Journal of Hydrology 510 (2014) 513–529

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Hydrologic implications of errors in bias-corrected regional reanalysis data for west central Florida

Syewoon Hwang [a,b,*], Wendy D. Graham [a], Jeffrey S. Geurink [c], Alison Adams [c]

[a] Water Institute, University of Florida, Gainesville, FL, United States
[b] Department of Agricultural Engineering, Gyeongsang National University (Institute of Agriculture and Life Science), Jinju 660-701, South Korea
[c] Tampa Bay Water, Clearwater, FL, United States

ARTICLE INFO

Article history:
Received 11 May 2012
Received in revised form 2 May 2013
Accepted 23 November 2013
Available online 12 December 2013
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of David J. Gochis, Associate Editor

Keywords:
Hydrologic implications of climate predictions
Regional reanalysis data
Integrated hydrologic model
Bias-correction

SUMMARY

This study investigated the limitations associated with using dynamically-downscaled, bias-corrected reanalysis data (i.e. regional reanalysis data) to predict hydrologic behavior of low-relief rainfall driven systems using an integrated surface/groundwater model. Four different sets of global reanalysis data (NCEP/NCAR-R1, NCEP-DOE-R2, ERA40, and 20CR) that were previously downscaled using two RCMs (MM5 and RSM) were obtained, bias-corrected on a daily basis using the CDF-mapping approach, and used to drive an integrated hydrologic model (INTB) that was previously calibrated and verified for the Tampa Bay region.

All raw dynamically-downscaled reanalysis datasets accurately estimated the annual cycle of daily maximum and minimum temperature, except the NCEP/NCAR R1+MM5 data which consistently underestimated daily maximum temperature. All raw regional reanalysis precipitation data significantly overestimated precipitation, particularly for the dry season. Bias-correction using the CDF-mapping approach effectively removed biases in the temporal mean and standard deviation of both the daily precipitation and temperature predictions. Biases in the mean monthly and mean annual precipitation totals were removed by CDF-mapping on a daily basis, but the standard deviation of the monthly and annual precipitation totals were not accurately reproduced. Furthermore inaccuracies in actual daily precipitation time series aggregated into monthly and annual rainfall total time series that showed significant and temporally persistent errors.

Precipitation timing errors produced by regional reanalysis data were propagated and enhanced by non-linear streamflow generation, groundwater flow and storage processes in the hydrologic model and produced significant errors in both actual and mean daily, monthly and annual streamflow and groundwater level predictions. These results show that improvement in large-scale reanalysis products and regional climate models may be required before dynamically downscaled bias-corrected reanalysis data can be used as a surrogate for observational data in hydrologic model applications for low-relief, rainfall driven systems.

1. Introduction

A common method of developing climate forecasts and future climate projections for quantitative climate variability and climate change impact assessments is to use results from general circulation model (GCM) experiments. Global climate modeling continues to be improved by incorporating more aspects of the complexities of the global system. GCMs are considered a robust tool for understanding present and past continental scale climates (Karl and Trenberth, 2003). However, the coarse resolution of existing GCMs (typically 200 km by 200 km) precludes the simulation of realistic circulation patterns and accurate representation of the small-scale spatial variability of climate variables (Christensen and Christensen, 2003; Giorgi et al., 2001; Jones et al., 2004; Lettenmaier et al., 1999; Wood et al., 2002) needed for impact assessments.

To overcome this limitation, a number of downscaling methods have been developed. The two main downscaling approaches are statistical downscaling techniques and dynamic downscaling methods. Whereas the former method uses empirical relations between features simulated by GCMs at grid scales and surface observations at subgrid scales (Hay et al., 2002; Wilby and Wigley, 1997), the latter uses regional climate models (RCMs) based on physical relationships between the climate at larger and smaller scales (Maraun et al., 2010; McGregor, 1997). Although

* Corresponding author. Address: Department of Agricultural Engineering, Gyeongsang National University (Institute of Agriculture and Life Science), Jinju 660-701, South Korea. Fax: +82 055 772 1939.
E-mail address: swhwang78@gmail.com (S. Hwang).

0022-1694/$ - see front matter © 2014 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jhydrol.2013.11.042

computationally more expensive, dynamic downscaling techniques using RCMs are not limited by the statistical assumptions that the statistical downscaling approaches must make and thus have more potential to reproduce current and future local climatology. However, physically-based dynamic downscaling methods must be evaluated for diverse regions with a variety of climate variability characteristics so that the strengths and weaknesses of dynamic downscaling approaches and their hydrologic implications can be better understood (Fowler et al., 2007; Hong, 2003; Wang et al., 2004).

Generally, before using RCM results for climate forecasts or projections, it must be shown that the model simulates local present-day climate conditions accurately (Christensen et al., 1997; Hay and Clark, 2003). This is typically achieved by running the RCM using GCM scale reanalysis data as initial and boundary conditions for given historical periods and comparing the resulting predictions to observations (Hwang et al., 2011; Mearns et al., 2003). Use of the reanalysis data removes the confounding factors of potential biases related to GCM process simulation, and thus provides a more objective measure of the skill of the RCM downscaling accuracy (Maurer and Hidalgo, 2008; Maurer et al., 2010). Furthermore, as the performance of dynamic downscaling models has improved, climate scientists have developed dynamically downscaled reanalysis data products at high spatial resolution (hereinafter referred to as 'regional reanalysis') that have been proposed for use as a proxy for historic meteorological observation data (Mesinger et al., 2006; Kanamaru and Kanamitsu, 2007; Sato et al., 2007; Stefanova et al., 2012).

Despite the common acceptance of reanalysis data as the best estimate of many atmospheric variables, they must be used with caution because of their biases. While GCM-scale reanalysis data honor measurements by assimilating historical observations, low density, degradation or replacement of measurement instruments, and changes in methods of observation, may inject error into reanalysis data (Berg et al., 2003; Trenberth et al., 2001; Francis, 2002; Widmann and Bretherton, 2000). Additionally it should be noted that not all reanalysis outputs are constrained by observation; precipitation and surface evapotranspiration are obtained by running general circulation or numerical weather prediction models that may include systematic errors (Nigam and Ruiz-Barradas, 2006).

As a result, many researchers have found that RCM results that are produced using reanalysis data as boundary conditions also produce considerable biases in precipitation and temperature that may significantly affect hydrologic predictions (e.g., Christensen et al., 2008; Fujihara et al., 2008; Hay et al., 2002; Maurer and Hidalgo, 2008; Wilby et al., 2000). For this reason, bias-correction of RCM predictions using historic observations is typically necessary (Christensen et al., 2008; Murphy, 1999; Teutschbein and Seibert, 2010; Wood et al., 2004). A common method for bias-correction is to map a cumulative distribution function (CDF) of climate model results onto that of observations. This algorithm was introduced for statistical application for meteorological analysis by Panofsky and Brier (1968) and was adapted for adjusting general circulation model (GCM) results at the monthly time scale by Wood et al. (2002, 2004). Recently a similar bias-correction approach at the daily time scale has been applied to climate modeling results adjustment (Dosio and Paruolo, 2011; Hwang et al., 2011, 2013; Piani et al., 2010a, b), climate change impact assessments on hydrology (Chen et al., 2011), and wildfire applications (Abatzoglou and Brown, 2011). Previous studies on comparative evaluations of various bias-correction methods for precipitation and temperature have shown that the CDF-based distribution mapping approach performs better over other bias-correction methods (e.g., linear scaling, variance scaling, power transformation, etc.) because this approach matches the full distribution of model outputs to observed distribution and thus high moment statistical properties of observation are accurately reproduced (e.g., Gudmundsson et al., 2012; Johnson and Sharma, 2011; Teutschbein and Seibert, 2012a). Teutschbein and Seibert (2012a) provide a detailed summary and comparison of alternative bias-correction methods.

Bias-correction using CDF-mapping at the daily time scale has been shown to remove bias in the daily precipitation and temperature predictions as well as the tendency to under-predict dry days and over-predict the number of low volume rainfall days (Hwang et al., 2011; Teutschbein and Seibert, 2012a). However, while statistical bias-correction improves the climate modeling results by matching distributional characteristics to those of observations for the particular time-scale at which it is conducted, precipitation timing errors are not removed through the statistical correction process. For example, Hwang et al. (2011) examined dynamically downscaled and bias-corrected reanalysis data and showed that while the precipitation fields accurately predicted mean seasonal and inter-annual climatology and daily rainfall transition probabilities, significant error remained in predicting specific daily, monthly and annual total time series. Similarly Haerter et al. (2011) found that the improvements to the statistical properties of the bias-corrected results were limited to the specific timescale of the fluctuation that are considered in the correction process.

A number of recent studies used bias-corrected RCM results including regional reanalysis data to conduct hydrologic assessments of retrospective and future conditions over various regions (e.g., Chen et al., 2011; Fujihara et al., 2008; Graham et al., 2007; Middelkoop et al., 2001; Minville et al., 2009; Teutschbein and Seibert, 2012a). However, the hydrologic implications of the timing errors and cross-time scale biases in climate forcing data discussed above, which may be especially important when using regional reanalysis data as a proxy for observations, have not been clearly assessed. In addition most of previous studies have focused on evaluating streamflow regimes because of their sensitivity to changes in climate forcing. However in regions such as Florida,

Table 1
Description of regional reanalysis data that used in the study.

Reanalysis data		RCM	Resolution	Data period	References	Working Group
NCEP-NCAR reanalysis (R1)	Kalnay et al., 1996	Meso-scale climate model (MM5)	9×9 km^2	1986–2008	Hwang et al., 2011	Water Institute, University of Florida
NCEP-DOE 2 reanalysis (R2)	Kanamitsu et al., 2002			1979–2001	Stefanova et al., 2012	COAPS (Center for Ocean-Atmospheric Prediction Studies), Florida State University
ECMWF's reanalysis (ERA40)	Uppala et al., 2005	Regional Spectral Model (RSM)	10×10 km^2			
20 century reanalysis (20CR)	Compo et al., 2011			1901–2008	DiNapoli and Misra, 2012	

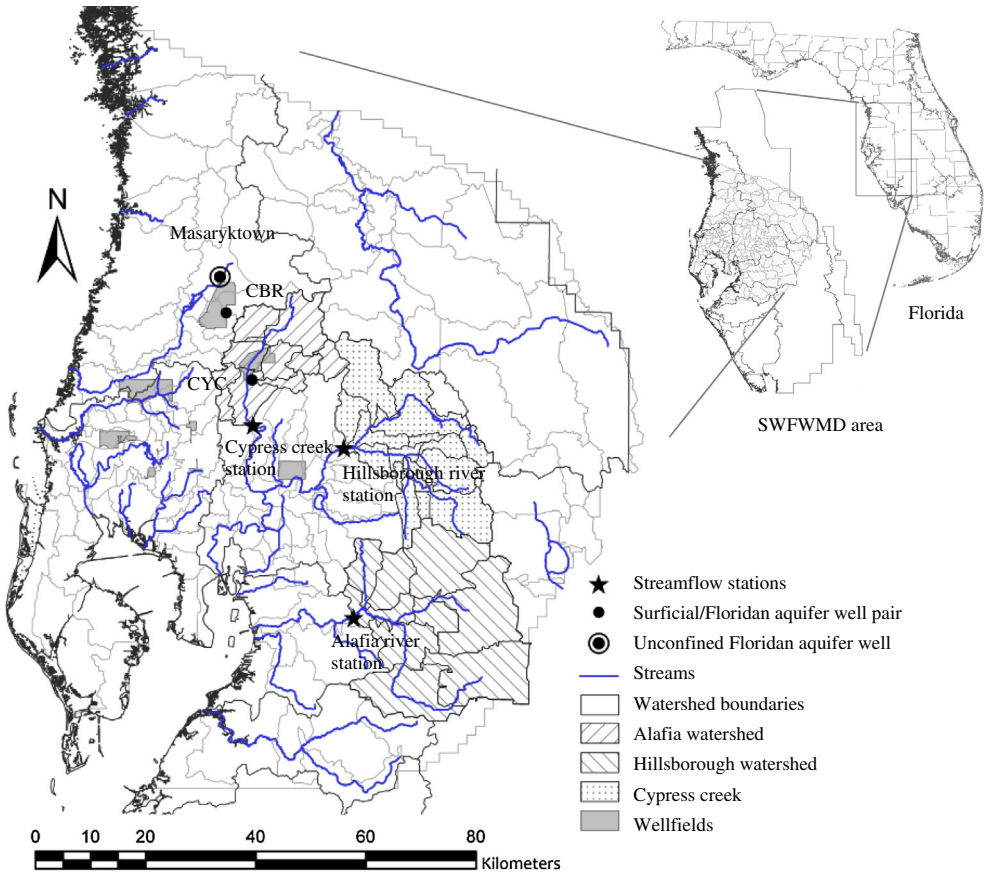

Fig. 1. Map of hydrologic modeling domain, sub-basins, and observation stations for streamflow and groundwater level. Marked areas indicate the contributing basins for each streamflow target station.

Table 2
Description of data used to calibrate hydrologic model.

Data	Source	Description
Streamflow discharge	USGS	Data for 38 streamflow stations were collected
Groundwater level	USGS SWFWMD TBW	Data were obtained from 236 surficial and 192 Floridan aquifer monitoring wells over the model domain
Diversion	SWFWMD TBW USGS	Surface-water diversions represent pumped surface-water sources, non-pumped surface-water transfers, and augmented water bodies. Daily pumping data for facilities pumping more than 1 mgd were acquired. Monthly pumping data were acquired for all others
Irrigation	SWFWMD	Agricultural irrigation flux rates were estimated from pumping records maintained by the SWFWMD and irrigated area from FLUCCS parcels classed as irrigated. Irrigation from potable or reclaimed water sources was not estimated and is not included in the model
Spring discharge	USGS, SWFWMD	Daily data were obtained from 6 springs. Periodic discharge records were obtained for one springs and linearly interpolated to produce daily estimates
Well pumping	TBW SWFWMD	Daily rates obtained for all of TBW production wells Monthly pumping rates for all other wells were obtained from SWFWMD.
Land-use data	SWFWMD	Coverage delineate areas of particular land use as classified by FLUCCS. Original 53 FLUCCS codes were reduced to 7 hydrologically unique classifications.

USGS: U.S. Geological survey.
SWFWMD: South West Florida Water Management District.
TBW: Tampa Bay Water.
FLUCCS: The Florida Land Use and Cover forms Classification System.

where groundwater is a major source of public water supply, simultaneously evaluating streamflow and groundwater predictions driven by climate modeling results is important for water resource management.

The goal of this study was to investigate the limitations associated with using dynamically-downscaled, bias-corrected reanalysis data to predict hydrologic behavior of low-relief rainfall driven systems. We collected four different regional reanalysis datasets available over west-central Florida and bias-corrected each of them at the daily timescale using the CDF-mapping approach with spatially distributed daily observation data. The raw and bias-corrected regional reanalysis were comprehensively evaluated for their skill in reproducing the spatiotemporal mean and variance of observed basin-scale precipitation and temperature at different timescales (e.g., daily, monthly, and annual). To determine the influence of errors in climate forcing data on

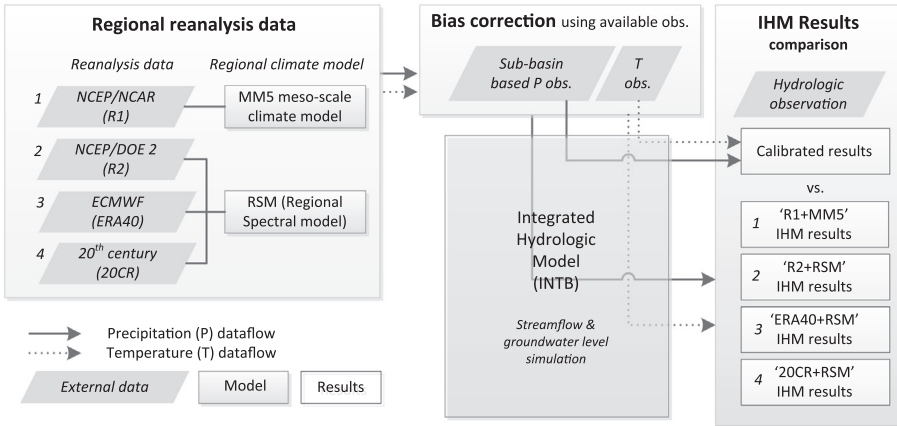

Fig. 2. Schematic representation of the study framework.

Table 3
Target stations for streamflow, surficial aquifer and Floridan aquifer.

	Name (data source)	Watershed	Lat.	Lon.	Drainage area, (km^2)
Streamflow stations	Alafia River at Lithia	Alafia	27.8719	−82.2114	867.3
	Hillsborough River near Zephyrhills	Hillsborough	28.1497	−82.2325	569.6
	Cypress Creek at Worthington Gardens	Hillsborough	28.1856	−82.4008	302.9
Surficial aquifer monitoring wells	CYC TMR-5 SH (SWFWMD)	Hillsborough	28.2057	−82.4680	–
	CBR-SERW-s (TBW)	Springs coast	28.3151	−82.5146	
Floridan aquifer monitoring wells	CYC TMR-5d (TBW)	Hillsborough	28.2053	−82.4680	–
	CBR-SERW-d (TBW)	Springs coast	28.3151	−82.5146	
	Masaryktown DP (SWFWMD)[a]	Springs coast	28.3740	−82.5262	

[a] Unconfined Floridan aquifer monitoring well. Corresponding surficial aquifer, therefore, does not exist.

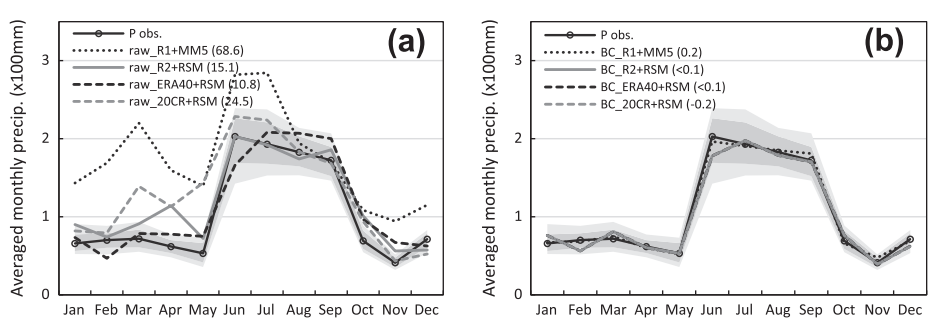

Fig. 3. Comparison of the mean monthly precipitation over the study period for (a) raw regional reanalysis data and (b) bias-corrected results to basin based observations (P obs.). Mean errors over the cycle in units of mm are represented on each legend. The bright and dark gray zones represent total data range and 5th to 95th percentile of P obs., respectively, reflecting spatial variation of the mean monthly precipitation over the 172 basins.

hydrologic behavior, the bias-corrected regional reanalysis data were used to drive the Integrated Northern Tampa Bay model (INTB) that was previously calibrated and verified for the Tampa Bay region. INTB results obtained using the four bias-corrected regional reanalysis datasets were compared to the calibrated model predictions and the observed hydrologic behavior.

2. Regional reanalysis data used in the study

Four different regional reanalysis datasets: (1) the National Center for Environmental Prediction (NCEP)- National Center for Atmospheric Research (NCAR) R1 reanalysis (Kalnay et al., 1996) downscaled using the meso-scale regional climate model MM5 (Hwang et al., 2011, hereinafter referred to as 'R1+MM5'), (2) the NCEP-Department of Energy (DOE) R2 reanalysis (Kanamaru and Kanamitsu, 2007) downscaled using the regional spectral model RSM (Stefanova et al., 2012, hereinafter 'R2+RSM'), (3) the European Centre for Medium-range Weather Forecasts – 40 year reanalysis (ECMWF, Uppala et al., 2005) downscaled using RSM (Stefanova et al., 2012, hereinafter 'ERA40+RSM'), and (4) The twentieth century reanalysis data (20CR, Compo et al., 2011) downscaled using RSM (DiNapoli and Misra, 2012, hereinafter '20CR+RSM'). Information regarding each data set is summarized in Table 1.

Hwang et al. (2011) dynamically downscaled NCEP/NCAR R1 reanalysis data over Florida using MM5 at a $9 \times 9\,km^2$ grid for the 23 year period from 1986 to 2008. They set up MM5 using the NCAR community model CCM2 radiation scheme (Kiehl et al., 1996), the Grell cumulus parameterization scheme (Grell et al., 1994), and the Simple ice scheme (Grell et al., 1994). Planetary boundary layer (PBL) physics was set to a non-local vertical

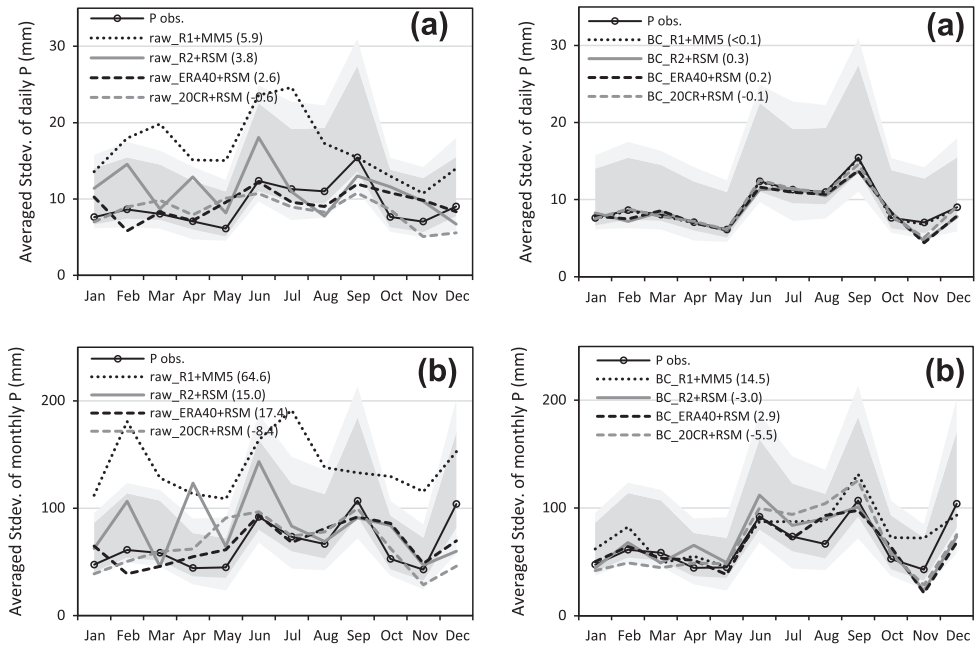

Fig. 4. The annual cycle of averaged standard deviation of (a) daily and (b) monthly precipitation (first column: raw regional reanalysis, second column: bias-corrected results) over the study domain. Mean errors over the cycle in units of mm are included in each legend. bright and dark gray zones represent total data range and 5th to 95th percentile of P obs. for 172 sub-basins, respectively.

diffusion scheme (Hong and Pan, 1996) and the 5-layer soil temperature model was used for land surface processes (Dudhia, 1996). They bias corrected the downscaled reanalysis data using data from 53 long term precipitation gages irregularly distributed over the Tampa Bay region. The bias-corrected dynamically downscaled reanalysis data showed reasonable skill in reproducing the observed seasonal cycle and spatiotemporal variability of precipitation.

Stefanova et al. (2012) dynamically downscaled the NCEP-DOE R2 and ERA40 reanalysis datasets over the Southeast United States using RSM at a 10×10 km^2 grid for the period from 1979 to 2001. For the RSM configuration, the Simplified Arakawa-Schubert Scheme (Pan and Wu, 1994) was used to parameterize deep convection and the Noah land surface scheme (Ek et al., 2003) was used for the land surface processes. They showed that R2+RSM and ERA40+RSM have reasonable skills in reproducing the seasonal cycle of precipitation over the Florida, but have systematic biases in precipitation volume.

Most recently, DiNapoli and Misra (2012) dynamically downscaled the 20CR reanalysis data from 1901 to 2008 using RSM with same grid, physics and parameterization schemes as Stefanova et al. (2012). They showed that raw 20CR+RSM also adequately reproduced the spatial pattern and long-term seasonal and diurnal variability of temperature and precipitation over Florida, but again included systematic biases in precipitation predictions.

3. Hydrologic model and observation data

3.1. Hydrologic modeling

In west central Florida the fresh groundwater flow system generally consists of a thin surficial aquifer underlain by the thick, highly productive carbonate rocks of the Floridan aquifer system. Most of the Floridan aquifer in the region is semi-confined, recharged by means of leakage from the overlying surficial aquifer. However, in the northern extent of the region some portions of the Floridan aquifer are unconfined, receiving direct recharge from vadose zone infiltration. The significant temporally variable flux and storage connection between surface and groundwater systems is caused by the near-surface water table condition that covers more than 50% of the region. In order to capture the dynamic interaction between surface and groundwater in this region an integrated hydrologic model is required.

Tampa Bay Water (TBW) and the local state regulatory agency for surface water and groundwater resources, the Southwest Florida Water Management District (SWFWMD), commissioned the development and application of an integrated surface water/groundwater model to gain an increased understanding of the surface and groundwater flow systems in the Tampa Bay region (Geurink et al., 2006a). From this effort the Integrated Hydrologic Model (IHM) was developed which integrates the EPA Hydrologic Simulation Program-Fortran (HSPF; Bicknell et al., 2001) for surface-water modeling with the US Geological Survey MODFLOW96 (Harbaugh and McDonald, 1996) for groundwater modeling.

IHM was designed to provide an advanced predictive capability of the complex interactions of surface water and groundwater features in shallow water-table environments. The model can be characterized as deterministic, semi-distributed-parameter, semi-implicit real-time formulation, with variable time steps and spatial discretization (Ross et al., 2004). The model components explicitly account for all significant hydrologic processes including precipitation, interception, evapotranspiration, runoff, recharge, streamflow, baseflow, groundwater flow, and all the component storages of surface, vadose and saturated zones (Ross et al., 2005). Climate input data requirements include time series for precipitation and maximum and minimum temperature to calculate reference evapotranspiration for each basin in the model domain (Geurink et al., 2006b).

3.2. INTB model domain and calibration

The Integrated Northern Tampa Bay (INTB) model was developed and calibrated using the IHM simulation engine. The INTB model domain is bordered by the Gulf of Mexico and inland groundwater flow lines. The area is located in the west central Flor-

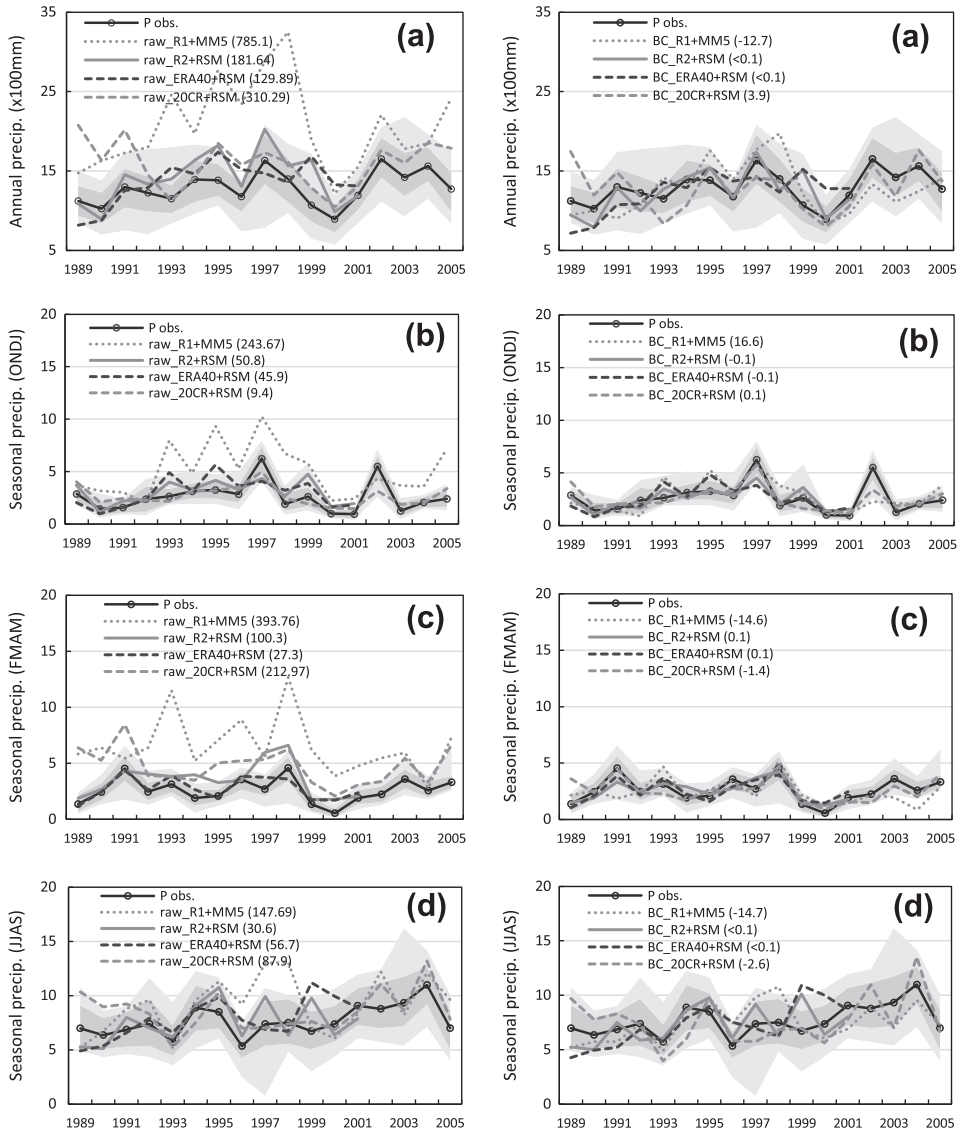

Fig. 5. Comparison of (a) annual total precipitation and seasonal total precipitation time series ((b) is for October to January, (c) is for February to May, and (d) is for June to September (wet season)) for IHM sub-basin observations (P obs.), raw regional reanalysis data (left column), and bias-corrected results (right column) using P obs. Mean errors over the study period in units of mm are represented on each legend. The bright and dark gray zones represent total data range and 5th to 95th percentile of P obs. indicating the spatial variability of observed annual and seasonal precipitation over the 172 sub-basins.

ida region and extends to the eastern boundary of SWFWMD as shown in Fig. 1. Tampa Bay is located in the southwest part of the domain. The north and east boundaries follow Floridan aquifer flow lines (i.e. no flux boundaries) 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., 2006a). In the study area, average annual rainfall for the model calibration (1989–1998) and verification periods (1999–2006) was 1295 mm and 1263 mm, respectively. In general, evapotranspiration (ET) accounts for approximately 70% of annual precipitation. Land cover over the domain is diverse, including urban, grassland, forest, agricultural, mined land, water, and wetlands. Open water and wetlands cover 25% of the region.

The surface-water of the component model domain is discretized into 172 basins based on surface drainage as shown in Fig. 1. For each basin, hydrologic processes are simulated within hydrologic response units (land segments) based on five upland landuse categories and two water-body categories (Ross et al., 2004). The groundwater component of the model domain is discretized into approximately 35,000 square and rectangular grid cells with cell dimension of one-quarter mile over the area of interest and expanding to one-mile in outer regions.

The INTB model was manually calibrated using hydrologic observations from 1989 to 1998 and verified using data from 1999 to 2005 (Geurink and Basso, 2011). Manual calibration efforts were significantly informed and enhanced by many applications of PEST (the automated Parameter Estimation software: Doherty, 2004) to the INTB model data. Hydrologic observations used as calibration targets included 38 streamflow monitoring stations, 200 locations each of surficial and Floridan aquifer wells, 7 springs, and long-term average annual target values of actual evapotranspiration assigned to landuse-depth to water table combinations. Table 2 details the data sources used by Geurink and Basso (2011) to develop and calibrate the INTB.

3.3. Meteorological data

Spatiotemporal distribution and intensity of precipitation are important in deterministic, physically-based hydrologic simulations. A short temporal resolution is required to adequately

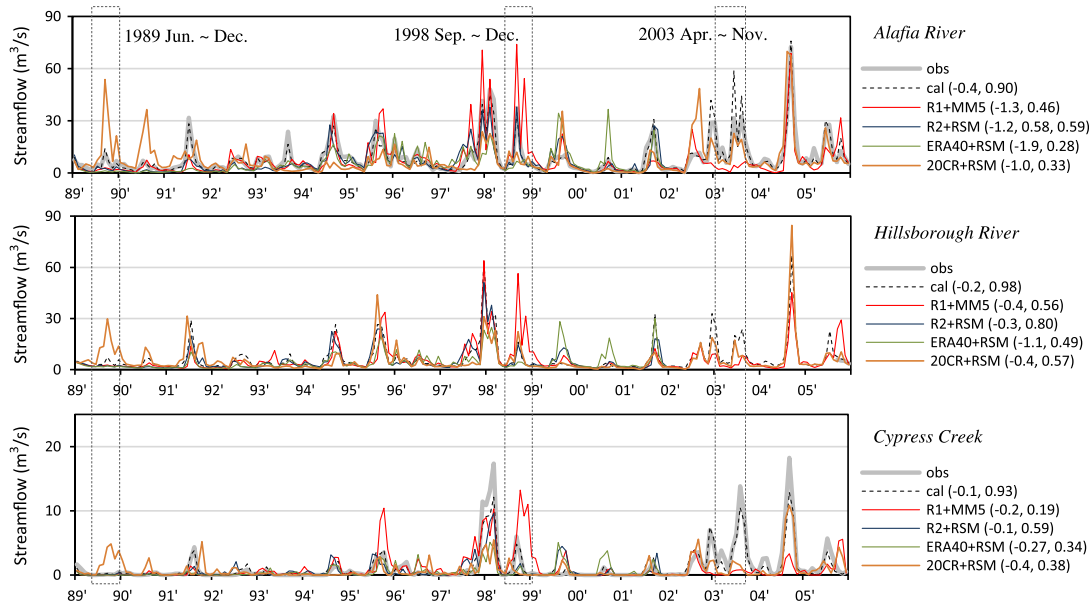

Fig. 6. Comparison of the monthly streamflow hydrographs of observations, calibrated streamflow predictions, and the streamflow predictions using bias-corrected regional reanalysis precipitation data. The vertical boxes indicate three periods for which streamflow significantly overestimated by 20CR+RSM data and underestimated by others (1989, first box), overestimated by R1+MM5 data (1998, second box), and underestimated by R1+MM5 and 20CR+RSM (2003, third box) for all stations. The values in the bracket indicate the mean error of the monthly streamflow (m^3/s) and R^2.

capture the effects of localized convective storms. Rokicki (2002) showed that hydrologic modeling with time steps larger than 15 min results in significant errors in runoff predictions for west-central Florida, and suggested discretizing precipitation into 15 min intervals or smaller for accurate simulation of runoff. The INTB model uses 15-min precipitation time series input for each of the 172 basins. For the calibrated INTB model precipitation data over the INTB model domain were obtained from 300 stations maintained by three different agencies including TBW, SWFWMD, and National Oceanic and Atmospheric Administration (NOAA). In order to estimate basin precipitation time series, available daily precipitation data within each basin were spatially distributed by Thiessen polygons and averaged over the basin (area weighted) to generate input for the hydrologic modeling. Daily precipitation values for each basin were temporally disaggregated using the pattern of the nearest NOAA station with 15 min observations which matched the daily rainfall volume.

For the calibrated INTB model six reference ET (Hargreaves and Samani, 1985) time series were developed from minimum and maximum temperature data at six NOAA stations and then spatially assigned to the nearest neighbor basins over the model domain to define the upper limit of evaporative demand. Daily reference ET values were temporally disaggregated into hourly values for the INTB model input using an annual profile of hourly reference ET values. Hourly values were generated with the FAO-56 PM method (Allen et al., 1998) using a data set with a shorter period of record but having a full suite of weather parameters.

In general, reanalysis data include a wide variety of climate variables at large scales and RCMs also produce most of them at the local scale. Various climate variables (e.g., solar radiation, wind, humidity, etc.) are available from regional reanalysis. However, for this research, only surface precipitation and maximum and minimum temperature were analyzed and used because the INTB model requires only these climate variables as input data. We bias-corrected the four regional reanalysis precipitation and temperature datasets using the basin-based daily precipitation data for the 172 model basins and temperatures data for the 6 stations used to calibrate the INTB model. The study periods for each dataset were chosen based on the overlap of the availability of regional reanalysis and the period for hydrologic modeling calibration and validation (i.e., 1989–2005 for R1+MM5 and 20CR+RSM, and 1989–2001 for R2+RSM and ERA40+RSM).

4. Methodology

A schematic representation of the framework for this study is shown in Fig. 2. Details of methods used to bias correct the climate predictions and evaluate the hydrologic predictions are summarized below.

4.1. Climate prediction adjustment

A cumulative distribution function (CDF) mapping approach (Panofsky and Brier, 1968; Wood et al., 2002; Ines and Hansen, 2006; Maurer and Hidalgo, 2008; Li et al., 2010; Dosio and Paruolo, 2011; Haerter et al., 2011) was used to bias-correct the raw downscaled daily reanalysis data using the following procedure: (1) CDFs of observed daily precipitation were created individually for each of the 172 model basins for each month using observed data. Thus, for each case, 12 different monthly CDFs were used for each basin for bias-correction of the daily predictions; (2) CDFs of simulated daily precipitation were created for the grid cell containing the centroid of each basin for each month; (3) daily grid cell predictions were bias-corrected at each basin centroid using CDF mapping that preserves the probability of exceedence of the simulated precipitation over the grid cell containing the basin centroid, but corrects the precipitation to the value that corresponds to the same probability of exceedence from the observed results at the centroid location. Thus bias-corrected rainfall $x'_{i,j}$ on day i for basin j was calculated as,

$$X'_{i,j} = F^{-1}_{obs,j}(F_{sim,j(x_{i,j})}) \quad (1)$$

where $F(x)$ and $F^{-1}(x)$ denote a CDF of daily precipitation x and its inverse, and subscripts *sim* and *obs* indicate downscaled simulation and observed daily rainfall, respectively. The use of the 172 basin observations for bias-correction allows bias-corrected results to

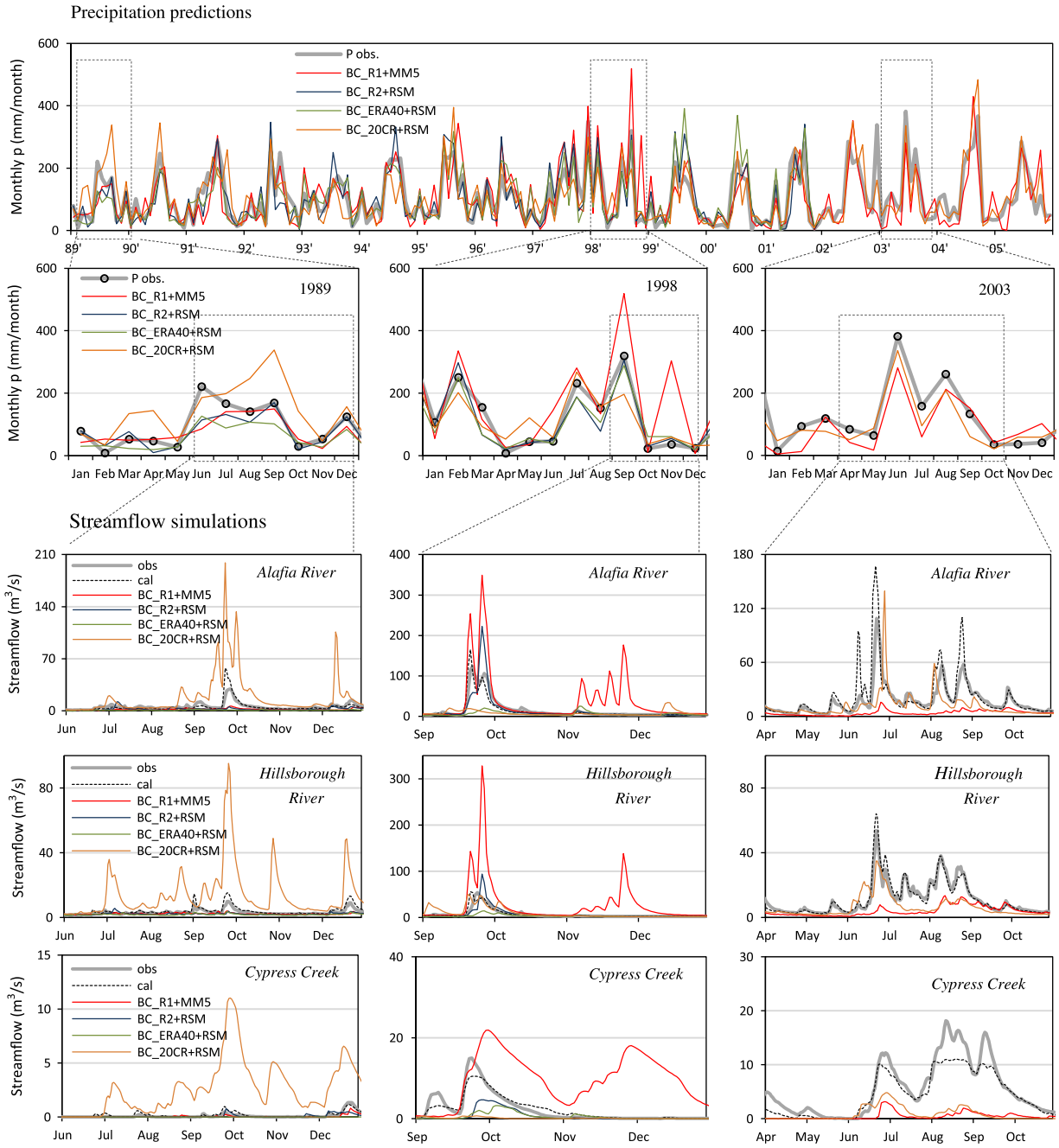

Fig. 7. Comparison of monthly observed and predicted precipitation over the study period and daily streamflow simulations for marked periods in the Fig. 7. The boxes indicate the dominant months for errors.

be directly used as input to INTB because those results have been bias corrected directly onto the model basin centroids.

To obtain bias-corrected maximum and minimum daily temperature predictions (hereafter, T_{max} and T_{min}) mean daily temperature ($T_{max} + T_{min}/2$; Hargreaves and Samani, 1985) was first bias corrected with point observations using the CDF mapping approach described above for precipitation. Then the observed T_{max} and T_{min} corresponding to the bias-corrected mean temperature were obtained from the historical record. This approach allows the bias-corrected results to preserve observed daily temperature range ($T_{max} - T_{min}$) and avoids potentially unrealistic results that might occur if T_{max} and T_{min} were bias corrected separately (for example the possibility that the corrected T_{min} would be greater than T_{max}).

4.2. Target stations for INTB model evaluation

Based on the importance to water supply management and variability of flow characteristics over the study area three streamflow stations on major rivers were chosen to evaluate hydrologic response to climate predictions (see Fig. 1 and Table 3). The Alafia and Hillsborough rivers have a mean discharge of 8.6 m^3/s and 6.2 m^3/s, respectively with very few no flow days whereas Cypress Creek has a mean discharge of less than 1.5 m^3/s. Furthermore Cypress Creek has a large percentage (approximately 20%) of no flow days. Investigating stations with large and small flow volumes is important to understand how different types of flow regimes are affected by changes in climate variables and/or errors in their estimates. Additionally, the flows at these stations are important for

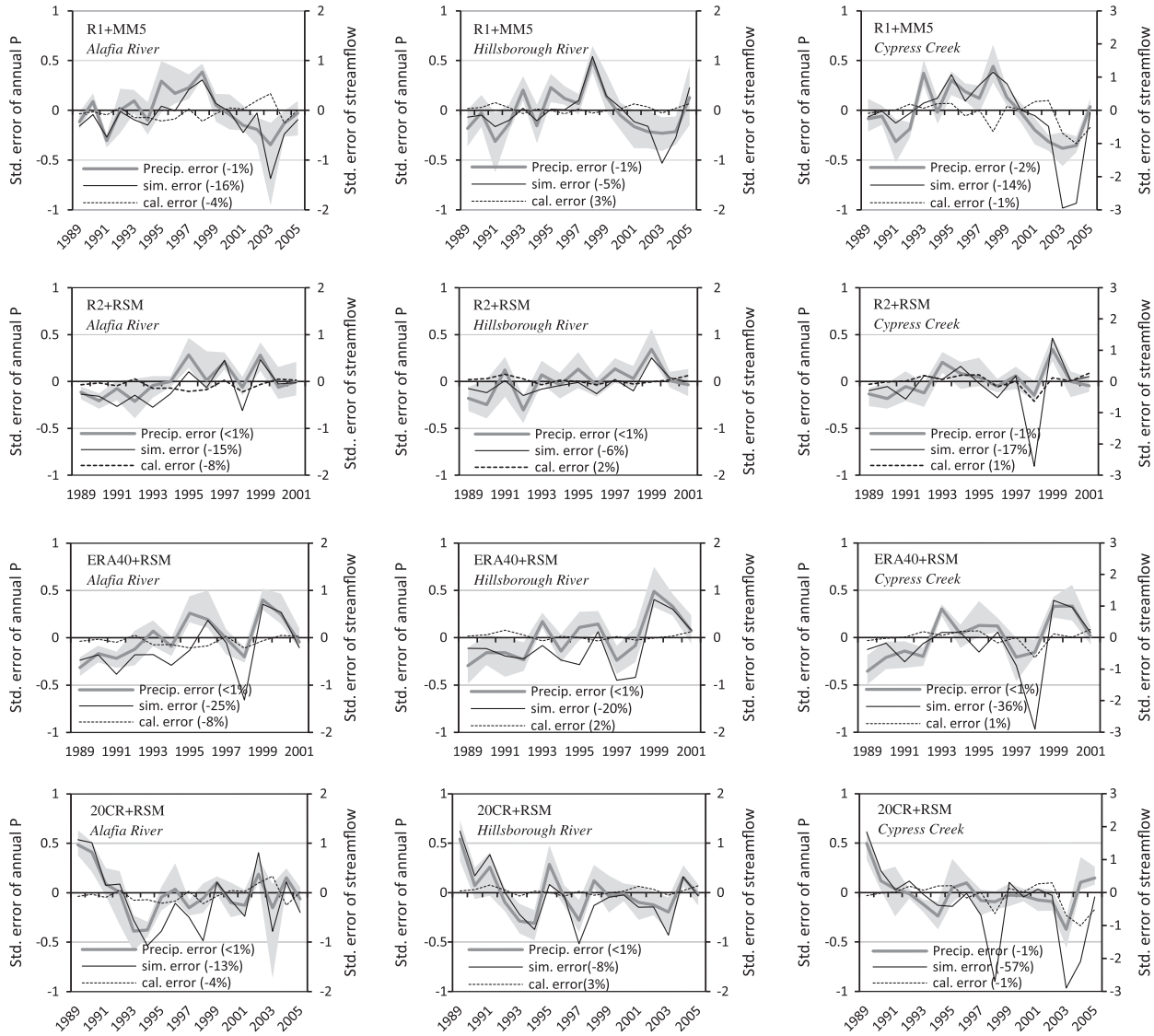

Fig. 8. Comparison of the standardized error of annual bias-corrected regional reanalysis precipitation data (each row) for the contributing basins to each streamflow station (each column) to standardized error of the calibrated and simulated annual streamflow simulations. The mean of standardized errors over the study period for precipitation and streamflow simulations are represented on each legend. The gray zones represent the range of standardized errors of annual precipitation over the contributing basins for each watershed.

water supply operations and management because they are either located near or downstream of wellfields or water is withdrawn from them directly to meet local water demand.

Groundwater levels were evaluated for both unconfined and semi-confined Floridan aquifer conditions. Two pairs of surficial and Floridan aquifer monitoring wells were chosen near each of two TBW wellfields. In addition one unconfined Floridan aquifer well in the northern part of the model domain was evaluated (see Fig. 1).

5. Results

5.1. Raw and bias-corrected regional reanalysis data

The raw R2+RSM, ERA40+RSM, and 20CR+RSM dynamically-downscaled reanalysis data estimated the annual cycle of both daily maximum and minimum temperature accurately with mean errors within less than ±0.4 °C. The raw R1+MM5 data estimated the daily minimum temperature within 0.5 °C, but consistently underestimated daily maximum temperature, with a mean error of −6.7 °C. After bias-correction mean errors for all dynamically downscaled reanalysis data were reduced to less than ±0.3 °C for both T_{min} and T_{max} (for graphical presentation of results by month, see Supplementary material).

Fig. 3 compares the mean monthly precipitation across the model domain and over the study period for the basin-based observations (P obs.), raw regional reanalysis, and bias-corrected regional reanalysis results. The figure confirms that all four regional reanalysis datasets tend to overestimate the mean monthly precipitation with mean errors (ME) ranging from 10.8 mm for the ERA40+RSM to 68.6 mm for R1+MM5. In particular, the raw R1+MM5, R2+RSM, and 20CR+RSM results showed significant overestimation of precipitation during the dry season. As expected, the bias-corrected mean monthly precipitation results for all four datasets closely match the mean of the observations that were used for bias-correction. The minor deviation of some of the bias-corrected results from the observations is a result of the slight differences in duration of the study period between the reanalysis datasets and the observation period.

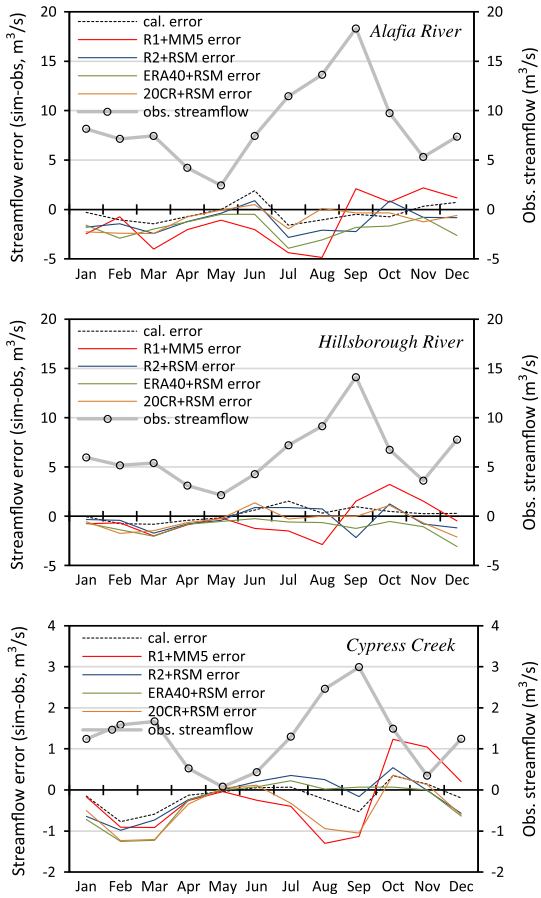

Fig. 9. Observed mean monthly streamflow and errors (sim.-obs.) in simulated mean monthly streamflow for each target station.

Fig. 4 compares the standard deviations of daily and monthly precipitation by month. The figure indicates that bias-correction also significantly improves the daily standard deviations (with ME less than 0.3 mm i.e., 7% of observed daily standard deviation), but that monthly standard deviation errors remain even after bias correction (for example, ME for R1+MM5 is 14.5 mm which is 22% of observed monthly standard deviation). These results confirm that while independent daily statistics are reproduced through daily bias correction, errors persist in rainfall statistics at other time-scales.

Fig. 5 shows the actual time series of observed, raw regional reanalysis, and bias-corrected regional reanalysis annual and seasonal totals over the study period. This figure confirms all raw downscaled reanalysis data are positively biased, with R1-MM5 showing the most pronounced bias, particularly in the dry seasons. After bias-correction, mean errors for all results were less than 1% of mean annual observed rainfall. However, persistent periods of overestimation (e.g., 1995–1999 for R1+MM5, 1989–1991 for 20CR+RSM, and 1999–2001 for ERA40+RSM) and underestimation (e.g., 1989–1992 for R1+MM5 and ERA40+RSM; 2001–2003 for R1+MM5 and 20CR+RSM; and 1993–1994 for 20CR+RSM) of annual rainfall totals occur (Fig. 5a), mainly due to persistent errors in wet season totals (Fig. 5d). These results emphasize that bias-correction through daily CDF-mapping only corrects statistical properties (e.g., mean, variance, CDF) at the daily scale. Errors in daily sequences of rainfall produce errors in both temporal sequences and higher-order statistics for longer (i.e. monthly and annual) time-scales.

5.2. Streamflow simulation results

Each of the bias-corrected regional reanalysis datasets were used to drive the INTB model. Streamflow simulations were compared to INTB calibrated model results to explore hydrologic implications due to errors in climatic forcing data from the regional reanalysis, and to observed data to explore total errors due to both climate and hydrologic model error. Daily and monthly time series, a comparison of the standardized error of mean annual precipitation to the standardized error of mean annual streamflow, the annual cycle of mean monthly discharge, and daily streamflow exceedence probabilities were used to evaluate streamflow results.

Fig. 6 shows the monthly streamflow hydrographs from 1989 to 2005 for the selected target stations. Calibrated streamflow simulation results, which used observed precipitation and temperature data, fit the observed streamflow observation well over the entire study period, with coefficients of determination (R^2) for monthly streamflow ranging from 0.90 for Alafia River basin to 0.98 for Hillsborough River basin, and mean errors in monthly streamflow ranging from $-0.4\ m^3/s$ for Alafia River basin to $-0.1\ m^3/s$ for Cypress Creek basin (see mean error and R^2 values indicated on Fig. 6). The streamflow predictions using bias-corrected reanalysis data generally agree with observations and calibrated predictions in timing, but the magnitude of monthly streamflow shows significant errors at several times which vary with the reanalysis data set. As a result coefficients of determination drop to 0.19–0.80, when the streamflow predictions using bias-corrected regional reanalysis are compared to the streamflow observations. It should be noted that mean errors deteriorate less dramatically to $-1.9\ m^3/s$ to $-0.1\ m^3/s$ when bias-corrected predictions are compared to observations and to $-1.5\ m^3/s$ to $0.1\ m^3/s$ when the results are compared to the calibrated results. Three time periods where major errors occur (i.e., in 1989, 1998, and 2003) are marked in Fig. 6.

Fig. 7 compares observed and predicted monthly precipitation predictions (averaged over the whole domain) and daily streamflow simulations at the three target stations to observations for the periods marked in Fig. 6. This figure shows that streamflow predictions are very sensitive to errors in the temporal patterns of actual precipitation and indicates that streamflow errors associated with long periods of rainfall underestimation (e.g., 1989 for ERA40+RSM and 2003 for R1+MM5) are more persistent than errors associated with long periods of rainfall overestimation (e.g., 1989 for 20CR+RSM and 1998 for R1+MM5). This is due to the slow, persistent depletion of soil water and groundwater storage during droughts which takes time to replenish when rainfall returns, versus the rapid triggering of surface runoff and streamflow that is lost from the domain when rainfall occurs during persistent wet periods.

Fig. 8 compares the standardized error of annual precipitation predictions (i.e. (annual reanalysis P-annual Pobs.)/mean annual Pobs. for each contributing area) to the standardized error of annual reanalysis streamflow predictions (i.e. (reanalysis streamflow – annual observed streamflow)/mean annual observed streamflow) and the standardized error for the calibrated model streamflow for each streamflow target station. These figures illustrate the relationship between the errors in precipitation input from bias-corrected regional reanalysis and streamflow simulation errors at the annual time scale. Errors in streamflow simulation closely follow errors in precipitation prediction in timing and pattern. However underestimation of precipitation produces larger and more persistent errors in streamflow than overestimation of precipitation, particularly for the larger streamflow stations (Alafia and Hillsborough Rivers). For example during the first four years of the simulation period (1989–1992) R1+MM5, R2+RSM, and ERA40+RSM all underestimate precipitation producing

Table 4
Comparison of observed versus predicted percentiles of daily streamflow and the percentage of days extraction would be permitted for Alafia River, Hillsborough River and Cypress Creek. Results were calculated using all available data (i.e., 1989 to 2005 for observations, calibrated results, R1+MM5 and 20CR+RSM and 1989 to 2001 for R2+RSM and 20CR+RSM). The values in brackets for observed and calibrated results are calculated for the period from 1989 to 2001 comparable with the R2+RSM and ERA40+RSM results.

(m^3/s)	Obs.	Cal.	R1+MM5	R2+RSM	ERA40+RSM	20CR+RSM
Alafia River at Lithia						
Avg. flow	8.56 (7.33)	8.20 (6.67)	7.27	6.15	5.46	7.59
25%	2.15 (1.95)	2.40 (2.12)	1.58	1.30	1.08	2.16
50%	4.33 (3.65)	4.26 (3.66)	3.42	3.11	3.09	3.83
75%	9.26 (7.65)	7.79 (6.62)	6.73	6.22	5.71	6.96
95%	28.88 (24.92)	25.78 (21.07)	21.58	21.07	18.17	25.15
Max.	278.1 (143.3)	251.5 (172.61)	349.0	234.08	222.17	281.44
No flow days (%)	–	0.4 (0.5)	0.6	2.3	2.6	2.1
Flow withdrawal range			3–22 m^3/s			
% Withdrawal days	56.2 (51.6)	59.5 (55.4)	49.5	46.76	47.81	56.72
Hillsborough River near Zephyrhills						
Avg. flow	6.23 (5.46)	6.42 (5.62)	5.86	5.12	4.39	5.79
25%	2.01 (1.87)	1.97 (1.82)	1.71	1.53	1.54	1.77
50%	2.97 (2.58)	3.05 (2.71)	2.66	2.33	2.26	2.67
75%	5.32 (4.45)	5.60 (4.48)	4.98	4.28	4.00	4.86
95%	21.97 (19.14)	23.77 (19.65)	18.73	18.83	15.52	20.95
Max.	195.4 (195.39)	190.8 (190.81)	328.30	185.33	114.14	345.96
No flow days (%)	–	–	–	–	–	–
Flow withdrawal range			1–18 m^3/s			
% Withdrawal days	91.6 (92.2)	90.2	87.2	87.5	85.3	85.0
Cypress Creek at Worthington Gardens						
Avg. flow	1.28 (0.77)	1.11 (0.76)	1.04	0.61	0.48	0.82
25%	–	–	–	–	–	–
50%	0.09 (0.03)	0.11 (0.14)	0.08	0.00	0.00	0.09
75%	0.82 (0.31)	0.87 (0.51)	0.77	0.37	0.28	0.62
95%	6.77 (3.26)	5.79 (3.96)	5.05	3.91	3.11	4.29
Max.	50.12 (36.25)	17.24 (13.77)	21.88	14.04	9.74	15.40
No flow days (%)	17.2 (21.7)	37.5 (44.3)	42.7	49.1	54.9	40.8
Flow withdrawal range			0.1–10 m^3/s			
% Withdrawal days	46.0 (37.13)	48.6 (41.83)	45.3	37.9	34.9	47.7

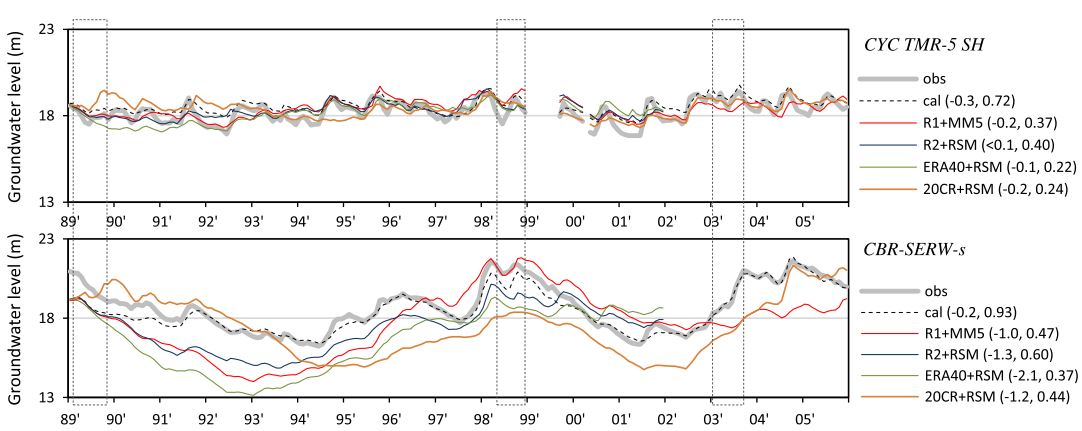

Fig. 10. Comparison of monthly averaged groundwater level predictions for each target station in surficial aquifer. The boxes indicate the periods of significant streamflow error that are marked in Fig. 7. The values in the brackets indicate the mean error of the monthly streamflow (m^3/s) and R^2.

underestimated streamflow and over-drying of the hydrologic system. In each case when precipitation is overestimated after this artificially dry period streamflow errors remain negative or close to zero for all stations indicating that excess precipitation is filling storage in the soils and aquifers of the basin rather than producing excess streamflow. Similarly the persistent underestimation of rainfall from 2000 to 2004 by R1+MM5 produces persistent underestimation of streamflow at all 3 target stations, with errors that are significantly magnified in 2003 and 2004. In contrast the over-predicted rainfall period observed from 1995 through 1999 for R1+MM5, produces over-predicted streamflows with approximately the same pattern of standardized errors as precipitation during this period, and the streamflow errors return immediately to zero in 2000 when precipitation is accurately estimated. Similar patterns occur during the period 1996 to 1997 when precipitation is over-predicted b EyRA40+RSM and from 1989 to 1993 when precipitation is over-predicted by 20CR+RSM. Errors are significantly smaller and less persistent for the calibrated model streamflow errors than the reanalysis streamflow errors.

Fig. 9 shows the observed mean monthly streamflow and the errors of monthly streamflow predictions using observation data (i.e., calibrated results) and bias-corrected reanalysis data at each of the streamflow locations. Even though mean monthly rainfall was accurately reproduced by all reanalysis datasets, because of

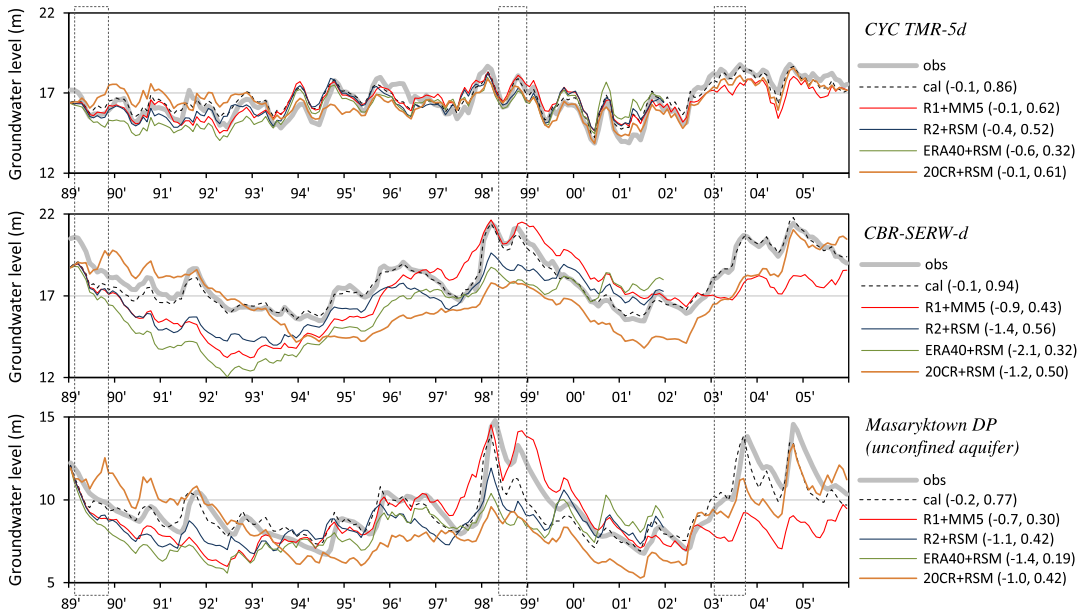

Fig. 11. Comparison of monthly averaged groundwater level predictions for 2 target stations in semi-confined Floridan aquifer . The boxes indicate the periods of significant streamflow error that are marked in Fig. 7. The values in the brackets indicate the mean error of the monthly streamflow (m^3/s) and R^2.

the higher sensitivity of the hydrologic system to persistent under-predicted rainfall than over-predicted rainfall, the mean monthly reanalysis streamflows are generally underestimated at all target stations except for late in the rainy season (i.e. September–October). At Cypress Creek, which has relatively low streamflow and significant surface flow attenuation due to conditionally connected wetlands, the magnitude of the streamflow error is larger (relative to the mean streamflow) than at the higher streamflow stations. Wetlands are modeled in INTB to contribute to the terminal discharge of a basin only when available storage within the wetland is filled to a specified water depth. Thus over-drying in the spring and early summer produces conditions that significantly delay streamflow generation processes, particularly in small watersheds.

For Tampa Bay Water extraction of surface water from the Hillsborough River, Alafia River, and Cypress Creek is permitted at times when streamflow is above permitted thresholds. For example, for the Alafia River streamflow withdrawals are permitted when daily flows are above approximately 3 m^3/s but are typically no longer required when daily streamflow exceeds 22 m^3/s because demand drops and reservoir storage capacity is generally at capacity. Table 4 compares streamflow corresponding to various flow percentiles at each target station as well as the percent of time that streamflow withdrawals are allowed (and required) from each station. For the larger streams (Alafia and Hillsborough River) all regional reanalysis data underestimate streamflow values at virtually all flow percentiles ($\leqslant$95%), indicating that flow is under-predicted more than 95% of the time. Furthermore the percentage of days when streamflows are within the operable range for withdrawals is underestimated, implying less surface water availability for domestic supply than the observed or calibrated cases would indicate.

5.3. Groundwater level simulation results

Similar to the streamflow analyses, time series of monthly groundwater level predictions, a comparison of the standardized error of mean annual precipitation to the standardized error of mean annual groundwater level predictions, and the annual cycle of mean monthly groundwater level were used to evaluate the sensitivity groundwater level predictions to errors in bias-corrected reanalysis precipitation data.

Figs. 10 and 11 show observed, calibrated, and bias-corrected regional reanalysis driven monthly groundwater level time series for the target wells in the surficial aquifer and Floridan aquifer respectively for two regions of the domain. The CYC TMR-5SH and CYC TMR-5d wells are located in a region where the Floridan aquifer is semi-confined and surface runoff is prevalent. Because precipitation exerts relatively little control on groundwater levels in the semi-confined region the reanalysis groundwater level predictions fit the observed groundwater levels quite well at the CYC TMR-5SH surficial and CYC TMR-5d Floridan wells (ME within ± 0.2 m compared to −0.3 m for the calibrated model for the surficial well and ME within ± 0.6 m for the Floridan well compared to −0.1 m for the calibrated model). In contrast the CBR-SERW-s and CBR-SERW-d wells are located in the area where the Floridan transitions to unconfined conditions. In this region relatively little surface runoff occurs and precipitation exerts a stronger control on groundwater levels. As a result the bias-corrected reanalysis data predictions for groundwater levels at the CBR-SERW-s surficial and CBR-SERW-d Floridan wells deteriorated significantly compared to the observations (ME ranging from −1.0 m to −2.1 m compared to −0.2 m for the calibrated results for the surficial well, and errors ranging from −0.9 m to −2.1 m compared to −0.1 m for the calibrated results for the Floridan well). Similarly the Masaryktown well located in the unconfined Floridan aquifer showed significant impact of precipitation variability (e.g., temporal fluctuation patterns) and errors in precipitation forcing data because surface runoff does not attenuate precipitation variability in this region.

Fig. 12 compares the standardized error of annual precipitation (i.e. (annual regional reanalysis P-annual Pobs.)/mean annual Pobs.) over the study domain to the standardized error of the annual groundwater level predictions (annual groundwater level simulation-annual observed groundwater level)/mean observed groundwater level) over the study period. This figure shows that bias corrected regional reanalysis data errors in the groundwater system lag errors in the precipitation inputs and persist for longer than in the surface water system in the unconfined region, due to

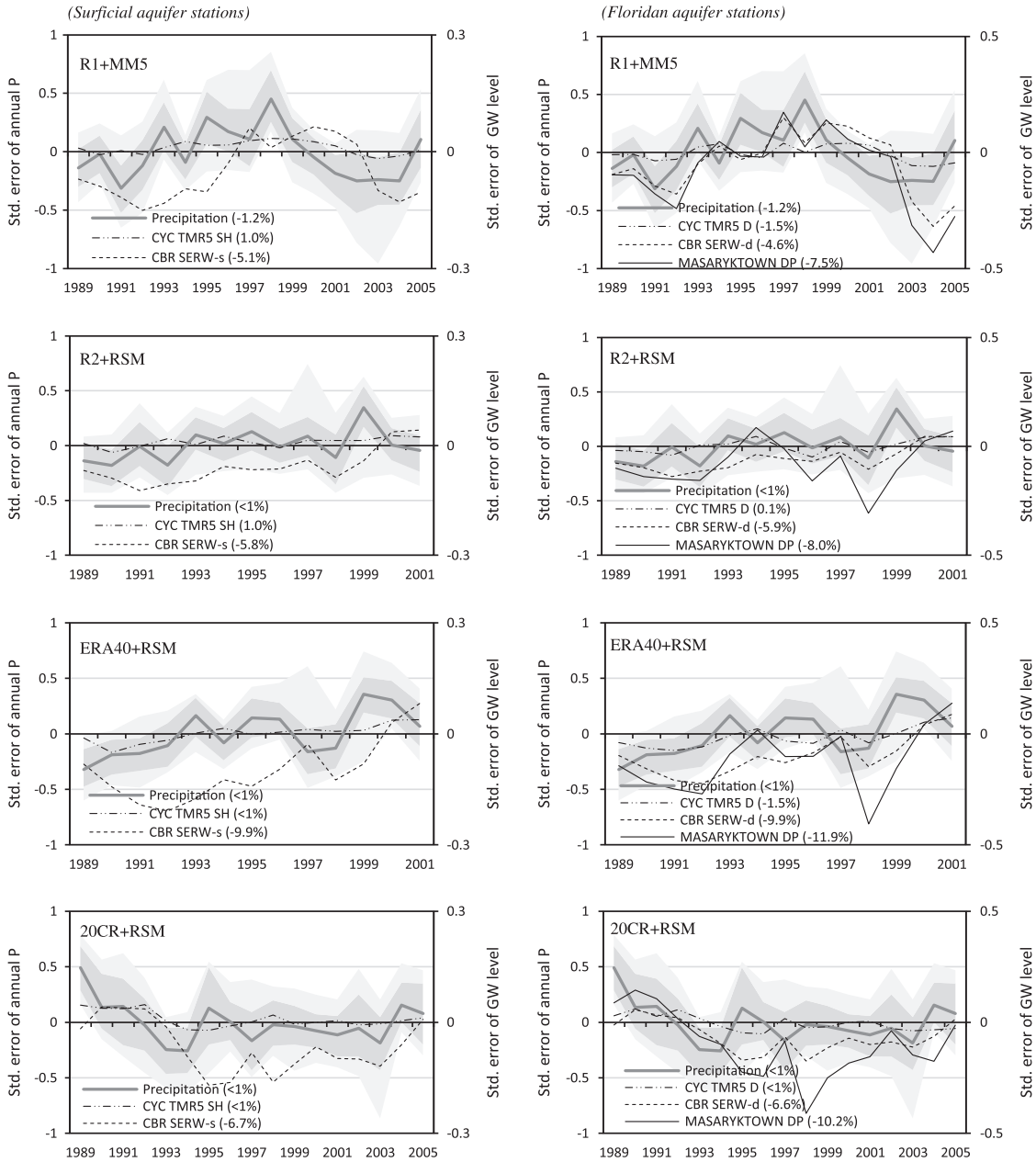

Fig. 12. Comparison of the standardized error of annual regional reanalysis precipitation data to annual groundwater level simulations using each reanalysis data (each row) for surficial (left column) and Floridan aquifer (right column). Mean of standardized errors over the study period are represented on each legend. The gray zones represent the range of precipitation errors over the 172 basins.

diffusive groundwater flow processes. In particular, during the first half of the study period, the underestimated R1+MM5, R2+RSM, and ERA40+RSM precipitation from 1989 to 1991 leads to underestimated groundwater levels through 1995 or later at the CBR-SERW wells and Masaryktown well. Furthermore, the significant over-prediction of R1+MM5 rainfall from 1995 through 1998 leads to over-prediction of groundwater levels which persisted until the end of 2000. In contrast while over-predictions of R2+RSM and ERA40 precipitation in 1993 and 1995 decrease the groundwater level underestimation slightly, levels remain underestimated compared to observations. These long periods of persistent underestimation of groundwater levels resulted in systematically low predictions of mean groundwater level over the study period for the CBR-SERW wells and Masaryktown well (Fig. 13). The standardized error of calibrated groundwater level predictions (i.e. (annual calibrated results – annual observed results)/mean annual observed results) are smaller and less persistent than the reanalysis errors, resulting in mean groundwater level predictions close to observed data, indicating good performance of calibrated INTB model.

6. Discussion

6.1. Errors in regional reanalysis data

All raw regional reanalysis data showed significant errors in mean daily, monthly and annual precipitation, which were removed by bias-correction using CDF mapping on a daily basis. Furthermore, because bias-correction was conducted individually for each basin in the model domain, the bias-corrected regional

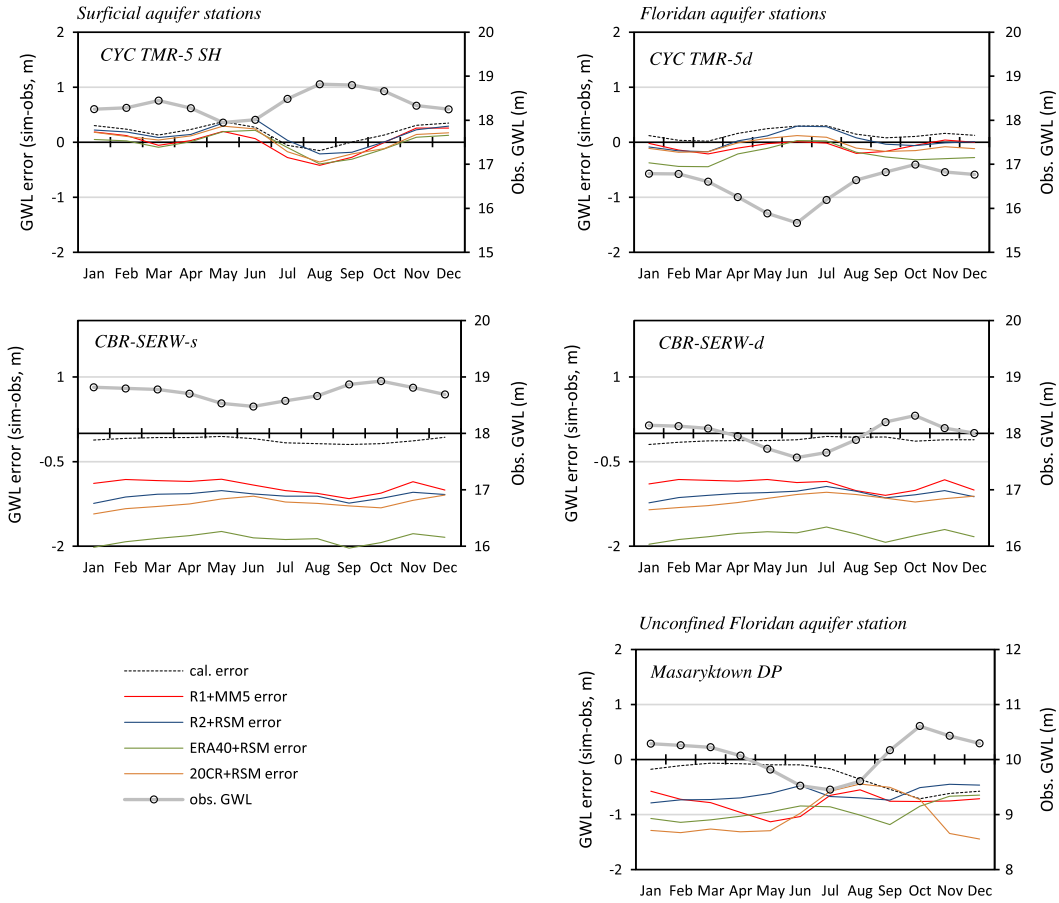

Fig. 13. Observed mean monthly groundwater level (GWL, secondary axis) and the errors of monthly GWL simulations using bias-corrected regional reanalysis (primary axis) for 2 pairs of surficial (left column)/Floridan aquifer (right column) target stations.

reanalysis data accurately reproduced the observed spatial pattern of the mean daily, monthly and annual precipitation fields over the model domain. Errors in the standard deviation of daily precipitation were also removed after bias-correction, but the standard deviation of the monthly and annual precipitation totals were not accurately reproduced even after bias-correction. Furthermore inaccuracies in actual raw daily precipitation time series were not corrected by bias-correction and thus bias-corrected daily precipitation times series aggregated into monthly and annual rainfall total time series that showed significant and temporally persistent errors. Similar limitations were found by Haerter et al. (2011) when bias-correcting coarse GCM outputs (~200 km^2 spatial resolution) on a daily basis.

Techniques have been developed to bias-correct climate model predictions at multiple time scales (e.g., nested bias-correction: Johnson and Sharma, 2012; cascade bias-correction: Haerter et al., 2011) in an effort to overcome this problem. However complex statistical adjustment of RCM outputs may eliminate the benefits of dynamical climate modeling by adjusting the climate model results without maintaining physical consistency among climate outputs (Johnson and Sharma, 2011; Ehret et al., 2012) and calls into question the utility of conducting computationally expensive dynamic downscaling instead of directly statistically downscaling and bias correcting the large-scale GCM output. Furthermore bias-correcting independently on a grid/point basis may disrupt regional spatial correlation patterns that may be predicted by RCMs. It is possible that bias-correction of the GCM or reanalysis forcing data prior to RCM simulation (e.g., Colette et al., 2012) could reduce this problem.

In order to account for the uncertainty and influence of bias-correction and downscaling techniques, researchers have recently proposed ensemble scenarios with different bias-correction methods, (e.g., Teutschbein and Seibert, 2010, 2012b), different downscaling methods (e.g., Chen et al., 2011; Dosio and Paruolo, 2011; Teutschbein et al., 2011), and different application models (e.g., Teutschbein et al., 2011; Velázquez et al., 2013). These studies have indicated that the uncertainty in regional climate projections due to different RCMs is greater than the uncertainty due to different bias-correction methods (Teutschbein and Seibert, 2012a) and as large as the uncertainty of due to different GCM forcing data (Giorgi, 2006). This study utilized 4 regional reanalysis products that were produced using 4 different GCM reanalysis products and two RCMs. Significant differences were found among the same RCM results forced by different GCM reanalysis data (i.e., R2+RSM, ERA40+RSM, and 20CR+RSM) indicating the important influence of potential errors or differences in the initial and boundary conditions used to force RCMs. In general the R2+RSM and ERA40+RSM products showed better skill in reproducing historical precipitation time series compared to R1+MM5 or 20CR+RSM; however, bias-correction reduced the variations among the results.

It should be noted that this study bias-corrected the regional reanalysis using an observation period that coincided with the hydrologic simulation period, which by construct produces a good match between the observed and bias-corrected CDF at the time-scale of correction (daily in this case). To test the accuracy of bias-correction methods the observation period is often split into two components, one of which is used to develop the CDFs for bias correction and the other to verify the process (e.g., differential

split-sample test: Piani et al., 2010a; Dosio and Paruolo, 2011; Teutschbein and Seibert, 2012b; generalized split-sample test: Coron et al., 2012, and bootstrapping approach: Li et al., 2010). In this case small errors in statistics can be expected even at the timescale of correction due to differences in observational records over the finite bias-correction and verification intervals. We did not conduct cross-validation in this work because the goal of the study was to examine the propagation of errors from bias-corrected regional reanalysis data into hydrologic predictions in order to evaluate the feasibility of using regional reanalysis data as a proxy for observations, without the confounding influences of additional bias-correction errors.

6.2. Hydrologic implications of errors in the regional reanalysis data

Even though mean monthly precipitation was accurately produced by all bias-corrected reanalysis datasets, daily precipitation timing errors were propagated and enhanced by non-linear streamflow generation processes in the model to produce errors in mean monthly streamflow throughout the domain. In fact, frequency analysis of monthly streamflow showed that all regional reanalysis datasets underestimated streamflow more than 95% of the time. As a result the bias-corrected reanalysis datasets predicted less streamflow available for extraction for public water supply than either the observed streamflow or calibrated model estimated, implying potential water resource planning implications of the regional reanalysis data streamflow errors. A similar frequency analysis of streamflow was derived from RCM results by Kleinn et al. (2005) for locations in central Europe. They also found significant errors for extreme streamflow conditions although they found that mean monthly streamflow showed good agreement.

Groundwater simulation results showed different response to the errors in precipitation input data over the study domain based on the degree of confinement of the aquifer from the surface system. In the portion of the model domain where the aquifer is confined to semi-confined both actual monthly groundwater levels and average monthly groundwater levels over the study period were accurately predicted by all bias-corrected reanalysis datasets. However in the portion of the domain where the Floridan aquifer is unconfined precipitation variability more directly influences groundwater level variability. In this area precipitation timing errors produced persistent errors in monthly groundwater levels which resulted in significant under-prediction of monthly averaged groundwater levels over the study period. These under-predictions would have significant water resource planning implications since they imply persistent violation of minimum aquifer levels regulated in the area.

Results of this study indicate that reproducing more detailed precipitation characteristics than the daily rainfall CDF (e.g., temporal persistence, inter-event durations, and variance of monthly and annual totals) is important to accurately capture hydrologic behaviors important for water resource planning in the Tampa Bay region. In other words, errors in the downscaled reanalysis data bias-corrected at the daily timescale produces operationally significant hydrologic errors because reproducing the individual daily frequency of precipitation events by itself does not constrain all of the characteristics important for determining hydrologic response (Frei et al., 2003; Seneviratne et al., 2002). Errors in predicting mean hydrologic behavior could possibly be improved if a longer study period allowed more temporal averaging to dampen the persistent timing errors; however a longer study period would not improve the model skill scores for the actual time series of hydrologic predictions.

It is possible that the skill of the INTB model could be improved if the model were re-calibrated using the bias-corrected regional reanalysis datasets (e.g., Chen et al., 2011). However the major differences between the observed and bias-corrected regional reanalysis precipitation patterns at monthly and annual timescales, and the strong correspondence between the observed precipitation patterns and observed flow regimes, indicates that the errors in the bias-corrected regional reanalysis results are not sufficiently small to be overcome by the hydrologic model through the parameter calibration process. Thus re-calibration of the INTB model was not conducted for this study.

7. Conclusions

The goal of this study was evaluate the feasibility of using currently available regional reanalysis datasets, generated from four large scale reanalysis datasets and two regional climate models, as surrogates for observational climatic data needed to drive integrated hydrologic models for low-relief rainfall-driven hydrologic systems in west central Florida. Results showed that, even after bias-correction, precipitation errors in the regional reanalysis datasets were propagated and enhanced by non-linear streamflow generation and groundwater flow and storage processes in the hydrologic model, producing low hydrologic model prediction skill and potential water resource planning implications. Streamflow exceedence probability predictions showed significant errors in predicting the duration of flow-ranges of water supply significance, implying under-prediction of surface water availability for public supply. Furthermore precipitation errors resulted in low mean monthly groundwater levels in the unconfined regions of the aquifer that could mistakenly predict persistent violation of minimum regulated groundwater levels that would trigger mandated reductions in groundwater pumping.

Findings of this study indicate that simple correction of dynamically downscaled regional reanalysis predictions using a daily CDF-mapping approach is insufficient for predicting hydrologic behavior with the accuracy needed to make reliable water resource planning decisions in low-relief rainfall-driven systems. Resolution of deficiencies in hydrologic predictions will require improvements large-scale reanalysis products used as boundary conditions for the regional models, improvement in the regional climate model physics and parameterization, or development of enhanced bias-correction techniques that correct not only daily statistics but also errors in precipitation correlation structures over time and space.

Acknowledgments

We would like to thank the anonymous reviewers for providing many constructive comments. This research was supported in part by the Sectoral Applications Research Program (SARP) of the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office and by Tampa Bay Water.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhydrol.2013.11.042.

References

Abatzoglou, T.J., Brown, J.T., 2011. A comparison of statistical downscaling methods suited for wildfire applications. Int. J. Climatol., http://dx.doi.org/10.1002/joc.2312.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, Rome.

Berg, A.A., Famiglietti, J.S., Walker, J.P., Houser, P.R., 2003. Impact of bias correction to reanalysis products on simulations of North American soil moisture and

hydrological fluxes. J. Geophys. Res. 108 (D16), 4490. http://dx.doi.org/10.1029/2002JD003334.
Bicknell, B., Imhoff, J.C., Kittle Jr., J.L., Jobes, T.H., Donigian Jr., A.D., 2001. Hydrologic Simulation Program-FORTRAN (HSPF): User's Manual for Version 12. U.S. Environmental Protection Agency, Athens, GA.
Chen, J., Brissette, F.P., Leconte, R., 2011. Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. J. Hydrol. 401, 190–202.
Christensen, J.H., Christensen, O.B., 2003. Severe summertime flooding in Europe. Nature 421, 805–806.
Christensen, J.H., Machenhauer, B., Jones, R.G., Schär, C., Ruti, P.M., Castro, M., Visconti, G., 1997. Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. Clim. Dyn. 13, 489–506.
Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys. Res. Lett. 35, L20709.
Colette, A., Vautard, R., Vrac, M., 2012. Regional climate downscaling with prior statistical correction of the global climate forcing. Geophys. Res. Lett. 39 (13). http://dx.doi.org/10.1029/2012GL052258.
Compo, G.P. et al., 2011. The twentieth century reanalysis project. Quart. J. Roy. Meteorol. Soc. 137, 1–28. http://dx.doi.org/10.1002/qj.776.
Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., Hendrickx, F., 2012. Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. Water Resour. Res. 48 (5), W05552. http://dx.doi.org/10.1029/2011WR011721.
DiNapoli, S.M., Misra, V., 2012. Reconstructing the 20th century high-resolution climate of the southeastern United States. J. Geophys. Res. 117, D19113. http://dx.doi.org/10.1029/2012JD018303.
Doherty, J., 2004. PEST: Model-independent Parameter Estimation, fifth ed. Watermark Numer. Comput., Brisbane, Queensl., Australia.
Dosio, A., Paruolo, P., 2011. Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: evaluation on the present climate. J. Geophys. Res. 116, D16106. http://dx.doi.org/10.1029/2011JD015934.
Dudhia, J., 1996. A multi-layer soil temperaturemodel for MM5. Preprints, the 6th PSU/NCARMesoscale Model Users Workshop, Boulder CO, July, National Center for. Atmos. Res., 49–50.
Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., Liebert, J., 2012. HESS Opinions "Should we apply bias correction to global and regional climate model data?". Hydrol. Earth Syst. Sci. 16 (9), 3391–3404. http://dx.doi.org/10.5194/hess-16-3391-2012.
Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., Tarpley, J.D., 2003. Implementation of Noah land surface model advances in the national centers for environmental prediction operational mesoscale Eta model. J. Geophys. Res. 108 (D22), 8851. http://dx.doi.org/10.1029/2002JD003296.
Fowler, H.J., 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.
Francis, J.A., 2002. Validation of reanalysis upper-level winds in the Arctic with independent rawinsonde data. Geophys. Res. Lett. 29 (9), 1315. http://dx.doi.org/10.1029/2001GL014578.
Frei, C., Christensen, J.H., Déqué, M., Jacob, D., Jones, R.G., Vidale, P.L., 2003. Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. J. Geophys. Res. 108, 4124. http://dx.doi.org/10.1029/2002JD002287.
Fujihara, Y., Tanaka, K., Watanabe, T., Nagano, T., Kojiri, T., 2008. Assessing the impacts of climate change on the water resources of the Seyhan river basin in Turkey: use of dynamically downscaled data for hydrologic simulations. J. Hydrol. 353, 33–48.
Geurink, J., Basso, R., 2011. 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., 2006a. 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, April 2-6, Reno, NV.
Geurink, J., Trout, K., Ross, M., 2006b. Introduction to the integrated hydrologic model. In: Proceedings of the Joint Federal Interagency Conference 2006, April 2-6, Reno, NV.
Giorgi, F., 2006. Regional climate modeling: status and perspectives. J. Phys. IV 139, 101–118. http://dx.doi.org/10.1051/jp4:2006139008.
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: Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dia, X., Maskell, K., Johnson, C.A. (Eds.), Climate Change 2001: The Scientific Basis. Cambridge University Press, Cambridge, 583–638, 739–768.
Graham, L.P., Hagemann, S., Jaun, S., Beniston, M., 2007. On interpreting hydrological change from regional climate models. Clim. Change 81, 97–122. http://dx.doi.org/10.1007/s10584-006-9217-0.
Grell, G.A., Dudhia, J., Stauffer, D.R., 1994. A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 117 and 122.
Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen-Skaugen, T., 2012. Technical note: downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. Hydrol. Earth Syst. Sci. 16 (9), 3383–3390. http://dx.doi.org/10.5194/hess-16-3383-2012.
Haerter, J.O., Hagemann, S., Moseley, C., Piani, C., 2011. Climate model bias correction and the role of timescales. Hydrol. Earth Syst. Sci. 15, 1065–1079. http://dx.doi.org/10.5194/hess-15-1065-2011.
Harbaugh, A.W., McDonald, M.G., 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, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1, 96–99.
Hay, L.E., Clark, M.P., 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.
Hay, L.E., Clark, M.P., Wilby, R.L., Gutowski, W.J., Leavesley, G.H., Pan, Z., Arritt, R.W., Takle, E.S., 2002. Use of regional climate model output for hydrologic simulations. J. Hydrometeorol. 3, 571–590.
Hong, J.S., 2003. Evaluation of the high-resolution model forecasts over the Taiwan area during GIMEX. Weather Forecast. 18, 836–846.
Hong, S.Y., Pan, H.L., 1996. Non-local boundary layer vertical diffusion in a medium-range forecast model. Mon. Weather Rev. 124, 2322–2339.
Hwang, S., Graham, W., Hernández, J.L., Martinez, C., Jones, J.W., Adams, A., 2011. Quantitative spatiotemporal evaluation of dynamically downscaled MM5 precipitation predictions over the Tampa Bay region, Florida. J. Hydrometeorol. 12 (6), 1447–1464.
Hwang, S., Graham, W., Adams, A., Geurink, J., 2013. Assessment of the utility of dyncamically-downscaled regional reanalysis data to predict streamflow in west central Florida using an integrated hydrologic model. Reg. Environ. Change in press, DOI: http://dx.doi.org/10.1007/s10113-013-0406-x.
Ines, A.V.M., Hansen, J.W., 2006. Bias correction of daily GCM rainfall for crop simulation studies. Agric. For. Meteorol. 138, 44–53.
Johnson, F., Sharma, A., 2011. Accounting for interannual variability: a comparison of options for water resources climate change impact assessments. Water Resour. Res. 47 (4), W04508. http://dx.doi.org/10.1029/2010WR009272.
Johnson, F., Sharma, A., 2012. A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. Water Resour. Res. 48 (1). http://dx.doi.org/10.1029/2011WR010464.
Jones, R.G., Noguer, M., Hassell, D.C., Hudson, D., Wilson, S.S., Jenkins, G.J., Mitchell, J., 2004. Generating High Resolution Climate Change Scenarios Using RECIS. Meteorological Office Hadley Centre, Exeter, 40 pp.
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteor. Soc. 77, 437–470.
Kanamaru, H., Kanamitsu, M., 2007. Fifty-seven-year reanalysis downscaling at 10 km (CaRD10). Part II: Comparison with North American Regional Reanalysis. J. Clim. 20, 5572–5592.
Kanamitsu, M., Evisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J.J., Fiorino, M., Potter, G.L., 2002. NCEP/DOE AMIP-II Reanalysis (R-2). Bull. Am. Meteor. Soc. 83, 1631–1643.
Karl, T., Trenberth, K., 2003. Modern global change. Science 302, 1719–1722.
Kiehl, J.T., Hack, J.J., Bonan, G.B., Boville, B.A., Briegleb, B.P., Williamson, D.L., Rasch, P.J., 1996. Description of the NCAR Community Climate Model (CCM3), NCAR/TN-420 STR, NCAR Technical Note.
Kleinn, J., Frei, C., Gurtz, J., Luthi, D., Vidale, P.L., Schar, C., 2005. Hydrologic simulations in the Rhine basin driven by a regional climate model. J. Geophys. Res. 110, D04102. http://dx.doi.org/10.1029/2004JD005143.
Lettenmaier, D.P., Wood, A.W., Palmer, R.N., Wood, E.F., Stakhiv, E.Z., 1999. Water resources implications of global warming: a U.S. regional perspective. Clim. Change 43, 537–579.
Li, H., Sheffield, J., Wood, E.F., 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. J. Geophys. Res. 115, D10101. http://dx.doi.org/10.1029/2009JD012882.
Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M., Thiele-Eich, I., 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. Rev. Geophys. 48, RG3003, http://dx.doi.org/10.1029/2009RG000314.
Maurer, E.P., Hidalgo, H.G., 2008. Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. Hydrol. Earth Syst. Sci. 12, 551–563.
Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., Cayan, D.R., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci. 14, 1125–1138.
McGregor, J.L., 1997. Regional climate modeling. Meteor. Atmos. Phys. 63, 105–117. http://dx.doi.org/10.1007/BF01025367.
Mearns, L.O., 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 on Climate Change.
Mesinger, F. et al., 2006. North American regional reanalysis. Bull. Am. Meteor. Soc. 87, 343–360.
Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J.C.J., Lang, H., Parmet, B.W.A.H., Schadler, B., Schulla, J., Wilke, K., 2001. Impact of climate change on hydrological regimes and water resources management in the Rhine basin. Clim. Change 49, 105–128.

Minville, M., Brissette, F., Krau, S., Leconte, R., 2009. Adaptation to climate change in the management of a Canadian water-resources system exploited for hydropower. Water Resour. Manage 23, 2965–2986.
Murphy, A.J., 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. J. Clim. 12, 2256–2284.
Nigam, S., Ruiz-Barradas, A., 2006. Seasonal hydroclimate variability over North America in global and regional reanalyses and AMIP simulations: varied representation. J. Clim. 19, 815–837.
Pan, H.-L., Wu, W.-S., 1994. Implementing a mass-flux convective parameterization package for the NMC medium range forecast model. In: Preprints, 10th Conference on Numerical Weather Prediction, Portland, OR, American Meteorological Society, pp. 96–98.
Panofsky, H.A., Brier, G.W., 1968. Some Applications of Statistics to Meteorology. The Pennsylvania State University, University park, PA, USA, 224 pp.
Piani, C., Haerter, J.O., Coppola, E., 2010a. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor. Appl. Climatol. 99, 187–192. http://dx.doi.org/10.1007/s00704-009-0134-9.
Piani, C., Weedon, G.P., Best, M., Gomes, S.M., Viterbo, P., Hagemann, S., Haerter, J.O., 2010b. Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. J. Hydrol. 395, 199–215. http://dx.doi.org/10.1016/j.jhydrol.2010.10.024.
Rokicki, R., 2002. Evaluation and performance of rainfall disaggregation methods for West-Central Florida. M.S. Thesis. University of South Florida.
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 Resourc. Assoc. (JAWRA) 41, 1013–1025.
Sato, T., Kimura, F., Kitoh, A., 2007. Projection of global warming onto regional precipitation over Mongolia using a regional climate model. J. Hydrol. 333, 144–154.
Seneviratne, S.I., Pal, J.S., Eltahir, E.A.B., Schär, C., 2002. Summer dryness in a warmer climate: a process study with a regional climate model. Clim. Dyn. 20, 69–85.
Stefanova, L., Misra, V., Chan, S., Griffin, M., O'Brien, J.J., Smith III, T.J., 2012. 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. http://dx.doi.org/10.1007/s00382-011-1230-y.
Teutschbein, C., Seibert, J., 2010. Regional climate models for hydrological impact studies at the catchment scale: a review of recent model strategies. Geogr. Compass 4 (7), 834–860. http://dx.doi.org/10.1111/j.1749-8198.2010.00357.x.
Teutschbein, C., Seibert, J., 2012a. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456–457, 12–29. http://dx.doi.org/10.1016/j.jhydrol.2012.05.052.
Teutschbein, C., Seibert, J., 2012b. Is bias correction of Regional Climate Model (RCM) simulations possible for non-stationary conditions? Hydrol. Earth Syst. Sci. Discuss. 9 (11), 12765–12795. http://dx.doi.org/10.5194/hessd-9-12765-2012.
Teutschbein, C., Wetterhall, F., Seibert, J., 2011. Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. Clim. Dynam. 37 (9–10), 2087–2105. http://dx.doi.org/10.1007/s00382-010-0979-8.
Trenberth, K.E., Stepaniak, D.P., Hurrell, J.W., Fiorino, M., 2001. Quality of reanalyses in the tropics. J. Clim. 14, 1499–1510.
Uppala, S.M. et al., 2005. The ERA-40 re-analysis. Quart. J. Roy. Meteorol. Soc. 131, 2961–3012.
Wang, Y., Leung, L.R., McGregor, J.L., Lee, D.K., Wang, W.C., Ding, Y., Kimura, F., 2004. Regional climate modeling: progress, challenges, and prospects. J. Meteor. Soc. Jpn. 82, 1599–1628.
Widmann, M., Bretherton, C.S., 2000. Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. J. Clim. 13, 1936–1950.
Wilby, R.L., Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. Prog. Phys. Geogr. 21, 530–548.
Wilby, R.L., Hay, L.E., Gutowski Jr., W.J., Arritt, R.W., Takle, E.S., Pan, Z., Leavesley, G.H., Clark, M.P., 2000. Hydrological responses to dynamically and statistically downscaled climate model output. Geophys. Res. Lett. 27, 1199–1202.
Wood, A.W., Maurer, E.P., Kumar, A., Letternmaier, D.P., 2002. Long-range experimental hydrologic forecasting for the eastern United States. J. Geophys. Res. 107, 4429. http://dx.doi.org/10.1029/2001JD000659.
Wood, A.W., Leung, L.R., Sridhar, V., Lettenmaier, D.P., 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Clim. Change 62, 189–216.