ABSTRACT
Small watersheds are ideal objects for studying the evolution of hydrological and water resources at small scales. Whether precipitation products can meet the runoff simulation of small watersheds is the main purpose of this study. With NOAA-CPC-US precipitation as a reference, in nine small watersheds of the United States, accuracy of the precipitation products such as PERSIANN, PERSIANN-CDR, TRMM-3B42V7, GPM-IMERG, StageIV, and ERA5 is analyzed. By driving the CREST hydrological model using these datasets, the runoff simulation effects were evaluated. Result shows the precipitation products match the NOAA-CPC-US from high to low in the order of StageIV, PERSIANN-CDR, GPM-IMERG, PERSIANN, ERA5, and finally TRMM-3B42V7. These datasets have relatively low accuracy in the northern high latitude area and the western mountains, while accuracy is better in the central, southern, and eastern parts of the United States. In the runoff simulation effectiveness evaluation, the daily runoff of watersheds is simulated during the same verification period. The results show that: NOAA-CPC-US and StageIV have good simulation effect in small watersheds. In the northern and western United States, the PERSIANN, PERSIANN-CDR, GPM-IMERG, and ERA5 for runoff simulation should be used with caution. TRMM-3B42V7 is not suitable for runoff simulation in small watershed.

Introduction
Precipitation, one of the main links in hydrological cycle, is also the main data input in hydrological model research. Precipitation plays an important role in the research, monitoring, and prediction of climate, meteorology, hydrological forecasting, and hydrological disasters (Islam, 2013). Affected by climatic factors and underlying surface factors, precipitation has uneven formation and distribution in space and time, which in turn affects water cycle processes of river basins such as land runoff, evapotranspiration, soil moisture, and groundwater. Therefore, accurate precipitation estimation and large-scale rapid acquisition of continuous precipitation observation data with high spatial and temporal resolution can not only help us better understand the water cycle, but also means important significance for the simulation and prediction of hydrological processes in the basin. At present, there are four methods for obtaining precipitation data: observation from ground precipitation stations, precipitation detection by ground radar, satellite remote sensing for precipitation and numerical simulation of precipitation based on climate system models (Tapiador et al., 2012). These four methods have respective advantages and disadvantages. Where observation by ground precipitation stations is direct and highly accurate. However, due to density of the stations and their spatial distribution, it is difficult to provide complex and variable spatial distribution of precipitation. Precipitation detection by ground radar can directly detect spatial distribution of precipitation, which can track the precipitation center and spatial changes in real time, but it is susceptible to terrain occlusion and radar ray uplift. Although satellite remote sensing for precipitation has a lower accuracy than ground observation of precipitation, it can quickly and conveniently acquire continuous precipitation data with certain space-time accuracy on a large scale. It can make up for the deficiency of ground observation of precipitation to a certain extent, which provides an effective source of precipitation data especially for areas with no data or few ground observation data. The climate system model provides an important way to simulate and predict climate change. However, due to the impact of global climate complexity, there are still uncertainties in the simulation results (Stocker et al., 2013). Current climate models can simulate the large-scale change features
of global precipitation, but there are still many deficiencies in the simulation of regional-scale precipitation (Ying & Chong-Hai, 2012; Zhen et al., 2015). The precipitation simulation data of some climate models cannot be directly used for hydrological utility assessment (Lü et al., 2016).

Recently, with the rapid development of satellite remote sensing technology, radar rain measurement technology, climate numerical models, as well as the application of machine learning algorithms in precipitation retrieval, a large number of precipitation products with different spatial and temporal resolutions have appeared (Hou et al., 2014). At the same time, many scholars have carried out a lot of research around evaluation of precipitation accuracy, analysis of precipitation changes at different time and space scales using precipitation products, and simulation of hydrological utility (Balsamo et al., 2018; Kumar et al., 2017; Maldonado, 2011; Nelson et al., 2016; Zhao et al., 2017). The global or large and medium regional-scale runoff simulation and real-time prediction system developed based on precipitation product are positively evaluated in the prediction and reproduction of multiple hydrological disasters (Omranian et al., 2018; Paska et al., 2017; Tian & Zou, 2018; Yang et al., 2017). However, there is still significant spatial variability in current precipitation products, and significant uncertainties still exist in hydrological applications at different temporal and spatial scales (Hong et al., 2006; Jiang et al., 2014, 2012; Krajewski et al., 2010; Mei et al., 2014; Sarachi et al., 2015). Large floods may also occur in small watersheds (Michaud et al., 2001); therefore, the ability in hydrological simulation and forecasting on small-watershed scale is still worth in-depth evaluation.

Small watersheds are ideal objects for studying the evolution rule of hydrological and water resource systems at small and micro scales. As important geographical units for soil erosion control and ecosystem restoration research, they are minimum units for calculating the production of water and sediment by rivers, which provide optimal geographical scale for research and management of hydrology and soil erosion. The importance of small watersheds can be illustrated by the relationship between cells and organisms, small tributaries, and large rivers. The use of different precipitation products to accurately simulate and predict runoff in small watersheds plays a very important supporting role for the above research work. This study selected small watersheds distributed in nine different geographic regions in the United States. Based on the precipitation reanalysis data of Climate Prediction Center, National Oceanic and Atmospheric Administration, United States (NOAA-CPC-US) ground precipitation stations, the accuracy of satellite precipitation products such as the precipitation estimation from remotely sensed information using artificial neural networks (PERSIANN), the precipitation estimation from remotely sensed information using artificial neural network climate data record (PERSIANN-CDR), the precipitation version 7 derived from tropical rainfall measuring mission 3B42 research version (TRMM-3B42V7), the precipitation of integrated multi-satellite retrievals for the global precipitation measurement (GPM-IMERG), radar precipitation Stage IV data of the National Centers for Environmental Prediction, United States (StageIV) and the precipitation product of the fifth-generation reanalysis dataset released by the European Center for Medium-Term Weather Forecasting (ERA5) is evaluated. Combining the coupled routing and excess storage (CREST) distributed hydrological model, the daily runoff process of the nine small watersheds is simulated and reproduced to evaluate the hydrological simulation effectiveness of multi-source precipitation products at the small-watershed scale, thereby providing references for research and applications of the above-mentioned multi-source precipitation product in the hydrological industry.

Materials and methods

Study area

There is no uniform definition for the concept of small watershed at home or abroad. Different countries and organizations have different watershed standards for different research purposes and perspectives. The small watersheds discussed herein are based on the U.S. Stream Classification System (USSCE) of the US Department of Energy’s Oak Ridge National Laboratory, USA. This system divides nearly 2.6 million rivers in the United States into 8 levels. From large to small, they are Great River, Large River, Mainstem, Medium River, Small River, Large Creek, Creek, and Headwater (McManamay & DeRolph, 2019). Indeed, there are many small watersheds in a climate zone. We used a random sampling method to select small watersheds in the middle of the climate zone without any restrictions. In this study, nine closed watersheds are selected from different geographic regions of the continental United States (Figure 1). The rivers in these small watersheds belong to headwater and creek level in the river classification system. See table (Table 1) for information like geographical zone, watershed area, and annual average flow of these nine small watersheds.
This study selects seven types of precipitation data: NOAA-CPC-US, PERSIANN, PERSIANN-CDR, TRMM-3B42V7, GPM-IMERG, StageIV, and ERA5. The data period is from 1 January 2002 to 30 June 2018. In the study, the time resolution was set to the daily scale, and the spatial resolution was 0.0083°×0.0083° (about 1 km²). Due to the different temporal and spatial resolutions of the above-mentioned 7 types of precipitation data, during the data preprocessing process, data with temporal resolution smaller than the daily scale is cumulatively converted to the daily scale, and the spatial resolution is unified to 0.0083°×0.0083° via bilinear interpolation.

**NOAA-CPC-US precipitation**
This dataset is a reanalyzed precipitation product of NOAA Climate Prediction Center based on ground observation of precipitation. This dataset collects precipitation observation information from tens of thousands of ground observation stations and compares it with radar, satellite observation and numerical model predicted precipitation to implement quality control. Considering the terrain effect, the data is optimally interpolated to create a daily precipitation field, finally forming a precipitation product with higher quantitative accuracy and consistency. Its spatial resolution is 0.25°, and the time coverage range is from 1 January 1979 to the present. The spatial range is 20.0°N–50.0°N, 230.0°E–305.0°E, covering the continental United States (https://www.esrl.noaa.gov/psd/data/gridded/data.unified.html).

**PERSIANN precipitation**
The PERSIANN precipitation dataset was developed by the Center for Hydrometeorology and Remote Sensing (CHRS) at the University of California, Irvine. Based on infrared brightness temperature images provided by geosynchronous satellites (GOES-8, GOES-10, GMS-5, Meteosat-6, Meteosat-7), it uses artificial neural networks to estimate 30-minute rainfall rate, then aggregates it to cumulative rainfall at 1, 3, 6-hour, and daily scales, and periodically amends it using low-orbit satellite (TRMM, NOAA-15, −16, −17, DMSP F13, F14, F15) precipitation estimation products. The data product covers the period from March 2000 to the present, with a spatial range of 60°S–60°N and a spatial resolution of 0.25°. The data download address is https://chrsdata.eng.uci.edu/.

**PERSIANN-CDR precipitation**
This dataset provides rainfall estimates from 1983 to the present with latitudes ranging from 60°S–60°N, spatial resolution of 0.25° at daily time scales. It is a precipitation product developed by the University of California; Irvine based on satellite remote sensing data. Different from other satellite precipitation products, it is generated on the GridSat-B1 infrared satellite data via artificial neural network algorithm (PERSIANN). Training of the artificial neural network algorithm is completed by the StageIV radar precipitation data of National Environmental Forecasting Center (NCEP); then, monthly precipitation product (GPCPv2.2) of the Global Precipitation Climatology Project (GPCP) is used for adjustment, finally forming this data product. The National Research Council (NRC) defines this data as a Climate Data Record, which has time series with sufficient length, consistency, and continuity to determine climate variability and change. The data download source is the same as PERSIANN.

**TRMM-3B42V7 satellite precipitation**
Due to the exhaustion of fuel in April 2015, TRMM satellite officially ended its observation mission and re-entered the Earth’s atmosphere in June of the
same year, but falling into the South Indian Ocean. Due to the success of TMPA precipitation product, TRMM team did not immediately stop the TMPA precipitation data product due to termination of satellite observation. Instead, various methods are adopted to extend TMPA precipitation products, which are planned to be updated until mid-2019 (Huffman, 2016). The selection of TRMM-3B42v7 precipitation product in this paper is also intended to assess the ability of this dataset to simulate hydrological utility under current conditions. TRMM-3B42 V7 data has a spatial resolution of 0.25°, a temporal resolution of 3 h, and a precipitation unit of mm/h. This data is downloaded from NASA’s TRMM satellite data website (https://pmm.nasa.gov/trmm/).

**GPM-IMERG satellite precipitation**
GPM is a follow-up global satellite precipitation observation program for TRMM led by the National Aeronautics and Space Administration (NASA). With the core observation platform launched on 28 February 2014, GPM can provide rain and snow observation data within 3 hours worldwide. Compared with TRMM, GPM covers a large area and extends to the poles of the earth. With high spatial-temporal resolution and superior sensor performance, it can accurately detect trace precipitation. For other characteristics of this precipitation product, please refer to the technical documentation provided by its portal (https://pmm.gsf.nasa.gov/GPM), and the Level 1–3 products of GPM data can also be downloaded from this website. This paper selects integrated multi-satellite retrievals for GPM (IMERG, version: 05B_Final) for global 30-minute rain and snow data product. The precipitation data was produced by the latest GPROF2017 algorithm, which calibrated and integrated all microwave, infrared, ground observation, and other possible sensor data of the GPM constellation (Huffman et al., 2015). It has a spatial resolution of 0.1°, a time resolution of 30 minutes and a precipitation unit of mm/h.

**NCEP-StageIV radar precipitation**
NCEP-StageIV is a Radar Quantitative Estimated Precipitation Product (QPE) of the National Weather Forecast Center of the United States Meteorological Service. The project started on 1 December 2001, and the data distribution started on 1 January 2002. The data incorporates precipitation estimates from approximately 150 Doppler new-generation weather radars across the United States and measurements from approximately 5,500 ground-based rain gauges, which is produced by the 12 river forecasting centers distributed throughout the continental United States by region. Data collection started at 33rd minute per hour, and cumulative calculations were performed with time steps of 1 h and 6 h, followed by iterative review and manual quality control, and finally the regional data were synthesized into quantitative radar precipitation estimation data for the continental United States (Lin, 2017). The data has a spatial resolution of 4 km, a temporal resolution of 1 h, 6 h, and 24 h, and a precipitation unit of mm. In this study, 6 h estimation data with StageIV time resolution is selected (www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/).

**ERA5 reanalysis dataset precipitation**
ERA5 is the fifth-generation reanalysis dataset released by the European Center for Medium-Term Weather Forecasting (ECMWF). Compared with ERA-Interim and ERA-40 re-analysis data of the previous generations, ERA5 adopts an improved data assimilation system (IFS Cycle 41r2), four-dimensional variational assimilation analysis (4D-Var) and radiation variation bias correction, thus systematically improving quality of this dataset. ERA5 provides an estimated global hourly value per day, and the hourly output resolution has a considerable improvement over ERA-Interim. The dataset includes 265 variables such as evaporation, humidity, air temperature, water vapor pressure, and precipitation, with time coverage ranging from 1979 to the present. With spatial resolution of 0.25°×0.25°, it has a spatial coverage of 89.142°S–89.142°N, 180°W–180°E. This study selects hourly scale precipitation data in this dataset. The data download address is https://cds.climate.copernicus.eu.

The above precipitation products have their own distinct specificities. NOAA-CPC-US precipitation mainly represents the huge traditional surface precipitation observation station data, and its precipitation on a single station has high accuracy. PERSIANN and PERSIANN-CDR precipitation products represent precipitation data retrieved by artificial intelligence algorithms based on many satellite observation data inputs. TRMM satellite precipitation data is very successful. It has created a milestone for humans to use satellite technology to predict precipitation, but it has failed. Its successor is the more excellent GPM satellite precipitation observation program. ERA5 represents the precipitation data output by the climate model. The precipitation products of climate models are also important data sources for runoff simulation and forecasting.

**Watershed runoff, topography and potential evapotranspiration data**
The watershed boundary, digital terrain DEM, and runoff data of the nine small watersheds herein adopts the Watershed Boundary Dataset V2.2.1 (US Geological Survey and US Department of Agriculture, Natural Resources Conservation Service, 2013) provided by the US Geological Survey (USGS), the US 10 m precision digital elevation model (DEM),
and US National Water Information System data. Based on the DEM data, GIS spatial analysis tools are used to extract the river network, river network flow direction, river network convergence information. In this study, a distributed hydrological model is used to evaluate hydrological simulation effectiveness of various precipitation data, and the required potential evapotranspiration (PET) data derives from USGS Famine Early Warning System (FEWS). Its daily global potential evapotranspiration is calculated using the Penman-Monteith equation based on climate data extracted from the global data assimilation system per 6 hours, and then summed to obtain a daily total. The data has a spatial resolution of 1.0°×1.0°. Data download source is https://earlywarning.usgs.gov/fews/datadownloads/Global/PET.

**Methodology**

The methodology consists of three major steps. First, the NOAA-CPC-US precipitation products were taken as a reference to conduct precipitation accuracy assessment for other precipitation products (StageIV, PERSIANN, PERSIANN-CDR, TRMM-3B42V7, GPM-IMERG, ERA5) of the same period. Second, the precipitation accuracy of different rain intensity levels was evaluated to indicate which precipitation products have good or poor estimation accuracy. Finally, the performance of hydrologic simulations in small watersheds was evaluated with multi-resource precipitation datasets by a distributed hydrologic model, Coupled Routing and Excess Storage (CREST).

**CREST distributed hydrological model**

The Coupled Routing and Excess Storage model (CREST) was jointly developed by the University of Oklahoma's Hydrometeorology and Remote Sensing Laboratory (HyDROS) and the NASA SERVIR project team. As a distributed hydrological model, it divides the whole basin into several grids, calculates runoff by grid using variable permeability curve, and calculates surface and groundwater confluence by grid using multi-linear reservoirs, and finally simulates and reproduces runoff process by coupling runoff elements and confluence structure (Wang et al., 2011). The model demonstrates excellent hydrological simulation and prediction capabilities at global scale, large regional scale, and small and medium scales (Khan et al., 2010; Meng et al., 2014; Shen et al., 2017; Xue et al., 2013). When the CREST model is running, the data that needs to be input include data watershed information (such as watershed boundary, DEM, flow direction, etc.), precipitation, potential evapotranspiration, watershed exit flow, initial conditions, and model parameters. Model output results include soil water content, soil moisture, and surface runoff. Please refer to the CREST model website of the HyDROS Laboratory for detailed information about model data preprocessing, model parameter value ranges, and so on (http://hydro.ou.edu/research/crest/).

**Statistical evaluation indicators**

Four statistical verification indicators were selected such as relative deviation (Bias), root mean square error (RMSE), mean absolute error (MAE), and correlation coefficient (CC) to quantitatively evaluate the accuracy of multi-source precipitation products. Standard deviation (SD), Nash efficiency coefficient (NSCE), relative deviation (Bias), and correlation coefficient (CC) were used to evaluate the hydrological simulation effectiveness of each precipitation product.

\[
Bias = \frac{\sum_{i=1}^{n} Sim_i - \sum_{i=1}^{n} Obs_i}{\sum_{i=1}^{n} Obs_i} \times 100
\]

Wang et al. (2011)

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Obs_i - Sim_i)^2}{n}}
\]

Xue et al. (2013)

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |Sim_i - Obs_i|
\]

Willmott et al. (2005)

\[
SD = \sqrt{\frac{\sum_{i=1}^{n} (Sim_i - \bar{Sim})^2}{n-1}}
\]

Willmott et al. (2005)

\[
CC = \frac{\sum_{i=1}^{n} (Obs_i - \bar{Obs})(Sim_i - \bar{Sim})}{\sqrt{\sum_{i=1}^{n} (Obs_i - \bar{Obs})^2 \sum_{i=1}^{n} (Sim_i - \bar{Sim})^2}}
\]

Wang et al. (2011)

\[
NSCE = 1 - \frac{\sum_{i=1}^{n} (Obs_i - Sim_i)^2}{\sum_{i=1}^{n} (Obs_i - \bar{Obs})^2}
\]

Wang et al. (2011)

In the above formula, Sim represents multi-source precipitation estimation data or runoff data simulated by CREST model, Obs represents reference precipitation data or runoff station observation data; n is the simulation value or the total observed value, and i is the i-th simulation value or observed value. When the four evaluation indicators meet Bias = 0%, RMSE = 0, CC = 1, NSCE = 1, it represents the optimal evaluation results. The NSCE value can be divided into 4 intervals to represent the utility of 4 levels of hydrological simulation: poor (NSCE ≤ 0.5), applicable...
(0.50 $\leq$ NSCE $\leq$ 0.65), good (0.65 $\leq$ NSCE $\leq$ 0.75), excellent (0.75 $\leq$ NSCE $\leq$ 1.00).

Results and discussion

Precipitation accuracy assessment

Accuracy of precipitation estimates by precipitation products has a significant impact on hydrological simulation effectiveness. Therefore, the precipitation accuracy of each precipitation product should be evaluated first. NOAA-CPC-US precipitation dataset collects precipitation observation information from tens of thousands of ground observation stations. Then, data is optimized for interpolation based on consideration to terrain effect, followed by quality control. Therefore, this study takes the NOAA-CPC-US precipitation products from 1 June 2014 to 30 June 2018 as a reference to conduct precipitation accuracy assessment for other precipitation products of the same period: StageIV, PERSIANN, PERSIANN-CDR, TRMM-3B42V7, GPM-IMERG, ERA5.

Precipitation accuracy

For each small watershed, the correlation coefficient between the precipitation products to be evaluated and NOAA-CPC-US precipitation was calculated at a grid scale of 1 km$^2$ (Figure 2, Figure 3). The vertical axis of Figures 2 and 3 is the correlation coefficient, and the horizontal axis is the precipitation product. The box in the figure represents the maximum, minimum and average value of the correlation coefficient. Among the six precipitation products, StageIV has the highest correlation coefficient with NOAA-CPC-US data, with an average value of 0.73. The correlation coefficient changes from 0.65 to 0.81 for small watersheds in different regions; the correlation coefficient is lower in the West and Southwest regions. Next comes PERSIANN-CDR and GPM-IMERG products, whose correlation coefficients with NOAA-CPC-US data are 0.51–0.69, 0.53–0.73, respectively. The correlation coefficients of these two precipitation products are smaller in Northwest and West regions. Next comes PERSIANN and ERA5 products, whose correlation coefficients with NOAA-CPC-US data are 0.49–0.66 and 0.49–0.62, respectively. Where, PERSIANN product has a low correlation coefficient in East North Central, West North Central and Northwest regions. ERA5 is a precipitation product of climate model. The average correlation coefficient in the nine small watersheds is 0.55. The regions with the lowest correlation coefficient are West, Northeast and West North Central. The precipitation product with the lowest correlation coefficient with NOAA-CPC-US data is TRMM-3B42V7, with a correlation coefficient of 0.44–0.59, and the average correlation coefficient in nine small watersheds is only 0.52. The correlation coefficient is the lowest in the Northeast, West North Central, East North Central, and Northwest regions. The precipitation product with the lowest correlation coefficient with NOAA-CPC-US data is TRMM-3B42V7, with a correlation coefficient of 0.44–0.59, and the average correlation coefficient in nine small watersheds is only 0.52. The correlation coefficient is the lowest in the Northeast, West North Central, East North Central, and Northwest regions. The precipitation products have slightly lower estimate accuracy in regions such as high latitude area (northeast) and western mountains of the United States. The precipitation products match the NOAA-CPC-US precipitation from high to low in the order of Stage IV radar precipitation, followed by PERSIANN-CDR and GPM-IMERG, followed by PERSIANN and ERA5, and finally TRMM-3B42V7.
Accuracy of different precipitation intensity levels

The International Meteorological Organization (WMO) precipitation intensity level classification standard (Tan et al., 2015) is adopted to analyze the precipitation accuracy of various precipitation products in more detail on different precipitation intensity levels. Precipitation intensity is divided into 6 levels according to the amount of hourly precipitation, namely (1) no rain: precipitation <1 mm, (2) drizzle: 1 mm ≤ precipitation < 2 mm, (3) light rain: 2 mm ≤ precipitation < 5 mm, (4) Medium rain: 5 mm ≤ precipitation < 10 mm, (5) Heavy rain: 10 mm ≤ precipitation < 20 mm, (6) rainstorm: 20 mm ≤ precipitation < 50 mm, (7) downpour: precipitation ≥ 50 mm. The frequency error of each precipitation product is calculated at different levels of precipitation intensity.

From the statistical results in Figure 4, it can be seen that relative to NOAA-CPC-US precipitation, occurrence frequency of different precipitation intensity differs for StageIV, PERSIANN, PERSIANN-CDR, TRMM-3B42V7, GPM-IMERG, and ERA5 precipitation products in nine small watersheds of different regions. There are overestimates and underestimates to varying degrees, but the errors do not exceed ±10%. In general, most precipitation products are underestimated for no rain (precipitation<1 mm) and drizzle (1 mm ≤ precipitation<2 mm) levels, and are overestimated for heavy rain (10 mm ≤ precipitation<50 mm) and rainstorm (precipitation≥50 mm). StageIV has higher estimation accuracy than PERSIANN, PERSIANN-CDR, TRMM-3B42V7, GPM-IMERG, ERA5 for each precipitation intensity level.

Hydrological simulation effectiveness assessment

After assessing the precipitation accuracy, this paper uses the above seven precipitation products to drive the distributed hydrological model CREST, respectively, so that runoff processes in nine small watersheds in different regions of the United States can be simulated, and hydrological simulation effectiveness of each precipitation product can be assessed. In run-off simulation scenarios, 1 January 2002 to 31 December 2003 is set as the model warm-up period, with 15 January 2004 to 30 Juneto 2014 as the model rate period, 1 July 2014 to 30 Juneto 2018 as the model verification period. In the same calibration period, the CREST parameters were respectively calibrated using seven types of precipitation products, and the runoff was simulated during the same verification period based on the parameter sets calibrated by the respective precipitation products. Statistical indicators of NSCE, Bias, SD, and CC were used to evaluate the hydrological simulation effectiveness. This scenario setting helps independent assessment of the simulation results of the basin hydrological process for each precipitation data.

Figures 5, 6 and 7) illustrate the daily runoff calibration and simulation results of the precipitation products in 9 small watersheds in different regions of the United States, which are relative to the standard deviation SD of the measured runoff, the root mean square error RMSE, and the correlation coefficient CC, Nash efficiency coefficient NSCE, and relative deviation Bias.

From Figure 5, Figure 6, Figure 7, and Table 3, the hydrological simulation results of the multi-source precipitation products are analyzed for the small watersheds of each region, and the precipitation product has better runoff simulation performance in calibration period than in verification period, and each statistical
These precipitation products have different runoff simulation performance in small watersheds in different regions of the United States. NOAA-CPC-US and StageIV precipitation maintain good runoff simulation characteristics, which reproduces the daily runoff process during the flood discharge of the basin. PERSIANN, PERSIANN-CDR, GPM-IMERG, and ERA5 precipitation are less capable of simulating runoff than NOAA-CPC-US and StageIV precipitation. The simulation of peak discharge has obvious fluctuations and lacks accuracy. The simulation performance is better in central, southern United States, but poor in northern high latitude area, west, and southwest of United States. TRMM-3B42V7 is not ideal for runoff simulation in multiple small watersheds due to sensor accuracy and resolution.

### Conclusion

It is a main trend to simulate and predict hydrological processes at different scales by obtaining precipitation estimation data via remote sensing technology and driving distributed hydrological models in the study of hydrological and water resources in river basins. The typical advantages of the remote sensing precipitation data lie in fast acquisition speed, large spatial range, continuous time, and certain accuracy guarantee. These data are also important for the prediction of natural disasters such as floods, hurricanes, and landslides; it is also an effective source of precipitation data for areas with no data or few ground observation data.

Compared with large-scale watersheds, small watersheds also play a very important role in the earth’s hydrosphere, biosphere, and lithosphere. Small watersheds have complete water cycle process, which are ideal objects for studying small and micro-scale hydrology, water resources, and water environment. It is also an important means to find out the evolution rules of small and micro hydrological water resource systems and find sustainable water resource utilization methods. In recent years, scholars have carried out many researches on hydrological and water resources on the global scale.
scale, intercontinental scale, and large-watershed scale based on remote sensing precipitation products. Scholars have done a lot of research work on small-watershed scale based on surface observation of precipitation around runoff coefficient, surface runoff simulation, and soil water movement (Hu et al., 2015; Lévesque et al., 2008; Liu & Li, 2008; Qiu et al., 2012). However, remote sensing precipitation estimation data still have insufficient research and application, especially in small watersheds.

Therefore, this study takes NOAA-CPC-US precipitation as a reference, comparatively analyzes the accuracy of Stage IV, PERSIANN, PERSIANN-CDR, GPM-IMERG, ERA5, TRMM-3B42 V7 precipitation, and drives the CREST distributed hydrological model using these precipitation products to evaluate their hydrological simulation effectiveness.

Judging from the results of precipitation accuracy assessment, the precipitation products have slightly lower estimate accuracy in the northern high latitude area and western mountains of the United States, while good accuracy is shown in the central, southern, and eastern regions. The precipitation products match NOAA-CPC-US precipitation from high to low in the order of Stage IV radar precipitation, followed by PERSIANN-CDR and GPM-IMERG, followed by PERSIANN and ERA5, and finally TRMM-3B42V7. From the perspective of the ability to capture different precipitation intensities, PERSIANN, PERSIANN-CDR, GPM-IMERG, and TRMM-3B42V7 satellite precipitation products are underestimated to varying degrees in the case of no rain and drizzle, but over-estimated in heavy rain and rainstorm. In the evaluation of hydrological simulation effectiveness, in general, NOAA-CPC-US and StageIV radar precipitation have a better effect in hydrological simulation of small watersheds, followed by PERSIANN-CDR, GPM-IMERG, PERSIANN, and ERA5’s runoff simulation capabilities, while TRMM- 3B42V7 has not ideal simulation effect. Based on the hydrological simulation scenarios set up in this paper, the above conclusions indicate that NOAA-CPC-US precipitation and StageIV radar precipitation in the United States can basically meet the needs of hydrological simulation in small watersheds. In the northern and western United States, the use of PERSIANN, PERSIANN-CDR, GPM-IMERG, ERA5 precipitation for hydrological simulation requires caution. TRMM-3B42V7 may be insufficient to support small-watershed hydrological simulation due to spatial resolution and sensor accuracy considerations.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the Special Project of First-Class University in western China for the Construction of State Key Laboratory of Ningxia University (No. GZXM2017001); The National Key R&D Project(No. 2018YFC0408104); The Scientific Research Project of Ningxia Colleges and Universities, China (No. NGY2020006); The University First-class Discipline Construction Project of Ningxia, China (No. NXYLXK2017A03); The Priority Research and Development Projects for Ningxia (No. 2020BEG03021).

References

Balsamo, G., Dutra, E., Alberge, C., Munier, S., Calvet, J.-C., Munoz-Sabater, J., & de Rosnay, P. (2018). ERA-5 and ERA-interim driven ISBA land surface model simulations: Which one performs better? Hydrology and Earth System Sciences, 22(6), 3515–3532. https://doi.org/10.5194/hess-22-3515-2018

Hong, Y., Hsu, K. L., Moradkhani, H., & Sorooshian, S. (2006). Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. Water Resources Research, 42, W08421. https://doi.org/10.1029/2005WR004398

Hou, A. Y., Kakar, R. K., Neck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., … Iguchi, T. (2014). The global precipitation measurement mission. Bulletin of the American Meteorological Society, 95(5), 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

Hu, W., She, D., Shao, M. A., Chun, K. P., & Si, B. (2015). Effects of initial soil water content and saturated hydraulic conductivity variability on small watershed runoff simulation using LISEM. Hydrological Sciences Journal, 60(6), 1137–1154. https://doi.org/10.1080/02626667.2014.903332

Huffman, G. J. (2016): The transition in multi-satellite products from TRMM to GPM (TMPA to IMERG) Technical Report. NASA.

Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2015). Integrated multi-satellite retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code, 612, 47. https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/TMPA-to-IMERG_transition.pdf; https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_doc_06.pdf

Islam, M. A. (2013). Evaluation of TRMM rainfall products for hydrological uses at different scales. Universitat Politècnica de Catalunya.

Jiang, S., Ren, L., Hong, Y., Yang, X., Ma, M., Zhang, Y., & Yuan, F. (2014). Improvement of multi-satellite real-time precipitation products for ensemble streamflow simulation in a middle latitude basin in South China. Water Resources Management, 28(8), 2259–2278. https://doi.org/10.1007/s11269-014-0612-4

Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M. (2012). Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. Journal of Hydrology, 452, 213–225. https://doi.org/10.1016/j.jhydrol.2012.05.055

Khan, S. I., Hong, Y., Wang, J., Yilmaz, K. K., Gourley, J. J., Adler, R. F., Brakenridge, G. R., Policelli, F., Habib, S., &
driving hydrological model for identifying flood events: A case study in the Xiangjiang River Basin, China. *Natural Hazards*, 87(3), 1489–1505. https://doi.org/10.1007/s11069-017-2836-0

Ying, X., & Chong-Hai, X. (2012). Preliminary assessment of simulations of climate changes over China by CMIP5 multi-models. *Atmospheric and Oceanic Science Letters*, 5(6), 489–494. https://doi.org/10.1080/16742834.2012.11447041

Zhao, Y., Xie, Q., Lu, Y., & Hu, B. (2017). Hydrologic evaluation of TRMM multisatellite precipitation analysis for Nanliu river basin in Humid Southwestern China. *Scientific Reports*, 7(1), 1–10. https://doi.org/10.1038/s41598-017-02704-1

Zhen, S., Shaofeng, J., Aifeng, L., Wenbin, Z., & Yan-chun, G. (2015). Assessment on precision of temperature simulated by the IPCC AR5 GCMs in China, 1996-2005. *Progress in Geography*, 34(10), 1229–1240. https://doi.org/10.18306/dlkxjz.2015.10.003