Journal of Hydrology 519 (2014) 50–68

Contents lists available at ScienceDirect

Journal of Hydrology

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

Estimation of the terrestrial water budget over northern China by merging multiple datasets

Yunjun Yao [a,*], Shunlin Liang [a,b], Xianhong Xie [a], Jie Cheng [a], Kun Jia [a], Yan Li [c], Ran Liu [c]

[a] State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing 100875, China
[b] Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA
[c] Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, China

ARTICLE INFO

Article history:
Received 26 September 2013
Received in revised form 13 June 2014
Accepted 25 June 2014
Available online 12 July 2014
This manuscript was handled by Laurent Charlet, Editor-in-Chief, with the assistance of Thierry Pellarin, Associate Editor

Keywords:
Terrestrial water budget
Merging technique
Linear weighting method
Multiple datasets
Northern China

SUMMARY

The terrestrial water budget over northern China, which plays an important role in water resource management, has experienced great changes during the past decades. However, its spatiotemporal variations in the past calculated from individual datasets remain quite uncertain. In this study, we improve the accuracy of evapotranspiration (E), precipitation (P) and runoff (R) estimates by merging remote sensing, reanalysis, data assimilation datasets and ground observations, and further analyze the spatiotemporal characterization of the terrestrial water budget at 0.25° over northern China during the period of 1984–2010. The results illustrate that using any of the individual datasets, there is significant uncertainty and an obvious seasonal cycle in the terrestrial water budget. Large differences exist among the different datasets, and the merged E, P and R outperform the individual datasets. The root mean square errors (*RMSEs*) from cross-validation are 8.4–14.2 mm, 15.9–27.3 mm and 4.1–14.2 mm for the monthly merged E, P and R at the site scale of the different basins, respectively. The spatial patterns of the merged annual E and R are consistent with that of P due to the water limitations mainly controlled by P. The interannual variations in these hydrological variables indicate a slight increase in the variables from 1984 to 1998, with a large El Niño event, and a larger decline thereafter as a result of a large-scale drought. However, decadal trends in terrestrial water storage changes (*TWSC*) over all five basins inferred from the merged products tend to increase to some extent with climate warming over the studied time period. The Budyko curve reveals that an increase in vegetation coverage increases the evaporation ratio (E/P) to some extent, but climate change is the dominant driver for the variations in the hydrological variables in these regions.

1. Introduction

The terrestrial water budget over northern China plays an important role in water resource management. Within the last few decades, because of the acceleration of reforestation caused by human activity and the frequent occurrence of large-scale drought across northern China, water resources have attracted widespread attention from scientists and governments (Cao, 2008; Dai et al., 2004; Liu et al., 2008; Ma and Fu, 2006; Wang et al., 2011; Yang et al., 2009). An accurate quantification of the variations in the water budget components of the regional terrestrial hydrological cycle is crucial to determining the linkages between climate change and changes in water resources. Point measurements from both meteorological and hydrological stations illustrate that since the 1960s, increases in annual mean precipitation (P) and runoff (R) have occurred in northwestern and eastern China, whereas decreases in annual mean P and R has occurred in northeastern China (Chen et al., 2006; Li and Ding, 2012; Meng and Mo, 2012; Wang and Zhou, 2005; Wang et al., 2012). These changes, along with altered P patterns, have likely led to the significant variations in both evapotranspiration (E) and streamflow volumes; the underlying evaluation of these changes is extremely important for proper water resource management (Liu et al., 2013a; Wang and Zhou, 2005).

Many scholars have attempted to estimate the terrestrial water budget over northern China and over the Northern Hemisphere to understand variations in the regional and global hydrological cycles. However, there is large uncertainty in estimates of spatial and temporal variations in the terrestrial water budget (Adam et al., 2007; Sahoo et al., 2011; Troy et al., 2011; Yang et al., 2012; Zhai et al., 2005; Zhang et al., 2010). For P and E, although there are some meteorological, hydrological point and flux tower measurements available, sparse observations and field experiments

* Corresponding author. Tel.: +86 10 5880 3002.
E-mail address: boyyunjun@163.com (Y. Yao).

http://dx.doi.org/10.1016/j.jhydrol.2014.06.046
0022-1694/

prohibit the accurate characterization of the spatial and temporal patterns of *P* or *E* over large spatial scales (Serreze et al., 2005; Sun et al., 2005). Reanalysis and remote sensing datasets can map *P* at various spatial scales, ranging from small "point" to large "continental" scales. Bosilovich et al. (2008) compared five reanalysis *P* fields and found that the 40-year European Centre for Medium-Range Weather Forecasts (*ECMWF*) Reanalysis (*ERA*-40) produces reasonable observed *P* over Northern Hemisphere continents. In addition, *ERA-Interim* provides a global integrated and coherent water resources estimate, which is superior to the *ERA*-40 in several aspects, such as annual cycle and drought detection (Belo-Pereira et al., 2011; Betts et al., 2003). Currently, the increasing availability of large-scale *P* datasets also highlights the differences in trend analysis between different models rather than the uncertainty of the true *P* value (Fan et al., 2011).

In contrast to the large number of ground-measurements for *P* that are available, there are very few direct measurements of actual *E* over northern China. Although prognostic assessments of variations in *E* and discharge over northern China during the past few decades have been explored using different models and conventional hydrological, meteorological, or ground-measurement datasets (Brutsaert and Parlange, 1998; Gao et al., 2007; Liu et al., 2011; Su et al., 2006; Walter et al., 2004; Wang et al., 2010a,b; Wu et al., 2012a,b; Yang et al., 2011; Yao et al., 2012), a better understanding of the mean state and variability of the energy and water budget components over these decades remains elusive because they are highly heterogeneous both spatially and temporally (Yao et al., 2014a,b). Multiple data sources, such as remote sensing, reanalysis datasets and hydrologic models, can bridge this gap. Recently, Liu et al. (2012) used hydrologic model and daily meteorological data and found an obviously increasing trend in *E* over eastern China during the period of 1961–2005. Other satellite-based datasets, such as the Gravity Recovery and Climate Experiment (*GRACE*) datasets, are widely used to quantify variations in water budget components at regional and global scales (Güntner et al., 2007; Moiwo et al., 2012; Munier et al., 2012; Papa et al., 2008; Troy et al., 2011).

Because of the large uncertainty in single datasets, the use of these datasets does not satisfy the systematic requirements for climate study and other applications. For example, *E* from reanalysis datasets such as the National Centers for Environmental Prediction (*NCEP*)/National Center for Atmospheric Research (*NCAR*) Reanalysis (*NNR*) data is not considered reliable because it is constrained not by *P* and radiation but by circulation and temperature observations (Betts et al., 2003; Huffman et al., 2001; Kalnay et al., 1996; Liang et al., 2010; Ruiz-Barradas and Nigam, 2005). Fortunately, multiple datasets merging technique can effectively reduce the uncertainty in terrestrial water budget estimations. A few studies concerning terrestrial water budgets have focused on merging multiple datasets. Outside of northern China, Alavi et al. (2006) applied a Kalman filter to estimate state variables and thus correct measurements, further filling gaps in *E* data over northwestern Guelph, Ontario, Canada. Recently, Troy et al. (2011) applied both multiple datasets and observed *R* to calculate components of the terrestrial water cycle over northern Eurasia. However, in the regions of northern China, there is a lack of similar studies that investigate the terrestrial water cycle by merging multiple datasets. As a result, little is accurately known about the spatiotemporal characterization of the response of the terrestrial water budget over northern China to climate change on large spatial scales and over long time periods.

In this study, we improve the accuracy of evapotranspiration (*E*), precipitation (*P*) and runoff (*R*) estimates by merging remote sensing, reanalysis, data assimilation datasets and ground observations, and further analyze the spatiotemporal characterization of the terrestrial water budget (1984–2010) over northern China. Our study has three major objectives. First, we evaluate the discrepancies in remote sensing, reanalysis and data assimilation datasets by analyzing the seasonal cycle of the terrestrial water budget. Second, we aim to develop a linear weighting method to improve the accuracy of the three hydrological components by merging multiple datasets and ground observations. Finally, we generate regional long-term (1984–2010) monthly merged *E*, *P* and *R* records to analyze the spatiotemporal characterization of the terrestrial water budget on a basin scale. The study area includes five river basins: the Heilongjiang, Liaohe, Haihe, Yellow River, and Inland River basins, as shown in Fig. 1.

2. Data

The datasets used in this study include remote sensing retrievals, reanalysis products, data assimilation datasets and ground-based observations. *E* has been estimated using several methods from remote sensing retrieval, reanalysis products, and data assimilation datasets. *P* has been extracted from reanalysis and remote sensing data. *R* has been derived from reanalysis products and data assimilation datasets. Moreover, terrestrial water storage changes (*TWSC*) has been calculated based on the water budget equation using merged *E*, *P* and *R*. The required datasets and ground-based observations are listed in Tables 1 and 2 and are described briefly below.

2.1. Remote sensing datasets

Remote sensing datasets can be used to retrieve *E* and *P*. Two remote sensing products of *E* are used in this study. One such product is derived from the Penman–Monteith (*PM*) equation (Shuttleworth, 1993), forced by inputs from the Global Energy and Water Cycle Experiment (*GEWEX*) products (Mu et al., 2011; Zhang et al., 2009) and the monthly Normalized Difference Vegetation Index (*NDVI*) products at 8-km spatial resolution (1984–2010) from the Global Inventory Modeling and Mapping Studies (*GIMMS*) group at the National Aeronautics and Space Administration (*NASA*) Goddard Space Flight Center (Tucker 2005) (referred to as the *GEWEX-PM* dataset). The other dataset uses the modified satellite-based Priestly–Taylor (*PT*) algorithm (*MS-PT*), driven by monthly Moderate Resolution Imaging Spectroradiometer (*MODIS*) products, including albedo, Land Surface Temperature (*LST*), surface emissivity, *NDVI* and incent solar radiation from the Japan Aerospace Exploration Agency (*JAXA*) (Yao et al., 2013) (called *MODIS-PT* dataset). To calculate *P*, we use the gauge-corrected satellite dataset from the Global Precipitation Climatology Project (*GPCP*) (Huffman et al., 2001) and two satellite-only products, the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (*PERSIANN*) (Hong et al., 2004) and the Climate Prediction Center Morphing Technique (*CMORPH*) (Joyce et al., 2004). The satellite *GRACE* product (Center for Space Research Release 4: *CSR RL*04) (Swenson and Wahr, 2002) is acquired to compare and analyze the water storage changes. The monthly Palmer Drought Severity Index (*PDSI*) products (Dai et al., 2004), derived from the *NCAR CGD*'s Climate Analysis Section dataset, with a 2.5° spatial resolution are also used to analyze variations in land surface drought.

2.2. Reanalysis and data assimilation datasets

We use two reanalysis products to characterize the water budget components. The first is the *ERA-Interim* (Simmons et al., 2006), a global reanalysis product from *ECMWF* with a 4*D* variational assimilation system at *T*255 horizontal resolution. The second dataset is the Modern Era Retrospective-analysis for Research