HYDROLOGICAL PROCESSES
Hydrol. Process. **26**, 1390–1404 (2012)
Published online 26 September 2011 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.8277

Effects of spatial aggregation of soil spatial information on watershed hydrological modelling

Runkui Li,[1,2] A-Xing Zhu,[2,3] Xianfeng Song,[1]* Baolin Li,[2] Tao Pei,[2] and Chengzhi Qin[2]

[1] *Graduate University of Chinese Academy of Sciences, 19A, Yuquan Road, Shijingshan District, Beijing 100049, China*
[2] *State Key Lab. of Resources & Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, No. 11A Datun Road, An Wai, Beijing 100101, China*
[3] *Department of Geography, University of Wisconsin, 550 North Park Street, Madison, WI 53706, USA*

Abstract:

Many researchers have examined the impact of detailed soil spatial information on hydrological modelling due to the fact that such information serves as important input to hydrological modelling, yet is difficult and expensive to obtain. Most research has focused on the effects at single scales; however, the effects in the context of spatial aggregation across different scales are largely missing. This paper examines such effects by comparing the simulated runoffs across scales from watershed models based on two different levels of soil spatial information: the 10-m-resolution soil data derived from the Soil-Land Inference Model (SoLIM) and the 1 : 24000 scale Soil Survey Geographic (SSURGO) database in the United States. The study was conducted at three different spatial scales: two at different watershed size levels (referred to as full watershed and sub-basin, respectively) and one at the model minimum simulation unit level. A fully distributed hydrologic model (WetSpa) and a semi-distributed model (SWAT) were used to assess the effects. The results show that at the minimum simulation unit level the differences in simulated runoff are large, but the differences gradually decrease as the spatial scale of the simulation units increases. For sub-basins larger than 10 km^2 in the study area, stream flows simulated by spatially detailed SoLIM soil data do not significantly vary from those by SSURGO. The effects of spatial scale are shown to correlate with aggregation effect of the watershed routing process. The unique findings of this paper provide an important and unified perspective on the different views reported in the literature concerning how spatial detail of soil data affects watershed modelling. Different views result from different scales at which those studies were conducted. In addition, the findings offer a potentially useful basis for selecting details of soil spatial information appropriate for watershed modelling at a given scale. Copyright © 2011 John Wiley & Sons, Ltd.

KEY WORDS distributed watershed modelling; resolution of soil spatial information; SWAT model; WetSpa; spatial aggregation; scale

Received 26 October 2010; Accepted 28 June 2011

INTRODUCTION

Geographic information systems (GIS) allow easy incorporation of spatially detailed heterogeneous watershed information, such as land use, elevation, and soil data, into hydrological models. However, the spatial resolution of soil data is usually lower than other input information, such as Digital Elevation Model (DEM) and vegetation data (Band and Moore, 1995; Zhu, 1997) due not only to the large amounts of resources required but also the overall difficulty of producing soil spatial information at detailed spatial scales. Therefore, it is important to evaluate the potential benefits of using detailed soil spatial information in hydrological modelling, particularly meso-scale watershed hydrological modelling, to compare improvement of model performance *versus* the costs of detailed soil spatial information production. Moreover, it is necessary to examine the underlying mechanism for detailed soil spatial information to affect model performance.

The effects of the resolution of soil spatial information on hydrological modelling have been the focus of many studies, but findings from these studies have not been consistent (Mednick *et al*., 2008). Several studies have reported differences in simulated stream flow based on different soil maps but have not drawn any firm conclusions concerning their accuracy (Levick *et al*., 2004; Peschel *et al*., 2006; Kumar and Merwade, 2009). Some researchers have argued that detailed soil data have the potential to improve simulation accuracy (Bosch *et al*., 2004; Di Luzio *et al*., 2004; Anderson *et al*., 2006). Conversely, other studies have shown that varying soil resolution has a limited effect on stream-flow predictions (Cotter *et al*., 2003; Chaplot, 2005; Di Luzio *et al*., 2005; Moriasi and Starks, 2010; Mukundan *et al*., 2010). At the same time, researchers have reported that the effects of resolution of soil spatial information on model results vary with environmental conditions (Zhu and Mackay, 2001; Quinn *et al*., 2005; Geza and McCray, 2008). For example, in evaluating the effects of detailed soil spatial data from the soil-land inference model (SoLIM) (Zhu *et al*., 2001) on watershed modelling in comparison to the Soil Survey Geographic (SSURGO) database

* Correspondence to: Xianfeng Song, Graduate University of Chinese Academy of Sciences, 19A, Yuquan Road, Shijingshan District, Beijing 100049, China. E-mail: song.osgeo@gmail.com

using a RHESSys model, Zhu and Mackay (2001) found that the effects of detailed soil spatial information vary with soil moisture conditions and slope aspects. Quinn *et al*. (2005) found that the effects vary with the size of the hill slope partitions used as basic unit for simulation. Geza and McCray (2008) concluded that the State Soil Geographic (STATSGO) Database performs better in SWAT modelling relative to SSURGO before calibration, while SSURGO performs better after calibration. Both were considered to be at the same satisfactory level of performance.

The effect of watershed size on model response to soil spatial detail has not been adequately considered as existing studies have been mainly conducted at fixed watershed sizes. Watershed response is often highly non-linear and dominated by different processes at different scales. Robinson *et al*. (1995) showed that at small scales the process is dominated by the hill slope response, but the response at large scales is dominated by channel network hydrodynamics. Therefore, special consideration is required to understand the effects of scale on model response to spatial details in soil data in watershed modelling.

Only a few studies have reported the effects of watershed size on hydrological response to input resolution of spatial data, with Wang and Melesse (2006) and Shrestha *et al*. (2006) among them. Wang and Melesse (2006) reported that discrepancies between the simulated discharges of STATSGO and SSURGO in the SWAT model are larger at upstream locations compared to those farther downstream within the study area. They attributed this mainly to the fact that soil merely influences overland hydrologic processes. When moving downstream, the relative importance of channel processes increases and the role of soil in hydrological processes decreases. Shrestha *et al*. (2006) noted the important effects of watershed size on the selection of meteorological data resolution for modelling applications. Thus, their criterion for data selection was based on the influence of meteorological data at the macro-scale and, as a result, might not be suitable for the selection of detailed geographic data (e.g. soil data) at the meso-scale level.

From what has been discussed in the literature, selection of an appropriate level of soil spatial data for modelling a new watershed still poses difficulties. Thus, the important, but inadequately investigated, effects of spatial scale on hydrologic response to levels of detail of soil spatial information need to be examined.

This study evaluates the effects of detailed soil spatial information on hydrologic modelling under different spatial scales, that is, two at different watershed size levels (referred to as full watershed and sub-basin, respectively), and one at the model minimum simulation unit level. Two models, a semi-distributed Soil and Water Assessment Tool, or SWAT model (Arnold and Fohrer, 2005) and a fully distributed Water and Energy Transfer between Soil, Plants and Atmosphere, or WetSpa model (Liu *et al*., 2003; Safari *et al*., (in press)) were used to examine the effects. Two different levels of soil spatial information were used in this exercise: the widely used and most detailed traditional soil survey database in the United States, SSURGO, at a scale of 1 : 24 000, and the more detailed soil spatial data with a spatial resolution of 10 m generated from the SoLIM approach (Zhu *et al*., 2001). This study was conducted in the northwest of the state of Wisconsin's Dane County in the mid-western US.

MATERIALS AND METHODS

The study area

Description. Brewery Creek in western Dane County between Middleton and Cross Plains in the US state of Wisconsin is an agricultural catchment area with a rich spatial database, including detailed soil spatial data (Figure 1). The study area covers about 19·5 km^2 and the elevation of the watershed ranges from about 273 m to 381 m above sea level. This area has a somewhat dissected topography due to its location between the glaciated eastern Dane County and the un-glaciated part of western Dane County, known as the Driftless Area (Graczyk, *et al*., 2003). Brewery Creek flows through outwash and alluvium composed of sandstone and some shale, with most of the bedrock in the watershed being dolomite. The soils are silt loams poorly drained in valley bottoms and highly erodible in the uplands (Glocker and Patzer, 1978). The bed material of the stream channel is mostly composed of soft silt and clay. Agriculture constitutes the major activity and land use of this watershed with the most prevalent crop being hay grown on 30% of the watershed area; other crops include corn and mixed row crops grown on 18 and 4% of the watershed area, respectively. In addition, a notable amount of deciduous forest (22%) and grassland (16%) exists with the remainder area comprising a mix of evergreen forest, wetlands, and urban areas.

Average annual precipitation for the past 30 years was approximately 30 inches (780 mm) per year based on data from the National Oceanic Atmospheric Administration (NOAA) weather station at the Dane County Airport, Madison, Wisconsin, located close to the study area.

Data collection. DEM provided by the US Geological Survey (USGS) with a 10 × 10 m grid size was used for computing slope gradients, extracting stream networks, and delineating sub-basins in the watershed. A SoLIM soil map at 30 feet resolution produced in the digital soil mapping project of Dane County conducted by the US Department of Agriculture (USDA) using the approach developed by Zhu *et al*. (2001) was used as the detailed soil spatial information (Figure 2). A SSURGO soil map at 1 : 24 000 scale (Figure 3) was used as a source of coarser soil spatial information after its attribute tables were converted to a SWAT soil database by a pre-processing extension (Peschel *et al*., 2006). Land-use data were derived from the WISCLAND Land Cover data provided by the Wisconsin Department of

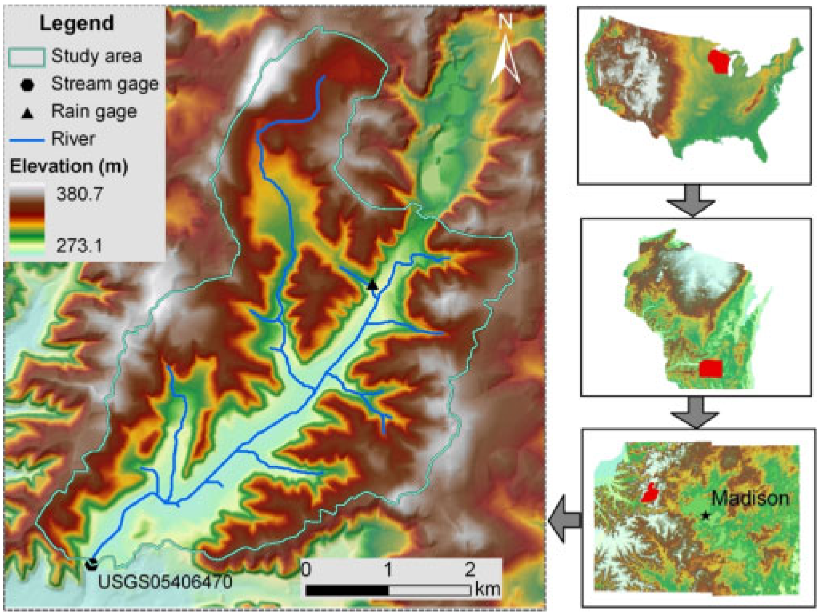

Figure 1. Location of the study area

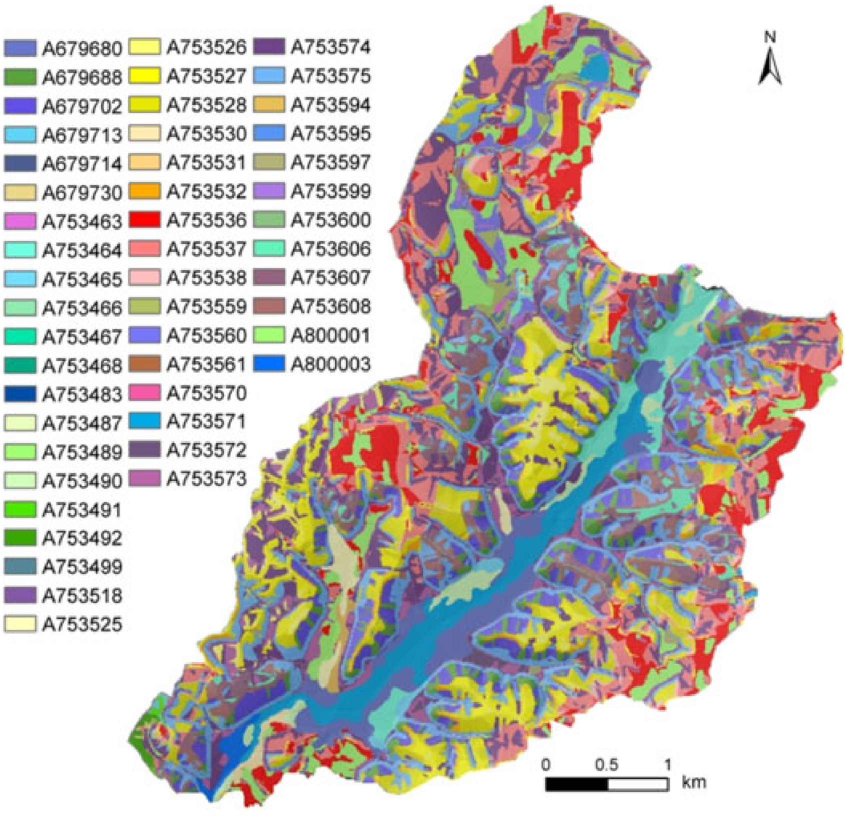

Figure 2. SoLIM soil map (Numbers with initial character of 'A' in the legend are names of soil map units defined by United States Department of Agriculture)

Natural Resources (WiDNR), Madison, Wisconsin, with a resolution of 30 m.

Daily precipitation data on the days above 0° from 1992 to 1996 were collected from a USGS rain gauge located in the centre of the Brewery Creek watershed. Snowfall and daily temperature data were obtained from the Charmany Research Farm Station located about 15 km to the southeast of the watershed outlet. The daily stream flow data observed and used for model calibration were obtained from USGS Gauging Station number 05 406 470, which is located in lower Brewery Creek near Cross Plains (Figure 1).

Soil data and its differences. SSURGO is the most detailed level of soil mapping conducted by the Natural Resources Conservation Service (NRCS) using the

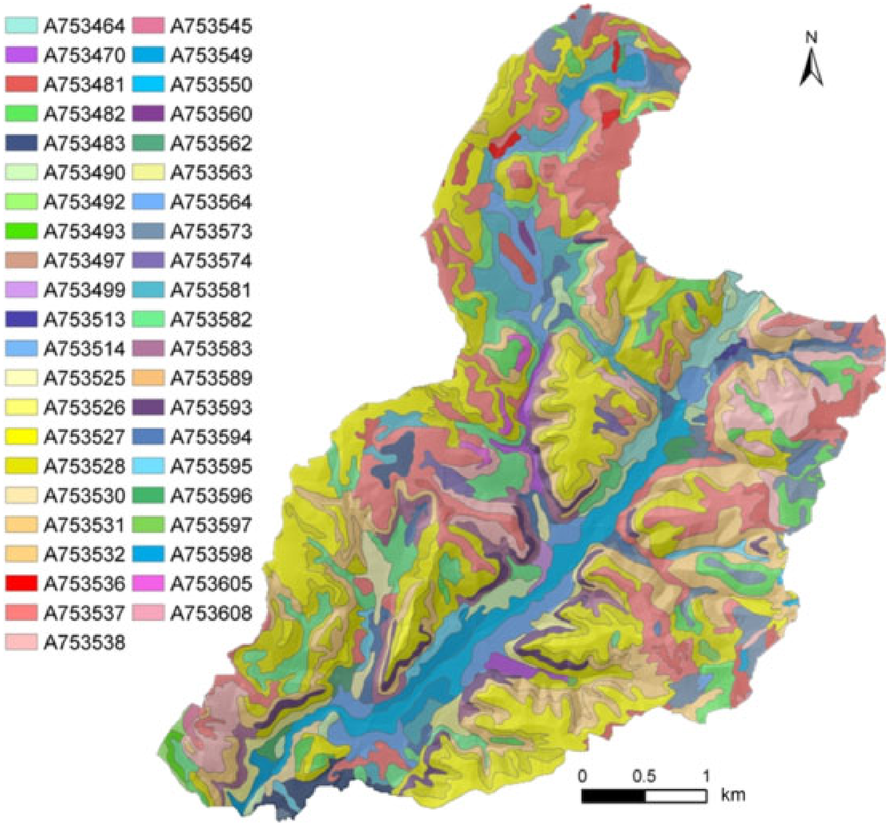

Figure 3. SSURGO soil map (Numbers with initial character of 'A' in the legend are names of soil map units defined by United States Department of Agriculture)

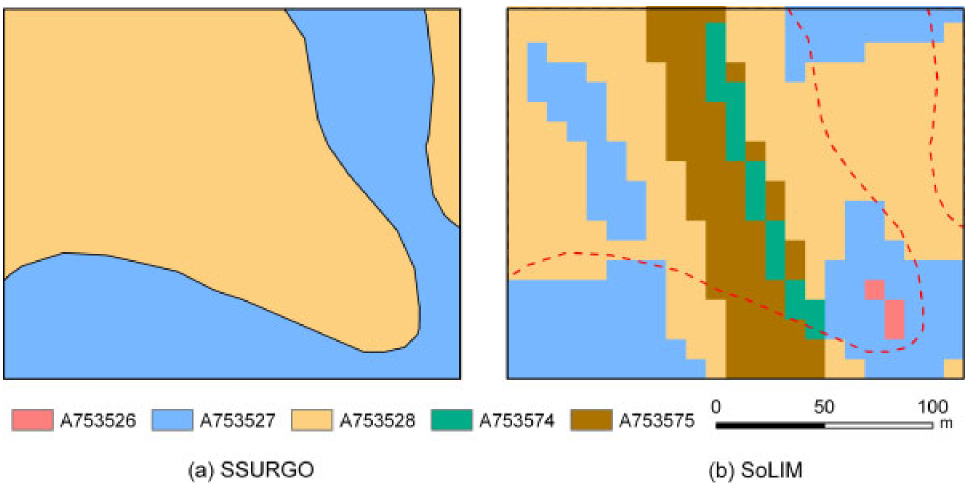

Figure 4. Differences in soil spatial details in Brewery Creek: (a) SSURGO soil map; (b) SoLIM soil map (The dashed line in (b) represents the soil type boundary for the SSURGO map as shown in (a))

traditional soil survey method (USDA, 1993). SSURGO delineates different soil map units through manual drawing of polygons. The smallest soil map unit is usually slightly larger than approximately 2 ha (Geza and McCray, 2008).

SoLIM was developed to overcome the limitations of conventional soil surveys (Zhu and Band, 1994; Zhu *et al.*, 1996). This approach combines knowledge of soil-landscape relationships with GIS techniques under fuzzy logic to map soils at a finer spatial detail and higher attribute accuracy (Zhu *et al.*, 2001). The fuzzy logic concept in representing spatial information makes SoLIM data superior to traditional soil maps in terms of retaining soil spatial heterogeneity (Zhu, 1997). The accuracy of a conventional soil map is about 60% while that of a SoLIM soil map is about 80% (Zhu *et al.*, 2001).

Parts of the SSURGO and SoLIM soil maps from the same area inside Brewery Creek are shown in Figure 4. The SSURGO soil map (Figure 4(a)), which was generated through the manual drawing of soil polygons, only shows the distribution of soil types occupying large areas. It is hard to capture detailed soil variation over space and, thus, only presents the general distribution of two major soil types in Figure 4(a). Owing to the enhanced ability of SoLIM in capturing and retaining spatial detail, the SoLIM soil map (Figure 4(b)) captures soil spatial information at a much finer resolution. As a result, five soil types are depicted over the area shown in Figure 4. As is shown in Figure 4(a) and (b), spatial differences between the conventional SSURGO soil map and the detailed SoLIM soil map are significant.

Description of the models

The SWAT model is a watershed-scale model for long-term assessment. It has proven to be an effective tool in the assessment of water resources and non-point pollution problems over a wide range of scales and environmental conditions around the globe (Gassman *et al.*, 2007). A detailed description of SWAT is given in Neitsch *et al.* (2005), and a comprehensive review of SWAT development, its application, and model analysis can be found in Gassman *et al.* (2007).

SWAT simulates the hydrological process based on the spatial characteristics of climate, topography, soil properties, land use and management practices. It uses a semi-distributed approach to represent the spatial variability of the watershed by subdividing it into a number of sub-basins. Each sub-basin is further subdivided into hydrological response units (HRUs) to reflect the spatial differences in evapotranspiration and other hydrologic conditions and reactions. HRUs with small areas are usually merged with other larger HRUs and, thus, often neglected in reducing complexity and modelling time. In each sub-basin, one HRU might consist of many spatially disconnected patches formed by the same land use and soil type. The water balance for each HRU is represented by storage in snow, soil, shallow aquifers, and deep aquifers. The soil profile is subdivided into soil layers with homogeneous properties. The soil water balance is a key component of the model that includes evaporation, infiltration, plant uptake, surface runoff, lateral flow, and percolation to lower layers (Neitsch, *et al.*, 2005; Arnold and Allen, 1996).

Runoff for each HRU is area weighted and totaled to attain an aggregate runoff within each sub-basin. SWAT then routes water through the stream network to the outlet of the watershed. Stream flow in the sub-basin and the watershed outlet consists of surface runoff, lateral subsurface flow, and base flow.

WetSpa is a grid-based, distributed hydrologic model that simulates water and energy transfer between soil, plants, and the atmosphere originally developed by Wang *et al.* (1996) and adapted for flood prediction over an hourly time scale by De Smedt *et al.* (2000). A detailed description of WetSpa is given in Liu and De Smedt (2004). WetSpa is able to predict peak discharge and hydrographs for any location in a channel network and can estimate flood runoff composition and contributions from certain land use classes. Four layers, or zones, are considered in the vertical direction for each grid cell: the vegetation, root, transmission, and saturated zones. The hydrologic processes considered within each cell are precipitation, snow melt, interception, depression, surface runoff, infiltration, evapotranspiration, percolation, interflow, groundwater flow, and water balance in the root and the saturated zones. Soil moisture content is a crucial factor in the model because it affects the hydrologic processes of surface runoff, actual evapotranspiration, interflow, and percolation from the root zone. Soil hydraulic parameters can be extracted from the 12 USDA soil texture classes provided by Rawls *et al.* (1982) and Cosby *et al.* (1984).

Runoff from different cells in the watershed is routed to the watershed outlet, depending upon flow velocity and wave damping coefficients, using the diffusive wave approximation method. An approximate solution proposed by De Smedt *et al.* (2000) in the form of an instantaneous unit hydrograph (IUH) was used to relate the discharge at the outlet to the available water at the start of the flow path.

Experiment design

The objective of this study is to examine the spatial scale effects on response of simulated water yield to soil spatial details. To fulfill a comprehensive investigation, two different hydrological models are adopted and different soil datasets were fed into identical model settings for simulation of different spatial levels. Simulated water yields from the two soil datasets with model parameters uncalibrated or calibrated either with SoLIM or with SSURGO were compared. Each of the calibrated parameter sets was used for two model runs: one for SoLIM and the other for SSURGO. Water yield from different soil datasets under the same parameter set was then compared. Thus, the difference of model performance was merely caused by the difference of soil data.

Figure 5 shows the effects of spatial scale on the hydrologic response to soil spatial detail investigated by separately comparing the simulated water yield based on different soil data at the minimum modelling unit level

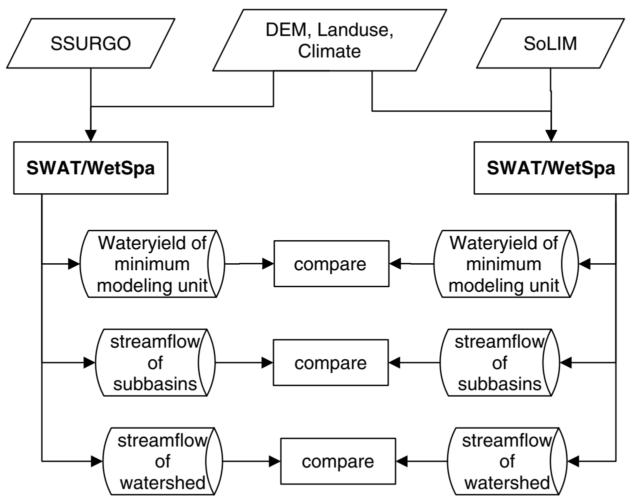

Figure 5. Comparisons of simulated water yield based on SoLIM and SSURGO data under different spatial scales

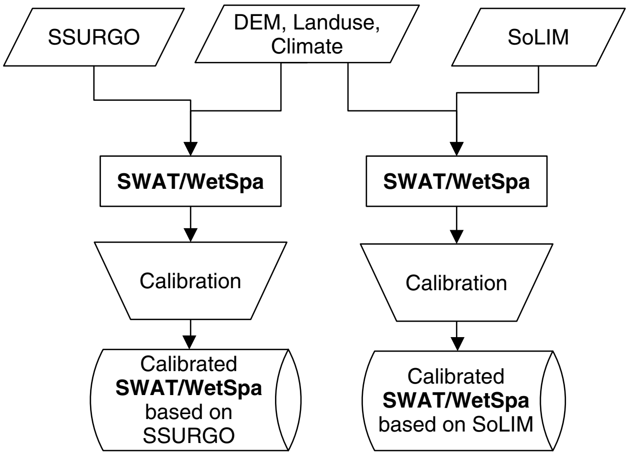

Figure 6. Model calibration procedure: derivation of parameter sets calibrated with SoLIM or SSURGO

(HRU level for SWAT; the cell unit level for WetSpa), sub-basin level, and Brewery Creek watershed level.

The response of a model to soil information may vary with the settings of its parameters. Some parameterisation may obscure the effects of soil representation on the models. In terms of a multivariate parameterisation, calibration of some parameters must be compensating for some other adjustments made in other parameters as well as changes to the spatial characterisation of soil properties. To escape from the risk of specific findings that were generated by one parameter setting and could never be verified in other cases, we repeated the experiments in three different parameter sets using two hydrologic models. The following three parameter sets were prepared for each model: (1) A default parameter set before model calibration; (2) A parameter set after calibration with SSURGO soil data; (3) A parameter set after calibration with SoLIM soil data. Among these calibrated sets of parameters, the parameter category is the same, but their values vary from each other to some extent. The model calibration procedures are shown in Figure 6.

The SoLIM and SSURGO soil data was switched as input under each of the three parameter sets. Simulated water yields from different soil datasets were compared under a fixed parameter set to guarantee the water yield differences are solely from the input soil data. The effects of soil spatial detail on watershed hydrological modelling were then investigated under each of the three different parameter sets mentioned above across different scales.

Parameter specification and parametrisation

ArcSWAT version 1.0.6 with the SWAT2003 executive program was used in this study. Threshold for land use and soil type area proportion to form an HRU in a sub-basin were set to 0 to retain HRUs with small area sizes in order not to lose soil spatial information. The widely used Soil Conservation Service (SCS) runoff curve number method adjusted for soil moisture conditions (Arnold, *et al.*, 1993) was used to estimate surface runoff.

WetSpa (Version 2004) was used in this study. The SoLIM and SSURGO soil maps were converted to 12 USDA soil texture classes based on the top soil layer textural properties. The hydrologic parameter values were derived for each soil texture by the model. Default textural hydraulic values provided by the model were used except for the soil field capacity (FC), the saturated water content (SAT), and the wilting point water content (WP). These three important soil water parameters were derived from the soil profile of the respective soil type maps, not the derived soil texture maps, to maintain their original spatial information. The land use map was reclassified into 14 classes for extraction of vegetation-related parameters and then further reclassified into 6 hydrologic land use classes for simulation of storm runoff partitions from different land use types, i.e. crops, grassland, forest, urban areas, bare soil, and surface water.

Model calibration

Nash-Sutcliffe coefficient (Nash and Sutcliffe, 1970), which is commonly used to report model performance in watershed modelling (Moriasi *et al.*, 2007) and is widely integrated into software packages as objective function for auto-calibration, is applied in this study as the objective function during calibration processes. It is worth noting that recent works reveal that Nash-Sutcliffe coefficient summarizes model performance relative to an extremely weak benchmark (the observed mean output) and does not measure model quality in absolute terms (Schaefli and Gupta, 2007; Gupta *et al.*, 2008).

The objective function is shown as follows:

$$Objective\ Function = \text{maximize} \left(1 - \frac{\sum_{i=1}^{N} (observed_i - predicted_i)^2}{\sum_{i=1}^{N} (observed_i - \overline{observed})^2} \right) \quad (1)$$

where $observed_i$ and $predicted_i$ are daily measured values and simulated values, respectively, and N is the number of days of the modelling period.

The parameters were separately calibrated using SoLIM and SSURGO soil data for both SWAT and WetSpa by daily observed stream flow data from the outlet of the watershed (USGS gauge 05 406 470) over the years 1993–1996. The simulation period was for 5 years: 1 January 1992 to 31 December 1996. The first year, 1992, was used to initialize the model (warm-up period).

For a thorough investigation of the effects of soil data spatial resolution on simulated water yield, soil parameters were not calibrated to retain their original differences between two soil datasets. Seven calibrated parameters and their details are given in Table I for the SWAT model. Calibration was carried out with the help of the SWAT-CUP software package (Abbaspour, 2007). Nine global parameters for WetSpa (Table II) were calibrated using PEST (Doherty, 2004).

Table III shows the Nash-Sutcliffe efficiency coefficient, or NSE, before and after calibration. Before model

Table I. Parameters and their optimal values based on SoLIM and SSURGO data for SWAT

Parameter	Description	Default value	Range	Calibrated value	
				SoLIM	SSURGO
CN2	Initial SCS CN with normal soil moisture	D[a]	(0·8 ∼ 1·2)D[a]	0·8D[a]	0·81 D[a]
ESCO	Soil evaporation compensation factor	0·95	0·01 ∼ 1	0·14	0·40
SMTMP	Snow melt base temperature (degree)	0·5	−5 ∼ 5	2·34	3·9
ALPHA_BF	Baseflow alpha factor (days)	0·048	0 ∼ 1	0·019	0·041
GWQMN	Threshold depth of water in shallow aquifer for return flow to occur (mm)	0·0	0 ∼ 5000	49·1	40·2
RCHRG_DP	Deep aquifer percolation fraction	0·05	0 ∼ 1	0·24	0·29
CH_K2	Effective hydraulic conductivity in main channel (mm/hr)	0·0	0 ∼ 150	67·5	115·2

[a] D: Default parameter values in SWAT, the values may vary with Hydrologic Response Units.

Table II. Calibrated global parameters and their optimal values for WetSpa

Parameter	Description	Default value	Calibrated value	
			SoLIM	SSURGO
Ki	Interflow scaling factor	2·5	15	8·1
Kg	Baseflow recession coefficient	0·01	0·001	0·0012
K_ep	PET correction factor	1·0	0·86	1·0
G0	Initial groundwater storage (mm)	30	100	235
G_max	Maximum groundwater storage in depth (mm)	120	1000	1028
T0	Base temperature for estimating snow melt (degree)	0	−0·084	−0·96
K_snow	Temperature degree-day coefficient for calculating snow melt (mm/degree/day)	2·0	2·5	1·7
K_run	Exponent reflecting the effect of rainfall intensity on runoff coefficient when the rainfall intensity is very small	3·0	6	4·2
P_max	Rainfall intensity threshold for adjusting runoff coefficient (mm/day)	60	203	300

Table III. Nash-Sutcliffe efficiency (NSE) before and after model calibration

Soil data input	Parameter sets for WetSpa			Parameter sets for SWAT		
	Default	Calibrated based on SoLIM	Calibrated based on SSURGO	Default	Calibrated based on SoLIM	Calibrated based on SSURGO
SoLIM	−2·01	0·49	0·48	−13·96	0·32	0·26
SSURGO	−1·41	0·28	0·57	−15·34	0·15	0·29

calibration, NSE is negative for each model and each soil data. This implies a large deviation of the simulated stream flow from the observed values using default model parameters. After model calibration, higher NSE values indicate improved model performance. NSE after model calibration is somewhat low, which is partly because soil parameters are not included for calibration. However, the purpose of this study is to examine the simulated water yield difference between the two soil datasets, therefore, fitness of simulated values *versus* observed values would be less important.

Evaluation indices of the simulated differences

Four indices, total volume difference (TD), relative difference (RD), root mean squared difference ($RMSD$), and consistency coefficient (CC) were used to measure the difference in magnitude between simulated stream flows based on SoLIM and SSURGO soil data, respectively. RD and relative mean absolute error ($R\text{-}MAE$) were used to measure spatial differences between simulated water yield maps.

TD was used to measure the yearly average volume difference between simulated stream flows

$$TD = \frac{1}{N_{\text{year}}}\left(\sum_{\mathrm{i}=1}^{N} Q^{i}_{\text{SoLIM}} - \sum_{\mathrm{i}=1}^{N} Q^{i}_{\text{SSURGO}}\right) \qquad (2)$$

where Q^{i}_{SoLIM} and Q^{i}_{SSURGO} are the simulated stream flows on day i using SoLIM or SSURGO, respectively; N_{year} is the number of years; and N is number of days simulated.

RD was used to measure the relative deviation in simulated water volume based on SoLIM compared to that from SSURGO

$$RD = \frac{\sum_{i=1}^{N} Q^{i}_{\mathrm{SoLIM}} - \sum_{i=1}^{N} Q^{i}_{\mathrm{SSURGO}}}{\sum_{i=1}^{N} Q^{i}_{\mathrm{SSURGO}}} \times 100\% \quad (3)$$

RMSD was used to measure the average daily difference in simulated stream flows between simulation based on SoLIM soil data and that based on SSURGO soil data

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Q^{i}_{\mathrm{SoLIM}} - Q^{i}_{\mathrm{SSURGO}})^2} \quad (4)$$

The Nash-Sutcliffe efficiency, NSE, is very commonly used, and was found to be the best objective function for reflecting the overall fit of a hydrograph (Sevat and Dezetter, 1991). The *CC* used in this study to measure consistency between simulated results based on SoLIM and SSURGO, being approximate to NSE, was calculated as follows:

$$CC = 1 - \frac{\sum_{i=1}^{N}(Q^{i}_{\mathrm{SoLIM}} - Q^{i}_{\mathrm{SSURGO}})^2}{\sum_{i=1}^{N}(Q^{i}_{\mathrm{SSURGO}} - \overline{Q_{\mathrm{SSURGO}}})^2} \quad (5)$$

The theoretical range for the *CC* is the same as that for NSE, i.e. from negative infinity to one. A low value (e.g. $<0{\cdot}5$) would indicate poor consistency between the two simulated stream flow data series. A good consistency would have a *CC* close to 1·0.

R-MAE was used to describe the average difference between the values of the two maps, considering both the positive and negative differences

$$R\text{-}MAE = \frac{\sum_{j=1}^{M}|Q^{j}_{SoLIM} - Q^{j}_{SSURGO}|}{\sum_{j=1}^{M} Q^{j}_{SSURGO}} \times 100\% \quad (6)$$

where Q^{j}_{SoLIM} and Q^{j}_{SSURGO} are the simulated yearly average total water yield volumes at grid cell j based on SoLIM and SSURGO, respectively, and M is the number of total grid cells. The difference between *RD* and *R-MAE* is that the former represents a total volume difference while the latter represents an average 'local' difference.

RD and *R-MAE* (Equations (3) and (6)) were used to describe the equilibration effect of the negative and positive differences. *R-MAE* between the two maps was designated as the 'averaged spatial difference'; *RD* between two maps as 'spatially balanced difference'. Therefore, a difference between *R-MAE* and *RD* shows quantitatively the cancelation effect (an instance of spatial aggregation effect) at that spatial scale. *R-MAE* and *RD* were derived following three steps: First, raster maps of the average annual water yield of the study watershed were derived at the HRUs in the SWAT model and at the cells in the WetSpa model (water yield here refers to the average annual volume of water generated from each modelling unit before routing); Second, the boundary of each sub-basin was extracted; Third, *R-MAE* and *RD* were calculated for each sub-basin based on the simulated water yield of each raster cell within the sub-basin using SoLIM and SSURGO.

RESULTS

Simulation differences at the minimum modelling unit level

A simulated water yield map for WetSpa was generated from the modelling results of each cell unit. The simulated water yield map for the SWAT model was derived by allocating the water yield of each HRU to its spatial location. Simulated annual water yield difference maps between SoLIM and SSURGO soil data were derived for each model under each parameter set.

Table IV shows the statistics of the simulated water yield difference map. As shown in the table, the area for annual water yield differences above 10% is larger than 55% of the study area when WetSpa is used for each parameter set. The area for annual water yield differences above 10% for SWAT model varies from about 9·6% before model calibration to 57% after calibration. The large area for differences above 10% indicates that obvious differences in simulation exist over a large part of the study area owing to the differences of input soil information.

The spatial distribution of the simulated annual water yield differences for parameter set calibrated with SoLIM is shown in Figure 7. As can be seen in the figure, the large water yield differs at many local areas, with positive and negative differences appearing at various locations. The results from either WetSpa (Figure 7(a)) or SWAT (Figure 7(b)) show that large water yield differences originate from differences between SoLIM and SSURGO soil information.

Simulated stream flow differences while varying sub-basin sizes

The effect of varying sub-basin sizes on simulated stream flow differences was investigated by plotting the stream flow differences against the sub-basin area. Differences of the yearly total stream flow volume

Table IV. Proportion for area with simulated annual water yield difference larger than 10% due to difference of input soil information

Parameter set	WetSpa (%)	SWAT (%)
Default	55·35	9·6
Calibrated with SSURGO	57·77	57·0
Calibrated with SoLIM	61·83	57·1

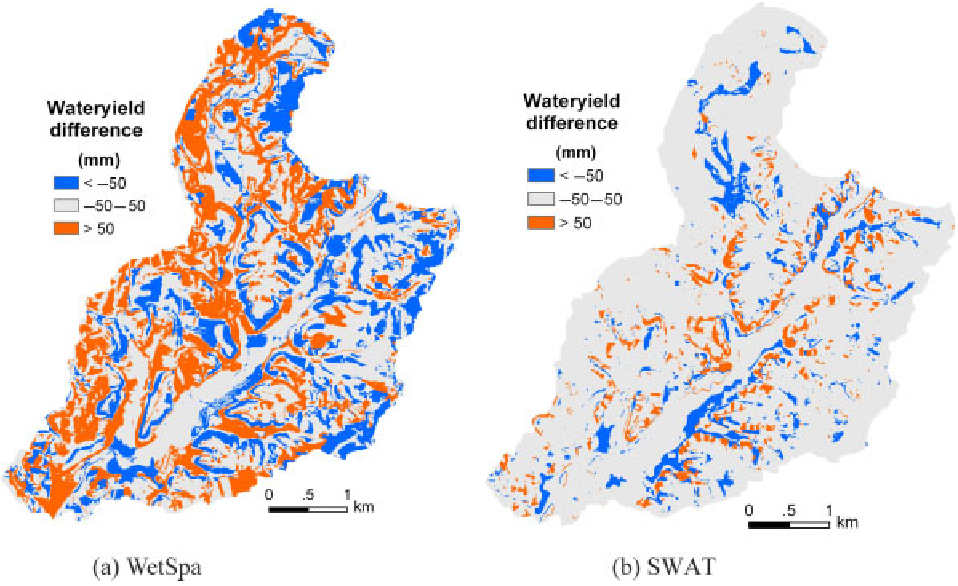

Figure 7. Differences between simulated annual water yield (WYLD) based on SSURGO and SoLIM with parameters set calibrated with SoLIM soil data: (a) WetSpa model; (b) SWAT model (calculated by subtracting SSURGO results from the SoLIM results)

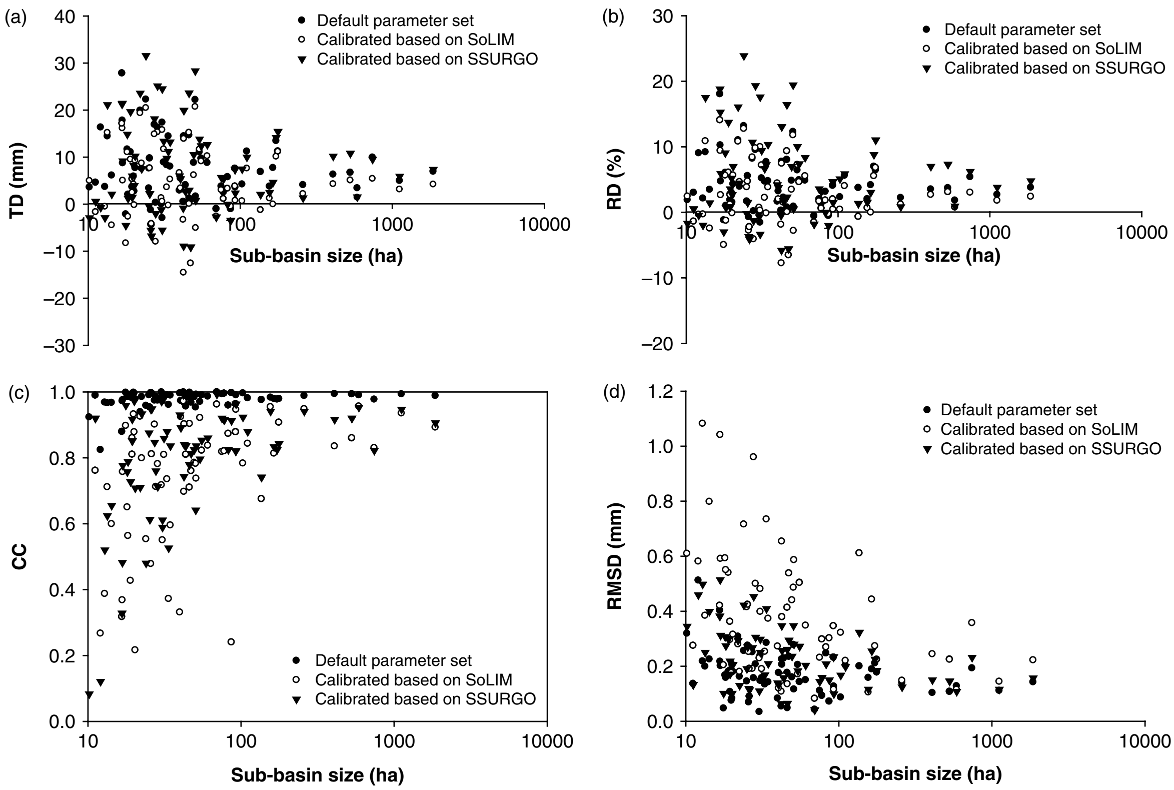

Figure 8. Variation of the differences between WetSpa-simulated stream flows based on SSURGO or SoLIM with sub-basin size: (a) yearly average volume difference; (b) relative difference in yearly volume; (c) consistency coefficient; (d) root mean squared difference

and daily stream flow between different soil spatial information for each sub-basin are shown in Figures 8 and 9.

As shown in the figures, the yearly average volume differences (*TD* and *RD*) and daily differences (*RMSD*) generally decrease with an increase in the size of sub-basin. Consistency of simulated daily stream flow based on SSURGO and SoLIM is generally higher at larger sub-basin sizes (Figure 8(c) and 9(c)). Large differences at small sub-basin size indicate that inconsistent simulated results appeared at small sub-basin levels for different soil spatial information. Smaller differences at larger sub-basin sizes indicate an increased consistency in the simulated results using SSURGO and SoLIM soil information.

Variation in the differences with sub-basin size shows that large differences generally appear locally and differences gradually become smaller as the sub-basin size

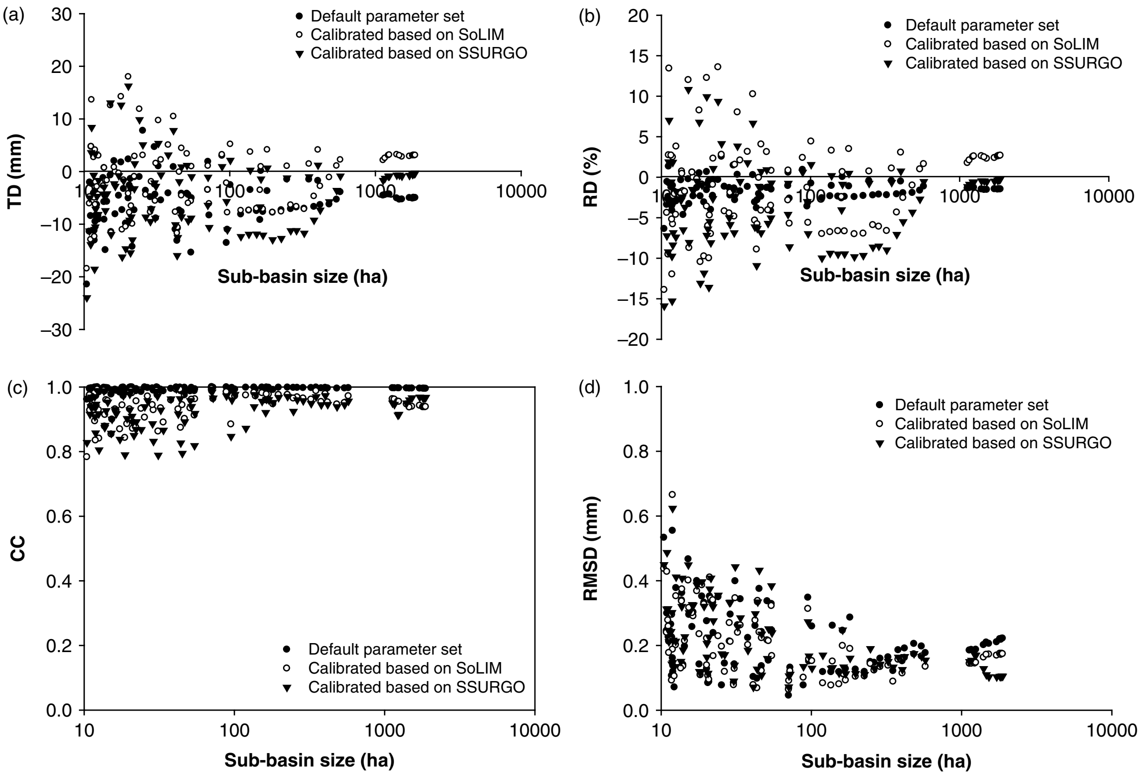

Figure 9. Variation of the differences between SWAT simulated stream flow based on SSURGO or SoLIM with sub-basin size: (a) yearly average volume difference; (b) relative difference in yearly volume; (c) consistency coefficient; (d) root mean squared difference

increases. The simulated results differ very little when the area is larger than 1000 ha (Figures 8 and 9), and, finally, the simulations become consistent at the outlet of the watershed (a total area of nearly 2000 ha).

Differences at the watershed outlet

The simulated daily stream flows based on SSURGO or SoLIM soil data for the whole study area before and after model calibration are shown in Figure 10 and their difference is listed in Table V.

Time series of simulated daily stream flow at watershed outlet from different soil data are shown in Figure 10. Discernible differences between the simulated water yield of SoLIM and SSURGO appear at peak runoffs, especially for calibrated models when simulated water yield is decreased to a comparable volume to that of observed values. Which of the two soil datasets can produce a higher peak runoff may vary with parameter settings. The two models generally produced comparable runoff using SoLIM to that of SSURGO under default parameter values (Figure 10(a) and (d)), though simulated water yield was much higher than observed values. After model calibration, simulated runoff peaks were comparable in magnitude to observed peaks, and WetSpa using SoLIM produces higher runoff than using SSURGO under parameter set calibrated based on SSURGO (Figure 10(c)), while SWAT using SSURGO produces higher runoff than using SoLIM under parameter set calibrated based on SoLIM (Figure 10(e)). Under the same parameter set, which soil dataset can produce higher peak runoff may vary with events (Figure 10(b) and (f)). Figure 10 shows that, difference between simulated runoff of the two soil datasets varies with models, parameter settings and events. However, the consistency coefficients between the simulated runoff of SoLIM and SSURGO exceed 0·90 for most parameter sets for either WetSpa or SWAT model, except for 0·89 for WetSpa calibrated with SoLIM. The largest *RD* observed in yearly average stream-flow volumes was 4·79 and 2·68% for WetSpa and SWAT, respectively. The overall small difference between stream flows based on SSURGO and SoLIM at the watershed outlet suggests that the simulated daily stream flow series from 1993 to 1996 based on SoLIM are similar to that based on SSURGO at Brewery Creek watershed scale.

The reason for the small differences of simulated water yield based on SoLIM and SSURGO soil dataset at the watershed level should not be attributed to model insensitivity because many soil parameters (e.g. *SOL_AWC, SOL_Z, SOL_K*) are important for simulating runoff (van Griensven *et al*., 2006; Xu *et al*., 2010). At the same time, the simulated differences were large at a local scale (Figures 7, 8, and 9). Therefore, aggregation effects during watershed routing are the main reason for small differences at the watershed level.

DISCUSSION

In this study, effects of parameter settings on model response to soil data were considered by using three different parameter sets. Parameters calibrated using observed stream flow at watershed outlet may not reach

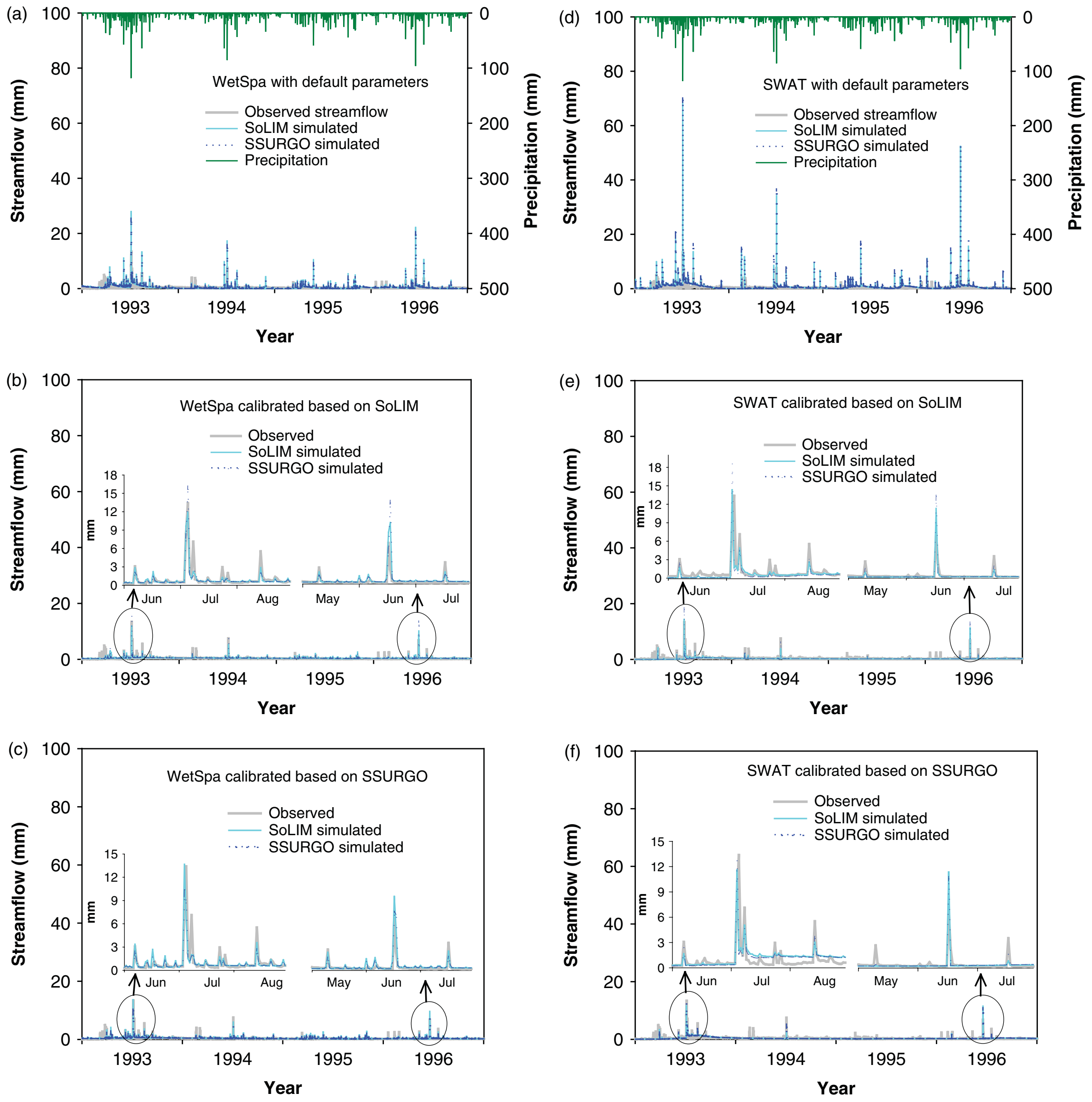

Figure 10. Comparison of simulated daily stream flow of Brewery Creek based on SoLIM and SSURGO soil data

Table V. Difference in simulated stream flows at watershed outlet from 1993 to 1996 based on SSURGO or SoLIM soil data using either the WetSpa or SWAT models with different parameter sets. *CC*: consistency coefficient of daily stream flow (with *CC* = 1 for nil difference); *RD*: relative difference between yearly averaged volume (with *RD* = 0 for nil difference)

Parameter set	Index	WetSpa	SWAT
Default	*CC*	0·99	1·0
	RD (%)	3·77	−1·48
Calibrated with SSURGO	*CC*	0·91	0·97
	RD (%)	4·79	−0·21
Calibrated with SoLIM	*CC*	0·89	0·94
	RD (%)	2·38	2·68

their virtual values at sub-basin level, and this may affect model response to soil dataset at sub-basin and modelling unit levels. However, the three parameter settings used in this study vary from each other to some extent and the results would probably reflect the general response of model to different soil datasets. The fact that simulated differences are small at watershed outlet (Table V) and larger at smaller units (Figures 7, 8, and 9) would not be caused by calibrating different soil data to the same observed streamflow, but originating from the water modelling processes which aggregated soil differences, because water yield from models calibrated based on one soil dataset was not compared to that calibrated based on the other soil dataset.

Differences in spatial resolution between SSURGO and SoLIM soil maps generate spatial disagreements of soil spatial information at local scale level. The discrepancy of soil distribution in the two maps subsequently leads to the spatial discrepancy in the simulated water yields. However, soil properties might have either negative or positive differences, thus leading to an increase or decrease in water yield in simulation. As areal units are added during the watershed routing process, these local differences essentially cancel each other, resulting in only small differences observed at the watershed outlet.

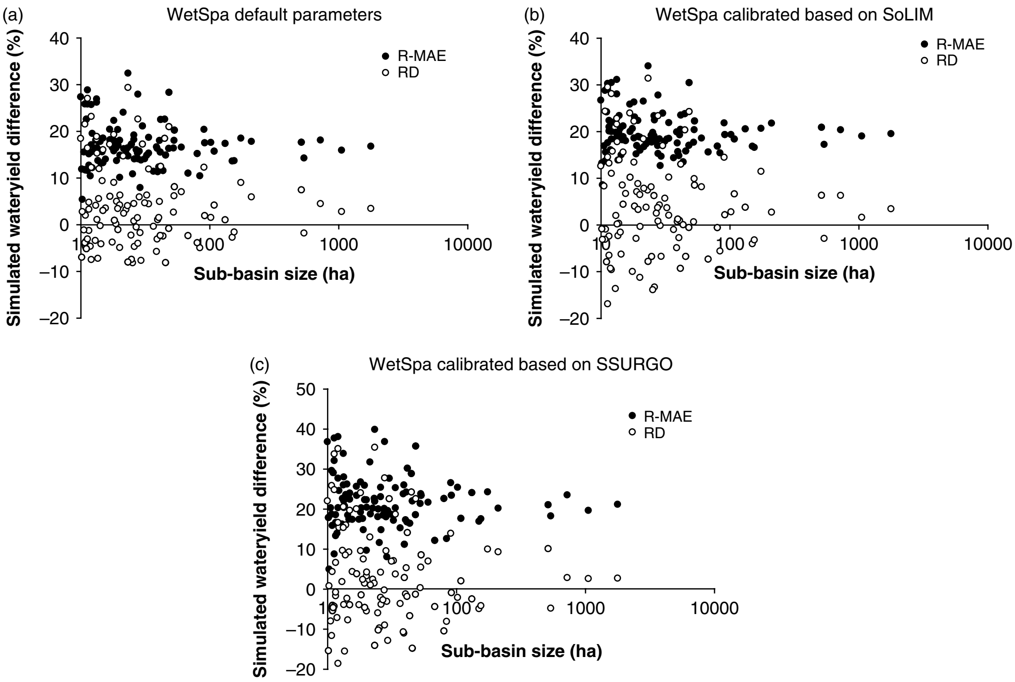

Figure 11. Difference between *R-MAE* and *RD* for different sub-basin sizes, WetSpa-simulated water yields: (a) default parameter set; (b) parameter set calibrated with SoLIM; (c) parameter set calibrated with SSURGO

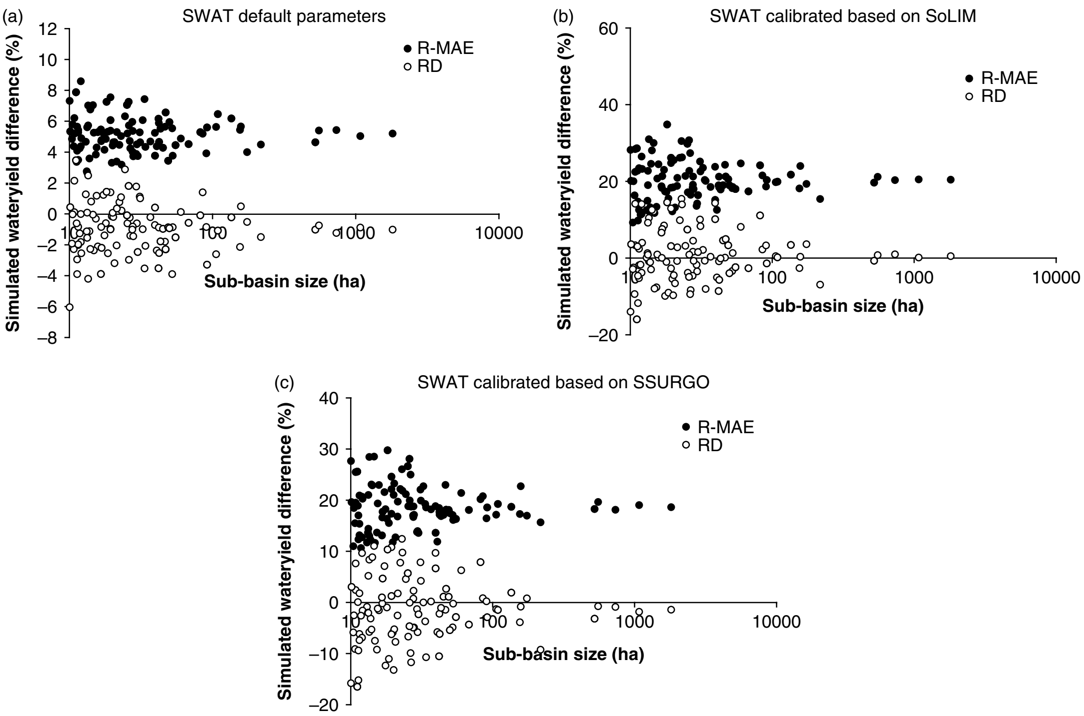

Figure 12. Difference between *R-MAE* and *RD* for different sub-basin sizes, SWAT-simulated water yields: (a) default parameter set; (b) parameter set calibrated with SoLIM; (c) parameter set calibrated with SSURGO

R-MAE and *RD* values calculated for the sub-basins in the study area are shown in Figures 11 and 12. *R-MAE* varies considerably when the sub-basin size is small, but gradually stabilizes when the sub-basin is larger than 1000 ha. The final *R-MAE* for the whole watershed stabilizes at about 20% using WetSpa, and 5 and 20% using SWAT before and after model calibration, respectively. Large *R-MAE* indicates large spatial differences of simulated water yield exists using different soil data.

RD also varies considerably over small-sized sub-basins and approaches stability when the sub-basin is larger than 1000 ha. The trend is the same for both WetSpa and SWAT. The decrease in *RD* with increased sub-basin size shows the approximation of spatially averaged mean simulated water yields based on SoLIM and SSURGO at larger spatial scales.

R-MAE is almost steady and far from 0 when the sub-basin is larger than approximately 1000 ha; however, *RD* is much smaller and generally closer to 0 at this scale. The large difference between *R-MAE* and *RD* at areas larger than 1000 ha shows the spatial balancing effects of local differences at this spatial scale. Spatial balancing effects are caused by the cancelation of negative and positive differences during the aggregation (routing) process of modelling units inside the sub-basin.

The spatial balancing effect becomes larger with an increase in sub-basin size, the reason why stream flows simulated with SoLIM and SSURGO soil data become similar. The differences between simulated stream flows are much smaller when the sub-basin size is larger than approximately 1000 ha for Brewery Creek watershed as shown in Figures 11 and 12.

General response patterns are same for SWAT and WetSpa to different soil data across varying sub-basin scales, but the magnitudes of responses are different for the two models (Figure 8 *versus* 9, Figure 11 *versus* 12). It is clear that WetSpa is more sensitive to soil data than SWAT, by comparison SWAT shows little sensitivity to soil data source, and almost too little to be of concern at the outlet of watershed. The differences of responses between SWAT and WetSpa can be analysed from procedures for the two models to derive hydrologic parameters from soil data. As multiple HRUs were used in each sub-basin for full consideration of soil information representation during HRU delineations, there was no loss of soil spatial detail from soil map to HRU representation. Grid cell representation of WetSpa can also fully represent soil spatial distribution. Thus, difference in spatial representation scheme between models would not be the reason for difference of model response. However, differences exist in the soil parameter extracting process. For SWAT model, one of the most important input parameters in SWAT model to simulate runoff is CN2 which is determined by hydrologic soil group and land use type, and may further be adjusted by terrain slope. Hydrologic soil group is a general soil hydrologic attributes describing the ability of a soil to generate runoff, and all types of soils are classified into merely four groups (A/B/C/D). Generalising of many different soil types into four hydrologic soil groups causes loss of detailed soil information. As a result, most of the study area was classified as group B using either SoLIM or SSURGO. This leads to the fact that over 90% of the study area has the same CN2 based on the two soil datasets; For WetSpa, soil types were classified into 12 soil texture classes, and soil hydrologic parameters were extracted from the corresponding soil textures for each modelling cell. Finally, 38·5% of the study area was classified as different soil texture based on SSURGO and SoLIM soil dataset. The high consistency of CN2 in SWAT model using the two soil datasets and relative low consistency of soil texture for WetSpa using the two soil datasets is probably an important reason for the different magnitudes of model responses to soil data.

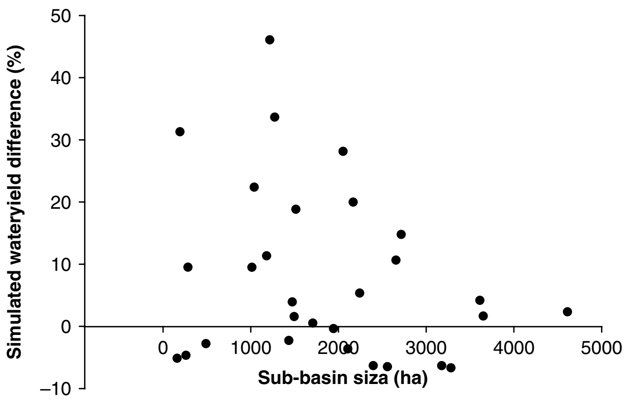

Figure 13. Variation in daily average water yield difference, STATSGO *versus* SSURGO data, with sub-basin area, derived from data in Table V of Peschel *et al.* (2006)

Data derived from other studies (Table V in Peschel *et al.*, 2006) show the same trend in differences of daily average water yield using STATSGO and SSURGO soil data as sub-basin size increases (Figure 13). The clear trend of decrease in simulation difference with the increase of sub-basin size shows an obvious spatial aggregation effect.

Variation of simulated stream flow differences between different soil spatial information with sub-basin size suggests watershed size is an important factor when evaluating effects of spatial resolution of input data on hydrologic modelling. The use of detailed soil spatial information does not significantly change the simulated stream flow compared to coarsely defined soil information after the watershed size exceeds a certain threshold. This relationship with watershed characteristics deserves a separate study in itself.

Spatial aggregation effects also explain different views reported in the literature concerning the impact of spatial detail of soil data on watershed modelling. Summarising the data from former studies, e.g. spatial scale of the different soil data used, the watershed size and the corresponding simulated average stream flow differences (Table VI), a general trend can be observed in which simulated average stream flow differences become smaller as the study area increases. Table VI provides an opportunity to verify the spatial aggregation effect emphasized in this paper and shows its potential guidance for selecting the required spatial detail of geographic data in a new watershed.

CONCLUSIONS

This study assessed the scale effects of using the SoLIM soil map at 10 m spatial resolution *versus* the SSURGO map at a scale of 1 : 24 000 as the inputs for stream flow simulations with SWAT and WetSpa models. Simulations based on the different soil data were compared at three

Table VI. Summary of the results derived from related studies, with the cases ordered by drainage area in ascending sequence

Reference	Model	Coarse soil	Detailed soil	Drainage area (km^2)	Average Stream flow difference[a] (%)
Di Luzio *et al*. (2005)	SWAT	1 : 250 000	1 : 24 000	21·3	12·1
Chaplot (2005)	SWAT	1 : 250 000	1 : 25 000	21·8	19·4
Wang & Melesse (2006)	SWAT	1 : 250 000	1 : 24 000	515·4	7·4
Peschel *et al*. (2006)	SWAT	1 : 250 000	1 : 24 000	540·0	5·2
Tang *et al*.(2000)	ADAPT	1 : 250 000	1 : 24 000	1400·0	2·6

[a] The average difference of simulated stream flows during the entire modelling period was used here. Average stream flow difference (%) = abs($\text{Streamflow}_{\text{Detailed soil}} - \text{Streamflow}_{\text{Corase soil}}$)/$\text{Streamflow}_{\text{Corase soil}}$ 100%. This table is not exhaustive since some studies did not provide values of simulated stream flow differences between different soil data.

spatial levels: the meso-scale watershed level, the sub-basin level, and the model minimum simulation unit level. Large differences between the simulated water yields exist at the local scale (model minimum simulation unit level and small sized sub-basin level). However, the simulated differences generally decrease with an increase in sub-basin size. A threshold area was found to be approximately 1000 ha larger than which SoLIM and SSURGO would generate similar stream flow simulations in the Brewery Creek watershed. Differences in simulated stream flows from different soil spatial information were shown to be largely affected by the size of modelling area. The effect of spatial scale detected mainly appears to originate from spatial aggregation that balances out differences in local runoff.

The unique findings in this paper provide an important and unified perspective on different views reported in the literature concerning the impact of detailed soil information on watershed modelling. In addition, the findings offer a useful basis for selecting the level of detail required for watershed modelling at different scales, important because detailed soil spatial information is expensive and difficult to obtain over large areas. These results are initial and obtained from the investigation of a specific watershed. The threshold area obtained here should be verified in other areas and with additional factors, such as the relief of the topography and climate, among other factors. Other response variables (e.g., sediments, nutrients) should also be considered in future studies to see if similar effects exist for these response variables.

ACKNOWLEDGMENTS

This study is financially supported by National Basic Research Program of China (2007CB407207), National Natural Science Foundation of China (Project No.: 40971236 and Project No.: 41023010), the National High Technology Research and Development Program ("863" Program) of China (Project No.: 2011AA120305), Program of International S & T Cooperation, MOST of China (Project No.: 2010DFB24140). The Vilas Associate Award and the Hammel Faculty Fellow Award from the University of Wisconsin-Madison, and Innovation Grant from the State Key Laboratory of Resources and Environmental Information Systems. The authors would like to acknowledge Randy J Hunt and David J Graczyk in the US Geological Survey—WRD for providing the groundwater aquifer information of the study area. The anonymous reviewers of this paper are highly acknowledged by the authors.

REFERENCES

Abbaspour KC. 2007. SWAT-CUP, SWAT Calibration and Uncertainty Programs. A User Manual, EAWAG. Zurich, Switzerland.

Anderson RM, Koren VI, Reed SM. 2006. Using SSURGO data to improve Sacramento Model a priori parameter estimates. *Journal of Hydrology* **320**: 103–116. DOI: 10.1016/j.jhydrol.2005.07.020.

Arnold JG, Allen PM. 1996. Estimating hydrologic budgets for three Illinois watersheds. *Journal of Hydrology* **176**: 57–77.

Arnold JG, Allen PM, Bernhardt G. 1993. A Comprehensive Surface-Groundwater Flow Model. *Journal of Hydrology* **142**(1): 47–69.

Arnold JG, Fohrer N. 2005. SWAT2000: current capabilities and research opportunities in applied watershed modeling. *Hydrological Processes* **19**(3): 563–572.

Band LE, Moore ID. 1995. Scale: Landscape attributes and geographical information systems. *Hydrological Processes* **9**(3): 401–422.

Bosch DD, Sheridan JM, Batten HL, Arnold JG. 2004. Evaluation of the SWAT model on a coastal plain agricultural watershed. *Transactions of the ASAE* **47**(5): 1493–1506.

Chaplot V. 2005. Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO_3-N loads predictions. *Journal of Hydrology* **312**(1): 207–222.

Cosby BJ, Hornberger GM, Clapp RB, Ginn TR. 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resources Research* **20**(6): 682–690.

Cotter AS, Chaubey I, Costello TA, Soerens TS, Nelson MA. 2003. Water quality model output uncertainty as affected by spatial resolution of input data. *Journal of the American Water Resources Association (JAWRA)* **39**(4): 977–986. DOI: 10.1111/j.1752-1688.2003.tb04420.x.

De Smedt F, Liu YB, Gebremeskel S. 2000. Hydrologic modelling on a catchment scale using GIS and remote sensed land use information. In *Risk Analysis II* Brebbia CA (ed). WTI Press: Southampton, Boston; 295–304.

Di Luzio M, Arnold JG, Srinivasan R. 2004. Integration of SSURGO maps and soil parameters within a geographic information system and nonpoint source pollution model system. *Journal of Soil and Water Conservation* **59**(4): 123–133.

Di Luzio M, Arnold JG, Srinivasan R. 2005. Effect of GIS data quality on small watershed stream flow and sediment simulations. *Hydrological Processes* **19**(3): 629–650.

Doherty J. 2004. *PEST: Model Independent Parameter Estimation User Manual*, 5th edn. Watermark Numerical Computing: Brisbane, Australia.

Gassman PW, Reyes MR, Green CH, Arnold JG. 2007. Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions. *Transactions of the ASABE* **50**(4): 1211–1250.

Geza M, McCray JE. 2008. Effects of soil data resolution on SWAT model stream flow and water quality predictions. *Journal of Environmental Management* **88**(3): 393–406.

Glocker CL, Patzer RA. 1978. Soil Survey of Dane County, US Dept. of Agriculture, Soil Conservation Service. 193.

Graczyk DJ, Walker JF, Horwatich JA, Bannerman RT. 2003. Effects of Best-Management Practices in the Black Earth Creek Priority

Watershed, Wisconsin, 1984-98, Water-Resources Investigations Report 2003–4163. US Geological Survey, Middleton, WI.

Gupta HV, Wagener T, Liu Y. 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes* **22**(18): 3802–3813.

Kumar S, Merwade V. 2009. Impact of Watershed Subdivision and Soil Data Resolution on SWAT Model Calibration and Parameter Uncertainty. *Journal of the American Water Resources Association* **45**(5): 1179–1196. DOI: 10.1111/j.1752-1688.2009.00353.x.

Levick LR, Semmens DJ, Guertin DP, Burns IS, Scott SN, Unkrich CL, Goodrich DC. 2004. Adding Global Soils Data to the Automated Geospatial Watershed Assessment Tool (AGWA), Proceedings 2nd SAHRA, (Sustainability of Semi-Arid Hydrologic and Riparian Areas), University of Arizona, International Symposium on Transboundary Water Management, Tucson, AZ. 16–19 November 2004.

Liu YB, De Smedt F. 2004. WetSpa Extension, Documentation and User Manual. Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Belgium.

Liu YB, Gebremeskel S, De Smedt F, Hoffmann L, Pfister L. 2003. A diffusive transport approach for flow routing in GIS-based flood modeling. *Journal of Hydrology* **283**(1): 91–106.

Mednick AC, Sullivan J, Watermolen DJ. 2008. Comparing the use of STATSGO and SSURGO soils data in water quality modeling: A literature review. Bureau of Science Services. Wisconsin Department of Natural Resources. Issue 60. http://www.dnr.state.wi.us/org/es/science/publications/PUB_SS_760_2008.pdf.

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50**(3): 885–900.

Moriasi DN, Starks PJ. 2010. Effects of the resolution of soil dataset and precipitation dataset on SWAT2005 streamflow calibration parameters and simulation accuracy. *Journal of Soil and Water Conservation* **65**(2): 163–178. DOI: 10.2489/jswc.65.2.63.

Mukundan R, Radcliffe DE, Risse LM. 2010. Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. *Journal of Soil and Water Conservation* **65**(2): 92–104. DOI: 10.2489/jswc.65.2.92.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models, Part I: A discussion of principles. *Journal of Hydrology* **10**(3): 282–290.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR, 2005. Soil and Water Assessment Tool Theoretical Documentation, Version 2005. Temple, TX: USDA-ARS Grassland, Soil and Water Research Laboratory. Available at: www.brc.tamus.edu/swat/doc.html. Accessed 1 November 2006.

Peschel JM, Haan PK, Lacey RE. 2006. Influences of soil dataset resolution on hydrologic modeling. *Journal of the American Water Resources Association* **42**(5): 1371–1389. DOI: 10.1111/j.1752-1688.2006.tb05619.x.

Quinn T, Zhu AX, Burt JE. 2005. Effects of detailed soil spatial information on watershed modeling across different model scales. *International Journal of Applied Earth Observation and Geoinformation* **7**(4): 324–338.

Rawls WJ, Brakensiek DL, Saxton KE. 1982. Estimation of soil water properties. *Transactions of the ASAE* **25**(5): 1316–1320.

Robinson JS, Sivapalan M, Snell JD. 1995. On the Relative Roles of Hillslope Processes, Channel Routing, and Network Geomorphology in the Hydrologic Response of Natural Catchments. *Water Resources Research* **31**(12): 3089–3101. DOI: 10.1029/95wr01948.

Safari A, De Smedt F, Moreda F. (In press). WetSpa model application in the Distributed Model Intercomparison Project (DMIP2). *Journal of Hydrology*. DOI: 10.1016/j.jhydrol.2009.04.001.

Schaefli B, Gupta HV. 2007. Do Nash values have value? *Hydrological Processes* **21**(15): 2075–2080.

Sevat E, Dezetter A. 1991. Selection of calibration objective functions in the context of rainfall-runoff modeling in a Sudanese savannah area. *Hydrological Science Journal* **36**(4): 307–330.

Shrestha R, Tachikawa Y, Takara K. 2006. Input data resolution analysis for distributed hydrological modeling. *Journal of Hydrology* **319**(1): 36–50.

Tang WL, Chen M, Ward AD, Desmond E, Shang H, White D. 2000. Comparison of ADAPT Model Between Different Scale Soils Data Bases on Predicted Hydrologic Responses of America Ohio Darby Creek Watershed. *Journal of Soil and Water Conservation* **14**(2): 15–18, 35. (In Chinese).

USDA. 1993. Soil Survey Manual. Soil Conservation Service. US Department of Agriculture Handbook 18.

van Griensven A, Meixner T, Grunwald S, Bishop T, Diluzio M, Srinivasan R. 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology* **324**: 10–23.

Wang X, Melesse AM. 2006. Effects of STATSGO and SSURGO as inputs on SWAT model's snowmelt simulation. *Journal of American Water Resources Association* **42**(5): 1217–1236.

Wang ZM, Batelaan O, De Smedt F. 1996. A distributed model for water and energy transfer between soil, plants and atmosphere (WetSpa). *Physics and Chemistry of the Earth* **21**(3): 189–193.

Xu HM, Taylor RG, Kingston DG, Jiang T, Thompson JR, Todd MC. 2010. Hydrological modeling of River Xiangxi using SWAT2005: A comparison of model parameterizations using station and gridded meteorological observations. *Quaternary International* **226**(1): 54–59.

Zhu AX. 1997. A similarity model for representing soil spatial information. *Geoderma* **77**: 217–242.

Zhu AX, Band LE. 1994. A knowledge-based approach to data integration for soil mapping. *Canadian Journal of Remote Sensing* **20**(4): 408–418.

Zhu AX, Band LE, Dutton B, Nimlos TJ. 1996. Automated Soil Inference Under Fuzzy Logic. *Ecological Modeling* **90**(22): 123–145.

Zhu AX, Hudson B, Burt JE, Lubich K, Simonson D. 2001. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal* **65**(5): 1463–1472.

Zhu AX, Mackay DS. 2001. Effects of spatial detail of soil information on watershed modeling. *Journal of Hydrology* **248**(1): 54–77. DOI: 10.1016/S0022-1694(01)00390-0.