Water Balance Data Fusion Applied to River Basins in China

MSc thesis

Xinyue Yang

Dr. Gerrit Schoups Dr.Markus Hrachowitz Dr.Ronald van Nooyen

Water Management

Department of Civil Engineering and Geosciences

TU Delft

Table of contents

Abstract

To obtain and fully use long series of high-precision observations of water balance components, it is necessary to quantify and reduce random and systematic data errors. This thesis applies and evaluates a previously developed data fusion methodology for bias-correcting and noise-filtering remote sensing observations of water balance variables that results in a consistent set of estimates that close the water balance. The method combines monthly water balance constraints and probabilistic data models for each water balance variable (precipitation, evaporation, river discharge and water storage), and uses Markov chain Monte Carlo sampling and iterative smoothing to estimate data errors and water balance variables without the need for any ground-truth. The methodology is evaluated here by application to three river basins in China located in different climate zones (humid to semi-arid). The evaluation assesses (i) how the method performs in the case study basins and to what extent data error assumptions in the probabilistic data models are satisfied, (ii) how sensitive the results are to changes in the datasets, and (iii) whether the use of dataset ensembles rather than dataset pairs (as in the original method) changes the results and further improves the overall data fit. The findings for these three research questions are as follows. First, the posterior water balance estimate for humid and arid basins with average standard error of 6-10 mm/month for precipitation, 4-6 mm/month for evaporation, 8-14 mm/month for water storage. Significant increase in both precipitation and evaporation uncertainty during wet summer period and the data error assumption is violated for precipitation in Wuding basin. Second, the results of replacing precipitation are more sensitive for Baihe humid basin, with both precipitation and evaporation increased, while for Wuding basin, replacing evaporation datasets is more sensitive.Third, significant increases in likelihood value by fusing all precipitation and evaporation datasets, with posterior uncertainty of water balance components increased. And the data error assumption is satisfied for Wuding basin.

Therefore, the use of dataset ensembles as opposed to dataset pairs is recommended in further applications of the data fusion methodology. Additional applications may focus on applying the methodology across a wider range of river basins, using a wider range of ensemble datasets (including different GRACE solutions), as well as comparison of different data error models.

1 Introduction

1.1 Background and problem statement

Global climate change caused by the growing atmospheric concentration of carbon dioxide and other trace gases has become evident in recent years (Intergovernmental Panel in Climate Change IPCC, 1995; Houghton et al, 2001). One manifestation, for example, were the extremely dry conditions in in the Yangtze River basin in the summer of 2006 and the spring of 2011 (Zhang et al, 2015). Therefore, climate change has great implications for the terrestrial water cycle and water resources planning, which is expected to alter the timing and magnitude of runoff. As a result, it's necessary and important to obtain long time series and accurate estimates of water balance variables for water-stressed regions or data-poor areas, e.g. using hydrological and water balance modeling (Xiong et al, 1999).

Various satellite datasets are available for evaluating and calibrating hydrological and water balance models. A big challenge is how to quantify and reduce errors (bias and noises) in these datasets. For example, satellite sensors can provide nearly all-weather global precipitation estimates with near-global coverage, but generally with a long time lag and low spatial resolution. Hence, it's necessary to select, apply and correct multi-source precipitation estimates from remote sensing datasets to improve the input data for water balance modeling in water-stressed regions and data-poor areas. (Wu, 2020).

Due to their limited number and uneven spatial distribution, ground stations cannot represent spatial-temporal distributions and variations of large-scale precipitation, especially considering the complex variability in topography and climate across China. Currently, retrieval technology based on satellite-based remote sensing has become the main source for a wide range of spatial-temporal continuous precipitation information. Several studies have evaluated and compared the performance of various precipitation products across China. These products include not only satellite-based products such as IMERG, CHIRPS, SSEBop, and GLEAM etc, but also reanalysis products such as ERA5 and MERRA2, etc.(Tang et al, 2020)

In the study of Xu(2020), the spatial trend of the IMERG and FY2G QPE precipitation products was similar compared with the ground observations, showing a decreasing trend from the southeast to the northwest in China. IMERG had complete spatial coverage, while FY2G QPE had incomplete spatial coverage. The spatial distributions of the indicators showed that the indicators of IMERG and FY2G QPE in northwestern China and southwestern China were generally inferior to other regions except for individual areas, the hourly and daily scale indicators of FY2G QPE were better than the indicators of IMERG; IMERG generally overestimated the precipitation at meteorological scale. (Xu, 2020). IMERG had good accuracy performance on a monthly scale, but the accurate performance at a meteorological scale was unstable. However, IMERG had excellent spatial-temporal coverage and resolution. In addition, FY2G QPE showed excellent estimation accuracy and precipitation capture ability at the meteorological scale due to the fusion of meteorological ground observations.

In the study of Tang et al(2020), PCDR and CHIRPS had good performance during winter due to the advantage of infrared data, whereas CMORPH almost loses the capability of detecting precipitation occurrence in a cold climate. SM2RAIN performs relatively better in arid regions where soil moisture is seldom saturated. Reanalysis products are generally worse than most gauge-adjusted satellite products in China. ERA5 performs better than ERA-interim and MERRA2(Tang et al.2020). IMERG products do not perform well at high latitudes, high altitudes and in arid regions, especially in areas with low precipitation intensity arid regions. Kim et al.(2017) evaluated the IMERG dataset in different landscapes and seasons in East Asia, and the study showed that the accuracy of IMERG data was not good in a complex landscape, especially in areas along the Yellow River basin, which is anomalously overestimated, also the IMERG data performed worse in identifying rain-free days along the coastal area.

In the study of Kong et al(2017), the GPM IMERG and CGDPA both show accurate estimation of the spatial distribution pattern of precipitation in China, which has relatively small errors in the eastern area and relatively large errors in the western area. The IMERG product was closer to the ground-observed values, and had a high correlation with CGDPA. IMERG datasets are relatively stable over the year, while the RMSE of CMORPH and PERSIAN products in winter is large and fluctuates throughout the year.

In addition to precipitation, evapotranspiration (ET), which includes interception, vegetation transpiration, and soil evaporation, is also an important component of the energy and water balance at the surface. Estimates of evaporation are mainly based on model simulation. Various methods and products have been evaluated to estimate evaporation across China. Gao et al(2015) used the Penman-Monteith method to calculate and analyze the spatial and temporal variability of potential evapotranspiration from 1956 to 2000 on the surface of China and the factors. In the study of Tian Jing et al(2019), the researchers used the NOAH land surface process model to simulate the changes of surface hydrothermal process parameters on mainland China from 1986 to 2008 and analyze the spatial and temporal variation of surface evapotranspiration in China. He et al. used MODIS global evapotranspiration product (MOD16) to analyze the spatial and temporal patterns of land surface evapotranspiration in China from 2001 to 2010. However, due to the different calculation methods and factors of evapotranspiration, the difference between the estimates of surface evapotranspiration in China from different models is more than 1/3, and the simulation effects in different regions are also different. For example, the GLEAM product has high accuracy in the semi-arid grassland region but performs moderately in the semi-humid region. While the MOD16 data has high accuracy in the plain region, it cannot simulate well in the arid region(Yang et al,2019). Wang YJ et al.(2020) suggested that the high-temperature datasets provided by GLDAS may be overestimated due to the large simulated evapotranspiration values. Therefore, there are some limitations that the study may have if it’s based on a single-source of evapotranspiration data.

In China, most of the remote sensing evapotranspiration estimation studies are based on NOAA/AVHRR, EOS/MODIS, LANDSAT/TM, etc. as the main data source. FY3/VIRR data performs well and can reflect the distribution patterns and characteristics of evaporation in different landscapes. By comparing the surface evapotranspiration estimated by FY-3/VIRR

satellite inversion and the evapotranspiration estimated by MODIS data, the correlation of these estimated values is high (R>0.99) which means these two datasets have a good consistency.

Apart from the mentioned remote sensing datasets of the water balance component, water storage changes can be obtained from the GRACE(Gravity Recovery and Climate Experiment) dataset and is commonly used in water balance studies. Since China has complex terrain covering a range of climate and levels of human interventions, evaluation of the performance of different GRACE datasets in China has been done by Yao et al(2019), who analyzed relative uncertainties in GRACE-derived TWS changes from five solutions over mainland China based on a generalized three-cornered hat (TCH) method, including CSR, GFZ, GRGS, HUST, and JPL GRACE mascon solution. The results showed that compared to the monthly scale, the uncertainties of each solution were lower at the seasonal and annual scales. At the basin scale, except for the Yangtze River basin, CSR showed the lowest uncertainties for the 13 river basins over mainland China, while GRGS showed relatively large uncertainties. In addition, GRGS-based TWS showed larger variability than other GRACE solutions and two hydrological models (Global Land Data Assimilation System, GLDAS, and WaterGAP Global Hydrology Model, WGHM) in the temperate continental climate region; CSR and JPL were less affected by the surrounding hydrological conditions, climate settings, size and geometry of the basins. Compared with GLDAS and WGHM hydrological models, the GRACE datasets can have a better reflection on the changes of water storage in different climatic regions in mainland China, with the largest mean annual amplitude in the tropical monsoon climate region, followed by subtropical monsoon climate region, and the smallest mean annual amplitude in temperate continental climate region.

Many studies on multi-source data fusion with GRACE have been done in recent years. Döll et al calculated the rate of groundwater decay by fusing the GHMs model dataset with GRACE data, indicating that the two datasets had a great consistency, which means fusing data can have a better ability to analyze changes in water storage. Long Di et al(2017) analyzed the correlation between GRACE changes of water storage and meteorological data(e.g precipitation and temperature), and reconstructed a historical dataset from 1980 to 2002 by using artificial neural networks. (Long et al,2017) Long et al also used a three-cornered hat method to analyze the uncertainty of three types of data product, i.e. storage from GRACE, LSMs, and GHMs at basin-scale, and fused the GRACE dataset by using a Bayesian model based on the global hydrological model WGHM, which indicated that multi-source data fusion can improve the accuracy of remote sensing datasets. However, they found it may also lead to relatively large errors in different regions. It was suggested that fusing GRACE, GHMs, and LSMs datasets can improve the accuracy of monthly GRACE data and reduce data uncertainty.

Despite the multitude of remote sensing datasets on water balance components, there is still a big challenge in properly accounting for and quantifying the data errors in these datasets in the absence of a reference ground-truth datasets. Indeed, a common method to estimating the bias and random errors of each water balance component is to compare the dataset to a reference ground-truth dataset. For example, in the study of Massari and Maggioni(2019), satellite-based precipitation estimates were evaluated by rain gauge data as ground truth. Moreira et al.(2019) evaluated the uncertainty in precipitation and evaporation with observations from rain gauges and

eddy covariance flux towers, respectively. Another method is to create a reference dataset using the water balance equation with other water balance components assumed known. For example, Moreira et al.(2019) use remote sensing data to calculate terrestrial water storage changes (TWSC) as residuals in the simplified water balance equation. Although the reference data can be considered as in-situ observation or calculated from the water balance equation, the 'true' value could never be known or measured since no dataset or estimate is completely error-free. Traditional ground-based sensors observations, e.g stream gauges/rain gauges, are limited in capturing variability across a large area and therefore the dataset may lack representative and consistency, while spaceborne sensors observation, are suffering from uncertainties in converting electromagnetic signals into water balance variable estimates. (Schoups and Nasseri,2021)

Without requiring the 'true' value as reference, another alternative error estimation approach is to build an ensemble of datasets for a specific water balance variable and estimate errors based on variability across the ensemble or based on the three-cornered hat method (Massari et al.,2017). For example, Long et al.(2014) used the three-cornered hat method to estimate errors of ET without knowing the true ET value.

Goal of the current study is to bring together multi-source, multi-sensor data on water balance components (precipitation, evaporation, river discharge, and water storage changes), which includes data from ground-based sensors (rain gauges/stream-flow discharge gauges) and spaceborne sensors(satellite-based data on precipitation/evaporation/water storage), with the aim of estimating and correcting systematic and random errors in the data. For this a recently developed water balance data fusion methodology is used to fuse monthly water balance data, and quantify the estimated data errors and error-corrected water balance variables. For each water balance variable, probabilistic data error models are used in combination with monthly water balance constraints. (Schoups and Nasseri, 2021).

1.2 Research questions and Objectives

To fully benefit from space-borne sensors(satellite-based data on land surface evaporation and precipitation) and ground-based sensors(stream gauges/rain gauges) to obtain improved and accurate water balance variables datasets, the study aims to apply and evaluate a recently developed water balance data fusion methodology for different climatic and topographic river basins across China at monthly scale and to propose possible improvements to the methodology. The following research questions are investigated :

- How does the methodology perform in river basins across China under different climatic and topographic conditions? To what extent are data error assumptions in the probabilistic data models satisfied in the case study basins?
- How do the estimated results change for different remote sensing datasets?
- Does the use of dataset ensembles rather than dataset pairs (as in the original method) lead to a better characterization of prior uncertainty, and to what extent does it change the results and further improve the overall data fit?

To answer these research questions, the following objectives will be pursued:

1. Apply and evaluate the existing data fusion methodology of (Schoups and Nasseri, 2021) to quantify and reduce systematic and random errors in water balance variables of monthly basin-scale datasets across multi-climatic basins in the Yellow River and Yangtze River basins in China. The same (global) data sources as used in (Schoups and Nasseri, 2021) will be used in this step.

2. Reapply the data fusion methodology with alternative combinations of precipitation and evaporation datasets, and compare the results with results from the previous step.

3. Reapply the data fusion methodology with dataset ensembles rather than dataset pairs (as in the original method), and compare the results with results from the previous two steps. Typically, more than two precipitation (or evaporation) datasets are available, which can potentially lead to a better characterization of prior uncertainty, and a change in posterior estimates of the water balance variables.

1.3 Thesis Structure

The paper starts with the methodology used in this study, and introduces the Bayesian hierarchical model that fuses monthly water balance data and probabilistic data error models in chapter 2. Chapter 3 introduces the three river basins in China and datasets used in the study. Chapter 4 presents the results of using the methodology and evaluates data errors for all basins. This is followed by a discussion and summary of the findings.

2 Methodology: Water balance data fusion

To address the research question and to achieve the objectives outlined, this chapter focuses on the main steps of the methodology. The approach is based on a probabilistic water balance model that combines multi-source datasets, quantifies their systematic and random errors, and results in consistent estimates for all water variables that close the monthly water balance.

2.1 Probabilistic water balance Model

The terrestrial water budget of a river basin is composed of fluxes of precipitation (rain and snowfall), evapotranspiration (soil and canopy water evaporation, plant transpiration), water storage(wetlands, lakes, rivers), together with runoff(surface and subsurface flow) on the land surface. (Sheffield J et al. 2009). Their relation is given by equation 1.

$$S_t = S_{t-1} + P_t - E_t - Q_t \qquad (1)$$

Where the parameters, S_t and S_{t-1} are water storage(surface and subsurface) at the start and end of month t, P_t and E_t are average precipitation and actual evapotranspiration for the monthly basin scale and Q_t is river discharge at the outlet of the basin. Every water balance term is normalized and expressed in water depth units(mm).
From equation 1, the lateral groundwater flow into or out of the basin is neglected for simplicity, and so are the upstream inflows and inter-basin water transfers.

In principle, each estimated term can lead to water balance errors. As a result, the water balance does not close by ending up with too much or too little water. The underlying true water balance variables do lead to water balance closure while each dataset contains errors, including systematic(bias) and random(noise) errors. It's necessary to adjust the data values within a range depending on the magnitude of data errors closer to their "true" values, where data with larger errors could be adjusted more than data with smaller errors. These adjustments can be done automatically and simultaneously for all datasets and water balance variables. Since the magnitude of data errors is not known a priori, the parameters which quantify the magnitude of bias and noise can be treated as random variables with prior distributions. The resulting error model can be viewed as a Bayesian hierarchical model with two levels of uncertainty, one for error parameters and the other for water balance variables. (Schoups and Nasseri, 2021)

In this study, a probabilistic error model for each water balance component with error parameters will be given to quantify the magnitude of unknown systematic and random errors. With specified error parameters, data adjustments are done to close the water balance and compute the likelihood of given error parameters by using Kalman smoothing methodology. The optimized error parameters value can be obtained by computing the largest likelihood. (Schoups and Nasseri, 2021)

2.1.1 Precipitation error model

The precipitation error model is shown by equations 2 to 5, which states the two observed and true precipitation datasets.

$$m_{P,t} = (1-\omega_P)P_{obs1,t} + \omega_P P_{obs2,t} \quad (2)$$

$$S_{P,t} = max\left(\sigma_{P,t}, \frac{1}{2} r_P \left|P_{obs1,t} - P_{obs2,t}\right|\right) \quad (3)$$

$$P_t \sim N\left(m_{P,t}, S_{P,t}^2\right) \quad (4)$$

$$P_t \geq 0 \quad (5)$$

The first equation models bias by describing prior mean precipitation $m_{p,t}$ in month t, which is the weighted average of two observed datasets $P_{obs1,t}$ and $P_{obs2,t}$. The parameter ω_p is the weight, which takes on an unknown value between 0 and 1. Specifically, a logit-normal prior with location parameter μ = 0 and scale parameter σ = 1.4 is used to reflect prior uncertainty.

The second equation models random errors by prescribing prior standard deviation $S_{p,t}$ of precipitation in month t, based on the largest of the standard error of the first precipitation dataset $\sigma_{p,t}$ and the scaled absolute difference between the two precipitation datasets in each month t. r_p represents a scaling parameter, which takes an unknown value between 0 and 1. When r_p is equal to 1, the prior standard deviation is half the absolute difference value between two datasets, and when r_p is close or equal to 0, the two precipitation datasets can be in close agreement. In order to avoid a situation like small prior uncertainty, the value of $S_{p,t}$ is not allowed to be less than $\sigma_{p,t}$.

The last two equations in the precipitation error model assume true precipitation P_t in month t follows a truncated normal distribution. True precipitation P_t is considered to be non-negative. (Schoups and Nasseri, 2021)

2.1.2 Evaporation error model

The evaporation error model is shown by equation 6 to 9,which states the two observed and true evaporation datasets.

$$m_{E,t} = f_E\left[(1-\omega_E)E_{obs1,t} + \omega_E E_{obs2,t}\right] \quad (6)$$

$$S_{E,t} = max\left(0.1\, m_{E,t}, \frac{1}{2} r_E \left|E_{obs1,t} - E_{obs2,t}\right|\right) \quad (7)$$

$$E_t \sim N\left(m_{E,t}, S_{E,t}^2\right) \quad (8)$$

$$E_t \geq 0 \quad (9)$$

The model bias is described by the prior mean evaporation, which is a weighted average $E_{obs1,t}$ and $E_{obs2,t}$ multiplied by f_E. Parameter f_E represents a scaling factor that provides an additional degree of freedom, to e.g account for bias outside the range of the two datasets; parameter f_E has a log-normal prior with mode at 1 and a coefficient of variation CV of 50%. Random errors are modeled using the same approach as for precipitation. Parameter r_E controls to what extent prior uncertainty scales with the absolute difference between the two evaporation datasets. Parameters ω_E and r_E take on unknown values between 0 and 1. Specifically, flat logit-normal priors between 0 and 1 are used with location parameter μ = 0 and scale parameter σ = 1.4.

If the two datasets are in close agreement, a minimum relative error of 10% is assumed by setting

$S_{E,t}$= 0.1$m_{E,t}$. Also, the true evaporation E_t in month t is assumed as a random draw from a truncated normal distribution, where truncation at zero constrains evaporation to be non-negative. (Schoups and Nasseri, 2021)

2.1.3 River discharge error model

The river discharge error model is given by equations 10 to 13, which relates runoff data to underlying true discharge value Q. Q_{obs} is a monthly river discharge dataset assumed to be measured at the outlet of each basin, possibly with missing data.

$$\mathrm{m}_{Q,t} = N(Q_{obs,t}, v_{Q_{obs,t}}) \tag{10}$$

$$\mathrm{s}_{Q,t} = a_Q Q_{obs,t} + b_Q \tag{11}$$

$$Q_{\mathrm{t}} \sim N(m_{Q,t}, s^2_{Q,t}) \tag{12}$$

$$Q_{\mathrm{t}} \geq 0 \tag{13}$$

For months with observations, $v_{Qobs,t}$ can be set as 0, the first equation becomes $m_{Q,t}$ = $Q_{obs,t}$,which means the mean monthly river discharge is equal to the unbiased observation. For months with missing data, $Q_{obs,t}$ and $v_{Qobs,t}$ are set equal to the mean and variance of river discharge observed across the entire observation record for those months. This method only works when only a few observations are missing.

Random errors are described by $S_{Q,t}$ as a linear function of observed discharge, which assumes that observation random errors increase linearly with discharge, and a time-invariant parameter is added. Parameter a_Q has a log-normal prior with mode at 0.1(relative error of 10%) and CV of 1%, while b_Q has a log-normal prior with mode at 0.001 and CV of 1%. The monthly discharge Q is constrained to be non-negative. (Schoups and Nasseri, 2021)

2.1.4 Water storage error model

The water storage error model is given by equations 14 to 16. The observed water storage data and true storage values (note: both are actually storage anomalies) typically have a seasonal cycle, with possibly different amplitudes and phases, which motivates the following noisy sine wave error model for quantifying the differences between GRACE water storage and underlying true storage S_t.(Schoups and Nasseri, 2021)

$$\mathrm{m}_{S,t} = S_t + A\sin(\omega(\frac{t}{12} - \delta)) \tag{14}$$

$$\mathrm{s}_{S,t} = \sigma_S \tag{15}$$

$$S_{\mathrm{obs},t} \sim N(m_{s,t}, s^2_{S,t}) \tag{16}$$

In equation 14, A represents the amplitude(mm), ω represents frequency of a wave(radians per year), and δ is the phase in the year. The first equation describes the systematic differences in amplitude and phase between the observed GRACE dataset with true values utilizing time-invariant parameters A and ω. Parameter σ_S describes the magnitude of random errors for

basin-scale data.

The value of ω is fixed at 2π radians per year, yielding a sine wave with a period of 12 months. Parameters A, σ_S, and δ all have vague priors to reflect large prior uncertainty in the value of these parameters. Specifically, A has a log-normal prior with mode at 30 mm and a CV of 200%, os has a log-normal prior between 0 and 1 year with location parameter μ= 0 and scale parameter σ = 1.4.

Note that parameter δ represents the phase of errors and is not to be interpreted as a phase difference between the observed and true signals. (Schoups and Nasseri, 2021)

2.2 Water balance estimation

The precipitation, evaporation, river discharge, and water storage error models define a joint distribution over the data and all unknown variables, with a total of 10 unknown error parameters, as summarized in Table 1. In addition to the error parameters, all monthly water balance variables(S_0, P_t, E_t, Q_t, S_t) are also unknown. Here, S_0 is the initial basin water storage (anomaly) at the start of the first month. If N represents the number of months, then there are $4N + 1$ monthly water balance variables to be estimated. The joint distribution of the model can be written as $p(x, \theta, S_{obs})$ where x represents all the water balance variables, θ represents the vector of 10 parameters and S_{obs} represent the entire time series of storage observations.

Table 1 Error model parameters

	Parameter	Unit	Value	Note
Precipitation error model	ω_p	-	0-1	Interpolation weight for precipitation
	r_p	-	0-1	Random error scaling factor for precipitation
Evaporation error model	f_E	-	Around 1	Bias scaling factor for evaporation
	ω_E	-	0-1	Interpolation weight for evaporation
	r_E	-	0-1	Random error scaling factor for evaporation
River discharge error model	a_Q	-	0.1	/
	b_Q	mm	/	/
Water storage error model	A	mm	0-30	amplitude
	ω	1/yr	2π	frequency
	δ	yr	0-1	phase

2.2.1 Posterior distributions

The posterior of vector θ can be written in equation 17:

$$p(\theta \mid S_{obs}) \propto p(\theta) p(S_{obs} \mid \theta) \qquad (17)$$

Where $p(\theta)$ represents the prior distribution of the parameters, and $p(S_{obs}|\theta)$ represents the likelihood, which is obtained by computing the normalizing constant of the conditional water balance posterior $p(x|S_{obs}, \theta)$. The likelihood defines a scoring function for the error parameters, i.e. it quantifies to what extent a specific set of error parameter values results in a match between

storage predicted from the water balance and storage from observation (GRACE).

The joint posterior for water balance variables x can be written in equation 18:

$$\mathrm{p}(x \mid S_{\mathrm{obs}}) = \int p(x \mid S_{obs}, \theta)\, p(\theta \mid S_{obs})\, d\theta \qquad (18)$$

Where $p(x|S_{obs}, \theta) = p(x, S_{obs}|\theta)/p(S_{obs}|\theta)$ is the conditional posterior distribution of x, and the normalizing constant of this posterior is equal to the parameter likelihood function $p(S_{obs}|\theta)$ in Eq. 17.

Computing all the conditional posteriors $p(x, S_{obs}|\theta)$ is done with an iterative Kalman smoothing algorithm, as detailed in Schoups and Nasseri (2021).

3 Case Study

3.1 Study area

China can be divided into different sub-regions according to topographic and climatic conditions. Three typical river basins are chosen in China, as shown in figure 3.1. These three basins are selected because of their availability of sufficient river discharge data, under different climates and landscapes, and relatively large areas ($>10^4$ km^2).

The Wuding(WD) river basin has a continental arid and semi-arid climate and complex topography. In the middle reaches of the Yellow River basin in China, the terrain is mainly loess hills and ravines, where the terrain is high in the west and low in the east. the surface vegetation in the basin is low in cover. The precipitation is mainly concentrated in the flood season. The average multi-year precipitation is 369.9mm/year, decreasing from southeast to northwest, with uneven distribution of precipitation during the year, 65% of which is concentrated in July to September and consists of high-intensity precipitation.

Wuding River runoff is mainly recharged by precipitation and its groundwater, with an average runoff of 1.53 billion cubic meters per year. The desert area located in the northern part of the basin, due to strong ground leakage, groundwater recharge accounts for a larger proportion, generally up to 80% ~ 90%. Most of the basin are loess hills and ravines, mainly precipitation recharge, base flow accounts for about 30% of annual runoff.

The Beiluo(BL) basin has a semi-humid continental climate and is located in the warm temperate zone in the middle of the Yellow River basin, semi-arid area, with complex geological landscape in this basin, mostly hilly and Loess Plateau.

In the middle reaches of the Beiluo River and its tributary Hulu River, the vegetation of forest is well covered, the annual runoff coefficient is relatively low so the runoff depth is less than 30 mm, and the runoff depth and runoff coefficient are smaller than those in the upper and lower reaches of the Beiluo River. Due to shortage of surface water supply, a large number of water diversion and irrigation projects have been built, and groundwater and external water transfers(0.07 billion per year, about 2.25mm/month) are used to meet water demand, and groundwater recharge accounts for about 35% of the annual water supply. Precipitation is mainly concentrated in July-October and consists of heavy precipitation. (Liu et al, 2001)

The Baihe(BH) basin has a humid subtropical monsoon climate and is located in the upper reaches of the Hanjiang River in the Yangtze River basin. Most of its tributaries are short and steep. There are mountain ranges above 2000 m in the upper reaches of the Hanjiang River from north to south. Baihe basin is mainly mountainous with 62% of the basin covered by forest, while crops are generally distributed at altitudes of 300-1000 m. Annual precipitation in the upper Hanjiang River is 815 mm, and average annual runoff is 1.52 billion cubic meters. 70% of the precipitation is mainly concentrated from May to October. The annual runoff coefficient in the basin is highly variable, ranging from 0.37 - 0.86. The mountainous floods in the basin have short transit times

and large peak volumes.(Mao et al, 2008)

Precipitation increases from north to south for these three basins due to the prevailing monsoon climate, ranging from 409 mm to 815 mm. Characteristics of stream gauges located at the outlet of each basin are shown in table 1.

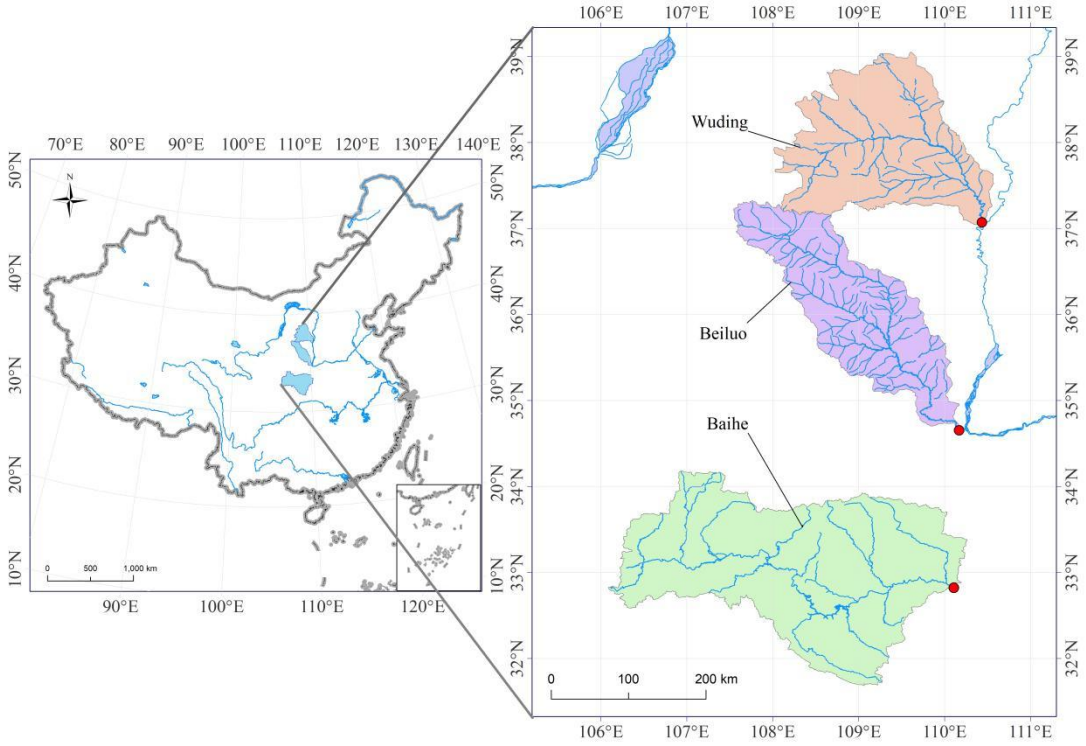


Figure 3.1 Map of China with the location of river basins and outlets

Table 2 River basin characteristics

ID	Region	Basin	ID	Stream gauge	Longitude (°E)	Latitude (°N)	Area ($\times 10^4$ km^2)	Elevation(m)	$\frac{E_p}{P}$*	$\frac{Q}{P}$*
1	Yellow River	Wuding	40801220	Baijiachuan	110.25	37.14	2.966	581~2116	0.97	0.87
2	Yellow River	Beiluo	41300700	Zhuangtou	109.50	35.02	2.56	297~1886	0.19	0.46
3	Yangtze River	Baihe	61801700	Baihe	110.105	32.83	5.911	~3400	0.28	0.87

* P, Q and E_p are average precipitation, river discharge and potential evaporation.

In addition to natural hydrological processes, there is also human water management in the basins. Figure 3.2 shows the area equipped for irrigation expressed as a percentage of total area in the three river basins, where the color ranges from light blue to dark blue, indicating higher dependence on irrigation (with larger percentage). The Baihe basin has a higher demand for irrigation water than the Beiluo and Wuding River basins.

Figure 3.3 shows areas irrigated with groundwater expressed as a percentage of total area equipped for irrigation. The trend of irrigation water demand in the three basins from south to north is mainly increasing, with groundwater irrigation rates of 31.94, 33.53 and 41.01

respectively. The groundwater recharge demand should be relatively higher, which also explains to a certain extent that the soil water content of Beiluo and Wuding basins show more obvious decreasing trend, but the Baihe basin has no obvious changes.

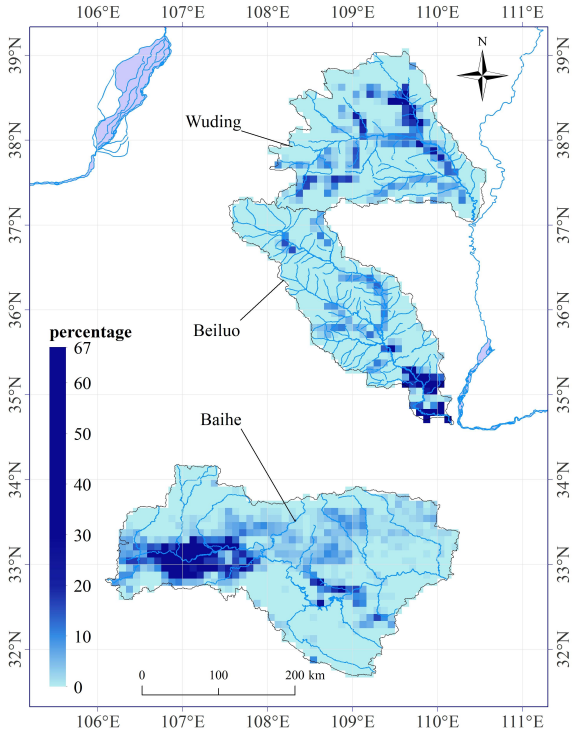


Figure 3.2 Map of Irrigation Areas (area equipped for irrigation expressed as a percentage of total area) for three basins.

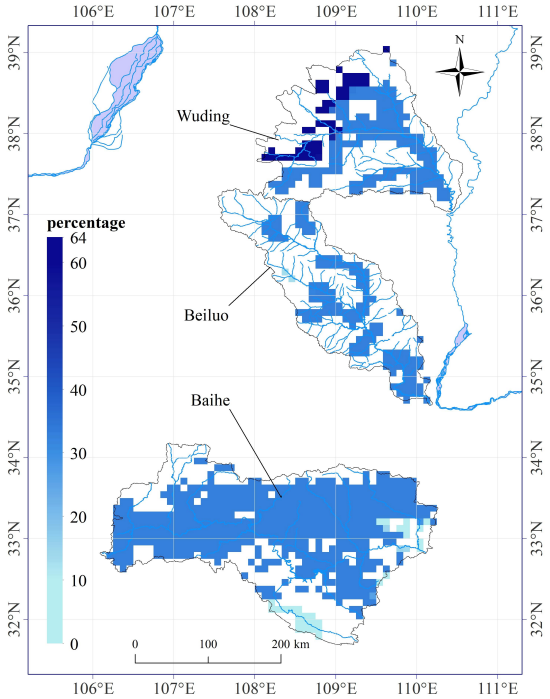


Figure 3.3 Map of Irrigation Areas (area irrigated with groundwater expressed as a percentage of total area equipped for irrigation) for three basins.

3.2 Remote sensing datasets descriptions

In this section, available remote sensing datasets are introduced, their time duration and spatial-temporal resolutions are given in Table 3.

3.2.1 Precipitation

Six commonly used satellite datasets available in the study area are shown in Table 3.

The monthly IMERG precipitation datasets merges satellite-based estimates with the GPCC rain gauge dataset, which belongs to GPM products, combined with the DMSP F17/18 and GCOM-W1, GCOM-W1 NOAA 18/19, Met Op B/C Met Op B/C, and many other satellites to produce data with spatial resolution 0.1°, time resolution 0.5 hour global coverage. IMERG precipitation products can be divided into three classes according to the data production time frame, which is delayed by about 4 hours Early-Run and delay of about 12 hours Late-Run, delay of about 3.5 months Final-Run. The quality of precipitation data is improved sequentially with the three levels of the precipitation dataset. (Huffman et al. 2019). In this paper, the final run product is used since it includes monthly-scale ground station observation.

The CHIRPS (climate hazards group infrared precipitation with station data) is a precipitation product with spatial coverage from 50N° to 50S°, 180W° to 180E°, and spatial resolution of 0.05°.Chirps is precipitation product produced by fusing multiple sources of datasets, including CHPclim (Monthly precipitation climatology) based on precipitation station data from FAO(food and agriculture organization) and GHCN (global historical climatology network), CPC (climate prediction center and NCDC), and NCDC (NOAA national climate data center). (NOAA national climate data center) thermal infrared data, and TRMM 3B42 rainfall data of V7 version, etc.

The CMORPH dataset (Joyce et al., 2004) combines microwave-based precipitation estimates with infrared imagery to produce 8 km (at the equator) fields every half hour, with 0.25 degree data every 3 hours available back to 2002. The infrared images are used to propagate the data in time between available retrievals of relatively high-quality passive microwave estimates, while morphing the shape and intensity of the precipitation features through interpolation. (Sheffield J et al.,2009)

The SM2RAIN-CCI is a global daily precipitation dataset that provides data for 1998-2015 at a spatial resolution of 0.25°. This precipitation product is developed by the European Space Agency's (ESA) climate change program. The algorithm is a new technique that estimates precipitation directly from soil moisture measurements through soil moisture balance equations(Brocca et al, 2014)

Persiann-CDR dataset provides precipitation data from 1983 to present, with a latitude range of 60° S to 60° N, with spatial resolution of 0.25° in daily scale. It is based on GridSat-B1 infrared satellite data using an artificial neural network algorithm (PERSIANN). The artificial neural network algorithm (PERSIANN) was used to generate the precipitation estimates. The training of the PERSIANN algorithm is done using the NCEP (National Center for Environmental Prediction)

StageIV radar precipitation data, then precipitation product is adjusted by using the Global Precipitation Climatology Project (GPCP) monthly precipitation product (GPCP v2.2)(Feng, et al, 2021)

ERA5 is a reanalysis product base on estimate using microwave images on polar-orbiting satellites and infrared images on geostationary satellites as a monthly mean value of precipitation in global scale.and is developed by the Global precipitation climatology project(GPCP). The GPCP daily product is tied to GPCC indirectly via its calibration with GPCP monthly product. Here the monthly precipitation datasets from 1979 to present at a spatial resolution of 0.25° is used.

3.2.2 Evapotranspiration

Four available evaporation satellite datasets are shown in Table3. The SSEBop datasets is based on the simplified surface energy balance(SSEBop) model from the period 2003 to the present(Senay et al, 2012) with a unique parameterization for operational applications. It combines ET fractions generated from remotely sensed MODIS thermal imagery, acquired every 8 days with reference ET using a thermal index approach. SSEBop uses Penman-Monteith for potential evaporation and estimate actual evaporation based on a surface energy balance and remotely sensed land surface temperature. (Schoups and Nasseri, 2021)

The GLEAM dataset uses Priestley-Taylor for potential evaporation and estimates actual evaporation as a function of microwave vegetation optical depth and soil moisture, in combination with a root-zone water balance, developed to estimate terrestrial evaporate fluxes at a spatial resolution of 0.25. It separates the terrestrial components of ET into canopy transpiration, soil and open water evaporation, interception loss, and sublimation.

The Mod16 dataset is based on the Penman-Monteith equation and considers the surface energy partitioning and atmospheric drivers on ET at 8-day and monthly time scales. Both GLEAM and Mod16 are based on 8-day averages for computing uncertainties, and on a monthly time-scale to compute water balance at global-scale.

Except precipitation, ERA5 can predict a wide range of atmospheric parameters, including atmospheric temperature, pressure and wind at different altitudes, as well as important parameters such as soil water content and evaporation. Here evaporation data with spatial resolution of 0.25° is used in monthly scale.

3.2.3 Terrestrial water storage

The GRACE(Gravity Recovery and Climate Experiment)mission was launched in 2002. It observes temporal variations of Earth's gravitational potential, with the main purpose of mapping variations in the Earth's gravitational field by measuring the distance between two orbiting satellites(Landerer and Swenson, 2012). After atmospheric and oceanic effects are accounted for, the remaining signal on monthly to inter annual timescales is mostly related to variations of terrestrial water storage (TWS). (Landerer and Swenson, 2012) Five different research centers process GRACE data: CSR, JPL, GFZ, GRGS, and HUST.For the basins studied in this paper, research reveals that the uncertainty of CSR is smaller than other models in mainland China, and

the uncertainty of CSR and JPL is less affected by the type of climate, hydrological characteristics, and the area and land shape of the basins. (Yao et al.,2019) A summary of the remote sensing datasets used to assess water storage (anomalies) is presented in Table 3. And the original datasets is selected in table 4 as used in Schoups and Nasseri 2021.

Table 3. Summary of remote sensing datasets available in this study for closing water balance in China

Dataset	Full name of the dataset	Resolution	Period	Reference
Precipitation datasets				
IMERG	Integrated multi-satellite retrievals for GPM final run V06B	0.1°/1month	2006-present	Huffman et al.(2019)
CHIRPS	climate Hazards group infrared precipitation with stations	0.05°/1month	1981-present	Funk et al.(2015)
SM2RAIN	SM2RAIN based on ESA climate change initiative(CCI)	0.25°/1d	1998-2015	Ciabatta et al.(2018)
CMORPH	climate prediction center(CPC) MORPHing technique bias corrected(CRT)	0.25°/1d	1998-present	Joyce et al.(2004)
T3B42	TRMM multi-satellite precipitation Analysis(TMPA) 3B42 V7	0.25°/1month	1998-2016	Huffman et al.(2007)
PCDR	PERSIANN Climate data record	0.25°/1d	1983-present	Ashouri et al.(2015)
MSWEP	Multi-source weighted-ensemble precipitation	0.1°/1month	1979-present	Beck et al.(2017)
Evaporation datasets				
GLEAM	Global Land Evaporation – the Amsterdam Method ET product	0.25°	2003-2017	Miralles et al.(2011);Moreira et al.(2019)
SSEBop V4	SSEB	0.01°	2000-present	Senay et al.(2020)
MOD-16	MODIS Global Evapotranspiration Project	500m	2000-2014	Mu et al.(2011);Moreura et al.(2019)
GLDAS Noah	Global Land Data Assimilation System	0.25˚	1948-present	Rodell et al.(2004)
Terrestrial water storage datasets(GRACE)				
CSR	Center for Space Research	1°	2002-2017	Feng et al(2021)
JPL	Jet Propulsion Laboratory	1°	2002-2017	Feng et al(2021)
GFZ	German Research Center for Geoscience	1°	2002-2017	Feng et al(2021)
GRGS	Groupe de Recherches de geodesie spatiale	1°	2002-2017	Yao et al(2019)
HUST	Huazhong University of Science and Technology	1°	2002-2017	Yao et al(2019);Zhao H et al.(2019)

Table 4 Monthly water balance data (as used in Schoups and Nasseri 2021)

Variable	Symbol	Data source	Resolution	Reference
Precipitation	P_{obs1}	GPM IMERG Final 06B	0.1°	Huffman et al.(2019)
	P_{obs2}	CHIRPS v2.0	0.05°	Funk et al.(2015)
Evaporation	E_{obs1}	SSEBop v5	0.01°	Senay et al.(2020)
	E_{obs2}	GLEAM v3.5b	0.25°	Miralles et al.(2011);Moreira et al.(2019)
River discharge	Q_{obs}	stream gauges	Basin	Changjiang Water Resources Commission
Water storage	S_{obs}	GRACE JPL mascon RL06v02	3°	Feng et al(2021);Wiese et al(2018)

3.3 Ground observation

3.3.1 River discharge measurement

Time series of monthly river discharge in the Yangtze River basin(Baihe Basin) were obtained from the national hydrological yearbook by the Yangtze River Water Resources Committee, Ministry of Water Resource (MWR) for the years 2000 to 2015. River discharge data for the Yellow River basin(Wuding and Beiluo Basin) were obtained from the national hydrological yearbook by the Yellow River Conservancy Commission, MWR. Only stream gauges at the outlet of each basin are used. The location of stream gauges can be seen in Figure 3.1 and basic information on these stream gauges is shown in Table 2.

4 Results

The chapter introduces respectively the detailed results of all three basins in this study and evaluates how results are affected when changing one of the datasets. Section 4.1 uses the same remote sensing datasets as used in Schoups and Nasseri, i.e. IMERG GPM and CHIRPS provide monthly precipitation datasets, while SSEBop and GLEAM are used for evaporation datasets. Next, section 4.2 evaluates how the results and estimated water balance variables change when different datasets are used. Finally, section 4.3 tests whether using more than two datasets for each variable (precipitation, evaporation) leads to better water balance estimates, in terms of data fit (likelihood) and consistency with the data error assumptions made in the model. In each case, GRACE JPL provides the water storage data.

4.1 Water balance data fusion for three basins in China

4.1.1 Baihe Basin

Baihe basin is the most humid region among the three studied basins in this study. Before presenting results of the data fusion methodology, we first assess to what extent the water balance does not close when using the original datasets. Calculated water storage and observed water storage are shown in figure 4.1. Here, total water storage is calculated via the monthly water balance equation using observed precipitation, evaporation, and discharge as inputs. The water balance equation(19) is used to compute water budget imbalance. (Tang Long et al, 2022) However, evaluating water balance closure can improve the understanding of data uncertainty of water budget components.

$$P - ET - Q - \Delta TWS = \varepsilon \qquad (19)$$

Where P is precipitation, ET is evaporation, Q is runoff and ΔTWS is terrestrial water storage variation, ε is the water budget imbalance. For Baihe basin, the average monthly water budget imbalance is about -8.93 mm computed from equation(19). The water imbalance of precipitation is 11.59% compare this to Tang et al. (2022) where 79% of catchments were found to have a water budget imbalance of less than 20% of precipitation). Clearly, the water balance does not close in Baihe basin.

From figure 4.2, IMERG tends to have higher precipitation than CHIRPS during wet summer months. And some study across China illustrates for CHIRPS the overestimation of monthly precipitation of low intensity and underestimation of monthly precipitation of high intensity. (Ren YJ et al, 2019)

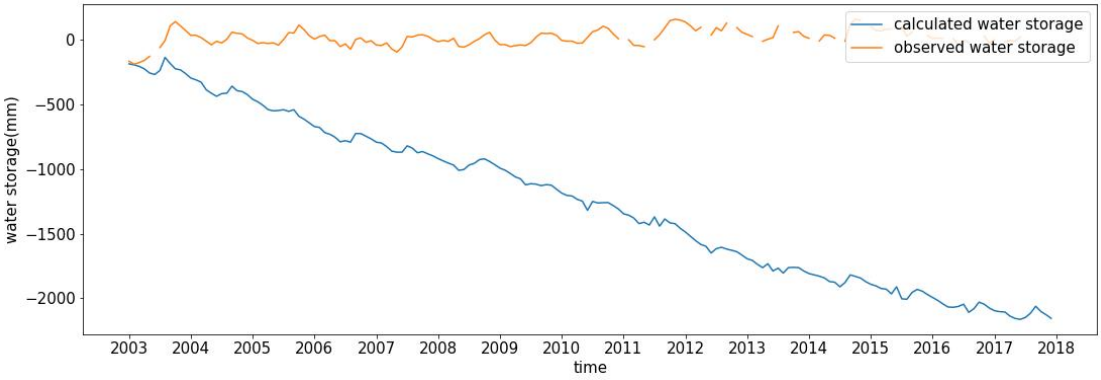


Figure 4.1 Calculated and observed water storage for Baihe basin. Values are in mm/month

Two different remote sensing evaporation datasets are used in this section, GLEAM, and SSEBop. With the dataset of GLEAM, the Priestley-Taylor method is used for estimating actual evaporation as a function of microwave vegetation optical depth and soil moisture with root-zone water balance. On the other hand, SSEBop is based on a simplified land surface energy balance model(SSEB), by using meteorological data to apply in Penman-Monteith for potential evaporation and estimate actual evaporation based on surface energy balance and remotely sensed land surface temperature. For the Baihe basin in this study, the SSEBop actual evaporation tends to have a higher value than potential evaporation from GLEAM during summer, and SSEBop evaporation tends to have a higher value than GLEAM evaporation during wet summer months (Figure 4.2), while in dry cold winter, the GLEAM data tend to have a higher value than SSEBop. These differences result in larger prior uncertainty in evaporation during summer.

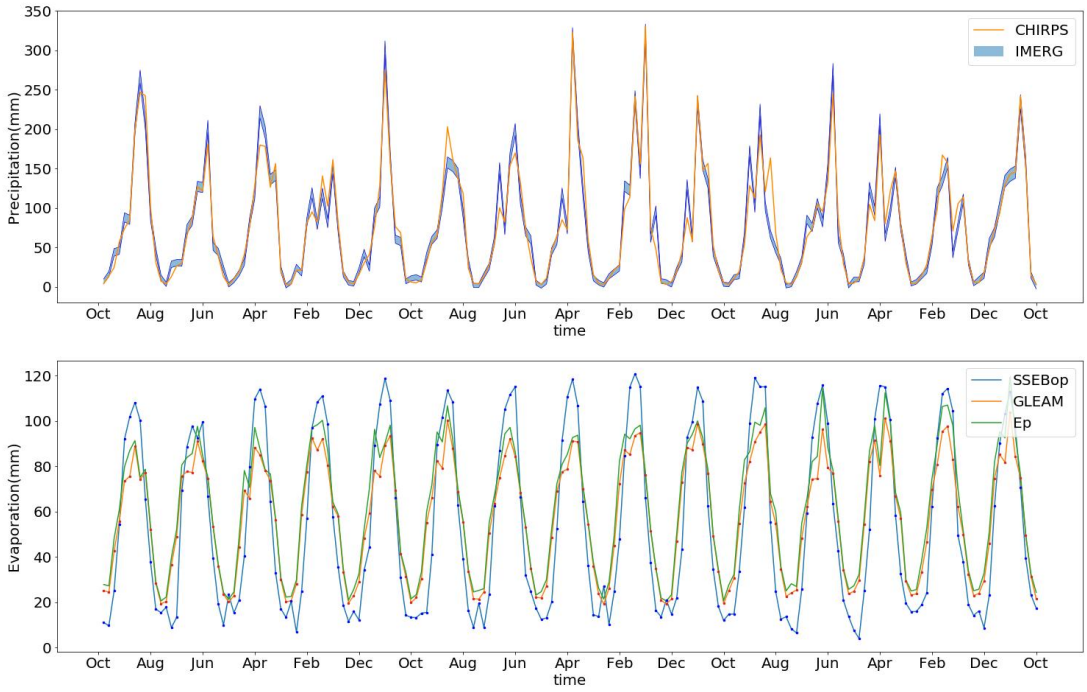


Figure 4.2 Monthly precipitation and evaporation data for Baihe Basin from the year of 2003 to 2017. Potential evaporation from the GLEAM dataset is shown as Ep.

Next, results are presented of applying the data fusion methodology to Baihe basin. The water balance posteriors for Baihe basin are shown in figure 4.3 and the error parameter posteriors are shown in figure 4.4. From figure 4.3, inferred precipitation tends to follow more closely the IMERG GPM than CHIRPS data, as confirmed by the estimated value for error parameter $\omega_p(0.23)$ which indicates that IMERG GPM receives a greater weight for this basin than CHIRPS. The posterior uncertainty in precipitation is not significantly different from the prior uncertainty in precipitation.

The posterior uncertainty in evaporation as shown in figure 4.3, is slightly different from prior uncertainty in evaporation. And the inferred evaporation tends to follow GLEAM than SSEBop since the weighted error parameter ω_E is 0.31, which indicates that SSEBop has a greater weight than GLEAM. The scaling factor f_E is 0.81 which means additional bias outside the range has been considered. When f_E is around 1, there is no additional bias. From figure 4.3, the posterior uncertainty is largest during wet summer seasons when the difference between the two datasets are also the largest.

River discharge of Baihe basin has smaller posterior uncertainty bands than uncertainty of precipitation and evaporation. The relative error a_Q assumed here is 0.1, i.e., 10%, and b_Q can be assumed negligible, as a result, the posterior uncertainty is similar to the prior uncertainty of river discharge. Note that the precipitation in the upper reaches of the Han River basin was significantly higher in 2003, and a 20-year flood occurred in the Han River basin in 2011.(Changjiang Water Resources Commission)

From the bottom plot in figure 4.3, it is shown that the inferred water storage uncertainty band follows the GRACE observations, and there are several large increases in amplitude in posteriors. For example in the year 2003, after the end of the summer, persistent high temperature and dry weather occurred in the Han River basin with severe drought, and the precipitation in the autumn flood season was large, intense, and concentrated, resulting in a large variation in water storage. (Changjiang Water Resources Commission) Compared to other water balance variables, the error parameters A,δ,σ_s all have relatively small uncertainty and are well defined. The inferred value for σ_S is about 20mm.

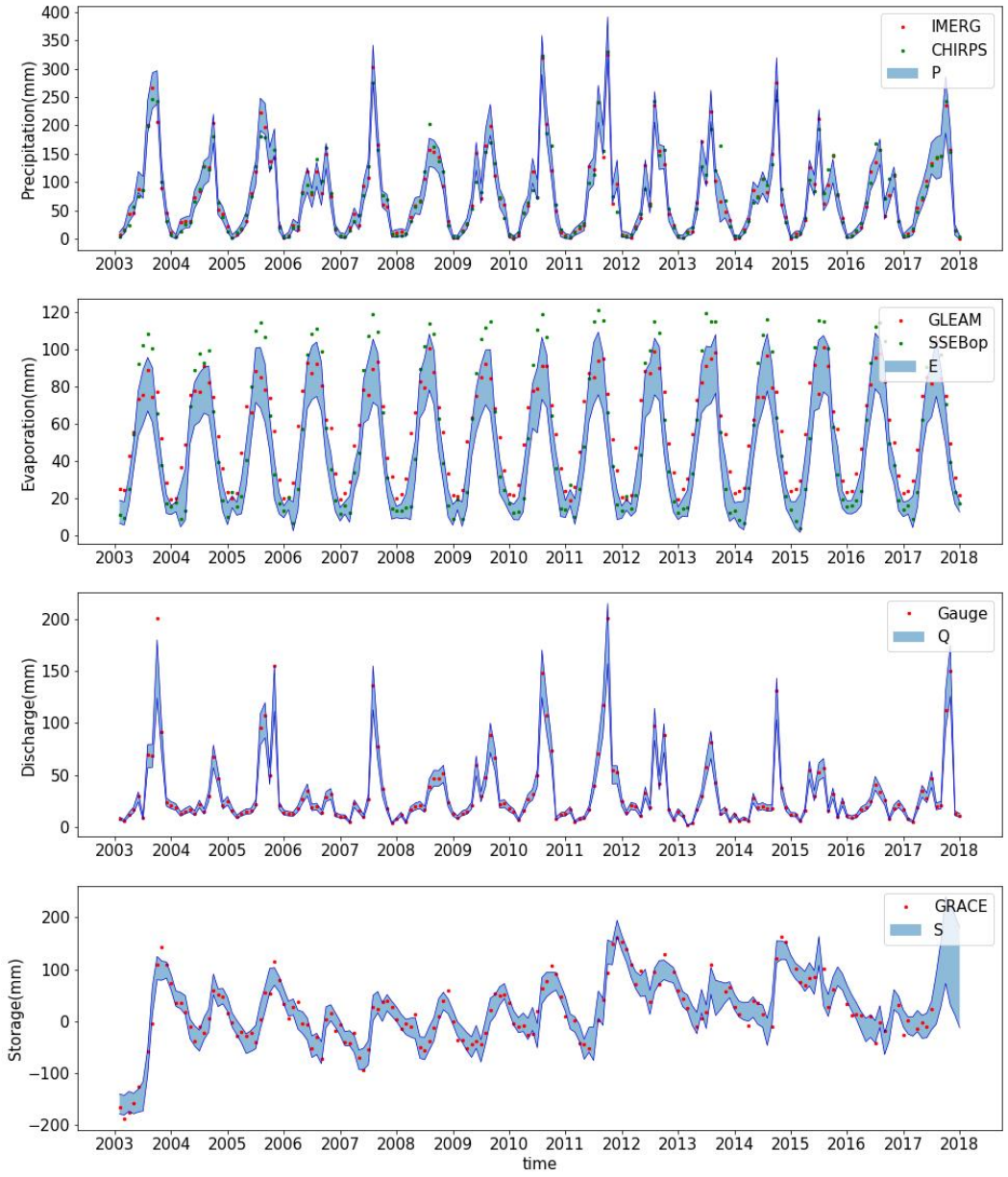


Figure 4.3 Monthly water balance estimates for Baihe Basin, shown as 90% posterior uncertainty bands. All values are in mm/month.

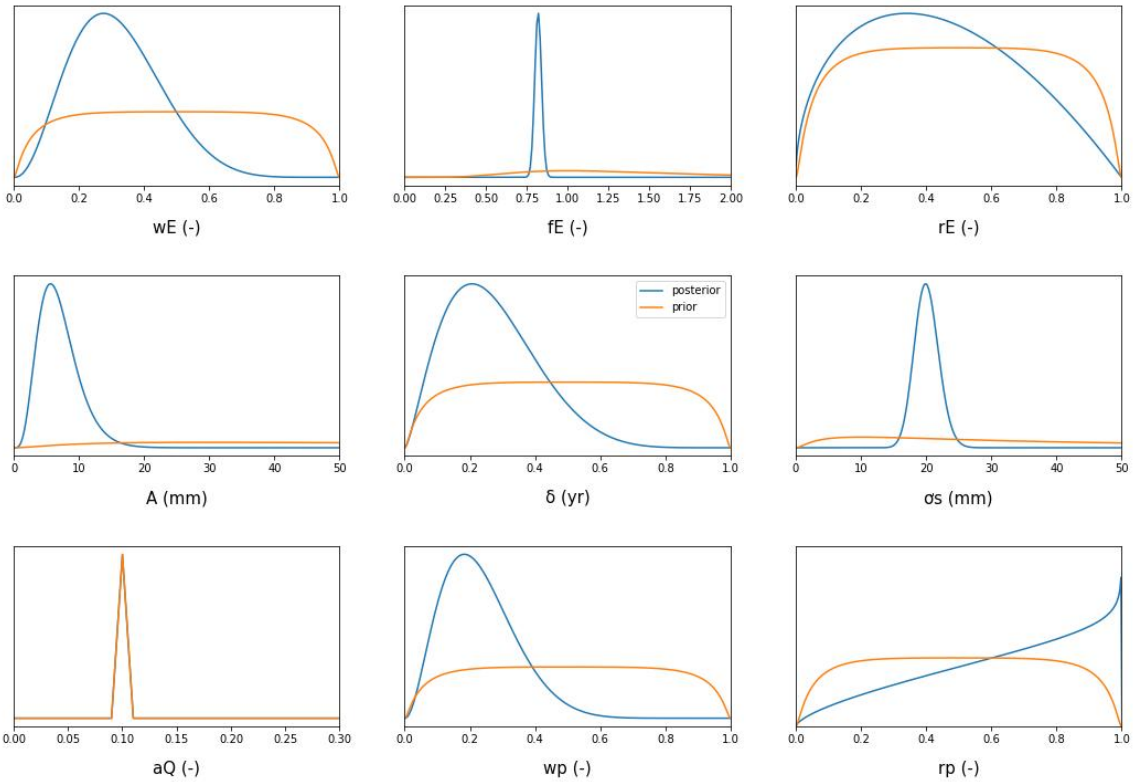


Figure 4.4 Prior and posterior density of error parameters for Baihe Basin.

4.1.2 Beiluo Basin

Beiluo basin can be considered as a transition zone among three regions which is semi-humid. In figure 4.5, the observed water storage shows a downward trend while calculated water storage increases. Similarly, the average monthly water imbalance budget for Beiluo basin is about 7.42 mm from equation(19), which is about 14.8% of precipitation. It is clear that the water balance also does not close for Beiluo basin.

As seen in Figure 4.6, IMERG has larger peaks during wet summers compared to CHIRPS, and also similarly has larger values than CHIRPS in dry winters.

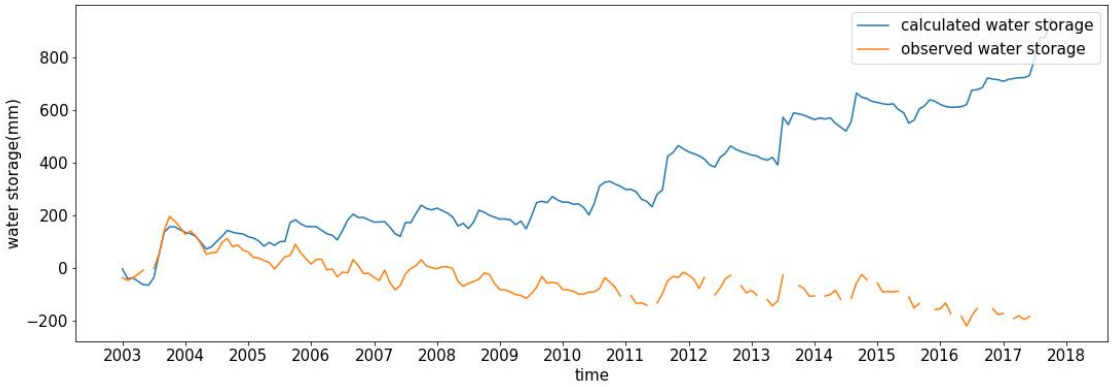


Figure 4.5 The calculated water storage and observed water storage for Beiluo basin. Values are in mm/month

Similarly, SSEBop and GLEAM are chosen as the first and second evaporation dataset for the Beiluo River basin, with discrepancies between the two that are qualitatively similar as in the Baihe basin. As seen in Figure 4.7, in the first five years, the potential evaporation of GLEAM remained smaller than the actual evaporation of SSEBop in the wet summer months, especially near the peak. In contrast, for years 2008 and 2013, the difference between the potential evaporation E_p and the actual evaporation from SSEBop decreases evidently, then in the last three years, the potential evaporation from GLEAM is already greater than the actual evaporation measured by SSEBop. Similarly, for the actual evaporation, the evaporation values of GLEAM are smaller than those of SSEBop during wet summers, and in contrast, the evaporation values of GLEAM are larger than SSEBop in winter months.

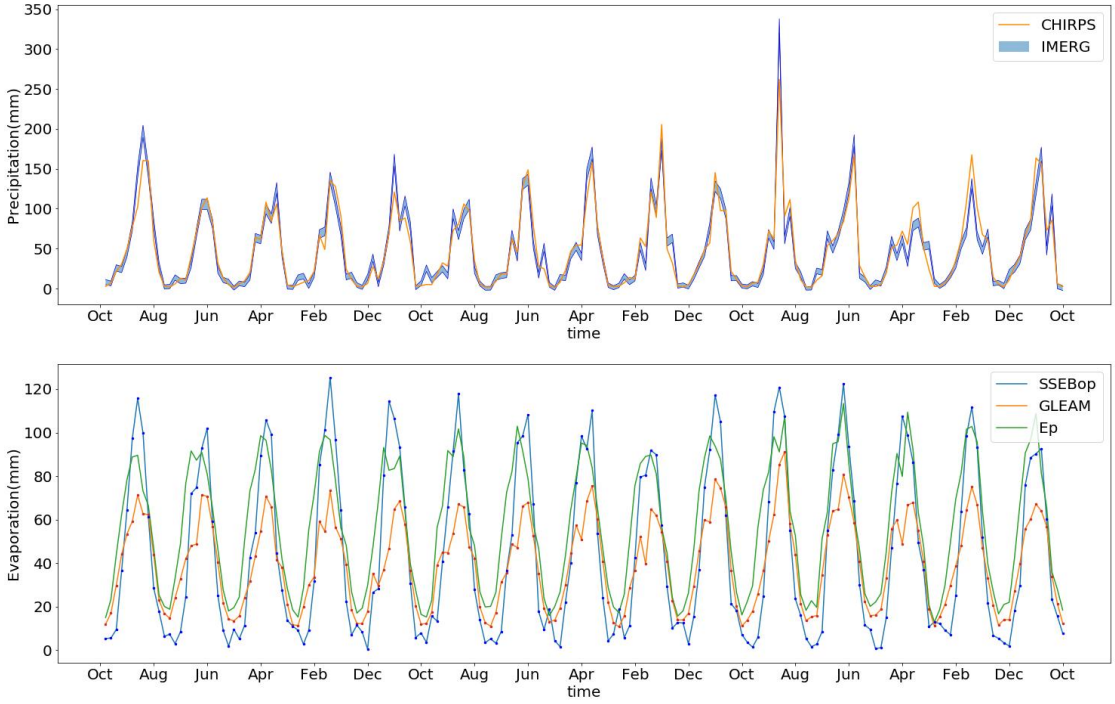


Figure 4.6 Monthly precipitation and evaporation data for Beiluo Basin from the year of 2003 to 2017. Potential evaporation from the GLEAM dataset is shown as Ep.

Water balance posteriors for Beiluo Basin are shown in Figure 4.7 and error parameter posteriors for Beiluo Basin are shown in Figure 4.8. For the inferred precipitation, unlike the Baihe basin, IMERG and CHIRPS have nearly equal weight with a weight error parameter ω_p of 0.54. Also, the posterior uncertainty in precipitation is not largely different from prior uncertainty shown in Figure 4.7.

As for the posterior uncertainty in evaporation for Beiluo Basin, the inferred evaporation tends to follow GLEAM instead of SSEBop with the weighted error parameter ω_E as 0.58, and the bias can be neglected as f_E is around 1. Similar to Baihe basin, the posterior uncertainty increases during wet summer seasons when the differences between the two datasets are largest.

From the bottom one of Fig 4.7, the inferred water storage still follows the GRACE observations but shows a downward trend in the Beiluo basin. The three water storage error parameters (A, δ, σ_s) have less posterior uncertainty than in Baihe basin and the inferred value of σ_s is around 14 mm.

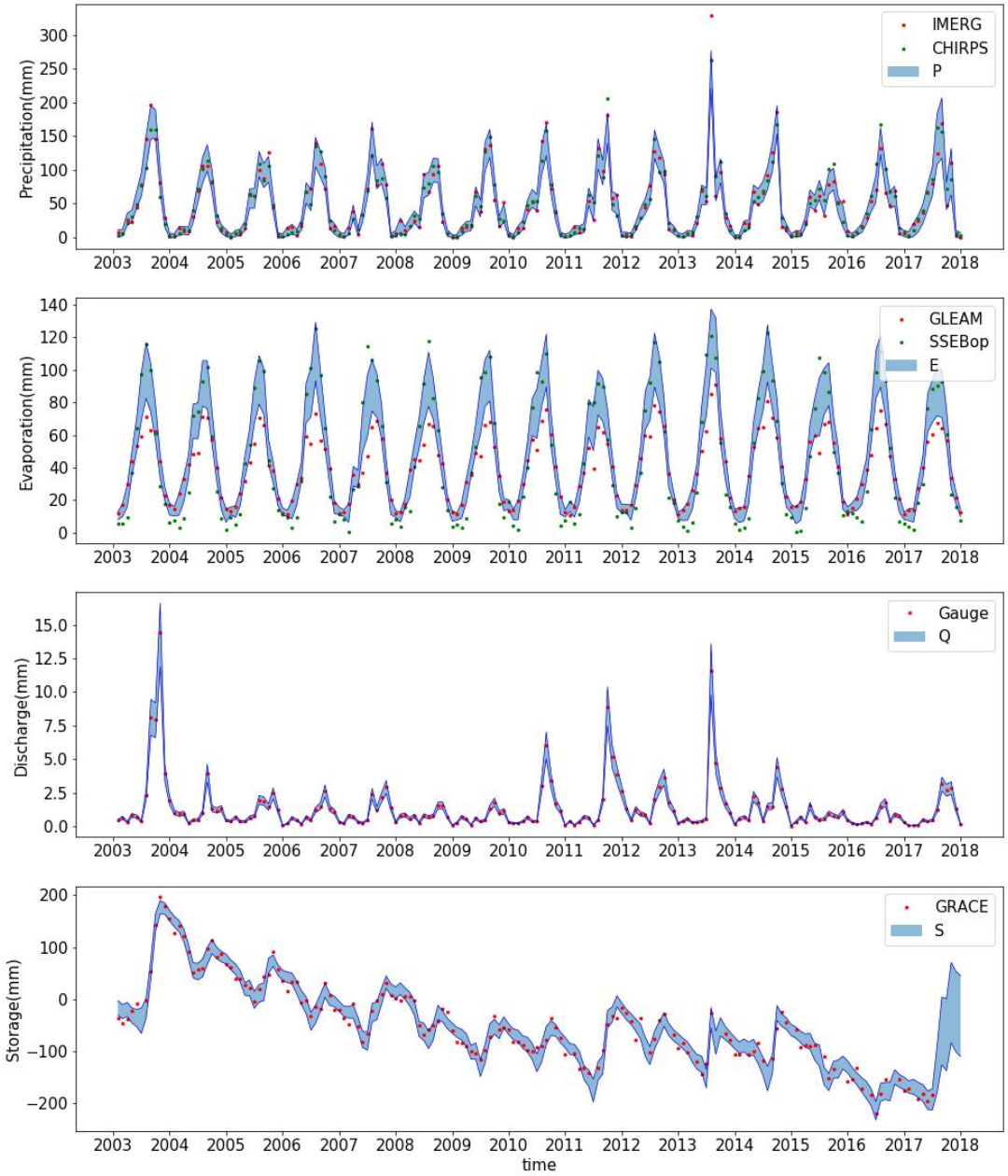


Figure 4.7 Monthly water balance estimates for Beiluo Basin from the year of 2003 to 2017. shown as 90% posterior uncertainty bands. all values are in unit of mm/month

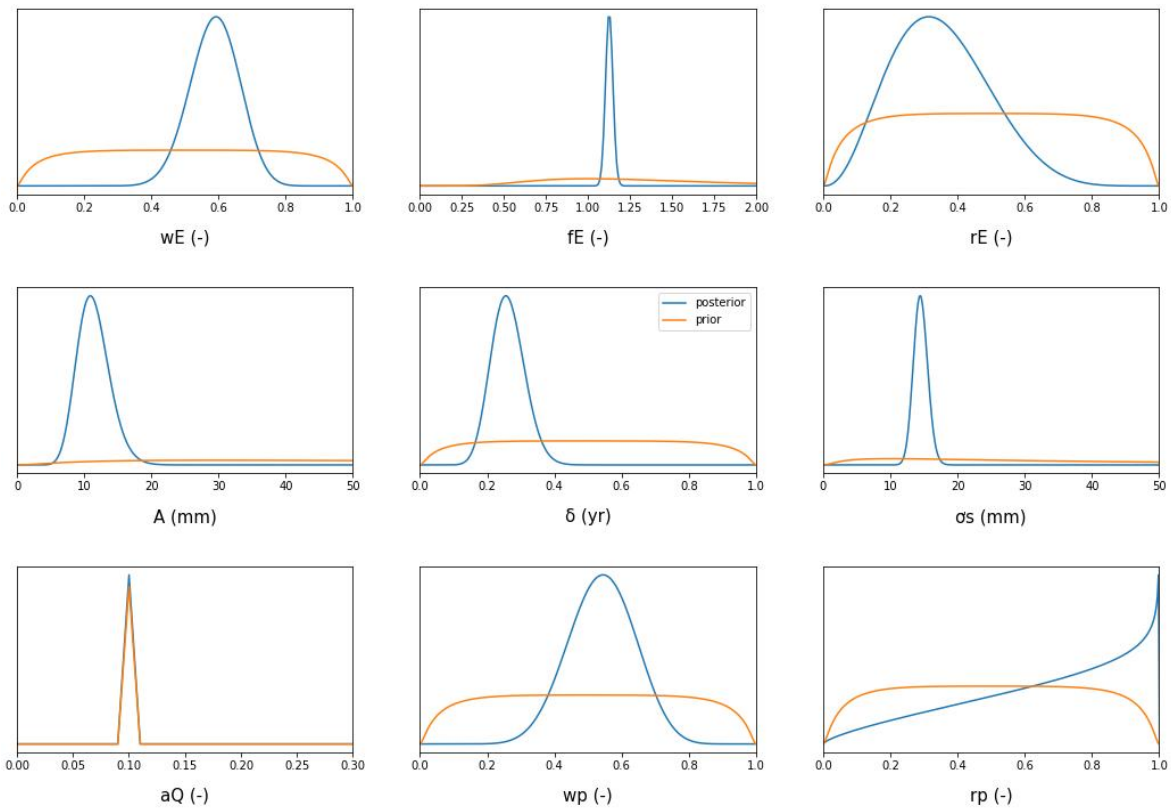


Figure 4.8 Prior and posterior density of error parameters for Beiluo Basin.

4.1.3 Wuding Basin

Wuding Basin, located in an arid climate zone of China, is the most arid region among the three basins in the study, which leads to a significant difference between the Wuding River basin and the two other basins. Figure 4.9 shows the calculated water storage and observed water storage in Wuding basin. The observed water storage shows downward trend and calculated water storage remains generally stable despite fluctuations, followed by an increasing trend starting around 2010. The average monthly water imbalance budget is 2.73 mm, which equals 7.5% of precipitation. Clearly, the water balance data for Wuding basin do not close the water balance either. As seen in Figure 4.10, IMERG precipitation peak values tend to be lower than those of CHIRPS.

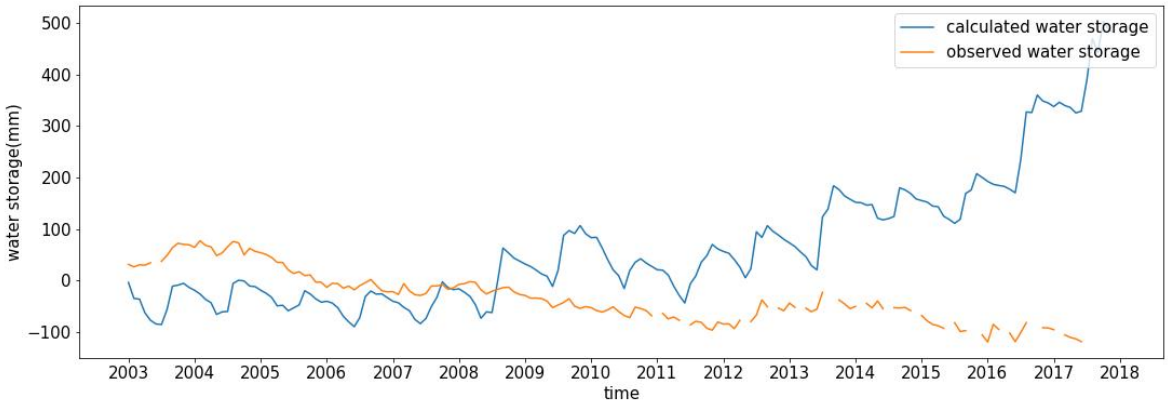


Figure 4.9 Monthly calculated water storage and observed water storage for Wuding basin. Values are in mm/month

As for evaporation, unlike the previous two basins, the potential evaporation E_p for GLEAM in the Wuding River basin is greater than the actual evaporation for both SSEBop and GLEAM, due to the difference in the datasets caused by the different basin climates. In wet summer months, the value of SSEBop evaporation is larger than that of GLEAM, and in contrast, in dry winter, GLEAM values are larger than evaporation of SSEBop.

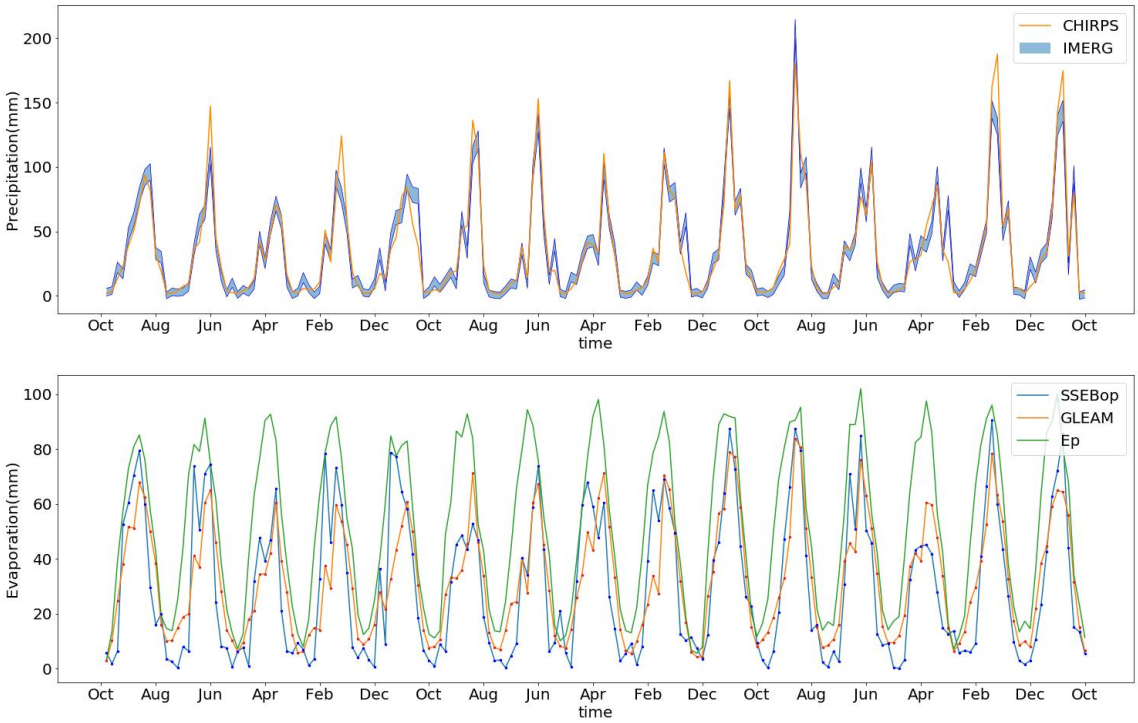


Figure 4.10 Monthly precipitation and evaporation data for Wuding Basin from 2003 to 2017. potential evaporation from the GLEAM dataset is shown as Ep.

Water balance posteriors for Wuding Basin are shown in Figure 4.11 and error parameter posteriors for Wuding Basin are shown in Figure 4.12. For the inferred precipitation in Wuding Basin, also the posterior uncertainty shows little difference with the prior uncertainty in figure 4.11. Since the posterior uncertainty band could not cover peak values during summer for both precipitation datasets, the posterior violates the prior assumption that true precipitation lies between the two datasets. Hence these two datasets can not reflect the basin well, which will be further explored in section 4.2. CHIRPS tends to have greater weight in this basin than IMERG due to parameter ω_p being equal to 0.59, which is completely different from that of the humid Baihe Basin. Especially, both r_p and ω_p have well defined posterior distributions compared to the vague priors in Figure 4.12.

Similar to previous basins, inferred evaporation tends to follow GLEAM instead of SSEBop, with estimated value of weight parameter ω_E equal to about 0.8, which is the largest value among the three basins, indicating that in the more arid region the GLEAM datasets may be more reliable.

The bias also can be neglected since f_E is equal to 1, and the posterior uncertainty increases during wet summer months when differences between the two datasets are largest.

As the Wuding River is an arid basin, it has the smallest flow magnitude of the three basins, with a maximum monthly average discharge of only about 10 mm from 2003 to 2018, and with an assumed 10% relative error, the prior and posterior uncertainty is quite small.

From the last row of figure 4.11, the inferred water storage dynamics does not quite follow the GRACE observations, suggesting there are potential leakage errors (bias) in the GRACE observations for this basin. All three storage error parameters have well-defined posterior distributions compared to their vague priors (Fig. 4.12).

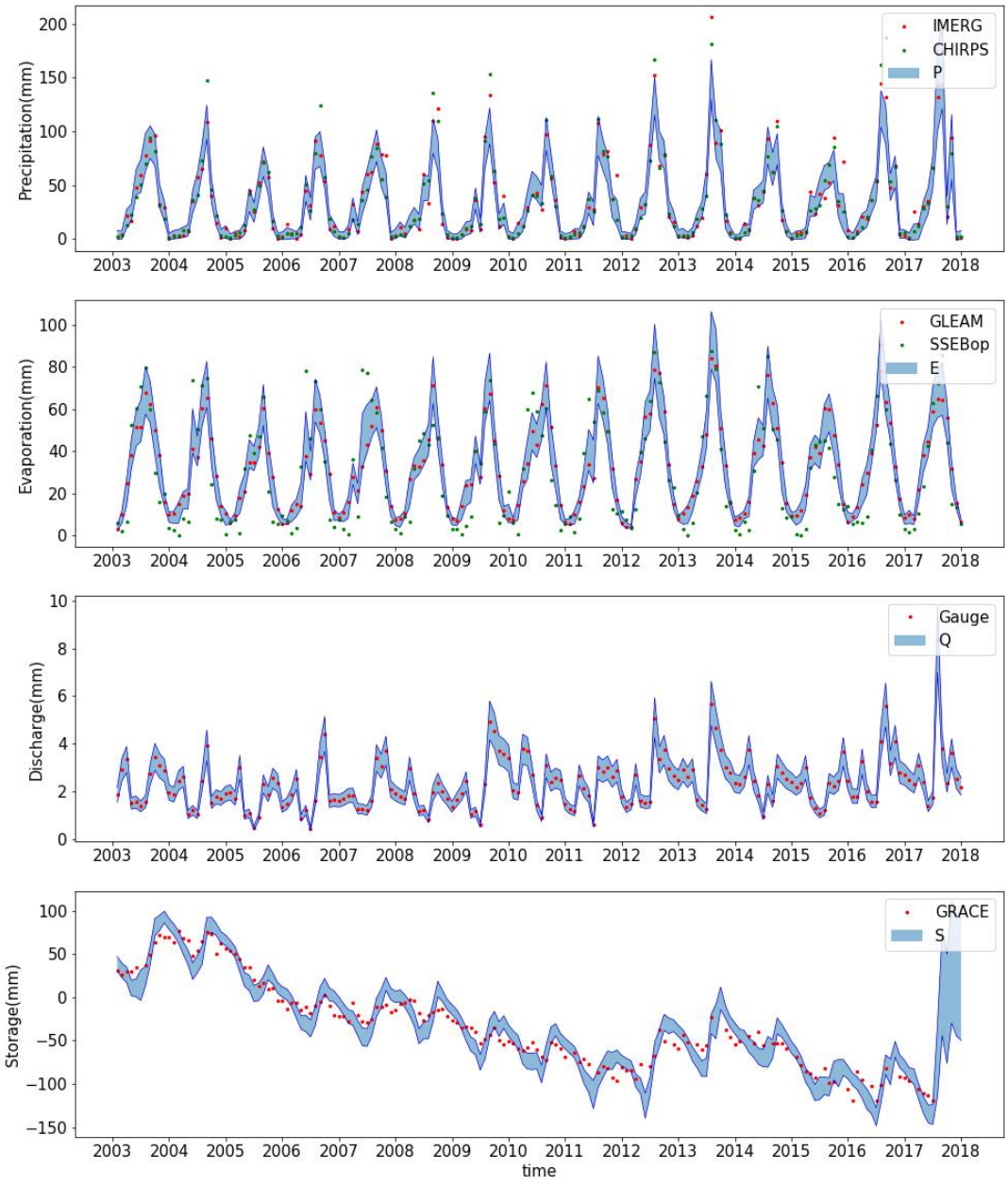


Figure 4.11 Monthly water balance estimates for Wuding Basin. shown as 90% posterior

uncertainty bands. All values are in mm/month

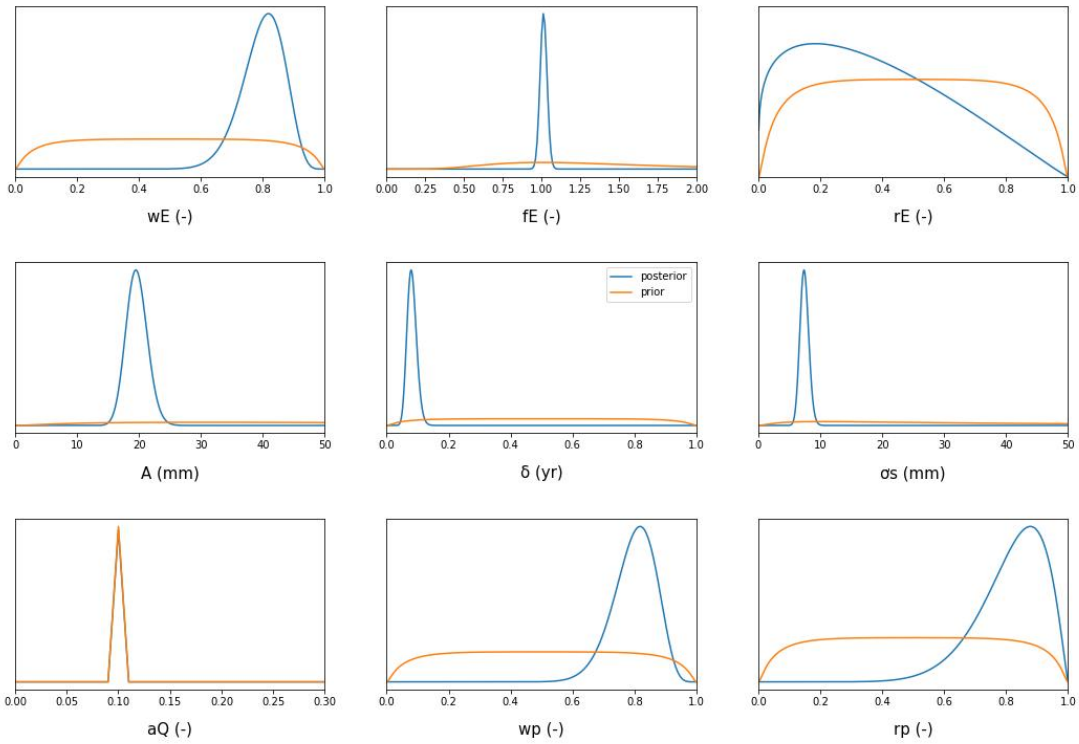


Figure 4.12 Prior and posterior density of error parameters for Wuding Basin

4.1.4 Analysis

Figure 4.13 illustrates posterior plots and posterior predictive distributions (i.e. posteriors for predicting S_{obs}) with GRACE data for all three basins. The posterior predictive distribution is obtained by using the inferred posterior mean of S_t in each month and applying it to the sine wave model (Eq.14) to get predictive distribution for the corresponding observation in that month. From the figure of Baihe and Beiluo, there is no significant change in phase between posteriors and GRACE observation data. While comparing with the posterior for S, the posterior predictive distribution S_{obs} have smaller amplitude in these two basins.

In the Wuding basin, as can be seen from the third line of figure 4.13, the inferred storage posteriors are shifted earlier than GRACE posterior predictive distribution and it's clear that with larger amplitude, the observed GRACE datasets do not fit with other water balance observations to get water balance closure. The model accounts for leakage error in GRACE data by shifting amplitude and phase of the storage dynamics. These leakage errors could be real, but they could also be an artefact caused by unaccounted for errors in the precipitation or evaporation data. We already saw in fig. 4.11 that the precipitation data are not well fit in this basin, and that the posterior lies below both precipitation datasets thereby violating assumptions of the precipitation error model. Repeating the analysis with different precipitation and evaporation datasets (see section 4.2) will provide additional insight.

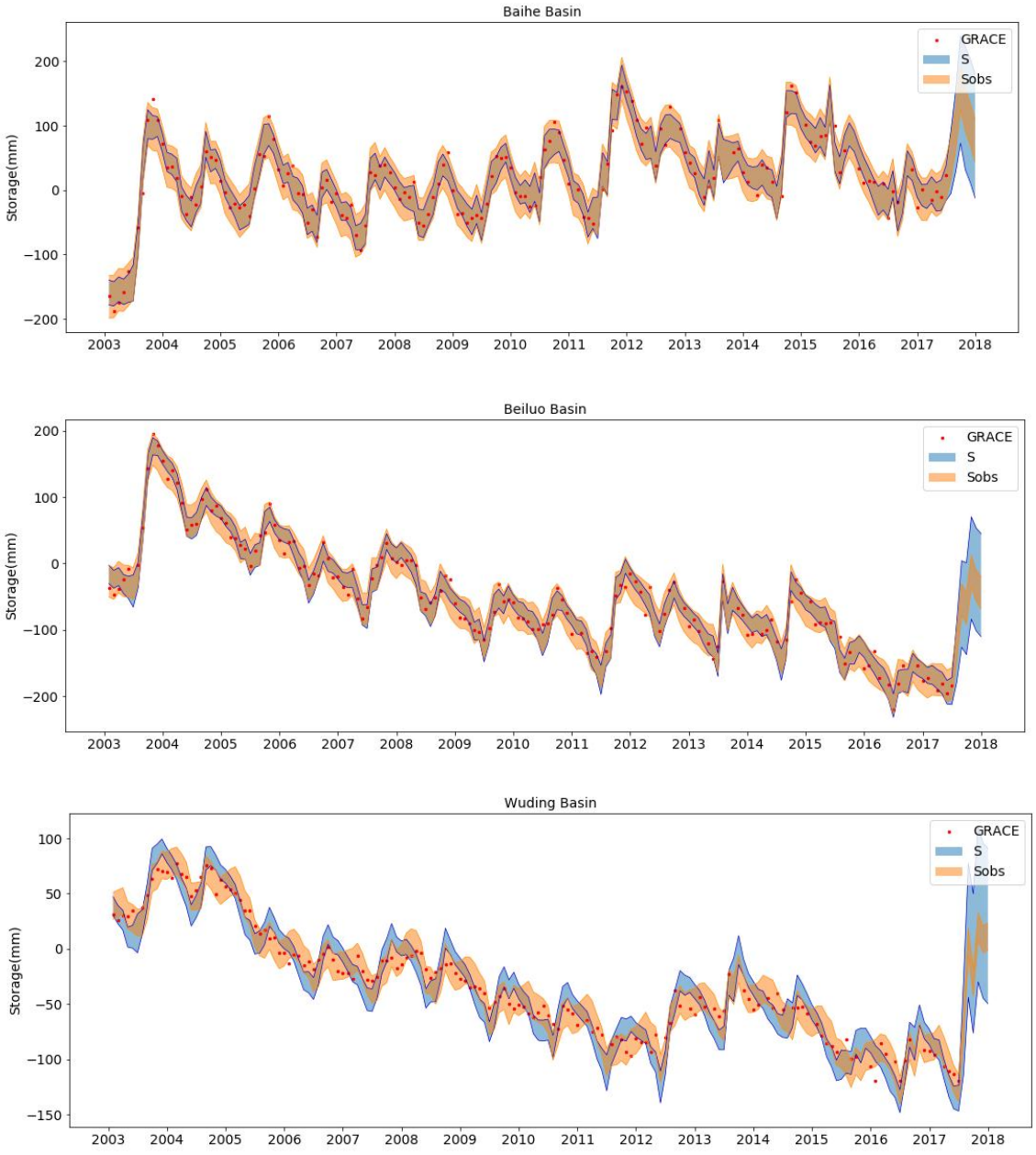


Figure 4.13 90% uncertainty bands of storage posteriors(S) and GRACE posterior predictive distribution(Sobs) with GRACE data (red dots) for Baihe, Beiluo, and Wuding basin.

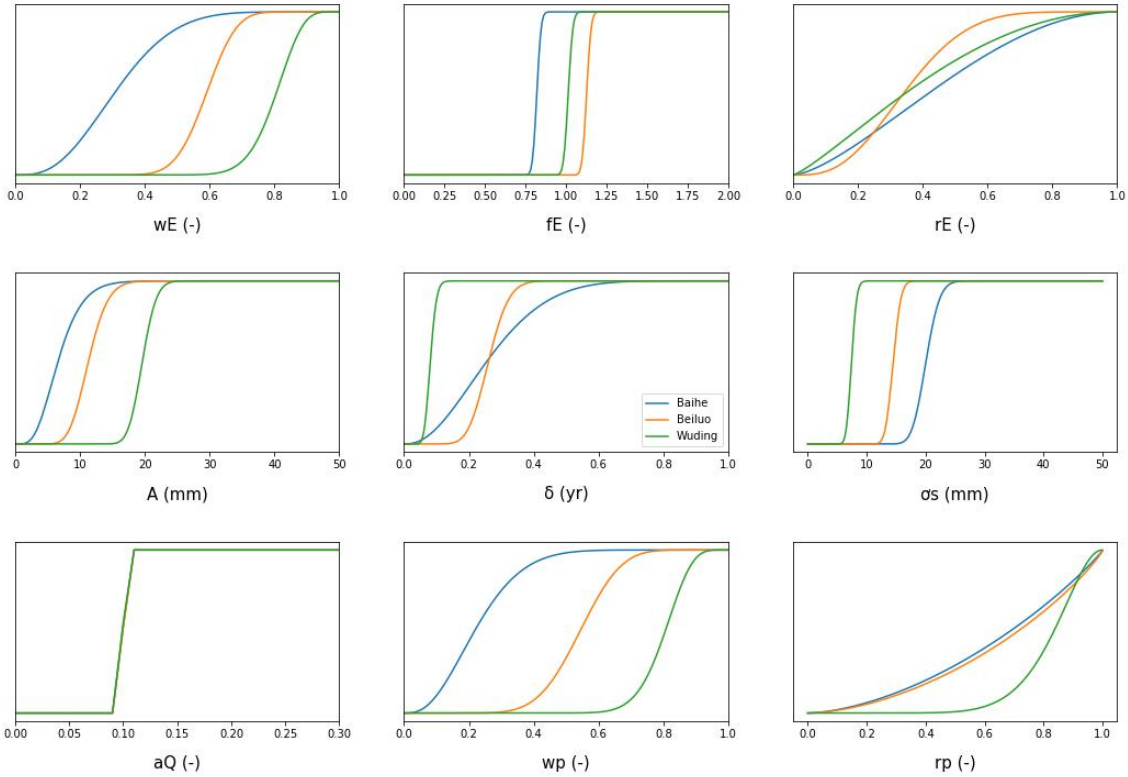


Figure 4.14 posterior error parameter distribution for Baihe, Beiluo, and Wuding basin.

Figure 4.14 shows the error parameter posterior distributions for the three basins. As shown in the third row, in the humid Baihe basin IMERG fits better with water balance data than CHIRPS while in the semi-humid Beiluo basin, IMERG and CHIRPS have nearly equal weight of importance in fitting data. For the arid Wuding basin, CHIRPS does have a better performance than IMERG since ω_p is greater than 0.5. As for scaling parameter r_p, which is observed in all three basins, have an upward trend toward a value of 1.

The first row shows the posterior distributions of the evaporation error parameters. In the humid Baihe basin, SSEBop receives more weight than the GLEAM dataset ($\omega_E < 0.5$) while in the semi-humid Beiluo basin, SSEBop and GLEAM provide nearly equal weight in fitting the data. In the arid Wuding basin, GLEAM provides a better fit. The inferred value for bias parameter f_E ranges from 0.8 to 1.2, indicating that in more arid regions, the scaling factor f_E tends to get close to 1, and hence there is no bias outside the range of two datasets.With all inferred values of r_E smaller than 0.5, the prior evaporation uncertainty reduces in all basins.

The second row in the figure shows the storage error parameter distributions which are all well identified, judging from the relatively sharp posteriors. The standard deviation σ_S in the humid Baihe basin is around 20 mm. In the semi-humid Beiluo basin, the σ_S is close to 15mm and in the arid Wuding basin, the standard deviation is less than 10 mm.

4.2 Effect of using different remote sensing datasets

This section presents a detailed analysis of how the results are affected by changing some of the remote sensing datasets in the data fusion methodology. The appendix contains posterior plots for all the basins and Wuding basin's result are highlighted then followed by a summary of results of the other two basins.

The results in this thesis are based on three kinds of remote sensing datasets, precipitation, evaporation and water storage. Improved results can be obtained by simply changing one or two of the remote sensing data, and three cases are given below, i.e replacing the precipitation remote sensing data, replacing the evaporation remote sensing datasets, or replacing both the precipitation and evaporation remote sensing data. This thesis does not consider replacing different GRACE solutions, since the average RMS values of water storage product uncertainty produced by JPL-mascon, CSR and GFZ are around 25 mm/month in China, which is pretty close to each other, especially the water storage change uncertainty of JPL RL05 and GFZ RL05 is about only 0.1mm in Yangtze River basin, however can be neglected.(Yao et al, 2019)

4.2.1 Effect of using different Precipitation datasets

Figure 4.15 shows the six precipitation remote sensing data time series from 2003 to 2018. Two remote sensing datasets with relatively larger differences, ERA5 precipitation and PCDR (Persiann Climate Data Record), are selected to replace the original IMERG and CHIRPS precipitation remote sensing datasets for all catchments. From equation (3) it is known that the larger absolute difference between the precipitation datasets, the larger prior standard deviation and the greater the prior uncertainty will be, and greater the probability that the posterior distribution will fall within the prior distribution, and hence a smaller chance that the posterior will violate the prior assumptions of the precipitation error model (as happened for Wuding basin in the previous section).

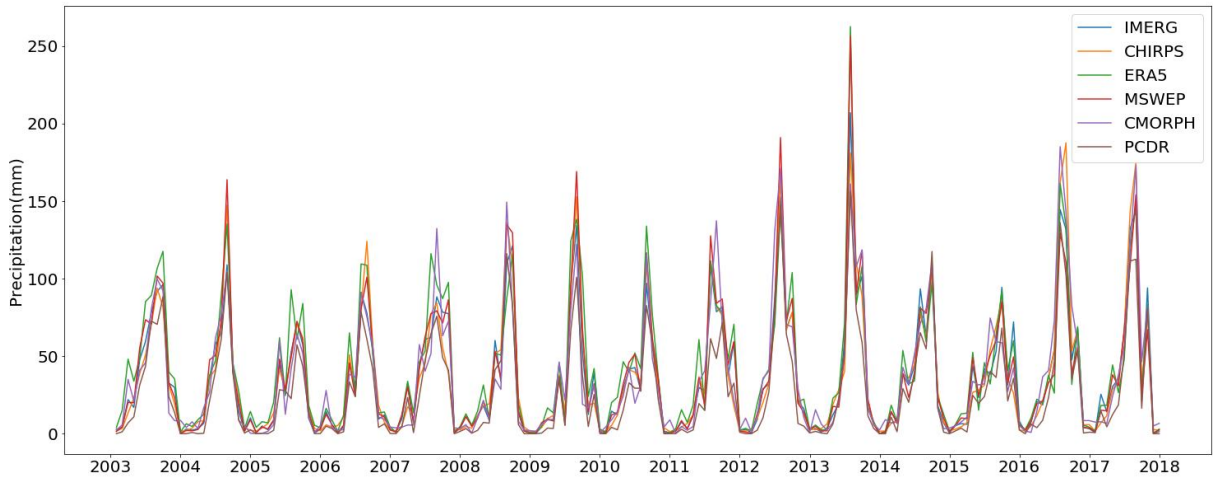


Figure 4.15 Monthly precipitation data of IMERG, CHIRPS, ERA5, MSWEP, CMORPH and PCDR for Wuding basin during 2003-2018. all values are in mm/month.

Table 4 Original and changed monthly water balance data used for each basin

Variable	Cases		Symbol	Original	Baihe	Beiluo	Wuding
Precipitation	Comb3	Comb1	P_{obs1}	GPM IMERG	ERA5	ERA5	ERA5
			P_{obs2}	CHIRPS	PCDR	PCDR	PCDR
Evaporation		Comb2	E_{obs1}	SSEBOP	ERA5	SSEBop	ERA5
			E_{obs2}	GLEAM	SSEBop	GLEAM	GLDAS
River Discharge			Q_{obs}	Stream gauges			
Storage			S_{obs}	GRACE JPL mascon RL06v02			

Table 5 Posterior mean value of each error parameter for original case and other 3 cases: (i)comb1:only precipitation datasets changed (ii)comb2:only evaporation datasets changed and (iii)comb3:both precipitation and evaporation datasets changed. Comb#:original combination.

Basin	Baihe				Beiluo				Wuding			
combinations	comb#	comb1	comb2	comb3	comb#	comb1	comb2	comb3	comb#	comb1	comb2	comb3
ω_E	0.317	0.322	0.220	0.543	5.887	0.684	5.887	0.684	0.799	0.770	0.220	0.141
f_E	0.819	0.966	1.068	0.869	1.127	1.014	1.127	1.014	1.013	0.892	1.068	0.951
r_E	0.432	0.366	0.223	0.370	0.351	0.346	0.351	0.346	0.374	0.336	0.223	0.288
A	6.999	24.810	9.410	23.69	11.430	12.770	11.430	12.770	19.640	17.370	9.410	8.169
$Delta$	0.270	0.425	0.061	0.406	0.262	0.037	0.262	0.037	0.082	0.088	0.061	0.054
$SStd$	20.080	18.530	7.703	19.23	14.510	13.910	14.510	13.910	7.434	7.362	7.703	7.209
a_Q	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.099	0.100	0.100	0.100
b_Q	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
ω_p	0.230	0.426	0.585	0.493	0.541	0.595	0.541	0.595	0.596	0.643	0.585	0.705
r_p	0.643	0.638	0.704	0.488	0.664	0.306	0.664	0.306	0.817	0.526	0.704	0.331
$Likelihood$	-767	-755	-766	-752	-702	-697	-702	-697	-666	-657	-648	-647

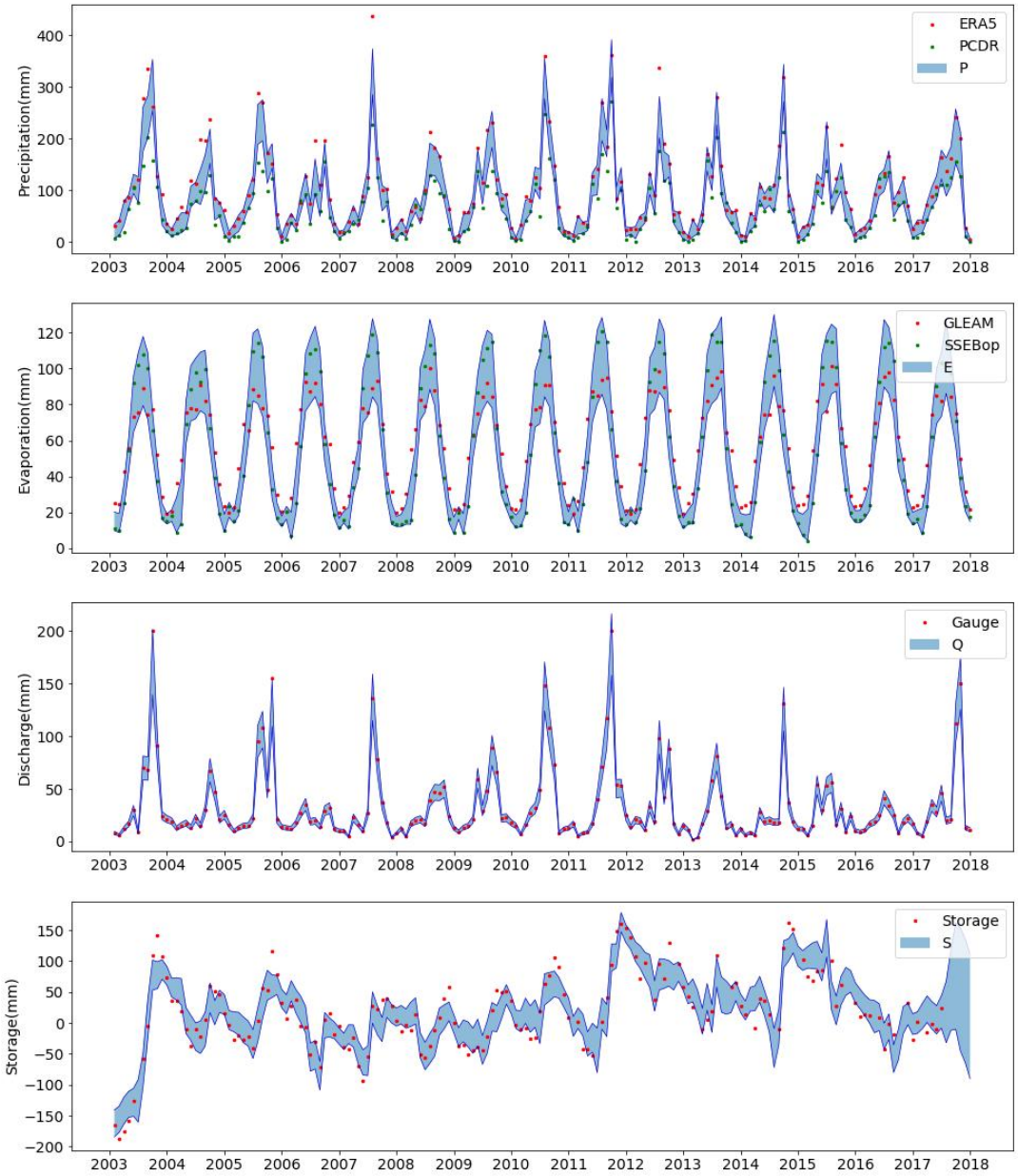


Figure 4-16 Monthly water balance estimates for Baihe basin with precipitation datasets changed.(ERA5 and PCDR) All values in mm/month.

It can be illustrated that for different basins with posterior plot and error parameters, the results of changing different remote sensing data are quite different. For the Baihe basin, the effect of changing precipitation datasets is better than that of changing evaporation remote sensing datasets because the value of log likelihood gets larger from -767 to -755. Differently, for the Wuding basin, the effect of changing evaporation remote sensing data is much greater than that of changing only precipitation data since the value of likelihood increases from -666 to -648, while there is no case of changing evaporation basin in the Beiluo basin, and changing both evaporation and precipitation remote sensing data is proven to be the best among these three cases according to log likelihood.

Figure 4-16 shows the monthly posterior distribution of water balance variables in Baihe basin with only the precipitation data changed (comb1), and Figure 4-17 shows the resulting precipitation and evaporation estimates together with those for the original case (comb#). The difference of the changed combination of precipitation remote sensing data gets larger, hence the uncertainty of the prior distribution becomes larger. The posterior estimates for comb1 tend to be larger compared to comb#, especially for peaks in summer. The inferred precipitation estimate tends to more closely follow ERA5 than PCDR, the ω_p value increases with respect to the original combination, and the r_p value does not differ much. Correspondingly, the posterior estimate of evaporation with changing datasets becomes smaller in summer in order to compensate for errors of other water balance variables, while in winter the estimate does not change much. Parameter f_E is closer to 1 (no additional bias), and r_E is smaller. For the water storage parameters, the amplitude A becomes larger and the random error becomes smaller, which means in this case the new combination of remote sensing datasets yield smaller noise and provides a better fit than the original one.

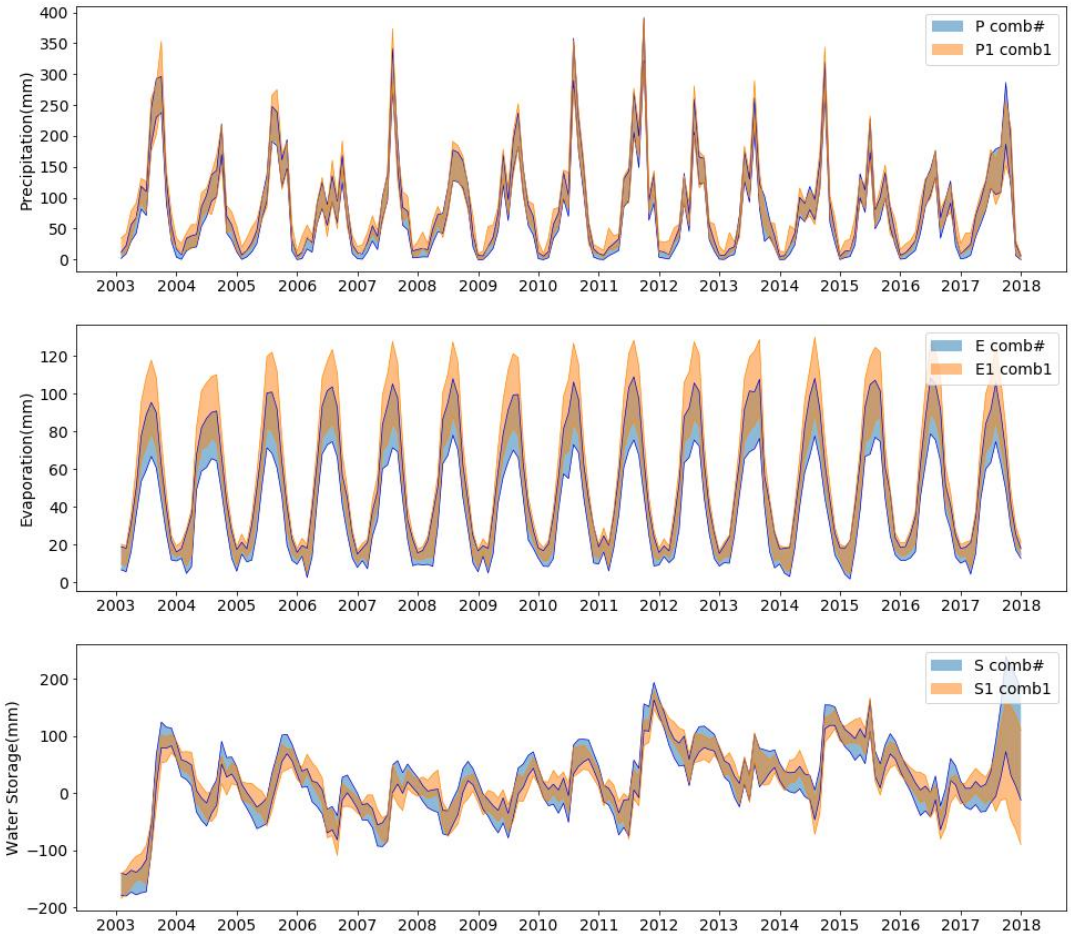


Figure 4-17 Monthly precipitation, evaporation and water storage estimates of comb# and comb1for Baihe basin. All values are in mm/month.

4.2.2 Effect of using different Evaporation datasets

Similar to Figure 4-15, figure 4-18 shows four evaporation remote sensing data time series from

2003 to 2018. The two remote sensing datasets with the largest difference in the figure were selected. Hence, different combinations of evaporation remote sensing data will be achieved for different catchments as shown in Table 4.

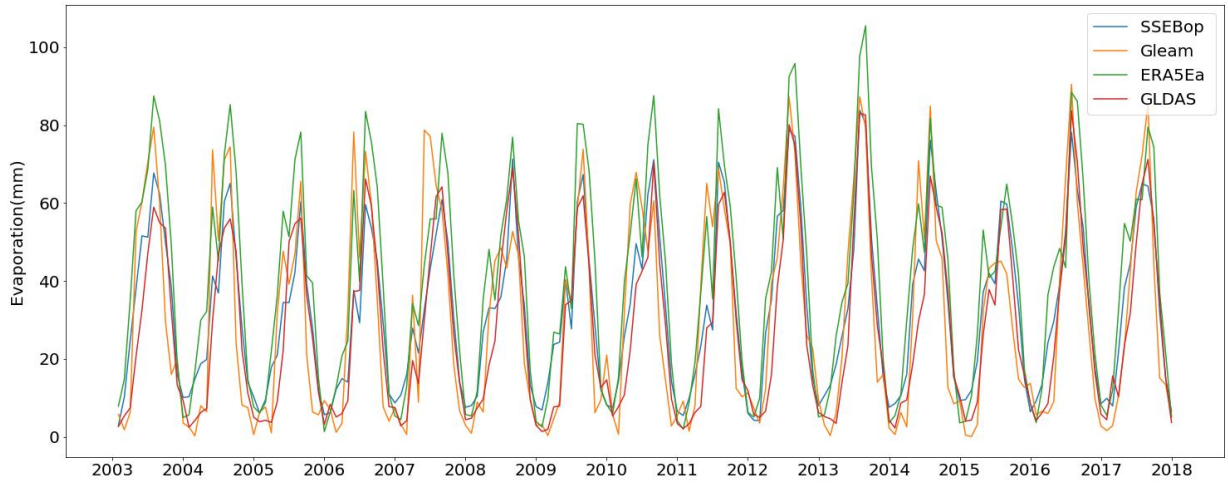


Figure 4-18 Monthly evaporation data of SSEBop, Gleam,ERA5 Ea and GLDAS for Wuding basin during 2003-2018. all values are in mm/month.

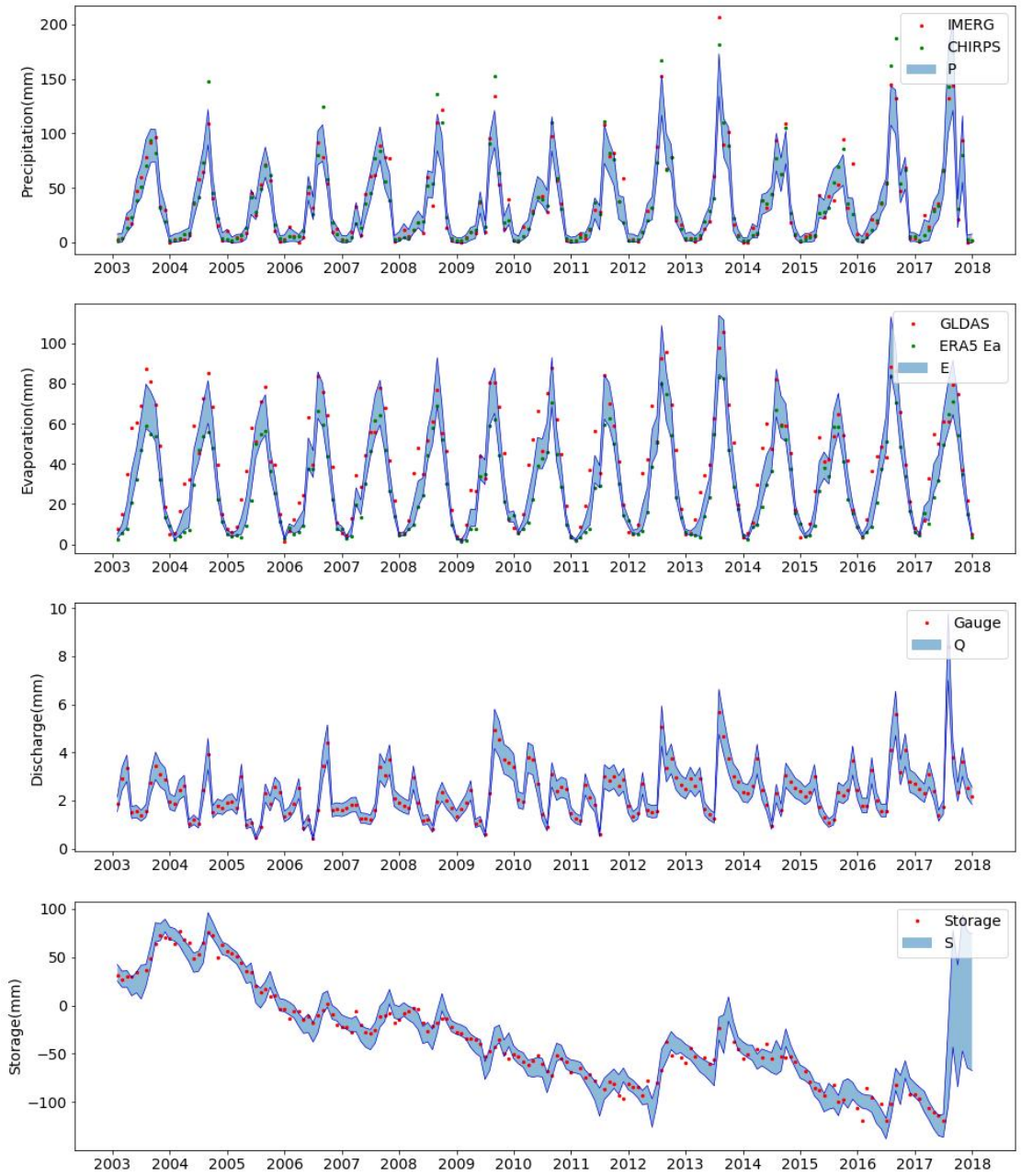


Figure 4-19 Monthly water balance estimates for Wuding basin with evaporation datasets changed.(ERA5 Ea and GLDAS) All values in mm/month.

Figure 4-18 shows the time series of four evaporation remote sensing datasets from 2003 to 2018. As mentioned in the previous section, changing evaporation data in the Wuding River basin is more effective than changing precipitation remote sensing datasets. Two evaporation datasets, ERA5 Ea and GLDAS, are chosen here. It can be seen from Figure 4-19 the precipitation prior distribution where the peak point is located does not fall within the posterior distribution. The error parameter ω_p decreases compared with the original combination (Table 5), which means that CHIRPS plays a greater weight in the error model, and r_p value increases, which means that the standard deviation of the prior distribution becomes smaller and the uncertainty of the posterior distribution decreases. From figure 4-20, the new posterior estimate for precipitation is

close to original precipitation posterior estimate.

For the evaporation posterior distribution, the new posterior is larger than the original estimate especially for peaks during 2006 - 2018 and the posterior probability gets larger (the second row of figure 4-20). The value of the ω_E parameter is significantly reduced, which means that ERA5 Ea has a higher weight, f_E is also closer to 1, indicating that there is almost no additional bias, and the value of r_E is also relatively reduced. The random error of water storage is not much different compared to the original combination.

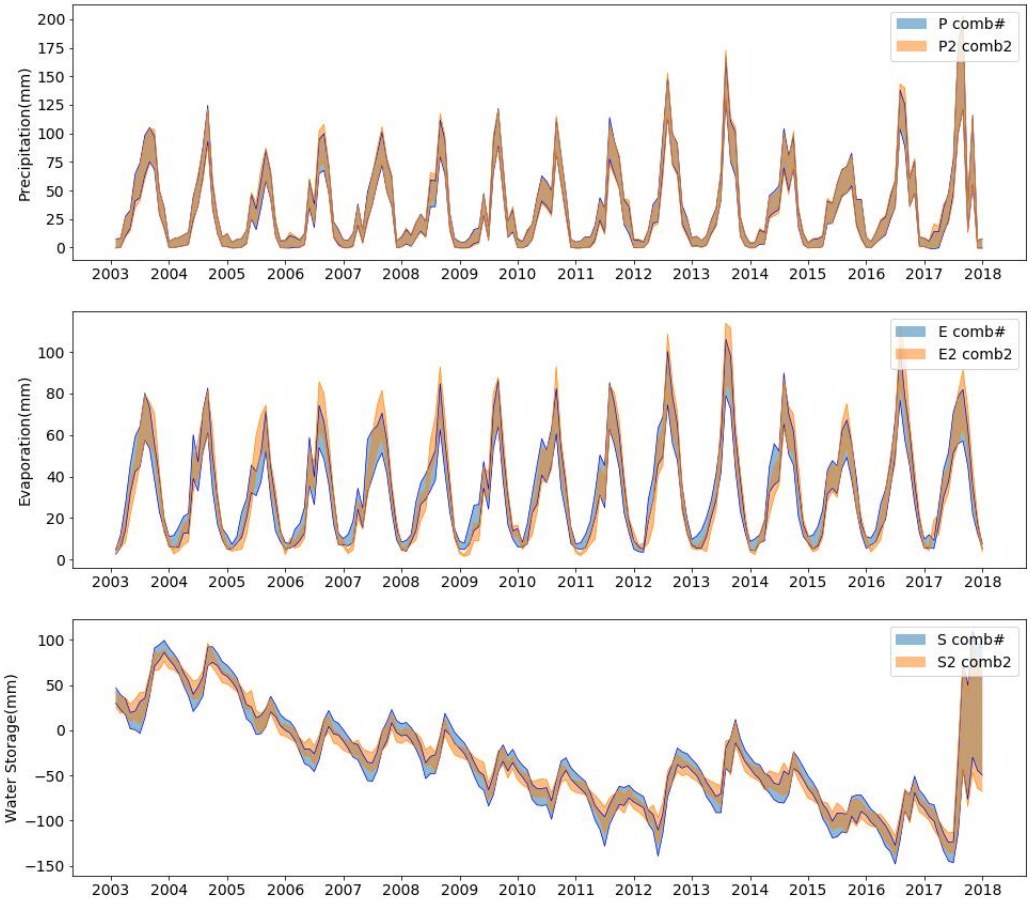


Figure 4-20 Monthly precipitation, evaporation and water storage estimates of comb# and comb2 for Wuding basin. All values are in mm/month.

4.2.3 Effect of using different Precipitation and Evaporation datasets

Compared with the first two cases, changing both precipitation and evaporation remote sensing datasets for the three catchments can improve the results. The larger difference between the two datasets of precipitation and evaporation can bring a correspondingly larger prior distribution, and the evaporation error parameter ω_E is reduced from 0.77 to 0.141, which means ERA5 has greater weight. The f_E is closer to 1 so additional error can be neglected, and r_E is reduced accordingly. The precipitation error parameter ω_p increases slightly and the value of r_p decreases significantly. For the water storage parameters, the amplitude decreases and the random errors are not significantly different.

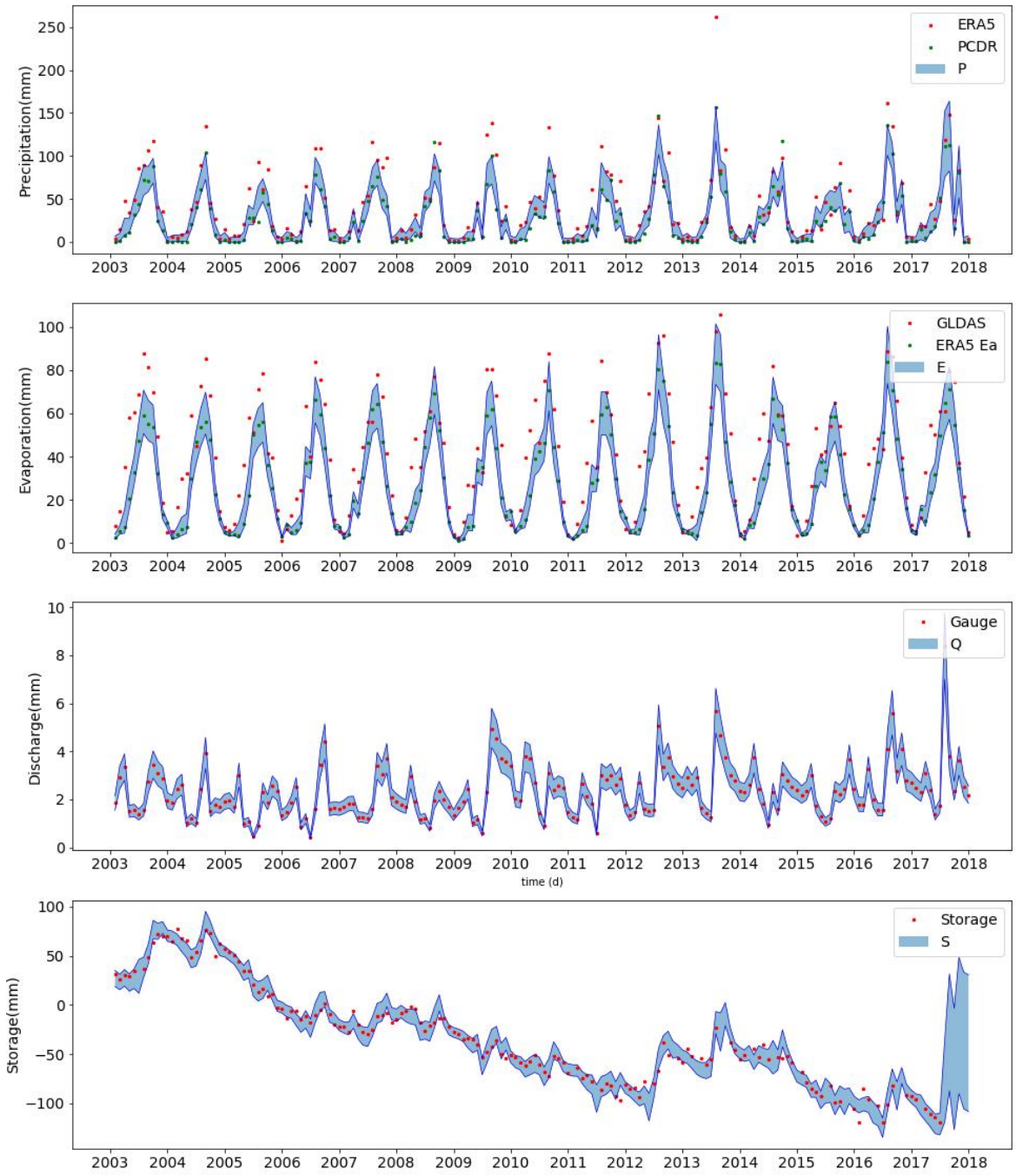


Figure 4-21 Monthly water balance estimates for Wuding basin with precipitation datasets changed(ERA5 and PCDR) and evaporation datasets changed(ERA5 Ea and GLDAS). All values in mm/month.

The situation of Baihe basin is more similar with that of the Beiluo basin, both of which have precipitation datasets playing a more important role in the improvement. The evaporation error parameter ω_E both show a tendency to become larger, implying that PCDR is heavily weighted in both basins, with no major difference in r_E values, while f_E values also both increase, but are about 0.9 in the Baihe basin. The precipitation error parameter ω_p increases in both, indicating that SSEBop and GLDAS are heavily weighted in both catchments respectively, and r_p decrease

in both. The random error of water storage significantly decreases in the Baihe basin, while the difference in the Wuding basin is not significant.

4.3 Effect of combining all available remote sensing datasets

The previous sections always used two precipitation and evaporation datasets for closing the water balance. However, since there are more than two datasets available, it is interesting to test whether the use of datasets ensembles (consisting of all available datasets) rather than datasets pairs (as in the previous sections) leads to a better characterization of prior uncertainty, and whether that then changes the water balance estimates. Therefore, this section presents results based on fusing all six precipitation datasets and all four evaporation datasets at the same time. Since the data fusion method still expects two time-series (e.g. P_{obs1} and P_{obs2}), the question then is how to collapse six precipitation datasets into two precipitation time series. The method used here is to compute the minimum and maximum value in each month of the six precipitation datasets and thus generate a time series of monthly minimum values (P_{obs1}) and monthly maximum values (P_{obs2}). The same procedure is applied to the four evaporation datasets yielding monthly minimum values (E_{obs1}) and monthly maximum values (E_{obs2}) respectively. as shown in formula 20 -23.

$$P_{\text{obs1}} = \min\left(P_{IMERG}, P_{CHIRPS}, P_{PCDR}, P_{ERA5}, P_{MSWEP}, P_{CMORPH}\right) \quad (20)$$

$$P_{\text{obs2}} = \max\left(P_{\text{IMERG}}, P_{CHIRPS}, P_{PCDR}, P_{ERA5}, P_{MSWEP}, P_{CMOPRH}\right) \quad (21)$$

$$E_{\text{obs1}} = \min\left(E_{SSEBop}, E_{GLEAM}, E_{ERA5}, E_{GLDAS}\right) \quad (22)$$

$$E_{\text{obs2}} = \max\left(E_{SSEBop}, E_{GLEAM}, E_{ERA5}, E_{GLDAS}\right) \quad (23)$$

After fusing all data, the prior distribution of precipitation is greatly increased, which ensures that the posterior distribution falls within the prior distribution range, and the error parameters ω_p and r_p after fusing all data are reduced compared to the original combination, so the uncertainty of the prior distribution is increased and the uncertainty of the posterior distribution of precipitation is correspondingly increased. ω_E and r_E are also reduced significantly, and the prior uncertainty is also reduced. The random error of water storage is only slightly decreased, while the amplitude is reduced to some extent.

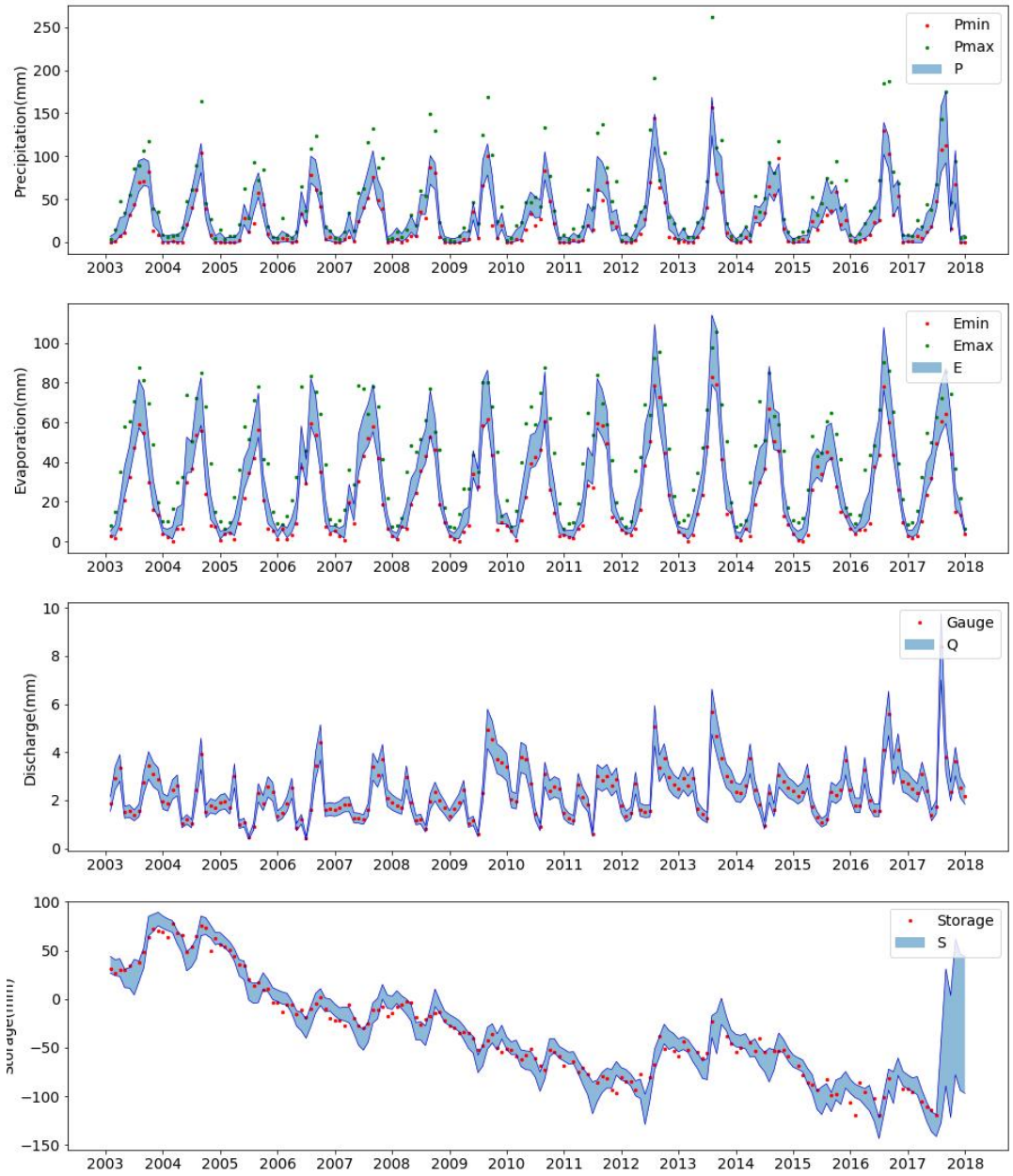


Figure 4-22 Monthly water balance estimates for Wuding basin when merging all datasets. All values in mm/month.

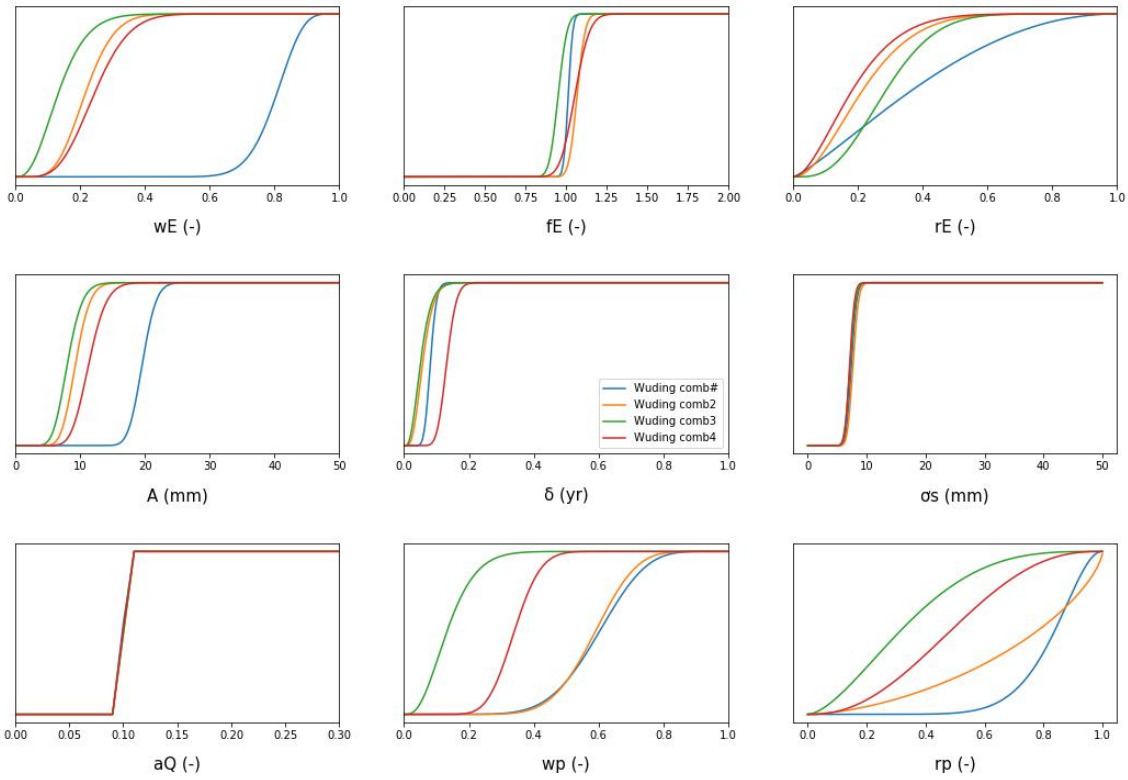


Figure 4-23 Posterior error parameter distribution for different cases for Wuding basin. Comb4:by merging all datasets.

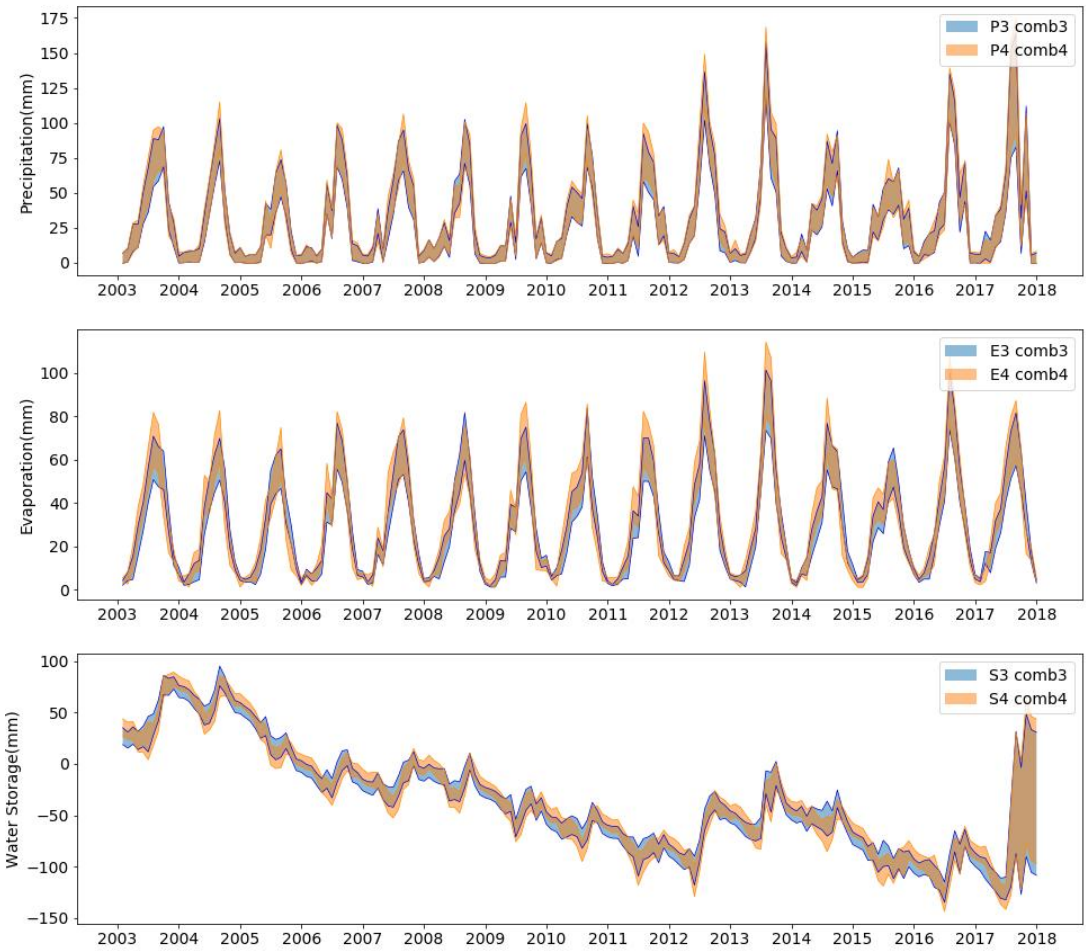


Figure 4- 24 Posterior estimate of precipitation, evaporation and water storage estimate for comb3 and comb4 for Wuding basin.All values are in mm/month.

5 Discussion and conclusions

In this thesis, a probabilistic model for estimating monthly basin-scale precipitation, evaporation, terrestrial water storage and river discharge was applied to evaluate the performance of an existing water balance data fusion method under different climatic conditions in China, i.e, wet, arid and semi-humid semi-arid basins.The fusion method differs from previous methods in that the data errors are not fixed a prior but are treated as unknown random variables that are estimated from the data. The fusion method is based on a Bayesian hierarchical model that calculates the posterior distribution of all unknown parameters by linking all water balance variables together with the error parameters in the error model, where random and systematic errors are calculated. The posterior distributions provide error-filtered and bias-corrected estimates of all water balance variables resulting in a hydrologically consistent water balance datasets that closes the water balance (Schoups and Nasseri, 2021).

In the probabilistic error model, the bias of precipitation and evaporation data are treated as weighted averages of two different datasets, where the weight factor ω_p and ω_E are treated as unknown parameters. Furthermore, for evaporation an additional scaling parameter f_E is added for additional bias correction, if necessary, and random errors are modeled as proportional to the difference between two input datasets. The error of the terrestrial water storage is mainly expressed by a noisy sine wave error model, where amplitude, phase and noise are treated as unknown parameters. River discharge data from gauging stations are assumed to be unbiased, with random errors assumed to increase linearly with the discharge using a 10% relative error for monthly discharge.

Application of the water balance data fusion methodology to river basins in China was used to investigate three research questions. The findings are summarized below and compared to other studies found in the literature, where appropriate.

- How does the methodology perform in river basins across China under different climatic and topographic conditions? To what extent are data error assumptions in the probabilistic data models satisfied in the case study basins?

In all three basins studied here, the posteriors of precipitation and evaporation in wet summer periods are larger than those in dry winter times. In general, the posterior uncertainty of evaporation is significantly smaller than its prior uncertainty, while the posterior uncertainty of precipitation is not significantly different from the prior uncertainty. Meanwhile, the observed GRACE data were corrected for amplitude and phase, and the underlying water storage estimates were obtained. The precipitation posterior of the Wuding basin is outside the prior range, which violates the prior model assumptions, suggesting that the water balance estimates may be less reliable in the arid Wuding basin than in the to other more humid basins. The time average standard errors of the Wuding basin are 8.9 mm/month for water storage, 6.5 mm/month for precipitation, 4.2 mm/month for evaporation, and 0.2 mm/month for river discharge, which are all lower than those of the humid and semi-humid regions. By comparison, Hao zhen(2017) used a

multi-source data fusion method to improve the accuracy of water storage inversion in humid and semi-humid areas, and applied Landsat8 and environmental remote sensing data to obtain water storage inversion results with an absolute error of 2.2 and a relative error of 11.42%. In Zhao's study, a multi-source precipitation information fusion scheme using ground precipitation datasets, CHIRPS 2.0 and Persiann-CDR and DEM in arid areas, which effectively improved the accuracy of precipitation estimates, reducing the average absolute error within 25 mm in monthly scale range and improving the KGE efficiency coefficient to 0.86, which result is also within the valid range as the average standard error in the thesis based on using part of the same datasets especially for humid and arid basins.

- How do the estimated results change for different remote sensing datasets?

Since the fusion methodology is data-driven and does not involve any hydrological processes, the analysis here optimizes and improves the results by replacing different datasets. The two different precipitation and evaporation datasets are selected based on the time series plots of six precipitation and four evaporation datasets, which requires the largest difference of the two datasets and applied to these three basins. Hence, ERA5 and PCDR were selected for all basins and for Baihe basin SSEBop and ERA5 were selected, for Wuding basin ERA5 and GLDAS were selected, while Beiluo basin remains the same.

Here the posterior estimate obtained under the three scenarios of changing only precipitation remote sensing datasets, changing only evaporation remote sensing datasets and changing both precipitation and evaporation remote sensing data are analyzed separately. The likelihood value increases significantly in the Baihe basin when changing precipitation datasets and in the Wuding basin when changing evaporation datasets, while replacing the precipitation-evaporation remote sensing data at the same time can effectively increase the likelihood value, which indicates better results, i.e. a better data fit. The evaporation posterior of the Wuding basin lies within prior distribution which is consistent with the prior assumption, thus improving on the original datasets used, while the evaporation posterior increases significantly in the wet summer period, and the precipitation posterior prediction also increases slightly in the summer peak time. This also proves that the replacement of both precipitation and evaporation datasets can effectively improve the estimated results of the method. As the posterior estimates of precipitation and evaporation increases, the posterior estimates of water storage also shows a decrease trend in winter and increase trend in wet summer seasons, but with no obvious change of phase.

- Does the use of dataset ensembles rather than dataset pairs (as in the original method) lead to a better characterization of prior uncertainty, and to what extent does it change the results and further improve the overall data fit?

In order to make full use of the remote sensing datasets, the monthly minimum and maximum values were calculated to create two different time-series for both precipitation and evaporation, based on all six precipitation datasets and all four evaporation datasets respectively, while the GRACE water storage datasets remained unchanged. It was found that using an entire ensemble of datasets, consisting of all available precipitation and evaporation datasets, further improves the

data fit, as measured by the likelihood values, which reach their largest value of all tested data fusion experiments. The likelihood values of all three basins were significantly larger compared to the original datasets combination.

In contrast with the other data fusion experiments, the precipitation posterior for the Wuding basin is now within the prior range, which is consistent with the prior assumptions in the model, thus increasing reliability of the estimated water balance variables. Furthermore, for Wuding basin the posterior estimates of evaporation are increased, and the uncertainty is increased, especially in summer time. For precipitation, the uncertainty increases for the summer peaks, and fusing of all remote sensing data results in increased uncertainty in the posterior values of precipitation and evaporation,the average standard errors increases from 4.1 mm/month for comb 3 to 4.4 mm/month for comb 4, also the likelihood increases, which is significantly better than the case that replaces only part of the datasets.

The posterior estimate obtained by fusing all precipitation and evaporation datasets have larger likelihood values and are better than single dataset or simple fusion of two types of precipitation-evaporation datasets. In Wang wen's study(Wang et al, 2020), for humid, semi-humid and arid basins, GLEAM/GLDAS/SSEBop was evaluated using ground flux observations and data fusion was performed based on four drought indices to obtain the results that evaporation datasets based on the fusion of the three evaporation data with better correlation with the flux observations and with higher accuracy than that of single dataset. Similarly, there is a general agreement with the previous research results for the arid and humid regions, the GLEAM is more accurate in semi-arid regions, and the GLDAS provides a high temperature datasets, which leads to large evapotranspiration simulation results. Hence, there are limitations in the study based on a single source dataset, and the wetness of the climate in central China has increased in recent years due to increased amount of precipitation, while the Yellow-Huai Plain region continues to become dry recent years.

The data fusion method can be extended to use more remote sensing datasets according to different climates or topography, and can be targeted to use remote sensing datasets with better accuracy and spatial distribution continuity in some areas. For example, IMERG has significantly better precipitation estimation accuracy in some eastern China than in the west, where precipitation has greater uncertainty. More data source types can also be used to enrich the number of datasets, such as introducing domestic satellites in future studies and making full use of the advantages of domestic satellite data sources. Also the corrected water balance posteriors can be applied in independent evaluation and calibrating hydrological models not only for discharge but also for precipitation, evaporation. Some data error models can also be used to correct errors on the datasets in different climates and land types. In this thesis, the study of multi-remote sensing data fusion is carried out only for precipitation and evaporation, which can be followed by multi-source data fusion for water storage such as adding different types and versions of updated GRACE datasets for fusion analysis.

Bibliography

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A. I., Weedon, G. P., ... & Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrology and Earth System Sciences, 21(12), 6201-6217.

Döll, P., Mueller Schmied, H., Schuh, C., Portmann, F. T., & Eicker, A. (2014). Global‐scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. Water Resources Research, 50(7), 5698-5720.

Feng Zheying, Yue Linwei, Shen Huanfeng. Accuracy Correction of GRACE Water Storage based on Multi-source Hydrological Data. Remote Sensing Technology and Application[J], 2021, 36(3): 605-617 doi:10.11873/j.issn.1004-0323.2021.3.0605

Feng Kepeng, Hong Yang, Tian Juncang, Tang Guoqiang, Kan Guangyuan, & Luo Xiangyu. (2020). Assessment of the utility of hydrological simulation of small watersheds with multi-source precipitation data. Arid Zone Geography, 43(5), 1179-1191.

Gao Ge, & Xu Chongyu. (2015). Characteristics of spatial and temporal variation of water surplus and deficit in ten major watersheds in China from 1961 to 2010. ACTA GEOGRAPHICA SINICA, 70(3).

Guo, S., Wang, J., Xiong, L., Ying, A., & Li, D. (2002). A macro-scale and semi-distributed monthly water balance model to predict climate change impacts in China. Journal of hydrology, 268(1-4), 1-15.

Hao Zhen. (2017). Research of soil water moisture inversion technique based on multi-source data collaboration (Master's thesis, Lanzhou Jiaotong University).

He Tian, Shao Quanqin. Based on MOD16 products in China during 2001-2010 Evapotranspiration Analysis of spatial pattern changes［J］. Journal of Geo-Information Science, 2014, 16(6): 979-988.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., ... & Stocker, E. F. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of hydrometeorology, 8(1), 38-55.
in the globe and typical regions (Doctoral dissertation, Tsinghua University).

Kim, K., Park, J., Baik, J., & Choi, M. (2017). Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. Atmospheric Research, 187, 95-105

Kong Y. (2017). Evaluation of accuracy of GPM/IMERG products over mainland China (Doctoral dissertation, Nanjing: Nanjing University of Information Engineering. Kong Y).

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. Water resources research, 48(4).

Liu B, Ran DC, Luo QH, Zhang ZP, & Wang CunR. (2001). Analysis of soil and water conservation measures in the Bei Luo River Basin for water and sand reduction (Doctoral dissertation).

Liu Bin, Ran Dachuan, Luo Quanhua, Zhang Zhiping, Wang Cunrong. (2001). Analysis of water and sediment reduction effects of soil and water conservation measures in Beiluo Region (PhD thesis).

Long D， Pan Y， Zhou J， et al. Global Analysis of Spatiotemporal Variability in Merged Total Water Storage Changes Using Multiple GRACE Products and Global Hydrological Models［J］. Remote Sensing of Environment， 2017， 192：198-216.

Long, D., Longuevergne, L., & Scanlon, B. R. (2014). Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. Water Resources Research, 50(2), 1131-1151.

Mao Hongmei, TRMM rainfall applied in flow prediction of LSHM in upper Hanjiang River basin, Express water resources&hydropower information,29(8),22-26

Massari, C., Crow, W., & Brocca, L. (2017). An assessment of the performance of global rainfall estimates without ground-based observations. Hydrology and earth system sciences, 21(9), 4347-4361.

Mo Xingguo, Liu Suxia, Lin Zhonghui, Chen Dan, Zhao Weimin. (2004). Simulating the water balance of Wuding River basin in Loess Plateau with a distributed Eco-hydrological model. ACTA GEOGRAPHICA SIHICA 59(3) 341-348

Mo, X. G., Liu, S. X., Lin, Z. H., Chen, D., & Zhao, W. (2004). Simulating the water balance of the wuding river basin in the Loess Plateau with a distributed eco-hydrological model. Acta Geographica Sinica, 59(3), 341-348.

Moreira, A. A., Ruhoff, A. L., Roberti, D. R., de Arruda Souza, V., da Rocha, H. R., & de Paiva, R. C. D. (2019). Assessment of terrestrial water balance using remote sensing data in South America. Journal of Hydrology, 575, 131-147.

Peng Z H, Li Y Z, Yu W J, et al. Research on the applicability of remote sensing precipitation products in different climatic regions of China[J]. Journal of Geo-information Science, 2021,23(7):1296-1311.

Ren YJ, Yong B, Lu DK, & Chen HQ. (2019). Evaluation of the Integrated Multi-satellite Retrievals (IMERG) for Global Precipitation Measurement (GPM) mission over the Mainland

China at multiple scales.Lake Science, 31(2), 560-572.

Schoups, G., & Nasseri, M. (2020). GRACEfully closing the water balance: a data‐driven probabilistic approach applied to river basins in Iran. Water Resources Research, e2020WR029071.

Sheffield J, Ferguson C R, Troy T J, et al. Closing the terrestrial water budget from satellite remote sensing[J]. Geophysical Research Letters, 2009, 36(7).

Tan, X., Liu, B., Tan, X., & Chen, X. Long‐term water imbalances of watersheds resulting from biases in hydroclimatic datasets for water budget analyses. Water Resources Research, e2021WR031209.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., & Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. Remote sensing of environment, 240, 111697.

Tang, G.. (2019). Validation, application and improvement of satellite remote sensing precipitation Tian Jing, Su Hongbo, Chen Shaohui, Sun Xiaomin, Chen Qingmei. (2012).The temporal and spatial changes of carbon dioxide evapotranspiration in China for the past 20 years. Resources Science, 34(7), 1277-1286.

Wang Wen, Yang Jiahui, Hua Tiantian, & Peng Min. (2020). Fusion of Multi-source evapotranspiration data and its application in drought monitoring. Yangtze River, 51(8), 19-26.

Wu Chenchen. (2020). Multi-source communication research information and the use of lectures in flood news (Master's thesis, Dalian University of Technology).

Xiong, L., & Guo, S. (1999). A two-parameter monthly water balance model and its application. Journal of hydrology, 216(1-2), 111-123.

Xu Jintao. Accuracy assessment and merging calibration for IMERG precipitation products over mainland China[D]. Zhejiang University, 2020.

Yang Xiuqin, Wang Guojie, Pan Xin, etc. Spatiotemporal changes of land surface evapotranspiration in China from 1980 to 2011 based on GLEAM remote sensing model［J］. Journal of Agricultural Engineering, 2015 (21): 132-141.

Yang, Y. D., & Zou, N. (1997). (1997). Preliminary analysis of the hydrological characteristics of the upper Han River. Hydrology, (2), 54-56.

Yang, Y. D., Zou, N., Guo, H., & Hu, Q.. (1997). Preliminary analysis of the hydrological characteristics of the upper Han River. Hydrology, (2), 54-56.

Yao C L, Li Q, Luo Z C, et al. 2019. Uncertainties in GRACE-derived terrestrial water storage changes over mainland China based on a generalized three-cornered hat method. Chinese J. Geophys. (in Chinese), 62(3): 883-897, doi: 10.6038/cjg2019L0454.

Yao Chao-Long, Li, Qiong, Luo, Zhi-Cai, Wang, Chang-Chi, Zhang, R., & Zhou, Bo-Yang. (2019). Uncertainty in GRACE-derived terrestrial water storage changes over mainland China based on a generalized three-cornered hat method . Journal of Geophysics, 62(3), 883-897.

Zhang Z， Chao B F， Chen J， et al. Terrestrial Water Storage Anomalies of Yangtze River Basin Droughts Observed by GRACE and Connections with ENSO［J］. Global and Planetary Change，2015， 126：35-45.

APPENDIX

Figure captions

Baihe

Figure 1 Monthly precipitation data of IMERG,CHIRPS,ERA5,MSWEP,CMORPH and PCDR for Baihe basin during 2003-2018. all values are in mm/month

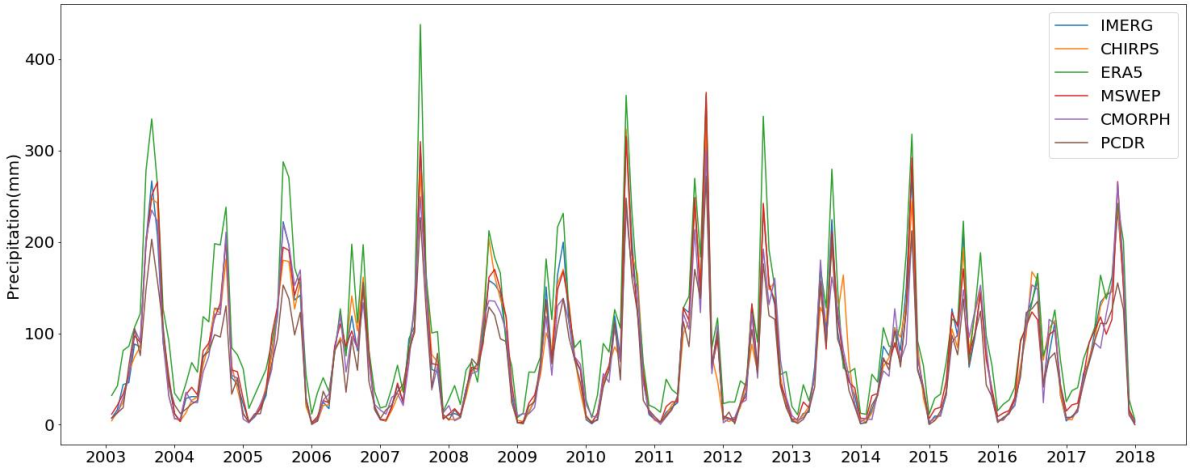
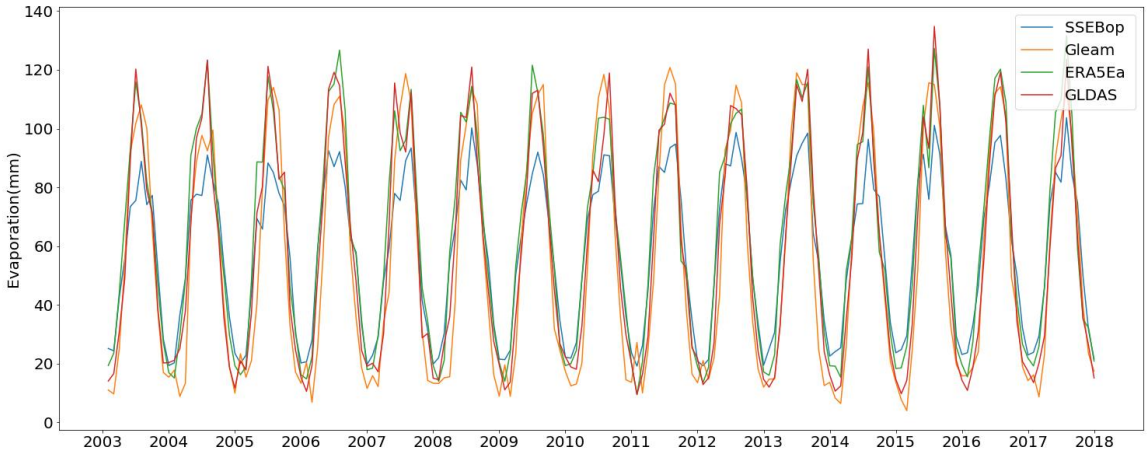


Figure 2 Monthly evaporation data of SSEBop, Gleam,ERA5 Ea and GLDAS for Baihe basin during 2003-2018. all values are in mm/month.



Beiluo

Figure 3 Monthly precipitation data of IMERG,CHIRPS,ERA5,MSWEP,CMORPH and PCDR for Beiluo basin during 2003-2018. all values are in mm/month

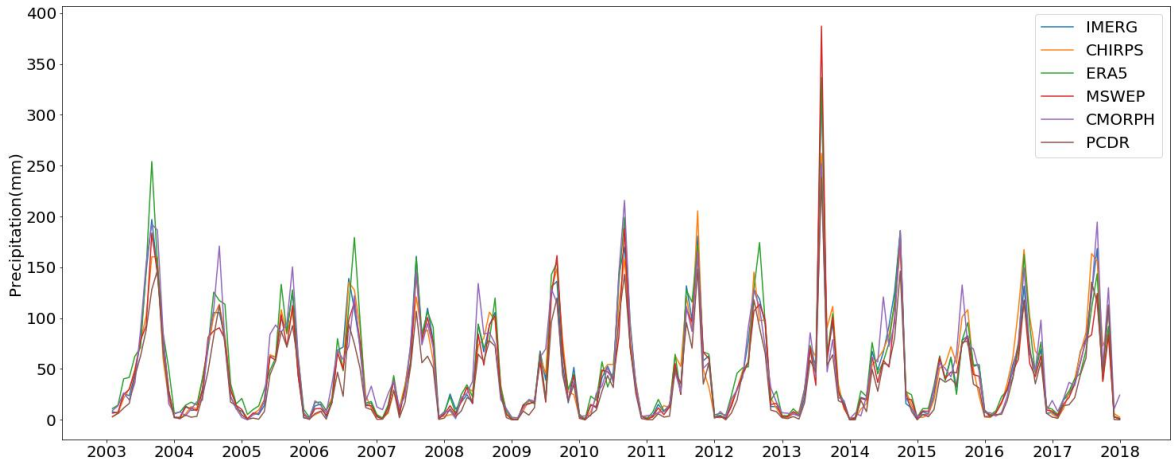


Figure 4 Monthly evaporation data of SSEBop, Gleam,ERA5 Ea and GLDAS for Baihe basin during 2003-2018. all values are in mm/month.