Quantifying the Space – Time Variability of Water Balance Components in an Agricultural Basin using a Process-Based Hydrologic Model and the Budyko Framework

Han Qiu[1], Jie Niu[2*], and Mantha S. Phanikumar[1*]

[1]Department of Civil and Environmental Engineering
Michigan State University, East Lansing, MI 48824
[2]Institute of Groundwater and Earth Sciences, Jinan University, Guangzhou 510632, China

*Corresponding Author email: phani@msu.edu, jniu@jnu.edu.cn

Abstract

Process-based distributed hydrologic models (PBHMs), which link watershed characteristics with process representations, are useful tools to evaluate both the distributed and ensemble hydrologic responses of a basin to climate inputs. However, complexities associated with parameter interactions and their spatial heterogeneities may produce high uncertainties in the parameterization of a PBHM. The Budyko curve framework presented in this work offers an effective approach for evaluating variabilities in the water balance components using a PBHM and explores the link between model performance with parameter heterogeneities and the Budyko curve characteristics. The PBHM was calibrated using a multi-site calibration strategy (MLT), which was built upon a step-wise calibration algorithm combined with multiple calibration targets including river discharges, evapotranspiration and ground water heads to reduce the compensation errors caused by component interactions. This strategy was tested for the Kalamazoo River watershed in Michigan, USA, with obvious physiographic and land surface heterogeneities. The Budyko framework characterized the water balance variability at the sub-watershed scale; two empirical methods are employed to evaluate calibrated parameters using Budyko-estimated values and to assess the physical relevance of parameters. The relative infiltration capacity is found to play an important role in affecting the spatial variability of the annual water balances of this watershed. This work brings out the importance of optimizing calibration strategies by linking catchment heterogeneities with processes reasoning in order to understand the underlying hydrologic controls.

Keywords: hydrologic models, model calibration, water budgets, Budyko curve

1. Introduction

Water resources management in an era of global change requires hydrologists to provide reliable predictions of water fluxes and to analyze and interpret their distributed and evolving roles (Wagener et al., 2010; McDonnell and Beven, 2014). At the catchment scale, it is of great interest to link the structure of a watershed to its response to climate variability and to evaluate water budget components and their variability (Sivapalan, 2006). Two approaches have been used in the past decades to study annual water budgets and their inter-annual variability: (a) the empirical approach and (b) process-based modeling (Sivapalan et al., 2011). One classic example of an empirical approach is Budyko's work (Budyko, 1974), which assumes that the partition of precipitation into evaporation and runoff could be determined from available water measured with precipitation (*P*) and available energy measured with potential evapotranspiration (*PET,* E_p). Based on the Budyko hypothesis, the ratio of actual evapotranspiration over precipitation (*ET/P*), i.e., the evaporation ratio, is fundamentally related to the ratio of potential evapotranspiration over precipitation (E_p /P), i.e., the climate dryness index (Budyko, 1974):

$$\frac{ET}{P} = \left[\frac{E_p}{P} \tanh\left(\frac{E_p}{P}\right) \left\{ 1 - \cosh\left(\frac{P}{E_p}\right) + \sinh\left(\frac{P}{E_p}\right) \right\} \right]^{0.5} \tag{1}$$

The Budyko framework was used to evaluate the inter-annual variability of annual water balances as well as water balances at seasonal time scales. For example, Yang et al. (2007) evaluated long time series of climate data and discharge in 108 non-humid catchments of China using the Budyko framework and found that the inter-annual variability of water-energy balance can be expressed with infiltration capacity, soil water storage capacity and the average ground surface slope. Wang (2012) studied effects of annual water storage changes on the inter-annual water balances of 12 watersheds in Illinois based on long-term soil moisture and groundwater level observations using the Budyko framework. The observed deviations of Budyko-type curves for different watersheds can reflect the ensemble effects of the climate fluctuations (Milly, 1994) and watershed characteristics such as storage (Fang et al., 2016) and soil moisture capacity, topography and soil properties (Yokoo et al., 2008), vegetation type and dynamics (Zhang et al., 2001; Oudin et al., 2008; Domohue et al., 2012) on the mean annual water balances. However, as with most empirical approaches, the derivation of Budyko-type curves lacks the explicit representations of interactions of climate inputs and various hydrologic processes. It is difficult for empirical approaches to distinguish the effects of different hydrologic processes, especially when considering the intra-annual (e.g. daily) variability of water balances (Chen et al., 2013; Fang et al., 2016).

PBHMs, on the other hand, describe the hydrologic processes explicitly and provide a direct link between catchment structures and response behaviors. Simulations based on PBHMs are suitable for distinguishing the distributed flow pathways (Beven, 2002), quantifying the water storage changes in hydrologic systems (Niu et al., 2014) and understanding the physical processes in controlling the hydrologic responses (Shen et al., 2013). They require, however, a large amount of input data in representing the physical processes and abundant observed data to calibrate the parameters (Beven and Binley, 1992; Ragettli and Pellicciotti, 2012). Co-evolution of water budget components in a PBHM with various parameter suits may contribute to similar ensemble predictions; this phenomenon has been famously generalized by (Beven, 1993; and Beven and Freer 2001) as the equifinality problem. Calibration against a single target (e.g., stream discharge) does not guarantee that internal processes are all correctly simulated; compensation errors could be produced with discrepancies coexisting in more than one hydrologic component representing the processes. For example, ungauged infiltration and lateral groundwater flow processes may cause underestimated (overestimated) groundwater supply while compensated by overestimated (underestimated) surface runoff to achieve comparable stream discharge results (Ragettli and Pellicciotti, 2012). An effective method for reducing uncertainty in parameter identification in PBHMs might be through evaluations against a number of responses representing different hydrological components (Anderton et al., 2002). Multistep, multi-site, and multivariable calibration methods are being widely applied by calibrating different internal processes to improve both the overall and distributed model performances (Ragettli and Pellicciotti, 2012; Stahl et al., 2008; Sutanudjaja et al., 2014; Choi et al., 2015), e.g., step-wise optimization using stream discharge, groundwater heads, and remotely-sensed soil moisture data as the optimization objectives; and calibration to data from multiple gauging stations in different sub-watersheds.

The objectives of this paper are to: (1) quantify the spatial and temporal variability of annual water balances in a semi-humid watershed in Michigan, USA using a PBHM; (2) use a multi-site calibration strategy combined with multiple criteria evaluations to understand how parameters, process representations and water budget results change across scales; and (3) evaluate calibrated parameters with Budyko-estimated values to assess the physical relevance of parameters and to identify deficiencies in each methodology. Briefly, we addressed the question of how different empirical equations for the parameter ω in the Budyko curve formulation compare with simulated results based on a PBHM in a Great Lakes watershed.

To address these questions, we use a PBHM, PAWS (Process-based Adaptive Watershed Simulator, *Shen and Phanikumar* (2010)). The model can simulate different hydrologic components and states including surface runoff, channel flow, groundwater, ET, soil moisture, soil temperature and changes in storage. Vegetation growth dynamics are also simulated by coupling PAWS with the land surface model CLM 4.0 as described in Shen et al. (2013; 2014). We first establish a calibration procedure to ensure that stream discharge and other important hydrologic components such as ET and groundwater are correctly simulated. A stepwise calibration method was applied to reduce the compensation errors. For reducing uncertainty in scaled heterogeneity, a multi-site calibration (MLT) method is employed and combined with multi-criteria evaluations. For the second effort, instead of attempting to elaborate on the heterogeneity within each process, we focus on inter-annual water balances at sub-watershed scale and study the ensemble catchment performances using the Budyko framework.

2. Methods

2.1 The Model

Governing equations and numerical details of the PAWS model have been explained in Shen and Phanikumar (2010) and presented in Table 2 of Niu et al. (2014). The coupling details of PAWS with CLM 4.0 have been extensively discussed and evaluated earlier (Shen et al., 2013; Riley and Shen, 2014; Niu et al., 2014; Niu and Phanikumar, 2015; Qiu et al., 2019).

Briefly, PAWS includes key hydrologic processes in the domains of surface flow, surface ponding, channel flow, unsaturated vadose zone and saturated groundwater flow. PAWS uses a structured grid and the finite-volume method to solve the governing partial differential equations in different hydrologic units. The overland flow governed by diffusive wave equation occurs in the surface flow domain while infiltration and evaporation happen in the ponding domain. Runoff occurs when the water depth of ponding domain is in excess of the interception depth. Water may also backfill into the ponding domain during flood conditions. The overland flow in the surface flow domain interacts and exchanges water with river segments. Channel flow is simulated using the diffusive wave equation and its exchange with groundwater is modeled based on the leakance concept (Gunduz and Aral, 2005). The vadose zone is simulated in 1-D columns connected to land surface cells at the top and saturated groundwater flow cells at the bottom. PAWS conceptualizes the unsaturated vadose zone as an array of vertical soil columns on the assumption that lateral fluxes in this domain are negligible. The saturated - unsaturated soil water flow is governed by the Richards equation with the vegetation uptake as a sink term. PAWS uses the concept of root water efficiency (Lai and Katul, 2000) to adjust the vegetation root water uptake fluxes along the soil column. The van Genuchten formulation is employed for soil water retention calculations. Field capacity, saturated water content and wilting point are set in correspondence to the soil properties. Phase change is also considered by applying the freezing-point depression formula in (Niu and Yang, 2006) to reduce hydraulic conductivities in freezing soils. The last computational cell of the soil column, whose thickness changes as the water table fluctuates, serves as the linkage between vadose zone and groundwater. Quasi- 3D Groundwater equation derived from Darcy's law is employed for solving groundwater flow. The vegetation dynamics, energy cycling, and carbon/nutrient cycling are incorporated through CLM (Oleson et al., 2010).

2.2 Sites and Data Sources

2.2.1 Site descriptions

The Kalamazoo River Watershed (KRW) is located in the southwest portion of the lower peninsula of Michigan (Figure 1). This watershed has a drainage area of approximately 2,020 square miles (5,200 km^2) and it drains portions of nine counties in Michigan. The Kalamazoo River stretches 130 miles (210 km) from the junction of its north and south branches to its outlet at Lake Michigan. Based on the 10-digit Hydrologic Unit Code (HUC), the KRW is divided into 9 sub-watersheds (Figure. 1). Substrates in the headwaters and upstream segment consist of mostly sand, gravel and some cobble. The substrates in the streams of the middle segment are dominated by gravel and cobble. In contrast, the substrates in the mouth segment streams are mostly composed of sand and silt. Annual mean precipitation in this area averages ~970 mm and there is an increase of annual snowfall from the head waters to the mouth area due to the lake effect (Wesley, 2005). Average growing season increases from ~150 days at the eastern end of the watershed to ~180 days near Lake Michigan. The land surface elevation ranges from 175 to 380 meters above the sea level and varies distinctly in different sub-basins. Seasonal ET demands vary throughout the year as the solar radiation and air temperature change. The bedrock is mainly cold-water shale, overlain by the glacial deposits composed of outwash sand and gravel, which form the unconfined aquifer. Dominance of the soil types in the watershed includes clay, silt, sand, and organic materials. The land use and land cover (LULC) types for this watershed are occupied by approximately 47% agriculture (dominated by corn and soybeans), 21% forest, 9% open land, and 7% urban (Figure 2). Varied topography and heterogeneous subsurface properties, diverse vegetation and land use types render the watershed well suited for our study. We use a relatively fine grid resolution of 400m×400m for horizontal discretization which produces a 247×366 mesh for the whole watershed and 20 vertical layers to simulate the vadose zone dynamics and 2 layers for the groundwater domain (unconfined and confined aquifers).

2.2.2 Data sources

Details of data assimilation and data integration algorithms of PAWS are available in (Shen et al., 2014) thus we simply introduce the basic data input and processing information in this section. For river network simulation, National Hydrography Dataset (NHD) (https://nationalmap.gov/hydro.html) from U.S. Geological Survey (USGS) is used and reorganized as 'river segments' with a length of 400m in correspondence to our grid resolution. The 30 m resolution National Elevation Dataset (NED) from USGS serves as the Digital Elevation Model (DEM) for topographic calculations (e.g. surface slope and overland flow). To avoid possible compensation errors from other hydrologic components resulting from reduced channel density (Wang and Wu, 2013), we keep a relatively high river network density and include up to level-5 rivers (Figure. 1). NHD is overlaid on the NED model to extract a profile of elevations simulated as the riverbank elevations. A 30 m resolution raster data provided by the Michigan Department of Natural Resources (MDNR), i.e., the Integrated

Forest Monitoring Assessment and Prescription (IFMAP) data set (*MDTMB: Michigan Department of Technology*, *Management & Budget*, 2016) is employed as the Land use and Land cover layout. PAWS regroups the land use and land cover using a hierarchical stochastic selection method to do reclassifications while preserving the proportionality between the land use types of the original dataset (Shen et al., 2014). Soil type and properties information are obtained from Soil Survey Geographic (SSURGO) (Soil Survey Staff) database from the U.S. Department of Agriculture, Natural Resources Conservation Services (NRCS). Spatially distributed soil parameters are processed by the pedotransfer functions provided in Rosetta (Schaap et al., 2001) with an Artificial Neutral Network method to provide information of van Genuchten parameters and to calculate soil water retention properties and unsaturated conductivities. Climate data (e.g., precipitation, snowfall, daily maximum temperature and minimum temperature, and wind speed) are acquired from National Climatic Data Center (NCDC, 2010) of the National Oceanic and Atmospheric Administration (NOAA) and Michigan Automated Weather Station Network (MAWN) (Enviro-weather, 2016). The nearest neighbor interpolation scheme is used for spatial interpolation of climate data sets. Locations of meteorological stations are shown on Figure 1, marked as NCDC and MAWN stations respectively. We downloaded the evapotranspiration data from Moderate Resolution Imaging Spectroradiometer (MODIS) Global Evapotranspiration Project (MOD16) (http://www.ntsg.umt.edu/project/mod16), which is part of a NASA/EOS project to estimate global terrestrial ET from earth land surface using remote sensing.

2.3 Step-wise calibration

Instead of calibrating to river discharge data exclusively, the calibration procedure employed in this work uses several state variables involved in the quantification of major hydrologic fluxes. Besides the parameters of the PAWS model described in Shen et al. (2013), we estimate several additional parameters that are listed in Table S1 in Supporting Information (SI). To estimate spatially varied parameters such as hydraulic conductivity and to honor geology and the raw data, we use a linear transformation of the form $\mathbf{y} = a\mathbf{x} + b$ to adjust the parameters where $\mathbf{x}$ represents the original parameters which vary spatially, $\mathbf{y}$ is the transformed variable and a and b are constants. Based on the physical meaning and scale effect of each parameter, the parameters are adjusted as shown in Table 1 using operators which are either pure multipliers noted with a × (non-zero value of a and $b = 0$ in the above equation) or purely additive constants added to the original value noted with a + (that is, $a = 1$ and a non-zero value of b). The parameters are separated into three groups according to their relevance to certain processes in controlling the water fluxes, following similar approaches used by Stahl et al. (2008), Huss et al. (2008), and Ragettli and Pellicciotti (2012).

This procedure starts with adjusting the annual ET outputs in correspondence to MOD16 products. The aim of this step is to constrain the largest water flux in the model first, as the annual ET is approximately 55% to 75% of the total annual precipitation in this region (Kalamazoo River Watershed Council, 2011). Since we do not have spatially distributed soil moisture observations to constrain our vadose zone simulations, we also employ the annual average ET for controlling land surface processes (e.g., interception depth) and the plant soil-uptake processes using a tunable parameter γ in root water uptake efficiency (Lai and Katul,

2000) and other soil properties. Since the data integration schemes of PAWS have already incorporated the heterogeneity of vegetation types and soil properties, the van Genuchten soil parameters are slightly adjusted based on the initial parameters generated by Rosetta (A newer version Rosetta 3 is also available, Zhang and Schaap, 2017). In the second step, we focus on improving the comparisons of stream discharge calibrated to the USGS gauging observations at the outlet for each sub-watershed by adjusting the values of river bed conductivity, length of flow paths for runoff contribution to overland flow domain, and river bed elevation, all of which have uncertainties associated with. The river bed conductivity parameter (K_r) is spatially heterogeneous, and is initially estimated as $K_r = \sqrt{K_1 K_s}$ for each river segment grid (Shen et al., 2016), where K_1 is the first layer groundwater hydraulic conductivity from the well logic database, K_s is the soil vertical saturated hydraulic conductivity derived from SSURGO database. The final step is to calibrate the model for the steady state groundwater heads by adjusting the groundwater hydraulic conductivities. The differential evolution algorithm (Price et al., 2005) is finally employed to optimize the parameters by minimizing the objective function f (x), which represents model errors relative to observed values (Eqs. (2) – (6)). Details of the calibration procedure are illustrated in Figure 3. The upper and lower limits of the parameters were constrained within the scope of reasonable physical reasoning and were gradually adjusted during the calibration. The model performance was evaluated using the Nash-Sutcliffe efficiency metric (NASH) (Eq. 3), the absolute bias (APB) (Eq. 7), and the root mean squared error (RMSE) (Eq. 8). The RNASH metric is used for calibrating stream discharge to emphasize the baseflow contribution (Shen and Phanikumar, 2010), as shown in Eq. (4). For ET and groundwater heads, NASH is used to calculate f_i, as shown in Eq. (3).

$$f(\mathrm{x}) = \sum_{i=1}^{N} w_i f_i \tag{2}$$

$$NASH = 1 - \frac{\sum_{j=1}^{n} (O_j - P_j)^2}{\sum_{j=1}^{n} (O_j - \overline{O}_j)^2} \tag{3}$$

$$RNASH = 1 - \frac{\sum_{j=1}^{n} (\sqrt{O}_j - \sqrt{P}_j)^2}{\sum_{j=1}^{n} (\sqrt{O}_j - \overline{\sqrt{O}}_j)^2} \tag{4}$$

$$f_i = 1 - NASH_i \tag{5}$$

or, alternatively $$f_i = 1 - \left[\frac{NASH_i + RNASH_i}{2} \right] \tag{6}$$

$$APB = \frac{\sum_{j=1}^{n} (O_j - P_j) \times 100}{\sum_{j=1}^{n} O_j} \tag{7}$$

$$RMSE = \sqrt{\sum_{j=1}^{n} \frac{1}{n} (O_j - P_j)^2} \tag{8}$$

Here x denotes the vector of parameters; f_i denotes the objective functions and w_i denotes the weights; i equals unity for ET and steady groundwater head or denotes the i-th stream gauging stations ($i = 1, 2, \ldots, N$) for the stream discharge. Oj and Pj denote observations and simulations respectively. j is the j-th year for ET and groundwater heads or the j-th simulated day for stream discharge.

2.4 Multi-site calibration

To further examine the effects of local heterogeneity on the overall model performance, we divided the whole KRW into 4 sub-areas (Figure 1) and regionalized the parameter groups accordingly. This division was based not on the drainage areas corresponding to each of the gauging stations but on the distinct geologic and hydrologic characteristics of the watershed described in section 2.2.1. Each of the four sub-areas is loosely referred to as sub-watershed (SW) in this paper since they were formed by combining different 10-digit Hydrologic Unit Code (HUC) sub-watersheds. North Branch and South Branch KRWs are grouped into one region (SW1). Battle Creek watershed formed SW2 while Morrow Lake and Spring Brook watersheds are grouped into SW3. Gun river watershed, Rabbit River watershed and small KRW are grouped as SW4. Each SW is marked with a distinct color in Figure 1. The MLT calibration for all of the four SWs followed the step-wise calibration procedure described above. In particular, the stream discharge of each SW was calibrated to the observations from the stream gauging stations within the SW domain with equal weights assigned to all gauges within the SW.

2.5 Empirical equation for ω

Different mathematical formulations based on climate and catchment characteristics have been developed for the Budyko framework in the past (Budyko, 1974; Fu, 1981; Choudhury, 1999; Wang et al., 2009; Donohue et al., 2012; Xu et al., 2013; Liang et al., 2015). *Fu*'s equation is used in this work, which provides a relation between the dryness index and the evaporation ratio with an adjustable parameter ω $(1 \le \omega \le \infty)$ that represents catchment characteristics (Fu, 1981):

$$\frac{ET}{P} = 1 + \frac{E_p}{P} - \left[1 + \left(\frac{E_p}{P} \right)^{\omega} \right]^{1/\omega} \tag{9}$$

Yang et al. (2007) found that the parameter ω in Fu's equation is closely correlated with three dimensionless landscape characteristics, i.e., the relative infiltration capacity (Berger and Entekhabi, 2001), the relative soil water storage and the average ground surface slope. Similarly, Xu et al. (2013) proposed an equation for ω based on data for 224 MOPEX (Model Parameter Estimation Experiment) watersheds.

In order to identify the controlling factors contributing to differences in the Budyko curves for the four SWs, we used empirical equations proposed by Yang et al. (2007), i.e. Eqs. (11) and (12) and Xu et al. (2013), i.e. Eq. (13), to estimate the ω values in Eq. (9). Three dimensionless variables were evaluated by *Yang et al.* (2007) as the key descriptors of a catchment to estimate the parameter ω, i.e., the relative infiltration capacity (Berger and Entekhabi, 2001), the relative soil water storage and the average ground surface slope. The relative infiltration capacity used in the Eqs. (11) and (12) in Yang et al. (2007) is defined as the ratio of saturated hydraulic conductivity K_s (mm hour^{-1}) to the mean precipitation intensity $\bar{i}_r$ (mm hour^{-1}) and represents infiltration excess. The mean precipitation intensity $\bar{i}_r$ is averaged over the rainy hours of the simulation period. To represent the effect of vegetation and soils on the annual water balance, the plant extractable water capacity, S_{max} (Dunne and Willmott, 1996) is employed and scaled by the mean annual potential evapotranspiration ($\overline{E_p}$) in a dimensionless form, i.e., the relative soil water storage ($S_{\max} / \overline{E_p}$). S_{max} is calculated as:

$$S_{\max} = (\theta_f - \theta_w) \times d_{root} \tag{10}$$

where θ_f and θ_w are the soil moisture contents at field capacity and wilting point respectively; $d_{root} = \min(d_{Top}, d_{rmax})$, where d_{Top} is the top soil depth and d_{rmax} is the maximum root depth of each vegetation type. d_{root} is also the most direct parameter representing the vegetation type affecting the ω value. Another dimensionless parameter is the average ground surface slope ($\tan\beta$). They used an empirical non-linear functional form as Eqs. (11) and (12), considering the correlations of the parameter ω with the three descriptors. A stepwise regression analysis of the data from 108 non-humid catchments in China was used to estimate the functional form and was generalized as Eq. (12).

$$\omega = 1 + f_1\left(\frac{K_s}{\bar{i}_r}\right) f_2\left(\frac{S_{\max}}{\bar{E}_0}\right) f_3(\tan\beta) \tag{11}$$

$$\omega = 1 + 8.652\left(\frac{K_s}{\bar{i}_r}\right)^{-0.368}\left(\frac{S_{\max}}{\bar{E}_0}\right)^{0.436}\exp(-4.464\tan\beta) \tag{12}$$

Xu et al. (2013) proposed the following equation for ω for basins with area 100 - 10,000 km^2:

$$\left.\begin{aligned} \omega &= 5.05722 - 0.09322(lat) + 0.13085(CTI) \\ &+ 1.31697(NDVI) + 0.00003(A) - 0.00018(elev) \end{aligned}\right\} \tag{13}$$

Where *lat* is the basin center latitude, $CTI = ln\left[A_s / tan\beta\right]$ is the compound topographic index, also called the topographic wetness index (Sørensen et al., 2006; Gessler et al., 1993), A_s is the specific catchment area (m^2) per unit width orthogonal to the flow direction and β is the slope angle in radians. *NDVI* is the normalized difference vegetation index, *A* is the catchment area (km^2) and *elev* is the elevation (m).

We used Eqs. (12) and (13) to estimate the ω values for the four SWs and the whole KRW. In addition, we used our simulated results to calculate the fitted values of ω following the Budyko framework in Eq. (9). The sets of ω values are then compared to assess the physical relevance of parameters to identify any deficiencies in each methodology. The values of θ_f and θ_w are directly obtained from the SSURGO database; for K_s we used the calibrated data sets. d_{Top} values are obtained from a 5 min resolution data set (Food and Agricultural Organization, 2003), following the same approach of Yang et al., (2007); d_{rmax} values are obtained for each vegetation as described in Zeng (2001) and weighted with the percentage of the corresponding vegetation. *NDVI* values are obtained from the NASA Earthdata website (https://earthdata.nasa.gov/).

3. Results and Discussion

In this section, we present the results for major hydrologic fluxes following the same order we used for the calibration procedures. Inter-annual water balances of a watershed can be described using the equation:

$$\Delta S = P - Q - E \tag{14}$$

where P is precipitation, Q is runoff and E is evapotranspiration. ΔS denotes the change of storage over the simulation period (i.e., the difference between the amount of water storage over the simulation period). All variables are annual average fluxes (mm yr^{-1}). The calibrated parameters are tabulated in Table 1. The parameters here are the multipliers or the additive or multiplicative constants used to change the initial model parameters. The real parameter values after the calibrations are tabulated in Table 2. For spatially heterogeneous parameters, we list the minimum, the maximum, the mean and the median values of the optimized parameters for each (sub-) watershed. For spatially homogeneous parameters, we simply list the values. All water balance components are expressed in mm yr^{-1} while conductivity values are expressed in mm hr^{-1}. The riverbed conductivity (K_r) values are low compared to the aquifer hydraulic conductivity values. This observation and the ranges of K_r values are supported by data based on geophysical surveys and temperature modeling reported for sites along the Great Miami River in Ohio (Wojnar et al., 2013).

3.1 The spatial and temporal ET results

Figures 4 (a) and (b) show the annual-average spatial maps of ET based on the simulations and MODIS16 data respectively. The spatial maps of ET from simulations and MODIS data generally follow a similar pattern. The results of the linear correlation analysis for the spatially distributed ET simulated values against the LULC types and soil types are summarized in section S1 and Table S2. The PAWS model outputs resolved the ET heterogeneity better than did the remotely sensed MODIS data. The major land cover in northwestern KRW is forest and there are many lakes and reservoirs located in the middle of the watershed. Therefore, we expect high ET values within this area as shown in simulated ET maps. The south-central areas of Kalamazoo (where the MODIS data are blank) are urban areas, which correspond to the low ET values in PAWS output (colored blue). Details related to the spatial variability of ET are further studied in the analysis based on Budyko framework in a later section.

Annual average ET of the 7-year simulation period is 583.43 mm yr^{-1}, which is comparable to the MODIS value, 559.89 mm yr^{-1}. To further evaluate simulated ET, the monthly ET time series (averaged over the entire watershed) is compared with MODIS16 data in Figure 5. The simulated monthly ET is similar compared with MODIS 16 data. The most obvious deviations are during winter months when the model underestimates, whereas in most summer months the model overestimates MODIS ET data. The mismatch between MODIS and simulation is probably due to the different algorithms used. PAWS (via CLM) uses a resistance approach to describe ET based on the two-big leaf model (Dai et al., 2004), while the MODIS product is based on the Penman-Monteith formulation (Mu et al., 2011).

3.2 Stream flow comparisons

Figure 6 shows the 7-year stream flow comparisons between simulations and observations from 6 different USGS gauging stations in the KRW. The NASH values range from 0.57 to 0.87, as tabulated in Table S3, which showed fairly well performance. At Gauging station 04105000, the optimized river bed conductivity values for all (sub-)watersheds are not significantly different compared with the first set of values used in (*Shen et al.* 2016), which has a log mean of 0.12 m/day in another watershed in Michigan. While *Hoaglund et al.* (2002) assigned a uniform value of 0.086 m/day to the riverbed conductivity for all rivers in a regional groundwater modeling study of Michigan. At the outlet gauging station 04108660 of the whole KRW, outputs show almost similar performance as the heaviest weight of the optimization strategy is laid on the outlet gauging station. Thus, the multi-site strategy helped in quantifying the overall stream flow values by preventing simulation compensation errors from other processes.

3.3 Steady state groundwater head comparisons

The plots of simulated versus observed depths to groundwater table from the Michigan Wellogic database (State of Michigan, 2016) for each computational grid cell are shown in Figure 7. The overall NASH values, as tabulated in Table 3, are over 0.91 showing a spatially good match. These values are comparable to or better than the reported NASH values for water table comparisons (e.g., Niu et al., 2014, Shen et al., 2013). The simulated annual groundwater recharge values are within the range of 176 - 351 mm yr^{-1} which was estimated with a tritium interface method by *Delcore and Larson* (1987) in the same watershed.

3.4 Soil Moisture and Soil Temperature Comparisons

Figures 8 and 9 show the 10 cm soil moisture and soil temperature comparisons at two MAWN stations. It should be noted that the observed data represent a point measurement as data were collected using a Campbell Scientific CS616 water content reflectometer (WCR) whereas our simulated results represent an average of a grid cell domain with area of 400×400 m^2. At station Albion (Figure 8 (a)), simulated soil moistures show almost the same trend comparing with observations but generally lower in winter and higher in summer. For example, around February 2004, the simulated soil moisture values are below 0.1 while the observed soil moisture values are between 0.2 and 0.25. Around July 2005, the simulated soil moistures are above 0.1 while the observations are slightly lower. At Michigan State University Kellogg Biological station (MSUKBS) (Figure 8 (b)), the relatively higher soil moistures could not be simulated accurately in February 2009, which may due to the underestimated rainfall intensity during this period. For the soil temperature simulations (Figure 9), the simulated results generally show a good performance at both stations compared with observations except that more fluctuations in simulated results are noted when temperature is below 0 ^oC during winter. The static temperatures around 0 ^oC measured by the WCRs during winter time are possibly because of the frozen soil. It was found by several researchers that the responses of WCRs such as the CS616 are sensitive to temperature and soil type (Benson et al., 2006; Saravanathiiban, 2014). Considering these differences, the comparisons of soil moisture and soil temperature are considered acceptable for vadose zone simulations.

3.5 Analysis based on the Budyko framework

Annual water budgets based on Eq. (14) for the whole KRW and four SWs are listed in Table 4. The plots of annual E/P versus E_p/P (termed Budyko pairs) in the Budyko framework for the four SWs and the whole KRW are shown in Figure 10 for the 7- year simulation period (2003 - 2009). All the Budyko curves in the section refer to the Budyko-type curves are generated by Eq. (9) using simulated results. For the whole KRW, the annual ET value is calculated using three methods (Chen et al., 2013; Condon and Maxwell, 2017): 1) direct ET, simulated ET values are used; 2) inferred ET, assuming the annual storage change is negligible, ET is computed as: $ET = P – Q$ in Eq. (9) and 3) effective precipitation, P is replaces by $P – \Delta S$ in Eq. (9). The alfalfa (mature, 40cm canopy height) reference ET values are calculated as the E_p values, using the Penman - Monteith equation (Dingman, 2008). In this figure, the horizontal straight line indicates the arid or water-limited conditions, while the 1 to 1 line indicates the humid or energy-limited upper bound. The fitted ω values and the R^2 values for the curve fitting using Eq. (9) are tabulated in Table 5. While the ω values using the three different methods are slightly changed due to the partitions of water storage (Istanbulluoglu et al. 2012; Wang, 2012), the ranks of ω values among the four SWs and whole KRW are the same. Therefore, the rest part of Budyko curve analyses are based on results of effective precipitation to simplify our discussion.

The average Mahalanobis distance D (*Mahalanobis*, 1936) of each SW was calculated using the whole KRW results as a reference sample (see Table 6) to represent the dissimilarity of each SW from the averaged pattern. The Mahalanobis distance is a measure of distance in a multidimensional parameter space, which is similar to the Euclidean distance but takes into account the covariance among dimensions of the reference sample. Large D indicates more dissimilarity. SW4 shows the most obvious dissimilarities from the KRW, and this is also apparent from Figure 10 - SW4 has consistently higher E/P values compared to other sub-basins, given similar E_p/P.

The ω values of the four SWs and the whole KRW calculated by Eqs. (12) and (13) and the parameters involved are tabulated in Table 6. All the parameters are calculated within each grid cell and are averaged across the watershed. The ω values calculated by Yang's method for the 4 SWs and whole KRW are generally lower compared with the fitted ω values in Budyko's framework for the 7-year simulation period, with an average deviation around 8.5%. Five major factors could be identified as the reasons for the uncertainties of calculating ω using *Yang*'s method here. First, there are errors in calculating the parameters (for example, there are uncertainties in calculating the average ground surface slope). Second, although *Yang*'s method has considered relative infiltration capacity and relative soil water storage, it overlooked the effects of groundwater flow which also play an important role in water storage and stream discharge contribution (Wang, 2012; Shen et al., 2013; Condon and Maxwell, 2017). Third, the climate conditions of 108 catchments used to generate the empirical equation of *Yang*'s method are different from the climate of KRM; there are uncertainties in the coefficients considering the climate variations. Fourth, there may be parameter inversion errors during the parameter generation processes of Yang's and Xu's methods. Fifth, open water ET plays an important role in SW4, however, this factor is not included in the three indicators suggested by *Yang*'s method. This is also an important reason ω is underestimated using *Yang*'s method for SW4. *Xu*'s method, however, overestimate the ω values with an average deviation around 10.5%, especially for the whole KRW. *Xu*'s method utilizes the catchment area as an indicator which is positively correlated with ω. This may create a discrepancy to differentiate the ω values between whole watershed and SWs. Since the area for the whole watershed is larger than each SW, which indicates larger ω value for the whole watershed, whereas its ω values should fall between the ranges of those of all SWs based on water budgets. In addition, the ω values obtained using *Xu*'s method for the four SWs are not as different as ω values obtained from using *Yang*'s method and the simulated data. This could be due to the climate conditions, i.e. the rainfall intensities, are not explicitly expressed in *Xu*'s method. Besides the possible uncertainties in estimating the values of ω in *Yang*'s and *Xu*'s methods, any errors in the model outputs of the ET and the estimation errors of E_p could also shift the Budyko pairs to some extent compared with the two empirical methods.

The ω values of the four sub-watersheds do not change much and generally fall within the range of 2 ~ 3, and one major reason is that the vegetation types and percentages do not vary significantly among the four SWs. This information is represented as the indicators of root depth in *Yang*'s method and *NDVI* in *Xu*'s method. Although the ω values produced by the three methods (two empirical equations and numerical simulation based on the PAWS model) are different, SW4 has the highest value of ω among all four sub-watersheds., and the most significant influencing factor is the mean precipitation intensity in *Yang*'s method and the watershed elevation in *Xu*'s method. Compared to the other SWs, SW4 had a relative higher precipitation intensity, which decreases the relative infiltration capacity accordingly. Relatively less infiltration capacity translates to more surface ponding, which produces higher actual ET. Some previous studies (e.g., Schenk and Jackson, 2002) suggested that rooting depth increases with precipitation (at least in water-limited ecosystems). Results shown in Table 6 suggest that SW 4 has the highest precipitation and root depth values, although the differences of root depth values are not significant. The lake effect may be responsible for the higher precipitation intensity at SW4, which is also implicitly reflected in the elevation indicator of *Xu*'s method. In contrast, SW3 shows the lowest ω value, and higher averaged soil saturated hydraulic conductivity increases the relative infiltration capacity, which tends to generate lower actual ET. In addition, SW3 has higher ground surface slope and accordingly lower CTI, which indicates higher potential to produce surface runoff with less water retained for evapotranspiration (Hjerdt et al., 2004, Yang et al., 2019). Another factor influencing the ET of SW3 could be its largest urban area ratio among all the SWs; this information was incorporated in the data integration algorithm during the model construction (Shen et al., 2013). These calculated ω values are also in correspondence to D values discussed above when considering the deviations of ω values of the four SWs from that of the whole KRW. Considering all the possible errors in estimating the values of ω, the dominant heterogeneous characteristics controlling the water budgets are effectively identified within the indicators of the two empirical methods at SW scale, which indicates the usefulness of the two empirical methods in predicting the interannual variability of regional water balances over a long period.

The advantage of the Budyko approach lies in its ability to predict changes in long term ET or water yields due to changes in vegetation (e.g. replacing traditional crops with biofuel crops) based on multiple observations. The Budyko approach could efficiently generate the general functional patterns of catchments and inform controlling hydrologic parameters based on empirical relationships. These results reveal the possibility of using Budyko approach to guide the calibration of PBHM models, to recognize the controlling processes, and to constrain individual processes in the integrated system.

4. Conclusions

In order to accurately quantify the spatial and temporal inter-annual water balances for a heterogeneous catchment, in this work, we used a step-wised calibration method combined with a multisite calibration strategy to optimize a PBHM. The calibration objectives are not limited to stream discharge exclusively, while also include ET and groundwater heads to resolve the equifinality issue. This calibration strategy successfully converged, and the calibrated results showed good comparisons with the observed data for different SWs. The Budyko curves based on the simulated water balance components and two empirical equations (Yang et al. (2007) and Xu et al, (2013)) are employed to quantify the variabilities of inter-annual water balances for different SWs. Although Yang's method (Yang et al., 2007) is more suitable for the KRW, the dimensionless landscape characteristics used in both empirical relations are found to be useful in characterizing the integrated hydrologic performances based on the Budyko framework.

Given its simplicity, the Budyko approach could efficiently generate the general functional patterns of hydrologic system at SW scale. The consistency of presenting the spatial variability of water budgets between PBHM and Budyko approach reveal the possibility of synthesizing Darwinian and Newtonian approaches, to deepen understanding of the hydrologic system (Harman and Troch, 2014).

References

Anderton, S., Latron, J., Gallart, F., 2002. Sensitivity analysis and multi-response, multi-criteria evaluation of a physically based distributed model. Hydrol. Process. 16, 333–353. doi:10.1002/hyp.336

Benson, C. H., and Wang, X, 2006. Temperature-Compensating Calibration Procedure for Water Content Reflectometers. Proc. TDR 2006, Purdue University, West Lafayette, USA, Sept. 2006, Paper ID 50, 16 p., https://engineering.purdue.edu/TDR/Papers

Berger, K.P., Entekhabi, D., 2001. Basin hydrologic response relations to distributed physiographic descriptors and climate. J. Hydrol. 247, 169–182. doi:10.1016/S0022-1694(01)00383-3

Beven, K.J., Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298. doi:10.1002/hyp.3360060305

Beven, K.J., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources, Research Perspectives in Hydrology 16, 41–51. doi:10.1016/0309-1708(93)90028-E

Beven, K.J., 2001. Uniqueness of place and process representations in hydrological modelling. Hydrol. Earth Syst. Sci. 4, 203–213. doi:10.5194/hess-4-203-2000

Beven, K.J., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249, 11–29. doi:10.1016/S0022-1694(01)00421-8

Beven, K.J., 2002. Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. Hydrol. Process. 16, 189–206. doi:10.1002/hyp.343

Budyko, 1974. Climate and life, English edition. Academic Press, New York.

Chen, X., Alimohammadi, N., Wang, D., 2013. Modeling interannual variability of seasonal evaporation and storage change based on the extended Budyko framework. Water Resour. Res. 49, 6067–6078. doi:10.1002/wrcr.20493

Choudhury, B., 1999. Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model. J. Hydrol. 216, 99–110. doi:10.1016/S0022-1694(98)00293-5

Condon, L.E., Maxwell, R.M., 2017. Systematic shifts in Budyko relationships caused by groundwater storage changes. Hydrol. Earth Syst. Sci. 21, 1117–1135. doi:10.5194/hess-21-1117-2017

Dai, Y., Dickinson, R.E., Wang, Y. P., 2004. A Two-Big-Leaf Model for Canopy Temperature, Photosynthesis, and Stomatal Conductance. J. Climate 17, 2281–2299. doi:10.1175/1520-0442(2004)017<2281:ATMFCT>2.0.CO;2

Delcore, M.R., Larson, G.J., 1987. Application of the tritium interface method for determining recharge rates to unconfined drift aquifers, II. Non-homogeneous case. J. Hydrol. 91, 73–81. doi:10.1016/0022-1694(87)90129-6

Dingman, S.L., 2008. Physical Hydrology. Waveland Press, Long Grove, IL., 3rd edition.

Donohue, R.J., Roderick, M.L., McVicar, T.R., 2012. Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model. J. Hydrol. 436 - 437, 35 -50. doi: 10.1016/j.jhydrol.2012.02.033.

Dunne, K.A., Willmott, C.J., 1996. Global Distribution of Plant-Extractable Water Capacity of Soil. International Journal of Climatology 16, 841–859. doi:10.1002/(SICI)1097-0088(199608)16:83.3.CO;2-8

Enviro-weather, 2016. Weather-based pest, natural resources and production management tools, Michigan State University Agbio Research. URL http://www.enviroweather.msu.edu/.

Fang, K., Shen, C., Fisher, J.B., Niu, J., 2016. Improving Budyko curve-based estimates of long-term water partitioning using hydrologic signatures from GRACE. Water Resour. Res. 52, 5537–5554, doi:10.1002/2016WR018748.

Food and Agricultural Organization, 2003. Map of world soil resources, Rome.

Fry, L.M., Hunter, T.S., Phanikumar, M.S., Fortin, V., Gronewold, A.D., 2013. Identifying streamgage networks for maximizing the e_ectiveness of regional water balance modeling. Water Resources Research 49, 2689-2700. doi:10.1002/wrcr.20233.

Fu, B. P., 1981. On the calculation of the evaporation from land surface. Chinese Journal of Atmospheric Sciences (in Chinese), 5(1), 23-31. doi:10.3878/j.issn.1006-9895.1981.01.03.

Gessler, P.E., Moore, I.D., McKenzie, N.J., Ryan, P.J., 1993. Soil-landscape modelling and spatial prediction of soil attributes. International Journal of Geographical Information Systems 9(4), 421-432. doi:10.1080/02693799508902047.

Gunduz, O., Aral, M.M., 2005. River networks and groundwater flow: a simultaneous solution of a coupled system. J. Hydrol. 301, 216–234. doi:10.1016/j.jhydrol.2004.06.034

Hjerdt K. N., McDonnell J. J., Seibert J. and Rodhe A., 2004. A new topographic index to quantify downslope controls on local drainage. Water Resources Research, 40(5), doi:10.1029/2004WR003130.

Hoaglund, J.R., Huffman, G.C., Grannemann, N.G., 2002. Michigan basin regional ground water flow discharge to three Great Lakes. Ground Water, 40(4), 390–405.

Huss, M., Bauder, A., Funk, M., Hock, R., 2008. Determination of the seasonal mass balance of four Alpine glaciers since 1865. J. Geophys. Res. 113, F01015. doi:10.1029/2007JF000803

Huss, M., Bauder, A., Funk, M., Hock, R., 2008. Determination of the seasonal mass balance of four alpine glaciers since 1865. J. Geophys. Res. 113(F01015.). doi:10.1029/2007JF000803.

Harman, C., Troch, P.A, 2013. Darwinian hydrology: can the methodology Charles Darwin pioneered help hydrologic science? Hydrology and Earth System Sciences Discussions. 10, 6407–6444, doi:10.5194/hessd-10-6407-2013.

Istanbulluoglu, E., Wang, T., Wright, O.M., Lenters, J.D., 2012. Interpretation of hydrologic trends from a water balance perspective: The role of groundwater storage in the Budyko hypothesis. Water Resour. Res. 48, W00H16. doi:10.1029/2010WR010100

Kalamazoo River Watershed Council., 2011. Kalamazoo River Watershed Management Plan. Prepared for the Michigan Nonpoint Source Program (Michigan Department of Environmental Quality and the United States Environmental Protection Agency).

Lai, C., Katul, G., 2000. The dynamic role of root-water uptake in coupling potential to actual transpiration. Adv. Water Resour 23(4), 427-439.

Liang, W., Bai, D., Wang, F., Fu, B., Yan, J., Wang, S., Yang, Y., Long, D., Feng, M., 2015. Quantifying the impacts of climate change and ecological restoration on streamflow changes based on a Budyko hydrological model in China's Loess Plateau. Water Resour. Res. 51, 6500–6519. doi:10.1002/2014WR016589

Mahalanobis, P.C., 1936. On the Generalised Distance in Statistics. On the Generalized Distance in Statistics 49–55.

McDonnell, J.J., Beven, K., 2014. Debates-The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. Water Resources Research 50(6), 5342-5350. doi:10.1002/2013WR015141.

MDTMB: Michigan Department of Technology, Management & Budget, 2016. 2001 IFMAP/GAP Lower peninsula land cover. URL https://www.mcgi.state.mi.us/mgdl/?rel=thext&action=thmname&cid=5&cat=Land+Cover+2001

Milly, P.C.D., 1994. Climate, interseasonal storage of soil water, and the annual water balance. Advances in Water Resources, MIT Colloquium on Hydroclimatology and Global Hydrology 17, 19–24. doi:10.1016/0309-1708(94)90020-5

Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sensing of Environment 115, 1781–1800. doi:10.1016/j.rse.2011.02.019

National Climatic Data Center (NCDC), 2010. Available at <http://www.ncdc.noaa.gov/oa/climate/climatedata.html#daily>.

Niu, G. Y., Yang, Z. L., 2006. Effects of Frozen Soil on Snowmelt Runoff and Soil Water Storage at a Continental Scale. J. Hydrometeor 7, 937–952. doi:10.1175/JHM538.1

Niu, J., Phanikumar, M.S., 2015. Modeling watershed-scale solute transport using an integrated, process-based hydrologic model with applications to bacterial fate and transport. J. Hydrol. 529, Part 1, 35–48. doi:10.1016/j.jhydrol.2015.07.013

Niu, J., Shen, C., Li, S. G., Phanikumar, M.S., 2014. Quantifying storage changes in regional Great Lakes watersheds using a coupled subsurface-land surface process model and

GRACE, MODIS products. Water Resour. Res. 50, 7359–7377. doi:10.1002/2014WR015589

Oleson, K.W., Lawrence, D.M., Gordon, B., Flanner, M.G., Kluzek, E., Peter, J., Levis, S., Swenson, S.C., Thornton, E., Feddema, J., others, 2010. Technical description of version 4.0 of the Community Land Model (CLM) (No. NCAR/TN-478+STR), NCAR Technical Note. National Center for Atmospheric Research, Boulder, Colorado.

Oudin, L., Andréassian, V., Lerat, J., Michel, C., 2008. Has land cover a significant impact on mean annual streamflow? An international assessment using 1508 catchments. J. Hydrol. 357, 303–316. doi:10.1016/j.jhydrol.2008.05.021

Price, K., Stone, R., and Lampinen, J. (2005), Differential Evolution: A Practical Approach to Global Optimization, Springer, Berlin.

Qiu, H., P. Blaen, S. Comer-Warner, D.M. Hannah, S. Krause, M.S. Phanikumar, Evaluating a coupled phenology – surface energy balance model to understand stream – subsurface temperature dynamics in a mixed-use farmland catchment, Water Resources Research, vol. 55, doi:10.1029 / 2018WR023644 (2019)

Riley, W.J., Shen, C., 2014. Characterizing coarse-resolution watershed soil moisture heterogeneity using fine-scale simulations. Hydrol. Earth Syst. Sci. 18, 2463–2483. doi:10.5194/hess-18-2463-2014

Ragettli, S., Pellicciotti, F., 2012. Calibration of a physically based, spatially distributed hydrological model in a glacierized basin: On the use of knowledge from glaciometeorological processes to constrain model parameters. Water Resour. Res. 48, W03509. doi:10.1029/2011WR010559

Saravanathiiban, D. S., 2014. Preferential flow through earthen landfill covers: Field evaluation of root zone water quality model and laboratory validation of lattice Boltzmann method, PhD dissertation, Dep. of Civ. Environ. Eng., Mich. State Univ., East Lansing.

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251, 163–176. doi:10.1016/S0022-1694(01)00466-8

Schenk, H., Jackson, R., 2002. Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems. Journal of Ecology 90(3), 480-494. doi:10.1046/j.1365-2745.2002.00682. x.

Shen, C., Niu, J., Phanikumar, M.S., 2013. Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface-land surface processes model. Water Resour. Res. 49, 2552–2572. doi:10.1002/wrcr.20189

Shen, C., Niu, J., Fang, K., 2014. Quantifying the effects of data integration algorithms on the outcomes of a subsurface–land surface processes model. Environmental Modelling & Software 59, 146–161. doi:10.1016/j.envsoft.2014.05.006

Shen, C., Phanikumar, M.S., 2010. A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling. Advances in Water Resources 33, 1524–1541. doi: 10.1016/j.advwatres.2010.09.002

Shen, C., Riley, W.J., Smithgall, K.R., Melack, J.M., Fang, K., 2016. The fan of influence of streams and channel feedbacks to simulated land surface water and carbon dynamics. Water Resour. Res. 52(2), 880 { 902. doi:10.1002/2015WR018086.

Sivapalan, M., 2006. Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. Encyclopedia of hydrological sciences, 193–219, John Wiley, Chichester, U. K., doi:10.1002/0470848944.hsa012.

Sivapalan, M., Yaeger, M.A., Harman, C.J., Xu, X., Troch, P.A., 2011. Functional model of water balance variability at the catchment scale: 1. Evidence of hydrologic similarity and space-time symmetry. Water Resour. Res. 47, W02522. doi:10.1029/2010WR009568

Sørensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: Evaluation of different methods based on field observations. Hydrology and Earth System Sciences 10, 101-112.

Soil Survey Staff, Survey Geographic (SSURGO) Database for Michigan. Natural Resources Conservation Service, United States Department of Agriculture. Available online at <http://soildatamart.nrcs.usda.gov> (accessed 06.01.10).

Stahl, K., Moore, R.D., Shea, J.M., Hutchinson, D., Cannon, A.J., 2008. Coupled modelling of glacier and streamflow response to future climate scenarios. Water Resour. Res. 44, W02422. doi:10.1029/2007WR005956

State of Michigan, 2016. Wellogic digital water well dataset, Michigan GIS open data portal, State of Michigan. URL http://gis.michigan.opendata.arcgis.com/datasets?q=wellogic&sort by=relevance. Accessed on June 21, 2016.

Sutanudjaja, E.H., van Beek, L.P.H., de Jong, S.M., van Geer, F.C., Bierkens, M.F.P., 2014. Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. Water Resour. Res. 50, 687–705. doi:10.1002/2013WR013807

Wagener, T., Sivapalan, M., Troch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Kumar, P., Rao, P.S.C., Basu, N.B., Wilson, J.S., 2010. The future of hydrology: An evolving science for a changing world. Water Resour. Res. 46, W05301. doi:10.1029/2009WR008906

Wang, D., 2012. Evaluating interannual water storage changes at watersheds in Illinois based on long-term soil moisture and groundwater level data. Water Resour. Res. 48, W03502. doi:10.1029/2011WR010759

Wang, D., Wu, L., 2013. Similarity of climate control on base flow and perennial stream density in the Budyko framework. Hydrol. Earth Syst. Sci. 17, 315–324. doi:10.5194/hess-17-315-2013

Wang, T., Istanbulluoglu, J.L., Scott, D., 2009. On the role of groundwater and soil texture in the regional water balance: An investigation of the Nebraska Sand Hills, USA. Water. Resour. Res. 45. doi:10.1029/2009WR007733.

Wesley, J.K., 2005. Kalamazoo River Assessment. Fisheries Division Special Report. Michigan Department of Natural Resources, Michigan, USA. Available online at http://www.michigan.gov/dnr/0,4570,7-153-10364_52259_19056-46270--,00.html

Wojnar, A.J., Mutiti, S., Levy, J., 2013. Assessment of geophysical surveys as a tool to estimate riverbed hydraulic conductivity. Journal of Hydrology 482, 40-56. doi:10.1016/j.jhydrol.2012.12.018.

Xu, X., Liu, W., Scanlon, B.R., Zhang, L., Pan, M., 2013. Local and global factors control water-energy balances within the Budyko framework. Geophys. Res. Lett. 40, 6123 - 6129.

Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., Lei, Z., 2007. Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis. Water Resour. Res. 43, W04426. doi:10.1029/2006WR005224

Yang, D., Yang, A., Qiu, H., Zhou, Y., Herrero, H., Fu, C.-S., Yu, Q. and Tang, J., 2019. A Citizen-Contributed GIS Approach for Evaluating the Impacts of Land Use on Hurricane-Harvey-Induced Flooding in Houston Area, Land, 8(2), 25, doi:10.3390/land8020025.

Yokoo, Y., Sivapalan, M., Oki, T., 2008. Investigating the roles of climate seasonality and landscape characteristics on mean annual and monthly water balances. J. Hydrol. 357, 255–269. doi:10.1016/j.jhydrol.2008.05.010

Zeng, X., 2001. Global Vegetation Root Distribution for Land Modeling. J. Hydrometeor 2, 525–530. doi:10.1175/1525-7541(2001)002<0525:GVRDFL>2.0.CO;2

Zhang, L., Dawes, W.R., Walker, G.R., 2001. Response of mean annual evapotranspiration to vegetation changes at catchment scale. Water Resour. Res. 37, 701–708. doi:10.1029/2000WR900325

Figures

Figure 1. Map of the Kalamazoo River watershed. Elevation is shown as the color gradient. National Hydrography Dataset (NHD) rivers, U.S. Geological Survey (USGS) gauges, National Climatic Data Center (NCDC) weather stations and Michigan Automatic Weather Network (MAWN) stations are shown.

Figure 2. Land Use and Land Cover map for the Kalamazoo River Watershed.

Figure 3. Flow chart of the calibration procedure. See Table S1 for an explanation of variables and their meaning.

Figure 4. Spatial map of yearly averaged evapotranspiration for the Kalamazoo River watershed for the 7-year period (2003–2009) of (a) simulated output and (b) MODIS data.

Figure 5. Monthly ET comparisons of simulated outputs with MODIS data for the 7-year simulation period (2003–2009). NASH is the Nash-Sutcliffe efficiency metric; APB is the absolute bias; RMSE is the root mean squared error.

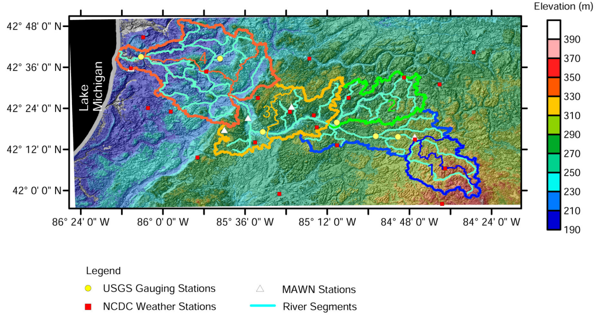
Figure 6. River discharge comparisons of simulated outputs with observations at different U.S. Geological Survey (USGS) gauge stations. Sim is the simulated; Obs is the USGS observations. NASH is the Nash-Sutcliffe efficiency metric. The model performance for each gauge is summarized in Table S3.

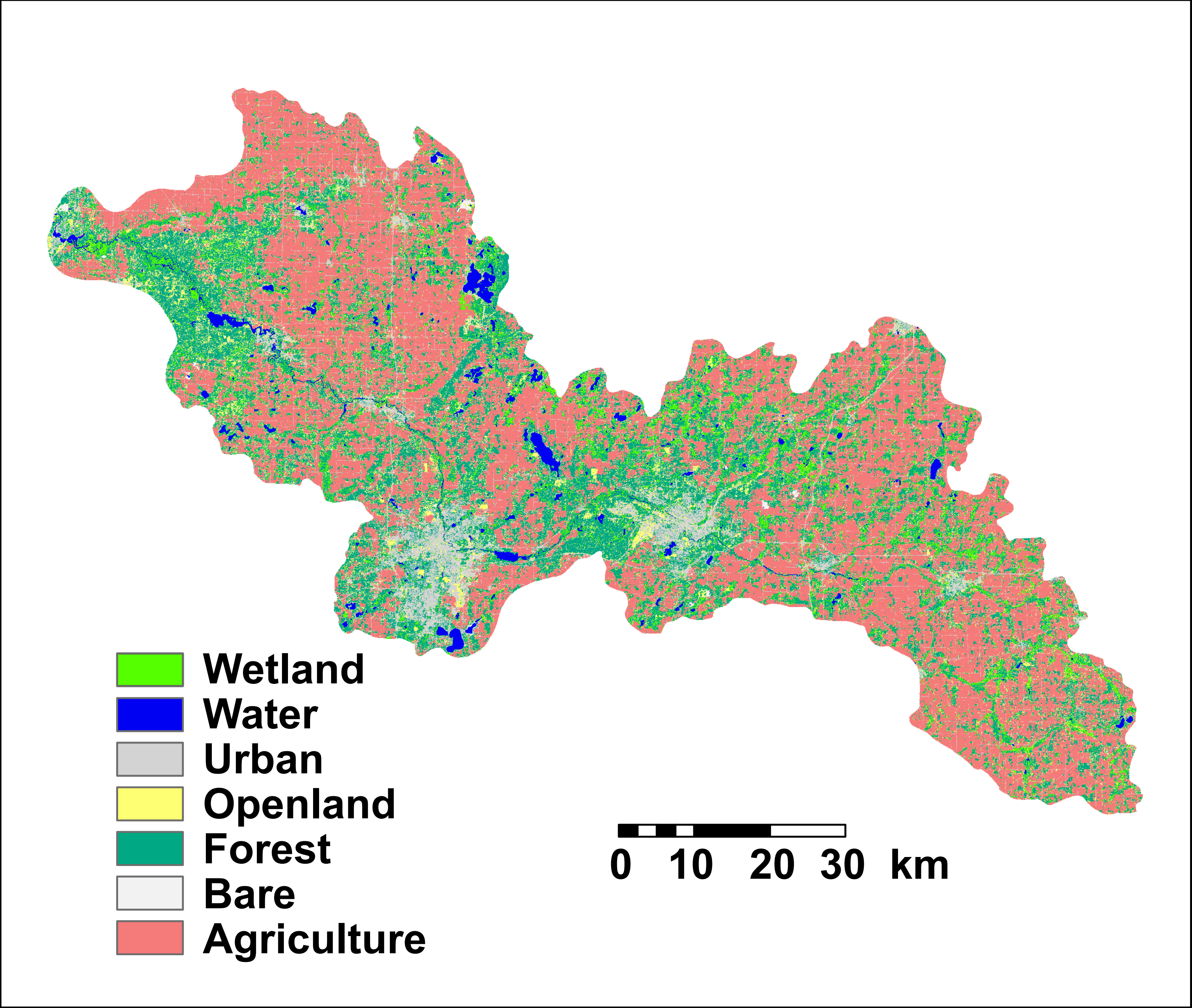
Figure 7. Plots of simulated versus observed depth to groundwater table (from Wellogic data set) for each computation grid cell. SW1, SW2, SW3, SW4 are the simulated results within each SW.

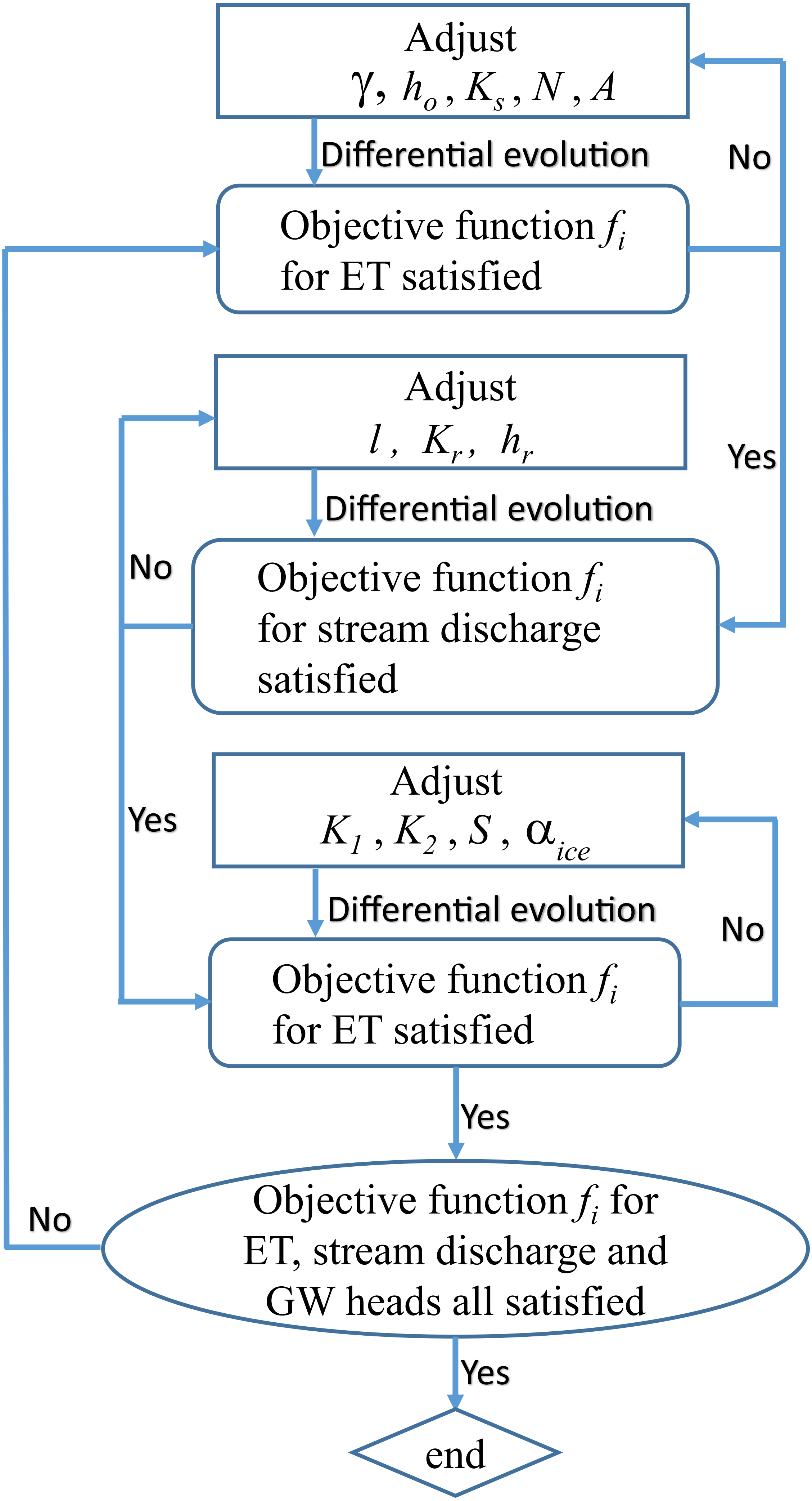
Figure 8. 10 cm Soil Moisture comparisons of simulated outputs with MAWN (Michigan Automatic Weather Network) station observations at (a) Albion and (b) MSUKBS. Sim is the simulated outputs; Obs is the MAWN station observations.

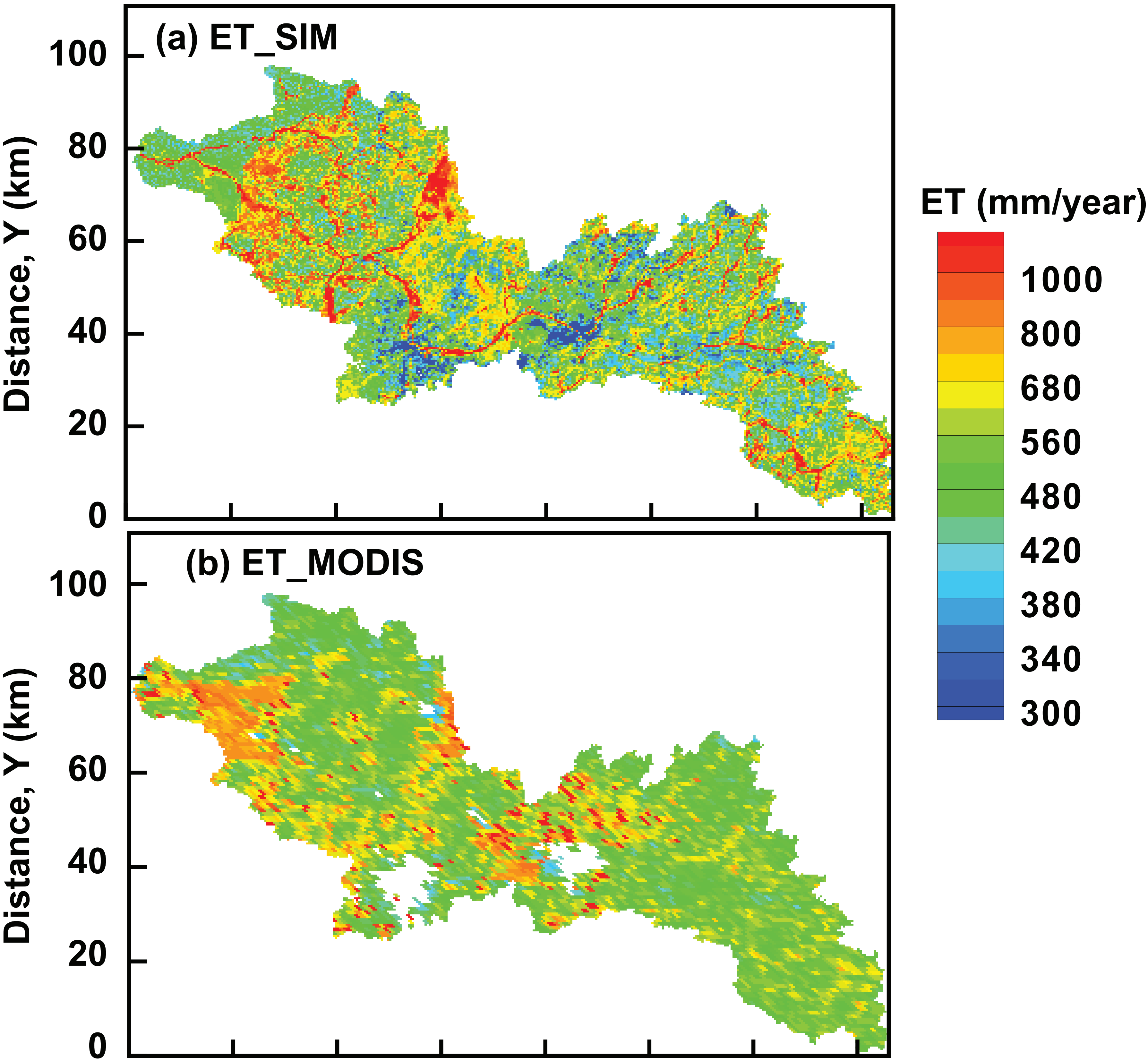
Figure 9. 10 cm Soil Temperature comparisons of GLB and MLT simulated outputs with MAWN (Michigan Automatic Weather Network) station observations at (a) Albion and (b) MSUKBS. Sim is the simulated outputs; Obs is the MAWN station observations.

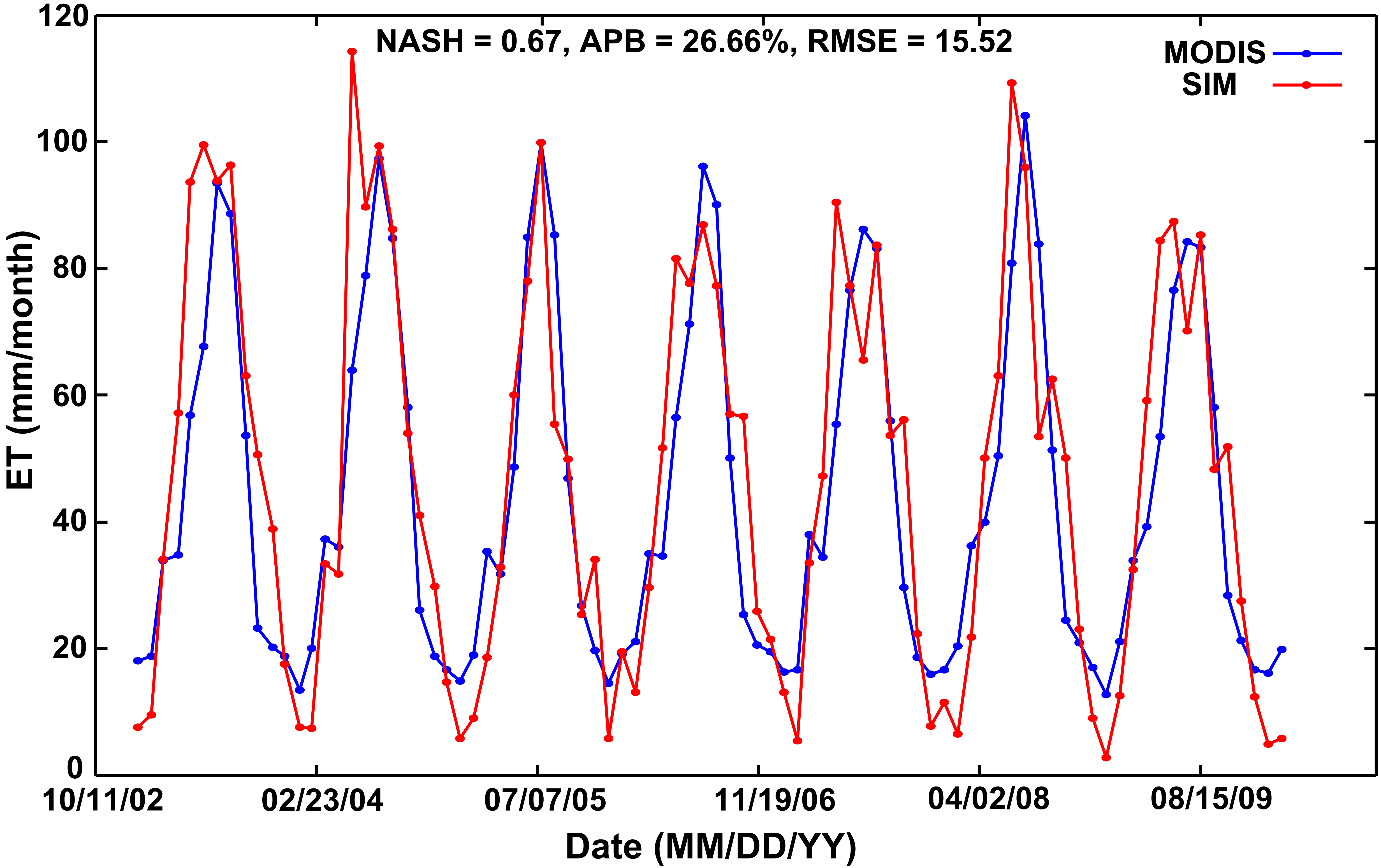
Figure 10. Budyko Curve Analysis for the 4 SWs and the whole Kalamazoo River watershed for a 7-year simulation period from 2003 to 2009 using a) direct ET, b) inferred ET and (c) effective precipitation.

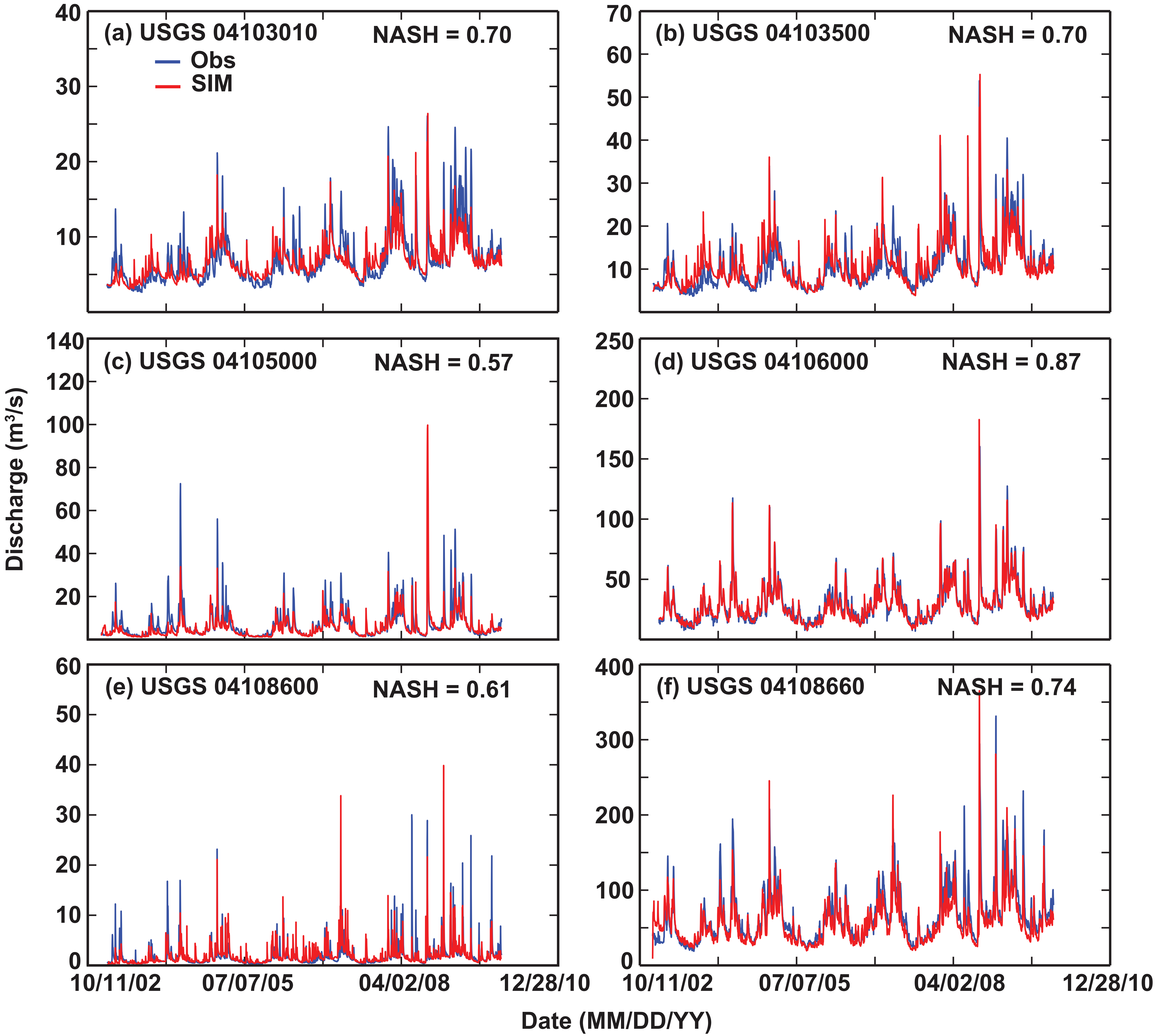

Elevation (m)
390
370
350
330
310
290
270
250
230
210
190
42° 48' 0" N
42° 36' 0" N
42° 24' 0" N
42° 12' 0" N
42° 0' 0" N
86° 24' 0" W
86° 0' 0" W
85° 36' 0" W
85° 12' 0" W
84° 48' 0" W
84° 24' 0" W
Lake Michigan
1
2
3
4
Legend
USGS Gauging Stations
NCDC Weather Stations
MAWN Stations
River Segments

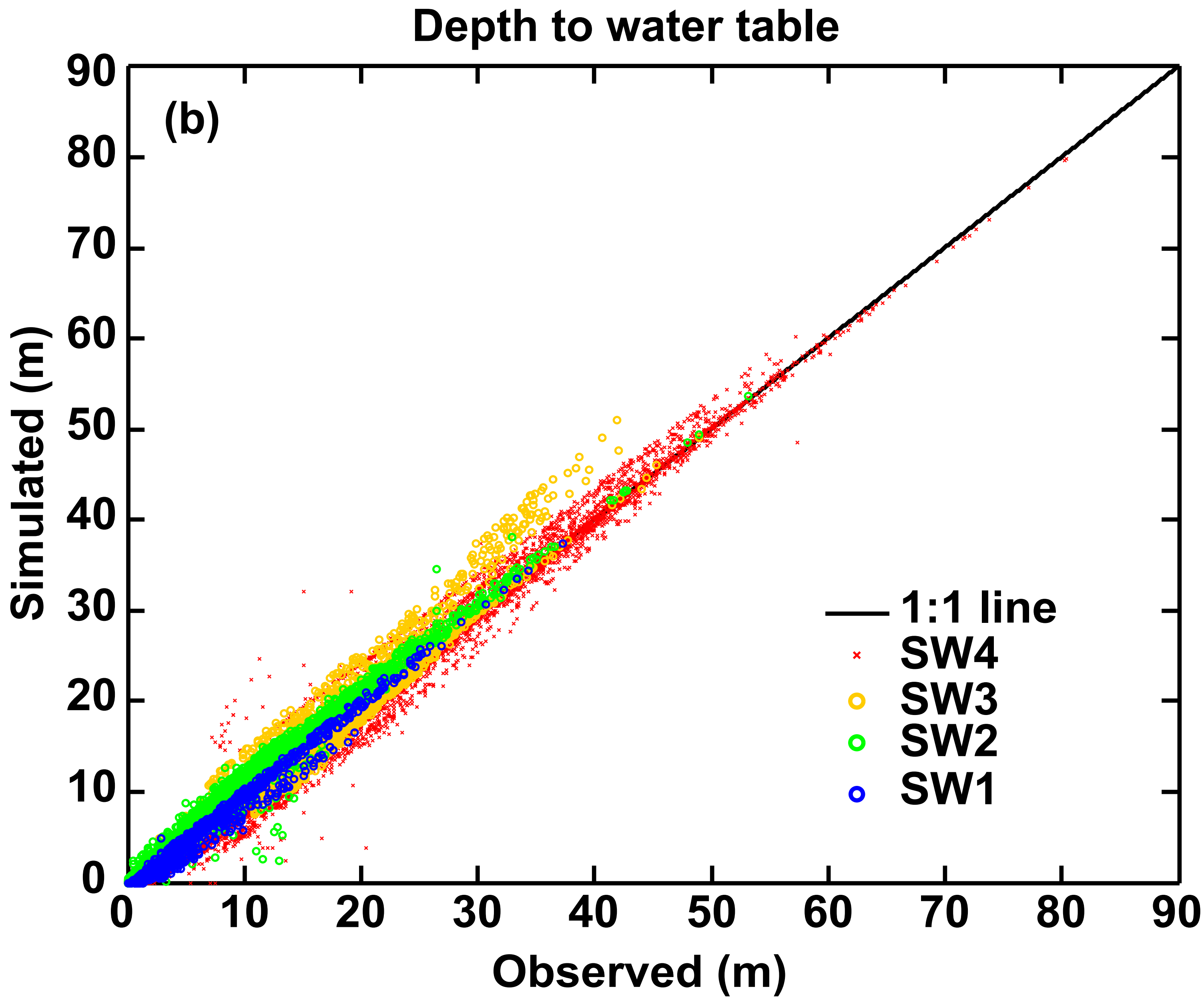

Wetland
Water
Urban
Openland
Forest
Bare
Agriculture
0 10 20 30 km

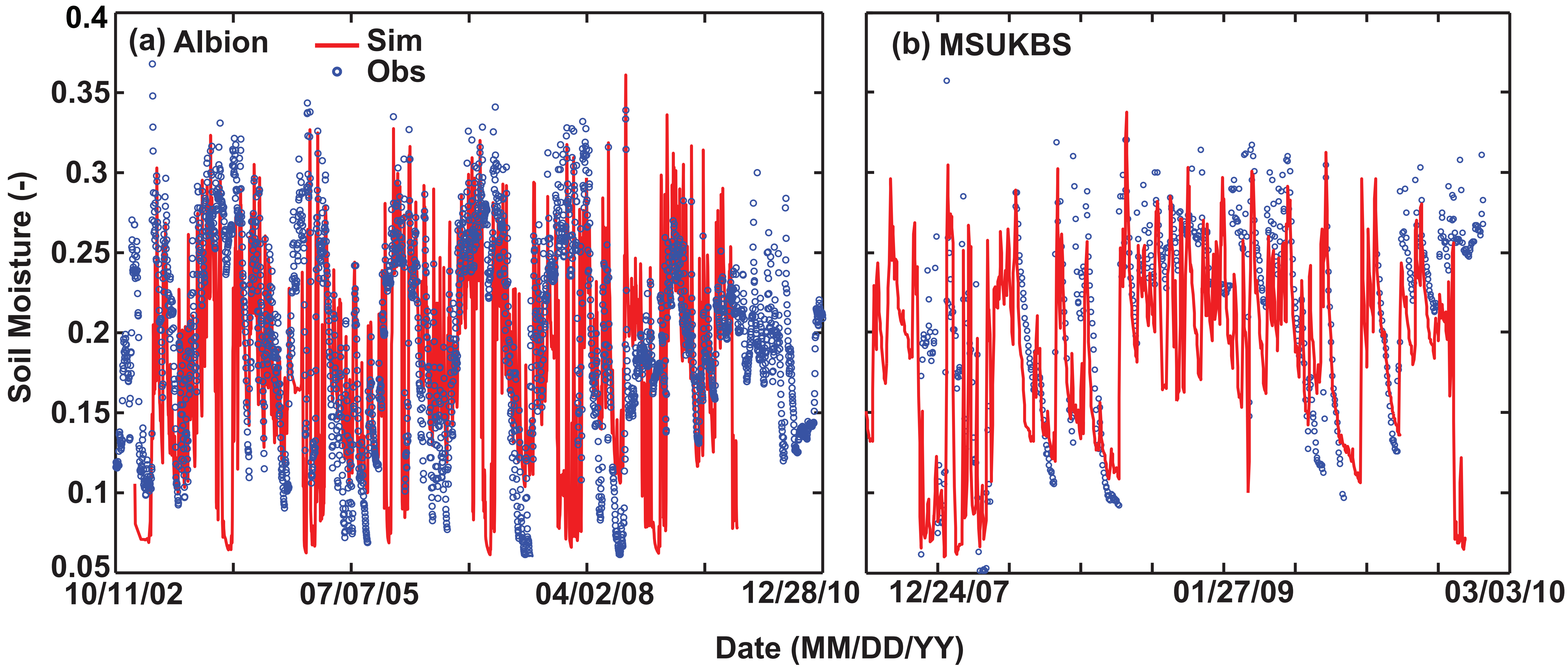

Adjust
γ, h_o, K_s, N, A
Differential evolution
No
Objective function f_i for ET satisfied
Adjust
l, K_r, h_r
Yes
Differential evolution
No
Objective function f_i for stream discharge satisfied
Yes
Adjust
$K_1, K_2, S, \alpha_{ice}$
Differential evolution
No
Objective function f_i for ET satisfied
Yes
No
Objective function f_i for ET, stream discharge and GW heads all satisfied
Yes
end


(a) ET_SIM
(b) ET_MODIS
Distance, Y (km)
0
20
40
60
80
100
ET (mm/year)
1000
800
680
560
480
420
380
340
300

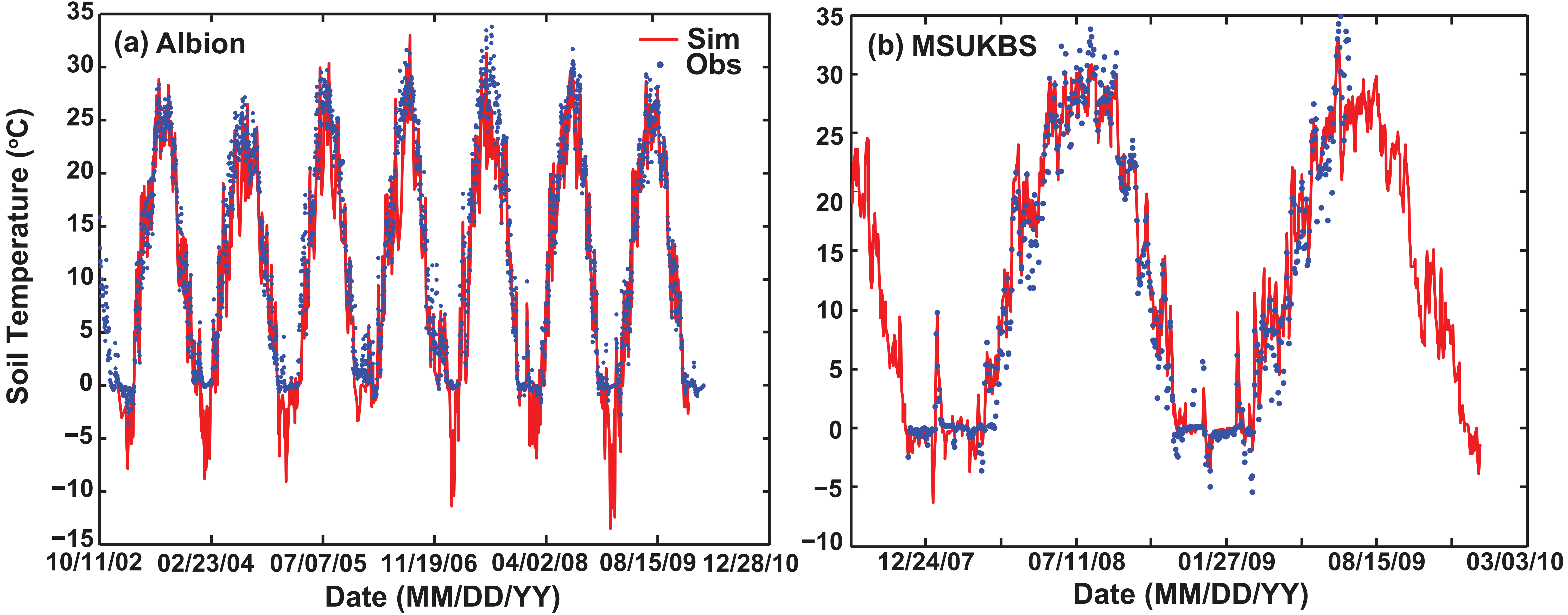

NASH = 0.67, APB = 26.66%, RMSE = 15.52
MODIS
SIM
ET (mm/month)
0
20
40
60
80
100
120
10/11/02
02/23/04
07/07/05
11/19/06
04/02/08
08/15/09
Date (MM/DD/YY)

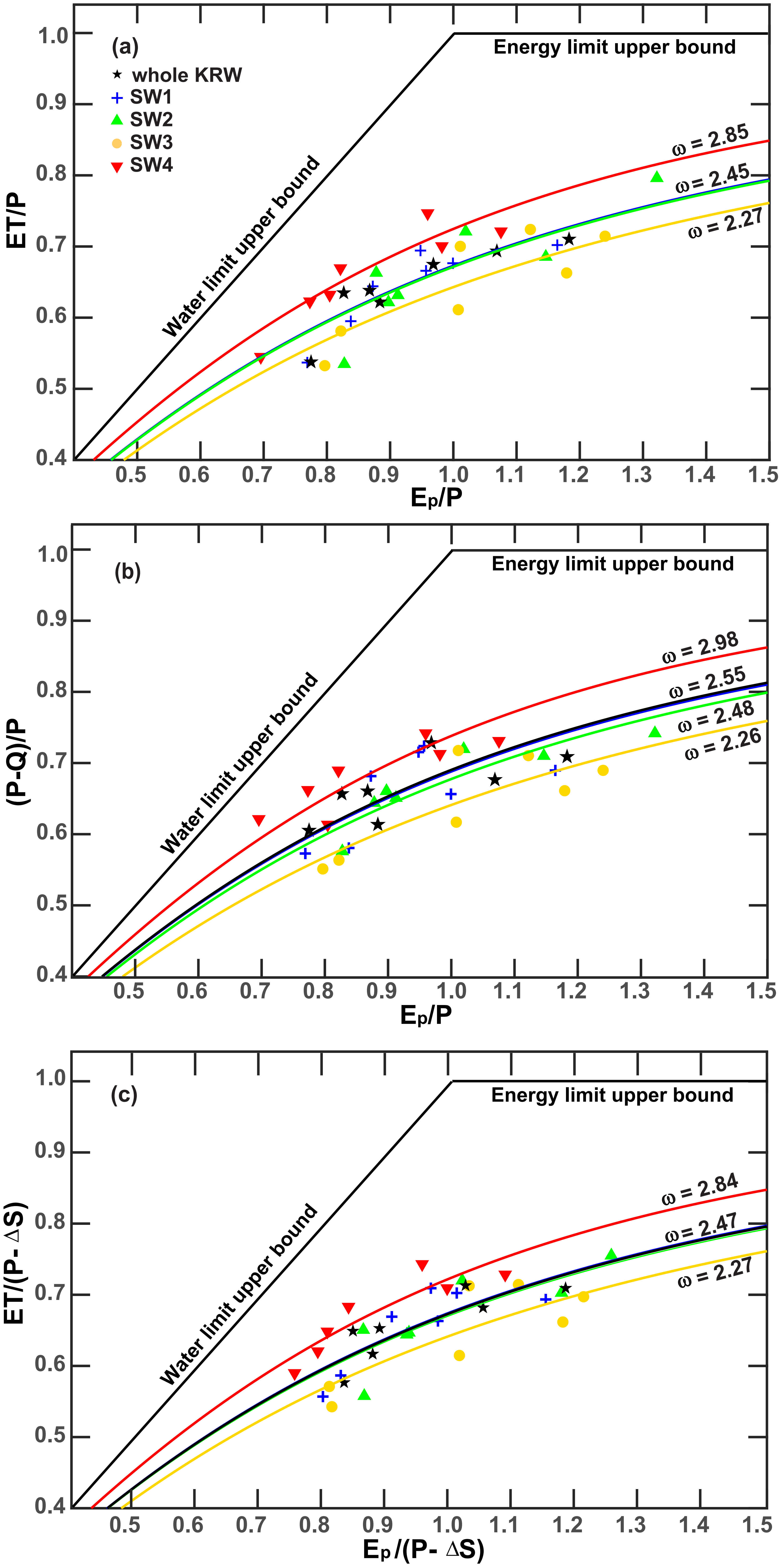

(a) USGS 04103010
NASH = 0.70
Obs
SIM
(b) USGS 04103500
NASH = 0.70
(c) USGS 04105000
NASH = 0.57
(d) USGS 04106000
NASH = 0.87
(e) USGS 04108600
NASH = 0.61
(f) USGS 04108660
NASH = 0.74
Discharge (m³/s)
10/11/02
07/07/05
04/02/08
12/28/10
Date (MM/DD/YY)


Depth to water table
(b)
Simulated (m)
Observed (m)
0
10
20
30
40
50
60
70
80
90
1:1 line
SW4
SW3
SW2
SW1



(a) Albion
Sim
Obs
(b) MSUKBS
Soil Moisture (-)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
10/11/02
07/07/05
04/02/08
12/28/10
12/24/07
01/27/09
03/03/10
Date (MM/DD/YY)


(a) Albion
Sim
Obs
Soil Temperature (°C)
35
30
25
20
15
10
5
0
−5
−10
−15
10/11/02
02/23/04
07/07/05
11/19/06
04/02/08
08/15/09
12/28/10
Date (MM/DD/YY)
(b) MSUKBS
35
30
25
20
15
10
5
0
−5
−10
12/24/07
07/11/08
01/27/09
08/15/09
03/03/10
Date (MM/DD/YY)



(a)
whole KRW
SW1
SW2
SW3
SW4
Energy limit upper bound
Water limit upper bound
ω = 2.85
ω= 2.45
ω = 2.27
ET/P
Ep/P
(b)
Energy limit upper bound
Water limit upper bound
ω = 2.98
ω = 2.55
ω = 2.48
ω = 2.26
(P-Q)/P
Ep/P
(c)
Energy limit upper bound
Water limit upper bound
ω = 2.84
ω= 2.47
ω = 2.27
ET/(P- ΔS)
Ep/(P- ΔS)

Table 1. Calibrated parameter operator values and the optimization types

		optimized operator values			
Symbol(Unit)	optimization type	SW1	SW2	SW3	SW4
γ	×	4.10×10^{-2}	4.30×10^{-2}	3.68×10^{-2}	1.68×10^{-2}
α_{ice}	+	0.45	-0.20	0.64	0.53
K_1	×	1.27	0.78	1.39	0.99
K_2	×	1.07	1.2	2.19	1.86
K_s	×	1.46	0.96	0.81	1.26
N	+	-0.14	-0.26	-0.15	0.08
A (1 m^{-1})	×	0.86	0.88	0.81	1.43
L (m)	+	-35	-6	-54	-44
h_o (m)	+	4.24×10^{-2}	4.45×10^{-2}	4.38×10^{-2}	3.98×10^{-2}
K_r (m day^{-1})	×	1.10×10^{-2}	7.39×10^{-3}	2.15×10^{-2}	1.01×10^{-1}
h_r (m)	+	0.14	0.21	0	0

Table 2. Calibrated parameter values for MLT methods. For spatially heterogeneous parameters we listed the minimum (min), the maximum (max), the mean and the median values within the specific simulated domain.

	SW1	SW2	SW3	SW4_
Parameter (Unit)	min-max(mean, median)	min-max(mean, median)	min-max(mean, median)	min-max(mean, median)
γ	4.1×10^{-2}	4.3×10^{-2}	3.68×10^{-2}	1.68×10^{-2}
α_{ice}	3.45	2.90	3.64	3.53
K_1 (m/day)	0.305-105.07(26.96,34.99)	0.249-62.86(17.41,13.58)	4.660-80.34(24.3,24.84)	0.097-75.05(19.15,18.27)
K_2 (m/day)	0.019-5.05(1.61,1.38)	0.017-6.18(0.947,1.30)	0.018-6.07(1.86,3.74)	0.004-1.96(0.10,0.16)
K_s (mm/hour)	1.25-263.72(73.18,78.83)	0.58-205.84(54.26,24.59)	0.38-275.15(113.53,98.64)	0.69-246.02(67.73,70.49)
N	1.05-2.26(1.30,1.31)	0.99-2.13(1.19,1.16)	1.06-1.93(1.31,1.31)	1.25-2.81(1.51,1.53)
A(1/m)	0.032-6.48(4.06,4.21)	0.038-6.37(3.01,3.22)	0.041-6.22(4.35,4.90)	0.051-6.99(4.68,5.06)
l(m)	365	394	346	356
h_o(m)	4.24×10^{-2}	4.45×10^{-2}	4.38×10^{-2}	3.98×10^{-2}
K_r(m/day)	0.011-0.206(0.109,0.110)	7.1×10^{-3}-0.113(0.059, 0.058)	0.077-0.318(0.174,0.177)	0.014-0.385(0.190,0.190)
h_r(m)	246.23-348.54 (289.84, 287.56)	247.12-287.05 (267.06,266.69)	232.71-285.50 (232.71,258.65)	172.56-285.44 (211.90,212.56)

Table 3. Comparison of observed (Wellogic data) steady state depth to groundwater table with simulation results based on the GLB and MLT methods.

	NASH	APB (%)	RMSE
SW1	0.97	8.19	1.09
SW2	0.91	6.07	1.68
SW3	0.98	0.42	1.64
SW4	0.98	0.53	1.70

Table 4. Water Budgets.

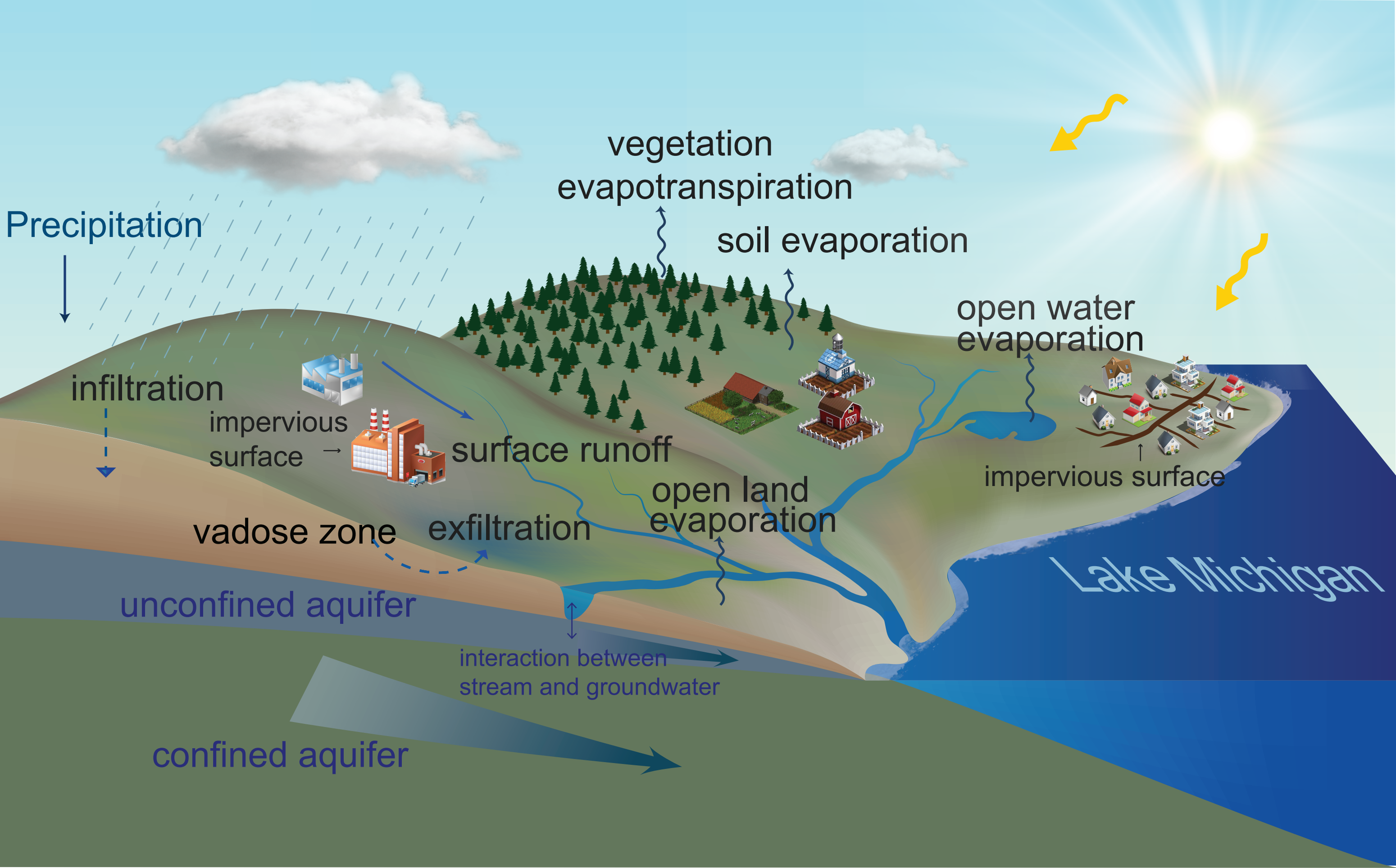
Water balance components (unit: mm yr-1)	SW1	SW2	SW3	SW4	whole KRW
Precipitation	868.54	860.67	876.29	1000.20	913.62
Percent %	100.00	100.00	100.00	100.00	100.00
Infiltration	448.67	463.37	485.43	484.52	471.16
Percent %	51.66	53.84	55.39	48.44	51.57
Recharge	237.29	205.54	198.37	252.00	228.76
Percent %	27.32	23.88	22.64	25.19	25.04
Overland Flow	219.02	207.36	249.60	235.01	233.46
Percent %	25.22	24.09	28.62	22.90	25.55
Groundwater contribution to streams	77.59	79.89	81.61	89.08	85.08
Percent %	8.93	9.28	81.61	9.01	9.31
Net Stream discharge	296.61	286.25	331.21	324.09	318.54
Percent %	34.15	33.26	37.93	31.91	34.87
Evapotranspiration	556.89	566.08	552.78	654.69	584.43
Percent %	64.12	65.77	62.73	65.14	63.97
Storage change	15.04	8.34	-7.70	21.42	10.65
Percent %	1.73	0.97	-0.66	2.95	1.16

Table 5. The fitted ω values and the R^2 for the curve fitting using Equation (9)

	direct ET		inferred ET		effective precipitation	
	ω	R^2	ω	R^2	ω	R^2
SW1	2.45	0.76	2.55	0.45	2.47	0.73
SW2	2.44	0.73	2.48	0.81	2.44	0.70
SW3	2.27	0.74	2.26	0.68	2.27	0.72
SW4	2.85	0.8	2.98	0.63	2.84	0.80
whole KRW	2.44	0.8	2.54	0.47	2.46	0.72

Table 6. Paramters invloved in plotting the Budyko pairs and for calculating the ω value in Equations (12) and (13)

	SW1	SW2	SW3	SW4	whole KRW
Area (km^2)	1395.81	726.41	1154.08	1983.92	5260.22
Elevation (m)	303.46	278.39	272.12	225.59	266.15
Average slope (%)	0.92	0.96	1.30	1.09	1.06
Basin Center Latitude	42.20	42.43	42.36	42.59	42.39
Compound Topographic Index (CTI)	9.18	9.09	8.82	9.19	9.14
Average NDVI	0.39	0.41	0.40	0.42	0.41
Forest (%)	21.32	20.93	22.46	22.82	21.47
Agricultural (%)	53.32	54.80	43.08	47.38	47.32
Urban area (%)	4.97	4.72	10.41	4.58	6.60
Mahalanobis distance	2.23	3.47	4.37	5.49	----
PET, Ep (mm/year)	807.19	848.71	887.35	857.81	846.52
Average K_S (mm hr^{-1})	73.18	54.26	113.53	67.73	80.83
Mean precip intensity i_r (mm hr^{-1})	12.69	9.3	13.83	19.17	11.25
S_{max} (mm)	48.56	48.46	51.77	50.98	50.09
$\theta_f - \theta_w$	0.35	0.34	0.36	0.35	0.35
d_{root} (mm)	140.34	142.10	144.74	147.38	144.73
ω based on simulation	2.47	2.44	2.27	2.84	2.46
ω calculated with Eq.(12)	2.30	2.26	2.12	2.51	2.16
ω calculated with Eq.(13)	2.81	2.8	2.78	2.86	2.95



vegetation
evapotranspiration
soil evaporation
Precipitation
open water
evaporation
infiltration
impervious
surface
surface runoff
impervious surface
open land
evaporation
vadose zone
exfiltration
Lake Michigan
unconfined aquifer
interaction between
stream and groundwater
confined aquifer