Characterized Flood Potential in the Yangtze River Basin from GRACE Gravity Observation, Hydrological Model, and In-Situ Hydrological Station

Nengfang Chao, Ph.D.[1]; and Zhengtao Wang[2]

Abstract: Comprehensive observations of total terrestrial water storage changes (TWSC) which include all the hydrological components (such as snow/glacier, surface water, soil moisture, and groundwater) are rarely available, so the predisposition of a region to flood is not fully clear. This paper combines the gravity recovery and climate experiment (GRACE) gravity observations, a hydrological model, and in-situ hydrological station data to establish the relationship between Yangtze River discharge and TWSC by a time-lagged autoregressive model and presents the TWSC data that apply for Yangtze River basin (YRB) flood forecasting. The TWSC in the YRB is inferred by the Lagrange multiplier method from GRACE gravity observations between April 2002 and December 2013. The root-mean-square error (RMSE) is optimal (2.1 cm) and the trend of TWSC in the YRB increased by 0.63 $\pm$ 0.11 cm/year. A case study of the flood catastrophe during summer 2010 is used to establish a relationship between river discharge from the Datong hydrological surveying station and basin water storage changes from GRACE by adopting a time-lagged autoregressive model, which shows that the total water storage changes from GRACE gravity observations can be used to estimate the tendency of a river basin to flood at 3–6 month lead times. This study concludes the basin scale of total water storage changes determined from satellite observations of time-variable gravity provides a new and effective tool for characterizing regional flood potential and may eventually lead to longer early flood warning times. **DOI: 10.1061/(ASCE)HE.1943-5584.0001547.**

Author keywords: Gravity recovery and climate experiment (GRACE); Yangtze river basin; Flood; Prediction.

Introduction

In nonfreezing areas, runoff hydrological processes can be divided into flow events, and show high regionalization in time and space due to the transient responses of increases in precipitation. The most fundamental signals of basic flow are the soil moisture and ground water stored in a basin (Appleby 1970). Because of observational limitations, many methods are able to predict floods only from the first flood phenomenon due to storms (Wetterhall et al. 2013; Guo et al. 2012). However, basic flow changes caused by ground water and its impact on the flooding are usually ignored. Therefore information such as precipitation forecasts, moisture in a basin (such as the estimation of the surface soil moisture using a model), and river levels are necessary in order to implement hydrological simulations. These may limit the time scale of flood forecasting ability relative to the weather forecast (3–10 days), or there may be more possibilities in flood forecasting due to the precipitation distribution and hydrological model (Siccardi et al. 2005). Currently, because these methods cannot contain all of the components of water storage, which is very important for the improvement of the early forecasting of flood conditions in basins, they limit the study of water resources. Because comprehensive observations are rarely made for these elements, the possibility of flooding occurring in basins remains unclear.

Only after the launch of the twin satellites of the gravity recovery and climate experiment (GRACE) (Tapley et al. 2004) did comprehensive observations of total terrestrial water storage changes (TWSC), which include snow/glacier, surface water, soil moisture, groundwater, and so on, become possible. The total water storage changes in large-scale areas can be obtained for each month by using the time-variable gravity field model provided by the GRACE satellite mission (Tapley et al. 2004). The GRACE observations not only can fill in the blanks of hydrological data but also can study the characteristics of the water circulation (Syed et al. 2008; Ramillien et al. 2008). Because the annual amplitude of water storage signals in a large basin exceeds the error range of the GRACE data, the GRACE data are very suitable for conducting limited studies of the hydrological process in a large catchment. In many areas, these annual signals are the main elements of the changes in water storage, and are consistent with the maximum water storage (Crowley et al. 2006). However, when these annual signals are larger than the maximum water storage, this can indicate flooding (Reager and Famiglietti 2009; Reager et al. 2014).

At present, there are many methods for discharge prediction. Wu et al. (2009) used moving average (MA), singular spectrum analysis (SSA), and wavelet multiresolution analysis (WMRA), coupled with an artificial neural network (ANN), to improve the estimate of daily flows. Chau and Wu (2010) used a hybrid model integrating ANNs and supported vector regression to develop daily rainfall prediction. Chen et al. (2015) compared the population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model. Gholami et al. (2015) combined dendrochronology, hydrogeological analysis, and ANN modeling to simulate groundwater level fluctuations during the past century.

[1]MOE Key Laboratory of Fundamental Physical Quantities Measurement, School of Physics, Huazhong Univ. of Science and Technology, Wuhan 430074, China; State Key Laboratory of Geodesy and Earth's Dynamics, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan 430077, China.

[2]Professor, School of Geodesy and Geomatics, Wuhan Univ., Wuhan 430079, China (corresponding author). E-mail: ztwang@whu.edu.cn

Note. This manuscript was submitted on December 16, 2016; approved on March 20, 2017; published online on July 5, 2017. Discussion period open until December 5, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699.

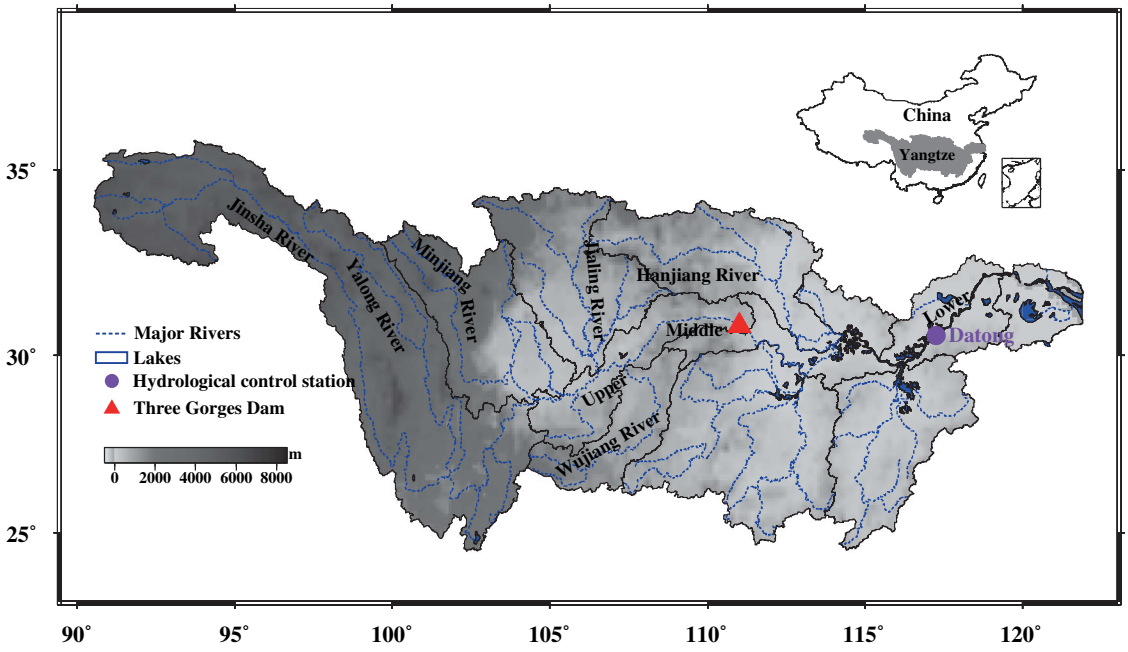

Fig. 1. Topography of Yangtze River basin, the major rivers, and in-situ Datong hydrological control station

Taormina and Chau (2015) derived input variable selection for rainfall–runoff modeling using binary-coded particle swarm optimization and extreme learning machines. Wang et al. (2015) improved forecasting accuracy of annual runoff time series using the autoregressive integrated moving average (ARIMA) model based on the ensemble empirical mode decomposition (EEMD); however, their idea of *flood potential* referred to extreme water storage changes which can increase runoff generation during future flood events several months in advance, and is clearly different from the traditional river flood forecasting (Reager et al. 2014, 2015).

Due to the insufficient spatial resolution (300–400 km) of the GRACE data, the determined surface quality changes are spatial smoothing rather than point measurements. Wahr et al. (1998) proposed a smoothing method based on a simple Gaussian filter. However, this method could not separate certain areas. In order to estimate regional surface quality changes using the GRACE data, corresponding technologies were required. Swenson and Wahr (2002) proposed that a regional quality change method, which reduces the impact of the GRACE observation errors, be utilized when separating the gravity signal of each area. This method mainly includes the following: (1) use of a Gaussian filter to smooth the real averaging kernel; (2) the overall errors (satellite measurement and signal leakage errors) are minimal; and (3) use of a Lagrange multiplier method, which is able to give the satellite errors in order to minimize the signal leakage errors and is able to give the signal leakage errors in order to minimize the satellite errors. Although a Gaussian filter can provide a simple and intuitive way of creating an averaging kernel, it decreases short-wavelength components and may not provide the most accurate estimate of the basin average. Moreover, the minimizing of the total error is used to determine an averaging kernel, which requires prior information, and if these values are untrustworthy, the resultant averaging kernel may not correspond to the signal characteristics (Swenson and Wahr 2002). Therefore this paper uses a Lagrange multiplier method in which the fixed satellite measurement error is used to minimize the signal leakage error to infer the TWSC in the Yangtse River basin (YRB) from GRACE satellite observations.

The TWSC can be determined from GRACE data, which have made a significant contribution to the field of hydrology in estimating large basin river discharge (Syed et al. 2009; Riegger et al. 2012; Riegger and Tourian 2014; Sneeuw et al. 2014) and in revealing large-scale groundwater depletion (Rodell et al. 2009; Famiglietti et al. 2011). Furthermore, there was a capacity limitation on using the water storage changes to estimate the regional flooding from previous studies (Crowley et al. 2006; Reager and Famiglietti 2009). This paper derives the total TWSC from monthly mean terrestrial water storage in the YRB from GRACE satellite observations by using a Lagrange multiplier method and then presents a case study of the YRB's flood in the summer of 2010. The paper also establishes the relationship of TWSC changes with river discharges provided by the Datong hydrological station (Fig. 1) and uses a time-lagged autoregressive model to analyze the possibility of flooding in the YRB, providing an earlier warning of future flood conditions.

The main work and contribution in this paper is (1) devising the Lagrange multiplier method to determine the TWSC in the YRB from GRACE gravity observations between April 2002 and December 2013; and (2) establishing the relationship between Yangtze River discharge from in-situ Datong hydrological surveying station river discharge and TWSC from GRACE by using a time-lagged autoregressive model to characterize the flood potential in the YRB. This paper develops a new and effective tool for characterizing regional flood potential and may ultimately result in longer lead times in flood warnings.

Study Area

The Yangtze River (Changjiang), one of the longest rivers in the world, originates from the Tuotuo River in the Tanggula Range of the Qinghai-Tibet Plateau, traverses 11 provinces and cities from west to east, and finally discharges in the East China Sea. The YRB has an area of approximately 1.8×10^6 km^2 which is approximately 20% of China's total mainland. The YRB is located in subtropical and temperate climate zones mainly affected by the

southeast monsoon winds. The mean air temperature and average annual precipitation are approximately 14°C and 1,100 mm, respectively. Most of precipitation (76%) occurs from April to September (Zhang et al. 2016).

The YRB is an important cultural and socioeconomic region for China and plays an unusual role in ecological conservation. However, it has suffered significant modification in climate and land cover, including the largest hydroelectric power station in the world—Three Gorges Dam (TGD). Climate change, extreme events (floods and droughts), and irrigation have deeply influenced water resources, thereby resulting in significant influences on natural and human systems. The YRB experienced three serious total-basin disastrous floods in 1931, 1956, and 1998. Although such similar catastrophic flooding has not occurred in the YRB since the 21st century began, the YRB has experienced many medium and small floods. For example, according to the statistics obtained from the Office of State Flood Control and Drought Relief Headquarters, the flooding which occurred in June, July, and August (JJA) of 2010 caused 28 provinces to suffer flood disaster conditions and affected 9,418,000 ha of crops. The total population affected by the flooding was 140 million, with a confirmed death toll of 1,057, 615 missing persons, and 1,090,000 collapsed buildings. The direct economic loss incurred was US$30,153,760,000.

Data Sets

This section presents the study area of the YRB and data sets of discharge from in-situ Datong measurements, GRACE gravity observations, and hydrological models [Global land data assimilation system (GLDAS), Climate Prediction Center (CPC), and the WaterGAP Global Hydrology Model (WGHM)].

In-Situ Datong Discharge of YRB

The Datong is the entire YRB control hydrological station, which is located approximately 500 km from the mouth of the Yangtze River and measures the contribution from an upstream area of approximately 1.7×10^6 km^2 (Fig. 1). The daily average runoff (R), precipitation (P) and actual evapotranspiration (ET_a) of the YRB during 2002–2013, acquired from the Datong control hydrological station, were used to determine the hydrological water storage change (ΔS_H) by the water balance method

$$\underbrace{\Delta S}_{\text{GRACE}} \leftrightarrow \underbrace{\Delta S_H = P - R - \text{ET}_a}_{\text{Hydrology}} \tag{1}$$

where $\underbrace{\Delta S}_{\text{GRACE}}$ is the water storage change from GRACE data.

GRACE

This paper used GRACE satellite data from the RL05 time-variable gravity field model provided by Jet Propulsion Laboratory (JPL), which included a total of 131 months (missing were June 2002, July 2002, June 2003, January 2011, June 2011, May 2012, October 2012, March 2013, August 2013, and September 2013) from April 2002 to December 2013. The gravity field could not be resolved in the case of GRACE satellite resonance, which therefore led to loss (Wagner et al. 2006). The GRACE monthly time-variable gravity field models consist of a set of spherical harmonic coefficients of the fully normalized external earth gravity field (Heiskanen and Moritz 1967), ($\bar{C}_{lm}$, $\bar{S}_{lm}$), up to degree and order 60. The second-order term of the GRACE time-variable gravity field model was replaced with the C_{20}, determined with satellite laser ranging (SLR) observation data (Cheng et al. 2011), the degree 1 harmonic coefficients (Earth's geocenter) were estimated from Swenson et al. (2008), and correction for glacial isostatic adjustment (GIA) was applied following Geruo et al. (2013).

Hydrological Models

This paper used the following hydrological models to verify the result of total water storage changes in the YRB from GRACE satellite observation by using a Lagrange multiplier method.

GLDAS

The GLDAS constrains land surface states through the observation data of the new-generation National Aeronautical and Spatial Administration (NASA) surface and space observation system (Rodell et al. 2004). The GLDAS produces optimal fields of land surface states and fluxes in near real time by combining satellite and ground-based observations into four land surface models. The GLDAS provides land surface models with spatial resolutions of $1 \times 1°$ and $0.25 \times 0.25°$, and time resolutions of 1 month and 3 h. This paper used the National Centers for Environmental Prediction/Oregon State University/Air Force/Hydrologic Research Laboratory NOAH) model monthly soil moisture states in a $1 \times 1°$ grid to estimate the water storage in the top 2-m soil layer along with the NOAH snowfield on the surface within the YRB.

CPC

The CPC at the National Oceanic and Atmospheric Administration (NOAA), which uses observed precipitation and temperature to create a hydrological model, has generated global monthly soil moisture estimates at $0.5 \times 0.5°$ resolution from 1948 to present (Dool et al. 2003).

WGHM

WaterGAP, a global water resources and use model, consists of the WGHM (Döll et al. 2003) and a number of water use models for irrigation, livestock, manufacturing, cooling of thermal power plants, and households. With a spatial resolution of $0.5 \times 0.5°$ and a daily time step, the time series of fast-surface and subsurface runoff, groundwater recharge, river discharge, and water storage changes in canopy, snow, soil, groundwater, lakes, reservoirs, wetlands, and rivers were simulated. Thus this model can quantify the total water resources as well as the renewable groundwater resources of a grid cell, river basin, or country. This paper applied WaterGAP version 2.2 (Müller et al. 2014) at a spatial resolution of 0.5° with global coverage in monthly time steps.

Method

This section introduces a Lagrange multiplier method to determine the total terrestrial water storage changes in the YRB from GRACE monthly solution, then establishes the relationship between river discharge and TWSC soil moisture with a time-lagged autoregressive model, which is used to analyze the possibility of flooding in the YRB and provide an earlier warning of future flood conditions (Fig. 2).

Lagrange Multiplier Method

The Lagrange multiplier method not only can decrease the satellite measurement errors but also can reduce the signal leakage errors in order to determine regional water storage changes. Its essence is that the ratio of the variance of error between the accurate average kernel and the approximate average kernel to the accurate average kernel is the lowest. Therefore this paper used a Lagrange multiplier method with minimum signal leakage error to determine the changes of water storage in the YRB on the basis of the fixed

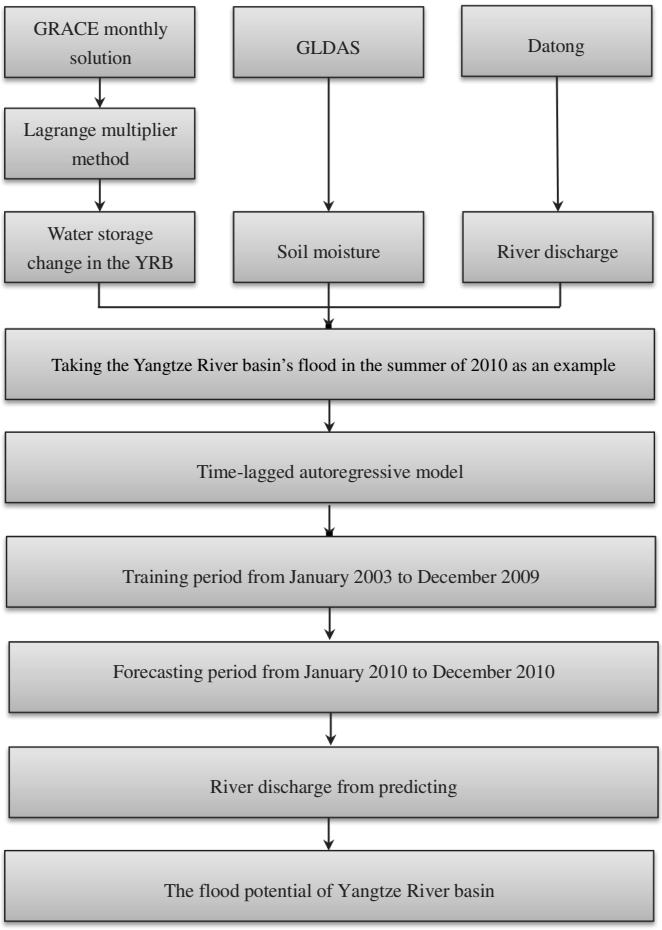

Fig. 2. Flowchart of the entire procedure

satellite measurement error. Swenson and Wahr (2002) described the method in detail. A brief description of the Lagrange multiplier method in this paper follows.

Wahr et al. (1998) concluded that the relationship between the local surface quality changes and the residual spherical harmonic is as follows:

$$\Delta\sigma(\theta,\phi)=\frac{a\rho_E}{3}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\frac{(2l+1)}{(1+k_l)}\tilde{P}_{lm}(\cos\theta)\{\Delta\tilde{C}_{lm}\cos m\phi+\Delta\tilde{S}_{lm}\sin m\phi\} \quad (2)$$

where θ = colatitude; ϕ = geocenter longitude; a = semimajor axis of the reference ellipsoid; (l,m) = degree and order of the spherical harmonic coefficient; $\tilde{P}_{lm}$ = fully normalized Legendre association function (e.g., Heiskanen and Moritz 1967); ρ_E = average earth density (5,517 kg/m^3); k_l = Love number (Farrell 1972); and $\Delta\tilde{C}_{lm}$ and $\Delta\tilde{S}_{lm}$ = residual spherical harmonic coefficients, which are obtained by removing the long-term mean of the Stokes coefficients from each of the monthly values.

On the premise that the spatial resolution is reduced, use spatial smoothing to improve the accuracy of the surface quality anomaly. Then assume that the accurate average kernel $\vartheta(\theta,\phi)$ is the function of the shape of the described area (river, ocean, ice sheet, political border, and so on)

$$\vartheta(\theta,\phi)=\begin{cases}0 & \text{outside the basin}\\ 1 & \text{inside the basin}\end{cases} \quad (3)$$

The vertical integration of water storage in any average area is

$$\overline{\Delta\sigma}_{\text{region}}=\frac{1}{\Omega_{\text{region}}}\int\Delta\sigma(\theta,\phi)\vartheta(\theta,\phi)d\Omega \quad (4)$$

where $d\Omega=\sin\theta d\theta d\phi$ is the solid angle element; and $\vartheta(\theta,\phi)$ = integration in the given spherical domain Ω_{region} (spherical area). From the Eq. (4)

$$\overline{\Delta\sigma}_{\text{region}}=\frac{a\rho_E}{3\Omega_{\text{region}}}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\frac{(2l+1)}{(1+k_l)}(\vartheta^c_{lm}\Delta C_{lm}+\vartheta^s_{lm}\Delta S_{lm}) \quad (5)$$

where ϑ^c_{lm} and ϑ^s_{lm} = spherical harmonic coefficients of $\vartheta(\theta,\phi)$

$$\vartheta(\theta,\phi)=\frac{1}{4\pi}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\tilde{P}_{lm}(\cos\theta)\{\vartheta^c_{lm}\cos m\phi+\vartheta^s_{lm}\sin m\phi\} \quad (6)$$

$$\begin{Bmatrix}\vartheta^c_{lm}\\ \vartheta^s_{lm}\end{Bmatrix}=\int\vartheta(\theta,\phi)\tilde{P}_{lm}(\cos\theta)\begin{Bmatrix}\cos m\phi\\ \sin m\phi\end{Bmatrix}d\Omega \quad (7)$$

Replacing the accurate average kernel $\vartheta(\theta,\phi)$ in the approximate expression area of equation with $\bar{W}(\theta,\phi)$ gives

$$\widetilde{\Delta\sigma}_{\text{region}}=\frac{1}{\Omega_{\text{region}}}\int\Delta\sigma(\theta,\phi)\bar{W}(\theta,\phi)d\Omega \quad (8)$$

where $\widetilde{\Delta\sigma}_{\text{region}}$ = approximate areal average; and $\bar{W}$ can be expanded to obtain the following:

$$\bar{W}(\theta,\phi)=\frac{1}{4\pi}\sum_{l=0}^{l_{\text{tmc}}}\sum_{m=0}^{l}\tilde{P}_{lm}(\cos\theta)\{W^c_{lm}\cos m\phi+W^s_{lm}\sin m\phi\} \quad (9)$$

The approximate areal average can be approximately expressed by using the Stokes coefficient

$$\widetilde{\Delta\sigma}_{\text{region}}=\sum_{l,m}\frac{K_l}{\Omega_{\text{region}}}(W^c_{lm}\Delta C_{lm}+W^s_{lm}\Delta S_{lm}) \quad (10)$$

where $K_l=(a\rho_E/3)[(2l+1)/(1+k_l)]$.

The variance of the corresponding satellite measurement error is

$$\text{var}(\varepsilon_{\text{sat}})=\frac{1}{\Omega^2_{\text{region}}}\sum_{l,m}\frac{K_l^2B_l^2}{2l+1}[W^{c2}_{lm}+W^{s2}_{lm}] \quad (11)$$

where $B_l^2=\frac{1}{n}\sum_{i=1}^{n}\sum_{m=0}^{l}\{[\Delta\delta^c_{lm}(t_i)]^2+[\Delta\delta^s_{lm}(t_i)]^2\}$.

A Lagrange multiplier method can be used to obtain the average kernel with the minimum limited satellite measurement error and signal leakage error not based on any prior information. Define a type of signal leakage as the ratio of the variance of accurate and approximate average kernel difference to the variance of the accurate average kernel

$$\text{var}(\varepsilon_{lkg})=\frac{\int[\bar{W}(\theta,\phi)-\vartheta(\theta,\phi)]^2d\Omega}{\int[\vartheta(\theta,\phi)]^2d\Omega}=\frac{1}{4\pi\Omega_{\text{region}}}\sum_{l,m}[(W^c_{lm}-\vartheta^c_{lm})^2+(W^s_{lm}-\vartheta^s_{lm})^2] \quad (12)$$

Let δ^2 be the variance of satellite measurement error, and $\Delta^2=\delta^2\Omega^2_{\text{region}}$ be the Lagrange multiplier multiplied by λ; W^c_{lm}, W^s_{lm}, and λ then can be determined through the minimum value of Eq. (13)

$$\xi = \sum_{l,m}[(W_{lm}^c - \vartheta_{lm}^c)^2 + (W_{lm}^s - \vartheta_{lm}^s)]^2 + \lambda\left\{\sum_{l,m}\frac{K_l^2 B_l^2}{2l+1}[W_{lm}^{c2} + W_{lm}^{s2}] - \Delta^2\right\} \quad (13)$$

Eq. (13) ignores $4\pi\Omega_{\text{region}}$, which is considered in λ. Seek the partial derivative of W_{clm} and W_{slm} of ξ and let them be zero to obtain

$$\begin{Bmatrix} W_{lm}^c \\ W_{lm}^s \end{Bmatrix} = \left[1 + \lambda\frac{K_l^2 B_l^2}{2l+1}\right]^{-1}\begin{Bmatrix} \vartheta_{lm}^c \\ \vartheta_{lm}^s \end{Bmatrix} \quad (14)$$

Seek the partial derivative of λ of ξ and let it be zero to obtain the satellite measurement error equivalent to Δ^2

$$\sum_{l,m}\frac{K_l^2 B_l^2}{2l+1}[W_{lm}^{c2} + W_{lm}^{s2}] = \Delta^2 \quad (15)$$

Through the combined Eq. (14)

$$\sum_{l,m}\frac{K_l^2 B_l^2}{2l+1}\frac{\vartheta_{lm}^{c2} + \vartheta_{lm}^{s2}}{\left[1 + \lambda\frac{K_l^2 B_l^2}{2l+1}\right]^2} = \Delta^2 \quad (16)$$

When λ is obtained through Eq. (16), use it in Eq. (14) to obtain W_{lm}^c and W_{lm}^s, then use them in Eq. (10) to infer the water storage change. This is the Lagrange multiplier method which uses the fixed satellite measurement error to minimize the signal leakage error.

The water storage change $[\Delta h(\theta, \phi, t_i)]$ of the YRB was fit by (Chao et al. 2016)

$$\Delta h(\theta, \phi, t_i) = a_0 + a_1(t_i - t_0) + a_2(t_i - t_0)^2 + \sum_{k=1}^{4}[A_k \cos(2\pi f_k t_i) + B_k \sin(2\pi f_k t_i)] \quad (17)$$

where t_i = time epoch; $f_k (k = 1, \ldots, 4)$ are the signal frequencies of the trend, acceleration signals (Ogawa et al. 2011), annual and semiannual seasonal signals, and S_2 tidal alias (cycle: 161 days) and K_2 tidal alias (cycle: 3.73 years) (Chen et al. 2009), respectively; A_k and B_k = signal amplitudes; a_1 = trend value; and a_2 = acceleration. The t_0 value was the middle moment of the calculation cycle.

Time-Lagged Autoregressive Model

The model framework was used to illustrate the ability of GRACE observations to estimate the flood potential by adopting the time-lagged autoregressive model from Reager et al. (2014).

The basic autoregressive model includes two terms as a function of river discharge at a previous time for forecasting river discharge

$$Q(t) = x_1 \cdot Q(t - \tau) + x_2 \cdot Q(t - 12) \quad (18)$$

where Q = river discharge at the time t; τ = lead time in months and the second term is the discharge in the same month of the previous year (because river discharge has strong seasonal characteristics); and x_1 and x_2 are constants.

The third term representing one of two potential regional flood predictors is

$$Q(t) = x_1 \cdot Q(t - \tau) + x_2 \cdot Q(t - 12) + y_1 \cdot \text{Soil}(t - \tau) \quad (19)$$

$$Q(t) = x_1 \cdot Q(t - \tau) + x_2 \cdot Q(t - 12) + y_2 \cdot \text{TWSC}(t - \tau) \quad (20)$$

Soil refers to the time series of soil moisture provided by the GLDAS model, TWSC is the time series of the total TWSC, and y_1 and y_2 are constants. A regression model is used to estimate the overall prediction time so that each point of the forecast time series is calculated through the time τ before its occurrence.

There are two metrics for the performance of autoregressive model (Reager et al. 2014): (1) the ability of the model to recognize a potential high discharge event is described by the forecast of a river discharge larger than the historic (2000–2013) 99th percentile for monthly discharge during the prediction period (54,664.4516 m^3/s for the data record of Datong); and (2) the model accuracy is described by the maximum absolute value error, the mean absolute value error (MAE), mean bias, and mean percentage relative error (MPRE) between model and observations.

Taking the flood of JJA 2010 in the YRB as an example, this paper divided the time series into the following: (1) training period from January 2003 to December 2009; and (2) forecast period from January 2010 to December 2010. The lead time τ was 1–12 months. This paper used a least-squares linear regression during the training period to solve the model coefficients, and then applied them to prediction during the forecast period.

Results

Comparison of GRACE and Hydrological Data in YRB

This paper determined the total water storage changes of YRB during 2002–2013 from data from the Center for Space Research (CSR), GeoForschungsZentrum Potsdam (GFZ), and Jet Propulsion Laboratory (JPL) through the Groupe de Recherche de Géodésie Spatiale (GRGS). The calculation of total water storage change adopted a DDK5 filtering (Kusche et al. 2009; Chao et al. 2016), Lagrange multiplier method, and the hydrological water storage change from Datong hydrological station through Eq. (1).

Fig. 3 shows the time series of GRACE-based TWSC and terrestrial water storage anomalies from different hydrological data spatially averaged over the YRB. The figure compares time series data from the four GRACE data sets and shows the distinctive seasonal variations with the minimum water storage usually found during the months of October–November and the maximum during the months of April–May.

As shown in Fig. 3, TWSC were estimated from different GRACE products, and the hydrological data were basically consistent with each other. However, because the method of gravity solution and the postprocessing of GRACE Level 2 data (Bettadpur 2012) were different, the results of water storage change were not fully consistent. Moreover, the water storage change in the YRB showed an increasing trend, with an annual observed increase rate of 0.63 ± 0.11 cm/year. The water storage change increased constantly since 2007.

In addition, this paper investigated statistics of water storage change from different GRACE and hydrology products with in-situ data from the Datong hydrological station (Fig. 4), which showed that the mean, RMSE, R-square, and P-value were basically consistent with each other, but the mean and RMSE were smallest and the R-square was the largest from the Lagrange multiplier method. Therefore the results of water storage change in the YRB from the Lagrange multiplier method were optimal.

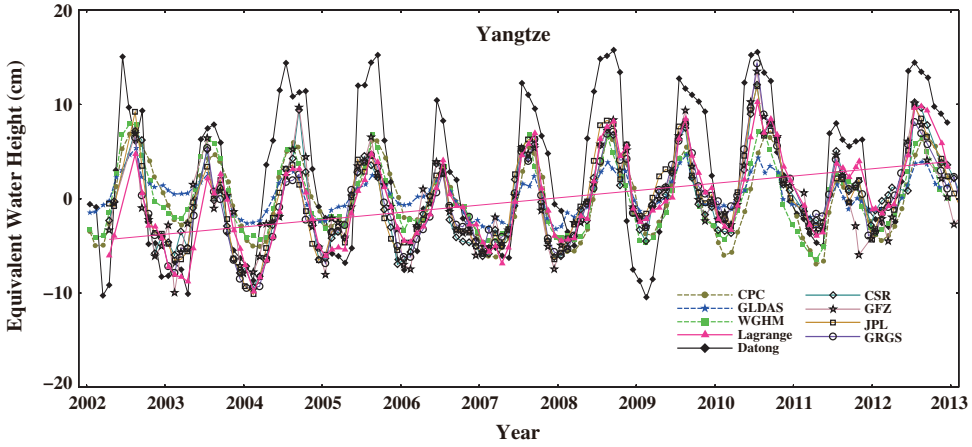

Fig. 3. Time series of water storage changes from different GRACE and hydrological data

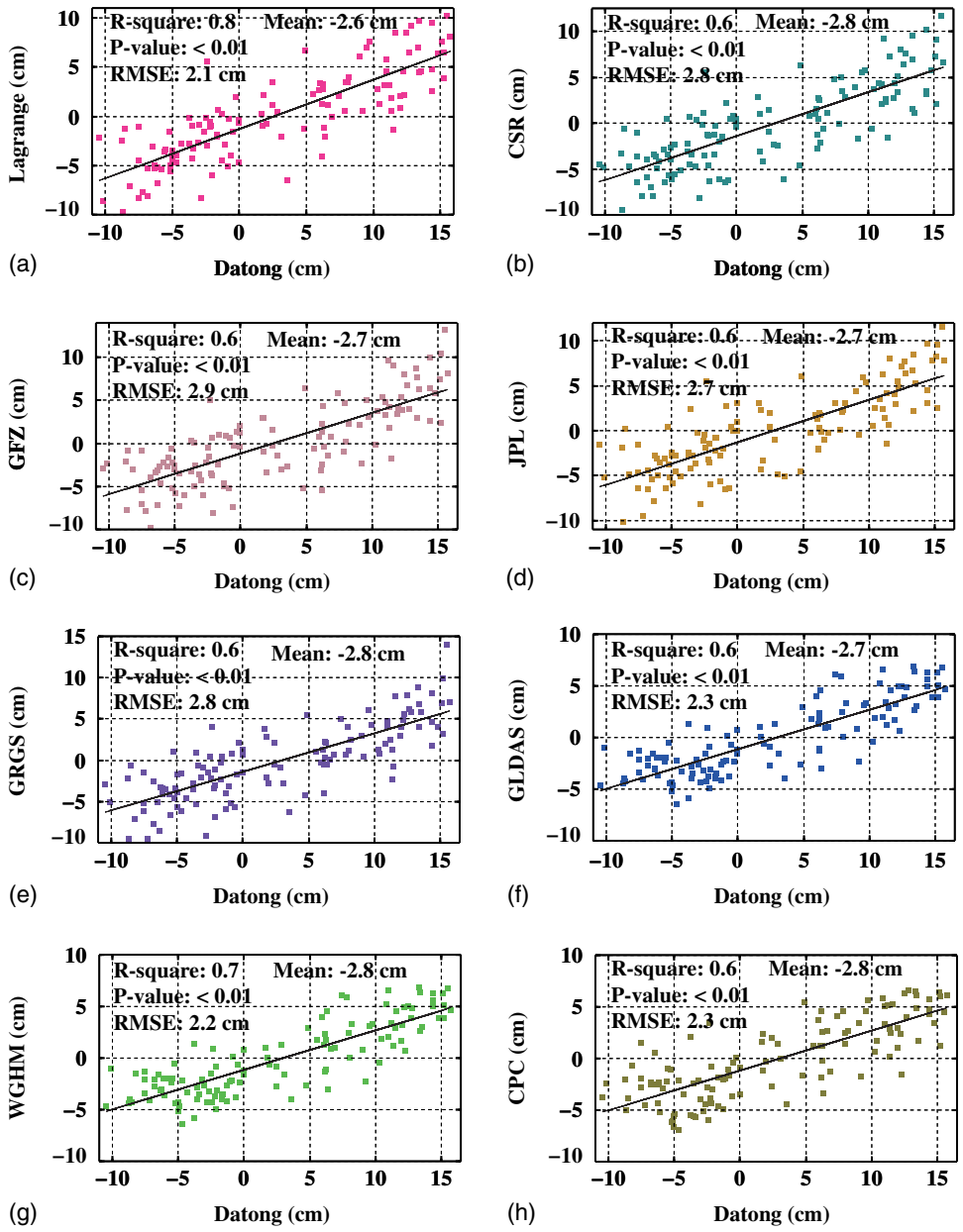

Fig. 4. Statistics of water storage changes from different data with in-situ Datong data (unit: centimeters): (a) Lagrange multiplier method; (b) CSR; (c) GFZ; (d) JPL; (e) GRGS; (f) GLDAS; (g) WGHM; (h) CPC

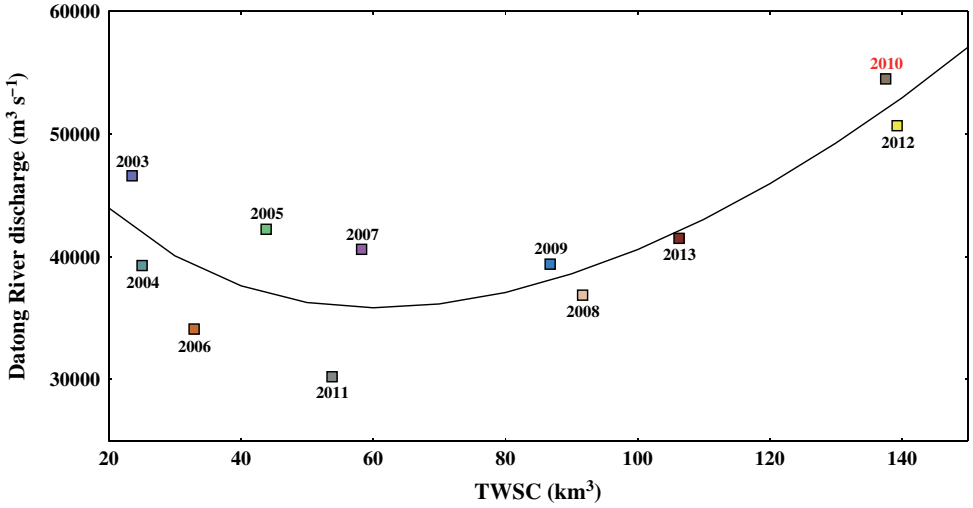

Fig. 5. Relationship between TWSC and river discharge for the mean of summer (June, July, and August) during 2003–2013 in Yangtze River basin

YRB Flood, 2010

Flooding in the YRB often occurs in the summer, so this paper used the relationship between TWSC and river discharge (T-D) in the mean of JJA during 2003–2013 (Fig. 5), mainly because the flow in the summer would not be impacted by the freezing of ice and snow. As shown in Fig. 5, the T-D relationship in the summer can be fitted by a least-squares exponent regression. Moreover, when the T-D relationship approached the regression line (i.e., high flow corresponding with high TWSC), it indicated local saturated river discharge, which was similar to the event of saturation exceeding the discharge. During these saturation-driving events, the river basin discharges its large area of storage water into the watercourse and its surrounding plain in order to release its saturation state. Therefore as time goes on, the TWSC and river discharge increase, so large-area flooding will occur at this time, as happened in 2010 (Fig. 5).

This paper used the YRB's flood in the summer of 2010 as an example and used a simple time-lagged autoregressive model to establish the relationship between river discharge and the total water storage changes from GRACE and compared this with basin observations from traditional operational measures such as soil moisture.

Fig. 6 shows the changes in the model-forecast discharge within the lead time from January to December. The results show that the compliance between the model-forecast discharge changes and the actual situation gradually decreased with the increase of the lead time. Over lead times from 1 to 12 months, the forecast of river discharge using three-term autoregressive model with added TWSC was closest to the actual situation when compared with the Soil three-term autoregressive model and the basic two-term autoregressive model. When the lead time was 1–6 months, the result of the forecast of river discharge from the Soil three-term autoregressive model was better than that of the basic two-term autoregressive model, but beyond 6 months, the basic two-term autoregressive model outperformed the Soil three-term autoregressive model. In particular, when the lead time exceeded 6 months, the maximum value of the model-forecast discharge changes decreased quickly.

Fig. 7 shows the properties of the lead time regression model as follows: Fig. 7(a) illustrates the relationship between the maximum flow during the forecast period and the lead time (1–12 months); the dashed line represents the 99th percentile value; Figs. 7(b-e) show the relationship between the forecast period and the maximum absolute value error, MAE, mean, and MPRE, respectively.

The three-term autoregressive model with the soil moisture data improved the MAE from 2-month to 7-month lead times compared with the basic two-term model (Fig. 7). However, it cannot forecast exceptional river discharge from 2 to 3 months before the event. In the trinomial regression model after adding TWSC, the maximum flow increased and the maximum value in all the lead times was the largest. Compared with the other data, only the single TWSC model had the response ability for extreme events with 2–3 months lead time (99th percentile larger than the historical data). When the lead time was 3–6 months, the forecasted river basin of the GRACE TWSC model approached but did not exceed the 99th percentile, which indicates that it has the potential to forecast the flood conditions in advance. Moreover, the maximum absolute value error and MAE with the addition of GRACE data were smaller than for any other tested data, and the mean bias and MPRE with GRACE data were the smallest in the whole prediction period, which shows that the GRACE data can improve the accuracy of the time-lagged autoregressive model. However, the accuracy decreased when the lead time was 1 month, as a result of early warning (overprediction).

Conclusions and Discussion

The major flooding which occurs in river basins requires high water storage and heavy rainfall. The aggravation of ice and snow melting, increases in frozen soil, increases in water levels, and so on, will lead to increases in river discharge, which thereby cause flooding. However, in addition to the contribution of these factors to the runoff in river basins, high-flow events can be observed by the GRACE mission. Comparing the traditional measurements, this paper showed that the basin-scale estimate of water storage changes inferred from GRACE satellite observations of time-variable gravity can characterize regional flood potential and ultimately result in longer lead times in flood warnings. Therefore it was shown to be an effective tool. However, earlier forecasting of flood conditions in river basins will require accurate and complete hydrological state information.

The primary aim of this work is not to design the best performance and optimal prediction model. Rather, it is the use of information in the GRACE-based total water storage changes from the

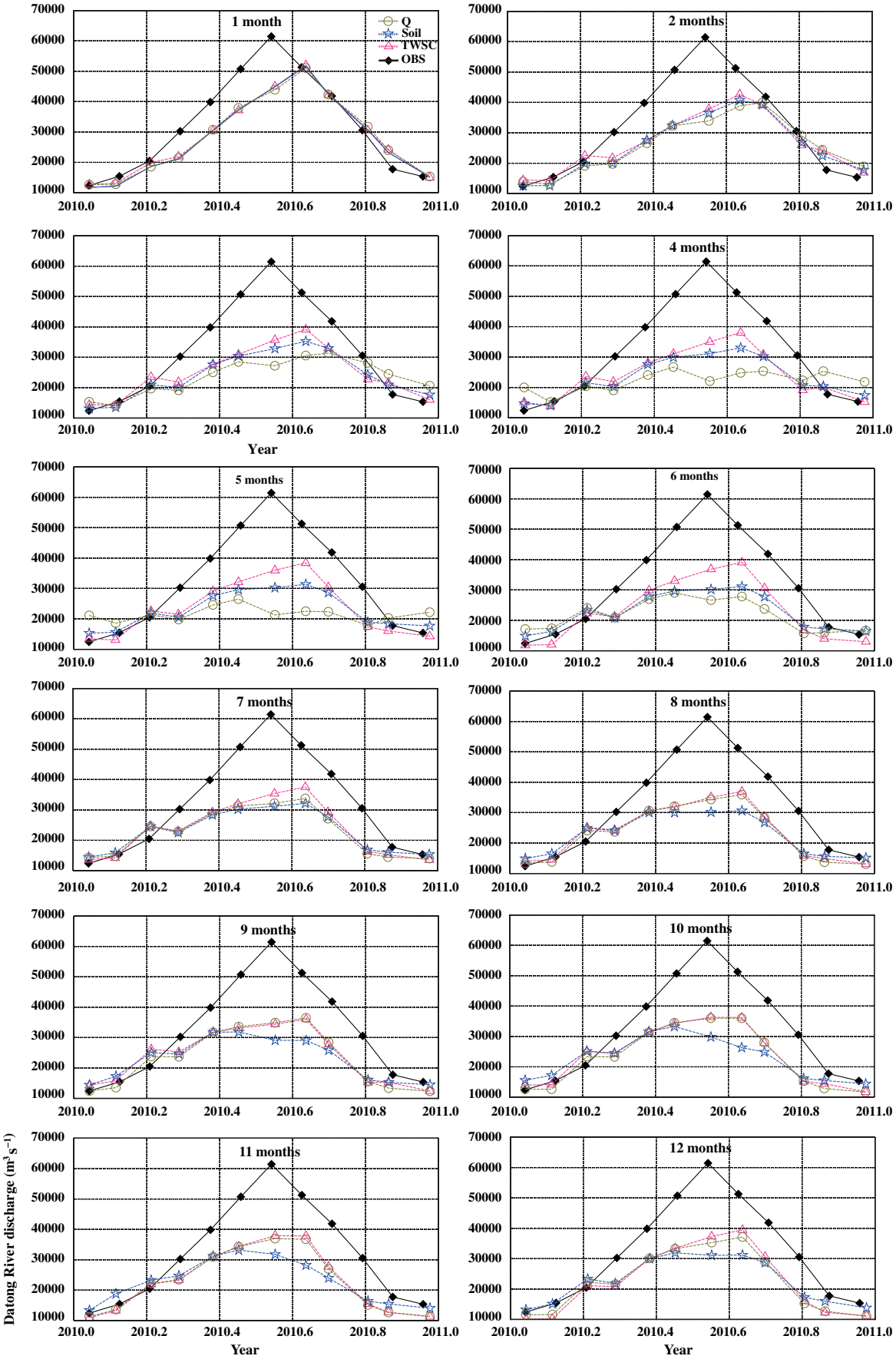

Fig. 6. Autoregressive model results for increasing lead times from 1 to 12 months; OBS is the actual river discharge observations from Datong, Q is the predicted river discharge from Eq. (18), Soil represents the predicted river discharge from Eq. (19) when the soil moisture data are included, TWSC represents the predicted river discharge from Eq. (20) when the total TWSC are included

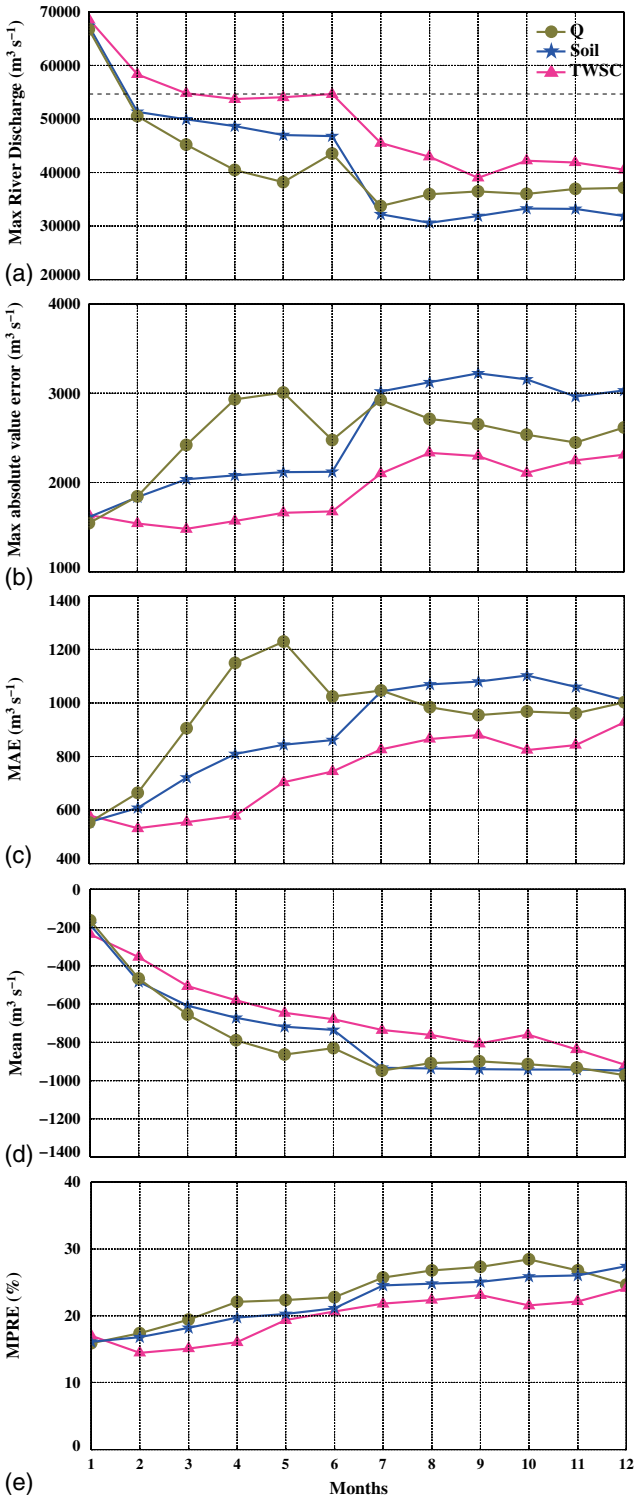

Fig. 7. Performance of autoregressive model: (a) maximum predicted river discharge during the forecast period versus model lead time for two-term basic discharge model and three-term model driven by soil moisture and GRACE TWSC, with the 99th percentile of historic river discharge (dashed line); (b) maximum absolute value error over the entire forecast period versus model lead time; (c) MAE over the entire forecast period versus model lead time; (d) mean bias between model and observations over the entire forecast period versus model lead time; (e) MPRE over the entire forecast period versus model lead time; *Q* is the the two-term basic autoregressive model, Soil is the three-term autoregressive model with added soil moisture data, and TWSC is the three-term autoregressive model with added total TWSC

Lagrange multiplier method to assess several months in advance the predisposition of a river basin to flooding. The authors hope that this work can be applied to regional flood forecasting. The GRACE observations can be best implemented for a flood prediction tool in a case of regional-scale saturation and the accumulation of storage in subsurface basin, such as the entire YRB. Therefore this work takes into account the spatial discrepancy of the YRB, which is divided into 11 subbasins (Jinsha River, Min River, Jialing River, Han River, Upper YRB, Middle YRB, Lower YRB, Wu River, Doting Lake, Poyang Lake, and Tai Lake) to study the flood potential.

In addition, the GRACE observational results and precipitation data can be combined to establish flood factors for the purpose of regional flood monitoring. Therefore, GRACE satellite observational results play an important role in flood monitoring and provide new technological means for the monitoring of flood conditions. With the accumulation of satellite gravity and hydrometeorology data as well as the gradual completeness of hydrological modeling, conditions will become more favorable for the separation and extraction of geophysics signals from satellite gravity-observation results. Moreover, this data not only can provide reliable geophysics information for studying the Earth's solid-matter migration and for refining and verifying global and local hydrological models, but also can forecast major disaster events, such as drought, floods, debris flow, earthquakes, and so on.

Acknowledgments

The authors thank the following data providers for making the data available: GRACE-CSR, JPL, GFZ, GRGS; GLDAS; CPC; and WGHM. This study is supported by the NSFC (China) under Grants 41274032, 41474018, and 41429401; by the National 973 Project of China under Grants 2013CB733301 and 2013CB733302; by the Basic Research Foundation 16-01-01 of the Key Laboratory of Geospace Environment and Geodesy of Ministry of Education, Wuhan University; and by the Open Research Fund Program of the State Key Laboratory of Geodesy and Earth's Dynamics (Grant No. SKLGED2017-2-2-E).

References

Appleby, V. C. (1970). "Recession and the base flow problem." *Water Resour. Res.*, 6(5), 1398–1403.

Bettadpur, S. (2012). "Level-2 gravity field product user handbook." *GRACE 327–734*, The GRACE Project, Center for Space Research, Univ. of Texas at Austin, Austin, TX.

Chao, N. F., Wang, Z. T., Jiang, W. P., and Chao, D. B. (2016). "A quantitative approach for hydrological drought characterization in southwestern China using GRACE." *Hydrogeol J.*, 24(4), 893–903

Chau, K. W., and Wu, C. L. (2010). "A hybrid model coupled with singular spectrum analysis for daily rainfall prediction." *J. Hydroinformatics*, 12(4), 458–473.

Chen, J. L., Wilson, C. R., and Seo, K. W. (2009). "S2 tide aliasing in GRACE time-variable gravity solutions." *J. Geod.*, 83(7), 679–687.

Chen, X. Y., Chau, K. W., and Busari, A. O. (2015). "A comparative study of population-based optimization algorithms for downstream river flow forecasting by a hybrid neural network model." *Eng. Appl. Artif. Intell.*, 46(Part A), 258–268.

Cheng, M. K., Ries, J. C., and Tapley, B. D. (2011). "Variations of the Earth's figure axis from satellite laser ranging and GRACE." *J. Geophys. Res.*, 116(B1), B01409.

Crowley, J. W., Mitrovica, J. X., Bailey, R. C., Tamisiea, M. E., and Davis, J. L. (2006). "Land water storage within the Congo Basin inferred from GRACE satellite gravity data." *Geophy. Res. Lett.*, 33(19), L19402.

Döll, P., Kaspar, F., and Lehner, B. (2003). "A global hydrological model for deriving water availability indicators: Model tuning and validation." *J. Hydrol.*, 270(1), 105–134.
Dool, H. V. D., Huang, H. J., and Fan, Y. (2003). "Performance and analysis of the constructed analogue method applied to US soil moisture applied over 1981–2001." *Geophys. Res. Lett.*, 108(D16), 8617.
Famiglietti, J. S., et al. (2011). "Satellites measure recent rates of groundwater depletion in California's Central Valley." *Geophys. Res. Lett.*, 38(3), L03403.
Farrell, W. E. (1972). "Deformation of the Earth by surface loads." *Rev. Geophys. Space Phys.*, 10(3), 761–797.
Geruo, A., Wahr, J., and Zhong, S. (2013). "Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: An application to glacial isostatic adjustment in Antarctica and Canada." *Geophys. J. Int.*, 192(2), 557–572.
Gholami, V., Chau, K. W., Fadaee, F., Torkaman, J., and Ghaffari, A. (2015). "Modeling of groundwater level fluctuations using dendrochronology in alluvial aquifers." *J. Hydrol.*, 529(3), 1060–1069.
GRGS. (2016). "The marvellous grace plotter." ⟨http://www.thegraceplotter.com/⟩ (Dec. 5, 2016).
Guo, Y., Liu, S., and Baetz, B. W. (2012). "Probabilistic rainfall-runoff transformation considering both infiltration and saturation excess runoff generation processes." *Water Resour. Res.*, 48(6), W06513.
Heiskanen, W. A., and Moritz, H. (1967). *Physical geodesy*, Freeman W H and Company, San Francisco.
Kusche, J., Schmidt, R., Petrovic, S., and Rietbroek, R. (2009). "Decorrelated GRACE time-variable gravity solutions by GFZ, and their validation using a hydrological model." *J. Geodesy*, 83(10), 903–913.
Müller, S. H., et al. (2014). "Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration." *Hydrol. Earth Syst. Sci.*, 18(9), 3511–3538.
Ogawa, R., Chao, B. F., and Heki, K. (2011). "Acceleration signal in GRACE time-variable gravity in relation to interannual hydrological changes." *Geophys. J. Int.*, 184(2), 673–679.
Ramillien, G., Famiglietti, J. S., and Wahr, J. (2008). "Detection of continental hydrology and glaciology signals from GRACE." *Surv. Geophys.*, 29(4–5), 361–374.
Reager, J. T., et al. (2015). "Assimilation of GRACE terrestrial water storage observations into a land surface model for the assessment of regional flood potential." *Remote Sens.*, 7(11), 14663–14679.
Reager, J. T., and Famiglietti, J. S. (2009). "Global terrestrial water storage capacity and flood potential using GRACE." *Geophys. Res. Lett.*, 36(23), L23402.
Reager, J. T., Thomas, B. F., and Famiglietti, J. S. (2014). "River basin flood potential inferred using GRACE gravity observations at several months lead time." *Nature Geosci.*, 7(8), 588–592.
Riegger, J., and Tourian, M. J. (2014). "Characterization of runoff-storage relationships by satellite gravimetry and remote sensing." *Water Resour. Res.*, 50(4), 3444–3466.
Riegger, J., Tourian, M. J., Devaraju, B., and Sneeuw, N. (2012). "Analysis of grace uncertainties by hydrological and hydro-meteorological observations." *J. Geodyn.*, 59–60, 16–27.
Rodell, M., et al. (2004). "The global land data assimilation system." *Bull. Am. Meteorol. Soc.*, 85(3), 381–394.
Rodell, M., Velicogna, I., and Famiglietti, J. S. (2009). "Satellite-based estimates of groundwater depletion in India." *Nature*, 460(7258), 999–1002.
Siccardi, F., Boni, G., Ferraris, L., and Rudari, R. (2005). "A hydrometeorological approach for probabilistic flood forecast." *J. Geophys. Res.*, 110(D5), D05101.
Sneeuw, N., et al. (2014). "Estimating runoff using hydro-geodetic approaches." *Survey Geophys.*, 35(6), 1333–1359.
Swenson, S., and Wahr, J. (2002). "Methods for inferring regional surface mass anomalies from gravity recovery and climate experiment (GRACE) measurements of time-variable gravity." *J. Geophys. Res.*, 107(B9), 2193.
Swenson, S. C., Chambers, D. P., and Wahr, J. (2008). "Estimating geocenter variations from a combination of GRACE and ocean model output." *J. Geophys. Res.*, 113(B8), B084.
Syed, T. H., Famiglietti, J. S., and Chambers, D. (2009). "GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales." *J. Hydrometeorol.*, 10(1), 22–40.
Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R. (2008). "Analysis of terrestrial water storage changes from GRACE and GLDAS." *Water Resour. Res.*, 44(2), W02433.
Taormina, R., and Chau, K. W. (2015). "Data-driven input variable selection for rainfall-runoff modeling using binary-coded particle swarm optimization and extreme learning machines." *J. Hydrol.*, 529(3), 1617–1632.
Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C. (2004). "The gravity recovery and climate experiment mission overview and early results." *Geophys. Res. Lett.*, 31(9), L09607.
Wagner, C., McAdoo, D., Klokočník, J., and Kostelecký, J. (2006). "Degradation of geopotential recovery from short repeat-cycle orbits: Application to GRACE monthly fields." *J. Geodesy*, 80(2), 94–103.
Wahr, J., Molenaar, M., and Bryan, F. (1998). "Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE." *J. Geophys. Res.*, 103(B12), 30205–30229.
Wang, W. C., Chau, K. W., Xu, D. M., and Chen, X. Y. (2015). "Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition." *Water Resour. Manage.*, 29(8), 2655–2675.
Wetterhall, F., et al. (2013). "Forecasters priorities for improving probabilistic flood forecasts." *Hydrol. Earth Syst. Sci. Discuss*, 10(2), 2215–2242.
Wu, C. L., Chau, K. W., and Li, Y. S. (2009). "Methods to improve neural network performance in daily flows prediction." *J. Hydrol.*, 372(1–4), 80–93.
Zhang, D., Zhang, Q., Werner, A. D., and Liu, X. M. (2016). "GRACE-based hydrological drought evaluation of the Yangtze River basin, China." *J. Hydrometeorol.*, 17(3), 811–828.