Environmental Modelling & Software 74 (2015) 48–65

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Setting up a hydrological model of Alberta: Data discrimination analyses prior to calibration

Monireh Faramarzi [a, *], Raghavan Srinivasan [b], Majid Iravani [a, c], Kevin D. Bladon [a, d], Karim C. Abbaspour [e], Alexander J.B. Zehnder [f, g], Greg G. Goss [a, h]

[a] Water Initiative, University of Alberta, Edmonton, Alberta, Canada
[b] Spatial Science Laboratory, Department of Ecosystem Science and Management, Texas A&M University, College Station, TX, USA
[c] Alberta Biodiversity Monitoring Institute, University of Alberta, Edmonton, Alberta, Canada
[d] Department of Forest Engineering, Resources, and Management, Oregon State University, Corvallis, OR, 97331, USA
[e] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland
[f] Alberta Innovates-Energy and Environment Solutions, Edmonton, Alberta, Canada
[g] Nanyang Technological University (NTU), Sustainable Earth Office, Singapore, 637459, Singapore
[h] Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

ARTICLE INFO

Article history:
Received 10 April 2015
Received in revised form
1 September 2015
Accepted 3 September 2015
Available online xxx

Keywords:
SWAT
Watershed modeling
Climate data
Hydrological processes
Geo-spatial data

ABSTRACT

Failure to setup a large-scale hydrological model correctly may not allow proper calibration and uncertainty analyses, leading to inaccurate model prediction. To build a model with accurate accounting of hydrological processes, a data discrimination procedure was applied in this study. The framework uses a hydrological model of Alberta built with the Soil and Water Assessment Tool (SWAT) program. The model was used to quantify the causes and extents of biases in predictions due to different types of input data. Data types represented different sources of errors, including input data (e.g., climate), conceptual model (e.g., potholes, glaciers), and control structure (e.g., reservoirs, dams). The results showed that accounting for these measures leads to a better physical accounting of hydrological processes, significantly improving the overall model performance. The procedure used in this study helps to avoid unnecessary and arbitrary adjustment of parameters to compensate for the errors in the model structure.

Software availability

SWAT program is available for use at the following URL http://swat.tamu.edu/

1. Introduction

Physically-based, distributed hydrological models have been widely used for water resources management and planning. They have been extensively applied to study the impact of climate change and landuse change on water quality and quantity, water related activities, and adaptation measures among others (Li et al., 2009; Faramarzi et al., 2010a, 2010b, Van Griensven et al., 2012; Faramarzi et al., 2013; Eum et al., 2014; Xue et al., 2014). The reliability of such applications depends on the accuracy of hydrological models in representing the physical processes (Beven, 2000; Muleta and Nicklow, 2005), correct input data, and proper model calibration. As such, a key challenge is initially to set up an accurate hydrological model, which correctly represents the site's actual physical processes (Gupta and Sorooshian, 1998; Perrin et al., 2001; Blasone et al., 2008; Moradkhani et al., 2012; Houska et al., 2014; Guse et al., 2014; Gabriel et al., 2014).

Calibration of distributed models is often difficult and subjective when there is a considerable simplification in model setup. It is standard practice in watershed modeling studies that the physical parameters are adjusted to achieve the optimal fit to the measured data. However, simplification of the models, especially in large scale watersheds (where a considerable heterogeneity exist in climate, vegetation, soil, physiography, and management activities), might result in a wrong parameter estimation (Schuol et al., 2008b; Faramarzi et al., 2009). In large scale models where a vast number of adjustable physical-parameters are allowed to vary within a broad range of values, a seemingly good simulation can be obtained with erroneous parameter values (Abbaspour et al., 2007). In other words, wrong model structure and inappropriate input data can be compensated by unrealistic model parameters. Such models could

* Corresponding author. University of Alberta Water Initiative, G-222 Biological Sciences Building, 11455 Saskatchewan Dr., Edmonton, Alberta, T6G 2E9, Canada.
E-mail address: faramarz@ualberta.ca (M. Faramarzi).

http://dx.doi.org/10.1016/j.envsoft.2015.09.006
1364-8152/

produce misleading results in scenario analyses, even though typical performance criteria are satisfied during calibration. One way to detect these over calibration problems is by validation of the model for a reasonable time period where major hydrological events (e.g., wet years, dry years) are presented.

A correct model setup, accurately representing the actual hydrological processes, can limit uncertainty in parameter estimation. In literature, to limit uncertainties in parameter estimation, various measures through automated calibration techniques have been examined. These include multi-variable calibration procedure (Gupta and Sorooshian, 1998; Xie et al., 2012; Qiao et al., 2013; Samuel et al., 2014), use of multiple calibration sites rather than only catchment integrated behavior (Abbaspour et al., 1999, 2007; Cao et al., 2006; Schuol et al., 2008a, 2008b), a multi-objective formulation by including different variables in the objective function (Gupta and Sorooshian, 1998; Madsen, 2003; White and Chaubey, 2005), and use of various techniques to increase the computational efficiency of the large scale hydrological models (Wu et al., 2013; Ercan et al., 2014). Although the schemes are beneficial in limiting uncertainties in the predictions, a more reliable result can be achieved through building an accurate model. Building a correct model, especially in large scale and complex watersheds, is an important practice to represent correct processes inside a watershed. A correct model is one that adheres to the principle of "correct neglect", where only unimportant processes are neglected in the model and all important processes should be included. Therefore, it is inevitable that large scale models should go through careful data discrimination scheme to ensure most of the important processes are represented prior to calibration. These include: (i) gathering and compiling appropriate input data (e.g., climate data in mountainous regions); (ii) including management control structures that can disrupt natural processes (e.g., dams that regulate downstream water flow); and (iii) incorporating local knowledge about the natural complexity and anthropogenic changes into watershed models. These are all key factors that can reduce the uncertainty in model predictions and avoid unnecessary and arbitrary adjustment of the parameters.

Overall, the majority of researchers have focused on elaboration of the importance of robust calibration schemes in parameter estimation (e.g., Joseph and Guillaume, 2013) and prediction uncertainty, while much fewer studies have addressed proper model setup and choice of appropriate input datasets. Later group are those that focused on modifying the existing climate datasets to better represent the effect of altitude on precipitation (Masih et al., 2011; Galvan et al., 2014) and those that examined the effect of input data quality and quantity on parameter estimation and model calibration (Getirana et al., 2011; Strauch et al., 2012; Yalew et al., 2013; Gabriel et al., 2014; Rouholahnejad et al., 2014; Yen et al., 2014; Abbaspour et al., 2015; Leta et al., 2015).

With an area of about 660,000 km^2, Alberta encompasses 17 river basins that principally originate from the east slopes of the Canadian Rocky Mountains and the majority drain east to Hudson Bay through the provinces of Saskatchewan and Manitoba and north to the Arctic Ocean. The heterogeneous hydro-climatic conditions and the diverse land management practices in combination with the scarcity of data, especially in the northern remote areas and western mountainous region, make hydrological modeling challenging in this region. To the best of our knowledge a high resolution and province-wide hydrological model has not been developed for Alberta. Most of the previous studies in Alberta have been conducted at a catchment (e.g., Kienzle et al., 2012; Marshall, 2014) or river basin (e.g., Islam and Gan, 2014; Eum et al., 2014) scale.

The model of choice for this project was "Soil and Water Assessment Tool" (SWAT) (Arnold et al., 1998). SWAT has been developed to quantify the impact of land management practices and climate on water, sediment, and agricultural chemical yields in large complex watersheds with varying soils, landuses, and management conditions over long periods of time. The program, therefore, lends itself easily to climate and landuse change analyses. SWAT is a valuable watershed-scale management tool and we chose this program for our purposes because: i) it integrates many components such as hydrology, climate, nutrient, soil, sediment, crop, pesticide, and agricultural management, ii) it has been successfully applied worldwide in many different climate and landuse situations (Arnold et al., 1999; Gosain et al., 2006; Schuol et al., 2008a,b; Rouholahnejad et al., 2014; Abbaspour et al., 2009, 2015), iii) the program is actively maintained and continuously updated with new and up-to-date knowledge of watershed processes, and iv) many side programs are written for SWAT from calibration and uncertainty analysis to graphic packages for visualization and animation of the results. Hence, over a 50-year period, a global consensus is built around the accuracy and usefulness of the program as there exist over 3000 scientific publications where SWAT has been used to address numerous watershed issues (Gassman et al., 2007, 2010).

We used the SWAT hydrological model of Alberta as an example to demonstrate that proper model setup could produce more accurate model outputs and represent most of the natural and anthropogenic processes. However, one hypothesis would be how a model with a better performance would guarantee that it will be actually the best option after calibration. We address in this paper the fact that building a correct model is a key step prior to calibration to avoid compensation through subjective and challenging parameter estimation and this will provide the best performance model.

Objectives of this paper are: (i) to build various SWAT projects to test the effects of including alternative climate and geo-spatial datasets available from global and regional sources; (ii) to evaluate the performance of the model predictions using combination of multiple datasets from different sources, (iii) to define the procedures by which raw datasets are evaluated for inclusion or exclusion in the model; and (iv) to calibrate and validate all of the model scenarios for the Athabasca River basin as an example hydrological region, thereby allow us to test how an accurate model will perform best after calibration. It is important to point out that the above SWAT models are tested against each other prior to calibration, as over calibration and over fitting of model parameters would mask the input data and model structure effects and will not allow a proper discrimination of initial model setups (Dile and Srinivasan, 2014; Abbaspour et al., 2015).

2. Materials and methods

2.1. Study area

Alberta, with an area of about 660,000 km^2, is located between 49–60 °N and 110–120 °W where altitude varies from 3747 m (Mount Columbia) to 152 m (Slave River-Wood Buffalo National Park) (Fig. 1a). Geographically, the province spans >1200 km from north to south and large-scale climate anomalies, originating from Pacific Ocean, have a considerable influence on climate diversity (Lapp et al., 2013). Air temperatures can drop to as low as −54 °C during the winter (northern Alberta), and rise to as high as 40 °C during the summer (southern Alberta). Average annual precipitation ranges from 300 mm in the southeast to 600 mm in the foothills of the Rocky Mountains (AENV-GA, 2008; Mwale et al., 2009).

The province has 17 river basins (Fig. 1a; AENV-GA, 2008) with the northern rivers of the province having comparatively larger areas and therefore higher discharge rates than the southern rivers that flow through regions that receive much lower annual precipitation. For instance, the average flow of Peace River in the north is

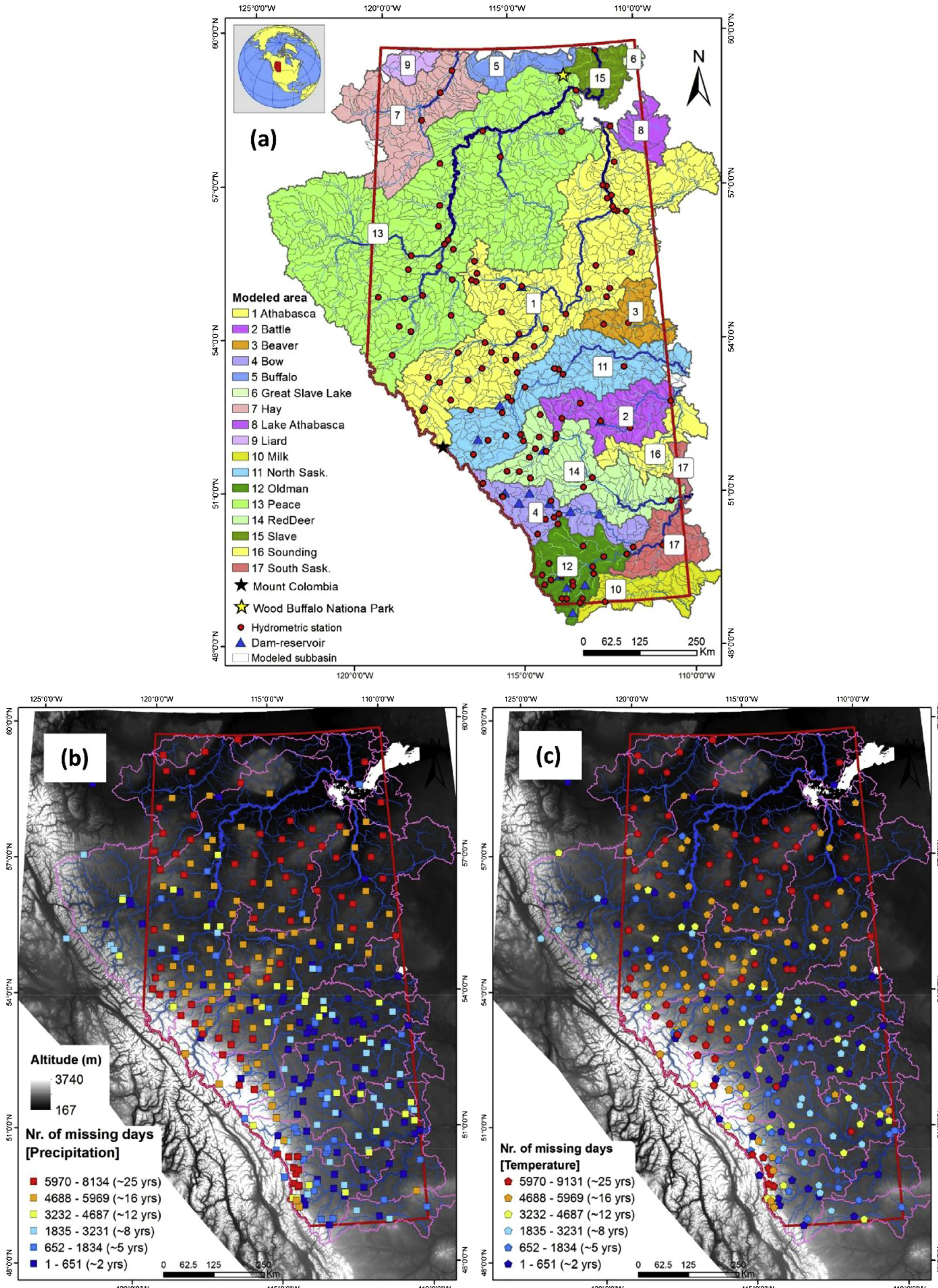

Fig. 1. Map of study area presenting geographic distribution of the main river basins, hydrometric stations, dams-reservoirs and the modeled sub-basins (a); and distribution of the meteorological stations in different river basins of Alberta (b, c). Different colors show the number of missing daily data during 1983–2007 in each station. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2161 $m^3 s^{-1}$ and the peak flow can exceed 5000 $m^3 s^{-1}$ at the outlet of Peace-Slave river basin, whereas the peak flow at the outlet of South Saskatchewan river basin in the south can reach 1200 $m^3 s^{-1}$ (see Supplementary Table A.1) (AENV-GA, 2006; AENV-GA, 2008). Similarly, landuse in each region varies considerably (see Supplementary Table A.2). Importantly, a large portion of the prairie landscape in the south and eastern parts of the province have a drainage network that is poorly developed resulting in many closed depressional areas (potholes). In addition, individual farmers are allowed to capture rainwater and snowmelt (sloughs) for on-farm storage resulting in undocumented water impoundments. Together, these natural and anthropogenic formations influence the contribution of precipitation to streamflows as they prohibit the drainage to the receiving stream. The Prairie Farm and Rehabilitation Administration (PFRA, 2012) of Agriculture and Agri-Food Canada (AAFC) has delineated and characterized these areas which are so called "non-contributing areas". In southern Alberta, the landuse is primarily medium- and large-scale agriculture; however, there is not enough rainfall and moisture to naturally sustain demands of agricultural crops in much of the region. As such, substantial dams, diversion channels, off-stream reservoirs, and irrigation systems have been constructed. Thirteen organized irrigation districts receive large quantities of water which are diverted from the tributaries of the South Saskatchewan River, primarily the Oldman (St. Mary, Waterton and Belly) and Bow Rivers (Sauchyn et al., 2011; AARD, 2013).

2.2. SWAT hydrologic model

We used the SWAT2012 model to simulate streamflow. SWAT model is a process-based, spatially distributed model that operates on a daily time step (Arnold et al., 1998). Spatial parameterization of the SWAT model is performed by delineating a watershed into sub-basins based on topography and into Hydrologic Response Units (HRUs) according to soil, landuse, and slope characteristics. SWAT simulates the watershed hydrology in two phases: 1) the land phase of the hydrologic cycle, which calculates the water balance of each HRU at a given time step, and 2) the routing phase, which routes the water through river network towards the basin outlet. The model uses daily climate data, such as precipitation, minimum and maximum temperatures. It assigns the nearest weather station to the centroid of each sub-basin to that sub-basin. It simulates streamflow, soil water, ground water recharge, potential and actual evapotranspiration, plant water uptake, transpiration, soil and canopy evaporation, and other hydrological components daily. A mass balance equation is used in SWAT to account for the snow hydrology based on whether the equivalent water content of the snowpack increases with more snowfall or decreases with snowmelt and sublimation. A weather generator module is accommodated to generate daily climate data or to fill in the gaps in measured records. Impoundments play an important role in water balance of a sub-basin. Four types of water bodies are simulated in SWAT: ponds, wetlands, potholes, and reservoirs. Water flows from sub-basin into these water bodies. A water balance equation is solved to initiate water impoundment which is a function of total inflow (e.g., runoff entering from the sub-basin, rainfall, ground water contribution) and total outflow from the water bodies (e.g., evaporation, seepage). Reservoirs are located on the main channels and receive water from all upstream sub-basins. The magnitude of water outflow from the reservoirs (dams) is defined by user. A more detailed description of the model is given by Neitsch et al. (2011).

2.3. Data and model setup

There are different sources of error in hydrological modeling. The most important sources are input data (e.g., climate data or spatial data), conceptual model (e.g., process simplifications), and the anthropogenic changes through management practices (Abbaspour et al., 2007). The first source can be accounted for by using the most relevant datasets through initial testing. The second and third sources cannot be quantified unless the modeler develops a clear understanding of the region of interest and the most important processes occurring across the region. While using multiple models of different complexities can help to identify key processes, this approach is costly and time consuming, especially for detailed large-scale studies. Instead, our approach was to use different SWAT model structures, in combination with local expert knowledge and testing of different data sources. This was an efficient way to model the study area and related processes.

Various data types, representing different sources of error, were used in this study to qualify their hydrological responses through simulation of the streamflow (Table 1). The results were compared with historical measured records at 130 hydrometric stations. A total of ten SWAT projects (S_1–S_{10}), corresponding to ten different datasets were constructed and the simulated monthly river discharges were compared with that of measured records obtained from the Environment Canada (http://www.ec.gc.ca/rhc-wsc/) for the period of 1986–2007. Here we give a brief description of the SWAT projects, data, and scenarios:

To build the hydrological model of Alberta (S1), we used the digital elevation model (DEM) at 90 m resolution (SRTM, Jarvis et al., 2008) for sub-basin delineation. A detailed stream network was initially delineated using a 10-m resolution DEM (AltaLIS, http://www.altalis.com/) and used for watershed delineation. Using a threshold drainage area of about 200 km^2 a total of 2255 sub-basins were delineated for the study area (Fig. 1a). With this threshold we made a balance between the resolution of the available data and the practical SWAT project size. The landuse map was obtained from the GeoBase Land Cover Product. (http://www.geobase.ca/geobase/en/data/landcover/csc2000v/description.html), which has a resolution of 30 m and distinguishes 36 landuse classes for Canada and 23 classes for our study area (see Supplementary Table A.2). The raw soils data was obtained from the Agriculture Agri-Food Canada, Soil Landscapes of Canada V3.2 (http://sis.agr.gc.ca/cansis/nsdb/slc/index.html) and modified to meet the input requirements of SWAT. This map represents a total of 364 soil types for Canada and 90 soil types for our study area, including the SWAT required physical parameters at a maximum of nine soil layers. A dominant soil, landuse, and slope were considered to characterize each sub-basin in this study. Provincial climate data from about 300 meteorological stations (MS) (Fig. 1b,c) were acquired from Environment Canada at a daily time step for our study period (1983–2007).

To elaborate on the climate data uncertainty, four additional SWAT projects (S2–S5) were built using gridded climate datasets. The gridded climate data were from the four widely utilized sources including that of the National Centers for Environmental Prediction's Climate Forecast System Reanalysis (CFSR), which provides daily climate data at a 0.3-degree grid resolution with a global coverage; the CRU TS2 (named CRU1 in this study) and CRU TS3.21 (named CRU2 in this study) from the Climate Research Unite (CRU), which both provide monthly climate data at 0.5-degree grid resolution with a global coverage; and the Natural Resources Canada (NRCan), which supplies daily gridded climate data at 10 km × 10 km resolution for Canada (see Table 1 for more specification).

To test the effect of soil and landuse data we built two other projects (S6 and S7). For these projects we replaced the landuse map used in S1 with the global landuse map of the US Geological Survey (USGS) (Table 1, S6) and the global soil map of the Food and

Table 1
Global and regional data sources used in this study.

Error sources		Scenarios/dataset	Time span	Resolution	Time step	Region	Nr. of stations/grids in study area	Reference
Input data	Climate	S1: Metrological stations	1983–2007	–	Daily	Local	320	Government of Canada, Path: http://climate.weather.gc.ca/
		S2: CFSR	1979–2010	0.3° grid	Daily	Global	1097	Fuka et al., 2013; Path: http://globalweather.tamu.edu
		S3: CRU1[a]	1900–2000[c]	0.5° grid	Monthly	Global	771	New et al., 2000; Mitchell and Jones 2005
		S4: CRU2[b]	1900–2012	0.5° grid	Monthly	Global	771	Harris et al., 2014
		S5: NRCAN	1910–2010	10 km × 10 km	Daily	Regional/Canadian	7543	McKenney et al., 2011
	Digital maps	S6: USGS Landuse/land cover map	1993	1 km × 1 km (1:1,000,000)	–	Global	100% coverage	USGS Global Land Use Land Cover Characterization (GLCC) database with a spatial resolution of 1 km and distinguishing 24 landuse/land cover classes. Path: http://edcsns17.cr.usgs.gov/glcc/glcc.html
		S7: FAO-Soil map	2005	10 km × 10 km (1:10,000,000)	–	Global	100% coverage	Food and Agriculture Organization of the United Nations (FAO, 1995), which provides data for 5000 soil types comprising two layers (0–30 cm and 30–100 cm depth) at a spatial resolution of 10 km.
Model conceptual		S8: Potholes	2012	Watershed (delineated for each hydrometric station)	–	Regional/Canadian	100% coverage	Prairie Farm and Rehabilitation Administration (PFRA), Agriculture Agri-food Canada (AAFC), 2012.
		S9: Glaciers	1985–2005	River Basin	Long-term monthly	Regional/Global	100% coverage	Raup et al., 2007; Marshall, 2014.
Management measures		S10: Reservoir/lake	Since compilation	–	Daily	Local	15 main reservoirs-lakes	AESRD, Alberta Environment Sustainable Resources Development: measured data at hydrometric stations. See Supplementary Table A.4.
Multiple dataset		S11: S1 + S10 + CFSR temp. replaced	1983–2007	–	Daily–monthly	Local-regional	–	–
		S12: S11 + FAO soil was replaced	1983–2007	–	Daily–monthly	Local-regional	–	–
		S13: S11 + S8 + S9	1983–2007	–	Daily–monthly	Local-regional	–	–

[a] CRU TS2.
[b] CRU TS3.21.
[c] SWAT weather generator (Neitsch et al., 2011) was used to fill the gaps between 2001 and 2007.

Agricultural Organization (FAO, 1995) of the United Nations (S7).

Additionally, three more SWAT projects (S8, S9, and S10) were built to simulate non-contributing areas and potholes (S8), to incorporate glaciers (S9), and to include regulatory dams-reservoirs and lakes (S10) (dataset information available in Table 1). To acquire this information, we held several expert meetings with relevant governmental organizations to discuss the factors that may alter our hydrological assessment and to understand the unique conditions of each river basin in Alberta.

In this study, we used the Hargreaves method to calculate the potential evapotranspiration; the Soil Conservation Service's curve number method to estimate surface runoff; and the variable storage routing method for the simulation of the channel processes. The aim of this study was to assess the performance of different input datasets in hydrological modeling prior to calibration. To compare the gridded climate datasets with the observed records we computed the following statistics for each sub-basin, river basin, and region accounting for the seasonal variation in the statistics during 1983–2007. The statistics used in this study were: linear correlation coefficient (CC), such that $-1 \leq CC \leq 1$ (unitless), mean absolute error (MAE), which ranges from 0 to ∞ with lower values indicating greater accuracy (mm and °C), and percent bias (PBias) with the optimal value of 0 (unitless):

$$CC = \frac{\sum_{i=1}^{n}[(x_i-\overline{x})(y_i-\overline{y})]}{\sqrt{\sum_{i=1}^{n}\left[(x_i-\overline{x})^2\right]\sum_{i=1}^{n}\left[(y_i-\overline{y})^2\right]}} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(x_i-y_i)| \tag{2}$$

$$PBias = 100 \times \frac{\sum_{i=1}^{n}(x_i-y_i)}{\sum_{i=1}^{n}y_i} \tag{3}$$

where x is the gridded dataset and y is the meteorological gauge dataset.

2.4. Model evaluation

The streamflow was simulated for the 1983–2007 period considering a three-year warm-up period (excluded from the analysis) to equilibrate the simulated physical processes to mitigate the unknown initial conditions. We compared the measured and simulated monthly discharges at 130 hydrometric stations (Fig. 1a) using a modified version of the efficiency criterion defined by Krause et al. (2005):

$$\phi = \left\{|b| \;\; R^2 \quad if \;\; |b| \leq 1; \;\; |b|^{-1} \;\; R^2 \quad if \;\; |b| > 1\right\}, \tag{4}$$

where, R^2 is the coefficient of determination between measured and simulated signals and b is the slope of the regression line. In large scale multi-site hydrological studies, where the stations' efficiency criteria are averaged for the watersheds, the bR^2 (ranging from 0 to 1) is widely used as a more efficient index compared to other criteria such as Nash-Sutcliffe Efficiency (*NSE*) or R^2. The *NSE* is a normalized statistic that determines the relative magnitude of the residual variance compared to the measured data variance and varies between -∞ and 1 (Nash and Sutcliffe, 1970). The *NSE* may be dominated by a few poorly simulated stations (with large negative values). A modified version of *NSE* has been recommended by Mathevet et al., 2006 in large scale studies, which varies between −1 and +1 and generates less skewed distribution. Similarly, the R^2 statistics represents the trend of the simulated results, but not the closeness to the measured data. In our simulations, we also calculated the bounded *NSE* (*BNSE*) using Eq. (5) (Mathevet et al., 2006) and used this criterion as an additional information to evaluate our province-wide model performance.

$$BNSE = \frac{NSE}{2-NSE} \tag{5}$$

In addition, we showed the average R^2 and *NSE* for the whole study area, as these are commonly used criteria for hydrological studies.

To test the effect of parameter adjustment (i.e., calibration) on scenario selection, we performed calibration only for the Athabasca River basin. The Athabasca River is the second largest river basin in Alberta (see Supplementary Table A.1,2 and Fig. 1). It originates from the glaciers of Rocky Mountains in Jasper National Park, has a drainage area of about 133,000 km^2, and flows for over 1230 km from the head waters to join Lake Athabasca in the east. The mean annual discharge rate of the basin is about 661 m^3 s^{-1}. The river flow regime is contingent on the seasonality of climate, reaching its minimum in winters and its maximum in warm summers, when snow and glacial melt waters from the river's head waters combine with runoff from localized snowmelt and rainfall throughout the basin. Furthermore, Lesser Slave Lake (LSL) significantly alters the hydrological regime of the downstream on the river while non-contributing areas (sloughs for on-farm storage) in upstream LSL serves as buffers for water flow in the region. The Athabasca River basin is a reasonably good choice for calibration since it represents most of the scenario attributes defined in this study (i.e., S1–S13, see Table 1).

We calibrated all of the model scenarios using the monthly data of 40 hydrometric stations in the basin. For calibration we ran a sensitivity analysis using the Sequential Uncertainty Fitting program (SUFI2) (Abbaspour et al., 2007; Faramarzi et al., 2009) to find the most sensitive parameters to river discharges. We found 22 parameters were generally sensitive to river discharges (Supplementary Table A.3). The parameters were further differentiated based on soil and landuse types to better represent the geospatial and hydrological characteristics (see Faramarzi et al., 2009). The parameterization was further regionalized for highlands, middle regions and lowlands in the basin, resulting in a total of 300 parameters. To perform parameter updates we provided a range for each parameter from which 500 Latin Hypercube samples were drawn and fed into the model for simulation. The parameter ranges were limited to a physically meaningful range (Neitsch et al., 2011; Abbaspour et al., 2007) to prevent over calibration of the models. We therefore performed 500 model runs for each model scenario and calculated the bR^2 for each individual station. The best simulation was found using the best parameter set sample which produced the largest bR^2 in the river basin.

3. Results and discussion

3.1. Climate data

The observed meteorological station (MS) climate data collected across the province varied widely in the number of missing data. Specifically, most MS stations in the northern regions and those located in the western mountainous regions of the study area had a large number of missing days (Fig. 1b,c). Therefore, we used the weather generator of SWAT to fill in the gaps using the closest data-rich stations and generated daily time series for the sub-basins where climate stations were sparse. This resulted in precipitation estimates ranging from 266 to 400 mm yr^{-1} in the southern and northern regions to 600–814 mm yr^{-1} in western mountainous

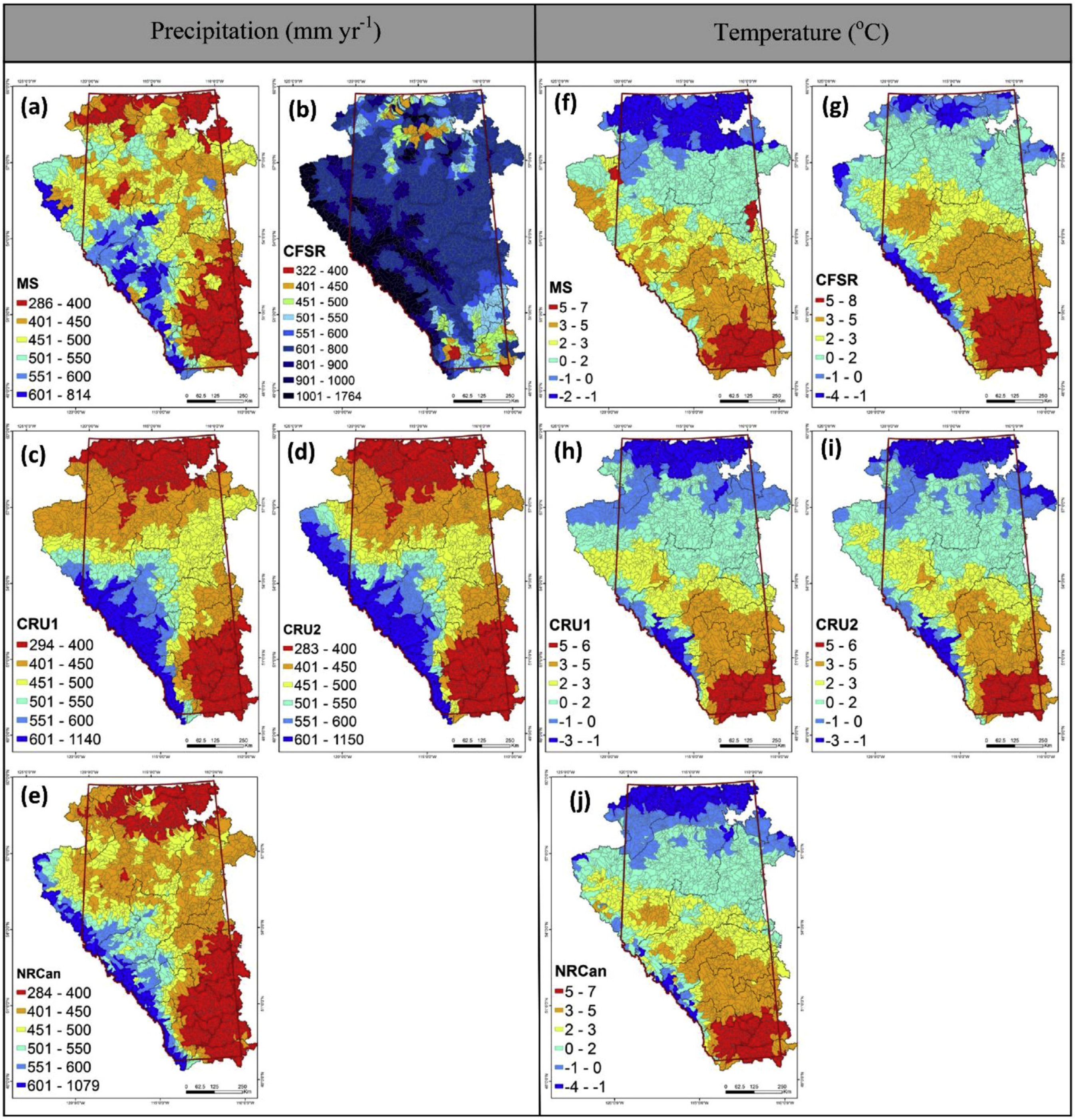

Fig. 2. Spatial distribution of the 25year (1983–2007) average precipitation (mm; left) and temperature (°C; right) across study area from the observed MS data and the four gridded climate datasets.

regions (Fig. 2a). Moreover, the 25-year mean temperature estimation using MS data varied from −2 °C to −1 °C in the northern parts of the province to a high of 5 °C–8 °C in the southern regions (Fig. 2f). The mean annual temperatures in the Rocky Mountains were above 0 °C using MS dataset, which are higher than the reported range of −7 °C to 0 °C by AENV-GA (2008).

Simulations using 4 other gridded datasets, produced differences in the spatial pattern of climate data, which also differed from the MS data results. The mean annual precipitation of the CFSR dataset (Fig. 2b) was significantly larger than observed MS data (Fig. 2a), ranging from 500 to 1764 mm yr^{-1} in most parts of the study area. Other gridded data (CRU1-Fig. 2c; CRU2-Fig. 2d; NRCan-Fig. 2e) produced similar spatial patterns, with the greatest precipitation occurring in the western mountainous areas (600–1100 mm yr^{-1}) and the lowest precipitation occurring in the southern and northern plains (280–400 mm yr^{-1}).

The long-term mean annual temperatures were within −4 °C to 5 °C for all datasets (Fig. 2f–j). While the spatial variation of the temperatures were generally consistent in all datasets, our statistical analysis showed temporal differences within and between the datasets (Table 2a,b). To better understand the statistical performances of gridded climate data relative to the observed MS, we also

Table 2
Seasonal statistics of the gridded compared to the observed MS data for precipitation (a) and temperature (b) during 1983–2007.

Statistic	Winter				Spring				Summer				Fall				Year			
	CFSR	CRU1	CRU2	NRCan	CFSR	CRU1	CRU2	NRCan	CFSR	CRU1	CRU2	NRCan	CFSR	CRU1	CRU2	NRCan	CFSR	CRU1	CRU2	NRCan
(a)																				
North: Athabasca, Peace/Slave, Hay, Buffalo, Lake Athabasca, Liard, Great Slave Lake																				
CC	0.09	0.13	0.14	0.17	0.45	0.48	0.48	0.69	0.49	0.28	0.27	0.63	0.17	0.13	0.12	0.15	0.38	0.28	0.28	0.45
MAE	72.72	18.33	18.6	19.79	72.49	33.76	33.68	27.16	58.83	43.57	40.93	32.47	65.74	24.98	25.1	24.99	67.45	30.16	29.58	26.10
PBias	151.7	1.51	5.81	18.78	55.9	−6.67	−5.39	−12.3	16.96	0.17	−1.32	−2.86	96.17	3.4	5.66	9.52	56.13	−1.74	−1.03	−1.36
Middle: Beaver, North Saskatchewan, Battle, Red Deer, Sounding																				
CC	0.35	0.36	0.34	0.6	0.56	0.48	0.45	0.8	0.52	0.34	0.35	0.76	0.53	0.56	0.53	0.7	0.52	0.37	0.38	0.77
MAE	67.12	20.39	20.11	14.77	100.9	40.54	41.64	25.33	55.54	45.06	46.39	28.42	56.77	19.42	19.88	14.47	70.07	31.35	32.01	20.75
PBias	143.3	14.19	10.91	1.97	63.71	8.8	6.1	−5.2	10.79	0.48	−0.41	−5.42	112.2	5.62	6.75	−0.42	55.73	5.07	3.57	−4.27
South: Bow, Oldman, South Saskatchewan, Milk																				
CC	0.42	0.46	0.32	0.65	0.63	0.55	0.56	0.77	0.7	0.68	0.72	0.82	0.56	0.58	0.58	0.74	0.63	0.57	0.56	0.78
MAE	50.96	26.35	26.67	18.68	73.14	48.4	48.8	34.29	49.02	39.87	37.48	27.8	44.76	25.71	25.31	19.17	54.47	35.08	34.57	24.99
PBias	99.57	8.37	1.09	6.02	33.8	0.54	−1.44	−1.54	−2.7	−0.97	−4.54	−5.05	80.71	4.83	5.07	4.88	35.46	1.45	−1.52	−1.02
(b)																				
North: Athabasca, Peace/Slave, Hay, Buffalo, Lake Athabasca, Liard, Great Slave Lake																				
CC	0.11	0.18	0.18	0.13	0.46	0.46	0.43	0.46	0.44	0.5	0.44	0.5	0.04	0.01	0.05	0.06	0.27	0.29	0.30	0.28
MAE	2.39	2.26	2.19	2.3	1.62	1.43	1.49	1.55	1.28	0.96	1.04	0.96	2.29	2.27	2.25	2.27	1.90	1.73	1.74	1.77
PBias	−1.58	1.53	1.14	0.27	−9.13	−6.82	−7.23	−9.22	−1.78	−1.92	−2.59	−0.82	−4.6	7.27	25.33	20.65	−21.31	−1.83	11.88	19.40
Middle: Beaver, North Saskatchewan, Battle, Red Deer, Sounding																				
CC	0.66	0.66	0.64	0.68	0.69	0.67	0.64	0.68	0.77	0.73	0.74	0.75	0.69	0.68	0.66	0.69	0.63	0.63	0.60	0.64
MAE	1.25	1.72	1.77	1.49	1.69	1.84	1.93	1.75	0.84	0.94	0.98	0.81	2.72	2.88	2.93	2.74	1.63	1.85	1.90	1.70
PBias	−5.82	13.12	14.6	10	−2.95	−5.12	−5.55	−3.04	6.23	0.97	−0.02	2.02	−10.4	−0.2	5.32	−5.63	−10.56	7.58	10.98	9.46
South: Bow, Oldman, South Saskatchewan, Milk																				
CC	0.82	0.79	0.78	0.83	0.8	0.78	0.73	0.79	0.84	0.83	0.83	0.8	0.74	0.72	0.71	0.73	0.81	0.78	0.74	0.79
MAE	1.15	1.48	1.6	1.16	1.18	1.28	1.44	1.25	0.98	1.03	1.13	0.88	1.6	1.12	1.43	1.21	1.23	1.23	1.40	1.13
PBias	−6.9	20.05	24.89	10.22	−4.99	−11.6	−13.7	−9.87	9.83	−1.04	−2.74	0.69	53.15	57.78	−72.3	−21.1	20.61	−39.92	−54.74	−20.75

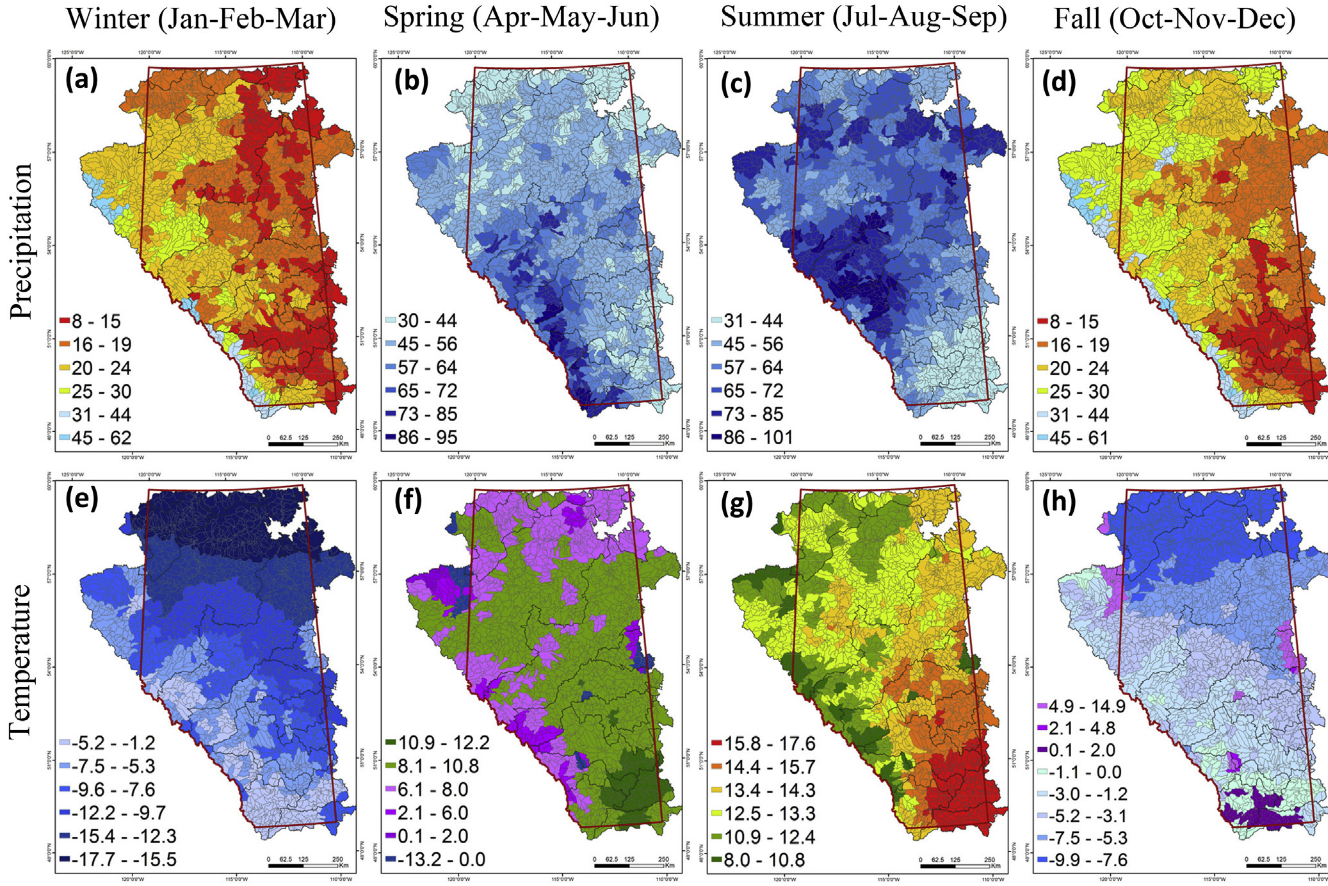

Fig. 3. Average monthly precipitation and temperature by seasons during 1983–2007 using the observed daily MS data.

illustrated seasonal variation of precipitation and temperature at 2255 sub-basins (Fig. 3). In general, the CC statistics of both temperature and precipitation in all databases increased from north to south in all seasons. The principal reasons for this trend are the denser observational stations and fewer missing data for the southern watersheds.

Statistical analysis of the CRU1 and CRU2 (Fig. 2c,d; Fig. 2h,i; Table 2a,b) datasets showed the performances of these two datasets were not consistent and varied depending on the season and climate variable. Across the Alberta watersheds, the CRU2 dataset did not perform better than CRU1, as indicated in a global study by Harris et al. (2014). The CC of precipitation using CRU2 indicated a slightly stronger relationship in spring and summer seasons and weaker relationship in winter and fall seasons compared to that of CRU1. The CC of temperature in CRU2 was slightly smaller than that of CRU1 for all seasons.

The NRCan precipitation data (Fig. 2e; Table 2a) showed the strongest relationships (highest CC) with observed MS data in all sub-basins, ranging from 0.15 to 0.74 for the fall and winter seasons and from 0.63 to 0.82 for the spring and summer seasons. Likewise, the temperature CC (Fig. 2j; Table 2b) was the greatest for the NRCan dataset, ranging from 0.06 to 0.83 in cold seasons and from 0.46 to 0.80 in warm seasons. Although the CC was generally the highest for NRCan, the MAE and PBias were different between seasons and watersheds.

Comparison of the performance of the gridded data in this study showed that the CFSR dataset performed well for temperature (a higher CC and lower MAE and PBias), especially in the middle to southern data-rich watersheds (Table 2b). In hydrological modeling, the accuracy of temperature data is important because it has direct effect on simulation of snow fall in cold seasons and snowmelt in warm seasons. Snow fall has significant but not an immediate contribution to streamflow. In upstream highlands, precipitation in the form of snow results in a temporal shift of the hydrograph, such that contributions to river flows occur in later seasons (spring and summer) when warmer temperature melts the snow. Occasionally, we find large PBias values in southern watersheds in the fall season because near-zero temperature values (Fig. 3h) are used as a denominator in Eq. (3). To overcome this problem we added 1 °C to all of the fall temperature data in these watersheds. The results were improved and the statistics were comparable with other data (Table 2b).

3.2. Scenario results

3.2.1. Comparison of SWAT model performance using various datasets

We evaluated streamflow in 10 different scenarios (S1–S10; Table 1) using SWAT. River discharge measured at each hydrometric station reflects system inflows (e.g., precipitation), outflows (e.g., evapotranspiration), water storage changes (e.g., in lakes and groundwater), and management measures (e.g., dam regulation) throughout the entire upstream area (Hunger and Doll, 2008). Therefore to improve performance of the streamflow simulations, the input data used for model setup, and a multi-gage evaluation procedure should characterize as many of the natural and anthropogenic processes in the catchment as possible. The MS simulation result (S1) at a monthly time step yielded a desirable performance with a bR^2 of up to 0.65 in the southern watersheds (Fig. 4, S1). However, the results showed a low bR^2 (ranging from 0 to 0.17) for

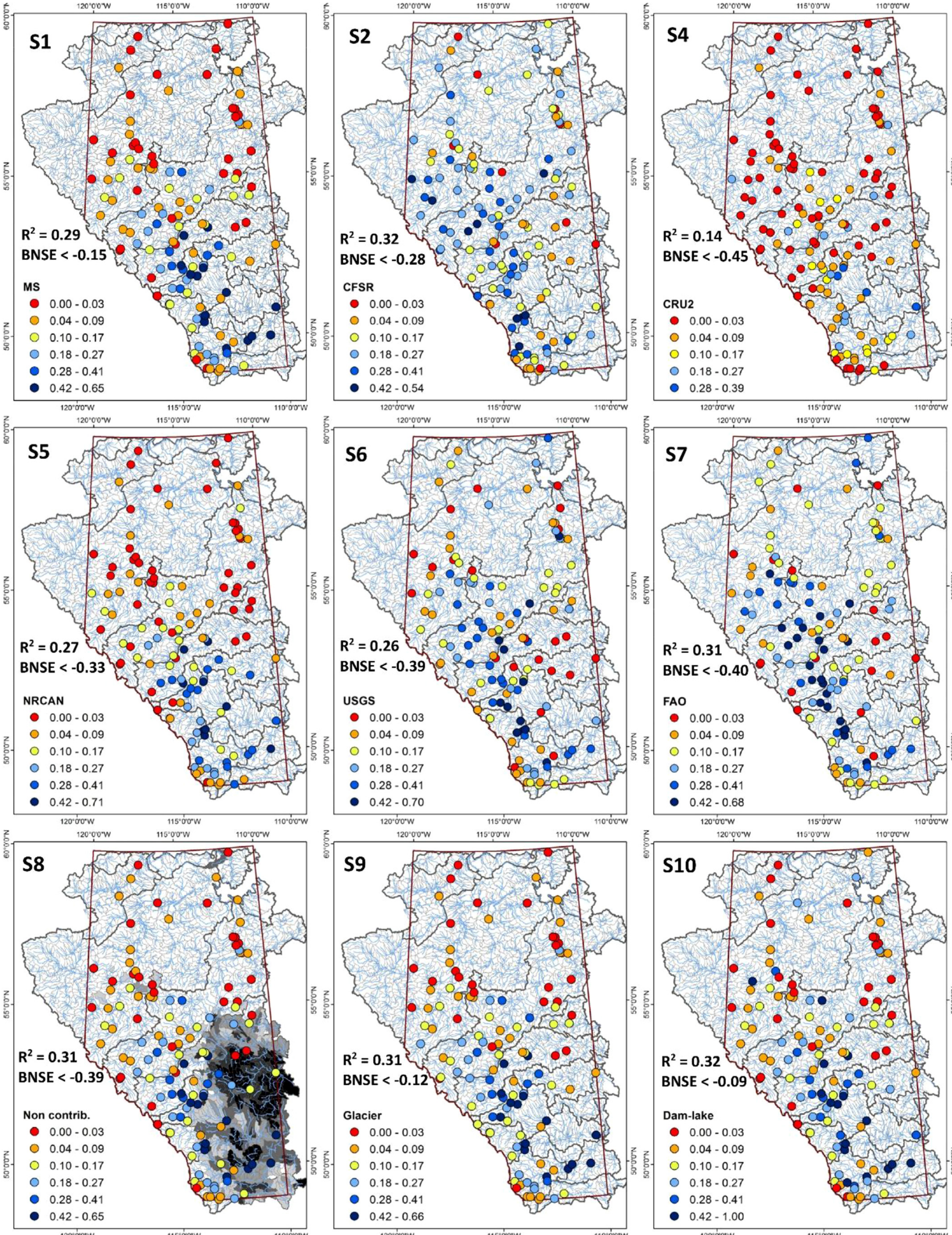

Fig. 4. Model performance of different scenarios for monthly streamflows at 130 hydrometric stations. Different colors show the bR^2 of the stations from the comparison of the measured versus simulated discharges during 1986–2007. The *BNSE* and R^2 are the mean values across the study area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

most of the northern watersheds. This poor performance was expected because the quantity and quality of the MS data were quite poor for these regions of the study area. In general, the average provincial *BNSE* and R^2 were about −0.15 and 0.29, respectively in this scenario.

Replacing the meteorological data with the CFSR dataset, the model performance was improved (Fig. 4, S2). However, the improvement was not consistent throughout the study area. The bR^2 was improved in the central, northern and mountainous regions (with a maximum of about 0.54) but it decreased in the south-eastern watersheds (from a maximum of 0.65 under S1 to a maximum of only 0.54 under S2). At the provincial scale, the mean *BNSE* was decreased from −0.15 under S1 to −0.28 under S2, while the R^2 was increased from 0.29 under S1 to 0.32 under S2. A comparison of the simulated monthly discharges with the observed records in S2 showed over estimation of the streamflow for most of the stations where bR^2 increased. Improvement of the bR^2 in this scenario were primarily driven by significant improvements in the trend of the simulated data (R^2) rather than the magnitude (b). This implies that large precipitation estimates of CSFR dataset (Fig. 2b) did not correspond well with the simulation of streamflows and caused an over estimation of the discharges in most of the hydrometric stations (with a small b). However, the CFSR temperature time series made a significant overall improvement in trend simulation (with large R^2). The performance of the CFSR temperature dataset in streamflow simulation is shown in the next section. Previous studies evaluating the utility of CFSR data for hydrological modeling have been conducted on one, or at most, a few small watersheds (Najafi et al., 2012; Smith and Kummerow, 2013; Fuka et al., 2013; Dile and Srinivasan, 2014). In this study we have evaluated how the CFSR dataset performs across a broad range of catchments, representative of diverse climatic and hydrological conditions.

Using either of the two CRU datasets (CRU1–Supplementary Fig. A.1; S3 or CRU2-Fig. 4, S4) resulted in poor simulations. The bR^2 values decreased for most of the hydrometric stations compared to S1. The average *BNSE* and R^2 decreased to about −0.64 and 0.11 for S3 and -0.45 and 0.14 for S4, respectively. Such low performance in streamflow simulation is in agreement with the MAE and PBias statistics (Table 2a,b), where the CRU1 and CRU2 datasets performed undesirable in most of the watersheds in the province.

The NRCan dataset (S5), which had the highest spatial resolution and a desirable statistical performance (Table 2a,b), did not produce the best performance for streamflow simulation. In S5, the average *BNSE* and R^2 of the study area were −0.33 and 0.27, respectively (Fig. 4). Analyses showed that both R^2 (trend) and b (closeness) decreased for most of the stations under this scenario when compared with S1.

Using different spatial maps, and related physical properties, may improve biases in streamflow prediction (S6 and S7). However, the use of the USGS global landuse map (S6) and the FAO global soils map (S7) did not significantly improve the simulation results. In scenarios we expect improvement not only in the tributaries and head waters but also in main streams and northern watersheds. The FAO soils map resulted in a slightly better mean R^2 (0.31) but a smaller mean *BNSE* (−0.40) for the study area, when compared to S1. The dominant landuse and dominant soil options were selected in this study to characterize the SWAT sub-basins; thus a more substantial change could occur if the simulations were conducted at HRU level, where more of the spatial resolution would be captured by the model.

Apart from climate data and geo-spatial maps, which are usually considered the major sources of error, large non-contributing areas in the southern portion of our study area were found to have considerable influences on streamflow predictions. It has been reported that the pothole topography and depressional areas (both natural and anthropogenic) in the southern prairies generally result in low runoff coefficients and water yields; however, the contributing areas may fluctuate greatly between wet and dry periods (Shaw et al., 2012; Kienzle and Mueller, 2013). Data from PFRA-AAFC (2012) was used to map potholes and sloughs in the southern prairies (Fig. 4 S8, darker shading indicates increasing share of non-contributing areas). Inclusion of this data and simulation of related physical processes considerably improved the simulation results for the affected areas. It must be pointed out that simulation of potholes allowed apportioning of the stored water in the impoundments into evaporation and infiltration to ground water. As such, the streamflow simulation for the hydrometric stations located at Beaver, North Saskatchewan, Battle, RedDeer, and Oldman river basins. However, for the province as a whole, inclusion of non-contributing areas did not significantly enhance our simulation with a mean *BNSE* of only −0.39 and an R^2 of 0.31.

The effect of including melt water runoff from glacierized subbasins of the Rocky Mountains was also examined as this can significantly affect the hydrological regime of the downstream subbasins, especially in warmer seasons and drier years. Disregarding these influences can result in erroneous parameter estimation. Glacial contribution to streamflow is not generally measured in Rocky Mountain headwater streams – as such, the long-term monthly glacial contribution to streamflow was estimated using data from Marshall (2014) for each individual river basin and distributed within the tributaries using the percent coverage of the glaciers obtained from the Global Land Ice Measurement from Space (GLIMS) map (see Table 1). The data were fed into the SWAT model through point sources in the upstream head waters where they were close to the glaciers. Although the yearly fluctuations of the melt water runoff were not considered in this study, the results were improved in some of the western hydrometric stations which were influenced by the glaciers. The total average *BNSE* and R^2 were improved to −0.12 and 0.31, respectively (Fig. 4, S9).

Finally, the effect of including major water management measures on streamflow simulation was examined. The monthly outflows of 14 dams-reservoirs (Supplementary Table A.4), which are mainly constructed on southern streams and managed to regulate downstreamflows, were fed into the SWAT model. In addition, there are several small lakes and natural reservoirs in Alberta which are located on the rivers and tributaries. These small lakes have negligible influences on downstreamflow regime. The LSL is the largest natural lake in the province, which significantly alters the hydrological regime of the downstream on the Athabasca River. We included this lake and treated it as a reservoir in the model. Overall, the S10 scenario considered the effect of 14 dams-reservoirs plus the LSL on Athabasca River. The bR^2 was improved by up to 0.99 in some hydrometric stations located downstream of the dams, demonstrating the importance of considering dams operation (Fig. 4, S10). However, for the province as a whole, the inclusion of this single measures did not significantly improve our predictions with the average R^2 improving only to 0.32 while the *BNSE* increased to −0.09 compared to −0.15 in S1.

Overall, multiple scenarios were generated to explicitly examine the effect of a range of specific datasets (i.e., observed meteorological data, gridded climate data, landuse data, soils data, glacial data, and dams-reservoirs and lakes data) on model performance. Results showed that model performance varied substantially for different watersheds depending on the input data, but the improvements were spatially heterogeneous, only occurring in specific catchments. The aggregated performance of the scenarios at each watershed level showed that the bR^2 was increased in some watersheds while it decreased in other watersheds (Fig. 5). Some

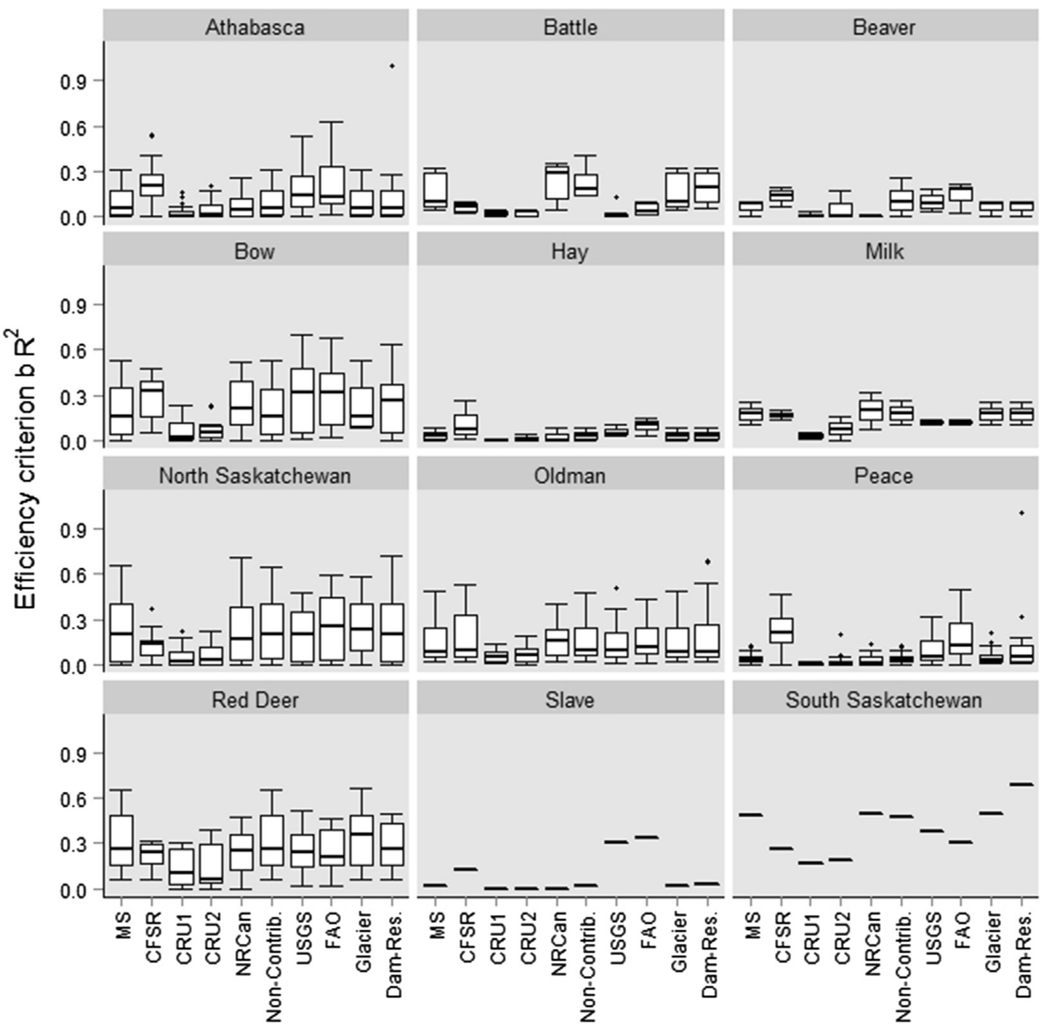

Fig. 5. Model performance of different scenarios for monthly streamflows at main river basins. The box plot of each scenario-dataset within each river basin shows the bR^2 of the simulated versus observed monthly discharges recorded in hydrometric stations of that river basin. Slave and South Saskatchewan River Basins are evaluated with one hydrometric station in each.

Table 3
Average criteria efficiency of different scenarios in the study area prior to calibration.

Scenarios	bR^2	Number of stations with $bR^2 > 0.40$	BNSE	NSE	R^2
S1: MS	0.16	16	−0.15	−2000	0.29
S2: CFSR	0.18	14	−0.28	−3000	0.32
S3: CRU1	0.05	0	−0.64	−6000	0.11
S4: CRU2	0.06	0	−0.45	−5000	0.14
S5: NRCan	0.13	12	−0.33	−3000	0.27
S6: USGS	0.17	13	−0.39	−5000	0.26
S7: FAO	0.19	20	−0.40	−4000	0.31
S8: non contrib.	0.17	18	−0.39	−4000	0.31
S9: glacier	0.18	19	−0.12	−2.82	0.31
S10: dam/lake	0.18	20	−0.09	−2000	0.32
S11: combined (1)	0.24	26	−0.04	−3.43	0.39
S12: combined (2)	0.18	17	−0.15	−100	0.29
S13: combined (3)	0.31	30	0.09	0.12	0.44

scenarios improved the simulation results, while others decreased the model performance. Table 3 summarizes performance of our scenarios (see S1–S10) at provincial level. Compared to S1, the number of stations with >0.40 was increased under S7–S10 scenarios followed by an increase in bR^2 and R^2. However a decrease in the performance of other stations resulted in an overall decrease in *BNSE* in S6 and S7. Overall, the performance of the model at a province scale did not appreciably improve to produce satisfactory results by incorporating any of the individual datasets. By "satisfactory results" we mean performance gain not only in small tributaries and head waters but also in main streams across the province. A statistical test is a way of quantifying significance of the performance gain among scenarios (Bennett et al., 2013). However, we did not perform statistical test on bR^2 values, because the statistics provided would be meaningless as stations do not have similar weights and are not equally important. In other word, our goal was to examine improvements in all rivers (Table 3) to be able to represent the actual processes not only in upstream tributaries but also in downstream main watersheds.

3.2.2. Use of multiple datasets to improve performance of the SWAT model

Given the lack of improvement in overall performance in the streamflow predictions for the province despite specific improvement for individual hydrometric stations, the effect of combined datasets on the performance of the model was also examined. Three scenarios were developed (S11, S12, S13, see Table 1) to combine data from the individual dataset scenarios where our simulation results were improved overall (Fig. 6a,b,c). We also included the effect of the dams/LSL in each of the new scenarios, since this was a major source of variability (Fig. 6d,e,f).

Using the S1 scenario as a base, the temperature data in the original MS simulation was replaced with data from the CFSR dataset and operation of the dams/LSRL was also included (S11). S11 resulted in a considerable improvement in streamflow simulation, especially in the northern watersheds where the quantity and quality of observed temperature data were poor (Fig. 6a). The bR^2 was increased to about 0.9 in some of the hydrometric stations which are close to the outlet of dams. The average *BNSE* and R^2 in this scenario were improved to −0.04 and 0.39, respectively. For the dams (Fig. 6d), we compared the simulated outflows with that of measured data and S11 demonstrated that model performed well for some dams but overall performance was not ideal since bR^2 was below 0.42 for 8 of the 14 stations recorded. Given the large effect of

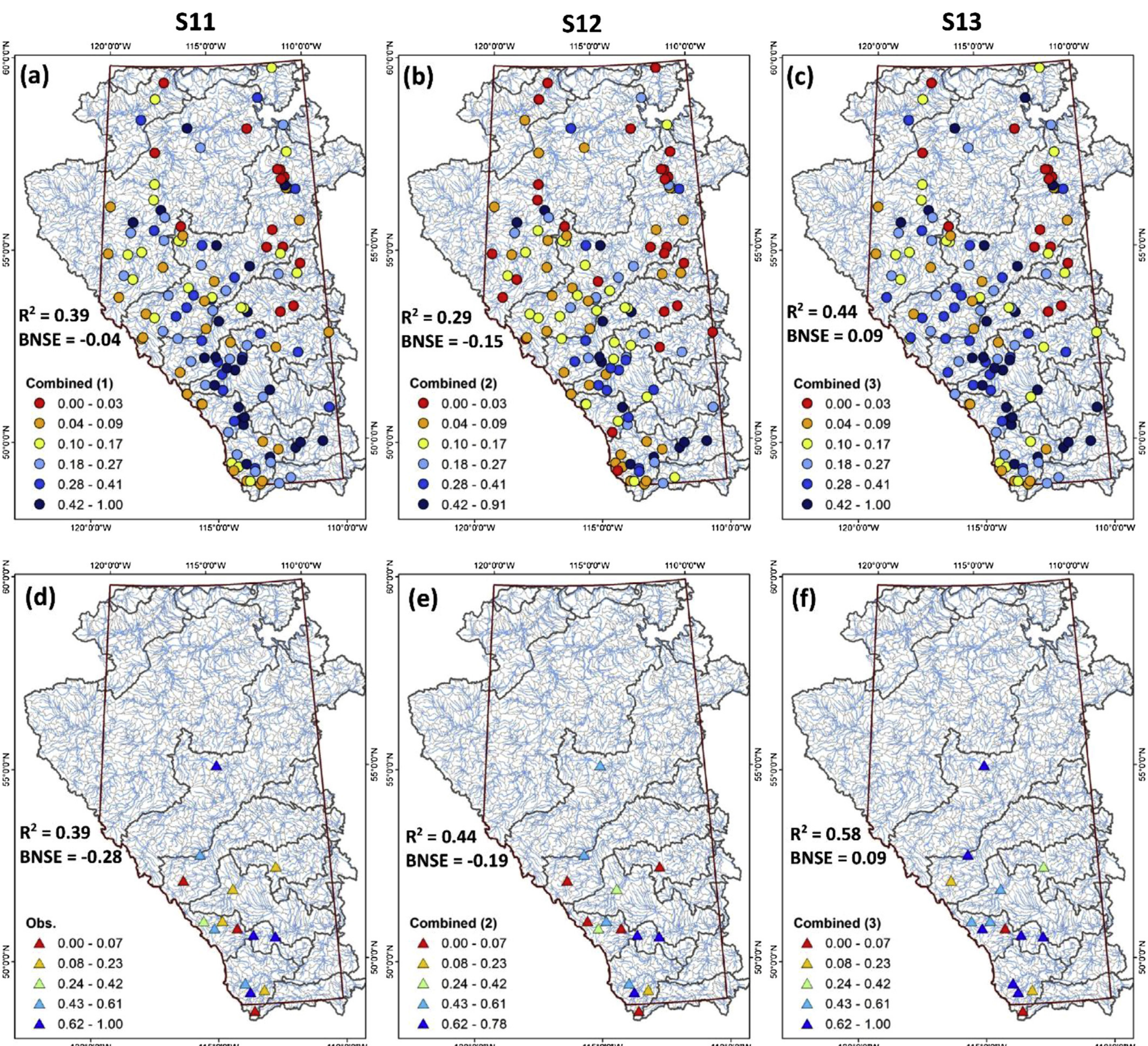

Fig. 6. Model performance of S11, S12, S13 scenarios for monthly streamflows at 130 hydrometric stations (a–c); and at the outlet of the 14 main dams and the Lesser Slave Lake (d–f). Different colors show the bR^2 of the stations and dams from the comparison of the measured versus simulated data. The *BNSE* and R^2 are the averaged values of the stations (a–c) and dams (d–f) for the whole study area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

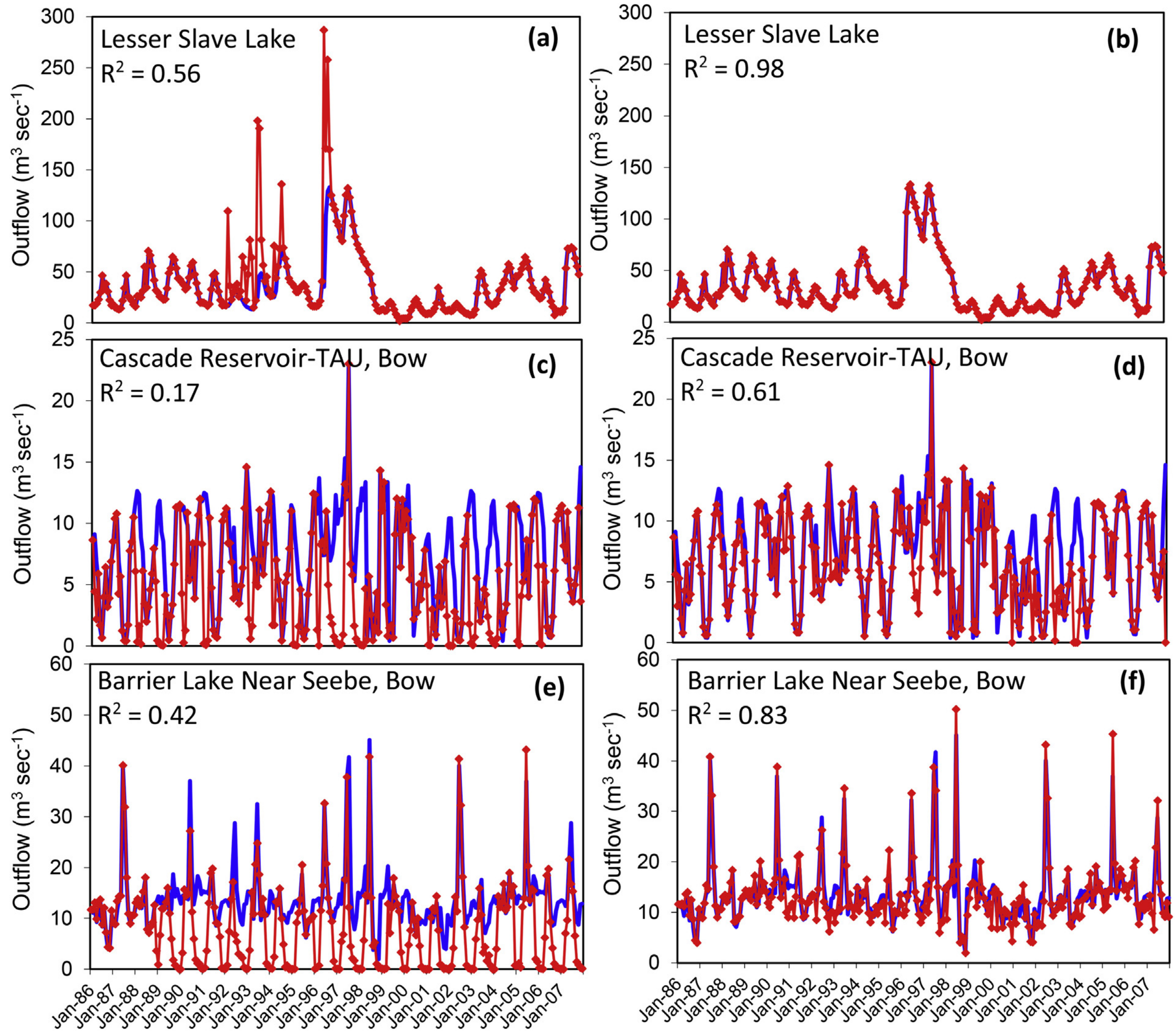

Fig. 7. Comparison of the monthly measured (blue line) and simulated (red line) discharges under S12 (left) and S13 (right) scenarios for Lesser Slave Lake and the two dams located in Bow River basin. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dams on the performance of the model, we sought to improve this performance by using additional datasets.

For S12, the Agriculture Agri-Food Canada, Soil Landscapes of Canada V3.2 data was replaced with the FAO soil data into a new simulation while also including the datasets from S11. However, S12 did not improve the performance of the model and the incorrect simulation of the inflows to the reservoirs caused improper simulation of the dam outflows (Fig. 6e; S12). This clearly demonstrates that inclusion of each dataset and evaluation of the model performance is required for proper set up of the SWAT model and this should occur prior to calibration.

In the final scenario (S13), the Agriculture Agri-Food Canada, Soil Landscapes of Canada V3.2 dataset was used and the dataset for non-contributing regions and glacier runoff were also included as these were shown previously to improve our streamflow simulations (see Table 1). S13 significantly improved model performance, both in streamflow prediction throughout the province (see Table 3) and in simulation of dam outflows (Fig. 6, S13). The mean *BNSE* and R^2 were improved to 0.09 and 0.44, respectively (Fig. 6c). Moreover, simulation of dam outflows improved substantially with the mean *BNSE* of 0.09 and R^2 of 0.58. Only 4 of the 14 reported stations had a bR^2 below 0.43 (Fig. 6f).

To better illustrate the effects of inclusion of different combined datasets on improvement in model predictions, we selected 3 hydrometric stations downstream of either lake outflows (Fig. 7a,b) or dams (Fig. 7c–f) and demonstrated the effect of S12 or S13 on the model performance. The stations are immediately after the reservoirs/lake and they are not influenced by any major tributaries. Hence, they represent the outflow of the dams and lake. Lesser Slave Lake exists in the northern part of the province and has some non-contributing areas due to extensive agriculture in the region and also has many large sand hills that likely serve as buffers for water flow. There is no inflow directly from glaciers in this upper watershed. Consequently, inclusion of non-contributing areas in S13 resulted in a near-perfect ($R^2 = 0.98$) simulation of water outflow from this natural impoundment (Fig. 7b). It is important to mention that although the monthly outflow data of the dams and LSL were used as input to the SWAT model but a perfect simulation of these outflows were not possible unless an accurate simulation of the upstream inflows to the reservoirs was obtained.

For man-made dams (Fig. 7c–f), we selected both the Cascade reservoir and Barrier Lake, two impoundments on the upper Bow River, as examples of the improvements in performance of the simulation associated with inclusion of the datasets in S13. In both

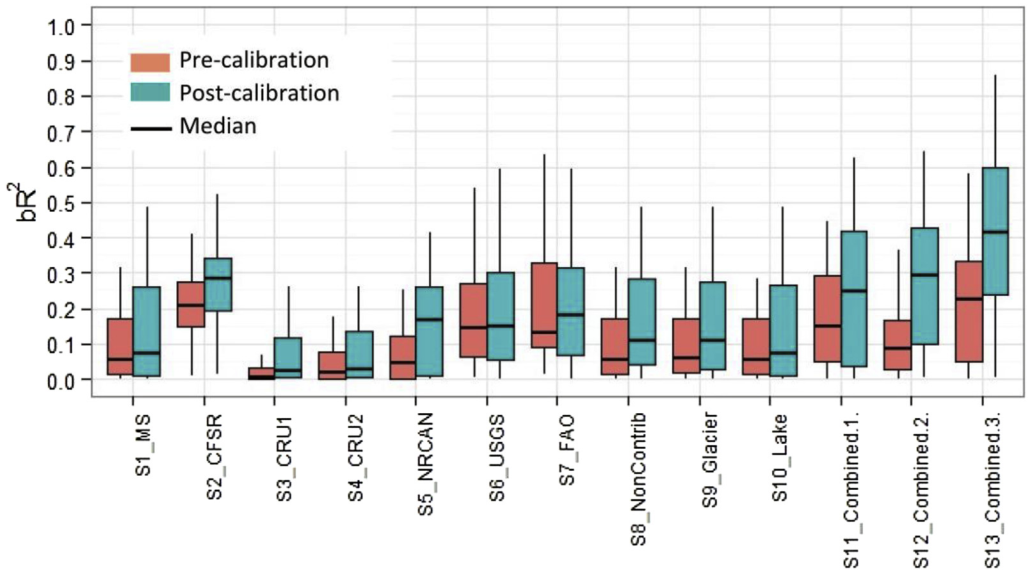

Fig. 8. Calibration performance of different scenarios for monthly streamflows in Athabasca River basin. The box plot of each scenario shows the bR^2 of the simulated versus observed monthly discharges in 40 hydrometric stations in Athabasca River basin.

cases, the inclusion of the non-contributing and glacier flows substantially improved the simulation of river flow when compared with measured values. An improper streamflow simulation at the upstream tributaries of the dams resulted in incorrect inflow to the reservoirs which then resulted in emptying or overflowing of the dams (S12, Fig. 7a,c,e). These events resulted in a poor simulation of the dams' outflow, which negatively affected the downstreamflow regime. We found that one of the main reasons for improvement noted in S13 (Fig. 7b,d,f) is the inclusion of CFSR temperature data compared with MS. The CFSR temperature data substantially improved simulation of the snow hydrology and snowmelt in the mountainous glacierized highlands which produced an accurate simulation of streamflow to the reservoirs behind the dams.

3.3. *Effects of calibration and parameter adjustment on the scenario selection*

To test how an accurate model setup performed prior to calibration serves as the best performing model after calibration, we calibrated all scenario models of Athabasca River basin (see Table 1). The calibration results showed an overall improvement in all scenarios at the river basin scale (See Fig. 8). However, the improvement was different across scenarios. In most of the less accurate scenarios (S1–S10), the overall trend of bR^2 in post-calibration step did not always mimic the trend in pre-calibration step. For example, the S1 ($bR^2 = 0.092$) and S4 ($bR^2 = 0.043$) had better performance in pre-calibration step compared to S5 ($bR^2 = 0.064$) and S3 ($bR^2 = 0.025$), respectively. However, adjustment of the parameters through calibration did not produce a better performance in S1 ($bR^2 = 0.147$) and S4 ($bR^2 = 0.068$) compared to S5 ($bR^2 = 0.165$) and S3 ($bR^2 = 0.082$). Nevertheless, the bR^2 trend was almost similar in pre and post-calibration steps in more accurate scenarios (S11–S13). In scenarios, where most appropriate input data were provided to represent most of the actual processes, the better performance models prior to calibration performed better after calibration, too. For example, in S11–S13 scenarios, where combination of best available data were provided to build the models, the S13 scenario served as the best performing model in both pre and post-calibration steps compared to all other scenarios (see Table 4). This followed by S11 and S12 as the second and third best scenarios in both pre and post calibration steps.

It must be pointed out that the parameters were optimized using bR^2 as the objective function in this study. The other efficiency criteria presented in Table 4 are based on the best

Table 4
Average criteria efficiency of different scenarios for pre-calibration and post-calibration steps in Athabasca River basin.

Scenario	bR^2		BNSE		NSE		R^2	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
S1: MS	0.092	0.147	−0.087	0.014	−0.363	−0.111	0.266	0.301
S2: CFSR	0.214	0.275	−0.632	−0.608	−8.767	−5.842	0.303	0.390
S3: CRU1	0.025	0.082	−0.224	−0.129	−0.743	−0.502	0.087	0.167
S4: CRU2	0.043	0.068	−0.191	−0.131	−0.656	−0.410	0.134	0.191
S5: NRCAN	0.064	0.165	−0.149	−0.001	−0.451	−0.135	0.229	0.367
S6: USGS	0.179	0.187	−0.348	−0.353	−3.086	−3.129	0.260	0.254
S7: FAO	0.208	0.226	−0.388	−0.253	−3.999	−2.080	0.287	0.300
S8: non contrib.	0.099	0.151	−0.072	0.015	−0.301	−0.090	0.266	0.308
S9: glacier	0.104	0.165	−0.061	0.032	−0.310	−0.097	0.291	0.322
S10: dam-Lake	0.108	0.147	−0.062	0.014	−0.337	−0.110	0.281	0.302
S11: combined (1)	0.202	0.314	0.019	0.107	−0.167	0.061	0.372	0.415
S12: combined (2)	0.156	0.281	−0.087	−0.035	−0.364	−0.355	0.255	0.378
S13: combined (3)	0.235	0.423	0.147	0.188	−0.024	0.112	0.394	0.535

performing parameter set to serve as side information in evaluation process. As also mentioned by Bennett et al. (2013), a single performing criterion may represent only specific aspects of model performance, which do not reproduce important features of a system. In addition, only streamflow data was considered for our calibration procedure. Given the fact that calibrated model parameters are "conditioned" on many factors including input databases, optimization algorithms, number of calibration variables (i.e., single or multiple variables), and all other assumptions, calibration of streamflow data using a single optimization algorithm should be treated with caution. Therefore, we emphasize that the "best" model in our calibration scheme subjects to the use of data, objective function (i.e., bR^2), and output variable (i.e., streamflow) used to calibrate model scenarios. Providing better quality data especially in mountainous regions where climate data are scarce and glaciers which play important role in the downstreamflow regime, might result in another best model. In addition, a best model that is calibrated against a single output variable (e.g. streamflow) may not perform best when comparing the other model outputs (e.g., evapotransiration, groundwater recharge, soil water etc.). Therefore, our obvious conclusion in this study is that building an accurate model using a database of higher quality is better than building less accurate models using many databases of questionable qualities where calibration is considered as a way to fill such gaps.

As indicated in Bennett et al. (2013), the method of choice for performance evaluations of the environmental models should be tailored to the model purpose and scale of the study. In our large scale study where measured ground water recharge or other water components are not available at the provincial scale we believe multi gauge calibration using the monthly river discharges of 130 hydrometric stations and providing good quality input data to the model will represent the upstream hydrological processes including system inflows, outflows, water storage changes, and management measures throughout the entire upstream area.

4. Conclusions

The process-based semi-distributed hydrologic model, SWAT, was used to quantify the causes and extents of biases in streamflow simulation due to the use of various input data. The various data types represented different sources of errors, including input data (e.g., climate, soils, and landuse), conceptual model error (i.e., the effect of glaciers and potholes, which were considered here as unknown processes altering hydrological regime and were explored through expert meetings), and land management measures (i.e., operation of large dams and lakes, which influence downstreamflow regime). We built 10 different SWAT projects, beginning from a base project and replacing individual datasets. We also built 3 scenarios using different combinations of multiple datasets. We qualified hydrological responses of the SWAT projects through simulation of the streamflow and comparison with the measured records at 130 hydrometric stations. The results showed that improvements due to single measures were local. However, a proper combination of input data, to better account for actual physical processes, considerably improved the overall model performance. Different scenarios demonstrated the importance of spatially representative temperature records and incorporation of glaciers runoff data.

Furthermore, simulation of potholes in southern prairies, and large reservoirs/lakes had a notable effect in producing more accurate simulation of streamflow. Specifically, we found that inclusion of the CFSR temperature data improved our simulation results in the data scarce northern watersheds and western snow dominated highlands but the precipitation data resulted in an over estimation of the streamflow. The procedure used in this study shows the importance of carefully scrutinizing and selecting databases that will most accurately represent the hydrological processes in the model prior to calibration.

Calibration of all scenarios in Athabasca River basin, revealed that an accurate model built with a database of higher quality performs better than a model where databases of questionable qualities were used. Providing better quality data in model setup will avoid unnecessary and arbitrary adjustment of the parameters and will ensure better performance when dealing with subjective and challenging calibration analysis. Overall, our findings recommend the use of various available data sources in hydrological modeling and qualifying them through alternative simulation scenarios prior to calibration of the model parameters.

Acknowledgments

This study was supported by the Alberta Innovates – Energy and Environment Solutions (AIEES, grant No. E309584). The authors are especially grateful to the Alberta Environment and Sustainable Resource Development, the Alberta Water Smart, the Alberta Geological Survey, the Alberta Energy Regulator, and the Alberta Biodiversity Monitoring Institute for their collaboration, provision of literature and data, and valuable comments and discussion of this paper. We are especially grateful to Mike Nemeth from Alberta Water Smart, and Jason Brisbois for aiding in data collection and linking us with the appropriate organizations throughout the Alberta.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2015.09.006.

References

Abbaspour, K.C., Sonnleitner, M., Schulin, R., 1999. Uncertainty in estimation of soil hydraulic parameters by inverse modeling: example Lysimeter experiments. Source Document Soil Sci. Soc. Am. J. 63, 501–509.

Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R., 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. J. Hydrol. 333, 413–430. http://dx.doi.org/10.1016/j.jhydrol.2006.09.014.

Abbaspour, K.C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Klöve, B., 2015. A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. J. Hydrol. 524, 733–752. http://dx.doi.org/10.1016/j.jhydrol.2015.03.027.

Alberta Agriculture and Rural Development (AARD), Government of Alberta, 2013. Alberta Irrigation Information Booklet 2012, Basin Water Management Branch, Irrigation and Farm Water Division. Path. http://www1.agric.gov.ab.ca/$department/deptdocs.nsf/all/irr7401.

Alberta Environment Golder Associates Ltd (AENV-GA), 2006. Hydrologic Regions of Alberta, p. 66.

Alberta Environment Golder Associates Ltd (AENV-GA), 2008. Water Supply Assessment for Alberta, Water for Life: Reliable, Quality Water Supplies for Sustainable Economy, p. 121.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment–Part 1: model development. J. Am. Water Resour. Assoc. 34, 73–89. http://dx.doi.org/10.1111/j.1752-1688.1998.tb05961.x.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Allen, P.M., 1999. Continental scale simulation of the hydrologic balance. J. Am. Water Resour. Assoc. 35 (5), 1037–1051. http://dx.doi.org/10.1111/j.1752-1688.1999.tb04192.x.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20. http://dx.doi.org/10.1016/j.envsoft.2012.09.011.

Beven, K.J., 2000. Uniqueness of place and process representations in hydrological modelling. Hydrol. Earth Syst. Sci. 4, 203–213. http://dx.doi.org/10.5194/hess-4-203-2000.

Blasone, R., Madsen, H., Rosbjerg, D., 2008. Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling. J. Hydrol. 353, 18–32. http://dx.doi.org/10.1016/j.jhydrol.2007.12.026.

Cao, W., Bowden, W.B., Davie, T., Fenemor, A., 2006. Multi-variable and multi-site

calibration and validation of SWAT in a large mountainous catchment with high spatial variability. Hydrol. Process. 20, 1057–1073. http://dx.doi.org/10.1002/hyp.5933.
Dile, Y.T., Srinivasan, R., 2014. Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile river basin. J. Am. Water Resour. Assoc. 50, 1226–1241. http://dx.doi.org/10.1111/jawr.12182.
Ercan, M.B., Goodall, J.L., Castronova, A.M., Humphrey, M., Beekwilder, N., 2014. Calibration of SWAT models using the cloud. Environ. Model. Softw. 62, 188–196. http://dx.doi.org/10.1016/j.envsoft.2014.09.002.
Eum, H., Yonas, D., Prowse, T., 2014. Uncertainty in modelling the hydrologic responses of a large watershed: a case study of the Athabasca River Basin, Canada. Hydrol. Process. 28, 4272–4293. http://dx.doi.org/10.1002/hyp.10230.
Faramarzi, M., Abbaspour, K.C., Vaghefi, S.A., Farzaneh, M.R., Zehnder, A.J.B., Srinivasan, R., Yang, H., 2013. Modeling impacts of climate change on freshwater availability in Africa. J. Hydrol. 480, 85–101. http://dx.doi.org/10.1016/j.jhydrol.2012.12.016.
Faramarzi, M., Abbaspour, K.C., Schulin, R., Yang, H., 2009. Modelling blue and green water resources availability in Iran. Hydrol. Process. 23, 486–501. http://dx.doi.org/10.1002/hyp.7160.
Faramarzi, M., Abbaspour, K.C., Schulin, R., Yang, H., 2010a. Modeling wheat yield and crop water productivity in Iran: implications of agricultural water management for wheat production. Agric. Water Manag. 97, 1861–1875. http://dx.doi.org/10.1016/j.agwat.2010.07.002.
Faramarzi, M., Yang, H., Mousavi, J., Schulin, R., Binder, R.C., Abbaspour, K.C., 2010b. Analysis of intra-country virtual water trade strategy to alleviate water scarcity in Iran. Hydrol. Earth Syst. Sci. 14, 1417–1433. http://dx.doi.org/10.5194/hess-14-1417-2010.
Food and Agriculture Organization (FAO), 1995. The Digital Soil Map of the World and Derived Soil Properties. CD-ROM, Version 3.5, Rome.
Fuka, D.R., Walter, M.T., MacAlister, C., Degaetano, A.T., Steenhuis, T.S., Easton, Z.M., 2013. Using the climate forecast system reanalysis as weather input data for watershed models. Hydrol. Process. 28, 5613–5623. http://dx.doi.org/10.1002/hyp.10073.
Gabriel, M., Knightes, C., Cooter, E., Dennis, R., 2014. The impacts of different meteorology data sets on nitrogen fate and transport in the SWAT watershed model. Environ. Model. Assess. 19, 301–314. http://dx.doi.org/10.1007/s10666-014-9400-z.
Galvan, L., Olias, M., Izquierdo, T., Ceron, J.C., de Villaran, R.F., 2014. Rainfall estimation in SWAT: an alternative method to simulate orographic precipitation. J. Hydrol. 509, 257–265. http://dx.doi.org/10.1016/j.jhydrol.2013.11.044.
Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. Trans. ASABE 50, 1211–1250.
Gassman, P.W., Arnold, J.G., Srinivasan, R., Reyes, M., 2010. The worldwide use of the SWAT model: technical drivers, networking impacts, and simulation trends. In: Proceeding of 21st Century Watershed Technology: Improving Water Quality and Environment, CD-Rom Proceedings, 21–24 February 2010, Costa Rica.
Getirana, A.C.V., Espinoza, J.C.V., Ronchail, J., Rotunno Filho, O.C., 2011. Assessment of different precipitation datasets and their impacts on the water balance of the Negro River Basin. J. Hydrol. 404, 304–322. http://dx.doi.org/10.1016/j.jhydrol.2011.04.037.
Gosain, A.K., Rao, S., Basuray, D., 2006. Climate change impact assessment on hydrology of Indian river basins. Curr. Sci. 90, 346–353.
Gupta, H.V., Sorooshian, S., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. Water Resour. Res. 34, 751–763. http://dx.doi.org/10.1029/97WR03495.
Guse, B., Reusser, D.E., Fohrer, N., 2014. How to improve the representation of hydrological processes in SWAT for a lowland catchment – temporal analysis of parameter sensitivity and model performance. Hydrol. Process. 28, 2651–2670. http://dx.doi.org/10.1002/hyp.9777.
Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. Int. J. Climatol. 34, 623–642. http://dx.doi.org/10.1002/joc.3711.
Houska, T., Multsch, S., Kraft, P., Frede, H.G., Breuer, L., 2014. Monte Carlo-based calibration and uncertainty analysis of a coupled plant growth and hydrological model. Biogeosciences 11, 2069–2082. http://dx.doi.org/10.5194/bg-11-2069-2014.
Hunger, M., Doll, P., 2008. Value of river discharge data for global-scale hydrological modeling. Hydrol. Earth Syst. Sci. 12, 841–861. http://dx.doi.org/10.5194/hess-12-841-2008.
Islam, Z., Gan, T.Y., 2014. Effects of climate change on the surface-water management of the South Saskatchewan River Basin. J. Water Resour. Planning ASCE 332–342.
Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled SRTM for the Globe Version 4, the CGIAR-CSI SRTM 90m Database. http://srtm.csi.cgiar.org.
Joseph, J.F., Guillaume, J.H.A., 2013. Using a parallelized MCMC algorithm in R to identify appropriate likelihood functions for SWAT. Environ. Model. Softw. 46, 292–298. http://dx.doi.org/10.1016/j.envsoft.2013.03.012.
Kienzle, S., Mueller, M., 2013. Mapping Alberta's surface water resources for the period 1971–2000. Can. Assoc. Geogr. 57, 506–518. http://dx.doi.org/10.1111/j.1541-0064.2013.12050.x.
Kienzle, S.W., Nemeth, M.W., Byrne, J.M., MacDonald, R.J., 2012. Simulating the hydrological impacts of climate change in the upper North Saskatchewan River basin, Alberta, Canada. J. Hydrol. 412–413, 76–89. http://dx.doi.org/10.1016/j.jhydrol.2011.01.058.
Krause, P., Boyle, D.P., Base, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97.
Lapp, S.L., St Jacques, J., Sauchyn, D.J., Vanstone, J.R., 2013. Forcing of hydroclimatic variability in the northwestern Great Plains since AD 1406. Quat. Int. 310, 47–61. http://dx.doi.org/10.1016/j.quaint.2012.09.011.
Leta, O.T., Nossent, J., Velez, C., Shrestha, N.K., van Griensven, A., Bauwens, W., 2015. Assessment of the different sources of uncertainty in a SWAT model of the River Senne (Belgium). Environ. Model. Softw. 68, 129–146. http://dx.doi.org/10.1016/j.envsoft.2015.02.010.
Li, Z., Liu, W., Zhang, X., Zheng, F., 2009. Impacts of land use change and climate variability on hydrology in an agricultural catchment on the Loess Plateau of China. J. Hydrol. 377, 35–42. http://dx.doi.org/10.1016/j.jhydrol.2009.08.007.
Madsen, H., 2003. Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. Adv. Water Resour. 26, 205–216. http://dx.doi.org/10.1016/S0309-1708(02)00092-1.
Marshall, S.J., 2014. Meltwater runoff from Haig Glacier, Canadian Rocky Mountains, 2002–2013. Hydrol. Earth Syst. Sci. 18, 5181–5200. http://dx.doi.org/10.5194/hess-18-5181-2014, 2014.
Masih, I., Maskey, S., Uhlenbrook, S., Smakhtin, V., 2011. Assessing the impact of areal precipitation input on streamflow simulations using the SWAT model. J. Am. Water Resour. Assoc. 47, 179–195. http://dx.doi.org/10.1111/j.1752-1688.2010.00502.x.
Mathevet, T., Michel, C., Anderassian, V., Perrin, C., 2006. A Bounded Version of the Nash-sutcliffe Criterion for Better Model Assessment on Large Sets of Basins. In: IAHS Red Books Series n°307, pp. 211–219.
McKenney, D.W., Hutchinson, M.F., Papadopol, P., Lawrence, K., Pedlar, J., Campbell, K., Milewska, E., Hopkinson, R.F., Price, D., Owen, T., 2011. Customized spatial climate models for North America. Am. Meteorol. Soc. 92, 1611–1623. http://dx.doi.org/10.1175/2011BAMS3132.1.
Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatol. 25, 693–712. http://dx.doi.org/10.1002/joc.1181.
Moradkhani, H., DeChant, C.M., Sorooshian, S., 2012. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. Water Resour. Res. 48, W12520. http://dx.doi.org/10.1029/2012WR012144.
Muleta, M.K., Nicklow, J.W., 2005. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. J. Hydrol. 306, 127–145. http://dx.doi.org/10.1016/j.jhydrol.2004.09.005.
Mwale, D., Yew Gan, T., Devito, K., Mendoza, C., Silins, U., Petrone, R., 2009. Precipitation variability and its relationship to hydrologic variability in Alberta. Hydrol. Process. 23, 3040–3056. http://dx.doi.org/10.1002/hyp.7415.
Najafi, M.R., Moradkhani, H., Piechota, T.C., 2012. Ensemble streamflow prediction: climate signal weighting methods vs. climate forecast system reanalysis. J. Hydrol. 442–443, 105–116. http://dx.doi.org/10.1016/j.jhydrol.2012.04.003.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, Part 1-A discussion of principles. J. Hydrol. 10, 282–290.
Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Water Assessment Tool. Theoretical Documentation: Version 2009, TWRITR-191. Texas Water Resources Institute, Texas A&M University System, College Station, Texas, pp. 77843–82118.
New, M., Hulme, M., Jones, P.D., 2000. Representing twentieth century space–time climate variability. Part 2: development of 1901–96 monthly grids of terrestrial surface climate. J. Clim. 13, 2217–2238. http://dx.doi.org/10.1175/1520 0442(2000)013<2217:RTCSTC>2.0.CO;2.
Perrin, C., Michel, C., Andreassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. J. Hydrol. 242, 275–301. http://dx.doi.org/10.1016/S0022-1694(00)00393-0.
Prairie Farm and Rehabilitation Administration (PFRA), Agriculture Agri-food Canada (AAFC), 2012. . ISO 19131 AAFC Watersheds Project - 2012-Data Product Specification, p. 34. Path. http://www.agr.gc.ca/atlas/supportdocument_documentdesupport/aafcWatersheds2012/en.
Qiao, L., Herrmann, R.B., Pan, Z., 2013. Parameter uncertainty reduction for SWAT using GRACE, streamflow, and groundwater table data for Lower Missouri River Basin. J. Am. Water Resour. Assoc. 49, 343–358. http://dx.doi.org/10.1111/jawr.12021.
Raup, B., Racoviteanu, A., Khalsa, S.J.S., Helm, C., Armstrong, R., Arnaud, Y., 2007. The GLIMS geospatial glacier database: a new tool for studying glacier change. Glob. Planet. Change 56, 101–110. http://dx.doi.org/10.1016/j.gloplacha.2006.07.018.
Rouholahnejad, E., Abbaspour, K.C., Srinivasan, R., Bacu, V., Lehmann, A., 2014. Water resources of the Black Sea Basin at high spatial and temporal resolution. Water Resour. Res. 50 http://dx.doi.org/10.1002/2013WR014132.
Samuel, J., Coulibaly, P., Dumedah, G., Moradkhani, H., 2014. Assessing model state and forecasts variation in hydrologic data assimilation. J. Hydrol. 513, 127–141. http://dx.doi.org/10.1016/j.jhydrol.2014.03.048.
Sauchyn, D., Vanstone, J., Perez-Valdivia, C., 2011. Modes and forcing of hydroclimatic variability in the upper North Saskatchewan River Basin since 1063. Can. Water Resour. Assoc. 36, 205–217. http://dx.doi.org/10.4296/cwrj3603889.
Schuol, J., Abbaspour, K.C., Sarinivasan, R., Yang, H., 2008a. Estimation of freshwater availability in the West African Sub-continent using the SWAT hydrologic model. J. Hydrol. 352, 30–42. http://dx.doi.org/10.1016/j.jhydrol.2007.12.025.
Schuol, J., Abbaspour, K.C., Yang, H., Srinivasan, R., Zehnder, A.J.B., 2008b. Modeling blue and green water availability in Africa. Water Resour. Res. 44, W07406. http://dx.doi.org/10.1029/2007WR006609.

Shaw, D.A., Vanderkamp, G., Conly, F.M., Pietroniro, A., Martz, L., 2012. The fill-spill hydrology of prairie wetland complexes during drought and deluge. Hydrol. Process. 26, 3147–3156. http://dx.doi.org/10.1002/hyp.8390.

Smith, R.A., Kummerow, C.D., 2013. A comparison of in situ, reanalysis, and satellite water budgets over the upper Colorado River Basin. J. Hydrometeorol. 14, 888–905.

Strauch, M., Bernhofer, C., Koide, S., Volk, M., Lorz, C., Makeschin, F., 2012. Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation. J. Hydrol. 414–415, 413–424. http://dx.doi.org/10.1175/JHM-D-12-0119.1.

Van Griensven, A., Ndomba, P., Yalew, S., Kilonzo, F., 2012. Critical review of SWAT applications in the upper Nile basin countries. Hydrol. Earth Syst. Sci. 16, 3371–3381. http://dx.doi.org/10.5194/hess-16-3371-2012.

White, K., Chaubey, I., 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. J. Am. Water Resour. Assoc. 41, 1077–1089. http://dx.doi.org/10.1111/j.1752-1688.2005.tb03786.x.

Wu, Y., Li, T., Sun, L., Chen, J., 2013. Parallelization of a hydrological model using the message passing interface. Environ. Model. Softw. 43, 124–132. http://dx.doi.org/10.1016/j.envsoft.2013.02.002.

Xie, H., Longuevergne, L., Ringler, C., Scanlon, B.R., 2012. Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data. Hydrol. Earth Syst. Sci. 16, 3083–3099. http://dx.doi.org/10.5194/hess-16-3083-2012.

Xue, C., Chen, B., MASCE, Wu, H., 2014. Parameter uncertainty analysis of surface flow and sediment yield in the Huolin Basin, China. J. Hydrol. Eng. 19, 1224–1236. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000909.

Yalew, S., van Griensven, A., Ray, N., Kokoszkiewicz, L., Betrie, G.D., 2013. Distributed computation of large scale SWAT models on the Grid. Environ. Model. Softw. 41, 223–230. http://dx.doi.org/10.1016/j.envsoft.2012.08.002.

Yen, h, Wang, X., Fontane, D.G., Harmel, R.D., Arabi, M., 2014. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. Environ. Model. Softw. 54, 211–221. http://dx.doi.org/10.1016/j.envsoft.2014.01.004.