-basins where the area is greater than unit area of the CLARReS10 grid (100 km^2)

averaged rainfall over the 172 irregularly sized basins (3.4–362.7 km^2; including 29 basins over 100 km^2 (the CLARReS10 grid area) as shown in Fig. 1) which comprise the study area. An areal weighting method was used to estimate daily precipitation for each of the 172 basins from the gridded precipitation data. For each basin, the grid cells that overlapped with the basin were identified (Fig. 1), and the fraction of the basin area covered by each grid cell was calculated. These fractions were used to weight the precipitation predicted for each grid cell, and the weighted precipitation predictions were then summed to get the total basin rainfall. This process was repeated for each of the 172 basins.

Statistical bias correction

Basin-based precipitation predictions estimated from the raw CLARReS10 R2 results (Downscaled Reanalysis R2, hereafter DR_R2) were bias-corrected using the CDF-mapping method (Panofsky and Brier 1968). This approach has been widely used for application of climate modeling outputs for impacts assessment (e.g., Wood et al. 2004; Piani et al. 2010; Hwang et al. 2011; Themeßl et al. 2012). The procedure is briefly described as follows: (1) daily basin-based precipitation observations and basin-based precipitation predictions were gathered for each calendar month; (2) CDFs of observations and predictions were

generated for each basin by month; (3) the percentile of each basin-based predicted value (from the predicted CDF) was mapped to the CDF of basin-based observations with the same percentile for the appropriate month; (4) the observed basin-based precipitation value for the predicted percentile was used as the 'bias-corrected value'. Steps (3) and (4) were repeated for each daily prediction, and the bias-correction procedure was conducted independently for each calendar month. These bias-corrected downscaled CLARReS10 results are referred to as BCDR_R2 in this paper. A schematic representation of bias-correction process and examples of CDFs for DR_R2 and BCDR_R2 results are presented in Online Resource 1.

For maximum and minimum daily temperature predictions (hereafter T_{max} and T_{min}), daily CLARReS10 predictions for the grid cells containing the six temperature observation stations were used. The same bias-correction process described for precipitation was conducted for the mean temperature (i.e., $T_{mean} = (T_{max} + T_{min})/2$) and the range of diurnal temperature (i.e., $T_{range} = (T_{max} - T_{min})$). Then, T_{max} and T_{min} were reconstructed from the bias-corrected T_{mean} and T_{range} as: bias-corrected $T_{max} =$ bias-corrected T_{mean} +bias-corrected $T_{range}/2$; and bias-corrected $T_{min} =$ bias-corrected $T_{mean} - T_{range}/2$. This approach avoids the possible, but unrealistic, result that bias-corrected T_{min} could be greater than bias-corrected T_{max} and also allows the bias-corrected results to preserve observed daily temperature range ($T_{max} - T_{min}$). This bias-correction method for temperature is similar to the one suggested by Piani et al. (2010).

Streamflow target stations

To compare the hydrologic simulations using the dynamically downscaled reanalysis data to those using the observed climatic data, we selected two streamflow stations on the Alafia River and Hillsborough River, based on their importance to water supply operations/management in the study area. Tampa Bay Water is permitted to withdraw surface water from the Hillsborough and Alafia Rivers for public water supply and reservoir restoration when streamflow is above ecologically acceptable thresholds. Tampa Bay Water uses monthly and seasonal streamflow estimates to plan their surface water withdrawals and reservoir operations. Daily to weekly operational surface water decisions are made based on actual streamflow measured at USGS stations, as specified in their permit. Thus, accurately predicting monthly streamflow response to climate variability and change is important for both monthly to annual and longer-term (decadal to multi-decadal) planning timeframes. Figure 1 shows the locations of target stations and their contributing drainage areas.

Evaluation criteria

To evaluate the CLARReS10 climate predictions, the following statistics were calculated for the raw and bias-corrected predictions as well as the observed data:

- The temporal mean and temporal standard deviation of daily precipitation, maximum temperature and minimum temperature by calendar month over all basins, and for the contributing areas of the two target stations, were calculated to evaluate the accuracy of the annual cycle of mean daily rainfall and temperature predictions.
- The percent bias (PBIAS: $100 \times E[\text{sim-obs}]/E[\text{obs}]$ where $E[x]$ is the expected value of variable x), Root Mean Square Error (RMSE) to observed STandard DEViation (STDEV) ratio (RSR: RMSE/STDEV[obs]), coefficient of determination R^2 and the Nash–Sutcliffe efficiency (NSE: 1- MSE[sim-obs]/MSE[obs] where MSE[x] is mean square error of variable x; Nash and Sutcliffe 1970) between observed and predicted monthly precipitation were calculated over the two target station drainage basins to evaluate the accuracy of monthly predictions.

To evaluate the hydrologic simulations, the following statistics were calculated:

- Observed and simulated monthly streamflow were plotted over the study period for the two target stations.
- The PBIAS, RSR, R^2 and NSE between observed and predicted monthly streamflow for the two target stations were calculated to evaluate the accuracy of monthly predictions. Based on the error statistics, hydrologic model performance for each climate input was judged as very good, good, satisfactory, or unsatisfactory using performance ratings suggested by Moriasi et al. (2007).
- Mean daily streamflow was calculated over the study period by calendar month to evaluate the accuracy of the annual cycle of mean monthly streamflow predictions for the two target stations.

In addition, to assess the relationship between errors in the precipitation and streamflow predictions, correlation coefficients between the following were calculated:

- Correlation between the errors in the annual cycle of mean daily precipitation and the errors in mean daily streamflow in each of the target drainage basins.
- Correlation between the errors in the annual cycle of the standard deviation of daily precipitation and the errors in mean daily streamflow in each of the target drainage basins.

Results and discussion

Precipitation

The raw DR_R2 results captured the observed annual cycle of mean daily precipitation over the 172 drainage basins and for each of the target drainage basins fairly well (left column in Fig. 2). However, as reported by Stefanova et al. (2011), mean daily precipitation was generally over-estimated (note that slight underestimation was found in July and August), with a mean error over the annual cycle of 0.5 mm (15 %) for all 172 drainage basins, approximately 0.6 mm (17 %) for the Alafia River basin and approximately 0.6 mm (17 %) for the Hillsborough River basin. By construct (i.e., because the prediction period coincides with the observation record used for bias correction in each month), the BCDR_R2 predicted annual cycle of mean daily precipitation is identical to the observed annual cycle of mean daily precipitation.

The raw DR_R2 results did not predict the annual cycle of the standard deviation of daily precipitation as well as the mean daily precipitation (right column in Fig. 2). The DR_R2 results showed a significant systematic over-prediction of the standard deviation of daily precipitation in February, April, June, October, and November and slight under-predictions in July and August for all 172 basins, and both of the target basins. The resulting mean error over the annual cycle was 2.5 mm (28 %) for all 172 drainage basins, 2.9 mm (35 %) for the Alafia drainage basin and 3.1 mm (36 %) for the Hillsborough River drainage basin. Again, the BCDR_R2 results for the annual cycle of the standard deviation of daily precipitation are identical to the observed results, due to the bias-correction methodology.

Although the bias-corrected results are guaranteed to reproduce the daily mean and standard deviation of precipitation on a monthly basis, they are not guaranteed to accurately reproduce actual timing and magnitude of observed precipitation events (Hwang et al. 2011). The error statistics for DR_R2 and BCDR_R2 actual monthly precipitation predictions for the two target drainage basins are compared in Fig. 3. The PBIAS for the DR_R2 monthly rainfall predictions was 17 % for the Alafia River basin and 15 % for the Hillsborough River basin. These errors were completely removed by bias correction. Bias correction reduced the RSR of monthly precipitation predictions by 21 % for the Alafia River basin and 22 % for the Hillsborough River basin and improved the NSE of the monthly precipitation predictions from approximately 0.28 to approximately 0.56. It is interesting to note, however, that the R^2 values for the monthly precipitation predictions were not significantly improved by bias correction. Thus, the CDF-mapping correction procedure completely removed the bias (i.e., PBIAS), reduces measures of mean

squared error (i.e., RSR and NSE), but does not improve the correlation of actual monthly predicted versus observed precipitation.

Temperature

Figure 4 compares the annual cycles of the mean and standard deviation of daily T_{max} and T_{min} averaged over the six stations for observations, DR_R2 and BCDR_R2 results. The DR_R2 predictions accurately reproduced both the mean daily T_{max} and T_{min} over the seasonal cycle (with mean errors less than 0.2 °C, or 1 %) and bias correction virtually removed these errors. The minor errors that remained in the mean daily T_{max} and T_{min} after bias correction were caused by the indirect bias correction of T_{mean} and T_{range} (instead of directly correcting T_{max} and T_{min}) as described in Sect. 5.2. Similarly, the temporal standard deviation of T_{min} was well predicted by both DR_R2 and BCDR_R2, with similar mean errors less than 0.1 °C (<1 %). In contrast, the temporal standard deviation of T_{max} was significantly overestimated by DR_R2, especially during the summer months. These errors were effectively removed by bias correction. Overall, the relative errors in DR_R2 temperature predictions (<1 %) were not as significant as for the precipitation results (>15 %) before bias correction.

Streamflow

In order to investigate hydrologic implication of the skills and deficiencies in the DR_R2 and BCDR_R2 CLAR-ReS10 predictions evaluated above, the calibrated INTB model was run using each of these climate inputs. Predictions of the time series of monthly streamflow and the annual cycle of mean monthly streamflow from each climate input were compared to the observed and calibrated model results for the two target stations. Note that comparing the simulations to the calibrated hydrologic simulations explores the hydrologic implications of errors/ differences in climatic forcing input. Comparing the simulations to observations explores the total errors due to climate input errors/differences and hydrologic model error.

Figure 5 compares the monthly hydrographs of streamflow predictions from 1989 to 2001. The figure shows that the DR-R2 climate inputs over-predicted monthly streamflow results for both the Alafia and Hillsborough Rivers especially during extreme events in 1995, 1997, 1998 and 1999. Bias correction of the precipitation and temperature inputs improved the streamflow predictions considerably, but not as much as it did for the climate inputs.

The error statistics for monthly streamflow predictions for the two target stations are compared in Fig. 3 (bottom row).
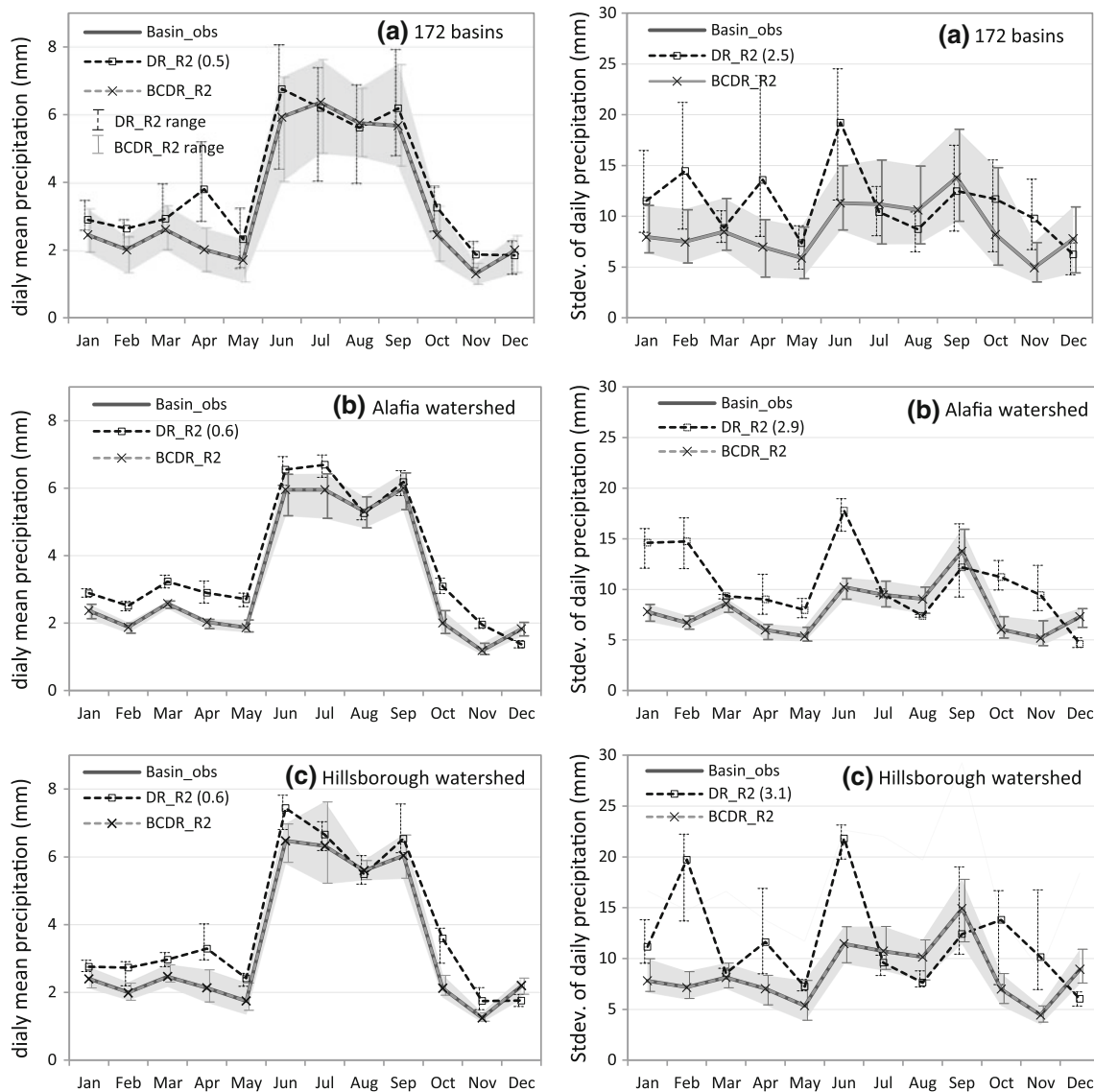
Fig. 2 Comparison of the annual cycles of monthly mean (*left column*) and standard deviation (*right column*) of daily precipitation from 1989 to 2001 for (**a**) all 172 basins, (**b**) Alafia River watershed and (**c**) Hillsborough River watershed. Mean error (unit: mm) over the cycle is represented in each legend. Note that the bias-corrected results match observed results exactly in both cases. The *gray* zones and *error bars* represent total range of observations and predictions, respectively, indicating spatial variation over contributing basins to streamflow discharge at each station

As expected, the calibrated model results fit the observed streamflow observations well over the entire study period and show relatively small error (PBIAS ($<\pm15$ %) and RSR (<0.41)) and high skill (R^2 (>0.88) and NSE (>0.83)) scores. These error statistics for calibrated results are categorized as 'good' (PBIAS) or 'very good' (RSR and NSE) hydrologic model performance according to Moriasi et al. (2007). As evident in Fig. 5, errors in the monthly DR_R2 precipitation predictions produce large errors in monthly streamflow predictions at both stations, resulting in significantly higher PBIAS (28 and 62 % for the Alafia River and Hillsborough

River, respectively), significantly higher RSR (1.36 and 1.30, respectively), significantly lower R^2 (0.29 and 0.42, respectively) and significantly lower NSE (-0.87 and -0.72, respectively). All these error statistics are categorized as 'unsatisfactory' hydrologic model performance by Moriasi et al. (2007). Use of bias-corrected climate data (i.e., BCDR_R2) improved the mean monthly streamflow results, reducing the average PBIAS by approximately 42 % and the average RSR by approximately 50 % over the DR_R2 results. Furthermore, the R^2 increased by approximately 76 % and NSE scores of BCDR_R2 results improved from

Error statistics for monthly precipitation
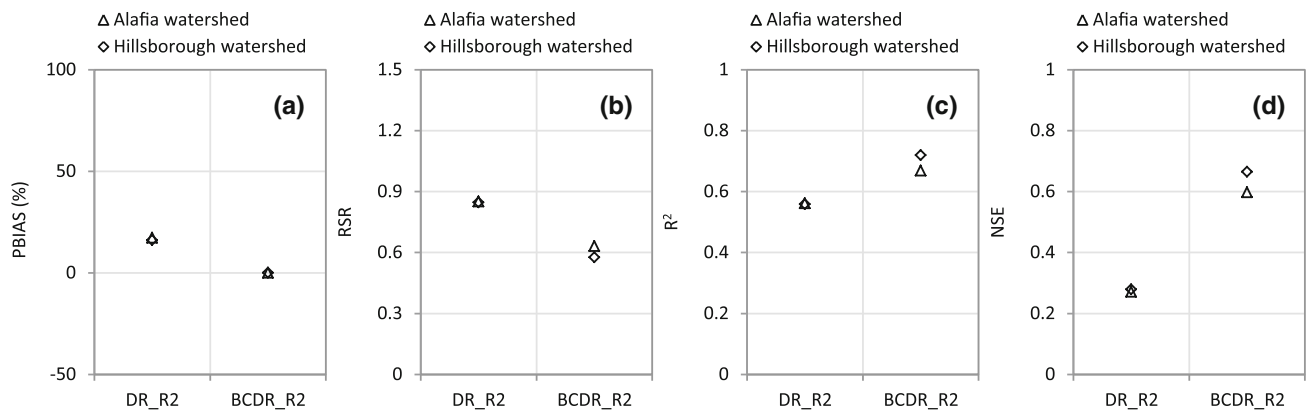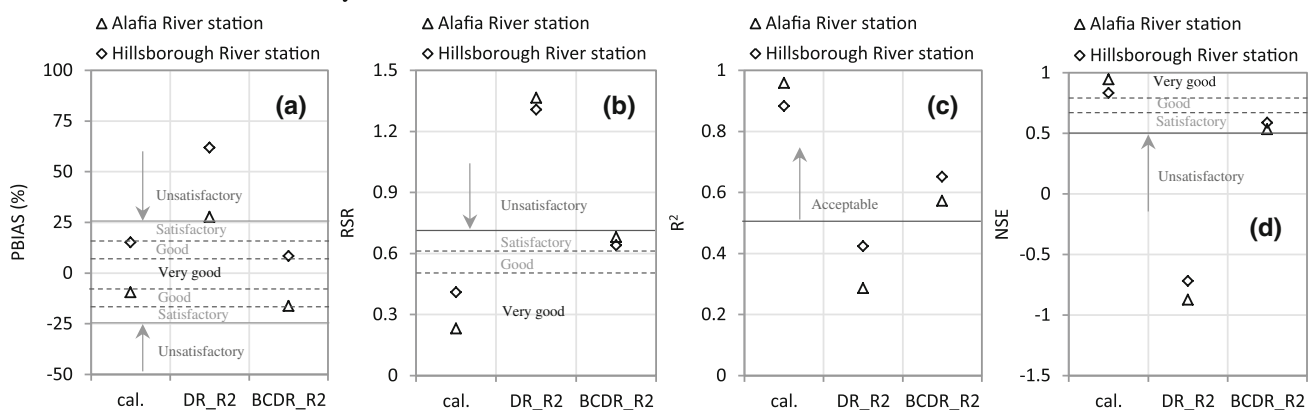

Error statistics for monthly streamflow

Fig. 3 Comparison of error statistics of monthly areal precipitation predictions (*upper row*) over the contributing area for each target streamflow station and monthly streamflow predictions (*bottom row*); (**a**) PBIAS, (**b**) RSR, (**c**) coefficient of determination (R^2), and (**d**) Nash–Sutcliffe efficiency (NSE). The range of performance ratings for hydrologic simulation recommended by Moriasi et al. (2007) is presented in the figures

an average of −0.80 to 0.56, indicating a significant improvement in the utility of the predictions. The streamflow error statistics produced using the bias-corrected climate data are categorized as 'satisfactory' by Moriasi et al. (2007). In general, Tampa Bay Water professionals consider that the streamflow predictions from their calibrated model, that are classified as good to very good, are suitable to guide monthly to annual water resource management decisions. The streamflow predictions from the bias-corrected CLARReS10 data (BCDR_R2), that are classified as satisfactory, are considered not sufficiently accurate for monthly decisions, but generally be suitable to guide long-term (seasonal to decadal) water resource planning. The streamflow predictions from the raw CLARReS10 data (DR_R2) are considered unsatisfactory, and therefore not recommended for use in hydrologic analyses in the Tampa Bay region.

In addition to evaluating the monthly precipitation and monthly streamflow prediction accuracies over the entire study period, the study period was disaggregated according

to ENSO phase, and the results were evaluated separately by ENSO phase. Durations of ENSO phase, as defined by the Oceanic Nino index (ONI) from the NOAA climate prediction center (http://ww.cpc.ncep.noaa.gov), are shown in Fig. 5. No significant differences in PBIAS, RSR, R^2 or NSE were found for either precipitation or streamflow predictions during *El-Niño*, neutral, or *La-Niña* phases, possibly because of the short duration of the study period. Time series plots of sea surface temperatures and corresponding ENSO classification, as well as plots of monthly precipitation and streamflow error statistics by ENSO phase, are presented in Online Resource 2.

The inset graphs in Fig. 5 compare the annual cycle of mean monthly streamflow predicted by the climate inputs to the observed and calibrated results for each target station. These results show that, even after averaging over the 13-year study period, the DR_R2 climate predictions produce mean monthly streamflow predictions that are significantly different from both the
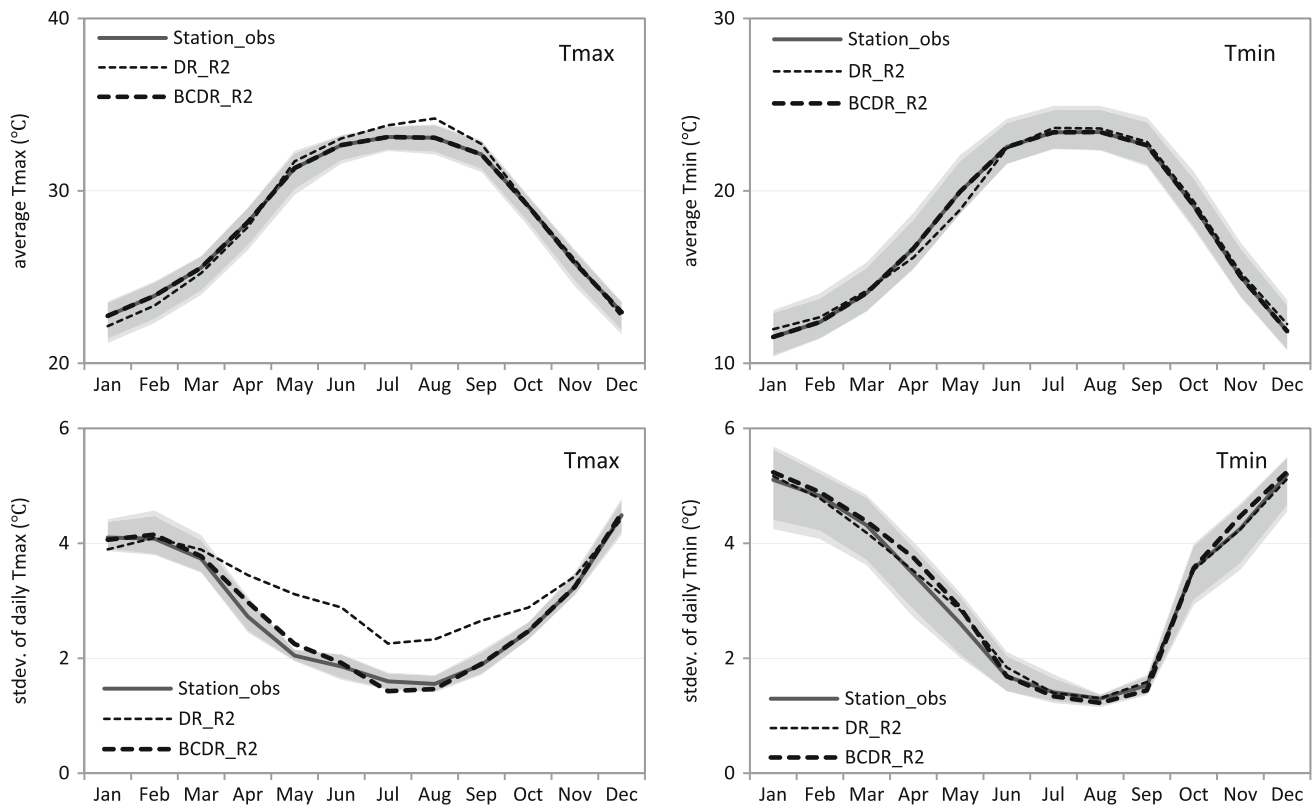
Fig. 4 Annual cycle of monthly mean (*top row*) and SD (*bottom row*) of daily T_{max} (*left column*) and T_{min} (*right column*). *Gray-shaded* area represents the observed range over the six stations

observed and calibrated results, even though the mean monthly precipitation results over the study period were quite similar to observations (see Fig. 2). The DR_R2 climate forcing data overestimated mean monthly flow throughout much of the year for both stations with mean error of 2.3 m³/s (31 %) for the Alafia River station and 4.2 m³/s (62 %) for the Hillsborough River station. Interestingly, the errors in prediction of mean streamflow at both stations are more strongly correlated with errors in the standard deviation of daily precipitation ($R^2 = 0.83$ for Alafia River and 0.71 for Hillsborough River) than errors in the mean of daily precipitation ($R^2 = 0.66$ for Alafia River and 0.61 for Hillsborough River) as can be seen by comparing the patterns in the right and left columns of Fig. 2 to the patterns in the insets of Fig. 4. This implies a significant nonlinear response of streamflow to precipitation, that is errors in the temporal variability of precipitation are more influential than errors in mean precipitation in causing errors in mean streamflow. Using bias-corrected climate inputs, however, produced satisfactory estimates of mean monthly streamflow over the entire annual cycle, with mean errors for monthly streamflow reduced to −1.2 m³/s (−16 %) for the Alafia River station and to 0.6 m³/s (8 %) for the Hillsborough River station.

Summary and conclusions

The goal of this study was to evaluate the ability of the CLARReS10 R2 downscaled reanalysis data to reproduce local-scale spatiotemporal precipitation and temperature data needed to predict streamflow in the Tampa Bay region of west central Florida. Results showed that raw CLAR-ReS10 R2 results (DR_R2) reproduced the mean annual cycle of daily minimum temperature, maximum temperature and precipitation fairly well, but systematically overestimated the mean and standard deviation of daily precipitation amounts for the entire study area. Biases in the temporal mean and standard deviation of daily precipitation and temperature predictions were effectively removed using a CDF-mapping approach with the observed spatially distributed daily precipitation and temperature observations used for calibration of the INTB model. Biases in the mean monthly precipitation totals were also removed by CDF mapping on a daily basis; however, errors in monthly precipitation totals and the standard deviation of the monthly precipitation totals remained after bias correction.

Precipitation errors produced by the raw DR_R2 data were propagated and enhanced by nonlinear processes in the hydrologic model, producing unacceptably high mean
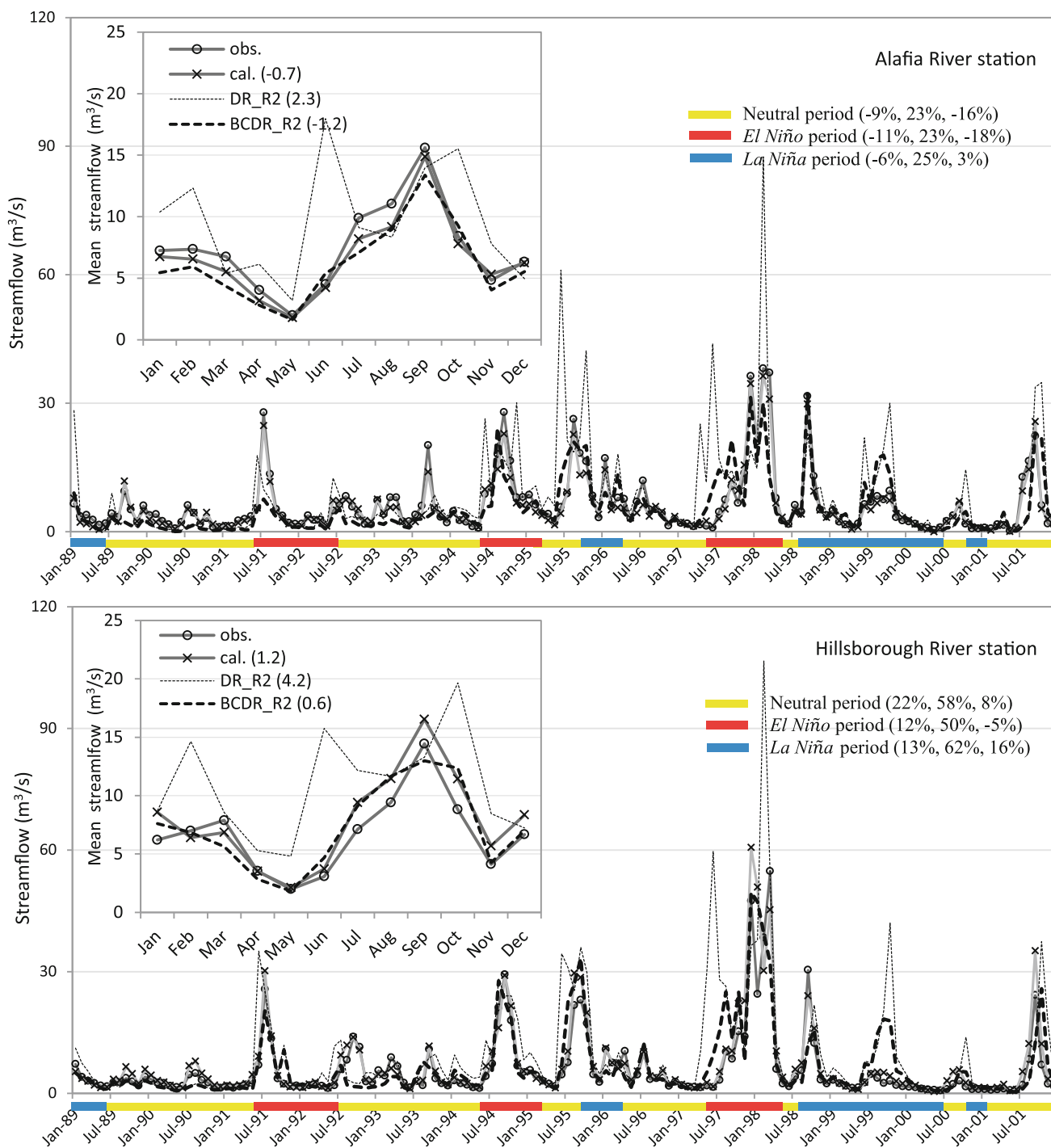
Fig. 5 Comparison of monthly streamflow time series of observed versus simulated (including calibrated results) for the Alafia River station (*upper*) and Hillsborough River station (*lower*). *Inset* graphs represent the mean monthly streamflow simulations over the study period from 1989 to 2001 for the target stations. Mean error (unit: m³/s) over the cycle is presented in each legend. *Color* bars near *x*-axis classify ENSO phase (i.e., *El-Niño* and *La-Niña* period) and PBIAS of calibrated, DR_R2, and BCDR_R2 monthly streamflow simulations for each ENSO period are presented

errors and low skill scores for monthly streamflow predictions over the 13-year study period. This indicates that the raw DR_R2 data cannot be considered as an acceptable

surrogate for historic precipitation and temperature measurements for use in hydrologic applications. Thus, further improvement in large-scale reanalysis data and RCMs is

needed before dynamically downscaled reanalysis data can be used directly to drive hydrologic models. The bias-corrected CLARReS10 R2 data (BCDR_R2) improved the mean error and skill scores significantly, bringing the hydrologic model performance into "satisfactory" range according to Moriasi et al. (2007). Thus, the accuracy of streamflow predictions produced using bias-corrected CLARReS10 dynamically downscaled climate predictions should be sufficient for use in seasonal to decadal water resources planning. Furthermore, the bias-corrected CLARReS10 R2 climate data could be used to extend the reanalysis simulation period for INTB from the 1989-2001 period, for which local historic observations are available, to the entire period of record available from the CLAR-ReS10 reanalysis data and local historic data, that is 1979–2006.

Finally, the accuracy of the bias-corrected dynamically downscaled climate and hydrologic predictions for the reanalysis period indicates that similarly bias-corrected dynamically downscaled retrospective and future GCM projections should be suitable for predicting potential hydrologic impacts of future climate change in the Tampa Bay region. However, until satisfactory improvements in GCMs, reanalysis data and RCMs occur, direct statistical downscaling of large-scale reanalysis data and GCMs using local historic data remains a viable, and computationally efficient, alternative to dynamic downscaling.

References

Beck A, Ahrens B, Stadlbacher K (2004) Impact of nesting strategies in dynamical downscaling of reanalysis data. Geophys Res Lett 31:L19101

Bicknell B, Imhoff JC, Kittle JL, Jobes TH, Donigian AD (2001) Hydrologic simulation program-FORTRAN (HSPF): user's manual for version 12. U.S Environmental Protection Agency, Athens, GA

Bromwich DH, Fogt RL (2004) Strong trend in the skill of the ERA-40 and NCEP-NCAR reanalyses in the high and midlatitudes of the southern hemisphere, 1958–2001. J Clim 17:4603–4619

El-Sadek A, Bleiweiss M, Shukla M, Guldan S, Fernald A (2011) Alternative climate data sources for distributed hydrological modeling on a daily time step. Hydrol Process 25:1542–1557

Fowler HJ, Blenkinsop S, Tebaldi C (2007) Linking climate change modeling to impacts studies: recent advances in downscaling techniques for hydrological modeling. Int J Climatol 27: 1547–1578

Geurink J, Basso R (2012) Development, calibration, and evaluation of the Integrated Northern Tampa Bay Hydrologic Model. Tampa Bay Water and the Southwest Florida Water Management District, Clearwater, FL

Geurink J, Basso R, Tara P, Trout K, Ross M (2006) Improvements to integrated hydrologic modeling in the Tampa Bay, Florida region: hydrologic similarity and calibration metrics. In: Proceedings of the joint federal interagency conference 2006, Reno, NV

Giorgi F, Hewitson B, Christensen J, Hulme M, Von Storch H, Whetton P, Jones R, Mearns L, Fu C (2001) Regional climate information -evaluation and projections. In: Climate change 2001: the scientific basis. Cambridge University Press, Cambridge, pp 583–638, 739–768

Haberlandt U, Kite GW (1998) Estimation of daily space-time precipitation series for macroscale hydrological modeling. Hydrol Process 12:1419–1432

Harbaugh AW, McDonald MG (1996) Programmer's documentation for MODFLOW-96, an update to the U.S. Geological Surbay modular finite-difference ground-water flow model, U.S. Geological Survey Open-File Report 96-486, Reston, VA

Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. Appl Eng Agric 1:96–99

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

Hwang S, Graham W, Hernández JL, Martinez C, Jones JW, Adams A (2011) Quantitative spatiotemporal evaluation of dynamically downscaled MM5 precipitation predictions over the Tampa Bay region, Florida. J Hydrometeorol 12(6):1447–1464

Iizumi T, Nishimori M, Dairaku K, Adachi AS, Yokozawa M (2011) Evaluation and intercomparison of downscaled daily precipitation indices over Japan in present-day climate: strengths and weaknesses of dynamical and bias correction-type statistical downscaling methods. J Geophys Res 116:D01111

Jasper K, Gurtz J, Lang H (2002) Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. J Hydrol 267:40–52

Jonathan TO, Meehl GA, Bony S, Easterling DR (2011) Climate data challenges in the 21st century. Science 331:700. doi:10.1126/science.1197869

Juang HMH, Kanamitsu M (1994) The NMC nested regional spectral model. Mon Weather Rev 122:3–26

Kanamitsu M, Kanamaru H (2007) Fifty-seven year California reanalysis downscaling at 10 km (CaRD10). Part I: system detail and validation with observations. J Clim 20:5553–5571

Kanamitsu M, Evisuzaki W, Woollen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL (2002) NCEP/DOE AMIP-II Reanalysis (R-2). Bull Am Meteorol Soc 83:1631–1643

Lim Y-K, Stefanova LB, Chan SC, Schubert SD, O'Brien JJ (2011) High-resolution subtropical summer precipitation derived from dynamical downscaling of the NCEP/DOE reanalysis: how much small-scale information is added by a regional model? Clim Dyn 37:1061–1080. doi:10.1007/s00382-010-0891-2

Lo JC, Yang Z, Pielke RA Sr (2008) Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model. J Geophys Res 113:D09112. doi:10.1029/2007JD009216

Mearns LO, Giorgi F, Whetton P, Pabon D, Hulme M, Lal M (2003) Guidelines for use of climate scenarios developed from regional climate model experiments. Data distribution centre of the Intergovernmental. Panel Clim Change. http://www.ipcc-data.org/guidelines/dgm_no1_v1_10-2003.pdf

Mesinger F, Janjić ZI, Nicković S, Gavrilov D, Deaven DG (1988) The step-mountain coordinate: model description and performance for cases of Alpine lee cyclogenesis and for a case of an Appalachian redevelopment. Mon Weather Rev 116:1493–1518

Mesinger F, DiMego G, Kalnay E, Shafran P, Ebisuzaki W, Jovic D, Woollen J, Mitchell K, Rogers E, Ek M, Fan Y, Grumbine R, Higgins W, Li H, Lin Y, Manikin G, Parrish D, Shi D (2006) North American regional reanalysis. Bull Am Meteorol Soc 87:343–360

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulation. Trans ASABE 50(3):885–900

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I: a discussion of principles. J Hydrol 10:282–290. doi:10.1016/0022-1694(70)90255-6

Panofsky HA, Brier GW (1968) Some applications of statistics to meteorology. The Pennsylvania State University, University park, p 224

Piani C, Weedon GP, Best M, Gomes SM, Viterbo P, Hagemann S, Haerter JO (2010) Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J Hydrol 395:199–215

Ross M, Trout K, Tara P, Said A, Geurink J (2004) A new discretization scheme for integrated surface and groundwater modeling. Hydrological science and technology. J Am Inst Hydrol (AIH) 21:143–156

Ross M, Geurink J, Said A, Aly A, Tara P (2005) Evapotranspiration conceptualization in the HSPF-MODFLOW integrated models. J Am Water Res Assoc 41(5):1012–1025

Stefanova L, Misra V, Chan S, Griffin M, O'Brien JJ, Smith JT III (2011) A proxy for high-resolution regional reanalysis for the southeast United States: assessment of precipitation variability in dynamically downscaled reanalyses. Clim Dyn 38(11):2449–2466. doi:10.1007/s00382-011-1230-y

Themeßl JM, Gobiet A, Heinrich G (2012) Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. Clim Change 112:449–468

Uppala SM et al (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131:2961–3012

Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr 21:530–548

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Clim Change 62:189–216